

# Lineare Algebra II

Sommersemester 2012

Universität Bayreuth

MICHAEL STOLL

## INHALTSVERZEICHNIS

18. Summen von Untervektorräumen, Komplemente, Kodimension	2
19. Polynome und Matrizen bzw. Endomorphismen	12
20. Die Jordansche Normalform	20
21. Simultane Diagonalisierbarkeit	36
22. Äquivalenzrelationen, Quotientenräume und affine Unterräume	40
23. Der Dualraum	50
24. Bilinearformen und quadratische Formen	60
25. Euklidische Vektorräume	68
26. Orthogonale Diagonalisierung	74
27. Klassifikation von Quadriken	79
28. Unitäre Vektorräume	86
29. Orthogonale Gruppen und Quaternionen	91
30. Äußere direkte Summe und Tensorprodukt	101
31. Symmetrische und alternierende Potenzen	111

## 18. SUMMEN VON UNTERVEKTORRÄUMEN, KOMPLEMENTE, KODIMENSION

Unser nächstes größeres Ziel ist die Vervollständigung der Klassifikation der Endomorphismen eines endlich-dimensionalen Vektorraums  $V$  (oder äquivalent: der Klassifikation von  $n \times n$ -Matrizen bis auf Ähnlichkeit) im Fall, dass das charakteristische Polynom in Linearfaktoren zerfällt (was über einem algebraisch abgeschlossenen Körper wie  $\mathbb{C}$  immer der Fall ist). Das wird auf die sogenannte „Jordan-Normalform“ führen. In diesem Zusammenhang wird es hilfreich sein, den Vektorraum  $V$  zu „zerlegen“, sodass der Endomorphismus auf den einzelnen „Teilen“ von  $V$  ein leicht überschaubares Verhalten zeigt. Dafür brauchen wir den Begriff der „direkten Summe“ von (Unter-)Vektorräumen.

Sei  $V$  ein Vektorraum. Wir erinnern uns daran, dass beliebige Durchschnitte von Untervektorräumen von  $V$  wieder Untervektorräume sind (Lemma 7.2), dass das im Allgemeinen aber nicht für Vereinigungen von Untervektorräumen gilt. Statt dessen können wir aber den kleinsten Untervektorraum betrachten, der alle betrachteten Untervektorräume (und damit ihre Vereinigung) enthält. Das führt auf folgende Definition.

\* **18.1. Definition.** Seien  $V$  ein Vektorraum und  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen von  $V$ . Dann heißt der von der Vereinigung  $\bigcup_{i \in I} U_i$  erzeugte Untervektorraum von  $V$  die *Summe* der Untervektorräume  $U_i$ ; wir schreiben

**DEF**  
Summe von  
Unter-VR

$$\sum_{i \in I} U_i = \left\langle \bigcup_{i \in I} U_i \right\rangle.$$

Im Fall  $I = \{1, 2, \dots, n\}$  schreibt man auch

$$U_1 + U_2 + \dots + U_n \quad \text{oder} \quad \sum_{i=1}^n U_i$$

statt  $\sum_{i \in I} U_i$ .

◇

Die Schreibweise erklärt sich durch die folgende Eigenschaft.

**18.2. Lemma.** Sei  $V$  ein Vektorraum.

**LEMMA**  
Elemente  
der Summe

(1) Sind  $U_1, U_2, \dots, U_n$  Untervektorräume von  $V$ , dann ist

$$U_1 + U_2 + \dots + U_n = \{u_1 + u_2 + \dots + u_n \mid u_1 \in U_1, u_2 \in U_2, \dots, u_n \in U_n\}.$$

(2) Ist  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen von  $V$ , dann ist

$$\sum_{i \in I} U_i = \left\{ \sum_{i \in J} u_i \mid J \subset I \text{ endlich, } u_i \in U_i \text{ für alle } i \in J \right\}.$$

*Beweis.* Es ist klar, dass die jeweils rechts stehende Menge in der links stehenden enthalten ist, denn ihre Elemente sind Linearkombinationen von Elementen von  $U_1 \cup \dots \cup U_n$  bzw.  $\bigcup_{i \in I} U_i$  (Satz 7.8). Wir zeigen, dass die rechts stehende Menge ein Untervektorraum von  $V$  ist. Dann folgt analog zum Beweis von Satz 7.8, dass sie das Erzeugnis der Vereinigung der  $U_i$  ist. Sei  $U$  die Menge auf der rechten Seite. Wir prüfen die drei Bedingungen für einen Untervektorraum nach (Definition 6.1). Dabei nutzen wir aus, dass die  $U_i$  Untervektorräume sind.

- $\mathbf{0} \in U$ : Wir können alle  $u_i = \mathbf{0}$  wählen (bzw. im zweiten Fall für  $J$  die leere Menge nehmen).

- Abgeschlossenheit unter der Addition: Im ersten Fall seien

$$u = u_1 + u_2 + \dots + u_n \quad \text{und} \quad u' = u'_1 + u'_2 + \dots + u'_n$$

zwei Elemente von  $U$  (mit  $u_i, u'_i \in U_i$  für alle  $i \in \{1, 2, \dots, n\}$ ). Dann ist auch

$$u + u' = (u_1 + u'_1) + (u_2 + u'_2) + \dots + (u_n + u'_n) \in U.$$

Im zweiten Fall seien  $J, J' \subset I$  endlich und

$$u = \sum_{i \in J} u_i \quad \text{und} \quad u' = \sum_{i \in J'} u'_i$$

zwei Elemente von  $U$ . Wenn wir  $u_i = 0$  (bzw.  $u'_i = 0$ ) setzen für  $i \in J' \setminus J$  (bzw.  $i \in J \setminus J'$ ), dann gilt

$$u + u' = \sum_{i \in J} u_i + \sum_{i \in J'} u'_i = \sum_{i \in J \cup J'} u_i + \sum_{i \in J \cup J'} u'_i = \sum_{i \in J \cup J'} (u_i + u'_i) \in U.$$

- Abgeschlossenheit unter der Skalarmultiplikation: Seien  $\lambda$  ein Skalar und  $u = u_1 + u_2 + \dots + u_n$  bzw.  $u = \sum_{i \in J} u_i$  ein Element von  $U$ . Dann ist

$$\lambda u = \lambda u_1 + \lambda u_2 + \dots + \lambda u_n \quad \text{bzw.} \quad \lambda u = \sum_{i \in J} \lambda u_i$$

wieder ein Element von  $U$ . □

### 18.3. Beispiele.

- Ist  $I = \emptyset$ , dann ist  $\sum_{i \in I} U_i = \{\mathbf{0}\}$  der Null-Vektorraum.
- Ist  $U \subset V$  ein Untervektorraum, dann gilt  $U + U = U$ .
- Ist  $V$  ein Vektorraum,  $I$  eine Menge und sind (für  $i \in I$ )  $A_i \subset V$  beliebige Teilmengen, dann gilt

$$\sum_{i \in I} \langle A_i \rangle = \left\langle \bigcup_{i \in I} A_i \right\rangle.$$

(Beweis als Übung.) Das bedeutet: Ist  $A_i$  ein Erzeugendensystem von  $U_i$  (für alle  $i \in I$ ), dann ist  $\bigcup_{i \in I} A_i$  ein Erzeugendensystem von  $\sum_{i \in I} U_i$ . ♣

Was kann man über die Dimension von  $U = U_1 + U_2$  sagen? Da  $U_1 \subset U$  und  $U_2 \subset U$ , gilt jedenfalls

$$\dim U \geq \max\{\dim U_1, \dim U_2\}$$

(und Gleichheit ist möglich, nämlich genau dann, wenn  $U_1 \subset U_2$  oder  $U_2 \subset U_1$ ). Wie groß kann  $\dim U$  höchstens werden? Wenn  $B_1$  eine Basis von  $U_1$  und  $B_2$  eine Basis von  $U_2$  ist, dann ist nach dem obigen Beispiel  $B_1 \cup B_2$  ein Erzeugendensystem von  $U$ , also gilt

$$\dim U \leq \#(B_1 \cup B_2) \leq \#B_1 + \#B_2 = \dim U_1 + \dim U_2.$$

Der folgende Satz gibt genauere Auskunft.

**BSP**  
Summen von  
Unter-VR

\* 18.4. **Satz.** Sei  $V$  ein Vektorraum mit Untervektorräumen  $U_1$  und  $U_2$ . Dann gilt

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

**SATZ**  
Dimension  
der Summe

Daran sieht man, dass  $\dim(U_1 + U_2) = \dim U_1 + \dim U_2$  genau dann gilt, wenn  $U_1$  und  $U_2$  den kleinstmöglichen Durchschnitt  $\{\mathbf{0}\}$  haben (vorausgesetzt, alle Dimensionen sind endlich).

*Beweis.* Ist  $\dim U_1 = \infty$  oder  $\dim U_2 = \infty$ , dann ist auch  $\dim(U_1 + U_2) = \infty$  (denn  $U_1, U_2 \subset U_1 + U_2$ ), und die Gleichung stimmt. Wir können also annehmen, dass  $U_1$  und  $U_2$  beide endlich-dimensional sind, etwa  $\dim U_1 = n_1$  und  $\dim U_2 = n_2$ . Sei  $m = \dim(U_1 \cap U_2) \leq \min\{n_1, n_2\}$ . Wir wählen eine Basis  $(b_1, b_2, \dots, b_m)$  von  $U_1 \cap U_2$ , die wir einerseits zu einer Basis  $(b_1, \dots, b_m, b'_{m+1}, b'_{m+2}, \dots, b'_{n_1})$  von  $U_1$  und andererseits zu einer Basis  $(b_1, \dots, b_m, b''_{m+1}, b''_{m+2}, \dots, b''_{n_2})$  von  $U_2$  ergänzen (Basisergänzungssatz mit Folgerung 8.16). Ich behaupte, dass

$$B = (b_1, \dots, b_m, b'_{m+1}, b'_{m+2}, \dots, b'_{n_1}, b''_{m+1}, b''_{m+2}, \dots, b''_{n_2})$$

eine Basis von  $U_1 + U_2$  ist. Daraus folgt die Gleichung im Satz, denn

$$\dim(U_1 + U_2) = \#B = m + (n_1 - m) + (n_2 - m) = n_1 + n_2 - m.$$

Es bleibt die Behauptung zu zeigen. Es ist klar, dass  $B$  ein Erzeugendensystem von  $U_1 + U_2$  ist, denn  $B$  enthält Erzeugendensysteme von  $U_1$  und von  $U_2$ . Wir müssen also noch nachweisen, dass  $B$  linear unabhängig ist. Seien also  $\lambda_i$  (für  $i \in \{1, 2, \dots, m\}$ ),  $\lambda'_i$  (für  $i \in \{m+1, \dots, n_1\}$ ) und  $\lambda''_i$  (für  $i \in \{m+1, \dots, n_2\}$ ) Skalare mit

$$\lambda_1 b_1 + \dots + \lambda_m b_m + \lambda'_{m+1} b'_{m+1} + \dots + \lambda'_{n_1} b'_{n_1} + \lambda''_{m+1} b''_{m+1} + \dots + \lambda''_{n_2} b''_{n_2} = \mathbf{0}.$$

Wir schreiben diese Gleichung als

$$u = \underbrace{\lambda_1 b_1 + \dots + \lambda_m b_m + \lambda'_{m+1} b'_{m+1} + \dots + \lambda'_{n_1} b'_{n_1}}_{\in U_1} = \underbrace{-\lambda''_{m+1} b''_{m+1} - \dots - \lambda''_{n_2} b''_{n_2}}_{\in U_2}.$$

Wir sehen, dass  $u \in U_1 \cap U_2$  ist, also ist  $u$  eine Linearkombination von  $b_1, \dots, b_m$ . Da  $b_1, \dots, b_m, b'_{m+1}, \dots, b'_{n_1}$  und  $b_1, \dots, b_m, b''_{m+1}, \dots, b''_{n_2}$  jeweils linear unabhängig sind (als Basen von  $U_1$  und  $U_2$ ), müssen

$$\lambda'_{m+1} = \dots = \lambda'_{n_1} = \lambda''_{m+1} = \dots = \lambda''_{n_2} = 0$$

sein; daraus folgt dann auch  $\lambda_1 = \dots = \lambda_m = 0$ . □

Man kann sich die Aussage ganz gut mit Hilfe der analogen Aussage über Kardinalitäten von Mengen merken:

$$\#(M_1 \cup M_2) + \#(M_1 \cap M_2) = \#M_1 + \#M_2.$$

Tatsächlich beruht obiger Beweis auf dieser Relation, wobei die Mengen Basen der vorkommenden Untervektorräume sind. Allerdings darf man diese Analogie auch nicht zu weit treiben: Die für Mengen gültige Relation

$$\begin{aligned} \#(M_1 \cup M_2 \cup M_3) + \#(M_1 \cap M_2) + \#(M_1 \cap M_3) + \#(M_2 \cap M_3) \\ = \#M_1 + \#M_2 + \#M_3 + \#(M_1 \cap M_2 \cap M_3) \end{aligned}$$

übersetzt sich nicht in eine analoge Dimensionsformel (Übung).

Besonders interessant ist der Fall  $U_1 \cap U_2 = \{\mathbf{0}\}$ .

18.5. **Lemma.** Sei  $V$  ein Vektorraum mit Untervektorräumen  $U_1$  und  $U_2$ . Dann sind die folgenden Aussagen äquivalent:

**LEMMA**  
Summe direkt

- (1)  $U_1 \cap U_2 = \{\mathbf{0}\}$ .
- (2) Jedes Element  $u \in U_1 + U_2$  lässt sich **eindeutig** schreiben als  $u = u_1 + u_2$  mit  $u_1 \in U_1$  und  $u_2 \in U_2$ .

Ist  $U_1 + U_2$  endlich-dimensional, dann sind beide Aussagen äquivalent zu

- (3)  $\dim(U_1 + U_2) = \dim U_1 + \dim U_2$ .

*Beweis.* Die Äquivalenz von (1) und (3) (unter der angegebenen Voraussetzung) folgt direkt aus der Dimensionsformel in Satz 18.4. Wir zeigen noch die Äquivalenz von (1) und (2).

Es gelte (1) und es sei  $u = u_1 + u_2 = u'_1 + u'_2$  mit  $u_1, u'_1 \in U_1$  und  $u_2, u'_2 \in U_2$ . Daraus folgt  $u_1 - u'_1 = u'_2 - u_2 \in U_1 \cap U_2 = \{\mathbf{0}\}$ , also  $u_1 = u'_1$  und  $u_2 = u'_2$ .

Jetzt gelte (2) und es sei  $u \in U_1 \cap U_2$ . Dann sind  $\mathbf{0} = \mathbf{0} + \mathbf{0} = u + (-u)$  zwei Darstellungen des Nullvektors; aus der Eindeutigkeit der Summendarstellung folgt also  $u = \mathbf{0}$ . □

18.6. **Definition.** Wenn die Aussagen in Lemma 18.5 gelten, dann heißt die Summe von  $U_1$  und  $U_2$  *direkt*. ◇

**DEF**  
direkte  
Summe  
zweier UVR

Eigenschaft (1) in Lemma 18.5 lässt sich auch so ausdrücken:

$$\forall u_1 \in U_1, u_2 \in U_2 : (u_1 + u_2 = \mathbf{0} \Rightarrow u_1 = u_2 = \mathbf{0}).$$

In dieser Form lässt sie sich verallgemeinern.

\* 18.7. **Definition.** Seien  $V$  ein Vektorraum und  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen von  $V$ . Dann heißt die Summe der  $U_i$  *direkt*, wenn für jede endliche Teilmenge  $J \subset I$  und beliebige Elemente  $u_i \in U_i$  für  $i \in J$  gilt

**DEF**  
direkte  
Summe

$$\sum_{i \in J} u_i = \mathbf{0} \implies \forall i \in J : u_i = \mathbf{0}.$$

Ist  $V = \sum_{i \in I} U_i$  und die Summe direkt, dann schreiben wir auch

$$V = \bigoplus_{i \in I} U_i$$

bzw.  $V = U_1 \oplus U_2 \oplus \dots \oplus U_n$ , wenn  $I = \{1, 2, \dots, n\}$  ist. ◇

Lemma 18.5 hat dann die folgende Verallgemeinerung.

18.8. **Lemma.** Seien  $V$  ein Vektorraum und  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen von  $V$ . Dann sind die folgenden Aussagen äquivalent:

**LEMMA**  
direkte  
Summe

- (1) Für jedes  $i \in I$  gilt  $U_i \cap \sum_{j \in I \setminus \{i\}} U_j = \{\mathbf{0}\}$ .
- (2) Die Summe der  $U_i$  ist direkt.

Ist  $I$  endlich und  $\sum_{i \in I} U_i$  endlich-dimensional, dann sind die Aussagen äquivalent zu

- (3)  $\dim \sum_{i \in I} U_i = \sum_{i \in I} \dim U_i$ .

*Beweis.* „(1)⇒(2)“: Sei  $J \subset I$  endlich und seien  $u_i \in U_i$  für  $i \in J$  mit  $\sum_{i \in J} u_i = \mathbf{0}$ . Sei  $i_0 \in J$ . Dann ist

$$u_{i_0} = \sum_{i \in J \setminus \{i_0\}} (-u_i) \in U_{i_0} \cap \sum_{i \in I \setminus \{i_0\}} U_i = \{\mathbf{0}\},$$

also ist  $u_{i_0} = \mathbf{0}$ . Da  $i_0 \in J$  beliebig war, müssen alle  $u_i = \mathbf{0}$  sein; damit ist die Summe direkt.

„(2)⇒(1)“: Sei  $i \in I$  und  $u \in U_i \cap \sum_{j \in I \setminus \{i\}} U_j$ . Dann gibt es  $J \subset I \setminus \{i\}$  endlich und Elemente  $u_j \in U_j$  für  $j \in J$ , sodass

$$\sum_{j \in J} u_j = u, \quad \text{also} \quad (-u) + \sum_{j \in J} u_j = \mathbf{0}$$

ist, wobei  $-u$  als Element von  $U_i$  betrachtet wird. Definition 18.7 besagt dann, dass  $u = \mathbf{0}$  sein muss.

„(2)⇒(3)“: Ist  $\sum_{i \in I} U_i$  endlich-dimensional, dann gilt das auch für alle  $U_i$  (denn sie sind in der Summe enthalten). Für jedes  $i \in I$  sei  $B_i$  eine Basis von  $U_i$ , dann ist  $B = \bigcup_{i \in I} B_i$  ein (endliches) Erzeugendensystem von  $\sum_{i \in I} U_i$ .

$B$  ist linear unabhängig: Sei  $B_i = \{b_{i1}, b_{i2}, \dots, b_{im_i}\}$  mit  $m_i = \dim U_i$  und seien  $\lambda_{ij}$  Skalare mit

$$\sum_{i \in I} \underbrace{\sum_{j=1}^{m_i} \lambda_{ij} b_{ij}}_{\in U_i} = \mathbf{0}.$$

Weil die Summe direkt ist, folgt  $\sum_{j=1}^{m_i} \lambda_{ij} b_{ij} = \mathbf{0}$  für alle  $i \in I$  und dann  $\lambda_{ij} = 0$  für alle  $j \in \{1, 2, \dots, m_i\}$ , weil  $B_i$  eine Basis ist. Als linear unabhängiges Erzeugendensystem ist  $B$  eine Basis von  $\sum_{i \in I} U_i$ , also gilt

$$\dim \sum_{i \in I} U_i = \#B = \sum_{i \in I} \#B_i = \sum_{i \in I} \dim U_i.$$

„(3)⇒(1)“: Es gilt

$$\begin{aligned} \sum_{i \in I} \dim U_i &= \dim \sum_{i \in I} U_i \leq \dim U_i + \dim \sum_{j \in I \setminus \{i\}} U_j \\ &\leq \dim U_i + \sum_{j \in I \setminus \{i\}} \dim U_j = \sum_{i \in I} \dim U_i, \end{aligned}$$

also muss überall Gleichheit herrschen. Aus Satz 18.4 folgt dann

$$U_i \cap \sum_{j \in I \setminus \{i\}} U_j = \{\mathbf{0}\}. \quad \square$$

**18.9. Beispiel.** Hier ist ein Beispiel, das zeigt, dass die vielleicht erst einmal näher liegende Version „ $\forall i, j \in I : i \neq j \Rightarrow U_i \cap U_j = \{\mathbf{0}\}$ “ für Bedingung (1) nicht ausreichend ist. **BSP**

Seien  $V = \mathbb{R}^2$ ,  $U_1 = \langle (1, 0) \rangle$ ,  $U_2 = \langle (0, 1) \rangle$  und  $U_3 = \langle (1, 1) \rangle$ . Dann gilt offenbar

$$U_1 \cap U_2 = U_1 \cap U_3 = U_2 \cap U_3 = \{(0, 0)\},$$

aber die Summe  $U_1 + U_2 + U_3$  ist *nicht* direkt. Zum Beispiel gilt

$$(0, 0) = (1, 0) + (0, 1) + (-1, -1)$$

als Summe je eines Elements von  $U_1$ ,  $U_2$  und  $U_3$ . ♣

Eine Zerlegung von  $V$  als direkte Summe,  $V = U_1 \oplus U_2$ , führt in natürlicher Weise zu zwei linearen Abbildungen  $\pi_1 : V \rightarrow U_1$  und  $\pi_2 : V \rightarrow U_2$ . Wir erhalten sie wie folgt: Jedes  $v \in V$  lässt sich eindeutig schreiben als  $v = u_1 + u_2$  mit  $u_1 \in U_1$  und  $u_2 \in U_2$ . Dann ist  $\pi_1(v) = u_1$  und  $\pi_2(v) = u_2$ . Diese Abbildungen sind linear, weil  $\lambda v = \lambda u_1 + \lambda u_2$  ist, und für  $v' = u'_1 + u'_2$  gilt  $v + v' = (u_1 + u'_1) + (u_2 + u'_2)$ . Außerdem sind  $\pi_1$  und  $\pi_2$  surjektiv, denn  $\pi_1|_{U_1} = \text{id}_{U_1}$  und  $\pi_2|_{U_2} = \text{id}_{U_2}$ .

**18.10. Definition.** Die Abbildungen  $\pi_1$  und  $\pi_2$  heißen die *Projektionen* von  $V$  auf  $U_1$  bzw.  $U_2$  bezüglich der Zerlegung  $V = U_1 \oplus U_2$ . **DEF**  
Projektionen  $\diamond$

Wenn wir mit  $p_1, p_2 : V \rightarrow V$  die Abbildungen bezeichnen, die durch  $p_i(v) = \pi_i(v)$  gegeben sind (sie unterscheiden sich von  $\pi_1$  und  $\pi_2$  nur durch den vergrößerten Wertebereich), dann gilt

$$p_1 \circ p_1 = p_1, \quad p_2 \circ p_2 = p_2 \quad \text{und} \quad p_1 + p_2 = \text{id}_V .$$

Umgekehrt gilt: Ist  $p : V \rightarrow V$  ein „Projektor“, d.h. eine lineare Abbildung mit  $p \circ p = p$ , dann gilt  $V = \text{im}(p) \oplus \text{ker}(p)$ , wobei  $\text{ker}(p) = \text{im}(\text{id}_V - p)$  ist (Übung). Mit  $p' = \text{id}_V - p$  gilt dann auch  $p' \circ p' = p'$  und  $p + p' = \text{id}_V$ .

Wir führen jetzt noch zwei Begriffe ein, die manchmal nützlich sind.

**\*** **18.11. Definition.** Seien  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum. Ein weiterer Untervektorraum  $U' \subset V$  heißt *komplementär* zu  $U$  oder ein *Komplement* von  $U$  in  $V$ , wenn  $V = U \oplus U'$  gilt. (Das bedeutet  $U \cap U' = \{0\}$  und  $U + U' = V$ .) **DEF**  
Komplement  $\diamond$

**18.12. Beispiele.** **BSP**  
Komplemente

- $V$  ist das einzige Komplement von  $\{0\}$  in  $V$  und  $\{0\}$  ist das einzige Komplement von  $V$  in  $V$ .
- Normalerweise gibt es aber viele Komplemente. Sei zum Beispiel  $V = \mathbb{R}^2$  und  $U = \langle (1, 0) \rangle \subset V$ . Dann sind die Komplemente von  $U$  gerade alle Untervektorräume der Form  $U' = \langle (a, 1) \rangle$  mit  $a \in \mathbb{R}$  beliebig.  $\clubsuit$

Gibt es immer ein Komplement?

**18.13. Satz.** Seien  $V$  ein endlich-dimensionaler Vektorraum und  $U \subset V$  ein Untervektorraum. Dann gibt es ein Komplement  $U'$  von  $U$  in  $V$ . Es gilt dann **SATZ**  
Existenz von  
Komplementen

$$\dim U + \dim U' = \dim V .$$

*Beweis.* Sei  $m = \dim U \leq \dim V = n$ . Wir wählen eine Basis  $(b_1, b_2, \dots, b_m)$  von  $U$  und ergänzen sie zu einer Basis  $(b_1, \dots, b_m, b_{m+1}, \dots, b_n)$  von  $V$ . Dann ist  $U' = \langle b_{m+1}, \dots, b_n \rangle$  ein Komplement von  $U$ :

$$U + U' = \langle b_1, \dots, b_m, b_{m+1}, \dots, b_n \rangle = V \quad \text{und} \quad U \cap U' = \{0\} ,$$

weil  $b_1, \dots, b_n$  linear unabhängig sind. Die Dimensionsformel folgt aus Lemma 18.5.  $\square$

Derselbe Beweis zeigt, dass es auch in beliebigen Vektorräumen stets Komplemente gibt, wenn man den Basisergänzungssatz für Mengen verwendet, der mit Hilfe des Zornschen Lemmas bewiesen wurde. Vergleiche die Diskussion nach Satz 8.14.

Wir sehen hier insbesondere, dass alle Komplemente von  $U$  dieselbe Dimension  $\dim V - \dim U$  haben. Das gilt ganz allgemein, auch wenn  $V$  unendliche Dimension hat.

**18.14. Lemma.** *Seien  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum. Seien weiter  $U'_1$  und  $U'_2$  zwei Komplemente von  $U$  in  $V$ . Dann sind  $U'_1$  und  $U'_2$  isomorph; insbesondere gilt  $\dim U'_1 = \dim U'_2$ .*

**LEMMA**  
Komplemente  
sind  
isomorph

*Beweis.* Wir betrachten die lineare Abbildung  $\phi = \pi \circ \iota : U'_1 \rightarrow U'_2$ ; dabei sei  $\iota : U'_1 \rightarrow V$  die Inklusionsabbildung und  $\pi : V \rightarrow U'_2$  die Projektion bezüglich der Zerlegung  $V = U \oplus U'_2$ . Dann gilt

$$\ker(\phi) = \ker(\pi) \cap U'_1 = U \cap U'_1 = \{0\},$$

also ist  $\phi$  injektiv. (Die erste Gleichheit folgt daraus, dass  $\iota$  injektiv ist.) Wir müssen noch zeigen, dass  $\phi$  auch surjektiv ist, dann ist  $\phi$  ein Isomorphismus und  $U'_1$  und  $U'_2$  sind isomorph. Sei dazu  $u_2 \in U'_2$ . Dann gibt es eindeutig bestimmte  $u \in U$  und  $u_1 \in U'_1$  mit  $u_2 = u + u_1$ . Wir können das auch als  $u_1 = (-u) + u_2$  lesen, woraus  $\phi(u_1) = u_2$  folgt.  $\square$

Damit ist folgende Definition sinnvoll.

\* **18.15. Definition.** Seien  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum, der ein Komplement  $U'$  in  $V$  hat. Dann heißt

**DEF**  
Kodimension

$$\text{codim}_V U = \dim U'$$

die *Kodimension* von  $U$  in  $V$ .  $\diamond$

Es gilt  $\dim U + \text{codim}_V U = \dim V$ : Ist die Kodimension klein, dann ist  $U$  „groß“, also nicht weit davon entfernt, ganz  $V$  zu sein.

**18.16. Beispiel.** Die Kodimension kann auch für unendlich-dimensionale Untervektorräume endlich sein. Sei zum Beispiel  $P$  der reelle Vektorraum der Polynomfunktionen, sei  $a \in \mathbb{R}$  und sei  $U_a = \{p \in P \mid p(a) = 0\} = \ker \text{ev}_a$ . Dann ist der eindimensionale Untervektorraum  $C = \langle x \mapsto 1 \rangle$  der konstanten Funktionen ein Komplement von  $U_a$  in  $P$  (für  $p \in P$  gilt eindeutig  $p = (p - p(a)) + p(a) \in U_a + C$ ), also ist  $\text{codim}_P U_a = 1$ . Dieselbe Überlegung zeigt, dass der Untervektorraum der in einem Punkt  $a$  verschwindenden Funktionen auch in anderen Funktionenräumen (alle Funktionen, stetige Funktionen,  $n$ -mal stetig differenzierbare Funktionen usw.) Kodimension 1 hat.

**BSP**  
Kodimension

Mit Polynomdivision (Satz 14.18) sieht man analog: Sind  $a_1, a_2, \dots, a_n \in \mathbb{R}$  paarweise verschieden und ist  $U_{a_1, \dots, a_n} = U_{a_1} \cap U_{a_2} \cap \dots \cap U_{a_n}$  der Untervektorraum der Polynomfunktionen, die in  $a_1, \dots, a_n$  verschwinden, dann ist der Untervektorraum  $P_{<n}$  der Polynomfunktionen vom Grad  $< n$  ein Komplement von  $U_{a_1, \dots, a_n}$  in  $P$ , also gilt  $\text{codim}_P U_{a_1, \dots, a_n} = n$ . Denn jedes Polynom  $p$  kann eindeutig geschrieben werden als

$$p(x) = q(x)(x - a_1)(x - a_2) \cdots (x - a_n) + r(x)$$

mit  $q \in P$  und  $r \in P_{<n}$ , und die Polynomfunktionen, die in  $a_1, \dots, a_n$  verschwinden, sind von der Form  $x \mapsto q(x)(x - a_1)(x - a_2) \cdots (x - a_n)$ .  $\clubsuit$

Der Begriff der Kodimension erlaubt eine etwas genauere Formulierung des „Rangsatzes“ 9.18. Zur Erinnerung: Der Satz besagt, dass für eine lineare Abbildung  $\phi : V \rightarrow W$  gilt

$$\dim \ker(\phi) + \operatorname{rk}(\phi) = \dim \ker(\phi) + \dim \operatorname{im}(\phi) = \dim V .$$

Wenn  $V$  unendlich-dimensional ist, dann ist das eine relativ schwache Aussage. Die folgende Version gibt zusätzliche Information, wenn der Rang von  $\phi$  endlich ist:

18.17. **Satz.** Sei  $\phi : V \rightarrow W$  eine lineare Abbildung mit  $\operatorname{rk}(\phi) < \infty$ . Dann gilt

$$\operatorname{codim}_V \ker(\phi) = \dim \operatorname{im}(\phi) = \operatorname{rk}(\phi) .$$

**SATZ**  
Rangsatz mit  
Kodimension

*Beweis.* Wir wählen eine Basis  $(b_1, b_2, \dots, b_m)$  von  $\operatorname{im}(\phi)$  (mit  $m = \dim \operatorname{im}(\phi)$ ). Seien weiter  $v_1, v_2, \dots, v_m \in V$  Urbilder von  $b_1, b_2, \dots, b_m$  unter  $\phi$ . Dann sind  $v_1, v_2, \dots, v_m$  linear unabhängig, denn aus

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_m v_m = \mathbf{0}$$

folgt

$$\mathbf{0} = \phi(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_m v_m) = \lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_m b_m$$

und damit  $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$ , weil  $b_1, b_2, \dots, b_m$  linear unabhängig sind. Wir setzen

$$U = \langle v_1, v_2, \dots, v_m \rangle \subset V ;$$

dann ist  $U$  ein Komplement von  $\ker(\phi)$  in  $V$  und  $\dim U = m$ , woraus die Behauptung folgt.

- $U + \ker(\phi) = V$ : Sei  $v \in V$ , dann gibt es Skalare  $\lambda_1, \dots, \lambda_m$  mit

$$\phi(v) = \lambda_1 b_1 + \dots + \lambda_m b_m .$$

Sei  $v' = v - (\lambda_1 v_1 + \dots + \lambda_m v_m)$ , dann ist

$$\phi(v') = \phi(v) - (\lambda_1 b_1 + \dots + \lambda_m b_m) = \mathbf{0} ,$$

also  $v' \in \ker(\phi)$  und  $v = v' + (\lambda_1 v_1 + \dots + \lambda_m v_m) \in \ker(\phi) + U$ .

- $U \cap \ker(\phi) = \{\mathbf{0}\}$ : Sei  $v \in U \cap \ker(\phi)$ , dann gibt es Skalare  $\lambda_1, \dots, \lambda_m$  mit  $v = \lambda_1 v_1 + \dots + \lambda_m v_m$ . Außerdem gilt

$$\mathbf{0} = \phi(v) = \lambda_1 b_1 + \dots + \lambda_m b_m ,$$

woraus  $\lambda_1 = \dots = \lambda_m = 0$  und damit  $v = \mathbf{0}$  folgt. □

Zum Abschluss dieses Abschnitts werden wir untersuchen, wann eine Zerlegung eines Vektorraums  $V$  als direkte Summe  $V = U_1 \oplus U_2 \oplus \dots \oplus U_n$  einer analogen Zerlegung eines Endomorphismus  $f$  von  $V$  entspricht. Dazu müssen wir erst einmal sagen, was Letzteres bedeutet.

**18.18. Lemma.** Sei  $V = U_1 \oplus U_2 \oplus \dots \oplus U_n$  eine Zerlegung eines Vektorraums  $V$  als direkte Summe von Untervektorräumen. Seien weiter

$$f_1 : U_1 \rightarrow U_1, \quad f_2 : U_2 \rightarrow U_2, \quad \dots, \quad f_n : U_n \rightarrow U_n$$

lineare Abbildungen. Dann gibt es einen eindeutig bestimmten Endomorphismus  $f$  von  $V$  mit  $f(v) = f_i(v)$  für alle  $i \in \{1, 2, \dots, n\}$  und alle  $v \in U_i$ .

**LEMMA**  
direkte  
Summe von  
Endo-  
morphis-  
men

*Beweis.* Jedes  $v \in V$  lässt sich eindeutig schreiben als  $v = u_1 + u_2 + \dots + u_n$  mit  $u_1 \in U_1, u_2 \in U_2, \dots, u_n \in U_n$ . Wenn  $f(u_i) = f_i(u_i)$  gelten soll, dann müssen wir  $f$  definieren durch

$$f(v) = f_1(u_1) + f_2(u_2) + \dots + f_n(u_n).$$

Dann gilt auch  $f(v) = f_i(v)$  für  $v \in U_i$ ; es bleibt nur noch zu zeigen, dass  $f$  linear ist. Das folgt aus

$$f = \iota_1 \circ f_1 \circ \pi_1 + \iota_2 \circ f_2 \circ \pi_2 + \dots + \iota_n \circ f_n \circ \pi_n,$$

wobei  $\pi_i : V \rightarrow U_i$  die Projektion auf  $U_i$  und  $\iota_i : U_i \rightarrow V$  die Inklusionsabbildung ist. (Alternativ kann man das auch wie vor Definition 18.10 nachrechnen.)  $\square$

**18.19. Definition.** Die Abbildung  $f$  in Lemma 18.18 heißt die *direkte Summe* von  $f_1, f_2, \dots, f_n$ ; wir schreiben

$$f = f_1 \oplus f_2 \oplus \dots \oplus f_n.$$

**DEF**  
direkte  
Summe von  
Endo-  
morphis-  
men

Wenn ein Endomorphismus  $f$  als direkte Summe von Endomorphismen  $f_i$  geschrieben werden kann, dann muss offenbar  $f(U_i) \subset U_i$  gelten. Wir geben dieser Eigenschaft einen Namen:

**18.20. Definition.** Seien  $V$  ein Vektorraum,  $U \subset V$  ein Untervektorraum und  $f$  ein Endomorphismus von  $V$ . Dann heißt  $U$  *f-invariant* oder *invariant unter  $f$* , wenn  $f(U) \subset U$  gilt.

**DEF**  
invarianter  
Untervektor-  
raum

**18.21. Satz.** Sei  $V = U_1 \oplus U_2 \oplus \dots \oplus U_n$  eine Zerlegung eines Vektorraums  $V$  als direkte Summe von Untervektorräumen. Sei weiter  $f \in \text{End}(V)$ . Dann lässt sich  $f$  genau dann als direkte Summe von Endomorphismen  $f_i \in \text{End}(U_i)$  schreiben, wenn alle Untervektorräume  $U_i$  invariant unter  $f$  sind.

**SATZ**  
Zerlegung  
von Endo-  
morphis-  
men

*Beweis.* Wir hatten schon gesehen, dass die Bedingung notwendig ist. Wir müssen noch zeigen, dass sie auch hinreichend ist. Es gelte also  $f(U_i) \subset U_i$  für alle  $i \in \{1, 2, \dots, n\}$ . Dann können wir  $f_i \in \text{End}(U_i)$  definieren durch  $f_i(u) = f(u)$  für  $u \in U_i$ ; damit gilt  $f = f_1 \oplus f_2 \oplus \dots \oplus f_n$ .  $\square$

18.22. **Beispiel.** Ist  $f$  ein Endomorphismus von  $V$  und  $\lambda$  ein Skalar, dann ist der Eigenraum  $E_\lambda(f)$  unter  $f$  invariant (denn  $f(v) = \lambda v$  für  $v \in E_\lambda(f)$ ).

Ist  $V$  endlich-dimensional und  $f \in \text{End}(V)$  diagonalisierbar mit paarweise verschiedenen Eigenwerten  $\lambda_1, \lambda_2, \dots, \lambda_m$ , dann gilt

$$V = E_{\lambda_1}(f) \oplus E_{\lambda_2}(f) \oplus \dots \oplus E_{\lambda_m}(f)$$

(denn  $V$  hat eine Basis aus Eigenvektoren von  $f$ ) und

$$f = f_1 \oplus f_2 \oplus \dots \oplus f_m$$

mit  $f_i = \lambda_i \text{id}_{E_{\lambda_i}(f)}$ . In diesem Fall lässt sich  $f$  also in besonders einfach gebaute Abbildungen zerlegen. 

**BSP**  
Zerlegung  
bei Diagona-  
lisierbarkeit

## 19. POLYNOME UND MATRIZEN BZW. ENDOMORPHISMEN

Als wir in der Linearen Algebra I über Polynome gesprochen haben, haben wir gesehen, dass man Elemente des zu Grunde liegenden Körpers in ein Polynom einsetzen kann (Definition 14.14). Man kann aber auch allgemeinere Objekte einsetzen, zum Beispiel Matrizen.

**19.1. Definition.** Sei  $K$  ein Körper, sei  $n \in \mathbb{N}$  und seien  $p \in K[X]$  ein Polynom und  $A \in \text{Mat}(n, K)$ . Wie üblich setzen wir  $A^0 = I_n$  (das Einselement des Matrizenrings  $\text{Mat}(n, K)$ ) und  $A^{k+1} = A \cdot A^k$  für  $k \in \mathbb{N}$ . Ist

$$p = a_0 + a_1X + a_2X^2 + \dots + a_mX^m,$$

dann definieren wir

$$p(A) = \sum_{j=0}^m a_j A^j = a_0 I_n + a_1 A + a_2 A^2 + \dots + a_m A^m \in \text{Mat}(n, K).$$

Ist  $V$  ein  $K$ -Vektorraum und  $f$  ein Endomorphismus von  $V$ , dann setzen wir analog  $f^0 = \text{id}_V$  und  $f^{\circ(k+1)} = f \circ f^{\circ k}$  und definieren

$$p(f) = \sum_{j=0}^m a_j f^{\circ j} = a_0 \text{id}_V + a_1 f + a_2 f^{\circ 2} + \dots + a_m f^{\circ m} \in \text{End}(V). \quad \diamond$$

Dann ist Folgendes klar: Ist  $A$  die Matrix von  $f$  bezüglich einer Basis  $B$  von  $V$ , dann ist  $p(A)$  die Matrix von  $p(f)$  bezüglich  $B$ .

Wir zeigen noch, dass die Abbildungen  $K[X] \rightarrow \text{Mat}(n, K)$ ,  $p \mapsto p(A)$ , und  $K[X] \mapsto \text{End}(V)$ ,  $p \mapsto p(f)$ , schöne Eigenschaften haben.

**19.2. Lemma.** *In der Situation von Definition 19.1 gilt für  $p, q \in K[X]$ :*

$$(p + q)(A) = p(A) + q(A) \quad \text{und} \quad (pq)(A) = p(A) \cdot q(A)$$

bzw.

$$(p + q)(f) = p(f) + q(f) \quad \text{und} \quad (pq)(f) = p(f) \circ q(f).$$

Außerdem ist  $1(A) = I_n$  und  $1(f) = \text{id}_V$ , wobei  $1$  das konstante Polynom  $1$  bezeichnet.

*Beweis.* Das folgt aus den Rechenregeln für Matrizen bzw. Endomorphismen (also daraus, dass  $\text{Mat}(n, K)$  und  $\text{End}(V)$  Ringe sind), zusammen mit

$$\lambda A = (\lambda I_n)A = A(\lambda I_n) \quad \text{und} \quad \lambda f = (\lambda \text{id}_V) \circ f = f \circ (\lambda \text{id}_V). \quad \square$$

Sind  $R$  und  $R'$  zwei Ringe, dann heißt eine Abbildung  $\phi : R \rightarrow R'$  ein *Ringhomomorphismus*, wenn  $\phi(1_R) = 1_{R'}$  und für alle  $r_1, r_2 \in R$  gilt  $\phi(r_1 + r_2) = \phi(r_1) + \phi(r_2)$  und  $\phi(r_1 r_2) = \phi(r_1) \phi(r_2)$  (d.h.,  $\phi$  bildet die Einselemente aufeinander ab und ist mit Addition und Multiplikation verträglich). Die Abbildungen  $K[X] \rightarrow \text{Mat}(n, R)$ ,  $p \mapsto p(A)$ , und  $K[X] \rightarrow \text{End}(V)$ ,  $p \mapsto p(f)$ , sind also Ringhomomorphismen. Weil sie dadurch gegeben sind, dass man etwas in die Polynome einsetzt, heißen sie auch *Einsetzungshomomorphismen*.

Mehr über Ringe im Allgemeinen und Polynomringe im Besonderen gibt es in der „Einführung in die Zahlentheorie und algebraische Strukturen“.

**DEF**  
Einsetzen  
von Matrizen  
und Endo-  
morphis-  
men  
in Polynome

**LEMMA**  
Einsetzungs-  
homomor-  
phismus

19.3. **Beispiel.** Sei  $A \in \text{Mat}(n, K)$ , sei  $v \in E_\lambda(A)$  und  $p \in K[X]$ . Dann gilt

$$p(A) \cdot v = p(\lambda)v.$$

**BSP**

Sei eine Matrix  $A \in \text{Mat}(n, K)$  gegeben. Dann kann man sich fragen, ob es stets ein Polynom  $p \in K[X]$  gibt mit  $p(A) = \mathbf{0}$  (und  $p \neq 0$ ), bzw. was man über die Menge solcher Polynome mit „Nullstelle“  $A$  aussagen kann.

Die erste Frage kann man leicht beantworten.

19.4. **Lemma.** Seien  $K$  ein Körper und  $A \in \text{Mat}(n, K)$ . Dann gibt es ein normiertes Polynom  $p \in K[X]$  mit  $\deg(p) \leq n^2$ , sodass  $p(A) = \mathbf{0}$  ist.

**LEMMA**

*Beweis.* Der Beweis ist eine schöne Anwendung grundlegender Resultate der linearen Algebra.  $\text{Mat}(n, K)$  ist ein  $K$ -Vektorraum der Dimension  $n^2$ , also müssen die  $n^2 + 1$  Elemente  $A^0, A^1, A^2, \dots, A^{n^2} \in \text{Mat}(n, K)$  linear abhängig sein. Es gibt also  $\lambda_0, \lambda_1, \dots, \lambda_{n^2} \in K$ , nicht alle null, mit  $\sum_{j=0}^{n^2} \lambda_j A^j = \mathbf{0}$ . Wir setzen  $m = \max\{j \mid \lambda_j \neq 0\}$ . Nach eventueller Multiplikation mit  $\lambda_m^{-1}$  können wir  $\lambda_m = 1$  annehmen. Die Behauptung folgt dann mit  $p = \sum_{j=0}^m \lambda_j X^j$ .  $\square$

Die zweite Frage kann man wie folgt beantworten:

19.5. **Lemma.** Seien  $K$  ein Körper und  $A \in \text{Mat}(n, K)$ . Sei

**LEMMA**

$$P(A) = \{p \in K[X] \mid p(A) = \mathbf{0}\}.$$

Dann gilt:

- (1)  $\mathbf{0} \in P(A)$ .
- (2) Aus  $p, q \in P(A)$  folgt  $p + q \in P(A)$ .
- (3) Aus  $p \in P(A)$ ,  $q \in K[X]$  folgt  $qp \in P(A)$ .

*Beweis.* Die erste Aussage ist klar. Die beiden anderen sieht man so:

$$(p+q)(A) = p(A) + q(A) = \mathbf{0} + \mathbf{0} = \mathbf{0} \quad \text{und} \quad (qp)(A) = q(A) \cdot p(A) = q(A) \cdot \mathbf{0} = \mathbf{0}.$$

 $\square$ 

Sei  $R$  ein kommutativer Ring. Eine Teilmenge  $I \subset R$  heißt ein *Ideal* von  $R$ , wenn  $I$  die obigen Eigenschaften hat:

$$0 \in I, \quad r, r' \in I \Rightarrow r + r' \in I, \quad r \in R, r' \in I \Rightarrow rr' \in I.$$

Ist  $\phi : R \rightarrow R'$  ein Ringhomomorphismus, dann zeigt derselbe Beweis wie oben, dass sein *Kern*  $\ker(\phi) = \{r \in R \mid \phi(r) = 0\}$  ein Ideal von  $R$  ist.

Die Tatsache, dass es im Polynomring eine „Division mit Rest“ gibt, führt zu einer einfachen Beschreibung von  $P(A)$ .

**19.6. Satz.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ . Dann gibt es ein eindeutig bestimmtes normiertes Polynom  $m_A \neq \mathbf{0}$  kleinsten Grades in  $P(A)$ , und

$$P(A) = \{qm_A \mid q \in K[X]\}$$

besteht genau aus den Vielfachen von  $m_A$ .

**SATZ**  
Minimal-  
polynom

*Beweis.* Nach Lemma 19.4 gibt es Polynome  $\neq \mathbf{0}$  in  $P(A)$ , also gibt es normierte (denn man kann durch den Leitkoeffizienten teilen) Polynome in  $P(A)$ . Sei  $m_A$  ein solches mit minimalem Grad. Ist  $p$  ein weiteres normiertes Polynom in  $P(A)$  mit  $\deg(p) = \deg(m_A)$ , dann ist  $p - m_A \in P(A)$  entweder das Nullpolynom (und damit  $p = m_A$ ) oder  $0 \leq \deg(p - m_A) < \deg(m_A)$ . Nach Division durch den Leitkoeffizienten würde man dann ein normiertes Polynom in  $P(A)$  mit kleinerem Grad als  $\deg(m_A)$  erhalten, im Widerspruch zur Wahl von  $m_A$ . Also ist  $m_A$  eindeutig bestimmt. Wir sehen auch, dass  $\deg(m_A) = \min\{\deg(p) \mid \mathbf{0} \neq p \in P(A)\}$  ist.

Nach Lemma 19.5 ist klar, dass  $\{qm_A \mid q \in K[X]\} \subset P(A)$  ist. Sei umgekehrt  $p \in P(A)$ . Dann gibt es  $q, r \in K[X]$  mit  $p = qm_A + r$  und  $\deg(r) < \deg(m_A)$  (Satz 14.18). Es folgt  $r \in P(A)$ , und wegen  $\deg(r) < \deg(m_A)$  muss  $r = \mathbf{0}$  sein. Damit ist  $p = qm_A$ , also hat man auch die umgekehrte Inklusion.  $\square$

**19.7. Definition.** Das Polynom  $m_A$  in Satz 19.6 heißt das *Minimalpolynom* von  $A$ . Ist  $f$  ein Endomorphismus eines endlich-dimensionalen Vektorraums, der bezüglich einer geeigneten Basis durch  $A$  beschrieben wird, dann heißt  $m_f = m_A$  das *Minimalpolynom* von  $f$ .  $\diamond$

**DEF**  
Minimal-  
polynom

Derselbe Beweis zeigt, dass jedes Ideal  $I$  von  $K[X]$  die Form  $I = \{pa \mid p \in K[X]\}$  hat mit einem geeigneten  $a \in K[X]$  ( $a = \mathbf{0}$  ist möglich, dann ist  $I = \{\mathbf{0}\}$ ). So ein Ideal, dessen Elemente genau die Vielfachen eines Elements  $a$  sind, heißt ein *Hauptideal*, und ein Ring, in dem jedes Ideal ein Hauptideal ist, ist ein *Hauptidealring*. Wir haben also gezeigt, dass  $K[X]$  ein solcher Hauptidealring ist. Ein anderes Beispiel für einen Hauptidealring ist der Ring  $\mathbb{Z}$  der ganzen Zahlen. Der Beweis ist im Wesentlichen der gleiche, nur dass man Division mit Rest von ganzen Zahlen verwendet statt der Polynomdivision.

**19.8. Beispiele.** Seien  $n \geq 1$  und  $A \in \text{Mat}(n, K)$ .

- (1) Sei  $A = \mathbf{0}$  die Nullmatrix. Dann ist  $m_A = X$ .
- (2) Ist  $A = \lambda I_n$ , dann ist  $m_A = X - \lambda$ . Die Umkehrung gilt ebenfalls.
- (3) Ist  $A$  diagonalisierbar und hat die (paarweise verschiedenen) Eigenwerte  $\lambda_1, \dots, \lambda_m$ , dann ist

$$m_A = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_m).$$

Das Minimalpolynom zerfällt also in Linearfaktoren und hat keine mehrfachen Nullstellen. Das sieht man so: Jedes Element  $v$  von  $K^n$  lässt sich schreiben als  $v = v_1 + \dots + v_m$  mit  $v_j \in E_{\lambda_j}(A)$ . Wegen  $Av_j = \lambda_j v_j$  ist dann mit  $p = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_m)$

$$\begin{aligned} p(A) \cdot v &= p(A) \cdot v_1 + p(A) \cdot v_2 + \dots + p(A) \cdot v_m \\ &= p(\lambda_1)v_1 + p(\lambda_2)v_2 + \dots + p(\lambda_m)v_m = \mathbf{0}, \end{aligned}$$

also ist  $p$  ein Vielfaches von  $m_A$ . Auf der anderen Seite gibt es für jedes  $j \in \{1, 2, \dots, m\}$  ein  $\mathbf{0} \neq v_j \in E_{\lambda_j}(A)$ , und es muss gelten

$$\mathbf{0} = m_A(A) \cdot v_j = m_A(\lambda_j)v_j, \quad \text{also ist } m_A(\lambda_j) = 0.$$

**BSP**  
Minimal-  
polynome

Das Minimalpolynom  $m_A$  muss also alle  $\lambda_j$  als Nullstellen haben und damit ist  $m_A$  ein Vielfaches von  $p$ . Weil beide normiert sind, müssen sie gleich sein.

Da das charakteristische Polynom von  $A$  die Form

$$(X - \lambda_1)^{e_1}(X - \lambda_2)^{e_2} \cdots (X - \lambda_m)^{e_m}$$

hat, wobei die  $e_j \geq 1$  (für  $j \in \{1, 2, \dots, m\}$ ) die algebraischen Vielfachheiten der Eigenwerte sind, sehen wir, dass das charakteristische Polynom ein Vielfaches des Minimalpolynoms ist. Hat  $A$   $n$  verschiedene Eigenwerte, dann sind die beiden Polynome gleich. ♣

**19.9. Definition.** Wir schreiben  $\chi_A = \det(XI_n - A)$  für das charakteristische Polynom einer Matrix  $A \in \text{Mat}(n, K)$  und analog  $\chi_f$  für das charakteristische Polynom eines Endomorphismus  $f$  eines endlich-dimensionalen Vektorraums.  $\diamond$

**DEF**  
 $\chi_A$

Das dritte Beispiel oben wirft zwei Fragen auf:

- Für  $A$  diagonalisierbar gilt, dass  $\chi_A$  ein Vielfaches von  $m_A$  ist. Ist das auch allgemein richtig?
- Gilt die Umkehrung der Beobachtung im Beispiel: Wenn  $m_A$  in Linearfaktoren zerfällt und keine mehrfachen Nullstellen hat, muss dann  $A$  diagonalisierbar sein?

Die erste Frage wird durch den Satz von Cayley-Hamilton beantwortet, den wir bald beweisen werden. Die zweite können wir gleich behandeln.

\* **19.10. Satz.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ . Dann ist  $A$  genau dann diagonalisierbar, wenn das Minimalpolynom  $m_A$  in Linearfaktoren zerfällt und keine mehrfachen Nullstellen hat.

**SATZ**  
Kriterium  
für Diagonalisierbarkeit

Die analoge Aussage gilt natürlich auch für Endomorphismen endlich-dimensionaler Vektorräume.

*Beweis.* Es ist nur noch die Rückrichtung zu zeigen. Es gelte also

$$m_A = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_m)$$

mit paarweise verschiedenen  $\lambda_1, \lambda_2, \dots, \lambda_m \in K$ . Für  $j \in \{1, 2, \dots, m\}$  sei

$$p_j = \prod_{i \neq j} \frac{X - \lambda_i}{\lambda_j - \lambda_i} \in K[X].$$

Dann gilt  $p_j(\lambda_i) = \delta_{ij}$  und  $\deg(p_j) \leq m - 1$ . Für die Summe  $p = p_1 + p_2 + \dots + p_m$  gilt also  $p(\lambda_j) = 1$  für alle  $j$  und  $\deg(p) < m$ . Es folgt  $p = 1$ , denn das Polynom  $p - 1$  vom Grad  $< m$  hat mindestens  $m$  Nullstellen. Außerdem ist  $(X - \lambda_j)p_j$  ein Vielfaches von  $m_A$ , also ist  $(A - \lambda_j I_n)p_j(A) = \mathbf{0}$ .

Sei  $U_j = \text{im}(p_j(A)) \subset K^n$ . Dann gelten die folgenden Aussagen:

(1)  $U_j \subset E_{\lambda_j}(A)$ : Sei  $u \in U_j$ , also  $u = p_j(A) \cdot v$  für ein  $v \in K^n$ . Es folgt

$$\mathbf{0} = (A - \lambda_j I_n)p_j(A) \cdot v = (A - \lambda_j I_n) \cdot u, \quad \text{also} \quad A \cdot u = \lambda_j u.$$

(2)  $U_1 + U_2 + \dots + U_m = K^n$ : Sei  $v \in K^n$ . Dann gilt

$$v = p(A) \cdot v = p_1(A) \cdot v + p_2(A) \cdot v + \dots + p_m(A) \cdot v \in U_1 + U_2 + \dots + U_m.$$

Außerdem wissen wir ganz allgemein, dass die Summe der Eigenräume  $E_{\lambda_j}(A)$  direkt ist (Satz 14.8). Zusammen mit den obigen beiden Aussagen folgt daraus  $K^n = E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus \dots \oplus E_{\lambda_m}(A)$ , also ist  $A$  diagonalisierbar.  $\square$

Damit wir dieses Kriterium anwenden können, müssen wir in der Lage sein, das Minimalpolynom zu bestimmen. Dabei wird uns die Antwort auf die erste Frage von oben helfen. Wir betrachten erst einmal  $2 \times 2$ -Matrizen (für  $1 \times 1$ -Matrizen ist die Aussage trivialerweise richtig).

19.11. **Beispiel.** Sei  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{Mat}(2, K)$ . Dann ist

$$A^2 = \begin{pmatrix} a^2 + bc & (a+d)b \\ (a+d)c & d^2 + bc \end{pmatrix} = (a+d)A - (ad - bc)I_2,$$

also  $\chi_A(A) = \mathbf{0}$ , denn  $\chi_A = X^2 - (a+d)X + (ad - bc)$ .  $\clubsuit$

**BSP**  
Cayley-  
Hamilton  
für  $n = 2$

Es ist jetzt vielleicht nicht mehr überraschend, dass das auch für beliebige  $n \times n$ -Matrizen gilt.

\* 19.12. **Satz.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ . Dann ist

$$\chi_A(A) = \mathbf{0}.$$

**SATZ**  
Cayley-  
Hamilton

Warum ist der folgende „Beweis“ nicht korrekt?

$$\chi_A(X) = \det(XI_n - A) \implies \chi_A(A) = \det(AI_n - A) = \det(\mathbf{0}) = 0.$$

Wenn wir an dieser Stelle die richtigen Hilfsmittel zur Verfügung hätten, dann könnten wir so argumentieren: Die Aussage gilt für diagonalisierbare Matrizen nach Beispiel 19.8(3). Diagonalisierbare Matrizen sind „dicht“ (in einem geeigneten Sinn) in allen  $n \times n$ -Matrizen; die Aussage folgt dann, weil die Abbildung  $\text{Mat}(n, A) \rightarrow \text{Mat}(n, A)$ ,  $A \mapsto \chi_A(A)$  stetig ist (wiederum in einem geeigneten Sinn). Das wird im Kleingedruckten unten genauer erklärt.

*Beweis.* Wir gehen hier anders vor. Wir stellen erst einmal fest, dass wir auch mit Matrizen über kommutativen Ringen (statt über Körpern) rechnen können, solange wir nicht durch Ringelemente dividieren müssen. Insbesondere können wir Determinanten bilden und damit auch die adjungierte Matrix (siehe Definition 13.12). Wir wenden das an auf den Ring  $K[X]$ , für den wir schon mit Determinanten gearbeitet haben, denn  $\chi_A$  ist ja definiert als  $\det(XI_n - A)$ . Sei also  $B = XI_n - A \in \text{Mat}(n, K[X])$  und  $\tilde{B}$  die adjungierte Matrix zu  $B$  (deren Einträge bis aufs Vorzeichen Determinanten von  $(n-1) \times (n-1)$ -Untermatrizen von  $B$  sind). Dann gilt (Satz 13.13)

$$(19.1) \quad \tilde{B}(XI_n - A) = \tilde{B}B = \det(B)I_n = \chi_A(X)I_n.$$

Die Einträge von  $\tilde{B}$  sind Polynome in  $X$  vom Grad  $< n$ , wie man sich leicht überlegt; wir können also schreiben

$$\tilde{B} = \left( b_{ij}^{(n-1)} X^{n-1} + \dots + b_{ij}^{(1)} X + b_{ij}^{(0)} \right)_{i,j}$$

mit geeigneten Koeffizienten  $b_{ij}^{(k)} \in K$ . (Wir schreiben den oberen Index in Klammern, um eine Verwechslung mit Potenzen zu vermeiden.) Nach den Rechenregeln für Matrizen können wir das auch schreiben als

$$\tilde{B} = \tilde{B}_{n-1} X^{n-1} + \dots + \tilde{B}_1 X + \tilde{B}_0$$

mit  $\tilde{B}_k = (b_{ij}^{(k)})_{i,j} \in \text{Mat}(n, K)$ . Sei außerdem

$$\chi_A(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0.$$

Wir setzen in (19.1) ein und erhalten

$$\begin{aligned} \tilde{B}_{n-1}X^n + (\tilde{B}_{n-2} - \tilde{B}_{n-1}A)X^{n-1} + \dots + (\tilde{B}_1 - \tilde{B}_2A)X^2 + (\tilde{B}_0 - \tilde{B}_1A)X - \tilde{B}_0A \\ = I_nX^n + a_{n-1}I_nX^{n-1} + \dots + a_2I_nX^2 + a_1I_nX + a_0I_n. \end{aligned}$$

Koeffizientenvergleich zeigt dann

$$\tilde{B}_{n-1} = I_n, \tilde{B}_{n-2} - \tilde{B}_{n-1}A = a_{n-1}I_n, \dots, \tilde{B}_0 - \tilde{B}_1A = a_1I_n, -\tilde{B}_0A = a_0I_n.$$

Wir multiplizieren diese Gleichungen von rechts mit  $A^n, A^{n-1}, \dots, A, I_n$  und summieren auf:

$$\begin{array}{rcl} \tilde{B}_{n-1}A^n & & = A^n \\ -\tilde{B}_{n-1}A^n + \tilde{B}_{n-2}A^{n-1} & & = a_{n-1}A^{n-1} \\ \quad \ddots & \ddots & \vdots \\ \quad \quad \quad \ddots & \ddots & \vdots \\ \quad \quad \quad \quad \quad -\tilde{B}_1A^2 + \tilde{B}_0A & = & a_1A \\ \quad \quad \quad \quad \quad \quad -\tilde{B}_0A & = & a_0I_n \\ \hline & & \mathbf{0} = \chi_A(A) \end{array}$$

□

Ein Beweis, wie er oben angedeutet wurde, könnte etwa wie folgt aussehen. Wir zeigen die Aussage erst einmal für Matrizen über  $\mathbb{C}$ . Sei also  $A \in \text{Mat}(n, \mathbb{C})$ . Ist  $A$  diagonalisierbar, dann ist  $\chi_A(A) = \mathbf{0}$  und wir sind fertig. Sonst können wir  $A$  beliebig wenig stören, sodass das charakteristische Polynom der gestörten Matrix  $A'$  keine mehrfachen Nullstellen hat. Dann ist  $A'$  diagonalisierbar, also gilt  $\chi_{A'}(A') = \mathbf{0}$ . Da  $\chi_A(A)$  eine stetige Funktion von  $A$  ist (d.h., die Einträge dieser Matrix hängen stetig von den Einträgen von  $A$  ab), folgt  $\chi_A(A) = \mathbf{0}$ .

Nun gilt ganz allgemein, dass die Einträge von  $\chi_A(A)$  Polynome in den Einträgen von  $A$  mit ganzzahligen Koeffizienten sind (diese Polynome haben den Grad  $n$  in dem Sinne, dass sie ganzzahlige Linearkombinationen von Produkten von jeweils  $n$  Einträgen von  $A$  sind) — daraus folgt auch die oben schon verwendete Stetigkeit. Wir haben gesehen, dass diese Polynome stets den Wert null annehmen, wenn man beliebige komplexe Zahlen einsetzt. Daraus folgt aber, dass die Polynome null sind (als Polynome). Das zeigt dann, dass  $\chi_A(A) = \mathbf{0}$  für Matrizen über beliebigen Körpern (oder sogar kommutativen Ringen) gilt.

**19.13. Folgerung.** *Das charakteristische Polynom  $\chi_A$  einer Matrix  $A$  ist ein Vielfaches des Minimalpolynoms  $m_A$ . Jede Nullstelle von  $\chi_A$  ist auch eine Nullstelle von  $m_A$ .*

**FOLG**  
 $m_A$  teilt  $\chi_A$

*Beweis.* Jedes Polynom  $p \in K[X]$  mit  $p(A) = \mathbf{0}$  ist ein Vielfaches von  $m_A$ , vgl. Satz 19.6, und  $\chi_A(A) = \mathbf{0}$ .

Ist  $\lambda$  eine Nullstelle von  $\chi_A$ , dann ist  $\lambda$  ein Eigenwert von  $A$ , es gibt also einen Eigenvektor  $\mathbf{0} \neq v \in K^n$  mit  $A \cdot v = \lambda v$ . Es folgt

$$\mathbf{0} = m_A(A) \cdot v = m_A(\lambda)v,$$

wegen  $v \neq \mathbf{0}$  also  $m_A(\lambda) = 0$ .

□

19.14. **Beispiel.** Folgerung 19.13 hilft uns, das Minimalpolynom zu bestimmen. Sei zum Beispiel

**BSP**  
Minimal-  
polynom

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix} \in \text{Mat}(3, \mathbb{R}).$$

Es ist

$$\chi_A = \begin{vmatrix} X & -1 & 0 \\ 1 & X-2 & 0 \\ -1 & 1 & X-1 \end{vmatrix} = (X(X-2)+1)(X-1) = (X-1)^3,$$

also ist  $m_A \in \{X-1, (X-1)^2, (X-1)^3\}$ . Wir probieren die Möglichkeiten der Reihe nach durch:

$$(X-1)(A) = A - I_3 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

ist nicht die Nullmatrix, aber  $(X-1)^2(A) = (A - I_3)^2 = \mathbf{0}$ , also ist  $m_A = (X-1)^2$  (und  $A$  ist nicht diagonalisierbar). ♣

Wir werden jetzt sehen, wie man eine Zerlegung  $p = p_1 p_2 \cdots p_m$  eines Polynoms  $p$  mit  $p(A) = \mathbf{0}$  in eine Zerlegung des zugehörigen Endomorphismus von  $K^n$  „übersetzen“ kann. Dazu müssen die Faktoren  $p_1, p_2, \dots, p_m$  allerdings eine Bedingung erfüllen.

19.15. **Definition.** Sei  $K$  ein Körper und seien  $p, q \in K[X]$ . Dann heißen  $p$  und  $q$  *relativ prim* oder *teilerfremd*, wenn es Polynome  $r, s \in K[X]$  gibt mit  $rp + sq = 1$ . ◇

**DEF**  
relativ prim

Man kann zeigen, dass zwei Polynome  $p$  und  $q$ , die keinen (nicht konstanten) gemeinsamen Teiler haben (es gilt also für alle  $a, b, c \in K[X]$ , dass aus  $p = ab$  und  $q = ac$  folgt, dass  $a$  konstant ist), relativ prim sind. Die Umkehrung ist einfach zu sehen. Das erklärt die Bezeichnung „teilerfremd“. Siehe die „Einführung in die Zahlentheorie und algebraische Strukturen“.

Hier ist eine typische Anwendung:

19.16. **Lemma.** Seien  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum und  $f$  ein Endomorphismus von  $V$ . Sei weiter  $p \in K[X]$  mit  $p(f) = \mathbf{0}$ ; es gebe eine Faktorisierung  $p = p_1 p_2$  mit  $p_1, p_2 \in K[X]$  relativ prim. Wir setzen  $f_1 = p_1(f)$  und  $f_2 = p_2(f)$ . Dann gilt:

**LEMMA**  
Zerlegung  
von Endo-  
morphis-  
men

- (1)  $U_1 = \text{im}(f_1)$  und  $U_2 = \text{im}(f_2)$  sind  $f$ -invariant.
- (2)  $U_1 = \text{ker}(f_2)$  und  $U_2 = \text{ker}(f_1)$ .
- (3)  $V = U_1 \oplus U_2$ .

*Beweis.* Nach Voraussetzung gibt es Polynome  $r_1, r_2 \in K[X]$  mit  $r_1 p_1 + r_2 p_2 = 1$ .

- (1) Sei  $u \in U_1$ , dann gibt es  $v \in V$  mit  $u = f_1(v)$ . Es folgt

$$f(u) = (f \circ f_1)(v) = ((Xp_1)(f))(v) = ((p_1 X)(f))(v) = f_1(f(v)) \in U_1,$$

also ist  $U_1$  unter  $f$  invariant. Die Aussage für  $U_2$  zeigt man analog.

(2) Es gilt  $f_2(u) = (f_2 \circ f_1)(v) = \mathbf{0}$  (denn  $f_2 \circ f_1 = p(f) = \mathbf{0}$ ), also ist  $U_1 \subset \ker(f_2)$ . Sei umgekehrt  $u \in \ker(f_2)$ . Dann gilt

$$u = ((p_1 r_1 + r_2 p_2)(f))(u) = f_1((r_1(f))(u)) + (r_2(f))(f_2(u)) = f_1((r_1(f))(u)) \in U_1,$$

also ist auch  $\ker(f_2) \subset U_1$ . Die Aussage für  $U_2$  zeigt man analog.

(3) Für  $v \in V$  gilt

$$v = \text{id}_V(v) = ((p_1 r_1 + p_2 r_2)(f))(v) = f_1((r_1(f))(v)) + f_2((r_2(f))(v)) \in U_1 + U_2.$$

Ist andererseits  $v \in U_1 \cap U_2 = \ker(f_2) \cap \ker(f_1)$ , dann ist  $f_2(v) = f_1(v) = \mathbf{0}$  und damit

$$v = \text{id}_V(v) = ((r_1 p_1 + r_2 p_2)(f))(v) = r_1(f)(f_1(v)) + r_2(f)(f_2(v)) = \mathbf{0}.$$

Damit gilt  $V = U_1 + U_2$  und  $U_1 \cap U_2 = \{\mathbf{0}\}$ . □

Um das anwenden zu können, brauchen wir Beispiele für relativ prime Polynome.

**19.17. Lemma.** *Seien  $K$  ein Körper,  $\lambda \in K$  und  $p \in K[X]$  mit  $p(\lambda) \neq 0$ . Dann sind die Polynome  $X - \lambda$  und  $p$  relativ prim.*

**LEMMA**  
relativ prime  
Polynome 1

*Beweis.* Wir schreiben  $p = q \cdot (X - \lambda) + \alpha$  mit  $\alpha = p(\lambda) \neq 0$  und  $q \in K[X]$ . Dann gilt

$$-\alpha^{-1} q \cdot (X - \lambda) + \alpha^{-1} \cdot p = 1,$$

also sind  $X - \lambda$  und  $p$  relativ prim. □

**19.18. Lemma.** *Seien  $K$  ein Körper und  $p, q_1, q_2 \in K[X]$ , sodass  $p, q_1$  und  $p, q_2$  jeweils relativ prim sind. Dann sind auch  $p$  und  $q_1 q_2$  relativ prim.*

**LEMMA**  
relativ prime  
Polynome 2

*Beweis.* Nach Voraussetzung gibt es Polynome  $r_1, s_1, r_2, s_2 \in K[X]$  mit

$$r_1 p + s_1 q_1 = 1 \quad \text{und} \quad r_2 p + s_2 q_2 = 1.$$

Multiplikation liefert

$$1 = (r_1 p + s_1 q_1)(r_2 p + s_2 q_2) = (r_1 r_2 p + r_1 s_2 q_2 + s_1 q_1 r_2) p + s_1 s_2 \cdot q_1 q_2,$$

also sind  $p$  und  $q_1 q_2$  relativ prim. □

**19.19. Folgerung.** *Sei  $K$  ein Körper, seien  $\lambda_1, \lambda_2, \dots, \lambda_m \in K$  paarweise verschieden und seien  $e_1, e_2, \dots, e_m \geq 1$ . Dann sind die Polynome  $(X - \lambda_j)^{e_j}$  für  $j \in \{1, 2, \dots, m\}$  paarweise relativ prim.*

**FOLG**

*Beweis.* Seien  $i, j \in \{1, 2, \dots, m\}$  mit  $i \neq j$ . Nach Lemma 19.17 sind  $X - \lambda_i$  und  $(X - \lambda_j)^{e_j}$  relativ prim (wir wenden das Lemma an mit  $\lambda = \lambda_i$  und  $p = (X - \lambda_j)^{e_j}$ ). Eine einfache Induktion unter Verwendung von Lemma 19.18 zeigt dann, dass auch  $(X - \lambda_i)^{e_i}$  und  $(X - \lambda_j)^{e_j}$  relativ prim sind (Anwendung mit  $p = (X - \lambda_j)^{e_j}$ ,  $q_1 = (X - \lambda_i)^{e_i}$ ,  $q_2 = X - \lambda_i$ ). □

20. DIE JORDANSCHEN NORMALFORM

Nach den Vorbereitungen aus dem letzten Abschnitt können wir den wesentlichen Schritt für die Klassifikation von Endomorphismen formulieren.

**20.1. Satz.** *Seien  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum und  $f \in \text{End}(V)$ . Seien weiter  $\lambda_1, \lambda_2, \dots, \lambda_m \in K$  paarweise verschieden und  $e_1, e_2, \dots, e_m \geq 1$ , sodass  $p(f) = \mathbf{0}$  ist, wobei*

$$p = (X - \lambda_1)^{e_1} (X - \lambda_2)^{e_2} \cdots (X - \lambda_m)^{e_m}.$$

Dann gibt es  $f$ -invariante Untervektorräume  $U_1, U_2, \dots, U_m$  von  $V$  mit

- (1)  $V = U_1 \oplus U_2 \oplus \dots \oplus U_m$ , und
- (2) für jedes  $j \in \{1, 2, \dots, m\}$  gilt  $(f - \lambda_j \text{id}_V)^{e_j}(U_j) = \{\mathbf{0}\}$ ; genauer ist  $U_j = \ker((f - \lambda_j \text{id}_V)^{e_j})$ .

**SATZ**  
Zerlegung  
von Endo-  
morphis-  
men

Das verallgemeinert die (interessante Richtung in der) Aussage von Satz 19.10: Sind alle Exponenten  $e_j = 1$ , dann erhält man die Zerlegung in Eigenräume wie dort.

*Beweis.* Induktion über  $m$ . Der Fall  $m = 1$  ist klar (und der Fall  $m = 0$  ist trivial; dann muss  $V = \{\mathbf{0}\}$  sein). Sei die Aussage für ein  $m$  bewiesen; wir wollen sie für  $m + 1$  zeigen. Dazu setzen wir

$$p_1 = (X - \lambda_1)^{e_1} (X - \lambda_2)^{e_2} \cdots (X - \lambda_m)^{e_m} \quad \text{und} \quad p_2 = (X - \lambda_{m+1})^{e_{m+1}}.$$

Nach Folgerung 19.19 und Lemma 19.18 sind  $p_1$  und  $p_2$  relativ prim, also können wir Lemma 19.16 anwenden. Das liefert uns  $f$ -invariante Untervektorräume  $U_{m+1}$  und  $U$  von  $V$  mit

$$U_{m+1} = \text{im}(p_1(f)) = \ker((f - \lambda_{m+1} \text{id}_V)^{e_{m+1}})$$

und

$$U = \text{im}((f - \lambda_{m+1} \text{id}_V)^{e_{m+1}}) = \ker(p_1(f)),$$

die  $f$ -invariant sind und  $V = U \oplus U_{m+1}$  erfüllen. Sei jetzt  $\tilde{f} \in \text{End}(U)$  die Einschränkung von  $f$  auf  $U$ . Dann gilt  $p_1(\tilde{f}) = \mathbf{0}$  (denn  $U = \ker(p_1(f))$ ), und wir können die Induktionsannahme anwenden. Das liefert  $\tilde{f}$ - und damit auch  $f$ -invariante Untervektorräume  $U_1, U_2, \dots, U_m$  von  $U$  (und damit von  $V$ ) mit  $U_j = \ker((\tilde{f} - \lambda_j \text{id}_U)^{e_j})$  und  $U_1 \oplus U_2 \oplus \dots \oplus U_m = U$ . Damit gilt bereits

$$V = U \oplus U_{m+1} = (U_1 \oplus U_2 \oplus \dots \oplus U_m) \oplus U_{m+1} = U_1 \oplus U_2 \oplus \dots \oplus U_m \oplus U_{m+1}.$$

Es bleibt zu zeigen, dass  $U_j = \ker((f - \lambda_j \text{id}_V)^{e_j})$  ist für  $j \leq m$ . Es gilt jedenfalls  $U_j = \ker((f - \lambda_j \text{id}_V)^{e_j}) \cap U$ , sodass es genügt zu zeigen, dass der Kern in  $U$  enthalten ist. Das folgt aber aus

$$\begin{aligned} v \in \ker((f - \lambda_j \text{id}_V)^{e_j}) &\implies ((X - \lambda_j)^{e_j}(f))(v) = \mathbf{0} \implies (p_1(f))(v) = \mathbf{0} \\ &\implies v \in \ker(p_1(f)) = U. \end{aligned} \quad \square$$

Das legt folgende Definition nahe. Wir bemerken zuerst, dass für zwei lineare Abbildungen  $V_1 \xrightarrow{f} V_2 \xrightarrow{g} V_3$  gilt  $\ker(f) \subset \ker(g \circ f)$  (denn aus  $f(v) = \mathbf{0}$  folgt  $(g \circ f)(v) = g(\mathbf{0}) = \mathbf{0}$ ). Insbesondere ist für  $f \in \text{End}(V)$

$$\{\mathbf{0}\} = \ker(\text{id}_V) \subset \ker(f) \subset \ker(f^{\circ 2}) \subset \ker(f^{\circ 3}) \subset \dots$$

eine aufsteigende Kette von Untervektorräumen von  $V$ .

\* **20.2. Definition.** Sei  $f$  ein Endomorphismus eines endlich-dimensionalen  $K$ -Vektorraums  $V$  und sei  $\lambda \in K$ . Dann heißt der Untervektorraum

**DEF**  
Hauptraum

$$H_\lambda(f) = \bigcup_{m \in \mathbb{N}} \ker((f - \lambda \operatorname{id}_V)^{\circ m}) \subset V$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von  $f$  zum Eigenwert  $\lambda$ . Analog definieren wir  $H_\lambda(A)$  für Matrizen  $A \in \operatorname{Mat}(n, K)$ .  $\diamond$

Nach Lemma 7.4 ist  $H_\lambda(f)$  als aufsteigende Vereinigung von Untervektorräumen ein Untervektorraum von  $V$ .

**20.3. Lemma.** In der Situation von Satz 20.1 oben gilt  $U_j = H_{\lambda_j}(f)$  für alle  $j \in \{1, 2, \dots, m\}$ .

**LEMMA**  
Zerlegung  
in Haupt-  
räume

*Beweis.* Da  $U_j = \ker((f - \lambda_j \operatorname{id}_V)^{\circ e_j})$  ist, müssen wir zeigen, dass für  $m \geq e_j$  aus  $(f - \lambda_j \operatorname{id}_V)^{\circ m}(v) = \mathbf{0}$  schon  $v \in U_j$  folgt. Sei  $q = p/(X - \lambda_j)^{e_j}$  (also  $p$  ohne den  $j$ -ten Faktor), dann sind  $q$  und  $(X - \lambda_j)^{m-e_j}$  relativ prim wie im Beweis von Satz 20.1. Es gibt also Polynome  $r, s \in K[X]$  mit  $r(X - \lambda_j)^{m-e_j} + sq = 1$  und damit

$$r(X - \lambda_j)^m + sp = (X - \lambda_j)^{e_j}.$$

Einsetzen von  $f$  und Anwenden auf  $v$  ergibt

$$\begin{aligned} (f - \lambda_j \operatorname{id}_V)^{\circ e_j}(v) &= (r(f))((f - \lambda_j \operatorname{id}_V)^{\circ m}(v)) + (s(f))((p(f))(v)) \\ &= (r(f))(\mathbf{0}) + (s(f))(\mathbf{0}) = \mathbf{0} \end{aligned}$$

(unter Verwendung von  $(f - \lambda_j \operatorname{id}_V)^{\circ m}(v) = \mathbf{0}$  und  $p(f) = \mathbf{0}$ ) wie gewünscht.  $\square$

Da  $V$  endlich-dimensional ist, kann die aufsteigende Kette von Untervektorräumen  $\ker((f - \lambda \operatorname{id}_V)^{\circ m})$  nicht unendlich oft echt aufsteigen, also muss es ein  $m \in \mathbb{N}$  geben mit

$$\ker((f - \lambda \operatorname{id}_V)^{\circ m}) = \ker((f - \lambda \operatorname{id}_V)^{\circ(m+1)}) = \ker((f - \lambda \operatorname{id}_V)^{\circ(m+2)}) = \dots$$

Auf  $H_\lambda(f)$  können wir also  $f$  schreiben als  $\lambda \operatorname{id}_{H_\lambda(f)} + g$  mit  $g^{\circ m} = \mathbf{0}$ . Diese Eigenschaft von  $g$  hat einen Namen.

\* **20.4. Definition.** Sei  $f$  ein Endomorphismus eines Vektorraums  $V$ .  $f$  heißt *nilpotent*, wenn es  $m \in \mathbb{N}$  gibt mit  $f^{\circ m} = \mathbf{0}$ . Analog heißt eine Matrix  $A \in \operatorname{Mat}(n, K)$  *nilpotent*, wenn es  $m \in \mathbb{N}$  gibt mit  $A^m = \mathbf{0}$ .  $\diamond$

**DEF**  
nilpotent

**20.5. Lemma.** Sei  $V = U_1 \oplus U_2 \oplus \dots \oplus U_n$  eine Zerlegung des Vektorraums  $V$  als direkte Summe und seien für  $i \in \{1, 2, \dots, n\}$  jeweils  $f_i, g_i \in \operatorname{End}(U_i)$ . Seien weiter  $f = f_1 \oplus f_2 \oplus \dots \oplus f_n$  und  $g = g_1 \oplus g_2 \oplus \dots \oplus g_n$ .

**LEMMA**

- (1) Sind alle  $g_i$  nilpotent, so ist auch  $g$  nilpotent.
- (2) Gilt  $f_i \circ g_i = g_i \circ f_i$  für alle  $i \in \{1, 2, \dots, n\}$ , dann gilt auch  $f \circ g = g \circ f$ .
- (3) Es gilt  $f + g = (f_1 + g_1) \oplus (f_2 + g_2) \oplus \dots \oplus (f_n + g_n)$ .
- (4) Es gilt  $f \circ g = (f_1 \circ g_1) \oplus (f_2 \circ g_2) \oplus \dots \oplus (f_n \circ g_n)$ .
- (5)  $\ker(f) = \ker(f_1) \oplus \ker(f_2) \oplus \dots \oplus \ker(f_n)$ .

*Beweis.*

- (1) Nach Voraussetzung gibt es zu jedem  $i$  ein  $m_i \in \mathbb{N}$  mit  $g_i^{m_i} = \mathbf{0}$ . Wir setzen  $m = \max\{m_1, m_2, \dots, m_n\}$ , dann gilt  $g_i^{om} = \mathbf{0}$  für alle  $i$ . Damit ist  $g^{om}(u_i) = g_i^{om}(u_i) = \mathbf{0}$  für  $u_i \in U_i$  und es folgt für  $v = u_1 + u_2 + \dots + u_n \in V$ , dass  $g^{om}(v) = \mathbf{0}$  ist. (Beachte  $g^{om} = g_1^{om} \oplus \dots \oplus g_n^{om}$  nach Teil (4) unten.) Also ist  $g^{om} = \mathbf{0}$  und  $g$  ist nilpotent.
- (2) Das folgt aus Teil (4).
- (3) Für  $u_i \in U_i$  gilt  $(f + g)(u_i) = f(u_i) + g(u_i) = f_i(u_i) + g_i(u_i) = (f_i + g_i)(u_i)$ , also ist  $f + g$  die direkte Summe der  $f_i + g_i$ .
- (4) Für  $u_i \in U_i$  gilt
 
$$(f \circ g)(u_i) = f(g(u_i)) = f(g_i(u_i)) = f_i(g_i(u_i)) = (f_i \circ g_i)(u_i),$$
 also ist  $f \circ g$  die direkte Summe der  $f_i \circ g_i$ .
- (5) Zunächst einmal ist klar, dass die Summe der Kerne rechts direkt ist: Sind  $u_1 \in \ker(f_1) \subset U_1, u_2 \in \ker(f_2) \subset U_2, \dots, u_n \in \ker(f_n) \subset U_n$  mit  $u_1 + u_2 + \dots + u_n = \mathbf{0}$ , dann folgt  $u_1 = u_2 = \dots = u_n = \mathbf{0}$ , weil die Summe der  $U_i$  direkt ist. Es ist also nur zu zeigen, dass

$$\ker(f) = \ker(f_1) + \ker(f_2) + \dots + \ker(f_n)$$

ist. Die Inklusion „ $\supset$ “ folgt aus  $\ker(f_i) \subset \ker(f)$  für alle  $i \in \{1, 2, \dots, n\}$  (denn für  $u_i \in U_i$  gilt  $f(u_i) = f_i(u_i)$ ), damit enthält der Untervektorraum  $\ker(f)$  auch die Summe der  $\ker(f_i)$ . Für die Inklusion „ $\subset$ “ sei jetzt  $v \in \ker(f)$ . Wir können  $v$  (eindeutig) schreiben als  $v = u_1 + u_2 + \dots + u_n$  mit  $u_i \in U_i$ , dann ist

$$\begin{aligned} \mathbf{0} &= f(v) = f(u_1 + u_2 + \dots + u_n) \\ &= f(u_1) + f(u_2) + \dots + f(u_n) = f_1(u_1) + f_2(u_2) + \dots + f_n(u_n). \end{aligned}$$

Aus  $f_i(u_i) \in U_i$  und daraus, dass die Summe der  $U_i$  direkt ist, folgt dann  $f_1(u_1) = f_2(u_2) = \dots = f_n(u_n) = \mathbf{0}$ , also  $u_i \in \ker(f_i)$  für alle  $i$ . Das bedeutet  $v \in \ker(f_1) + \ker(f_2) + \dots + \ker(f_n)$ . □

**20.6. Beispiel.** Wenn  $V$  unendlich-dimensional ist, dann braucht die aufsteigende Kette der Kerne nicht „stationär“ zu werden. Sei zum Beispiel  $V = K[X]$  der Polynomring und  $f \in \text{End}(V)$  die „Division durch  $X$  ohne Rest“, gegeben durch  $f(p) = (p - p(0))/X$ . Dann gilt  $\ker(f^{om}) = K[X]_{<m} = \langle 1, X, X^2, \dots, X^{m-1} \rangle$ , also werden diese Kerne immer größer. ♣

**BSP**  
unendlich  
aufsteigende  
Kerne

Wir können jetzt eine erste Version des Satzes von der Jordan-Normalform formulieren.

**\* 20.7. Satz.** Sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und sei  $f \in \text{End}(V)$  ein Endomorphismus, sodass das charakteristische Polynom  $\chi_f \in K[X]$  in Linearfaktoren zerfällt. Dann gibt es  $d, g \in \text{End}(V)$  mit  $f = d + g, d \circ g = g \circ d, d$  diagonalisierbar und  $g$  nilpotent.

**SATZ**  
Jordansche  
Normalform  
(schwach)

Die analoge Aussage gilt für Matrizen  $A \in \text{Mat}(n, K)$ : Zerfällt  $\chi_A$  in Linearfaktoren, dann gibt es Matrizen  $D, N \in \text{Mat}(n, K)$  mit  $A = D + N, DN = ND, D$  diagonalisierbar und  $N$  nilpotent.

*Beweis.* Wir beweisen die Version für Endomorphismen; der Beweis für Matrizen ist analog. Sei

$$\chi_f = (X - \lambda_1)^{e_1} (X - \lambda_2)^{e_2} \dots (X - \lambda_m)^{e_m}.$$

Nach Satz 19.12 gilt  $\chi_f(f) = \mathbf{0}$ . Satz 20.1 liefert dann  $f$ -invariante Untervektorräume  $U_1, U_2, \dots, U_m$  mit  $V = U_1 \oplus U_2 \oplus \dots \oplus U_m$ ; nach Lemma 20.3 gilt  $U_j = \ker((f - \lambda_j \text{id}_V)^{e_j}) = H_{\lambda_j}(f)$  für alle  $j$ . Die Einschränkung  $f_j$  von  $f$  auf  $U_j$  (als Endomorphismus von  $U_j$ ) hat also die Form  $f_j = d_j + g_j$  mit  $d_j = \lambda_j \text{id}_{U_j}$  und  $g_j^{e_j} = \mathbf{0}$ ; es gilt  $d_j \circ g_j = \lambda_j g_j = g_j \circ d_j$ . Mit  $d = d_1 \oplus d_2 \oplus \dots \oplus d_m$  und  $g = g_1 \oplus g_2 \oplus \dots \oplus g_m$  gilt dann  $f = d + g$ , und nach Lemma 20.5 gilt  $d \circ g = g \circ d$  und  $g$  ist nilpotent. Schließlich ist  $d$  diagonalisierbar, weil  $U_j = E_{\lambda_j}(d)$  ist; damit ist  $V$  die direkte Summe der Eigenräume von  $d$ .  $\square$

Um daraus eine stärkere Version abzuleiten, die eine Normalform für Matrizen ähnlich den Diagonalmatrizen ergibt, müssen wir uns die Struktur von nilpotenten Endomorphismen noch genauer ansehen. Sei also  $g \in \text{End}(V)$  nilpotent und sei  $m \in \mathbb{N}$  die kleinste natürliche Zahl mit  $g^{om} = \mathbf{0}$ . Dann haben wir die Kette

$$\{\mathbf{0}\} \subsetneq \ker(g) \subsetneq \ker(g^2) \subsetneq \dots \subsetneq \ker(g^{o(m-1)}) \subsetneq V$$

von Untervektorräumen von  $V$ . Dass die Inklusionen echt sind, ergibt sich aus dem folgenden Lemma.

**20.8. Lemma.** *Seien  $V$  ein Vektorraum,  $g \in \text{End}(V)$  und  $m \in \mathbb{N}$ . Dann folgt aus  $\ker(g^{o(m+1)}) = \ker(g^{om})$ , dass  $\ker(g^{on}) = \ker(g^{om})$  ist für alle  $n \geq m$ .*

**LEMMA**

*Beweis.* Übung.  $\square$

**\* 20.9. Definition.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A = (a_{ij}) \in \text{Mat}(n, K)$ . Dann heißt  $A$  eine *obere Dreiecksmatrix*, wenn gilt  $a_{ij} = 0$  für alle  $j < i$ , und eine *strikte obere Dreiecksmatrix*, wenn gilt  $a_{ij} = 0$  für alle  $j \leq i$ . Analog definiert man (*strikte*) *untere Dreiecksmatrizen* (mit  $j > i$  bzw.  $j \geq i$ ).  $\diamond$

**DEF**  
Dreiecks-  
matrix

In einer oberen Dreiecksmatrix sind also alle Einträge echt unterhalb der Diagonalen null, in einer strikten oberen Dreiecksmatrix sind zusätzlich die Einträge auf der Diagonalen null.

**20.10. Satz.** *Seien  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $g \in \text{End}(V)$  nilpotent. Dann gibt es eine Basis  $B$  von  $V$ , sodass  $\text{Mat}_{B,B}(g)$  eine strikte obere Dreiecksmatrix ist.*

**SATZ**  
Trigonalisierung  
nilpotenter  
Endo-  
morphisimen

*Analog gilt: Ist  $A \in \text{Mat}(n, K)$  nilpotent, dann ist  $A$  zu einer strikten oberen Dreiecksmatrix ähnlich.*

*Umgekehrt ist jede strikte obere Dreiecksmatrix  $A \in \text{Mat}(n, K)$  nilpotent; genauer gilt  $A^n = \mathbf{0}$ .*

*Beweis.* Sei  $U_j = \ker(g^{oj})$ , dann gilt  $\{\mathbf{0}\} = U_0 \subsetneq U_1 \subsetneq \dots \subsetneq U_m = V$ , wobei  $m$  minimal ist mit  $g^{om} = \mathbf{0}$ . Außerdem ist  $g(U_j) \subset U_{j-1}$  für alle  $j > 0$  (denn  $g^{oj}(v) = \mathbf{0}$  impliziert  $g^{o(j-1)}(g(v)) = \mathbf{0}$ ). Wir wählen eine Basis von  $U_1$ , die wir sukzessive zu Basen von  $U_2, U_3, \dots, U_m$  erweitern; sei  $B = (b_1, b_2, \dots, b_n)$  die resultierende Basis von  $V$ . Dann ist  $(b_1, b_2, \dots, b_{\dim U_j})$  eine Basis von  $U_j$ , und  $g(b_k)$  ist Linearkombination von Basiselementen  $b_i$  mit  $i < k$ , für alle  $k$ . Die  $k$ -te Spalte von  $A = \text{Mat}_{B,B}(g)$  hat also höchstens in den ersten  $k - 1$  Positionen

von oben Einträge ungleich null; das bedeutet gerade, dass  $A$  eine strikte obere Dreiecksmatrix ist.

Die Aussage über nilpotente Matrizen folgt aus der Aussage über Endomorphismen, indem man  $A$  als Endomorphismus von  $K^n$  betrachtet.

Die letzte Behauptung folgt aus Satz 19.12, denn für eine strikte Dreiecksmatrix  $A \in \text{Mat}(n, K)$  gilt  $\chi_A = X^n$ .  $\square$

Wir überlegen uns jetzt noch, wie die Matrix von  $f = f_1 \oplus f_2 \oplus \dots \oplus f_n$  bezüglich einer an die direkte Summenzerlegung angepassten Basis aussieht.

**20.11. Lemma.** *Sei  $V$  ein endlich-dimensionaler Vektorraum mit einer Zerlegung  $V = U_1 \oplus U_2 \oplus \dots \oplus U_n$ , seien  $f_i \in \text{End}(U_i)$  für  $i \in \{1, 2, \dots, n\}$  und sei  $f = f_1 \oplus f_2 \oplus \dots \oplus f_n$ . Seien  $B_i$  Basen von  $U_i$  und  $B$  die durch Aneinanderhängen von  $B_1, B_2, \dots, B_n$  gegebene Basis von  $V$ . Dann ist*

**LEMMA**  
Matrix  
einer  
direkten  
Summe

$$\text{Mat}_{B,B}(f) = \begin{pmatrix} M_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & M_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & M_n \end{pmatrix}$$

eine Block-Diagonalmatrix, wobei  $M_i = \text{Mat}_{B_i, B_i}(f_i)$  die Matrizen von  $f_i$  bezüglich der Basen  $B_i$  sind.

*Beweis.* Das folgt aus  $f(b) = f_i(b) \in U_i$  für Elemente  $b \in B_i$ : in den den Elementen von  $B_i$  entsprechenden Spalten von  $\text{Mat}_{B,B}(f)$  können nur die den Elementen von  $B_i$  entsprechenden Zeilen von null verschiedene Einträge enthalten, und die sich daraus ergebende Untermatrix ist die Matrix von  $f_i$  bezüglich  $B_i$ .  $\square$

Da die Determinante einer Block-Diagonalmatrix das Produkt der Determinanten der Diagonalblöcke ist, folgt daraus auch

$$\det(f) = \det(f_1) \det(f_2) \cdots \det(f_n)$$

und

$$\chi_f = \chi_{f_1} \chi_{f_2} \cdots \chi_{f_n}.$$

Damit können wir eine Aussage über die Dimension des Hauptraums  $H_\lambda(f)$  beweisen.

**20.12. Lemma.** *Seien  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $f$  ein Endomorphismus von  $V$ , sei weiter  $\lambda \in K$ . Dann ist die Dimension von  $H_\lambda(f)$  gleich der algebraischen Vielfachheit von  $\lambda$  als Eigenwert von  $f$  (also gleich der Vielfachheit von  $\lambda$  als Nullstelle von  $\chi_f$ ).*

**LEMMA**  
Dimension des  
Hauptraums

*Beweis.* Ist  $\lambda$  kein Eigenwert von  $f$ , dann sind beide Seiten der zu beweisenden Gleichung null. Anderenfalls ist  $\chi_f = (X - \lambda)^e \cdot p$  mit  $p \in K[X]$  relativ prim zu  $(X - \lambda)^e$ ; dabei ist  $e$  die algebraische Vielfachheit von  $\lambda$  als Eigenwert von  $f$  (dann ist  $p(\lambda) \neq 0$ ; nach Lemma 19.17 und Lemma 19.18 sind die Faktoren relativ prim). Nach Lemma 19.16 entspricht dieser Faktorisierung eine Zerlegung  $V = U_1 \oplus U_2$  mit  $U_1 = \ker((f - \lambda \text{id}_V)^e) = H_\lambda(f)$  und  $U_2 = \ker(p(f))$ . Sei  $f_1 = f|_{U_1} \in \text{End}(U_1)$  und  $f_2 = f|_{U_2} \in \text{End}(U_2)$ . Dann ist  $f_1 = \lambda \text{id}_{U_1} + g_1$  mit  $g_1$  nilpotent; nach Satz 20.10 ist dann  $f_1$  bezüglich einer geeigneten Basis von  $U_1$  durch eine obere Dreiecksmatrix dargestellt, deren Diagonaleinträge alle gleich  $\lambda$  sind. Es folgt  $\chi_{f_1} = (X - \lambda)^{\dim U_1}$ .

Auf der anderen Seite ist  $\lambda$  kein Eigenwert von  $f_2$  (ist  $v \in U_2$  mit  $f(v) = \lambda v$ , dann ist  $v \in U_1 \cap U_2 = \{\mathbf{0}\}$ , also  $v = \mathbf{0}$ ), also gilt  $\chi_{f_2}(\lambda) \neq 0$ . Aus der Gleichheit

$$(X - \lambda)^e \cdot p = \chi_f = (X - \lambda)^{\dim U_1} \cdot \chi_{f_2}$$

und  $p(\lambda) \neq 0$ ,  $\chi_{f_2}(\lambda) \neq 0$  folgt dann

$$\dim H_\lambda(f) = \dim U_1 = e$$

wie behauptet. □

Da der Eigenraum  $E_\lambda(f)$  stets im Hauptraum  $H_\lambda(f)$  enthalten ist, liefert dies auch einen weiteren Beweis der Aussage, dass die geometrische Vielfachheit eines Eigenwerts nicht größer als die algebraische Vielfachheit sein kann.

\* **20.13. Folgerung.** *Seien  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $f$  ein Endomorphismus von  $V$ , sodass das charakteristische Polynom  $\chi_f \in K[X]$  in Linearfaktoren zerfällt. Dann gibt es eine Basis  $B$  von  $V$ , sodass  $\text{Mat}_{B,B}(f)$  eine obere Dreiecksmatrix ist.*

**FOLG**  
Trigonalisierung von Endomorphismen

*Analog gilt: Ist  $A \in \text{Mat}(n, K)$  eine Matrix, sodass  $\chi_A$  in Linearfaktoren zerfällt, dann ist  $A$  ähnlich zu einer oberen Dreiecksmatrix.*

*Beweis.* Wie im Beweis von Satz 20.7 haben wir eine Zerlegung

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_m \quad \text{mit } U_j = H_{\lambda_j}(f)$$

in die Haupträume von  $f$ , sodass  $f$  auf  $U_j$  die Form  $\lambda_j \text{id} + g_j$  hat mit  $g_j$  nilpotent. Wir wählen Basen  $B_j$  von  $U_j$ , sodass  $\text{Mat}_{B_j, B_j}(g_j)$  eine strikte obere Dreiecksmatrix ist (das ist möglich nach Satz 20.10). Die Matrix bezüglich  $B_j$  der Einschränkung von  $f$  auf  $U_j$  ist dann  $M_j = \lambda_j I_{\dim U_j} + \text{Mat}_{B_j, B_j}(g_j)$ ; dies ist eine obere Dreiecksmatrix. Wir setzen die Basis  $B$  von  $V$  aus den Basen  $B_j$  zusammen. Nach Lemma 20.11 ist dann  $\text{Mat}_{B,B}(f)$  eine Block-Diagonalmatrix mit Blöcken  $M_j$ ; da die  $M_j$  obere Dreiecksmatrizen sind, gilt das auch für  $\text{Mat}_{B,B}(f)$ . Die Aussage für Matrizen folgt in der üblichen Weise. □

Die Voraussetzung, dass das charakteristische Polynom in Linearfaktoren zerfällt, ist stets erfüllt, wenn der Körper  $K$  algebraisch abgeschlossen ist, wie zum Beispiel  $K = \mathbb{C}$ .

Für viele Anwendungen sind die Aussagen von Satz 20.7 oder Folgerung 20.13 ausreichend. Manchmal möchte man aber eine im Wesentlichen eindeutige Normalform von Matrizen bis auf Ähnlichkeit haben. Dazu betrachten wir noch einmal nilpotente Endomorphismen und daran angepasste Basen.

Zur Vorbereitung noch ein Lemma zur Struktur von nilpotenten Endomorphismen. Darin leisten wir die Hauptarbeit für den Beweis der (starken) Jordanschen Normalform.

**20.14. Lemma.** *Sei  $V \neq \{\mathbf{0}\}$  ein endlich-dimensionaler Vektorraum und sei  $f \in \text{End}(V)$  nilpotent mit  $f^m = \mathbf{0}$  und  $m \in \mathbb{N}$  minimal mit dieser Eigenschaft. Sei weiter  $v \in V$  mit  $f^{\circ(m-1)}(v) \neq \mathbf{0}$  und  $U = \langle v, f(v), f^{\circ 2}(v), \dots, f^{\circ(m-1)}(v) \rangle \subset V$ . Dann gilt:*

**LEMMA**  
Struktur nilpotenter Endomorphismen

(1)  $\dim U = m$  (d.h.,  $v, f(v), f^{\circ 2}(v), \dots, f^{\circ(m-1)}(v)$  sind linear unabhängig).

(2) Es gibt ein  $f$ -invariantes Komplement  $U'$  von  $U$  in  $V$ .

*Beweis.* Wir zeigen zuerst, dass  $v, f(v), f^{\circ 2}(v), \dots, f^{\circ(m-1)}(v)$  linear unabhängig sind. Seien dazu  $\lambda_0, \lambda_1, \dots, \lambda_{m-1}$  Skalare mit

$$\lambda_0 v + \lambda_1 f(v) + \lambda_2 f^{\circ 2}(v) + \dots + \lambda_{m-1} f^{\circ(m-1)}(v) = \mathbf{0}.$$

Wenn wir  $f^{\circ(m-1)}$  auf diese Gleichung anwenden und beachten, dass  $f^{\circ m} = \mathbf{0}$  ist, dann erhalten wir  $\lambda_0 f^{\circ(m-1)}(v) = \mathbf{0}$ , wegen  $f^{\circ(m-1)}(v) \neq \mathbf{0}$  also  $\lambda_0 = 0$ . Durch Anwenden von  $f^{\circ(m-2)}$  bekommen wir dann analog  $\lambda_1 = 0$ , und in der gleichen Art dann nacheinander  $\lambda_2 = 0, \dots, \lambda_{m-1} = 0$ . Also sind die Vektoren linear unabhängig.

Wir zeigen nun die Existenz eines  $f$ -invarianten Komplements von  $U$  in  $V$ . Dazu sei  $U'$  ein Untervektorraum von  $V$  maximaler Dimension mit den beiden Eigenschaften  $U \cap U' = \{\mathbf{0}\}$  und  $U'$  invariant unter  $f$  (der Null-Vektorraum ist ein Untervektorraum mit diesen Eigenschaften, also gibt es so ein  $U'$ ). Wir müssen noch zeigen, dass  $V = U + U'$  ist. Sei anderenfalls  $w \in V \setminus (U + U')$ . Dann gibt es ein minimales  $k \in \mathbb{N}$  mit  $f^{\circ k}(w) \in U + U'$ ; es gilt  $k \geq 1$ , denn  $w \notin U + U'$ , und  $k \leq m$ , denn  $f^{\circ m}(w) = \mathbf{0} \in U + U'$ . Wir können dann schreiben

$$f^{\circ k}(w) = \lambda_0 v + \lambda_1 f(v) + \dots + \lambda_{m-1} f^{\circ(m-1)}(v) + u'$$

mit Skalaren  $\lambda_0, \lambda_1, \dots, \lambda_{m-1}$  und  $u' \in U'$ . Anwenden von  $f^{\circ(m-k)}$  liefert (unter Beachtung von  $f^{\circ m} = \mathbf{0}$ )

$$\begin{aligned} \mathbf{0} &= f^{\circ m}(w) \\ &= \underbrace{\lambda_0 f^{\circ(m-k)}(v) + \lambda_1 f^{\circ(m-k+1)}(v) + \dots + \lambda_{k-1} f^{\circ(m-1)}(v)}_{\in U} + \underbrace{f^{\circ(m-k)}(u')}_{\in U'}. \end{aligned}$$

Die lineare Unabhängigkeit der noch vorkommenden Vektoren (beachte dafür, dass  $U \cap U' = \{\mathbf{0}\}$  ist) erzwingt dann  $\lambda_0 = \lambda_1 = \dots = \lambda_{k-1} = 0$ , also gilt

$$\begin{aligned} f^{\circ k}(w) &= \lambda_k f^{\circ k}(v) + \lambda_{k+1} f^{\circ(k+1)}(v) + \dots + \lambda_{m-1} f^{\circ(m-1)}(v) + u' \\ &= f^{\circ k}(\lambda_k v + \lambda_{k+1} f(v) + \dots + \lambda_{m-1} f^{\circ(m-1-k)}(v)) + u'. \end{aligned}$$

Sei

$$u = \lambda_k v + \lambda_{k+1} f(v) + \dots + \lambda_{m-1} f^{\circ(m-1-k)}(v) \in U$$

und  $w' = w - u$ , dann ist  $w' \notin U + U'$  (denn  $w - w' = u \in U \subset U + U'$ ) und  $f^{\circ k}(w') = u' \in U'$ . Außerdem ist (wir verwenden hier  $k \geq 1$ , also  $k - 1 \geq 0$ )

$$w'' = f^{\circ(k-1)}(w') = f^{\circ(k-1)}(w) - f^{\circ(k-1)}(u) \notin U + U',$$

da  $f^{\circ(k-1)}(w) \notin U + U'$ . Sei  $U'' = U' + \langle w'' \rangle$ . Dann gilt:

- $\dim U'' > \dim U'$ , denn  $w'' \notin U'$ , also ist  $U' \subsetneq U''$ .
- $U'' \cap U = \{\mathbf{0}\}$ , denn sei  $u'_1 + \lambda w'' = u_1 \in U$  mit  $u'_1 \in U'$  und einem Skalar  $\lambda$ , dann folgt  $\lambda w'' = u_1 - u'_1 \in U + U'$ , aber  $w'' \notin U + U'$ . Daraus ergibt sich, dass  $\lambda = 0$  ist; die Behauptung folgt dann, weil  $U' \cap U = \{\mathbf{0}\}$  ist.
- $U''$  ist  $f$ -invariant, denn sei  $u'' = u'_1 + \lambda w'' \in U''$  mit  $u'_1 \in U'$  und einem Skalar  $\lambda$ , dann ist

$$f(u'') = f(u'_1) + \lambda f(w'') = f(u'_1) + \lambda f^{\circ k}(w') = f(u'_1) + \lambda u' \in U' \subset U'',$$

denn  $f(u'_1) \in U'$ , weil  $U'$  unter  $f$  invariant ist.

Insgesamt sehen wir, dass  $U''$  ein  $f$ -invarianter Untervektorraum von  $V$  ist, der  $U'' \cap U = \{\mathbf{0}\}$  erfüllt, aber größere Dimension als  $U'$  hat. Das ist ein Widerspruch zur Wahl von  $U'$ , also ist die Annahme  $U + U' \neq V$  falsch. Also muss  $U + U' = V$  gelten; damit ist  $U'$  das gesuchte  $f$ -invariante Komplement von  $U$  in  $V$ .  $\square$

Der obige Beweis ist, so wie er formuliert ist, nicht konstruktiv. Man kann das Argument, das zum Widerspruch führt, aber auch dazu verwenden, ein  $f$ -invariantes Komplement schrittweise (ausgehend von  $\{\mathbf{0}\}$ ) zu konstruieren.

**20.15. Definition.** Sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum. Ein Endomorphismus  $f$  von  $V$  heißt *nilzyklisch*, wenn  $f$  nilpotent ist und es  $v \in V$  gibt, sodass  $V = \langle v, f(v), f^{\circ 2}(v), \dots \rangle$  ist.

**DEF**  
nilzyklisch

Nach Lemma 20.14 sind  $v, f(v), \dots, f^{\circ(m-1)}(v)$  linear unabhängig, bilden also eine Basis von  $V$ , wobei  $m \in \mathbb{N}$  die kleinste Zahl ist mit  $f^{\circ m} = \mathbf{0}$ . Wir ordnen um und setzen  $B = (f^{\circ(m-1)}(v), \dots, f(v), v)$ , dann ist

$$\text{Mat}_{B,B}(f) = J_m = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \in \text{Mat}(m, K).$$

Die Matrizen  $J_m$  heißen ebenfalls *nilzyklisch*. ◇

**20.16. Satz.** Sei  $f$  ein nilpotenter Endomorphismus eines endlich-dimensionalen Vektorraums  $V$ . Dann gibt es eine Zerlegung  $V = U_1 \oplus U_2 \oplus \dots \oplus U_m$  in  $f$ -invariante Untervektorräume, sodass  $f|_{U_j}$  nilzyklisch ist für alle  $j \in \{1, 2, \dots, m\}$ .

**SATZ**  
Struktur  
nilpotenter  
Endo-  
morphismen

*Beweis.* Induktion über  $\dim V$ . Im Fall  $\dim V = 0$ , also  $V = \{\mathbf{0}\}$ , ist nichts zu zeigen ( $m = 0$ ). Sei also  $\dim V > 0$ . Dann gibt es nach Lemma 20.14 eine Zerlegung  $V = U_1 \oplus V'$  in  $f$ -invariante Untervektorräume mit  $U_1 \neq \{\mathbf{0}\}$ , sodass  $f|_{U_1}$  nilzyklisch ist ( $U_1$  ist  $U$  in Lemma 20.14,  $V'$  ist  $U'$ ). Nach Induktionsannahme (beachte  $\dim V' = \dim V - \dim U_1 < \dim V$ ) gibt es eine Zerlegung  $V' = U'_1 \oplus \dots \oplus U'_{m'}$  in  $f$ -invariante Untervektorräume, sodass  $f|_{U'_j}$  nilzyklisch ist für alle  $j \in \{1, 2, \dots, m'\}$ . Wenn wir  $m = m' + 1$  und (für  $j \in \{2, 3, \dots, m\}$ )  $U_j = U'_{j-1}$  setzen, dann erhalten wir insgesamt die gewünschte Zerlegung von  $V$ . □

**20.17. Folgerung.** Sei  $V$  ein endlich-dimensionaler Vektorraum und sei  $f$  ein nilpotenter Endomorphismus von  $V$ . Dann gibt es eine Basis  $B = (b_1, b_2, \dots, b_n)$  von  $V$  und Zahlen  $m_1, m_2, \dots, m_k \in \mathbb{N}_+$  mit  $m_1 + m_2 + \dots + m_k = n$ , sodass

**FOLG**  
Normalform  
für nilpotente  
Endo-  
morphismen

$$\text{Mat}_{B,B}(f) = \left( \begin{array}{c|c|c|c} J_{m_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & J_{m_2} & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & J_{m_k} \end{array} \right)$$

ist. Die Zahlen  $m_1, m_2, \dots, m_k$  sind bis auf ihre Reihenfolge eindeutig bestimmt. Analog gilt: Jede nilpotente Matrix  $A \in \text{Mat}(n, K)$  ist zu einer Matrix der obigen Form ähnlich. Diese Matrix ist bis auf die Reihenfolge der  $J_{m_i}$  eindeutig bestimmt.

*Beweis.* Nach Satz 20.16 gibt es eine Zerlegung

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_k$$

in  $f$ -invariante Untervektorräume mit  $f|_{U_i}$  nilzyklisch. Sei  $B_i$  für  $i \in \{1, 2, \dots, k\}$  eine Basis von  $U_i$  wie in Definition 20.15; sei  $B$  die durch Aneinanderhängen

von  $B_1, B_2, \dots, B_k$  gegebene Basis von  $V$ . Die erste Behauptung folgt dann mit Lemma 20.11; dabei ist  $m_i = \dim U_i$ .

Zur Eindeutigkeit: Für  $j \in \mathbb{N}$  gilt

$$\dim \ker(f^{\circ j}) = \sum_{i=1}^k \dim \ker(f|_{U_i}^{\circ j}) = \sum_{i=1}^k \min\{j, m_i\}$$

(die erste Gleichung folgt aus  $\ker(f^{\circ j}) = \bigoplus_{i=1}^k \ker(f|_{U_i}^{\circ j})$ , siehe Lemma 20.5(5), die zweite gilt, weil  $\ker(f|_{U_i}^{\circ j}) = \langle f^{\circ(m_i-j)}(v_i), f^{\circ(m_i-j+1)}(v_i), \dots, f^{\circ(m_i-1)}(v_i) \rangle$  ist für  $j \leq m_i$  und  $\ker(f|_{U_i}^{\circ j}) = U_i$  für  $j \geq m_i$ , wenn  $U_i = \langle v_i, f(v_i), \dots, f^{\circ(m_i-1)}(v_i) \rangle$  ist) und damit

$$\begin{aligned} \dim \ker(f^{\circ(j+1)}) - \dim \ker(f^{\circ j}) &= \sum_{i=1}^k (\min\{j+1, m_i\} - \min\{j, m_i\}) \\ &= \sum_{i=1}^k \begin{cases} 1, & \text{falls } m_i > j \\ 0, & \text{falls } m_i \leq j \end{cases} \\ &= \#\{i \in \{1, 2, \dots, k\} \mid m_i > j\}, \end{aligned}$$

also auch (für  $j \geq 1$ )

$$\begin{aligned} &(\dim \ker(f^{\circ j}) - \dim \ker(f^{\circ(j-1)})) - (\dim \ker(f^{\circ(j+1)}) - \dim \ker(f^{\circ j})) \\ &= \#\{i \in \{1, 2, \dots, k\} \mid m_i \geq j\} - \#\{i \in \{1, 2, \dots, k\} \mid m_i > j\} \\ &= \#\{i \in \{1, 2, \dots, k\} \mid m_i = j\}; \end{aligned}$$

damit sind die Zahlen  $m_i$  bis auf ihre Reihenfolge eindeutig festgelegt.

Die Aussage für Matrizen folgt in der üblichen Weise. □

Man kann sich also einen nilpotenten Endomorphismus  $f$  so vorstellen:

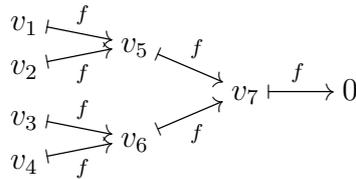
$$\begin{array}{ccccccccccc} U_1 : & & \bullet & \xrightarrow{f} & \bullet & \xrightarrow{f} & \cdots & \xrightarrow{f} & \bullet & \xrightarrow{f} & \bullet & \xrightarrow{f} & 0 \\ U_2 : & & \bullet & \xrightarrow{f} & \bullet & \xrightarrow{f} & \cdots & \xrightarrow{f} & \bullet & \xrightarrow{f} & \bullet & \xrightarrow{f} & 0 \\ U_3 : & & & & \bullet & \xrightarrow{f} & \cdots & \xrightarrow{f} & \bullet & \xrightarrow{f} & \bullet & \xrightarrow{f} & 0 \\ \vdots & & & & & & & & \vdots & & & & \vdots \\ U_k : & & & & & & & & \cdots & \xrightarrow{f} & \bullet & \xrightarrow{f} & 0 \end{array}$$

Die Punkte stehen dabei für die Basiselemente. Der Kern von  $f^{\circ j}$  wird dann erzeugt von den Basiselementen in den letzten  $j$  Spalten (außer der Nullspalte am Ende). Daraus kann man Folgendes ablesen:

Sei  $m$  die kleinste Zahl mit  $f^{\circ m} = \mathbf{0}$ . Für  $j \in \{1, 2, \dots, m\}$  sei  $V_j$  ein Komplement von  $\ker(f^{\circ(j-1)}) + f(\ker(f^{\circ(j+1)}))$  in  $\ker(f^{\circ j})$ . Sei  $B_j$  eine Basis von  $V_j$ . Für  $b \in B_j$  ist dann  $B'_b = (f^{\circ(j-1)}(b), \dots, f(b), b)$  eine Basis des  $f$ -invarianten Untervektorraums  $\langle b, f(b), \dots \rangle$ ; Hintereinanderhängen dieser Basen  $B'_b$  für alle  $b \in B_1 \cup B_2 \cup \dots \cup B_m$  liefert eine Basis  $B$  von  $V$ , bezüglich derer  $f$  durch eine Matrix wie in Folgerung 20.17 gegeben ist: Die Elemente von  $B_j$  entsprechen den Punkten im Diagramm in der  $j$ ten Spalte von rechts (außer der Nullspalte), an denen eine Kette  $\bullet \mapsto \bullet \mapsto \dots$  beginnt.

20.18. **Beispiel.** Sei  $V$  ein Vektorraum mit Basis  $(v_1, v_2, \dots, v_7)$  und  $f \in \text{End}(V)$  gegeben durch

**BSP**  
Normalform  
für nilpotenten  
Endo-  
morphismus



Dann ist  $f^3 = \mathbf{0}$  (und 3 ist die kleinste Zahl  $m$  mit  $f^m = \mathbf{0}$ ), also ist  $f$  nilpotent. Wir finden eine Basis von  $V$  wie in Folgerung 20.17. Dazu bestimmen wir erst einmal die „höheren Kerne“  $K_j = \ker(f^j)$  für  $j = 0, 1, 2, 3, \dots$ . Offenbar ist  $K_0 = \{\mathbf{0}\}$ . Aus

$$f(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_7 v_7) = (\lambda_1 + \lambda_2)v_5 + (\lambda_3 + \lambda_4)v_6 + (\lambda_5 + \lambda_6)v_7$$

folgt

$$K_1 = \langle v_1 - v_2, v_3 - v_4, v_5 - v_6, v_7 \rangle.$$

Weiter ist

$$f^2(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_7 v_7) = (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)v_7$$

und damit

$$K_2 = \langle v_1 - v_2, v_3 - v_4, v_1 - v_3, v_5, v_6, v_7 \rangle.$$

Schließlich ist  $K_j = V$  für  $j \geq 3$ . Wir müssen Komplemente  $V_j$  von  $K_{j-1} + f(K_{j+1})$  in  $K_j$  wählen. Für  $j = 1$  ist

$$K_0 + f(K_2) = f(K_2) = \langle v_5 - v_6, v_7 \rangle;$$

ein Komplement in  $K_1$  ist zum Beispiel gegeben durch

$$V_1 = \langle v_1 - v_2, v_3 - v_4 \rangle.$$

Für  $j = 2$  ist

$$K_1 + f(K_3) = \langle v_1 - v_2, v_3 - v_4, v_5, v_6, v_7 \rangle;$$

ein Komplement in  $K_2$  ist etwa

$$V_2 = \langle v_1 - v_3 \rangle.$$

Für  $j = 3$  schließlich ist

$$K_2 + f(K_4) = K_2 + f(V) = K_2;$$

ein Komplement ist zum Beispiel

$$V_3 = \langle v_1 \rangle.$$

Nach dem oben beschriebenen Rezept können wir als Basis wählen:

$$\begin{array}{ll}
 b_7 = v_1 - v_2 & \text{mit } f(b_7) = \mathbf{0} \\
 b_6 = v_3 - v_4 & \text{mit } f(b_6) = \mathbf{0} \\
 b_5 = v_1 - v_3 & \\
 b_4 = f(b_5) = v_5 - v_6 & \text{mit } f(b_4) = \mathbf{0} \\
 b_3 = v_1 & \\
 b_2 = f(b_3) = v_5 & \\
 b_1 = f(b_2) = v_7 & \text{mit } f(b_1) = \mathbf{0}
 \end{array}$$

Mit  $B = (b_1, b_2, b_3, b_4, b_5, b_6, b_7)$  ist dann

$$\text{Mat}_{B,B}(f) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \left( \begin{array}{c|c|c|c} J_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & J_2 & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & J_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & J_1 \end{array} \right)$$

Das zugehörige Diagramm sieht so aus:

$$\begin{array}{ccccccc} b_3 & \xrightarrow{f} & b_2 & \xrightarrow{f} & b_1 & \xrightarrow{f} & 0 \\ & & b_5 & \xrightarrow{f} & b_4 & \xrightarrow{f} & 0 \\ & & & & b_6 & \xrightarrow{f} & 0 \\ & & & & & & b_7 & \xrightarrow{f} & 0 \end{array}$$

Die Dimensionen von  $K_0, K_1, K_2, K_3, \dots$  sind  $0, 4, 6, 7, 7, \dots$ , die Zuwächse der Dimensionen also  $4, 2, 1, 0, 0, \dots$  und die Differenzen der Zuwächse sind  $2, 1, 1, 0, 0, \dots$ . Diese Zahlen geben die Häufigkeiten der nilzyklischen Kästchen  $J_1, J_2, J_3, \dots$  in der Matrix an, vergleiche den Beweis der Eindeutigkeitsaussage in Folgerung 20.17.



\* 20.19. **Definition.** Sei  $K$  ein Körper, seien  $\lambda \in K$  und  $m \in \mathbb{N}_+$ . Die Matrix

$$J_m(\lambda) = \lambda I_m + J_m$$

heißt das *Jordan-Kästchen* der Größe  $m$  zum Eigenwert  $\lambda$ .

**DEF**  
Jordan-  
Kästchen



\* 20.20. **Satz.** Seien  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $f$  ein Endomorphismus von  $V$  mit in Linearfaktoren zerfallendem charakteristischem Polynom  $\chi_f$ . Dann gibt es eine Zerlegung

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_k$$

in  $f$ -invariante Untervektorräume, sodass  $f|_{U_i} = \lambda_i \text{id}_{U_i} + g_i$  ist mit  $g_i$  nilzyklisch, für alle  $i \in \{1, 2, \dots, k\}$ . Insbesondere gibt es eine Basis  $B$  von  $V$ , sodass  $\text{Mat}_{B,B}(f)$  eine Block-Diagonalmatrix ist, deren Blöcke die Jordan-Kästchen  $J_{\dim U_i}(\lambda_i)$  sind. Die Jordan-Kästchen sind bis auf ihre Reihenfolge eindeutig bestimmt.

**SATZ**  
Jordansche  
Normalform  
(stark)

Analog gilt: Ist  $A \in \text{Mat}(n, K)$  mit zerfallendem charakteristischem Polynom  $\chi_A$ , dann ist  $A$  ähnlich zu einer Block-Diagonalmatrix, deren Blöcke Jordan-Kästchen sind, und die Jordan-Kästchen sind bis auf ihre Reihenfolge eindeutig bestimmt.

*Beweis.* Nach Satz 20.1 können wir  $V$  als direkte Summe in die verschiedenen Haupträume  $H_\lambda(f)$  zerlegen; auf  $H_\lambda(f)$  ist  $f = \lambda \text{id} + g$  mit  $g$  nilpotent. Nach Satz 20.16 gibt es eine Zerlegung von  $H_\lambda(f)$  als direkte Summe von Untervektorräumen, auf denen  $g$  nilzyklisch ist, dort ist also  $f = \lambda \text{id} + \text{nilzyklisch}$ . Wir erhalten die gewünschte Zerlegung von  $V$ , indem wir die Zerlegungen der Haupträume kombinieren. Wie in Definition 20.15 können wir Basen  $B_i$  der  $U_i$  so wählen, dass  $g_i$  auf  $U_i$  durch  $J_{\dim U_i}$  gegeben ist, dann ist

$$\text{Mat}_{B_i, B_i}(f|_{U_i}) = \lambda_i I_{\dim U_i} + J_{\dim U_i} = J_{\dim U_i}(\lambda_i).$$

Setzen wir diese Basen zu einer Basis  $B$  von  $V$  zusammen, erhalten wir die Block-Diagonalmatrix für  $f$  wie angegeben. Die Eindeutigkeit folgt daraus, dass die

Summe der Größen der Jordan-Kästchen zu einem gegebenen Eigenwert  $\lambda$  gleich der algebraischen Vielfachheit von  $\lambda$  als Eigenwert von  $f$  sein muss (vergleiche Lemma 20.12), und aus der Eindeutigkeitsaussage in Folgerung 20.17.

Die Aussagen für Matrizen folgen wie üblich aus denen für Endomorphismen: Wir wenden die Aussage auf den Endomorphismus  $f : K^n \rightarrow K^n$ ,  $\mathbf{x} \mapsto A \cdot \mathbf{x}$ , an (wobei die Elemente von  $K^n$  als Spaltenvektoren betrachtet werden). Dann ist  $A = \text{Mat}_{E,E}(f)$  mit der Standard-Basis  $E$  von  $K^n$ , und mit  $P = \text{Mat}_{B,E}(\text{id}_{K^n})$  hat dann  $\text{Mat}_{B,B}(f) = P^{-1}AP$  die Form wie im Satz.  $\square$

\* 20.21. **Definition.** Die Matrix in Satz 20.20 heißt die *Jordansche Normalform* von  $f$  bzw.  $A$ .

**DEF**  
 $\diamond$  Jordansche  
 Normalform

Die Jordansche Normalform liefert also eine vollständige Klassifikation der Matrizen mit zerfallendem charakteristischem Polynom bis auf Ähnlichkeit. Zum Beispiel gibt es genau drei Ähnlichkeitsklassen von Matrizen in  $\text{Mat}(3, K)$  mit charakteristischem Polynom  $(X - \lambda)^3$ , denn die Jordan-Normalform kann die Jordan-Kästchen  $J_3(\lambda)$  oder  $J_2(\lambda), J_1(\lambda)$  oder  $J_1(\lambda), J_1(\lambda), J_1(\lambda)$  haben.

Wie kann man die Jordansche Normalform einer gegebenen Matrix  $A \in \text{Mat}(n, K)$  bestimmen?

Wenn man nur wissen möchte, wie die Jordansche Normalform aussieht, dann bestimmt man zuerst die Eigenwerte (indem man das charakteristische Polynom faktorisiert) und berechnet dann für jeden Eigenwert  $\lambda$  die Dimensionen der Kerne von  $(A - \lambda I_n)^m$  für  $m = 1, 2, 3, \dots$ . Aus diesen Dimensionen ergeben sich die Größen der vorkommenden Jordan-Kästchen  $J_k(\lambda)$  wie im Beweis von Folgerung 20.17.

Braucht man zusätzlich die Matrix  $P \in \text{GL}(n, K)$ , sodass  $P^{-1}AP$  in Jordan-Normalform ist, dann muss man für jeden Hauptraum  $H_\lambda(A)$  eine Basis wie in Folgerung 20.17 bestimmen (zu  $g = (f - \lambda \text{id})|_{H_\lambda(A)}$ ) und diese Basen dann zu einer Basis von  $K^n$  zusammensetzen. Die Basiselemente bilden dann die Spalten von  $P$ .

Wir führen das in einem Beispiel durch:

20.22. **Beispiel.** Wir betrachten

$$A = \begin{pmatrix} 5 & -1 & 0 & 0 & -10 \\ 4 & -1 & 0 & 0 & -8 \\ 1 & 2 & 1 & -2 & -1 \\ 5 & 0 & 0 & -1 & -9 \\ 2 & -1 & 0 & 0 & -3 \end{pmatrix} \in \text{Mat}(5, \mathbb{R}).$$

**BSP**  
 Jordansche  
 Normalform

Das charakteristische Polynom ist

$$\begin{aligned} \chi_A &= \begin{vmatrix} X-5 & 1 & 0 & 0 & 10 \\ -4 & X+1 & 0 & 0 & 8 \\ -1 & -2 & X-1 & 2 & 1 \\ -5 & 0 & 0 & X+1 & 9 \\ -2 & 1 & 0 & 0 & X+3 \end{vmatrix} \\ &= (X-1)(X+1) \begin{vmatrix} X-5 & 1 & 10 \\ -4 & X+1 & 8 \\ -2 & 1 & X+3 \end{vmatrix} \\ &= (X-1)(X+1) \cdot ((X-5)(X+1)(X+3) - 16 - 40 - 8(X-5) + 20(X+1) + 4(X+3)) \\ &= (X-1)(X+1)(X^3 - X^2 - X + 1) = (X-1)^3(X+1)^2. \end{aligned}$$

Es sind also die beiden Haupträume  $H_1(A)$  (Dimension 3) und  $H_{-1}(A)$  (Dimension 2) zu betrachten. Für die Kerne erhalten wir (wir schreiben die Elemente von  $\mathbb{R}^5$  als Spaltenvektoren):

$$\begin{aligned} \ker(A - I_5) &= \langle (0, 0, 1, 0, 0)^\top, (3, 2, 0, 3, 1)^\top \rangle \\ \ker((A - I_5)^2) &= \langle (0, 0, 1, 0, 0)^\top, (1, 1, 0, 1, 0)^\top, (0, -1, 0, 0, 1)^\top \rangle \\ \ker(A + I_5) &= \langle (0, 0, 1, 1, 0)^\top \rangle \\ \ker((A + I_5)^2) &= \langle (0, 0, 1, 1, 0)^\top, (2, 2, -2, 0, 1)^\top \rangle \end{aligned}$$

Für die höheren Potenzen bleiben die Dimensionen gleich, da sie bereits den Dimensionen der Haupträume entsprechen. Daraus ergeben sich die Größen der Jordan-Kästchen:

$$\begin{aligned} \lambda = 1 : & \quad 0, 2, 3, 3, \dots \longrightarrow 2, 1, 0, \dots \longrightarrow 1, 1, 0, \dots \longrightarrow J_1(1), J_2(1) \\ \lambda = -1 : & \quad 0, 1, 2, 2, \dots \longrightarrow 1, 1, 0, \dots \longrightarrow 0, 1, 0, \dots \longrightarrow J_2(-1) \end{aligned}$$

Die Jordansche Normalform von  $A$  hat also die folgende Gestalt:

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

Um eine Transformationsmatrix  $P$  zu finden, gehen wir analog zu Beispiel 20.18 vor: Ein Komplement von  $\ker(A + I_5)$  in  $\ker((A + I_5)^2)$  wird zum Beispiel erzeugt von  $v_1 = (2, 2, -2, 0, 1)^\top$ . Ein Komplement von  $\ker(A - I_5)$  in  $\ker((A - I_5)^2)$  wird erzeugt von  $v_2 = (1, 1, 0, 1, 0)^\top$ , und ein Komplement von

$$(A - I_5)(\ker((A - I_5)^2)) = \langle (3, 2, 1, 3, 1)^\top \rangle$$

in  $\ker(A - I_5)$  wird erzeugt von  $v_3 = (0, 0, 1, 0, 0)^\top$ . Eine geeignete Basis ist damit  $B = (v_3, (A - I_5) \cdot v_2, v_2, (A + I_5) \cdot v_1, v_1)$ , entsprechend der Matrix

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



20.23. **Beispiel.** Als eine Anwendung der Klassifikationsaussage von Satz 20.20 wollen wir untersuchen, welche der Matrizen

**BSP**  
Testen auf  
Ähnlichkeit

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{und} \quad B(x, y) = \begin{pmatrix} -1 & -1 & x \\ 1 & 1 & y \\ 0 & 0 & 3 \end{pmatrix}$$

(für  $x, y \in \mathbb{R}$ ) in  $\text{Mat}(3, \mathbb{R})$  ähnlich zueinander sind. Dazu berechnen wir zunächst die charakteristischen Polynome; das ergibt

$$\chi_A = \chi_{B(x, y)} = X^2(X - 3).$$

Das schließt noch keine Ähnlichkeitsrelationen aus. Deshalb bestimmen wir die Jordansche Normalform der Matrizen. Für den Eigenwert 3 der algebraischen Vielfachheit 1 gibt es nur die Möglichkeit eines Jordan-Kästchens  $J_1(3)$ . Für den Eigenwert 0 gibt es die beiden Möglichkeiten  $J_2(0)$  und  $J_1(0), J_1(0)$ . Um sie zu unterscheiden, bestimmen wir die Dimension des Kerns von  $A$  bzw. von  $B(x, y)$ :

$$\dim \ker(A) = 3 - \text{rk}(A) = 3 - 1 = 2$$

und

$$\dim \ker(B(x, y)) = 3 - \text{rk}(B(x, y)) = 3 - 2 = 1.$$

Daraus ergibt sich, dass die Jordansche Normalform von  $A$  die Jordan-Kästchen  $J_1(0), J_1(0), J_1(3)$  hat (insbesondere ist  $A$  diagonalisierbar), während die Jordan-Normalform der Matrizen  $B(x, y)$  die Kästchen  $J_2(0), J_1(3)$  hat (insbesondere sind die  $B(x, y)$  nicht diagonalisierbar). Wir sehen also, dass  $A$  zu keiner der Matrizen  $B(x, y)$  ähnlich ist, dass aber alle  $B(x, y)$  zueinander ähnlich sind. ♣

Die hauptsächlich praktische Anwendung der Jordanschen Normalform besteht darin, dass sie die Berechnung von Potenzen einer Matrix vereinfacht: Sei etwa  $J = P^{-1}AP$  die Jordan-Normalform einer Matrix  $A$ , dann ist  $A = PJP^{-1}$  und  $A^k = PJ^kP^{-1}$  für alle  $k \in \mathbb{N}$ . Da  $J$  eine Block-Diagonalmatrix ist, ist  $J^k$  ebenfalls eine Block-Diagonalmatrix, deren Blöcke die  $k$ -ten Potenzen der Blöcke  $J_m(\lambda)$  sind. Nun ist  $J_m(\lambda) = \lambda I_m + J_m$ , also

$$J_m(\lambda)^k = \lambda^k I_m + \binom{k}{1} \lambda^{k-1} J_m + \binom{k}{2} \lambda^{k-2} J_m^2 + \dots + J_m^k,$$

und die Potenzen  $J_m^k$  haben eine sehr einfache Gestalt. Zum Beispiel ist

$$J_3(\lambda)^2 = \begin{pmatrix} \lambda^2 & 2\lambda & 1 \\ 0 & \lambda^2 & 2\lambda \\ 0 & 0 & \lambda^2 \end{pmatrix}$$

und allgemeiner

$$J_3(\lambda)^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{pmatrix}.$$

Eine Anwendung, die Sie in der „Einführung in die gewöhnlichen Differentialgleichungen“ kennenlernen werden, ist die Berechnung von  $e^{tA}$  für Matrizen  $A \in \text{Mat}(n, \mathbb{R})$  und  $t \in \mathbb{R}$ . Die Exponentialfunktion für Matrizen ist definiert wie für Zahlen:

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k;$$

man kann zeigen, dass die Reihe stets konvergiert (die Partialsummen sind Matrizen, sodass die Folge der Einträge an jeder gegebenen Position konvergiert). Ist  $A = PJP^{-1}$

wie oben, dann gilt  $e^{tA} = Pe^{tJ}P^{-1}$ , und  $e^{tJ}$  ist eine Block-Diagonalmatrix mit Blöcken der Form

$$e^{tJ_m(\lambda)} = e^{\lambda t} e^{tJ_m} = e^{\lambda t} \left( I_m + tJ_m + \frac{t^2}{2} J_m^2 + \dots \right).$$

Die Wichtigkeit der Funktion  $t \mapsto e^{tA}$  kommt aus folgendem Resultat:

Ist  $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top : \mathbb{R} \rightarrow \mathbb{R}^n$  ein  $n$ -Tupel differenzierbarer Funktionen, das das System von Differentialgleichungen

$$\mathbf{x}'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = A \cdot \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = A \cdot \mathbf{x}(t)$$

erfüllt, dann gilt  $\mathbf{x}(t) = e^{tA} \cdot \mathbf{x}(0)$ . Mit der Exponentialfunktion  $e^{tA}$  kann man also solche Differentialgleichungssysteme lösen. Für die Matrix  $A$  aus dem Beispiel 20.22 oben etwa hat man

$$e^{tA} = P \cdot \begin{pmatrix} e^t & 0 & 0 & 0 & 0 \\ 0 & e^t & te^t & 0 & 0 \\ 0 & 0 & e^t & 0 & 0 \\ 0 & 0 & 0 & e^{-t} & te^{-t} \\ 0 & 0 & 0 & 0 & e^{-t} \end{pmatrix} \cdot P^{-1} \\ = \begin{pmatrix} 3e^t - 2e^{-t} & (3t-2)e^t + 2e^{-t} & 0 & 0 & -(6t+2)e^t + 2e^{-t} \\ 2e^t - 2e^{-t} & (2t-1)e^t + 2e^{-t} & 0 & 0 & -(4t+2)e^t + 2e^{-t} \\ e^t - (t+1)e^{-t} & te^t + te^{-t} & e^t & -e^t + e^{-t} & -2te^t + te^{-t} \\ 3e^t - (t+3)e^{-t} & (3t-2)e^t + (t+2)e^{-t} & 0 & e^{-t} & -(6t+2)e^t + (t+2)e^{-t} \\ e^t - e^{-t} & (t-1)e^t + e^{-t} & 0 & 0 & -2te^t + e^{-t} \end{pmatrix}.$$

Wie sieht die Klassifikation von Matrizen bis auf Ähnlichkeit aus über  $\mathbb{R}$ , wo ja nicht jedes Polynom in Linearfaktoren zerfällt?

Sei  $A \in \text{Mat}(n, \mathbb{R})$ . Dann zerfällt  $\chi_A$  jedenfalls in  $\mathbb{C}[X]$  in Linearfaktoren. Diese können die Form  $X - \lambda$  haben mit  $\lambda \in \mathbb{R}$  oder die Form  $X - (\lambda + \mu i)$  mit  $\lambda, \mu \in \mathbb{R}$  und  $\mu \neq 0$ . Dann ist neben  $\lambda + \mu i$  auch  $\lambda - \mu i$  eine Nullstelle von  $\chi_A$ , denn sei

$$\chi_A = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0 \quad \text{mit } a_0, a_1, \dots, a_{n-1} \in \mathbb{R},$$

dann ist für  $z = a + bi \in \mathbb{C}$  mit komplex konjugierter Zahl  $\bar{z} = a - bi$

$$\overline{\chi_A(z)} = \overline{z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0} = \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \dots + a_1\bar{z} + a_0 = \chi_A(\bar{z}).$$

(Dabei haben wir benutzt, dass für  $w, z \in \mathbb{C}$  gilt  $\overline{w+z} = \bar{w} + \bar{z}$  und  $\overline{wz} = \bar{w}\bar{z}$ , sowie  $\bar{\bar{a}} = a$  für  $a \in \mathbb{R}$ .) Aus  $\chi_A(\lambda + \mu i) = 0$  folgt daher  $\chi_A(\lambda - \mu i) = \overline{\chi_A(\lambda + \mu i)} = 0$ . Man kann den Faktor

$$(X - \lambda - \mu i)(X - \lambda + \mu i) = X^2 - 2\lambda X + \lambda^2 + \mu^2$$

abdividieren; eine einfache Induktion zeigt dann, dass  $\lambda + \mu i$  und  $\lambda - \mu i$  dieselbe Vielfachheit als Nullstelle von  $\chi_A$  haben. Damit haben die zugehörigen Haupträume  $H_{\lambda + \mu i}(A)$  und  $H_{\lambda - \mu i}(A)$  in  $\mathbb{C}^n$  dieselbe Dimension. Sei  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$  eine Basis von  $H_{\lambda + \mu i}(A)$ , sodass  $A \cdot \mathbf{x}_j = (\lambda + \mu i)\mathbf{x}_j$  oder  $(\lambda + \mu i)\mathbf{x}_j + \mathbf{x}_{j-1}$  ist (also eine Basis, die zu den Jordan-Kästchen für den Eigenwert  $\lambda + \mu i$  gehört). Für einen Vektor  $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{C}^n$  sei  $\bar{\mathbf{y}} = (\bar{y}_1, \dots, \bar{y}_n)$ . Dann ist  $(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_m)$  eine Basis von  $H_{\lambda - \mu i}(A)$ , und

$$(\mathbf{x}_1 + \bar{\mathbf{x}}_1, \mathbf{i}^{-1}(\mathbf{x}_1 - \bar{\mathbf{x}}_1), \mathbf{x}_2 + \bar{\mathbf{x}}_2, \mathbf{i}^{-1}(\mathbf{x}_2 - \bar{\mathbf{x}}_2), \dots, \mathbf{x}_m + \bar{\mathbf{x}}_m, \mathbf{i}^{-1}(\mathbf{x}_m - \bar{\mathbf{x}}_m))$$

ist eine Basis von  $H_{\lambda + \mu i}(A) \oplus H_{\lambda - \mu i}(A)$ , deren Elemente in  $\mathbb{R}^n$  liegen. Die Matrix bezüglich dieser Basis des durch  $A$  gegebenen Endomorphismus dieses Untervektorraums

ist dann eine Block-Diagonalmatrix mit Blöcken der Form

$$J_{2m}(\lambda, \mu) = \begin{pmatrix} \lambda & -\mu & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \mu & \lambda & 0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \lambda & -\mu & 1 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \mu & \lambda & 0 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \ddots & \lambda & -\mu & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \mu & \lambda & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \lambda & -\mu \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \mu & \lambda \end{pmatrix} \in \text{Mat}(2m, \mathbb{R}).$$

Diese Matrix entsteht aus  $J_m(\lambda + \mu i)$ , indem jeder Eintrag  $a + bi$  (mit  $a, b \in \mathbb{R}$ ) durch die  $2 \times 2$ -Matrix  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$  ersetzt wird (dies ist die Matrix des  $\mathbb{R}$ -linearen Endomorphismus  $z \mapsto (a + bi)z$  von  $\mathbb{C}$  bezüglich der  $\mathbb{R}$ -Basis  $(1, i)$  von  $\mathbb{C}$ ).

Daraus ergibt sich der folgende Satz (formuliert für Matrizen):

**Satz.** Sei  $A \in \text{Mat}(n, \mathbb{R})$ . Dann ist  $A$  ähnlich zu einer Block-Diagonalmatrix, deren Blöcke die Form  $J_m(\lambda)$  (mit  $\lambda \in \mathbb{R}$ ) oder  $J_{2m}(\lambda, \mu)$  (mit  $\lambda \in \mathbb{R}, \mu \in \mathbb{R}_{>0}$ ) haben. Diese Blöcke sind bis auf ihre Reihenfolge eindeutig bestimmt.

**SATZ**  
Reelle JNF

## 21. SIMULTANE DIAGONALISIERBARKEIT

Bis jetzt haben wir immer nur *einen* Endomorphismus oder *eine* Matrix betrachtet. In diesem Abschnitt wollen wir untersuchen, wann sich mehrere Endomorphismen oder Matrizen „gleichzeitig“ diagonalisieren lassen. Die folgende Definition sagt, was genau damit gemeint ist.

\*

**21.1. Definition.** Sei  $V$  ein endlich-dimensionaler Vektorraum und sei  $(f_i)_{i \in I}$  eine Familie von Endomorphismen von  $V$ . Dann heißen die  $f_i$  *simultan diagonalisierbar*, wenn es eine Basis von  $V$  gibt, sodass jeder Basisvektor Eigenvektor aller  $f_i$  (zu möglicherweise verschiedenen Eigenwerten) ist.

**DEF**  
Simultane  
Diagonalisierbarkeit

Eine Familie  $(A_i)_{i \in I}$  von Matrizen in  $\text{Mat}(n, K)$  heißt *simultan diagonalisierbar*, wenn es eine invertierbare Matrix  $P \in \text{GL}(n, K)$  gibt, sodass  $P^{-1}A_iP$  für alle  $i \in I$  eine Diagonalmatrix ist.  $\diamond$

Die Matrizen  $A_i$  sind genau dann simultan diagonalisierbar, wenn die durch sie beschriebenen Endomorphismen  $f_i : \mathbf{x} \mapsto A_i\mathbf{x}$  simultan diagonalisierbar sind.

**21.2. Lemma.** Ist  $(f_i)_{i \in I}$  eine Familie simultan diagonalisierbarer Endomorphismen eines endlich-dimensionalen Vektorraums  $V$ , dann ist auch jeder Endomorphismus, der aus den  $f_i$  durch Bilden von Linearkombinationen und Verknüpfung von Abbildungen konstruiert werden kann, diagonalisierbar.

**LEMMA**

*Beweis.* Ist  $b \in V$  ein Eigenvektor aller  $f_i$ , dann ist  $b$  auch ein Eigenvektor aller wie in der Aussage des Lemmas konstruierten Endomorphismen  $f$ . Ist also  $B$  eine Basis von  $V$ , deren Elemente Eigenvektoren aller  $f_i$  sind, dann ist  $B$  auch eine Basis, die aus Eigenvektoren von  $f$  besteht.  $\square$

Bevor wir das wesentliche Ergebnis formulieren und beweisen können, brauchen wir noch zwei Hilfsaussagen.

**21.3. Lemma.** Sei  $V$  ein  $K$ -Vektorraum, seien  $f, g \in \text{End}(V)$  mit  $f \circ g = g \circ f$  und sei  $\lambda \in K$ . Dann ist der  $\lambda$ -Eigenraum von  $f$  unter  $g$  invariant und dasselbe gilt (falls  $V$  endlich-dimensional ist) für den  $\lambda$ -Hauptraum von  $f$ .

**LEMMA**  
Invarianz  
von Eigen-  
räumen

*Beweis.* Sei  $v \in E_\lambda(f)$ . Wir müssen zeigen, dass  $g(v) \in E_\lambda(f)$  ist:

$$f(g(v)) = (f \circ g)(v) = (g \circ f)(v) = g(f(v)) = g(\lambda v) = \lambda g(v).$$

Der Hauptraum  $H_\lambda(f)$  ist  $\ker((f - \lambda \text{id}_V)^{\circ m})$  für ein geeignetes  $m \in \mathbb{N}$ . Es gilt dann wie eben für  $v \in H_\lambda(f)$ :

$$\begin{aligned} (f - \lambda \text{id}_V)^{\circ m}(g(v)) &= ((f - \lambda \text{id}_V)^{\circ m} \circ g)(v) = (g \circ (f - \lambda \text{id}_V)^{\circ m})(v) \\ &= g((f - \lambda \text{id}_V)^{\circ m}(v)) = g(\mathbf{0}) = \mathbf{0}. \end{aligned}$$

Wir haben dabei verwendet, dass  $g$  auch mit  $f - \lambda \text{id}_V$  kommutiert (dass also  $(f - \lambda \text{id}_V) \circ g = f \circ g - \lambda g = g \circ f - \lambda g = g \circ (f - \lambda \text{id}_V)$  gilt).  $\square$

Allgemeiner gilt unter der Voraussetzung  $f \circ g = g \circ f$ , dass  $\ker(p(f))$  unter  $g$  invariant ist für jedes Polynom  $p \in K[X]$  (Übung). Das Lemma ist der Spezialfall  $p = X - \lambda$  bzw.  $(X - \lambda)^m$ .

**21.4. Lemma.** Sei  $V$  ein endlich-dimensionaler Vektorraum, sei  $f \in \text{End}(V)$  und sei  $U \subset V$  ein  $f$ -invarianter Untervektorraum. Ist  $f$  diagonalisierbar, dann ist auch  $f|_U \in \text{End}(U)$  diagonalisierbar.

**LEMMA**

*Beweis.* Nach Satz 19.10 zerfällt das Minimalpolynom  $m_f$  von  $f$  in Linearfaktoren und hat keine mehrfachen Nullstellen. Es gilt  $m_f(f|_U) = m_f(f)|_U = \mathbf{0}|_U = \mathbf{0}$ , also ist das Minimalpolynom von  $f|_U$  ein Teiler von  $m_f$ . Dann muss auch  $m_{f|_U}$  in Linearfaktoren zerfallen und kann keine mehrfachen Nullstellen haben, also ist wieder nach Satz 19.10 auch  $f|_U$  diagonalisierbar.  $\square$

\*

**21.5. Satz.** Sei  $V$  ein endlich-dimensionaler Vektorraum und sei  $(f_i)_{i \in I}$  eine Familie von Endomorphismen von  $V$ . Die  $f_i$  sind genau dann simultan diagonalisierbar, wenn jeder Endomorphismus  $f_i$  diagonalisierbar ist und die  $f_i$  paarweise kommutieren:  $f_i \circ f_j = f_j \circ f_i$  für alle  $i, j \in I$ .

**SATZ**  
Kriterium für  
simultane  
Diagonalisierbarkeit

*Analog gilt:* Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $(A_i)_{i \in I}$  eine Familie von Matrizen in  $\text{Mat}(n, K)$ . Dann sind die  $A_i$  genau dann simultan diagonalisierbar, wenn jede Matrix  $A_i$  diagonalisierbar ist und  $A_i A_j = A_j A_i$  gilt für alle  $i, j \in I$ .

*Beweis.* „ $\Rightarrow$ “: Nach Voraussetzung hat  $V$  eine Basis  $B = (b_1, b_2, \dots, b_n)$ , sodass jeder Basisvektor  $b_k$  ein Eigenvektor für alle  $f_i$  ist:  $f_i(b_k) = \lambda_{ik} b_k$ . Dann ist natürlich auch jedes  $f_i$  diagonalisierbar (denn  $B$  ist eine Basis von  $V$ , die aus Eigenvektoren von  $f_i$  besteht). Außerdem gilt für alle  $k \in \{1, 2, \dots, n\}$  und alle  $i, j \in I$ :

$$\begin{aligned} (f_i \circ f_j)(b_k) &= f_i(f_j(b_k)) = f_i(\lambda_{jk} b_k) = \lambda_{jk} f_i(b_k) = \lambda_{jk} \lambda_{ik} b_k \\ &= \lambda_{ik} \lambda_{jk} b_k = \lambda_{ik} f_j(b_k) = f_j(\lambda_{ik} b_k) = f_j(f_i(b_k)) = (f_j \circ f_i)(b_k). \end{aligned}$$

Da eine lineare Abbildung durch die Bilder der Basiselemente eindeutig bestimmt ist, folgt  $f_i \circ f_j = f_j \circ f_i$  für alle  $i, j \in I$ .

„ $\Leftarrow$ “: Induktion über die Dimension von  $V$ . Im Fall  $\dim V = 0$  oder  $\dim V = 1$  ist nichts zu zeigen (jeder Vektor  $\neq \mathbf{0}$  ist Eigenvektor jedes Endomorphismus). Sei also  $\dim V > 1$ . Wenn alle  $f_i$  die Form  $f_i(v) = \lambda_i v$  haben, dann ist jeder Vektor  $\mathbf{0} \neq v \in V$  Eigenvektor aller  $f_i$ , also tut es jede beliebige Basis. Es gebe also jetzt  $i_0 \in I$ , sodass  $f_{i_0}$  wenigstens zwei verschiedene Eigenwerte hat. Dann können wir  $V$  als direkte Summe

$$V = E_{\lambda_1}(f_{i_0}) \oplus E_{\lambda_2}(f_{i_0}) \oplus \dots \oplus E_{\lambda_m}(f_{i_0}) = V_1 \oplus V_2 \oplus \dots \oplus V_m$$

von Eigenräumen von  $f_{i_0}$  schreiben, wobei alle  $V_l = E_{\lambda_l}(f_{i_0})$  kleinere Dimension als  $V$  haben. Nach Lemma 21.3 sind alle  $V_l$  invariant unter allen  $f_i$ . Nach Lemma 21.4 ist  $f_i|_{V_l}$  diagonalisierbar für alle  $i \in I$  und alle  $l \in \{1, 2, \dots, m\}$ . Außerdem gilt natürlich auch  $f_i|_{V_l} \circ f_j|_{V_l} = f_j|_{V_l} \circ f_i|_{V_l}$ . Wir können also die Induktionsannahme auf  $V_l$  und die Familie  $(f_i|_{V_l})_{i \in I}$  anwenden: Es gibt für jedes  $l \in \{1, 2, \dots, m\}$  eine Basis  $B_l$  von  $V_l$ , die aus Eigenvektoren für alle  $f_i$  besteht. Wir setzen die  $B_l$  zu einer Basis  $B$  von  $V$  zusammen; dann besteht auch  $B$  aus Eigenvektoren aller  $f_i$ , also sind die  $f_i$  simultan diagonalisierbar.

Die Aussage für Matrizen folgt aus der für Endomorphismen.  $\square$

Wir können das auf *zyklische Matrizen* anwenden; das sind Matrizen der Form

$$Z(a_0, a_1, \dots, a_{n-1}) = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix} \in \text{Mat}(n, \mathbb{C})$$

mit  $a_0, a_1, \dots, a_{n-1} \in \mathbb{C}$ . Der Name kommt daher, dass jede Zeile durch zyklische Permutation der vorigen Zeile entsteht.

Sei  $Z = Z(0, 1, 0, 0, \dots, 0)$ , dann sieht man leicht, dass  $Z^m = Z(0, \dots, 0, 1, 0, \dots, 0)$  ist mit der Eins an Position  $m$  (dabei ist die erste Position „Position 0“), für alle  $m \in \{0, 1, \dots, n-1\}$ . Also ist

$$\begin{aligned} Z(a_0, a_1, \dots, a_{n-1}) &= a_0 I_n + a_1 Z + a_2 Z^2 + \dots + a_{n-1} Z^{n-1} \\ &= (a_0 + a_1 X + a_2 X^2 + \dots + a_{n-1} X^{n-1})(Z). \end{aligned}$$

Das bedeutet, dass alle zyklischen Matrizen miteinander kommutieren. Außerdem ist  $Z$  diagonalisierbar, denn  $\chi_Z = X^n - 1$  hat keine mehrfachen Nullstellen: Die Nullstellen sind gerade die Potenzen  $1, \omega, \omega^2, \dots, \omega^{n-1}$  von  $\omega = e^{2\pi i/n} = \cos(2\pi/n) + i \sin(2\pi/n)$ . Es folgt, dass alle zyklischen Matrizen diagonalisierbar sind; damit sind sie sogar simultan diagonalisierbar. Eine Basis aus Eigenvektoren aller zyklischen Matrizen ist gegeben durch eine Basis aus Eigenvektoren von  $Z$ , nämlich zum Beispiel

$$\begin{aligned} \mathbf{b}_0 &= (1, 1, 1, \dots, 1), \\ \mathbf{b}_1 &= (1, \omega, \omega^2, \dots, \omega^{n-1}), \\ &\vdots \\ \mathbf{b}_j &= (1, \omega^j, \omega^{2j}, \dots, \omega^{(n-1)j}), \\ &\vdots \\ \mathbf{b}_{n-1} &= (1, \omega^{n-1}, \omega^{n-2}, \dots, \omega); \end{aligned}$$

es gilt  $Z \cdot \mathbf{b}_j = \omega^j \mathbf{b}_j$ . (Die auf den ersten Blick von der für  $\mathbf{b}_j$  gegebenen allgemeinen Form abweichende Darstellung von  $\mathbf{b}_{n-1}$  ergibt sich daraus, dass  $\omega^n = 1$  ist. Man hat dann nämlich

$$\omega^{k(n-1)} = \omega^{kn-k} = \omega^{(k-1)n+(n-k)} = (\omega^n)^k \omega^{n-k} = \omega^{n-k}.)$$

Es folgt:

**Satz.** Die Eigenwerte (mit Vielfachheit) der zyklischen Matrix  $Z(a_0, a_1, \dots, a_{n-1})$  sind wie folgt:

**SATZ**  
zyklische  
Matrizen

$$\begin{aligned} \lambda_0 &= a_0 + a_1 + a_2 + \dots + a_{n-1} \\ \lambda_1 &= a_0 + a_1 \omega + a_2 \omega^2 + \dots + a_{n-1} \omega^{n-1} \\ &\vdots \\ \lambda_j &= a_0 + a_1 \omega^j + a_2 \omega^{2j} + \dots + a_{n-1} \omega^{(n-1)j} \\ &\vdots \\ \lambda_{n-1} &= a_0 + a_1 \omega^{n-1} + a_2 \omega^{n-2} + \dots + a_{n-1} \omega \end{aligned}$$

Insbesondere ist

$$\det(Z(a_0, a_1, \dots, a_{n-1})) = \lambda_0 \lambda_1 \cdots \lambda_{n-1} = \prod_{j=0}^{n-1} \sum_{k=0}^{n-1} a_k \omega^{jk}.$$

Zum Abschluss beweisen wir noch, dass der diagonalisierbare Endomorphismus  $d$  und der nilpotente Endomorphismus  $g$  in Satz 20.7 (schwache Jordansche Normalform) eindeutig bestimmt sind.

**21.6. Satz.** *In Satz 20.7 sind die Endomorphismen  $d$  und  $g$  eindeutig bestimmt.*

**SATZ**  
Eindeutigkeit  
der JNF-  
Zerlegung

*Beweis.* Es sei  $f = d + g = d' + g'$  mit  $d \circ g = g \circ d$ ,  $d' \circ g' = g' \circ d'$  und  $d, d'$  diagonalisierbar,  $g, g'$  nilpotent; dabei seien  $d$  und  $g$  wie im Beweis von Satz 20.7 konstruiert. Wir zeigen zuerst, dass  $d'$  und  $g'$  mit  $d$  und  $g$  kommutieren. Wir bemerken, dass  $d'$  und  $g'$  jedenfalls mit  $f$  kommutieren (denn

$$d' \circ f = d' \circ (d' + g') = d' \circ d' + d' \circ g' = d' \circ d' + g' \circ d' = (d' + g') \circ d' = f \circ d'$$

und analog für  $g'$ ). Nach Lemma 21.3 sind die Haupträume von  $f$  also unter  $d'$  und  $g'$  invariant. Auf dem Hauptraum  $H_\lambda(f)$  hat  $d$  die Form  $\lambda \text{id}$ , also kommutiert  $d$  mit  $d'$  und  $g'$  auf  $H_\lambda(f)$ . Da  $d$  und  $d'$  (bzw.  $g'$ ) direkte Summen von Endomorphismen dieser Haupträume sind, folgt  $d \circ d' = d' \circ d$  und  $d \circ g' = g' \circ d$ . Da  $d'$  und  $g'$  mit  $f = d + g$  kommutieren, kommutieren sie auch mit  $g$ . Da  $d$  und  $d'$  beide diagonalisierbar sind und miteinander kommutieren, sind sie nach Satz 21.5 simultan diagonalisierbar, also ist nach Lemma 21.2 auch  $d - d' = g' - g$  diagonalisierbar. Auf der anderen Seite ist  $g' - g$  nilpotent (da  $g$  und  $g'$  nilpotent sind und miteinander kommutieren — Übung). Ein Endomorphismus, der gleichzeitig diagonalisierbar und nilpotent ist, muss aber die Nullabbildung sein (das Minimalpolynom hat keine mehrfachen Nullstellen und ist von der Form  $X^m$ , also ist es  $X$  (oder 1; dann ist  $V = \{\mathbf{0}\}$ ) und damit ist die Abbildung  $\mathbf{0}$ ). Es folgt  $d = d'$  und  $g = g'$  und damit die behauptete Eindeutigkeit.  $\square$

## 22. ÄQUIVALENZRELATIONEN, QUOTIENTENRÄUME UND AFFINE UNTERRÄUME

Wir erinnern uns daran, dass der Kern jeder linearen Abbildung  $f : V \rightarrow V'$  ein Untervektorraum von  $V$  ist. Hier wollen wir gewissermaßen der umgekehrten Frage nachgehen: Ist jeder Untervektorraum der Kern einer linearen Abbildung?

Im Fall, dass  $V$  endlich-dimensional ist, können wir mit unseren Kenntnissen über direkte Summen und Komplemente recht leicht zeigen, dass die Antwort „Ja“ lautet: Sei  $U$  ein Untervektorraum des endlich-dimensionalen Vektorraums  $V$ , dann hat  $U$  ein Komplement  $U'$  in  $V$  (Satz 18.13), es ist also  $V = U \oplus U'$ . Die zu dieser Zerlegung gehörende Projektion  $\pi : V \rightarrow U'$  hat dann  $U$  als Kern.

Dieses Argument ist aus zwei Gründen etwas unbefriedigend. Zum Einen verwendet es die Existenz von Basen (genauer: den Basisergänzungssatz), die wir nur für endlich-dimensionale Vektorräume gezeigt haben. Zum Anderen ist das Komplement  $U'$  im Normalfall weit davon entfernt, eindeutig bestimmt zu sein; wir müssen bei der Konstruktion der linearen Abbildung also eine Wahl treffen.

In diesem Abschnitt werden wir eine Konstruktion kennen lernen, die diese Nachteile vermeidet: Sie funktioniert für jeden Untervektorraum jedes Vektorraums und erfordert keine Auswahlen. Die Art dieser Konstruktion des „Quotientenraums“ und des zugehörigen „kanonischen Epimorphismus“ ist recht typisch für die Methoden der Algebra und wird in sehr ähnlicher Form im Rahmen der Vorlesungen „Einführung in die Zahlentheorie und algebraische Strukturen“ und „Einführung in die Algebra“ wieder auftauchen, dann für andere algebraische Strukturen wie zum Beispiel Ringe und Gruppen.

Sei also  $V$  ein (beliebiger)  $K$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Wenn es eine lineare Abbildung  $f : V \rightarrow V'$  gibt mit  $\ker(f) = U$ , dann gibt es auch eine surjektive solche Abbildung, denn wir können einfach die im Wertebereich eingeschränkte Abbildung  $f : V \rightarrow \text{im}(f)$  betrachten. Wir nehmen jetzt an, dass wir so eine surjektive lineare Abbildung  $f : V \rightarrow V'$  mit Kern  $U$  haben. Wie können wir dann den Vektorraum  $V'$  beschreiben?

- Die **Elemente** von  $V'$  können wir durch Elemente von  $V$  repräsentieren; dabei wird  $v' \in V'$  durch jedes  $v \in V$  mit  $f(v) = v'$  repräsentiert (das ist möglich, weil  $f$  surjektiv ist). Zwei Elemente  $v_1$  und  $v_2$  von  $V$  stellen genau dann dasselbe Element von  $V'$  dar, wenn  $f(v_1) = f(v_2)$  ist. Das ist äquivalent zu  $f(v_1 - v_2) = \mathbf{0}$ , also zu  $v_1 - v_2 \in \ker(f) = U$ .
- Die **Addition** und **Skalarmultiplikation** auf  $V'$  kann unter Zuhilfenahme der Linearität von  $f$  ebenfalls über die entsprechenden Operationen von  $V$  erfolgen: Sind  $v_1$  und  $v_2$  Repräsentanten von  $v'_1 = f(v_1)$  und  $v'_2 = f(v_2)$ , dann ist  $v_1 + v_2$  ein Repräsentant von  $v'_1 + v'_2$ , denn  $f(v_1 + v_2) = f(v_1) + f(v_2)$ . Ebenso ist für  $\lambda \in K$  auch  $\lambda v_1$  ein Repräsentant von  $\lambda v'_1$ .

Wir schreiben  $[v]$  (statt  $f(v)$ ) für das von  $v \in V$  repräsentierte Element von  $V'$ . Dann können wir unsere Überlegungen wie folgt zusammenfassen: Falls  $V'$  existiert, dann

- (1) besteht  $V'$  aus allen  $[v]$  mit  $v \in V$ ;
- (2) es gilt  $[v_1] = [v_2] \iff v_1 - v_2 \in U$
- (3) und  $[v_1] + [v_2] = [v_1 + v_2]$ ,  $\lambda[v] = [\lambda v]$ .

Es liegt also nahe,  $V'$  auf diese Weise zu *definieren*; dann wäre  $f : V \rightarrow V', v \mapsto [v]$ , die passende surjektive lineare Abbildung mit Kern  $U$  (denn  $[v] = \mathbf{0}$  genau dann, wenn  $v \in U$ ). Dafür müssen wir nachweisen, dass diese Vorgehensweise zu keinen Widersprüchen führt.

Der erste Punkt dabei ist, sich zu überlegen, dass die Gleichheit der Elemente von  $V'$  sinnvoll definiert ist. Eine sinnvolle Definition von Gleichheit muss sicher die folgenden Eigenschaften haben:

- (1) Jedes Element ist gleich zu sich selbst.
- (2) Wenn  $a$  und  $b$  gleich sind, dann sind auch  $b$  und  $a$  gleich.
- (3) Wenn sowohl  $a$  und  $b$ , als auch  $b$  und  $c$  gleich sind, dann sind auch  $a$  und  $c$  gleich.

Wir gießen das in eine formale Definition. Dafür brauchen wir den Begriff der *Relation*.

\* 22.1. **Definition.** Seien  $X$  und  $Y$  beliebige Mengen. Eine *Relation zwischen  $X$  und  $Y$*  ist eine Teilmenge  $R \subset X \times Y$ . Man sagt,  $x \in X$  steht in der Relation  $R$  zu  $y \in Y$  (oder  $x$  und  $y$  stehen in der Relation  $R$ ), wenn  $(x, y) \in R$  gilt. Manchmal schreibt man dafür abkürzend  $x R y$ . **DEF**  
Relation

Im Fall  $X = Y$  spricht man von einer *Relation auf  $X$* . ◇

22.2. **Beispiele.** **BSP**  
Relationen  
Auf jeder Menge  $X$  gibt es die *Gleichheitsrelation*  $\{(x, x) \mid x \in X\}$  und die *Allrelation*  $X \times X$ .

Auf  $\mathbb{R}$  gibt es die Vergleichsrelationen  $\{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \leq y\}$  und analog für  $<$ ,  $\geq$ ,  $>$ .

Auf  $\mathbb{Z}$  gibt es die *Teilbarkeitsrelation*  $\{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid \exists c \in \mathbb{Z} : ac = b\}$ , deren Bestehen als  $a \mid b$  notiert wird.

Zwischen einer Menge  $X$  und ihrer Potenzmenge  $\mathcal{P}(X)$  gibt es die Element-Relation  $\{(x, T) \in X \times \mathcal{P}(X) \mid x \in T\}$ . ♣

\* 22.3. **Definition.** Seien  $X$  eine Menge und  $R$  eine Relation auf  $X$ . Dann heißt  $R$  eine *Äquivalenzrelation*, wenn folgende Bedingungen erfüllt sind: **DEF**  
Äquivalenz-  
relation

- (1)  $\forall x \in X : x R x$  (Reflexivität).
- (2)  $\forall x, y \in X : (x R y \implies y R x)$  (Symmetrie).
- (3)  $\forall x, y, z \in X : (x R y \wedge y R z \implies x R z)$  (Transitivität). ◇

Die Gleichheitsrelation und die Allrelation sind Äquivalenzrelationen auf jeder Menge  $X$ . Dagegen sind die Vergleichsrelationen auf  $\mathbb{R}$  (außer der Gleichheit) und die Teilbarkeitsrelation auf  $\mathbb{Z}$  keine Äquivalenzrelationen, denn (z.B.) aus  $a \leq b$  folgt nicht unbedingt  $b \leq a$  und aus  $a \mid b$  folgt nicht unbedingt  $b \mid a$ .

Man kann eine Äquivalenzrelation als eine „vergrößerte“ Version von Gleichheit verstehen: Man betrachtet Elemente als gleich, obwohl sie nicht unbedingt identisch sind, aber so, dass die wesentlichen Eigenschaften der Gleichheit erfüllt sind. Das führt zu einer Einteilung von  $X$  in Klassen als untereinander gleich betrachteter Elemente:

**22.4. Satz.** Sei  $X$  eine Menge und sei  $\sim$  eine Äquivalenzrelation auf  $X$ . Ist  $x \in X$ , dann schreiben wir  $[x]$  für die Menge  $\{y \in X \mid y \sim x\}$  und nennen  $[x]$  die Äquivalenzklasse von  $x$  (bezüglich  $\sim$ ). Für  $x, y \in X$  sind dann die folgenden Aussagen äquivalent:

**SATZ**  
Äquivalenz-  
klassen

- (1)  $y \sim x$ .
- (2)  $y \in [x]$ .
- (3)  $[x] = [y]$ .
- (4)  $[x] \cap [y] \neq \emptyset$ .

Insbesondere sind je zwei Äquivalenzklassen entweder gleich oder disjunkt.

*Beweis.* Die Äquivalenz von (1) und (2) ist nichts Anderes als die Definition von  $[x]$ .

„(1)  $\Rightarrow$  (3)“: Für  $z \in X$  gilt (unter der Voraussetzung  $y \sim x$ , also auch  $x \sim y$ ):

$$z \in [x] \iff z \sim x \iff z \sim y \iff z \in [y],$$

also sind  $[x]$  und  $[y]$  gleich. Die mittlere Äquivalenz benutzt die Transitivität von  $\sim$ .

„(3)  $\Rightarrow$  (4)“ ist trivial, denn  $x \in [x] = [y]$ .

„(4)  $\Rightarrow$  (1)“: Sei  $z \in [x] \cap [y]$ , dann gilt  $z \sim x$  und  $z \sim y$ . Die Symmetrie von  $\sim$  impliziert  $y \sim z$ , die Transitivität dann  $y \sim x$ .  $\square$

**22.5. Definition.** In der Situation von Satz 22.4 schreiben wir

$$X/\sim = \{[x] \mid x \in X\} \subset \mathcal{P}(X)$$

**DEF**  
Quotienten-  
menge

für die Menge der Äquivalenzklassen und nennen  $X/\sim$  die *Quotientenmenge* von  $X$  bezüglich  $\sim$ . Die Abbildung  $X \rightarrow X/\sim, x \mapsto [x]$ , ist surjektiv; sie heißt die *kanonische Surjektion*.  $\diamond$

Wir können das auf unser Problem anwenden.

\* **22.6. Lemma.** Seien  $V$  ein  $K$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Die wie folgt definierte Relation  $\equiv_U$  auf  $V$  ist eine Äquivalenzrelation. Statt  $v \equiv_U v'$  schreiben wir  $v \equiv v' \pmod{U}$  (gesprochen „ $v$  ist kongruent zu  $v'$  modulo  $U$ “).

**LEMMA**  
Kongruenz  
modulo  $U$  ist  
Äquivalenz-  
relation

$$v \equiv v' \pmod{U} \iff v - v' \in U.$$

Statt  $V/\equiv_U$  schreiben wir  $V/U$  für die Quotientenmenge. Für die Äquivalenzklassen gilt

$$[v] = \{v' \in V \mid v' - v \in U\} = \{v + u \mid u \in U\} = v + U.$$

*Beweis.* Wir müssen zeigen, dass die so definierte Relation reflexiv, symmetrisch und transitiv ist:

- Für  $v \in V$  gilt  $v - v = \mathbf{0} \in U$ , also  $v \equiv v \pmod{U}$ .
- Für  $v, v' \in V$  gelte  $v \equiv v' \pmod{U}$ , das bedeutet  $v - v' \in U$ . Dann ist auch  $v' - v = -(v - v') \in U$  und damit  $v' \equiv v \pmod{U}$ .
- Für  $v, v', v'' \in V$  gelte  $v \equiv v' \pmod{U}$  und  $v' \equiv v'' \pmod{U}$ , das bedeutet  $v - v', v' - v'' \in U$ . Dann ist auch  $v - v'' = (v - v') + (v' - v'') \in U$ , also gilt  $v \equiv v'' \pmod{U}$ .

Die Aussage über die Gestalt der Äquivalenzklassen ist klar (mit  $u = v' - v$ ).  $\square$

Wir wollen jetzt gerne  $V' = V/U$  setzen, mit der kanonischen Surjektion als linearer Abbildung. Dafür müssen wir nachweisen, dass die Definitionen der Addition,  $[v] + [v'] = [v + v']$ , und der Skalarmultiplikation,  $\lambda[v] = [\lambda v]$ , sinnvoll sind. Das wird durch zusätzliche Eigenschaften von  $\equiv_U$  sichergestellt.

**22.7. Lemma.** *Die Relation  $\equiv_U$  aus Lemma 22.6 hat zusätzlich folgende Eigenschaften:*

- (1) Für  $v_1, v_2, v'_1, v'_2 \in V$  gilt:  
*Aus  $v_1 \equiv v'_1 \pmod U$  und  $v_2 \equiv v'_2 \pmod U$  folgt  $v_1 + v_2 \equiv v'_1 + v'_2 \pmod U$ .*
- (2) Für  $v, v' \in V$  und  $\lambda \in K$  gilt: *Aus  $v \equiv v' \pmod U$  folgt  $\lambda v \equiv \lambda v' \pmod U$ .*

**LEMMA**  
 Kongruenz modulo  $U$  ist Kongruenzrelation

Eine Äquivalenzrelation auf einem Vektorraum mit diesen zusätzlichen Eigenschaften (also Verträglichkeit mit der Vektorraum-Struktur) wird auch als *Kongruenzrelation* bezeichnet.

*Beweis.* (1) Wir haben  $v_1 - v'_1, v_2 - v'_2 \in U$ , also auch  $(v_1 + v_2) - (v'_1 + v'_2) = (v_1 - v'_1) + (v_2 - v'_2) \in U$ .

(2) Aus  $v - v' \in U$  folgt  $\lambda v - \lambda v' = \lambda(v - v') \in U$ .  $\square$

\* **22.8. Satz.** *Seien  $V$  ein  $K$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Durch die Festlegungen*

$$[v] + [v'] = [v + v'] \quad \text{und} \quad \lambda \cdot [v] = [\lambda v]$$

*wird die Menge  $V/U$  zu einem  $K$ -Vektorraum.*

*Die kanonische Surjektion  $\pi : V \rightarrow V/U, v \mapsto [v]$ , ist dann eine lineare Abbildung mit  $\ker(\pi) = U$ .*

**SATZ**  
 Quotientenraum

Der Vektorraum  $V/U$  („ $V$  modulo  $U$ “) heißt der *Quotientenraum* von  $V$  modulo  $U$ , die lineare Abbildung  $\pi : V \rightarrow V/U$  der *kanonische Epimorphismus*.

*Beweis.* Zuerst ist zu zeigen, dass die Definitionen der Addition und Skalarmultiplikation sinnvoll („wohldefiniert“) sind: Da es im Allgemeinen viele Möglichkeiten gibt, ein Element von  $V/U$  in der Form  $[v]$  zu schreiben, müssen wir nachprüfen, dass die Definitionen nicht von der Auswahl der Repräsentanten abhängen. Es gelte also  $[v_1] = [v'_1]$  und  $[v_2] = [v'_2]$ , also  $v_1 \equiv v'_1 \pmod U$  und  $v_2 \equiv v'_2 \pmod U$ . Nach Lemma 22.7 folgt dann  $v_1 + v_2 \equiv v'_1 + v'_2 \pmod U$ , also  $[v_1 + v_2] = [v'_1 + v'_2]$ . Das zeigt, dass die Summe von  $[v_1]$  und  $[v_2]$  nicht von der Wahl der Repräsentanten abhängt. Auf analoge Weise zeigt man, dass auch die Definition der Skalarmultiplikation sinnvoll ist.

Als Nächstes müssen wir die Axiome für einen Vektorraum nachprüfen. Sobald klar ist, dass Addition und Skalarmultiplikation wohldefiniert sind, folgen diese aber direkt aus ihrer Gültigkeit für  $V$ , wobei man natürlich  $\mathbf{0} = [\mathbf{0}]$  und  $-[v] = [-v]$  setzt. Wir zeigen das am Beispiel eines der Distributivgesetze: Seien  $v, v' \in V$  und  $\lambda \in K$ . Dann gilt

$$\lambda([v] + [v']) = \lambda[v + v'] = [\lambda(v + v')] = [\lambda v + \lambda v'] = [\lambda v] + [\lambda v'] = \lambda[v] + \lambda[v'].$$

Die anderen Axiome zeigt man nach demselben Schema: Linke Seite als Restklasse eines Elements von  $V$  schreiben, dann das Axiom in  $V$  anwenden, dann in die rechte Seite umformen.

Dass die kanonische Surjektion  $\pi$  linear ist, folgt schließlich direkt aus der Definition von Addition und Skalarmultiplikation in  $V/U$ . Tatsächlich ist die Definition gerade so gemacht, *damit*  $\pi$  linear wird! Es gilt dann (man beachte  $\mathbf{0}_{V/U} = [\mathbf{0}_V]$ , also  $[v] = \mathbf{0}_{V/U} \iff v \in U$ )

$$\ker(\pi) = \{v \in V \mid [v] = \mathbf{0}\} = \{v \in V \mid v \in U\} = U. \quad \square$$

Damit ist die eingangs gestellte Frage positiv beantwortet.

Hat  $U$  ein Komplement  $U'$  in  $V$ , dann ist die Einschränkung des kanonischen Epimorphismus  $\pi : V \rightarrow V/U$  auf  $U'$  ein Isomorphismus  $U' \rightarrow V/U$  (Übung). Es folgt  $\text{codim}_V U = \dim U' = \dim V/U$ . Wir können also die Kodimension für beliebige Untervektorräume als  $\text{codim}_V U = \dim V/U$  definieren, ohne auf die Existenz eines Komplements angewiesen zu sein. Die Formel

$$\dim U + \text{codim}_V U = \dim V$$

ist dann nichts anderes als der „Rangatz“  $\dim V = \dim \ker(\pi) + \dim \text{im}(\pi)$  für den kanonischen Epimorphismus  $\pi$ .

Wir beweisen jetzt noch einige Eigenschaften von Quotientenraum und kanonischem Epimorphismus.

**22.9. Satz.** *Seien  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum; sei weiter  $\pi : V \rightarrow V/U$  der kanonische Epimorphismus. Dann sind die Abbildungen*

$$\begin{aligned} \{U' \subset V \mid U' \text{ UVR mit } U \subset U'\} &\longleftrightarrow \{W \subset V/U \mid W \text{ UVR}\} \\ U' &\longmapsto \pi(U') \\ \pi^{-1}(W) &\longleftarrow W \end{aligned}$$

**SATZ**  
UVR von  
 $V$  und  $V/U$

*zueinander inverse inklusionserhaltende Bijektionen.*

Das bedeutet, dass die Untervektorräume von  $V/U$  genau den Untervektorräumen von  $V$  entsprechen, die  $U$  enthalten (und zwar mitsamt der Inklusionen, die zwischen ihnen gelten: Eine Abbildung  $\phi$  ist inklusionserhaltend, wenn aus  $A \subset B$  folgt, dass  $\phi(A) \subset \phi(B)$  ist).

*Beweis.* Seien  $\Phi$  die Abbildung  $U' \mapsto \pi(U')$  und  $\Psi$  die Abbildung  $W \mapsto \pi^{-1}(W)$ . Da  $\pi$  linear ist, ist für jeden Untervektorraum  $U'$  von  $V$  die Bildmenge  $\pi(U')$  ein Untervektorraum von  $V/U$ , und für jeden Untervektorraum  $W$  von  $V/U$  ist die Urbildmenge  $\pi^{-1}(W)$  ein Untervektorraum von  $V$ , vergleiche Satz 9.9, wo auch gezeigt wurde, dass  $\Phi$  und  $\Psi$  zueinander inverse Bijektionen sind. Dass die beiden Abbildungen inklusionserhaltend sind, ist eine allgemeine Eigenschaft von Bild- und Urbildmengen.  $\square$

Als nächstes beantworten wir die Frage, wann eine lineare Abbildung  $f : V \rightarrow W$  eine lineare Abbildung  $\phi : V/U \rightarrow W$  „induziert“, wann es also so ein  $\phi$  gibt, sodass  $\phi \circ \pi = f$  ist, wobei  $\pi : V \rightarrow V/U$  der kanonische Epimorphismus ist: Gibt es  $\phi$ , sodass das folgende Diagramm „kommutiert“?

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \pi \downarrow & \nearrow \phi & \\ V/U & & \end{array}$$

Da für jedes  $u \in U$  gilt, dass  $\pi(u) = \mathbf{0}$  ist, muss auch  $f(u) = \phi(\pi(u)) = \mathbf{0}$  sein; das bedeutet  $U \subset \ker(f)$ . Wie der folgende Satz zeigt, ist diese Bedingung auch hinreichend.

**22.10. Satz.** *Seien  $V$  ein Vektorraum mit Untervektorraum  $U$  und  $\pi : V \rightarrow V/U$  der kanonische Epimorphismus. Sei weiter  $f : V \rightarrow W$  linear. Es gibt genau dann eine lineare Abbildung  $\phi : V/U \rightarrow W$  mit  $f = \phi \circ \pi$ , wenn  $U \subset \ker(f)$  ist. Diese Abbildung ist genau dann injektiv, wenn  $U = \ker(f)$  ist.*

**SATZ**  
lin. Abb.  
auf  $V/U$

*Beweis.* Dass die Bedingung notwendig ist, hatten wir uns bereits überlegt. Für die Gegenrichtung nehmen wir  $U \subset \ker(f)$  an. Wenn es  $\phi$  gibt, dann muss gelten

$$\phi([v]) = \phi(\pi(v)) = f(v);$$

die Frage ist nur, ob wir  $\phi$  tatsächlich so definieren können. Dazu müssen wir zeigen, dass  $f(v)$  nicht vom Repräsentanten von  $[v]$  abhängt. Es seien also  $v, v' \in V$  mit  $[v] = [v']$ , also  $v - v' \in U$ . Dann ist

$$f(v) = f((v - v') + v') = f(v - v') + f(v') = f(v'),$$

weil aus  $v - v' \in U \subset \ker(f)$  folgt, dass  $f(v - v') = \mathbf{0}$  ist. Damit ist  $\phi$  durch  $\phi([v]) = f(v)$  wohldefiniert, und es gilt jedenfalls  $\phi \circ \pi = f$ . Es bleibt zu zeigen, dass  $\phi$  linear ist. Das folgt aber aus der Linearität von  $\pi$  und von  $f$ :

$$\phi([v] + [v']) = \phi([v + v']) = f(v + v') = f(v) + f(v') = \phi([v]) + \phi([v'])$$

und

$$\phi(\lambda[v]) = \phi([\lambda v]) = f(\lambda v) = \lambda f(v) = \lambda \phi([v]).$$

$\phi$  ist genau dann injektiv, wenn  $\ker(\phi)$  trivial ist. Aus der Definition von  $\phi$  folgt  $\ker(\phi) = \pi(\ker(f))$ , also nach Satz 22.9  $\ker(f) = \pi^{-1}(\ker(\phi))$ . Aus  $U = \pi^{-1}(\{\mathbf{0}\})$  ergibt sich die Behauptung  $\ker(\phi) = \{\mathbf{0}\} \iff \ker(f) = U$ .  $\square$

Satz 22.10 zeigt, wie man lineare Abbildungen mit Definitionsbereich  $V/U$  konstruieren kann.

\*

**22.11. Satz.** *Sei  $f : V \rightarrow W$  eine lineare Abbildung. Seien  $\pi : V \rightarrow V/\ker(f)$  der kanonische Epimorphismus und  $\iota : \text{im}(f) \rightarrow W$  die Inklusionsabbildung. Dann gibt es einen eindeutig bestimmten Isomorphismus  $\phi : V/\ker(f) \rightarrow \text{im}(f)$ , sodass  $f = \iota \circ \phi \circ \pi$  ist:*

**SATZ**  
Homomorphiesatz für lineare Abb.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \pi \downarrow & & \uparrow \iota \\ V/\ker(f) & \xrightarrow[\cong]{\phi} & \text{im}(f) \end{array}$$

*Insbesondere sind  $V/\ker(f)$  und  $\text{im}(f)$  isomorph.*

*Beweis.* Nach Satz 22.10 gibt es eine lineare Abbildung  $\tilde{\phi} : V/\ker(f) \rightarrow W$  mit  $f = \tilde{\phi} \circ \pi$ . Es gilt  $\text{im}(\tilde{\phi}) = \text{im}(f)$ , also können wir  $\tilde{\phi}$  im Wertebereich einschränken zu  $\phi : V/\ker(f) \rightarrow \text{im}(f)$ ; es folgt  $f = \iota \circ \phi \circ \pi$ . Es bleibt zu zeigen, dass  $\phi$  ein Isomorphismus ist und dass  $\phi$  eindeutig bestimmt ist. Letzteres folgt aus

$$\phi([v]) = \iota(\phi([v])) = f(v).$$

$\phi$  ist injektiv nach Satz 22.10 und surjektiv wegen  $\text{im}(\phi) = \text{im}(\tilde{\phi}) = \text{im}(f)$ , also ist  $\phi$  ein Isomorphismus.  $\square$

**22.12. Beispiel.** Die rationalen Cauchy-Folgen bilden einen Untervektorraum  $C$  des Vektorraums  $\mathbb{Q}^{\mathbb{N}}$  der Folgen über  $\mathbb{Q}$ , denn Summen und skalare Vielfache von Cauchy-Folgen sind wieder Cauchy-Folgen. In  $C$  bilden die Nullfolgen einen Untervektorraum  $N$ . Jede Cauchy-Folge konvergiert in  $\mathbb{R}$  und jede reelle Zahl ist Grenzwert einer rationalen Cauchy-Folge. Das liefert uns eine surjektive  $\mathbb{Q}$ -lineare Abbildung

$$\lim : C \longrightarrow \mathbb{R}, \quad (a_n)_{n \in \mathbb{N}} \longmapsto \lim_{n \rightarrow \infty} a_n$$

mit Kern  $N$  („Nullfolge“ heißt ja gerade „Grenzwert null“). Aus dem Homomorphiesatz 22.11 folgt jetzt, dass  $C/N$  isomorph zu  $\mathbb{R}$  ist (als  $\mathbb{Q}$ -Vektorraum). Dies ist eine der Möglichkeiten, wie man die reellen Zahlen aus den rationalen Zahlen konstruieren kann. In der „Einführung in die Zahlentheorie und algebraische Strukturen“ werden wir lernen, dass die gleiche Konstruktion auch die Struktur von  $\mathbb{R}$  als Körper mitliefert. ♣

Weitere Anwendungen des Homomorphiesatzes sind durch die folgenden „Isomorphiesätze“ gegeben.

**22.13. Satz.** Seien  $V$  ein Vektorraum und  $U_1, U_2 \subset V$  zwei Untervektorräume. Dann ist die Abbildung

$$\phi : U_1/(U_1 \cap U_2) \longrightarrow (U_1 + U_2)/U_2, \quad u + (U_1 \cap U_2) \longmapsto u + U_2$$

ein Isomorphismus.

(Wir verwenden hier die präzisere Schreibweise  $v + U$  für die Äquivalenzklasse  $[v]$ , weil wir es mit zwei verschiedenen Quotientenräumen zu tun haben. In der Beschreibung von  $\phi$  ist  $u$  ein Element von  $U_1$ .)

*Beweis.*

$$\begin{array}{ccc} U_1 & \xrightarrow{\quad} & U_1 + U_2 \\ \downarrow & \searrow f & \downarrow \\ U_1/(U_1 \cap U_2) & \xrightarrow[\phi]{\cong} & (U_1 + U_2)/U_2 \end{array}$$

Wir betrachten die Verknüpfung  $f : U_1 \rightarrow (U_1 + U_2)/U_2$  der Inklusionsabbildung  $U_1 \rightarrow U_1 + U_2$  mit dem kanonischen Epimorphismus  $U_1 + U_2 \rightarrow (U_1 + U_2)/U_2$ . Dann ist  $\ker(f) = U_1 \cap U_2$ . Außerdem ist  $f$  surjektiv: Sei  $v + U_2 \in (U_1 + U_2)/U_2$  mit  $v \in U_1 + U_2$ , dann gibt es  $u_1 \in U_1$  und  $u_2 \in U_2$  mit  $v = u_1 + u_2$ . Es folgt  $v + U_2 = u_1 + U_2 = f(u_1)$ , da  $v - u_1 = u_2 \in U_2$ . Nach dem Homomorphiesatz 22.11 existiert der Isomorphismus  $\phi$  wie angegeben. □

**22.14. Satz.** Seien  $V$  ein Vektorraum und  $U_1 \subset U_2 \subset V$  Untervektorräume. Dann ist  $U_2/U_1$  ein Untervektorraum von  $V/U_1$  und die Abbildung

$$\phi : V/U_2 \longrightarrow (V/U_1)/(U_2/U_1), \quad v + U_2 \longmapsto (v + U_1) + U_2/U_1$$

ist ein Isomorphismus.

*Beweis.*

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U_1 \\ \downarrow & \searrow f & \downarrow \\ V/U_2 & \xrightarrow[\phi]{\cong} & (V/U_1)/(U_2/U_1) \end{array}$$

**BSP**  
Konstruktion  
von  $\mathbb{R}$  aus  $\mathbb{Q}$

**SATZ**  
Erster Iso-  
morphiesatz

**SATZ**  
Zweiter Iso-  
morphiesatz

Sei  $\pi : V \rightarrow V/U_1$  der kanonische Epimorphismus, dann ist  $U_2/U_1 = \pi(U_2)$  ein Untervektorraum von  $V/U_1$  nach Satz 22.9. Ähnlich wie eben betrachten wir die Abbildung  $f : V \rightarrow (V/U_1)/(U_2/U_1)$ , die die Komposition von  $\pi$  mit dem kanonischen Epimorphismus  $V/U_1 \rightarrow (V/U_1)/(U_2/U_1)$  ist. Da beide Epimorphismen surjektiv sind, gilt das auch für  $f$ . Außerdem ist  $\ker(f) = \pi^{-1}(U_2/U_1) = U_2$ . Die Behauptung folgt wieder aus dem Homomorphiesatz 22.11.  $\square$

Die Äquivalenzklassen  $[v] = v + U$ , die in diesem Zusammenhang auch *Nebenklassen* (von  $U$ ) oder *Restklassen* (modulo  $U$ ) heißen, haben auch eine geometrische Interpretation als „verschobene Untervektorräume“ (man verschiebt nämlich  $U$  um den Vektor  $v$ ). Dafür gibt es einen eigenen Namen.

\* 22.15. **Definition.** Sei  $V$  ein Vektorraum. Ein *affiner Unterraum* von  $V$  ist entweder die leere Menge oder eine Menge der Form  $v + U$  mit  $v \in V$  und einem Untervektorraum  $U$  von  $V$ . Die *Dimension* von  $v + U$  ist  $\dim(v + U) = \dim U$ , die Dimension des leeren affinen Unterraums wird als  $-\infty$  definiert.  $\diamond$

**DEF**  
Affiner  
Unterraum

Wir kennen affine Unterräume bereits als Lösungsmengen von linearen Gleichungen (siehe Satz 11.10): Die Lösungsmenge jeder linearen Gleichung  $f(x) = b$  (wobei  $f : V \rightarrow W$  eine lineare Abbildung ist) ist ein affiner Unterraum von  $V$ . Umgekehrt ist jeder affine Unterraum von  $V$  auch Lösungsmenge einer linearen Gleichung. Das ist klar für die leere Menge (wähle  $f = \mathbf{0} : V \rightarrow K$  und  $b = 1$ ); für  $A = v + U$  ist  $A = \pi^{-1}([v])$  für den kanonischen Epimorphismus  $\pi : V \rightarrow V/U$ .

Der Untervektorraum  $U$ , der zu einem nicht-leeren affinen Unterraum  $A$  gehört, ist durch  $A$  eindeutig bestimmt, denn es ist  $U = A - A = \{v - v' \mid v, v' \in A\}$ . Dagegen ist der „Aufpunkt“  $v$  nicht eindeutig bestimmt (außer  $U = \{\mathbf{0}\}$ ), denn jedes  $v \in A$  erfüllt  $A = v + U$ .

Wir können affine Unterräume durch eine Abgeschlossenheitseigenschaft charakterisieren.

\* 22.16. **Satz.** Seien  $V$  ein  $K$ -Vektorraum und  $A \subset V$  eine Teilmenge. Dann sind äquivalent:

**SATZ**  
Charakterisierung  
affiner  
Unterräume

- (1)  $A$  ist ein affiner Unterraum von  $V$ .
- (2)  $A$  ist unter affinen Linearkombinationen abgeschlossen:  
Für  $a_1, a_2, \dots, a_n \in A$  und  $\lambda_1, \lambda_2, \dots, \lambda_n \in K$  mit  $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$  gilt  $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n \in A$ .

*Beweis.* „(1) $\Rightarrow$ (2)“: Wenn  $A = \emptyset$  ist, ist nichts zu zeigen. Sei also  $A = v + U$  mit  $v \in V$  und einem Untervektorraum  $U$ . Dann ist  $a_j = v + u_j$  mit  $u_j \in U$  für alle  $j \in \{1, 2, \dots, n\}$ , also erhalten wir

$$\begin{aligned} \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n &= \lambda_1(v + u_1) + \lambda_2(v + u_2) + \dots + \lambda_n(v + u_n) \\ &= (\lambda_1 + \lambda_2 + \dots + \lambda_n)v + \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n \\ &= v + (\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n) \in v + U = A. \end{aligned}$$

„(2) $\Rightarrow$ (1)“: Wenn  $A = \emptyset$  ist, dann ist  $A$  ein affiner Unterraum. Wir können also  $A \neq \emptyset$  annehmen; sei  $v \in A$  fest gewählt und  $U = A - v = \{a - v \mid a \in A\} \subset V$ . Wir zeigen, dass  $U$  ein Untervektorraum von  $V$  ist, dann folgt, dass  $A = v + U$  ein affiner Unterraum ist.

- $\mathbf{0} \in U$ , da  $\mathbf{0} = v - v$  und  $v \in A$  ist.

- $U$  ist abgeschlossen unter der Addition: Seien  $u = a - v$  und  $u' = a' - v$  mit  $a, a' \in A$ . Nach Voraussetzung gilt  $a + a' - v \in A$  (das ist eine affine Linearkombination), also ist  $u + u' = (a + a' - v) - v \in U$ .
- $U$  ist abgeschlossen unter der Skalarmultiplikation: Seien  $u = a - v$  mit  $a \in A$  und  $\lambda \in K$ . Nach Voraussetzung gilt  $\lambda a - \lambda v + v \in A$ , also ist  $\lambda u = (\lambda a - \lambda v + v) - v \in U$ .  $\square$

Daraus folgt, dass Durchschnitte von affinen Unterräumen wieder affine Unterräume sind.

**22.17. Folgerung.** Sei  $V$  ein Vektorraum und sei  $(A_i)_{i \in I}$  eine nicht-leere (d.h.,  $I \neq \emptyset$ ) Familie von affinen Unterräumen von  $V$ . Dann ist  $\bigcap_{i \in I} A_i$  ebenfalls ein affiner Unterraum von  $V$ .

**FOLG**  
Durchschnitte  
von affinen  
Unterräumen

*Beweis.* Nach Satz 22.16 sind alle  $A_i$  abgeschlossen unter affinen Linearkombinationen. Seien jetzt  $a_1, a_2, \dots, a_n \in A = \bigcap_{i \in I} A_i$  und seien  $\lambda_1, \lambda_2, \dots, \lambda_n$  Skalare mit  $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$ . Dann ist  $a = \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n \in A_i$  für alle  $i \in I$ , also ist  $a \in A$ . Damit ist  $A$  unter affinen Linearkombinationen abgeschlossen, also ist  $A$  wiederum nach Satz 22.16 ein affiner Unterraum von  $V$ .  $\square$

Ist  $A_i = v_i + U_i$  und  $\bigcap_{i \in I} A_i = v + U \neq \emptyset$ , dann ist  $U = \bigcap_{i \in I} U_i$  (Übung).

**22.18. Beispiel.** Welche affinen Unterräume gibt es im  $\mathbb{R}^3$ ?

- Die leere Menge ist ein affiner Unterraum.
- Jede einelementige Menge  $\{\mathbf{x}\}$  ist ein affiner Unterraum der Dimension 0.
- Jede Gerade (nicht unbedingt durch den Nullpunkt) ist ein affiner Unterraum der Dimension 1.
- Jede Ebene (nicht unbedingt durch den Nullpunkt) ist ein affiner Unterraum der Dimension 2.
- $\mathbb{R}^3$  selbst ist der einzige affine Unterraum der Dimension 3.

**BSP**  
Affine  
Unterräume  
im  $\mathbb{R}^3$

Zwei Geraden können zusammenfallen, sich in einem Punkt (affiner Unterraum der Dimension 0) schneiden oder disjunkt sein (dann sind sie parallel oder windschief). Für eine Gerade  $g$  und eine Ebene  $E$  gibt es die folgenden Möglichkeiten:  $g \subset E$ ,  $g \cap E = \{P\}$  oder  $g \cap E = \emptyset$  (dann ist  $g$  parallel zu  $E$ ). Zwei Ebenen können übereinstimmen, sich in einer Geraden schneiden oder disjunkt sein (dann sind sie parallel).

Man kann affine Unterräume wahlweise in der Form  $A = v + U$  (wenn  $A \neq \emptyset$ ) oder als Lösungsmenge eines linearen Gleichungssystems beschreiben. Eine (affine) Gerade im  $\mathbb{R}^3$  kann also in der Form  $g = \mathbf{x}_0 + \langle \mathbf{y} \rangle$  beschrieben werden (mit „Aufpunkt“  $\mathbf{x}_0$  und „Richtungsvektor“  $\mathbf{y} \neq \mathbf{0}$ ) oder in der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned}$$

(mit linear unabhängigen Vektoren  $(a_{11}, a_{12}, a_{13})$  und  $(a_{21}, a_{22}, a_{23})$ ). Diese zweite Form kann man auch so interpretieren, dass man  $g$  als Schnitt zweier nicht paralleler Ebenen darstellt, denn jede der beiden Gleichungen beschreibt eine Ebene. ♣

Analog zur linearen Hülle kann man jetzt die *affine Hülle* einer Teilmenge  $T \subset V$  definieren als den kleinsten affinen Unterraum, der  $T$  enthält (formal: als Durchschnitt

aller affinen Unterräume, die  $T$  enthalten). Auf dieselbe Weise, wie wir gezeigt haben, dass die lineare Hülle von  $T$  genau aus allen Linearkombinationen von Elementen von  $T$  besteht, sieht man, dass die affine Hülle von  $T$  genau aus allen affinen Linearkombinationen von Elementen von  $T$  besteht. Es gilt  $\dim(\text{affine Hülle von } T) \leq \#T - 1$ . Zum Beispiel ist die affine Hülle von drei verschiedenen Punkten im  $\mathbb{R}^3$  entweder eine Gerade (wenn die drei Punkte auf einer Geraden liegen) oder eine Ebene, nämlich die durch die drei Punkte aufgespannte Ebene.

Ein anderes Beispiel ist die affine Hülle  $A$  der Vereinigung  $g_1 \cup g_2$  zweier Geraden im  $\mathbb{R}^3$ . Im Fall  $g_1 = g_2$  ist  $A = g_1 = g_2$ . Schneiden sich  $g_1$  und  $g_2$  in einem Punkt, dann spannen sie gemeinsam die Ebene  $A$  auf (die die Form  $A = \mathbf{x}_0 + \langle \mathbf{y}_1, \mathbf{y}_2 \rangle$  hat, wobei  $g_1 \cap g_2 = \{\mathbf{x}_0\}$  ist und  $\mathbf{y}_1, \mathbf{y}_2$  Richtungsvektoren von  $g_1$  und  $g_2$  sind). Sind  $g_1$  und  $g_2$  parallel, dann spannen sie ebenfalls eine Ebene auf (finden Sie eine Beschreibung dieser Ebene!). Sind  $g_1$  und  $g_2$  schließlich windschief, dann ist  $A = \mathbb{R}^3$ .

Sind  $A_1 = v_1 + U_1$  und  $A_2 = v_2 + U_2$  endlich-dimensionale und nicht-leere affine Unterräume eines Vektorraums  $V$  und ist  $A$  die affine Hülle von  $A_1 \cup A_2$ , dann kann man folgende Dimensionsformel zeigen:

$$\dim A = \begin{cases} \dim A_1 + \dim A_2 - \dim(A_1 \cap A_2), & \text{falls } A_1 \cap A_2 \neq \emptyset; \\ \dim A_1 + \dim A_2 - \dim(U_1 \cap U_2) + 1, & \text{falls } A_1 \cap A_2 = \emptyset. \end{cases}$$

## 23. DER DUALRAUM

Wir hatten im ersten Semester schon gesehen, dass die Menge aller Homomorphismen  $f : V \rightarrow W$  zwischen zwei  $K$ -Vektorräumen  $V$  und  $W$  selbst wieder die Struktur eines  $K$ -Vektorraums  $\text{Hom}(V, W)$  hat (Satz 9.21). Ein besonders wichtiger Spezialfall tritt auf, wenn  $W = K$  ist.

\* **23.1. Definition.** Sei  $V$  ein  $K$ -Vektorraum. Eine lineare Abbildung  $\phi : V \rightarrow K$  heißt auch eine *Linearform* auf  $V$ . Der Vektorraum  $V^* = \text{Hom}(V, K)$  heißt der *Dualraum* von  $V$ . **DEF**  
Linearform  
Dualraum  
◇

Die Elemente von  $V^*$  sind also gerade die Linearformen auf  $V$ . Wir erinnern uns an die Definition der Vektorraumstruktur von  $V^*$ : Für Linearformen  $\phi, \phi' \in V^*$  und  $\lambda \in K$  ist  $\phi + \phi'$  die Linearform  $v \mapsto \phi(v) + \phi'(v)$  und  $\lambda\phi$  ist die Linearform  $v \mapsto \lambda\phi(v)$ .

**23.2. Beispiele.** Auf dem Standardvektorraum  $K^n$  sind die *Koordinatenabbildungen* oder *Projektionen* **BSP**  
Linearformen

$$\text{pr}_j : (x_1, x_2, \dots, x_n) \mapsto x_j, \quad j \in \{1, 2, \dots, n\},$$

Linearformen.

Ist  $V$  ein Vektorraum von reellen Funktionen auf einer Menge  $X$ , dann ist für jedes  $x \in X$  die *Auswertungsabbildung*

$$\text{ev}_x : f \mapsto f(x)$$

eine Linearform auf  $V$ .

Ist  $V^*$  der Dualraum eines Vektorraums  $V$ , dann ist zu jedem  $v \in V$  die *Auswertungsabbildung*

$$\text{ev}_v : \phi \mapsto \phi(v)$$

eine Linearform auf  $V^*$ , also ein Element des *Bidualraums*  $V^{**} = (V^*)^*$ . ♣

Wir erinnern uns daran, dass eine lineare Abbildung durch ihre Werte auf einer Basis eindeutig bestimmt ist und dass diese Werte beliebig vorgegeben werden können (Satz 9.11). Daraus ergibt sich der folgende wichtige Satz.

\* **23.3. Satz.** Sei  $V$  ein endlich-dimensionaler Vektorraum mit Basis  $(v_1, v_2, \dots, v_n)$ . Dann gibt es eine eindeutig bestimmte Basis  $(v_1^*, v_2^*, \dots, v_n^*)$  des Dualraums  $V^*$ , so dass für alle  $i, j \in \{1, 2, \dots, n\}$  gilt **SATZ**  
Existenz und  
Eindeutigkeit  
der dualen  
Basis

$$v_i^*(v_j) = \delta_{ij} = \begin{cases} 1, & \text{falls } i = j; \\ 0, & \text{falls } i \neq j. \end{cases}$$

\* **23.4. Definition.** Die Basis  $(v_1^*, \dots, v_n^*)$  in Satz 23.3 heißt die zur Basis  $(v_1, \dots, v_n)$  **DEF**  
 duale Basis von  $V^*$ . ◇ duale Basis

Man beachte, dass jedes Element  $v_i^*$  der dualen Basis von *allen* Elementen  $v_1, \dots, v_n$  abhängt, nicht nur von  $v_i$ !

*Beweis.* Die Linearformen  $v_i^*$  sind durch die angegebene Bedingung eindeutig festgelegt, denn wir schreiben ihre Bilder auf einer Basis von  $V$  vor. Es bleibt zu zeigen, dass diese Elemente  $v_1^*, v_2^*, \dots, v_n^* \in V^*$  eine Basis bilden.

Wir zeigen zuerst, dass sie linear unabhängig sind. Seien dazu  $\lambda_1, \lambda_2, \dots, \lambda_n$  Skalare mit  $\lambda_1 v_1^* + \lambda_2 v_2^* + \dots + \lambda_n v_n^* = \mathbf{0}$ . Wir werten die links stehende Linearform auf  $v_1, v_2, \dots, v_n$  aus:

$$\begin{aligned} 0 &= \mathbf{0}(v_j) = (\lambda_1 v_1^* + \lambda_2 v_2^* + \dots + \lambda_n v_n^*)(v_j) \\ &= \lambda_1 v_1^*(v_j) + \dots + \lambda_{j-1} v_{j-1}^*(v_j) + \lambda_j v_j^*(v_j) + \lambda_{j+1} v_{j+1}^*(v_j) + \dots + \lambda_n v_n^*(v_j) \\ &= \lambda_1 \cdot 0 + \dots + \lambda_{j-1} \cdot 0 + \lambda_j \cdot 1 + \lambda_{j+1} \cdot 0 + \dots + \lambda_n \cdot 0 \\ &= \lambda_j. \end{aligned}$$

Also sind alle  $\lambda_j = 0$ , und die lineare Unabhängigkeit ist bewiesen.

Wir müssen noch zeigen, dass die  $v_i^*$  ein Erzeugendensystem von  $V^*$  sind. Sei dazu  $\phi \in V^*$  beliebig. Dann gilt

$$\phi = \phi(v_1)v_1^* + \phi(v_2)v_2^* + \dots + \phi(v_n)v_n^*,$$

denn beide Seiten sind Linearformen, die auf der gegebenen Basis von  $V$  dieselben Werte annehmen: Wie eben gilt

$$(\phi(v_1)v_1^* + \phi(v_2)v_2^* + \dots + \phi(v_n)v_n^*)(v_j) = \phi(v_j).$$

Das zeigt, dass  $\phi$  eine Linearkombination von  $v_1^*, v_2^*, \dots, v_n^*$  ist. □

Wenn  $V$  nicht endlich-dimensional ist, dann kann man zu einer Basis  $(b_i)_{i \in I}$  von  $V$  immer noch eine Familie  $(b_i^*)_{i \in I}$  in  $V^*$  konstruieren, die  $b_i^*(b_j) = \delta_{ij}$  erfüllt. Diese Familie ist linear unabhängig (mit demselben Beweis wie eben), aber kein Erzeugendensystem von  $V^*$ , denn jede Linearkombination (die ja immer nur endlich viele Vektoren involviert) der  $b_i^*$  nimmt nur auf endlich vielen Basiselementen  $b_j$  von null verschiedene Werte an. Es gibt aber zu *jeder* Wahl von Werten auf allen  $b_j$  eine zugehörige Linearform; zum Beispiel gibt es  $\phi \in V^*$  mit  $\phi(b_j) = 1$  für alle  $j \in I$ , aber  $\phi \notin \langle \{b_i^* \mid i \in I\} \rangle$ .

**23.5. Beispiel.** Die duale Basis zur Standardbasis  $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$  von  $K^n$  besteht gerade aus den Koordinatenabbildungen  $(\text{pr}_1, \text{pr}_2, \dots, \text{pr}_n)$ . **BSP**  
♣ duale Basis der Standardbasis

Aus dem Satz ergibt sich unmittelbar:

**23.6. Folgerung.** *Ist  $V$  ein endlich-dimensionaler Vektorraum, dann gilt*  

$$\dim V = \dim V^* .$$
 **FOLG**  
 $V \cong V^*$  für  $V$  endl.-dim.

*Insbesondere sind  $V$  und  $V^*$  isomorph.*

Die Aussage von Folgerung 23.6 ist für unendlich-dimensionale Vektorräume *falsch*. Das liegt daran, dass die Dimension (als Mächtigkeit einer Basis definiert) des Dualraums  $V^*$  „unendlicher“ ist als die Dimension von  $V$  selbst. Genauer bedeutet das: Es gibt zwar injektive, aber keine surjektiven Abbildungen von einer Basis von  $V$  in eine Basis von  $V^*$ . Diese Aussage ist verwandt mit dem Satz aus der Mengenlehre, dass die Potenzmenge  $\mathcal{P}(X)$  einer Menge  $X$  stets echt mächtiger ist als  $X$ : Es gibt keine surjektive Abbildung

$X \rightarrow \mathcal{P}(X)$ . Zum Beweis sei  $f : X \rightarrow \mathcal{P}(X)$  irgendeine Abbildung. Wir betrachten die Teilmenge

$$T = \{x \in X \mid x \notin f(x)\} \subset X.$$

(Als Element von  $\mathcal{P}(X)$  ist  $f(x)$  eine Teilmenge von  $X$ , also ist die Bedingung „ $x \notin f(x)$ “ sinnvoll. Die Konstruktion ist ähnlich wie in der Russellschen Antinomie, die am Ende des Abschnitts über Mengenlehre in der Linearen Algebra I im Kleingedruckten erwähnt wird.) Dann ist  $T \in \mathcal{P}(X)$  nicht im Bild von  $f$ . Denn wäre  $T = f(x)$  für ein  $x \in X$ , dann erhielte man den Widerspruch

$$x \in T \iff x \notin f(x) \iff x \notin T$$

(die erste Äquivalenz ist die Definition von  $T$ , die zweite folgt aus  $f(x) = T$ ). Der Zusammenhang ergibt sich so: Sei  $B$  eine Basis von  $V$ . Dann gibt es zu jeder Teilmenge  $T$  von  $B$  eine eindeutig bestimmte Linearform  $\phi_T \in V^*$  mit  $\phi_T(b) = 1$  für alle  $b \in T$  und  $\phi_T(b) = 0$  für alle  $b \in B \setminus T$ . Die Menge  $\mathcal{T} = \{\phi_T \mid T \subset B\} \subset V^*$  hat die Mächtigkeit von  $\mathcal{P}(B)$ . Die  $\phi_T$  sind zwar nicht linear unabhängig (zum Beispiel gilt  $\phi_T + \phi_{T'} - \phi_{T \cup T'} - \phi_{T \cap T'} = \mathbf{0}$ ), aber man kann zeigen, dass  $\mathcal{T}$  eine linear unabhängige Teilmenge gleicher Mächtigkeit enthält (das kommt daher, dass jede lineare Relation nur endlich viele  $\phi_T$  enthält). Es folgt, dass jede Basis von  $V^*$  echt mächtiger sein muss als  $B$ .

Ein Isomorphismus  $V \rightarrow V^*$  ist — nach Wahl einer Basis  $(v_1, v_2, \dots, v_n)$  von  $V$  — dadurch gegeben, dass man  $v_i$  auf  $v_i^*$  abbildet. Der Isomorphismus hängt von der Wahl der Basis ab (man kann leicht Beispiele finden, die das belegen), er ist also nicht „natürlich“ oder *kanonisch*. Im Unterschied dazu gibt es eine kanonische lineare Abbildung in den Bidualraum  $V^{**}$ .

Wir formulieren vorher noch eine Aussage, die wir später brauchen.

### 23.7. Lemma.

- (1) Seien  $V$  und  $W$  Vektorräume und  $U \subset V$  ein Untervektorraum. Ist außerdem  $f : U \rightarrow W$  eine lineare Abbildung, dann kann man  $f$  zu einer linearen Abbildung  $F : V \rightarrow W$  fortsetzen (es gilt also  $F|_U = f$ ).
- (2) Ist  $V$  ein Vektorraum und  $\mathbf{0} \neq v \in V$ , dann gibt es  $\phi \in V^*$  mit  $\phi(v) = 1$ . Allgemeiner gilt: Ist  $U \subset V$  ein Untervektorraum und  $v \in V \setminus U$ , dann gibt es  $\phi \in V^*$  mit  $\phi|_U = \mathbf{0}$  und  $\phi(v) = 1$ .

**LEMMA**  
Fortsetzung  
linearer  
Abbildungen

*Beweis.*

- (1) Wir verwenden, dass es ein Komplement  $U'$  von  $U$  in  $V$  gibt. Das haben wir nur für  $V$  endlich-dimensional bewiesen; es gilt jedoch auch allgemein. (Dafür braucht man den Basisergänzungssatz für unendliche Mengen und damit das Auswahlaxiom.) Jedes Element  $v$  von  $V$  lässt sich dann eindeutig schreiben als  $v = u + u'$  mit  $u \in U$  und  $u' \in U'$ ; wir definieren  $F$  durch  $F(v) = f(u)$ .  $F$  ist linear als Komposition der Projektion auf  $U$  (bezüglich der Zerlegung  $V = U \oplus U'$ ) und der linearen Abbildung  $f$ ; es ist klar, dass  $F|_U = f$  gilt.
- (2) Wir wenden den ersten Teil an auf  $U = \langle v \rangle$  und  $f : U \rightarrow K$ ,  $\lambda v \mapsto \lambda$ . Für die allgemeinere Aussage sei  $\pi : V \rightarrow V/U$  der kanonische Epimorphismus. Die eben bewiesene Aussage liefert eine Linearform  $\bar{\phi} : V/U \rightarrow K$  mit  $\bar{\phi}(\pi(v)) = 1$ . Dann leistet  $\phi = \bar{\phi} \circ \pi : V \rightarrow K$  das Gewünschte.  $\square$

\* **23.8. Satz.** *Sei  $V$  ein  $K$ -Vektorraum. Dann ist die folgende Abbildung ein injektiver Homomorphismus:*

$$\alpha_V : V \longrightarrow V^{**}, \quad v \longmapsto (\text{ev}_v : \phi \mapsto \phi(v)).$$

*Ist  $V$  endlich-dimensional, dann ist  $\alpha_V$  ein Isomorphismus.*

**SATZ**  
kanon. Abb.  
in den  
Bidualraum

*Beweis.* Es sind verschiedene Aussagen zu zeigen.

- $\text{ev}_v \in V^{**}$  (siehe Beispiel 23.2):  $\text{ev}_v$  ist eine Abbildung  $V^* \rightarrow K$ ; wir müssen zeigen, dass  $\text{ev}_v$  linear ist:

$$\text{ev}_v(\phi + \phi') = (\phi + \phi')(v) = \phi(v) + \phi'(v) = \text{ev}_v(\phi) + \text{ev}_v(\phi')$$

und analog für die Skalarmultiplikation. (Hier benutzen wir die Definition der Vektorraumstruktur von  $V^*$ .)

- $\alpha_V$  ist linear:  $\alpha_V(v+v')$  bildet  $\phi \in V^*$  auf  $\phi(v+v') = \phi(v) + \phi(v') = \text{ev}_v(\phi) + \text{ev}_{v'}(\phi)$  ab, hat also denselben Effekt wie  $\alpha_V(v) + \alpha_V(v')$ . Analog sehen wir, dass  $\alpha_V(\lambda v) = \text{ev}_{\lambda v}$  die Abbildung  $\phi \mapsto \phi(\lambda v) = \lambda\phi(v)$  ist und daher mit  $\lambda\alpha_V(v)$  übereinstimmt. (Hier benutzen wir, dass die Elemente von  $V^*$  lineare Abbildungen sind, und die Definition der Vektorraumstruktur von  $V^{**}$ .)
- $\alpha_V$  ist injektiv: Wir zeigen  $\ker(\alpha_V) = \{\mathbf{0}\}$ . Sei  $v \neq \mathbf{0}$ . Nach Lemma 23.7 gibt es  $\phi \in V^*$  mit  $(\alpha_V(v))(\phi) = \phi(v) = 1 \neq 0$ , also ist  $\alpha_V(v) \neq \mathbf{0}$  und damit  $v \notin \ker(\alpha_V)$ . Es bleibt also nur der Nullvektor als einzig mögliches Element von  $\ker(\alpha_V)$ .
- $\alpha_V$  ist Isomorphismus, falls  $\dim V < \infty$ : In diesem Fall gilt nach Folgerung 23.6  $\dim V = \dim V^* = \dim V^{**}$ . Als injektive lineare Abbildung zwischen endlich-dimensionalen Vektorräumen derselben Dimension muss  $\alpha_V$  dann ein Isomorphismus sein (Folgerung 9.14).  $\square$

Ist  $V$  endlich-dimensional, dann kann man also  $V$  und  $V^{**}$  durch den kanonischen Isomorphismus  $\alpha_V$  miteinander identifizieren und damit  $V$  als den Dualraum von  $V^*$  betrachten. Das wird zum Beispiel durch die nächste Aussage illustriert.

**23.9. Folgerung.** *Sei  $V$  ein endlich-dimensionaler Vektorraum; sei  $(v_1^*, \dots, v_n^*)$  eine Basis von  $V^*$ . Dann gibt es eine eindeutig bestimmte Basis  $(v_1, v_2, \dots, v_n)$  von  $V$ , sodass  $(v_1^*, v_2^*, \dots, v_n^*)$  die zu  $(v_1, v_2, \dots, v_n)$  duale Basis ist.*

**FOLG**  
Basis dual zu  
Basis von  $V^*$

*Beweis.* Sei  $(v_1^{**}, v_2^{**}, \dots, v_n^{**})$  die zu  $(v_1^*, v_2^*, \dots, v_n^*)$  duale Basis von  $V^{**}$ . Da  $\alpha_V$  ein Isomorphismus ist, gibt es eindeutig bestimmte  $v_1, v_2, \dots, v_n \in V$  mit  $\alpha_V(v_j) = v_j^{**}$  für alle  $j \in \{1, 2, \dots, n\}$ ;  $(v_1, v_2, \dots, v_n)$  ist eine Basis von  $V$ . Außerdem gilt für alle  $i, j \in \{1, 2, \dots, n\}$ :

$$v_i^*(v_j) = (\alpha_V(v_j))(v_i^*) = v_j^{**}(v_i^*) = \delta_{ji} = \delta_{ij},$$

also ist  $(v_1^*, v_2^*, \dots, v_n^*)$  die zu  $(v_1, v_2, \dots, v_n)$  duale Basis. Die Eindeutigkeit folgt aus der Eindeutigkeit der  $v_j^{**}$ .  $\square$

Sind  $(v_1, v_2, \dots, v_n)$  und  $(v_1^*, v_2^*, \dots, v_n^*)$  zueinander duale Basen von  $V$  und  $V^*$ , dann gilt:

$$\begin{aligned} \forall v^* \in V^* : \quad v^* &= v^*(v_1) \cdot v_1^* + v^*(v_2) \cdot v_2^* + \dots + v^*(v_n) \cdot v_n^* && \text{und} \\ \forall v \in V : \quad v &= v_1^*(v) \cdot v_1 + v_2^*(v) \cdot v_2 + \dots + v_n^*(v) \cdot v_n. \end{aligned}$$

Die erste Aussage haben wir im Beweis von Satz 23.3 verwendet, die zweite folgt durch Vertauschen der Rollen von  $V$  und  $V^*$ .

**23.10. Beispiel.** Sei  $V = K[X]_{<n}$  der Vektorraum der Polynome über  $K$  vom Grad  $< n$ . Seien  $a_1, a_2, \dots, a_n \in K$  paarweise verschieden. Dann wissen wir, dass die Auswertungsabbildungen  $\text{ev}_{a_i}$  für  $i \in \{1, 2, \dots, n\}$  linear unabhängig sind; sie bilden also eine Basis von  $V^*$ . Welche Basis von  $V$  ist dazu dual? Wenn diese Basis  $(p_1, p_2, \dots, p_n)$  ist, dann muss gelten  $p_i(a_j) = \delta_{ij}$ , also ist

**BSP**  
Interpolationspolynome als duale Basis

$$p_i = \prod_{j \neq i} \frac{X - a_j}{a_i - a_j}.$$

Die obige Relation liefert dann für  $p \in V$  beliebig, dass

$$p = p(a_1) \cdot p_1 + p(a_2) \cdot p_2 + \dots + p(a_n) \cdot p_n$$

ist — wir erhalten wieder die Lagrangesche Interpolationsformel, vergleiche Beispiel 9.15. ♣

Wir haben gesehen, wie man Vektorräume und Basen „dualisieren“ kann. Jetzt erweitern wir das auf lineare Abbildungen: Ist  $f : V \rightarrow W$  linear und  $\phi \in W^*$ , dann ist  $f^\top(\phi) = \phi \circ f : V \rightarrow K$  eine Linearform auf  $V$ :

$$\begin{array}{ccc} V & & \\ f \downarrow & \searrow f^\top(\phi) & \\ W & \xrightarrow{\phi} & K \end{array}$$

Wir erhalten eine Abbildung  $f^\top : W^* \rightarrow V^*$ .

**\* 23.11. Definition.** Ist  $f : V \rightarrow W$  linear, dann heißt  $f^\top : W^* \rightarrow V^*$ ,  $\phi \mapsto \phi \circ f$ , die zu  $f$  *duale* oder *transponierte* lineare Abbildung. ◇

**DEF**  
duale lineare Abbildung

Dass  $f^\top$  tatsächlich linear ist, folgt aus der Definition der Vektorraumstruktur auf  $W^*$ :

$$f^\top(\phi + \phi') = (\phi + \phi') \circ f = \phi \circ f + \phi' \circ f = f^\top(\phi) + f^\top(\phi')$$

und

$$f^\top(\lambda\phi) = (\lambda\phi) \circ f = \lambda(\phi \circ f) = \lambda f^\top(\phi).$$

Auch die Bezeichnung  $f^*$  ist gebräuchlich.

Die Notation  $f^\top$  erklärt sich durch die folgende Aussage.

**23.12. Satz.** Seien  $V$  und  $W$  endlich-dimensionale  $K$ -Vektorräume, seien  $B$  und  $B'$  Basen von  $V$  und  $W$  und seien  $B^*$  und  $B'^*$  die dazu dualen Basen von  $V^*$  und  $W^*$ . Sei  $f : V \rightarrow W$  linear und  $A = \text{Mat}_{B, B'}(f)$  die  $f$  bezüglich der Basen  $B$  und  $B'$  darstellende Matrix. Dann gilt

**SATZ**  
 $f^\top$  und  $A^\top$

$$\text{Mat}_{B'^*, B^*}(f^\top) = A^\top.$$

*Beweis.* Seien  $B = (b_1, b_2, \dots, b_n)$ ,  $B' = (b'_1, b'_2, \dots, b'_m)$  und  $A = (a_{ij})$ . Wir schreiben  $B^* = (b_1^*, b_2^*, \dots, b_n^*)$  und  $B'^* = (b_1'^*, b_2'^*, \dots, b_m'^*)$ . Dann ist

$$f(b_j) = a_{1j}b'_1 + a_{2j}b'_2 + \dots + a_{mj}b'_m,$$

also ist

$$a_{ij} = b_i^*(f(b_j)) = (b_i^* \circ f)(b_j) = (f^\top(b_i^*))(b_j).$$

Auf der anderen Seite gilt mit  $\text{Mat}_{B'^*, B^*}(f^\top) = (a'_{ij})$ :

$$f^\top(b_i^*) = a'_{1i}b_1^* + a'_{2i}b_2^* + \dots + a'_{ni}b_n^*;$$

Anwenden auf  $b_j$  ergibt

$$a_{ij} = (f^\top(b_i^*))(b_j) = a'_{ji},$$

also sind die beiden Matrizen zueinander transponiert.  $\square$

**23.13. Lemma.** Sind  $V$  und  $W$  zwei  $K$ -Vektorräume, dann ist

$$\Phi : \text{Hom}(V, W) \longrightarrow \text{Hom}(W^*, V^*), \quad f \longmapsto f^\top$$

eine injektive lineare Abbildung. Sind  $V$  und  $W$  beide endlich-dimensional, dann ist  $\Phi$  ein Isomorphismus.

**LEMMA**

$f \mapsto f^\top$

*Beweis.*  $\Phi$  ist linear, denn für  $\phi \in W^*$  und  $f, g \in \text{Hom}(V, W)$  und  $\lambda \in K$  gilt

$$(f + g)^\top(\phi) = \phi \circ (f + g) = \phi \circ f + \phi \circ g = f^\top(\phi) + g^\top(\phi) = (f^\top + g^\top)(\phi),$$

also ist  $\Phi(f + g) = (f + g)^\top = f^\top + g^\top = \Phi(f) + \Phi(g)$ , und

$$(\lambda f)^\top(\phi) = \phi \circ (\lambda f) = \lambda(\phi \circ f) = \lambda f^\top(\phi) = (\lambda f^\top)(\phi),$$

also ist  $\Phi(\lambda f) = (\lambda f)^\top = \lambda f^\top = \lambda \Phi(f)$ .

$\Phi$  ist injektiv, denn für  $f \in \text{Hom}(V, W)$  mit  $\Phi(f) = f^\top = \mathbf{0}$  gilt  $\phi \circ f = \mathbf{0}$  für alle  $\phi \in W^*$ . Da es zu jedem  $\mathbf{0} \neq w \in W$  ein  $\phi \in W^*$  gibt mit  $\phi(w) \neq 0$  (Lemma 23.7), folgt  $f(v) = \mathbf{0}$  für alle  $v \in V$ , also  $f = \mathbf{0}$ .

Sind  $V$  und  $W$  beide endlich-dimensional, dann gilt (Satz 9.23)

$$\dim \text{Hom}(W^*, V^*) = \dim W^* \cdot \dim V^* = \dim W \cdot \dim V = \dim \text{Hom}(V, W),$$

also ist  $\Phi$  als injektive lineare Abbildung zwischen zwei endlich-dimensionalen Vektorräumen derselben Dimension ein Isomorphismus.  $\square$

Wir zeigen noch einige weitere einfache Eigenschaften der transponierten Abbildung.

**23.14. Lemma.**

- (1) Ist  $V$  ein Vektorraum, dann gilt  $\text{id}_V^\top = \text{id}_{V^*}$ .
- (2) Sind  $V, V', V''$  Vektorräume und  $f : V \rightarrow V'$  und  $g : V' \rightarrow V''$  lineare Abbildungen, dann gilt  $(g \circ f)^\top = f^\top \circ g^\top$ .
- (3) Ist  $f : V \rightarrow W$  ein Isomorphismus, dann ist auch  $f^\top$  ein Isomorphismus und es gilt  $(f^\top)^{-1} = (f^{-1})^\top$ .

**LEMMA**

Eigenschaften von  $f^\top$

*Beweis.*

(1) Für  $\phi \in V^*$  ist  $\text{id}_V^\top(\phi) = \phi \circ \text{id}_V = \phi$ .

(2) Für  $\phi \in (V'')^*$  ist

$$(g \circ f)^\top(\phi) = \phi \circ (g \circ f) = (\phi \circ g) \circ f = g^\top(\phi) \circ f = f^\top(g^\top(\phi)) = (f^\top \circ g^\top)(\phi).$$

(3) Nach den beiden ersten Teilen gilt

$$f^\top \circ (f^{-1})^\top = (f^{-1} \circ f)^\top = \text{id}_V^\top = \text{id}_{V^*}$$

und

$$(f^{-1})^\top \circ f^\top = (f \circ f^{-1})^\top = \text{id}_W^\top = \text{id}_{W^*},$$

woraus die Behauptungen folgen.  $\square$

Der Beweis der folgenden Aussagen, die Zusammenhänge zwischen  $f^\top$  und der kanonischen Injektion  $\alpha_V$  aufzeigen, ist eine Übungsaufgabe.

**23.15. Lemma.**

**LEMMA**  
 $f^\top$  und  $\alpha_V$

(1) Sei  $V$  ein Vektorraum. Dann gilt  $\alpha_V^\top \circ \alpha_{V^*} = \text{id}_{V^*}$ .

(2) Sei  $f : V \rightarrow W$  linear. Dann kommutiert das folgende Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \alpha_V \downarrow & & \downarrow \alpha_W \\ V^{**} & \xrightarrow{f^{\top\top}} & W^{**} \end{array}$$

es gilt also  $f^{\top\top} \circ \alpha_V = \alpha_W \circ f$ .

Zum Abschluss dieses Abschnitts untersuchen wir noch, wie sich Untervektorräume unter Dualisierung verhalten.

**23.16. Definition.** Sei  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum. Dann heißt der Untervektorraum

**DEF**  
Annulator

$$U^\circ = \{\phi \in V^* \mid \phi(u) = 0 \text{ für alle } u \in U\} = \bigcap_{u \in U} \ker(\text{ev}_u) \subset V^*$$

der *Annulator* von  $U$  (in  $V^*$ ).  $\diamond$

Trivialbeispiele für Annulatoren sind  $\{\mathbf{0}_V\}^\circ = V^*$  und  $V^\circ = \{\mathbf{0}_{V^*}\}$ .

**23.17. Lemma.** Seien  $V$  ein Vektorraum und  $U_1, U_2 \subset V$  Untervektorräume. Dann gilt

**LEMMA**  
Eindeutigkeit  
von  
Annulatoren

$$U_1 \subset U_2 \iff U_1^\circ \supset U_2^\circ.$$

Insbesondere gilt auch  $U_1^\circ = U_2^\circ \Rightarrow U_1 = U_2$ .

*Beweis.* Die Richtung „ $\Rightarrow$ “ folgt unmittelbar aus der Definition von  $U^\circ$ : Ist  $\phi \in U_2^\circ$ , also  $\phi|_{U_2} = \mathbf{0}$ , dann gilt auch  $\phi|_{U_1} = \mathbf{0}$ , also  $\phi \in U_1^\circ$ . Für die Gegenrichtung zeigen wir die Kontraposition „ $U_1 \not\subset U_2 \Rightarrow U_1^\circ \not\supset U_2^\circ$ “. Nach Voraussetzung gibt es  $u \in U_1$  mit  $u \notin U_2$ . Dann gibt es nach Lemma 23.7 eine Linearform  $\phi \in V^*$  mit  $\phi|_{U_2} = \mathbf{0}$  und  $\phi(u) \neq 0$ . Das bedeutet  $\phi \in U_2^\circ$ , aber  $\phi \notin U_1^\circ$ , also folgt  $U_1^\circ \not\supset U_2^\circ$ .  $\square$

**23.18. Satz.** Sei  $V$  ein endlich-dimensionaler Vektorraum und  $U \subset V$  ein Untervektorraum. Dann gilt

$$\dim U + \dim U^\circ = \dim V \quad \text{und} \quad \alpha_V(U) = U^{\circ\circ}.$$

**SATZ**  
Dimension von  
Annulatoren

Wenn wir  $V$  und  $V^{**}$  via  $\alpha_V$  identifizieren, dann können wir schreiben  $U^{\circ\circ} = U$ .

*Beweis.* Sei  $(b_1, b_2, \dots, b_m, \dots, b_n)$  eine Basis von  $V$ , sodass  $(b_1, \dots, b_m)$  eine Basis von  $U$  ist, und sei  $(b_1^*, \dots, b_n^*)$  die dazu duale Basis von  $V^*$ . Für ein beliebiges Element  $\phi = \lambda_1 b_1^* + \dots + \lambda_n b_n^* \in V^*$  gilt

$$\begin{aligned} \phi \in U^\circ &\iff \phi(b_j) = 0 \quad \text{für alle } j \in \{1, 2, \dots, m\} \\ &\iff \lambda_j = 0 \quad \text{für alle } j \in \{1, 2, \dots, m\} \\ &\iff \phi \in \langle b_{m+1}^*, \dots, b_n^* \rangle. \end{aligned}$$

Es folgt  $U^\circ = \langle b_{m+1}^*, \dots, b_n^* \rangle$  und damit  $\dim U^\circ = n - m = \dim V - \dim U$ .

Aus den Definitionen folgt, dass  $\alpha_V(U) \subset U^{\circ\circ}$  ist: Für  $u \in U$  und  $\phi \in U^\circ$  gilt  $(\alpha_V(u))(\phi) = \phi(u) = 0$ . Da nach der ersten Aussage gilt

$\dim U^{\circ\circ} = \dim V^* - \dim U^\circ = \dim V - (\dim V - \dim U) = \dim U = \dim \alpha_V(U)$ ,  
muss Gleichheit gelten. □

Nun ergibt sich ein schöner Zusammenhang zwischen Kern und Bild, transponierter Abbildung und Annulatoren.

**23.19. Satz.** Sei  $f : V \rightarrow W$  eine lineare Abbildung. Dann gilt

$$(\ker(f))^\circ = \text{im}(f^\top) \quad \text{und} \quad (\text{im}(f))^\circ = \ker(f^\top).$$

Sind  $V$  und  $W$  endlich-dimensional, dann gilt außerdem

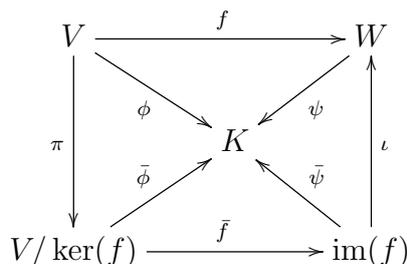
$$\text{rk}(f^\top) = \text{rk}(f).$$

**SATZ**  
lineare Abb.  
und  
Annulatoren

Wir erhalten also eine neue (und in gewisser Weise natürlichere) Begründung für die Aussage „Zeilenrang = Spaltenrang“ (Satz 11.15).

*Beweis.*

- $(\text{im}(f))^\circ = \ker(f^\top) \subset W^*$ : Sei  $\psi \in W^*$ . Dann ist  $\psi \in \ker(f^\top)$  genau dann, wenn  $f^\top(\psi) = \psi \circ f = \mathbf{0}$  ist; das bedeutet  $\psi|_{\text{im}(f)} = \mathbf{0}$ , was nach Definition genau  $\psi \in (\text{im}(f))^\circ$  heißt.
- $(\ker(f))^\circ = \text{im}(f^\top) \subset V^*$ : Wir betrachten folgendes Diagramm (vergleiche Satz 22.11):



Dabei ist  $\pi$  der kanonische Epimorphismus und  $\iota$  die Inklusionsabbildung;  $\bar{f}$  ist ein Isomorphismus.

Sei  $\phi \in V^*$ . Dann ist  $\phi \in (\ker(f))^\circ$  genau dann, wenn für alle  $v \in V$  gilt  $f(v) = \mathbf{0} \Rightarrow \phi(v) = 0$ . Das bedeutet genau  $\ker(f) \subset \ker(\phi)$ , und das

ist nach Satz 22.10 genau dann der Fall, wenn es eine lineare Abbildung  $\bar{\phi} : V/\ker(f) \rightarrow K$  gibt mit  $\phi = \bar{\phi} \circ \pi$ . Da  $\bar{f}$  ein Isomorphismus ist, ist das äquivalent zur Existenz von  $\bar{\psi} : \text{im}(f) \rightarrow K$  mit  $\phi = \bar{\psi} \circ \bar{f} \circ \pi$  (setze  $\bar{\psi} = \bar{\phi} \circ \bar{f}^{-1}$ ). Nach Lemma 23.7 ist das genau dann der Fall, wenn es  $\psi \in W^*$  gibt mit  $\phi = \psi \circ \iota \circ \bar{f} \circ \pi = \psi \circ f$ . Das heißt aber gerade, dass  $\phi \in \text{im}(f^\top)$  ist.

- $\text{rk}(f^\top) = \text{rk}(f)$ : Nach Satz 23.18 und dem eben Bewiesenen gilt

$$\text{rk}(f^\top) = \dim \text{im}(f^\top) = \dim(\ker(f))^\circ = \dim V - \dim \ker(f) = \text{rk}(f);$$

die letzte Gleichheit ist der Rangatz 9.18. □

**23.20. Beispiel.** Die erste Aussage in Satz 23.19 lässt sich so formulieren: Sei  $f : V \rightarrow W$  linear. Zu einer Linearform  $\phi : V \rightarrow K$  gibt es genau dann eine Linearform  $\psi : W \rightarrow K$  mit  $\phi = \psi \circ f$  (d.h.,  $\phi(v)$  „hängt nur von  $f(v)$  ab“), wenn  $\phi$  auf dem Kern von  $f$  die Nullabbildung ist. **BSP**  
**HDI**

Als Beispiel betrachten wir die Ableitungsabbildung  $D : \mathcal{C}^1(\mathbb{R}) \rightarrow \mathcal{C}^0(\mathbb{R})$ ,  $f \mapsto f'$ . Dabei bezeichnet  $\mathcal{C}^n(\mathbb{R})$  wie üblich den Vektorraum der  $n$ -mal stetig differenzierbaren reellen Funktionen auf  $\mathbb{R}$ . Die Ableitung einer Funktion verschwindet genau dann, wenn die Funktion konstant ist, also besteht  $\ker(D)$  gerade aus den konstanten Funktionen.

Seien  $a, b \in \mathbb{R}$  mit  $a < b$ . Dann verschwindet die Linearform  $\text{ev}_b - \text{ev}_a : \mathcal{C}^1(\mathbb{R}) \rightarrow \mathbb{R}$ ,  $f \mapsto f(b) - f(a)$ , auf den konstanten Funktionen. Nach unserem Satz muss es also eine Linearform  $\psi : \mathcal{C}^0(\mathbb{R}) \rightarrow \mathbb{R}$  geben, sodass  $f(b) - f(a) = \psi(f')$  gilt für alle  $f \in \mathcal{C}^1(\mathbb{R})$ . Tatsächlich sagt der Hauptsatz der Differential- und Integralrechnung, dass  $\psi$  gegeben ist durch

$$\psi(h) = \int_a^b h(x) dx, \quad \text{denn} \quad f(b) - f(a) = \int_a^b f'(x) dx. \quad \clubsuit$$

**23.21. Folgerung.** Sei  $f : V \rightarrow W$  eine lineare Abbildung. Dann gilt **FOLG**

$$f \text{ injektiv} \iff f^\top \text{ surjektiv} \quad \text{und} \quad f \text{ surjektiv} \iff f^\top \text{ injektiv}.$$

*Beweis.*  $f$  injektiv ist äquivalent mit  $\ker(f) = \{\mathbf{0}\}$ ; es folgt  $\text{im}(f^\top) = \{\mathbf{0}\}^\circ = V^*$ , also ist  $f^\top$  surjektiv. Umgekehrt folgt aus  $f^\top$  surjektiv, dass  $(\ker(f))^\circ = V^*$  ist; dann muss  $\ker(f) = \{\mathbf{0}\}$  sein (das folgt aus Lemma 23.17), also ist  $f$  injektiv.

$f$  surjektiv bedeutet  $\text{im}(f) = W$ , also ist  $\ker(f^\top) = W^\circ = \{\mathbf{0}\}$ , und  $f^\top$  ist injektiv. Ist umgekehrt  $f^\top$  injektiv, dann ist  $(\text{im}(f))^\circ = \{\mathbf{0}\}$ . Nach Lemma 23.17 folgt  $\text{im}(f) = W$ , also ist  $f$  surjektiv. □

**23.22. Beispiel.** Sei  $V \subset \mathbb{R}^{\mathbb{N}}$  der reelle Vektorraum der beschränkten Folgen reeller Zahlen. Wir können die hier bewiesenen Aussagen verwenden, um zu zeigen, dass es eine Linearform  $m : V \rightarrow \mathbb{R}$  gibt mit folgenden Eigenschaften:

$$m((1)_{n \in \mathbb{N}}) = 1 \quad \text{und} \quad m((a_{n+1})_{n \in \mathbb{N}}) = m((a_n)_{n \in \mathbb{N}}) \quad \text{für alle } (a_n)_{n \in \mathbb{N}} \in V.$$

Dies sind Eigenschaften, die ein „Mittelwert“ von  $(a_n)$  haben sollte.

Wir betrachten den „Verschiebungsoperator“  $T : V \rightarrow V$ ,  $(a_n)_{n \in \mathbb{N}} \mapsto (a_{n+1})_{n \in \mathbb{N}}$ , und den „Differenzenoperator“  $\Delta = T - \text{id}_V$ . Letzterer ist nicht surjektiv, denn die konstante Folge  $(1)_{n \in \mathbb{N}}$  ist nicht in  $\text{im}(\Delta)$  (eine beschränkte Folge kann nicht konstante Differenz 1 haben). Also ist  $\Delta^\top = T^\top - \text{id}_{V^*}$  nicht injektiv, was bedeutet, dass es Linearformen  $\mathbf{0} \neq \phi \in V^*$  gibt mit  $\phi \circ T = \phi$ . Das ist genau die zweite Bedingung an unseren Mittelwert  $m$ . Es bleibt zu zeigen, dass man  $\phi$  so wählen kann, dass  $c = \phi((1)_{n \in \mathbb{N}}) \neq 0$  ist, denn dann erfüllt  $m = c^{-1}\phi$  beide Bedingungen. Dazu beachten wir, dass aus der strikten Inklusion  $\text{im}(\Delta) \subsetneq \text{im}(\Delta) + \langle (1)_{n \in \mathbb{N}} \rangle$  nach Lemma 23.17 eine strikte Inklusion  $(\text{im}(\Delta) + \langle (1)_{n \in \mathbb{N}} \rangle)^\circ \subsetneq \ker(\Delta^\top)$  folgt. Das bedeutet, dass es  $\phi \in \ker(\Delta^\top)$  gibt mit  $\phi((1)_{n \in \mathbb{N}}) \neq 0$ , wie gewünscht. ♣

Dass die hier bewiesene Aussage über die Existenz so eines „Mittelwerts“ alles andere als trivial ist, werden Sie feststellen, wenn Sie versuchen, so einen Mittelwert explizit zu definieren, denn das wird Ihnen nicht gelingen. (Man kann geeignete Mittelwerte für mehr oder weniger große Teilmengen von Folgen konstruieren; zum Beispiel kann man für konvergente Folgen den Grenzwert nehmen. Aber man bekommt keine explizite Definition hin, die für *alle* beschränkten Folgen funktioniert.) Das hängt damit zusammen, dass wir das Auswahlaxiom verwendet haben (es steckt in der Existenz von Komplementen in beliebigen Vektorräumen, die wir für Lemma 23.7 gebraucht haben). Es steckt hinter vielen nicht-konstruktiven Existenzaussagen, wie zum Beispiel auch der Existenz einer  $\mathbb{Q}$ -Basis von  $\mathbb{R}$  (so eine Basis lässt sich auch nicht explizit hinschreiben).

**BSP**  
Mittelwert  
beschränkter  
Folgen

24. BILINEARFORMEN UND QUADRATISCHE FORMEN

Wir werden jetzt noch einmal etwas ausführlicher über Bilinearformen sprechen (wir hatten damit bereits in §16 begonnen). Zur Erinnerung hier noch einmal die wichtigsten Definitionen und Eigenschaften.

\*

**24.1. Definition.** (Vgl. Definition 16.1) Seien  $K$  ein Körper und  $V_1, V_2, W$  drei  $K$ -Vektorräume. Eine Abbildung  $\beta : V_1 \times V_2 \rightarrow W$  heißt  $(K-)$ bilinear, wenn  $\beta$  in jedem der beiden Argumente  $K$ -linear ist, also wenn für alle  $v_1, v'_1 \in V_1, v_2, v'_2 \in V_2$  und  $\lambda \in K$  gilt

**DEF**  
bilineare Abb.  
Bilinearform

$$\begin{aligned} \beta(v_1 + v'_1, v_2) &= \beta(v_1, v_2) + \beta(v'_1, v_2), & \beta(\lambda v_1, v_2) &= \lambda \beta(v_1, v_2) \\ \beta(v_1, v_2 + v'_2) &= \beta(v_1, v_2) + \beta(v_1, v'_2), & \beta(v_1, \lambda v_2) &= \lambda \beta(v_1, v_2). \end{aligned}$$

Ist  $W = K$ , dann heißt  $\beta$  eine  $(K-)$ Bilinearform oder auch *Paarung*. Gilt außerdem  $V_1 = V_2 = V$ , dann heißt  $\beta$  eine  $(K-)$ Bilinearform auf  $V$ . Wir bezeichnen den  $K$ -Vektorraum aller Bilinearformen  $V_1 \times V_2 \rightarrow K$  mit  $\text{Bil}(V_1, V_2)$ .

Ist  $\beta : V \times V \rightarrow K$  eine Bilinearform auf  $V$ , dann heißt  $\beta$  *symmetrisch*, wenn für alle  $v_1, v_2 \in V$  gilt, dass  $\beta(v_2, v_1) = \beta(v_1, v_2)$  ist.  $\beta$  heißt *alternierend*, wenn für alle  $v \in V$  gilt, dass  $\beta(v, v) = 0$  ist.  $\diamond$

Die folgende Definition ist etwas allgemeiner als Definition 16.3.

**24.2. Definition.** Seien  $V$  und  $W$  endlich-dimensionale  $K$ -Vektorräume. Wir setzen  $\dim V = m$  und  $\dim W = n$ . Sei  $\beta : V \times W \rightarrow K$  eine Bilinearform. Seien weiter  $B = (b_1, b_2, \dots, b_m)$  eine Basis von  $V$  und  $B' = (b'_1, b'_2, \dots, b'_n)$  eine Basis von  $W$ . Dann heißt

**DEF**  
Matrix  
einer  
Bilinearform

$$\text{Mat}_{B,B'}(\beta) = (\beta(b_i, b'_j))_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{pmatrix} \beta(b_1, b'_1) & \beta(b_1, b'_2) & \cdots & \beta(b_1, b'_n) \\ \beta(b_2, b'_1) & \beta(b_2, b'_2) & \cdots & \beta(b_2, b'_n) \\ \vdots & \vdots & \ddots & \vdots \\ \beta(b_m, b'_1) & \beta(b_m, b'_2) & \cdots & \beta(b_m, b'_n) \end{pmatrix}$$

die *Matrix von  $\beta$*  bezüglich  $B$  und  $B'$ . Im Fall  $V = W$  und  $B = B'$  schreiben wir auch  $\text{Mat}_B(\beta)$  statt  $\text{Mat}_{B,B}(\beta)$ .  $\diamond$

Sind  $v = x_1 b_1 + x_2 b_2 + \dots + x_m b_m \in V$  und  $v' = y_1 b'_1 + y_2 b'_2 + \dots + y_n b'_n \in W$ , dann ist  $\beta(v, v') = \sum_{i=1}^m \sum_{j=1}^n x_i y_j \beta(b_i, b'_j)$ , was sich in folgende Matrixmultiplikation übersetzen lässt (rechts steht eine  $1 \times 1$ -Matrix, die wir mit ihrem einzigen Eintrag identifizieren):

$$\beta(v, v') = (x_1, x_2, \dots, x_m) \text{Mat}_{B,B'}(\beta) (y_1, y_2, \dots, y_n)^\top.$$

Ganz analog wie in Lemma 16.4 sieht man, dass zwei Matrizen  $A$  und  $A'$  genau dann dieselbe Bilinearform  $V \times W \rightarrow K$  bezüglich geeigneter Basen beschreiben, wenn es invertierbare Matrizen  $P \in \text{GL}(m, K)$  und  $Q \in \text{GL}(n, K)$  gibt mit

$$A' = P^\top A Q.$$

Für Bilinearformen auf einem Vektorraum  $V$ , wo man nur eine Basis (von  $V$ ) wählen kann, muss dabei  $Q = P$  sein. In diesem Fall, also wenn es  $P \in \text{GL}(\dim V, K)$  gibt mit  $A' = P^\top A P$ , heißen  $A$  und  $A'$  auch *kongruent*.

Es gibt einen Zusammenhang zwischen Bilinearformen und Linearformen und Dualräumen.

## 24.3. Lemma.

(1) Seien  $V$  und  $W$  zwei  $K$ -Vektorräume und  $\beta : V \times W \rightarrow K$  eine Bilinearform. Dann ist für jedes  $v \in V$  die Abbildung  $W \rightarrow K$ ,  $w \mapsto \beta(v, w)$ , eine Linearform auf  $W$ , und für jedes  $w \in W$  ist die Abbildung  $V \rightarrow K$ ,  $v \mapsto \beta(v, w)$ , eine Linearform auf  $V$ .

(2) Die durch (1) gegebenen Abbildungen

$$\beta_L : V \longrightarrow W^*, \quad v \longmapsto (w \mapsto \beta(v, w))$$

und

$$\beta_R : W \longrightarrow V^*, \quad w \longmapsto (v \mapsto \beta(v, w))$$

sind linear.

(3) Die sich aus (2) ergebenden Abbildungen

$$\text{Bil}(V, W) \longrightarrow \text{Hom}(V, W^*), \quad \beta \longmapsto \beta_L$$

und

$$\text{Bil}(V, W) \longrightarrow \text{Hom}(W, V^*), \quad \beta \longmapsto \beta_R$$

sind Isomorphismen. Insbesondere sind  $\text{Hom}(V, W^*)$  und  $\text{Hom}(W, V^*)$  isomorph.

*Beweis.*

(1) Das folgt unmittelbar aus der Definition von „Bilinearform“.

(2) Das folgt ebenfalls direkt aus der Definition.

(3) Man rechnet nach, dass die Abbildungen linear sind. Wir zeigen, dass die erste Abbildung bijektiv ist mit Inverser  $f \mapsto ((v, w) \mapsto (f(v))(w))$ : Einerseits wird  $\beta \in \text{Bil}(V, W)$  wie folgt abgebildet:

$$\beta \longmapsto \beta_L \longmapsto ((v, w) \mapsto (\beta_L(v))(w) = \beta(v, w)) = \beta;$$

andererseits haben wir für  $f \in \text{Hom}(V, W^*)$

$$f \longmapsto ((v, w) \mapsto (f(v))(w)) \longmapsto (v \mapsto (w \mapsto (f(v))(w)) = f(v)) = f.$$

Die Bijektivität der zweiten Abbildung zeigt man analog. (Die Inverse ist  $f \mapsto ((v, w) \mapsto (f(w))(v))$ .)  $\square$

Ist  $B$  eine endliche Basis von  $V$  und  $B'$  eine endliche Basis von  $W$ , dann gilt

$$\text{Mat}_{B, B'}(\beta_L) = \text{Mat}_{B, B'}(\beta)^\top \quad \text{und} \quad \text{Mat}_{B', B^*}(\beta_R) = \text{Mat}_{B, B'}(\beta).$$

\*

24.4. **Definition.** Eine Bilinearform  $\beta : V \times W \rightarrow K$  heißt *nicht-ausgeartet*, wenn  $\beta_L : V \rightarrow W^*$  und  $\beta_R : W \rightarrow V^*$  Isomorphismen sind. Anderenfalls heißt  $\beta$  *ausgeartet*.  $\diamond$

**DEF**  
Bilinearform  
nicht-  
ausgeartet

Wenn man eine solche nicht-ausgeartete Bilinearform hat, dann kann man (via  $\beta_L$  und  $\beta_R$ )  $V$  als Dualraum von  $W$  und umgekehrt betrachten: Zu jeder Linearform  $\phi$  auf  $V$  gibt es genau ein Element  $w \in W$  mit  $\phi = \beta_R(w)$  (also sodass  $\phi(v) = \beta(v, w)$  ist für alle  $v \in V$ ), und zu jeder Linearform  $\psi$  auf  $W$  gibt es genau ein Element  $v \in V$  mit  $\psi = \beta_L(v)$  (also sodass  $\psi(w) = \beta(v, w)$  ist für alle  $w \in W$ ).

Man kann zeigen, dass es eine nicht-ausgeartete Bilinearform auf  $V \times W$  nur dann geben kann, wenn  $V$  und  $W$  endlich-dimensional sind.

Es folgt nämlich wegen  $\beta_R^\top \circ \alpha_V = \beta_L$  (Übung), dass  $\alpha_V$  ein Isomorphismus ist. Das ist aber nur für endlich-dimensionale Vektorräume  $V$  der Fall.

Dann müssen  $V$  und  $W$  dieselbe Dimension haben:  $\dim W = \dim W^* = \dim V$ .

**24.5. Beispiel.** Sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum. Dann ist die *Auswertungspaarung*

$$\text{ev} : V \times V^* \longrightarrow K, \quad (v, \phi) \longmapsto \phi(v)$$

nicht-ausgeartet. Für beliebiges  $V$  gilt (Übung)

$$\text{ev}_L = \alpha_V : V \longrightarrow V^{**} \quad \text{und} \quad \text{ev}_R = \text{id}_{V^*} : V^* \longrightarrow V^* .$$

**BSP**  
nicht-ausg.  
Bilinearform



**24.6. Lemma.** Seien  $V$  und  $W$  zwei  $K$ -Vektorräume derselben endlichen Dimension  $n$ , sei  $B$  eine Basis von  $V$ ,  $B'$  eine Basis von  $W$  und  $\beta \in \text{Bil}(V, W)$ . Wir setzen  $A = \text{Mat}_{B, B'}(\beta)$ . Dann sind folgende Aussagen äquivalent:

- (1)  $\beta$  ist nicht-ausgeartet.
- (2)  $\ker(\beta_L) = \{\mathbf{0}\}$ .
- (3)  $\ker(\beta_R) = \{\mathbf{0}\}$ .
- (4)  $\det(A) \neq 0$ .

**LEMMA**  
Kriterium  
für nicht-  
ausgeartet

*Beweis.* Dass aus (1) die Aussagen (2) und (3) folgen, ist klar nach Definition 24.4. Umgekehrt folgt aus (2) zunächst, dass  $\beta_L$  ein Isomorphismus ist (denn  $\dim V = n = \dim W^*$ ) und dann, dass  $\beta_R = \beta_L^\top \circ \alpha_W$  ebenfalls ein Isomorphismus ist. Genauso zeigt man „(3) $\Rightarrow$ (1)“. Schließlich ist (3) äquivalent dazu, dass  $\text{Mat}_{B', B^*}(\beta_R)$  invertierbar ist. Diese Matrix ist aber genau  $A$ , und „ $A$  invertierbar“ ist äquivalent zu „ $\det(A) \neq 0$ “.  $\square$

**24.7. Beispiel.** Ist  $V$  ein endlich-dimensionaler reeller Vektorraum mit einer positiv definiten symmetrischen Bilinearform  $\beta : V \times V \rightarrow \mathbb{R}$ , dann ist  $\beta$  nicht-ausgeartet: Wir zeigen  $\ker(\beta_L) = \{\mathbf{0}\}$ . Sei also  $v \in \ker(\beta_L)$ . Dann ist

$$0 = \mathbf{0}(v) = (\beta_L(v))(v) = \beta(v, v) .$$

Wäre  $v \neq \mathbf{0}$ , dann hätten wir  $\beta(v, v) > 0$ , also muss  $v = \mathbf{0}$  sein.



**BSP**  
pos. def.  $\Rightarrow$   
nicht-ausg.

Aus einer bilinearen Abbildung kann man eine „quadratische“ Abbildung machen: Ist  $\beta : V \times V \rightarrow W$  bilinear, dann hat die Abbildung

$$q : V \longrightarrow W, \quad v \longmapsto \beta(v, v)$$

folgende Eigenschaften:

$$\begin{aligned} q(\lambda v) &= \lambda^2 q(v) && \text{für alle } \lambda \in K, v \in V, \quad \text{und} \\ q(v + v') + q(v - v') &= 2q(v) + 2q(v') && \text{für alle } v, v' \in V \\ &&& \text{ („Parallelogramm-Gleichung“).} \end{aligned}$$

Außerdem ist  $q(v + v') - q(v) - q(v') = \beta(v, v') + \beta(v', v)$  eine (symmetrische) bilineare Abbildung. Wir untersuchen den Zusammenhang etwas genauer.

\*

**24.8. Definition.** Sei  $V$  ein  $K$ -Vektorraum. Eine *quadratische Form* auf  $V$  ist eine Abbildung  $q : V \rightarrow K$ , sodass

**DEF**  
quadratische  
Form

- (1)  $q(\lambda v) = \lambda^2 q(v)$  für alle  $\lambda \in K, v \in V$ , und
- (2)  $(v, w) \mapsto q(v + w) - q(v) - q(w)$  eine Bilinearform ist.

Die Menge aller quadratischen Formen auf  $V$  bildet in der üblichen Weise einen Vektorraum  $\text{Qu}(V)$ .

Zwei quadratische Formen  $q$  und  $q'$  auf  $V$  heißen *äquivalent*, wenn es einen Isomorphismus  $f : V \rightarrow V$  gibt, sodass  $q' = q \circ f$  ist.  $\diamond$

Die Parallelogramm-Gleichung folgt aus den beiden Eigenschaften in der Definition.

Analog zu symmetrischen Bilinearformen definiert man positive Definitheit usw. für quadratische Formen über  $\mathbb{R}$ .

**24.9. Definition.** Sei  $q$  eine quadratische Form auf einem reellen Vektorraum  $V$ .

**DEF**  
pos./neg.  
definit für  
qu. Formen

- (1)  $q$  heißt *positiv (negativ) definit*, wenn  $q(v) > 0$  ( $q(v) < 0$ ) für alle  $0 \neq v \in V$  gilt.
- (2)  $q$  heißt *positiv (negativ) semidefinit*, wenn  $q(v) \geq 0$  ( $q(v) \leq 0$ ) für alle  $v \in V$  gilt.
- (3)  $q$  heißt *indefinit*, wenn  $q$  weder positiv noch negativ semidefinit ist.  $\diamond$

Für das Folgende ist es wichtig, dass wir durch 2 teilen können. Deshalb noch eine Definition.

**24.10. Definition.** Sei  $K$  ein Körper. Ist  $n \cdot 1_K \neq 0_K$  für alle  $n \in \mathbb{Z}_{>0}$ , dann hat  $K$  *Charakteristik* 0. Sonst ist die *Charakteristik* von  $K$  die kleinste positive ganze Zahl  $p$  mit  $p \cdot 1_K = 0_K$ . Wir schreiben  $\text{char}(K)$  für die Charakteristik von  $K$ .  $\diamond$

**DEF**  
Charakteristik

Die Charakteristik ist entweder null oder eine Primzahl: Wäre  $\text{char}(K) = n$  keine Primzahl, dann könnten wir schreiben  $n = km$  mit  $1 \leq k, m < n$ . Aus  $n \cdot 1_K = 0_K$  folgt  $(k \cdot 1_K) \cdot (m \cdot 1_K) = 0_K$ , also  $k \cdot 1_K = 0_K$  oder  $m \cdot 1_K = 0_K$ , was ein Widerspruch dazu ist, dass  $n$  die kleinste solche Zahl ist.

Wenn  $K$  nicht Charakteristik 2 hat, dann ist  $2 \neq 0$  in  $K$  und damit invertierbar. Ein Körper der Charakteristik 2 ist zum Beispiel  $\mathbb{F}_2$ ; dort gilt ja  $1 + 1 = 0$ .

**24.11. Lemma.** Sei  $V$  ein  $K$ -Vektorraum mit  $\text{char}(K) \neq 2$ . Wir schreiben  $\text{Sym}(V)$  für den Vektorraum der symmetrischen Bilinearformen auf  $V$ . Dann ist

**LEMMA**  
symm. bil.  
= quadr.

$$\text{Sym}(V) \longrightarrow \text{Qu}(V), \quad \beta \longmapsto (v \mapsto \beta(v, v))$$

ein Isomorphismus.

Man kann also quadratische Formen mit den zugehörigen symmetrischen Bilinearformen identifizieren; insbesondere sind auch quadratische Formen durch symmetrische Matrizen beschrieben. Wir schreiben  $\text{Mat}_B(q)$  für diese Matrix; es gilt für  $v = x_1 b_1 + \dots + x_n b_n$  (wenn  $B = (b_1, \dots, b_n)$  ist)

$$q(v) = \mathbf{x}^\top \text{Mat}_B(q) \mathbf{x},$$

wobei  $\mathbf{x}$  der Spaltenvektor  $(x_1, \dots, x_n)^\top$  ist.

*Beweis.* Dass die angegebene Abbildung wohldefiniert und linear ist, ist klar. Wir zeigen, dass sie bijektiv ist, indem wir die Umkehrabbildung angeben:

$$q \mapsto ((v, w) \mapsto \frac{1}{2}(q(v+w) - q(v) - q(w)))$$

(hier verwenden wir  $\text{char}(K) \neq 2$ ). Nach Definition 24.8 ist das Bild eine (symmetrische) Bilinearform, also haben wir eine Abbildung  $\text{Qu}(V) \rightarrow \text{Sym}(V, V)$ . Wir prüfen nach, dass das tatsächlich die Inverse ist:

$$\begin{aligned} q \mapsto & ((v, w) \mapsto \frac{1}{2}(q(v+w) - q(v) - q(w))) \\ \mapsto & (v \mapsto \frac{1}{2}(q(2v) - 2q(v)) = q(v)) \\ & = q \end{aligned}$$

und

$$\begin{aligned} \beta \mapsto & (v \mapsto \beta(v, v)) \\ \mapsto & ((v, w) \mapsto \frac{1}{2}(\beta(v+w, v+w) - \beta(v, v) - \beta(w, w))) \\ & = \frac{1}{2}(\beta(v, w) + \beta(w, v)) = \beta(v, w) \\ & = \beta. \end{aligned}$$

□

Daraus folgt unmittelbar:

**24.12. Folgerung.** *Ist  $K$  ein Körper mit  $\text{char}(K) \neq 2$ , dann sind die quadratischen Formen auf  $K^n$  alle gegeben durch*

$$(x_1, x_2, \dots, x_n) \mapsto \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$$

**FOLG**  
qu. Formen  
auf  $K^n$

mit  $a_{ij} \in K$ . Die zugehörige Matrix hat Diagonaleinträge  $a_{ii}$  und Einträge  $a_{ij}/2$  an den Positionen  $(i, j)$  und  $(j, i)$ , wenn  $i < j$  ist.

Zwei quadratische Formen auf  $K^n$  sind äquivalent genau dann, wenn die zugehörigen symmetrischen Matrizen kongruent sind.

Die erste Aussage bleibt auch für Körper der Charakteristik 2 richtig; die Aussagen über die Matrizen haben in diesem Fall keinen Sinn.

Äquivalenz von zwei quadratischen Formen  $q$  und  $q'$  auf  $K^n$  bedeutet dann ganz konkret, dass

$$\begin{aligned} q'(x_1, x_2, \dots, x_n) = & q(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ & \dots, \\ & a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) \end{aligned}$$

gilt mit einer Matrix  $A = (a_{ij}) \in \text{GL}(n, K)$ .

**24.13. Definition.** Eine quadratische Form  $q$  auf  $K^n$  heißt *diagonal* oder eine *Diagonalform*, wenn sie die Form

$$q(x_1, x_2, \dots, x_n) = a_1 x_1^2 + a_2 x_2^2 + \dots + a_n x_n^2$$

hat mit geeigneten  $a_1, a_2, \dots, a_n \in K$ .

◇

Es treten also keine „gemischten Terme“  $x_i x_j$  (mit  $i \neq j$ ) auf, und die zugehörige Matrix ist eine Diagonalmatrix.

Seien  $A$  und  $A'$  die symmetrischen Matrizen zweier äquivalenter quadratischer Formen  $q$  und  $q'$  auf  $K^n$  (mit  $\text{char}(K) \neq 2$ ). Dann gibt es  $P \in \text{GL}(n, K)$  mit

**DEF**  
diagonale  
qu. Form

$A' = P^TAP$ ; insbesondere ist  $\text{rk}(A') = \text{rk}(A)$ . Das zeigt, dass folgende Definition sinnvoll ist.

**24.14. Definition.** Sei  $K$  ein Körper mit  $\text{char}(K) \neq 2$  und sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum. Sei weiter  $q \in \text{Qu}(V)$ . Ist  $B$  eine Basis von  $V$ , dann heißt  $\text{rk}(\text{Mat}_B(q))$  der *Rang* von  $q$ .  $\diamond$

**DEF**  
Rang einer  
qu. Form

Wir wollen jetzt quadratische Formen auf endlich-dimensionalen komplexen und reellen Vektorräumen klassifizieren. Das ist dazu äquivalent, symmetrische Matrizen bis auf Kongruenz zu klassifizieren. Wir beginnen mit einem Resultat, das für (fast) beliebige Körper gilt.

\* **24.15. Satz.** Sei  $K$  ein Körper mit  $\text{char}(K) \neq 2$ . Dann ist jede quadratische Form auf  $K^n$  äquivalent zu einer Diagonalform.

**SATZ**  
Diagonalisierung  
von  
qu. Formen

Dazu äquivalent sind folgende Aussagen:

- Seien  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $q$  eine quadratische Form auf  $V$ . Dann hat  $V$  eine Basis  $B$ , sodass  $\text{Mat}_B(q)$  eine Diagonalmatrix ist.
- Jede symmetrische Matrix  $A \in \text{Mat}(n, K)$  ist kongruent zu einer Diagonalmatrix.

*Beweis.* Der Beweis geht durch Induktion über  $n$ . Im Fall  $n = 1$  (oder  $n = 0$ ) ist nichts zu zeigen. Wir nehmen jetzt an, dass  $n > 1$  ist und die Aussage für  $n - 1$  gilt. Wir schreiben

$$q(x_1, x_2, \dots, x_n) = q'(x_1, x_2, \dots, x_{n-1}) + b_1x_1x_n + \dots + b_{n-1}x_{n-1}x_n + a_nx_n^2$$

mit  $b_1, \dots, b_{n-1}, a_n \in K$  und einer quadratischen Form  $q'$  auf  $K^{n-1}$ . Ist  $a_n \neq 0$ , dann ist („quadratische Ergänzung“)

$$q(x_1, \dots, x_{n-1}, x_n - \frac{1}{2a_n}(b_1x_1 + \dots + b_{n-1}x_{n-1})) = q''(x_1, x_2, \dots, x_{n-1}) + a_nx_n^2$$

mit

$$q''(x_1, \dots, x_{n-1}) = q'(x_1, \dots, x_{n-1}) - \frac{1}{4a_n}(b_1x_1 + \dots + b_{n-1}x_{n-1})^2;$$

das ist eine quadratische Form auf  $K^{n-1}$ . Nach Induktionsannahme ist  $q''$  äquivalent zu einer Diagonalform  $a_1x_1^2 + \dots + a_{n-1}x_{n-1}^2$ ; damit ist  $q$  äquivalent zu  $a_1x_1^2 + \dots + a_{n-1}x_{n-1}^2 + a_nx_n^2$ .

Es bleibt der Fall  $a_n = 0$  zu behandeln. Gilt  $b_1 = \dots = b_{n-1} = 0$ , dann können wir die Induktionsannahme direkt auf  $q'$  anwenden und sind fertig. Sei also jetzt  $b_m \neq 0$  für ein  $m \in \{1, 2, \dots, n-1\}$ . Dann ist  $q(x_1, \dots, x_n) = \alpha x_m^2 + b_m x_m x_n + R$ , wobei jeder Term, der in  $R$  vorkommt, eine Variable  $x_j$  mit  $j \notin \{m, n\}$  enthält. Wir ersetzen  $x_m$  durch  $x_m \pm x_n$  und erhalten

$$q(x_1, \dots, x_m \pm x_n, \dots, x_n) = \alpha x_m^2 + (b_m \pm 2\alpha)x_m x_n + (\alpha \pm b_m)x_n^2 + R'$$

Da  $b_m \neq 0$  ist, muss für wenigstens eine Wahl des Vorzeichens  $\alpha \pm b_m \neq 0$  sein (hier benutzen wir wieder  $\text{char}(K) \neq 2$ , also  $1_K \neq -1_K$ ). Wir sehen, dass  $q$  zu einer Form mit  $a_n \neq 0$  äquivalent ist; diesen Fall haben wir bereits behandelt.  $\square$

Die Koeffizienten der Diagonalform sind keineswegs eindeutig bestimmt. Wir können die Reihenfolge beliebig ändern durch Permutation der Variablen. Durch Skalieren der Koordinaten können wir außerdem die Koeffizienten mit beliebigen Quadraten  $\neq 0$  multiplizieren. Aber es gilt zum Beispiel auch, dass

$$2x_1^2 + 2x_2^2 \quad \text{und} \quad x_1^2 + x_2^2$$

über  $\mathbb{Q}$  äquivalent sind, obwohl 2 kein Quadrat in  $\mathbb{Q}$  ist:

$$(x_1 + x_2)^2 + (x_1 - x_2)^2 = 2x_1^2 + 2x_2^2.$$

**24.16. Beispiel.** Wie sieht eine zu  $q(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3$  äquivalente Diagonalform über  $\mathbb{Q}$  aus? Da kein Term  $x_j^2$  auftritt, müssen wir zunächst einen erzeugen:

$$q'(x_1, x_2, x_3) = q(x_1 + x_3, x_2, x_3) = x_1x_2 + x_1x_3 + 2x_2x_3 + x_3^2.$$

Jetzt können wir die quadratische Ergänzung durchführen:

$$q''(x_1, x_2, x_3) = q'(x_1, x_2, x_3 - \frac{1}{2}x_1 - x_2) = -\frac{1}{4}x_1^2 - x_2^2 + x_3^2.$$

Wir können  $x_1$  noch mit 2 skalieren und erhalten die etwas hübschere Form

$$q'''(x_1, x_2, x_3) = q''(2x_1, x_2, x_3) = -x_1^2 - x_2^2 + x_3^2. \quad \clubsuit$$

Im Körper  $\mathbb{C}$  der komplexen Zahlen hat (nach Satz 4.3) jedes Element eine Quadratwurzel. Da wir die Diagonaleinträge mit beliebigen Quadraten multiplizieren können, erhalten wir den folgenden Klassifikationssatz.

\* **24.17. Satz.** Jede quadratische Form  $q \in \text{Qu}(\mathbb{C}^n)$  ist äquivalent zu einer Form

$$Q_r(x_1, \dots, x_n) = x_1^2 + \dots + x_r^2.$$

Die Zahl  $r \in \{0, 1, \dots, n\}$  ist dabei eindeutig bestimmt.

**SATZ**  
Klassifikation  
qu. Formen  
über  $\mathbb{C}$

*Beweis.* Nach Satz 24.15 ist  $q$  äquivalent zu einer Diagonalform  $q'$ . Wir können annehmen (nach eventueller Permutation der Variablen), dass in  $q'$  genau die Terme  $x_1^2, \dots, x_r^2$  vorkommen. Durch Skalieren können wir erreichen, dass die Koeffizienten = 1 sind; damit haben wir die gewünschte Form. Als Rang von  $q$  ist  $r$  eindeutig bestimmt.  $\square$

In  $\mathbb{R}$  gilt nur noch, dass jede positive Zahl (und die Null) ein Quadrat ist. Das führt zum folgenden *Sylvesterschen Trägheitssatz* oder *Signatursatz*:

\* **24.18. Satz.** Jede quadratische Form  $q \in \text{Qu}(\mathbb{R}^n)$  ist äquivalent zu einer Form

$$Q_{r,s}(x_1, \dots, x_n) = x_1^2 + \dots + x_r^2 - x_{r+1}^2 - \dots - x_{r+s}^2.$$

Die Zahlen  $r, s \geq 0$  mit  $r + s \leq n$  sind eindeutig bestimmt.

**SATZ**  
Klassifikation  
qu. Formen  
über  $\mathbb{R}$

*Beweis.* Nach Satz 24.15 ist  $q$  äquivalent zu einer Diagonalform  $q'$ . Nach Permutation der Variablen können wir annehmen, dass  $q'(x_1, \dots, x_n) = \sum_{j=1}^{r+s} a_j x_j^2$  ist mit  $a_1, \dots, a_r > 0$  und  $a_{r+1}, \dots, a_{r+s} < 0$ . Durch Skalieren können wir die positiven Koeffizienten durch 1 und die negativen Koeffizienten durch  $-1$  ersetzen; damit haben wir die gewünschte Form.

Wie eben ist der Rang  $r+s$  eindeutig durch  $q$  bestimmt. Die Zahl  $r$  ist die maximale Dimension eines Untervektorraums, auf dem  $q$  positiv definit ist und ist damit ebenfalls eindeutig bestimmt: Diese Dimension ist für äquivalente quadratische

Formen offensichtlich gleich, also müssen wir diese Aussage nur für  $Q_{r,s}$  zeigen.  $Q_{r,s}$  ist auf dem  $r$ -dimensionalen Untervektorraum  $\langle \mathbf{e}_1, \dots, \mathbf{e}_r \rangle$  positiv definit. Ist  $U \subset \mathbb{R}^n$  mit  $\dim U > r$ , dann gilt mit  $U' = \langle \mathbf{e}_{r+1}, \dots, \mathbf{e}_n \rangle$ , dass

$$\dim U \cap U' = \dim U + \dim U' - \dim(U + U') \geq \dim U + (n - r) - n = \dim U - r > 0$$

ist, also gibt es einen Vektor  $\mathbf{0} \neq v \in U \cap U'$ . Es gilt dann aber  $Q_{r,s}(v) \leq 0$ , also ist  $Q_{r,s}$  auf  $U$  nicht positiv definit. Damit ist  $r$  die maximale Dimension eines Untervektorraums, auf dem  $Q_{r,s}$  positiv definit ist.  $\square$

\* 24.19. **Definition.** In der Situation von Satz 24.18 heißt  $r - s$  die *Signatur* von  $q$ . **DEF**  
 $\diamond$  Signatur

24.20. **Beispiel.** Die Signatur der quadratischen Form  $x_1x_2 + x_1x_3 + x_2x_3$  aus Beispiel 24.16 ist  $-1$ , denn wir haben  $r = 1$  und  $s = 2$ . **BSP**  
 $\clubsuit$  Signatur

24.21. **Folgerung.** Sei  $q \in \text{Qu}(\mathbb{R}^n)$  mit  $r$  und  $s$  wie in Satz 24.18. Dann gilt: **FOLG**  
Definitheit  
über  $r, s$

- (1)  $q$  positiv definit  $\iff r = n$ .
- (2)  $q$  negativ definit  $\iff s = n$ .
- (3)  $q$  positiv semidefinit  $\iff s = 0$ .
- (4)  $q$  negativ semidefinit  $\iff r = 0$ .
- (5)  $q$  indefinit  $\iff r, s > 0$ .

*Beweis.* Es ist klar, dass diese Äquivalenzen für die zu  $q$  äquivalente Diagonalform  $Q_{r,s}$  gelten. Weil die Definitheitseigenschaften unter Äquivalenz invariant sind, gelten sie dann auch für  $q$ .  $\square$

25. EUKLIDISCHE VEKTORRÄUME

Wir haben am Ende des letzten Abschnitts schon damit begonnen, von der „allgemeinen“ linearen Algebra über beliebigen Grundkörpern etwas wegzugehen und Resultate für die speziellen Körper  $\mathbb{R}$  und  $\mathbb{C}$  zu beweisen. Das setzen wir in diesem Abschnitt fort. Der Hintergrund dafür ist, dass wir *Geometrie* betreiben wollen: Wir wollen in der Lage sein, Abstände und Winkel zu messen. Dies wird in einem reellen Vektorraum durch eine positiv definite symmetrische Bilinearform ermöglicht.

\* **25.1. Definition.** Eine positiv definite symmetrische Bilinearform auf einem reellen Vektorraum heißt *euklidisches Skalarprodukt*. Ein reeller Vektorraum  $V$  zusammen mit einem euklidischen Skalarprodukt auf  $V$  ist ein *euklidischer Vektorraum*. Das Skalarprodukt in einem euklidischen Vektorraum wird häufig  $(v, w) \mapsto \langle v, w \rangle$  (oder auch  $v \cdot w$ ) geschrieben. ◇

**DEF**  
euklidisches  
Skalarprod.  
eukl. VR

Um Verwechslungen zu vermeiden, notieren wir den von einer Menge  $A$  erzeugten Untervektorraum als  $\langle A \rangle_{\mathbb{R}}$ .

25.2. Beispiele.

- Das *Standard-Skalarprodukt*  $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$  auf  $\mathbb{R}^n$  (mit Spaltenvektoren  $\mathbf{x}, \mathbf{y}$ ) ist ein euklidisches Skalarprodukt.  $\mathbb{R}^n$  mit diesem Skalarprodukt ist das Standardbeispiel für einen (endlich-dimensionalen) euklidischen Vektorraum.
- Seien  $a, b \in \mathbb{R}$  mit  $a < b$  und sei  $V = \mathcal{C}([a, b])$  der Vektorraum der stetigen reellen Funktionen auf  $[a, b]$ . Dann definiert

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

ein euklidisches Skalarprodukt auf  $V$ . ♣

**BSP**  
eukl. VR

\* **25.3. Definition.** Sei  $V$  ein euklidischer Vektorraum. Für einen Vektor  $v \in V$  heißt  $\|v\| = \sqrt{\langle v, v \rangle}$  die *Länge* von  $v$ . Gilt  $\|v\| = 1$ , dann heißt  $v$  ein *Einheitsvektor*. ◇

**DEF**  
Länge  
Einheitsvektor

Es gilt dann  $\|v\| \geq 0$  und  $\|v\| = 0 \iff v = \mathbf{0}$ .

Im Standardraum  $\mathbb{R}^n$  ist  $\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2}$  die übliche euklidische Länge eines Vektors. Die Standardbasis besteht aus Einheitsvektoren.

Wir beweisen einige Eigenschaften der Länge.

\* **25.4. Satz.** Sei  $V$  ein euklidischer Vektorraum.

- (1) Für  $v \in V$  und  $\lambda \in \mathbb{R}$  gilt  $\|\lambda v\| = |\lambda| \|v\|$ .
- (2) (Cauchy-Schwarzsche Ungleichung) Für  $v, w \in V$  gilt  $|\langle v, w \rangle| \leq \|v\| \|w\|$  mit Gleichheit genau dann, wenn  $v$  und  $w$  linear abhängig sind.
- (3) (Dreiecksungleichung) Für  $v, w \in V$  gilt  $\|v+w\| \leq \|v\| + \|w\|$  mit Gleichheit genau dann, wenn  $v = \lambda w$  oder  $w = \lambda v$  ist mit  $\lambda \geq 0$ .

**SATZ**  
Cauchy-  
Schwarz  
Dreiecksungl.

*Beweis.*

- (1)  $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \|v\|$ .  
 (2) Die Aussage ist klar für  $w = \mathbf{0}$ . Wir können also  $w \neq \mathbf{0}$  annehmen. Sei

$$v' = v - \frac{\langle v, w \rangle}{\|w\|^2} w;$$

dann ist

$$\langle v', w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\|w\|^2} \langle w, w \rangle = 0$$

und damit

$$0 \leq \langle v', v' \rangle = \langle v', v \rangle = \langle v, v \rangle - \frac{\langle v, w \rangle}{\|w\|^2} \langle w, v \rangle = \|v\|^2 - \frac{\langle v, w \rangle^2}{\|w\|^2},$$

was zur behaupteten Ungleichung äquivalent ist. Gleichheit gilt genau dann, wenn  $v' = \mathbf{0}$  ist, daraus folgt, dass  $v$  ein skalares Vielfaches von  $w$  ist. Ist umgekehrt  $v = \lambda w$ , dann ist  $v' = \mathbf{0}$  und es gilt Gleichheit in der Ungleichung.

- (3) Es gilt unter Verwendung der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

Die Ungleichung folgt. Gleichheit ist äquivalent zu  $\langle v, w \rangle = \|v\|\|w\|$ ; dafür müssen  $v$  und  $w$  linear abhängig sein und damit das Vorzeichen stimmt (lineare Abhängigkeit von  $v$  und  $w$  ist äquivalent zu  $\langle v, w \rangle = \pm \|v\|\|w\|$ ), muss der Skalarfaktor  $\geq 0$  sein.  $\square$

Die Eigenschaften (1) und (3) (zusammen mit  $\|v\| = 0 \implies v = \mathbf{0}$ ) besagen, dass  $\|\cdot\|$  eine *Norm* auf  $V$  ist. Daraus folgt insbesondere, dass

$$(v, w) \mapsto d(v, w) = \|v - w\|$$

eine *Metrik* auf  $V$  ist. Damit wird  $V$  in natürlicher Weise zu einem metrischen Raum (diese Begriffe wurden in der Analysis erklärt und studiert). Wir nennen  $d(v, w)$  den *Abstand* zwischen  $v$  und  $w$ .

\* 25.5. **Definition.** Sei  $V$  ein euklidischer Vektorraum.

- (1) Für zwei Vektoren  $v, w \in V$  mit  $v, w \neq \mathbf{0}$  ist der *Winkel* zwischen  $v$  und  $w$  die Zahl  $\alpha = \angle(v, w) \in [0, \pi]$  mit  $\|v\|\|w\| \cos \alpha = \langle v, w \rangle$ .  
 (2) Zwei Vektoren  $v, w \in V$  heißen *orthogonal* (oder *zueinander senkrecht*), wenn  $\langle v, w \rangle = 0$  ist. Wir schreiben dafür  $v \perp w$ .  
 (3) Sei  $U \subset V$  ein Untervektorraum. Dann heißt

$$U^\perp = \{v \in V \mid \forall w \in U : v \perp w\}$$

das *orthogonale Komplement* von  $U$  in  $V$ .

- (4) Eine Teilmenge  $A \subset V$  heißt *orthogonal*, wenn ihre Elemente paarweise orthogonal sind ( $\forall v, w \in A : v \neq w \implies v \perp w$ ).  $A$  heißt *orthonormal*, wenn zusätzlich alle Elemente von  $A$  Einheitsvektoren sind.  
 (5) Eine Basis  $B$  von  $V$  heißt eine *Orthonormalbasis* oder kurz *ONB* von  $V$ , wenn sie aus paarweise orthogonalen Einheitsvektoren besteht (wenn also die Menge der Vektoren in  $B$  orthonormal ist).  $\diamond$

**DEF**  
 Winkel  
 orthogonal  
 orthonormal  
 ONB

Die Cauchy-Schwarzsche Ungleichung stellt sicher, dass die Definition des Winkels sinnvoll ist, denn es gilt ja

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \|w\|} \leq 1.$$

Zwei Vektoren sind orthogonal genau dann, wenn wenigstens einer der Nullvektor ist oder der Winkel zwischen ihnen  $\pi/2$  (entsprechend  $90^\circ$ ) ist.

**25.6. Beispiel.** Klassische Sätze über Dreiecke lassen sich elegant durch Rechnen in euklidischen Vektorräumen beweisen: Seien  $V$  ein euklidischer Vektorraum und  $v, w \in V$ ; wir betrachten das Dreieck mit Eckpunkten  $\mathbf{0}$ ,  $v$  und  $w$ ; es hat Seitenlängen  $a = \|v\|$ ,  $b = \|w\|$ ,  $c = \|v - w\|$ ; der Winkel bei  $\mathbf{0}$  sei  $\gamma$ . Dann gilt  $\langle v, w \rangle = \|v\| \|w\| \cos \gamma$ , also

$$c^2 = \|v - w\|^2 = \|v\|^2 - 2\langle v, w \rangle + \|w\|^2 = a^2 - 2ab \cos \gamma + b^2;$$

das ist der *Cosinussatz*. Gilt  $v \perp w$  (dann ist  $\gamma$  ein rechter Winkel), dann vereinfacht sich das zum *Satz des Pythagoras*

$$c^2 = a^2 + b^2.$$

**BSP**  
Cosinussatz  
Pythagoras



Die Standardbasis ist eine Orthonormalbasis des Standardraums  $\mathbb{R}^n$ .

**25.7. Lemma.** Sei  $V$  ein euklidischer Vektorraum und seien  $v_1, v_2, \dots, v_n \in V$  paarweise orthogonal und von  $\mathbf{0}$  verschieden. Dann sind  $v_1, v_2, \dots, v_n$  linear unabhängig.

**LEMMA**  
orthogonal  
 $\Rightarrow$  lin.unabh.

*Beweis.* Seien  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$  mit  $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \mathbf{0}$ . Es folgt für alle  $j \in \{1, 2, \dots, n\}$ :

$$0 = \langle \mathbf{0}, v_j \rangle = \left\langle \sum_{i=1}^n \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^n \lambda_i \langle v_i, v_j \rangle = \lambda_j \|v_j\|^2,$$

und weil  $v_j \neq \mathbf{0}$  ist, muss  $\lambda_j = 0$  sein. □

Gibt es immer eine Orthonormalbasis? Der folgende wichtige Satz zeigt, dass man aus jeder endlichen Basis eine Orthonormalbasis konstruieren kann.

**\* 25.8. Satz.** Sei  $V$  ein euklidischer Vektorraum mit Basis  $B = (b_1, b_2, \dots, b_n)$ . Dann bilden die wie folgt sukzessive definierten Vektoren  $e_j$  eine ONB von  $V$ ; dabei gilt  $\langle e_1, e_2, \dots, e_j \rangle_{\mathbb{R}} = \langle b_1, b_2, \dots, b_j \rangle_{\mathbb{R}}$ .

**SATZ**  
Gram-Schmidt-Orthonormalisierung

$$\begin{aligned} e_1 &= \frac{1}{\|v_1\|} v_1 \quad \text{mit} \quad v_1 = b_1 \\ e_2 &= \frac{1}{\|v_2\|} v_2 \quad \text{mit} \quad v_2 = b_2 - \langle e_1, b_2 \rangle e_1 \\ e_3 &= \frac{1}{\|v_3\|} v_3 \quad \text{mit} \quad v_3 = b_3 - \langle e_1, b_3 \rangle e_1 - \langle e_2, b_3 \rangle e_2 \\ &\vdots \\ e_n &= \frac{1}{\|v_n\|} v_n \quad \text{mit} \quad v_n = b_n - \sum_{j=1}^{n-1} \langle e_j, b_n \rangle e_j \end{aligned}$$

*Beweis.* Aus der Konstruktion ist klar, dass  $\langle e_1, e_2, \dots, e_j \rangle_{\mathbb{R}} = \langle b_1, b_2, \dots, b_j \rangle_{\mathbb{R}}$  gilt für alle  $j \in \{1, 2, \dots, n\}$ .

Wir zeigen durch Induktion, dass  $\{e_1, \dots, e_n\}$  orthonormal ist. Nach Lemma 25.7 sind die  $e_j$  dann auch linear unabhängig, müssen also eine Basis des  $n$ -dimensionalen Vektorraums  $V$  bilden.

Sei  $1 \leq j \leq n$ . Wir nehmen an, dass  $\{e_1, \dots, e_{j-1}\}$  orthonormal ist. Dann gilt  $v_j \perp e_i$  für alle  $i < j$ , denn

$$\langle e_i, v_j \rangle = \langle e_i, b_j \rangle - \sum_{k=1}^{j-1} \langle e_k, b_j \rangle \langle e_i, e_k \rangle = \langle e_i, b_j \rangle - \langle e_i, b_j \rangle = 0.$$

Außerdem ist  $v_j \neq \mathbf{0}$ , denn  $b_j \notin \langle e_1, \dots, e_{j-1} \rangle_{\mathbb{R}} = \langle b_1, \dots, b_{j-1} \rangle_{\mathbb{R}}$ . Damit ist  $e_j$  definiert und ein Einheitsvektor und (als skalares Vielfaches von  $v_j$ ) ebenfalls orthogonal zu  $e_1, \dots, e_{j-1}$ . □

**25.9. Beispiel.** Wir erzeugen eine ONB aus der Basis  $b_1 = (1, 1, 1)$ ,  $b_2 = (1, -1, 1)$ ,  $b_3 = (1, 0, 0)$  von  $\mathbb{R}^3$ . Wir schreiben  $N(\mathbf{x})$  für  $\|\mathbf{x}\|^{-1}\mathbf{x}$  (für Vektoren  $\mathbf{x} \neq \mathbf{0}$ ). Wir erhalten **BSP ONB**

$$\begin{aligned} e_1 &= N(b_1) = \frac{1}{\sqrt{3}}(1, 1, 1) \\ e_2 &= N(b_2 - \langle e_1, b_2 \rangle e_1) = N\left(\left(\frac{2}{3}, -\frac{4}{3}, \frac{2}{3}\right)\right) = \frac{1}{\sqrt{6}}(1, -2, 1) \\ e_3 &= N(b_3 - \langle e_1, b_3 \rangle e_1 - \langle e_2, b_3 \rangle e_2) = N\left(\left(\frac{1}{2}, 0, -\frac{1}{2}\right)\right) = \frac{1}{\sqrt{2}}(1, 0, -1) \quad \clubsuit \end{aligned}$$

Wir rechtfertigen die Bezeichnung „orthogonales Komplement“ für  $U^\perp$ :

**25.10. Lemma.** Sei  $V$  ein euklidischer Vektorraum und  $U \subset V$  ein endlich-dimensionaler Untervektorraum. Dann ist  $U^\perp$  ein Komplement von  $U$  in  $V$  (also  $U + U^\perp = V$  und  $U \cap U^\perp = \{\mathbf{0}\}$ ). **LEMMA**  
 $U^\perp$  ist  
Komplement  
von  $U$

*Beweis.* Sei  $v \in U \cap U^\perp$ . Dann folgt aus der Definition von  $U^\perp$ , dass  $\langle v, v \rangle = 0$  und damit  $v = \mathbf{0}$  ist. Es folgt  $U \cap U^\perp = \{\mathbf{0}\}$ .

Sei jetzt  $v \in V$  beliebig und  $(e_1, \dots, e_n)$  eine ONB von  $U$ . Wir setzen

$$v_1 = \langle e_1, v \rangle e_1 + \langle e_2, v \rangle e_2 + \dots + \langle e_n, v \rangle e_n \in U \quad \text{und} \quad v_2 = v - v_1.$$

Dann gilt jedenfalls  $v = v_1 + v_2$ . Es bleibt zu zeigen, dass  $v_2 \in U^\perp$  ist. Sei dazu  $u = \lambda_1 e_1 + \dots + \lambda_n e_n \in U$ . Dann gilt

$$\langle u, v_2 \rangle = \sum_{j=1}^n \lambda_j \langle e_j, v - v_1 \rangle$$

und

$$\langle e_j, v - v_1 \rangle = \langle e_j, v \rangle - \sum_{i=1}^n \langle e_i, v \rangle \langle e_j, e_i \rangle = \langle e_j, v \rangle - \langle e_j, v \rangle = 0.$$

Also ist  $v_2 \perp u$  für alle  $u \in U$ , damit  $v_2 \in U^\perp$  und  $v \in U + U^\perp$ . □

Die Aussagen lassen sich noch etwas verfeinern.

\* 25.11. **Satz.** Sei  $V$  ein euklidischer Vektorraum mit ONB  $(e_1, e_2, \dots, e_n)$ .

**SATZ**  
Parsevalsche  
Gleichung

- (1) Für alle  $v \in V$  gilt  $v = \langle e_1, v \rangle e_1 + \langle e_2, v \rangle e_2 + \dots + \langle e_n, v \rangle e_n$ .
- (2) Für alle  $v \in V$  gilt  $\|v\|^2 = \langle e_1, v \rangle^2 + \langle e_2, v \rangle^2 + \dots + \langle e_n, v \rangle^2$ .
- (3) Für alle  $v, w \in V$  gilt  $\langle v, w \rangle = \langle e_1, v \rangle \langle e_1, w \rangle + \dots + \langle e_n, v \rangle \langle e_n, w \rangle$ .

*Beweis.* Sei die rechte Seite in der ersten Gleichung  $u$ , dann ist wie im Beweis von Lemma 25.10 (mit  $V = U$ )  $v - u \in V^\perp = \{\mathbf{0}\}$ , also  $v = u$ .

Die beiden weiteren Aussagen folgen aus (1), da  $\langle e_i, e_j \rangle = \delta_{ij}$ . □

Die letzte Aussage im obigen Satz besagt also, dass die Matrix der Bilinearform  $\langle \cdot, \cdot \rangle$  bezüglich einer ONB die Einheitsmatrix ist:

$$\langle x_1 e_1 + x_2 e_2 + \dots + x_n e_n, y_1 e_1 + y_2 e_2 + \dots + y_n e_n \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

\* 25.12. **Satz.** Sei  $V$  ein euklidischer Vektorraum und  $\{e_1, e_2, \dots, e_n\} \subset V$  eine orthonormale Menge. Dann gilt für alle  $v \in V$

**SATZ**  
Besselsche  
Ungleichung

$$\|v\|^2 \geq \langle e_1, v \rangle^2 + \langle e_2, v \rangle^2 + \dots + \langle e_n, v \rangle^2$$

mit Gleichheit genau für  $v \in \langle e_1, e_2, \dots, e_n \rangle_{\mathbb{R}}$ .

*Beweis.* Sei  $U = \langle e_1, e_2, \dots, e_n \rangle_{\mathbb{R}}$ . Wie im Beweis von Lemma 25.10 können wir  $v \in V$  schreiben als  $v = u + v'$  mit  $u = \sum_{j=1}^n \langle e_j, v \rangle e_j \in U$  und  $v' \in U^\perp$ . Dann gilt  $\|v\|^2 = \|u\|^2 + \|v'\|^2 = \sum_{j=1}^n \langle e_j, v \rangle^2 + \|v'\|^2$ , wobei wir Satz 25.11 verwendet haben. Daraus folgt die Ungleichung; Gleichheit ist äquivalent mit  $v' = \mathbf{0}$ , also mit  $v \in U$ . □

Isomorphismen zwischen euklidischen Vektorräumen, die zusätzlich das euklidische Skalarprodukt erhalten, haben einen besonderen Namen.

\* 25.13. **Definition.** Seien  $V$  und  $W$  zwei euklidische Vektorräume. Eine Abbildung  $f : V \rightarrow W$  heißt (*lineare*) *Isometrie*, wenn  $f$  ein Isomorphismus ist und zusätzlich für alle  $v, v' \in V$  gilt, dass  $\langle f(v), f(v') \rangle = \langle v, v' \rangle$  ist. (Hier steht links das euklidische Skalarprodukt von  $W$ , rechts das von  $V$ .) Gibt es so eine Isometrie, dann heißen  $V$  und  $W$  *isometrisch*. ◇

**DEF**  
Isometrie

Da sich  $\langle v, w \rangle$  durch die Längen von  $v$ ,  $w$  und  $v + w$  ausdrücken lässt:

$$2\langle v, w \rangle = \|v + w\|^2 - \|v\|^2 - \|w\|^2,$$

genügt es, statt der zweiten Bedingung nur zu fordern, dass  $\|f(v)\| = \|v\|$  ist für alle  $v \in V$ .

Man kann das so interpretieren, dass eine lineare Abbildung, die Längen erhält, auch Winkel erhalten muss. Das liegt daran, dass ein Dreieck durch die Längen seiner drei Seiten bis auf Kongruenz eindeutig bestimmt ist: Die Winkel sind durch die Längen festgelegt.

25.14. **Beispiel.** Ist  $V$  ein euklidischer Vektorraum mit ONB  $(e_1, e_2, \dots, e_n)$ , dann ist die Abbildung

**BSP**  
Isometrie

$$\mathbb{R}^n \longrightarrow V, \quad (x_1, x_2, \dots, x_n) \longmapsto x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

eine Isometrie. Das ist gerade der Inhalt von Satz 25.11.

So wie jeder  $n$ -dimensionale  $K$ -Vektorraum zum Standard-Vektorraum  $K^n$  isomorph ist, ist also jeder  $n$ -dimensionale euklidische Vektorraum zum euklidischen Standard-Vektorraum  $\mathbb{R}^n$  isometrisch.

Allgemein gilt: Ein Isomorphismus  $V \rightarrow W$  zwischen endlich-dimensionalen euklidischen Vektorräumen ist genau dann eine Isometrie, wenn er eine Orthonormalbasis auf eine Orthonormalbasis abbildet. ♣

## 26. ORTHOGONALE DIAGONALISIERUNG

Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum. Wir hatten schon in Beispiel 24.7 gesehen, dass eine positiv definite symmetrische Bilinearform auf einem endlich-dimensionalen reellen Vektorraum nicht-ausgeartet ist; dies lässt sich also auf das euklidische Skalarprodukt von  $V$  anwenden:

**26.1. Lemma.** *Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum und sei  $\phi \in V^*$  eine Linearform auf  $V$ . Dann gibt es ein eindeutig bestimmtes Element  $w \in V$  mit  $\phi(v) = \langle v, w \rangle = \langle w, v \rangle$  für alle  $v \in V$ .*

**LEMMA**  
Linearformen  
via  $\langle \cdot, \cdot \rangle$

*Beweis.* Wir schreiben  $\beta(v, w) = \langle v, w \rangle$  für das Skalarprodukt. Da  $\beta$  nicht-ausgeartet ist, ist  $\beta_R : V \rightarrow V^*$ ,  $w \mapsto (v \mapsto \langle v, w \rangle)$ , ein Isomorphismus. Dann ist klar, dass  $w = \beta_R^{-1}(\phi)$  als einziges Element von  $V$  die gewünschte Eigenschaft hat.  $\square$

Für einen endlich-dimensionalen euklidischen Vektorraum gibt es also einen *kanonischen* Isomorphismus  $V \rightarrow V^*$ . Ist  $(e_1, e_2, \dots, e_n)$  eine Orthonormalbasis von  $V$  und  $(e_1^*, e_2^*, \dots, e_n^*)$  die dazu duale Basis von  $V^*$ , dann identifiziert dieser Isomorphismus  $e_j$  mit  $e_j^*$ , denn

$$\langle e_i, e_j \rangle = \delta_{ij} = e_j^*(e_i).$$

**26.2. Beispiel.** Sei  $V = \mathbb{R}^3$  (Elemente als Spaltenvektoren) mit dem Standard-Skalarprodukt. Seien  $v_1, v_2 \in V$ . Dann ist  $v \mapsto \det(v_1, v_2, v)$  eine Linearform auf  $V$ , also gibt es nach Lemma 26.1 einen eindeutig bestimmten Vektor  $v_1 \times v_2 \in V$  mit

**BSP**  
Vektor-  
produkt

$$\langle v_1 \times v_2, v \rangle = \det(v_1, v_2, v) \quad \text{für alle } v \in V.$$

Dieser Vektor  $v_1 \times v_2$  heißt das *Vektorprodukt* oder *Kreuzprodukt* von  $v_1 \times v_2$ . Es hat folgende Eigenschaften (Beweis: Übung):

- (1) Die Abbildung  $V \times V \rightarrow V$ ,  $(v_1, v_2) \mapsto v_1 \times v_2$ , ist bilinear.
- (2) Sind  $v_1, v_2 \in V$  linear abhängig, dann ist  $v_1 \times v_2 = \mathbf{0}$ .
- (3) Für alle  $v_1, v_2 \in V$  gilt  $(v_1 \times v_2) \perp v_1$  und  $(v_1 \times v_2) \perp v_2$ .
- (4) Für alle  $v_1, v_2 \in V \setminus \{\mathbf{0}\}$  gilt  $\|v_1 \times v_2\| = \|v_1\| \|v_2\| \sin \angle(v_1, v_2)$ .

Das lässt sich so interpretieren, dass  $v_1 \times v_2$  der Nullvektor ist, wenn  $v_1$  und  $v_2$  linear abhängig sind; anderenfalls ist es ein Vektor, der auf der von  $v_1$  und  $v_2$  aufgespannten Ebene senkrecht steht und dessen Länge der Fläche des von  $v_1$  und  $v_2$  aufgespannten Parallelogramms entspricht. Dabei ist die Richtung so, dass  $v_1, v_2, v_1 \times v_2$  eine positiv orientierte Basis bilden („Rechte-Hand-Regel“, siehe Definition 17.1), denn

$$\det(v_1, v_2, v_1 \times v_2) = \langle v_1 \times v_2, v_1 \times v_2 \rangle = \|v_1 \times v_2\|^2 > 0. \quad \clubsuit$$

**26.3. Folgerung.** *Seien  $V$  und  $W$  euklidische Vektorräume mit  $\dim V < \infty$  und sei  $f : V \rightarrow W$  linear. Dann gibt es eine eindeutig bestimmte lineare Abbildung  $f^* : W \rightarrow V$  mit*

**FOLG**  
adjungierte  
Abbildung

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle \quad \text{für alle } v \in V \text{ und } w \in W.$$

*Beweis.* Sei zunächst  $w \in W$  fest gewählt. Dann ist  $v \mapsto \langle f(v), w \rangle$  eine Linearform auf  $V$ , also gibt es nach Lemma 26.1 ein eindeutig bestimmtes  $f^*(w) \in V$  mit  $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$  für alle  $v \in V$ . Das liefert uns eine Abbildung  $f^* : W \rightarrow V$ . Da  $f^*(w) + f^*(w')$  die definierende Gleichung für  $f^*(w + w')$  erfüllt, muss wegen der Eindeutigkeit  $f^*(w + w') = f^*(w) + f^*(w')$  gelten. Ebenso sieht man, dass  $f^*(\lambda w) = \lambda f^*(w)$  gilt; damit ist  $f^*$  linear.  $\square$

Die führt auf folgende Begriffsbildung.

\*

**26.4. Definition.** Seien  $V$  und  $W$  euklidische Vektorräume und sei  $f : V \rightarrow W$  linear.

**DEF**  
adjungierte  
Abbildung

- (1) Gibt es eine lineare Abbildung  $f^* : W \rightarrow V$ , sodass für alle  $v \in V$  und  $w \in W$  gilt  $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$ , dann heißt  $f^*$  die zu  $f$  adjungierte Abbildung.
- (2) Hat  $f \in \text{End}(V)$  eine adjungierte Abbildung  $f^*$  und gilt  $f = f^*$ , dann heißt  $f$  selbst-adjungiert. Das bedeutet also  $\langle f(v), v' \rangle = \langle v, f(v') \rangle$  für alle  $v, v' \in V$ .  $\diamond$

Folgerung 26.3 besagt, dass es für  $V$  endlich-dimensional stets adjungierte Abbildungen gibt.

**26.5. Lemma.** Seien  $V_1, V_2$  und  $V_3$  endlich-dimensionale euklidische Vektorräume und seien  $f, g : V_1 \rightarrow V_2$  und  $h : V_2 \rightarrow V_3$  linear und  $\lambda \in \mathbb{R}$ . Dann gilt:

**LEMMA**  
Eigenschaften  
von  $f^*$

- (1)  $(f + g)^* = f^* + g^*$  und  $(\lambda f)^* = \lambda f^*$ .  
(Das bedeutet, dass  $\text{Hom}(V, W) \rightarrow \text{Hom}(W, V)$ ,  $f \mapsto f^*$ , linear ist.)
- (2)  $(h \circ f)^* = f^* \circ h^*$ .
- (3)  $(f^*)^* = f$ .
- (4)  $f$  ist eine Isometrie  $\iff f$  ist ein Isomorphismus mit  $f^{-1} = f^*$ .

*Beweis.* (1)–(3): Übung.

(4) Es ist

$$\langle f(v), f(w) \rangle = \langle v, f^*(f(w)) \rangle,$$

und das ist dasselbe wie  $\langle v, w \rangle$  für alle  $v, w \in V_1$  genau dann, wenn  $f^* \circ f = \text{id}_{V_1}$ . Ist  $f$  eine Isometrie, dann ist also  $f$  ein Isomorphismus und es gilt  $f^* \circ f = \text{id}_{V_1}$ , also ist  $f^{-1} = f^*$ . Ist umgekehrt  $f$  ein Isomorphismus mit  $f^{-1} = f^*$ , dann folgt  $f^* \circ f = \text{id}_{V_1}$  und  $f$  ist eine Isometrie.  $\square$

Wir wollen uns jetzt mit der Frage beschäftigen, wann es für einen Endomorphismus  $f$  eines euklidischen Vektorraums  $V$  eine Orthonormalbasis von  $V$  gibt, die aus Eigenvektoren von  $f$  besteht. Man sagt dann,  $f$  sei *orthogonal diagonalisierbar*. Das Resultat wird sein, dass das genau für selbst-adjungierte Endomorphismen der Fall ist. Wir zeigen zuerst die einfachere Richtung. Wir wiederholen Definition 16.8.

\* **26.6. Definition.** Eine Matrix  $A \in \text{GL}(n, \mathbb{R})$  heißt *orthogonal*, wenn sie die Gleichung  $A^{-1} = A^\top$  erfüllt. Wir schreiben  $O(n)$  für die Gruppe (!) der orthogonalen  $n \times n$ -Matrizen. ◇

**DEF**  
orthogonale  
Matrix

Schreibt man die Bedingung  $A^\top A = AA^\top = I_n$  aus, dann sieht man, dass  $A$  genau dann orthogonal ist, wenn die Spalten (Zeilen) von  $A$  eine Orthonormalbasis von  $\mathbb{R}^n$  bilden.

**26.7. Lemma.** Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum mit Orthonormalbasis  $B = (e_1, e_2, \dots, e_n)$  und sei  $f : V \rightarrow V$  linear. Dann gilt  $\text{Mat}_B(f^*) = \text{Mat}_B(f)^\top$ ; insbesondere gilt

**LEMMA**  
Matrix für  
selbst-adj.  
Endom.

$$f \text{ selbst-adjungiert} \iff \text{Mat}_B(f) \text{ symmetrisch}$$

und

$$f \text{ ist eine Isometrie} \iff \text{Mat}_B(f) \text{ orthogonal.}$$

Die erste Aussage gilt auch allgemeiner für Matrizen bezüglich ONBen von linearen Abbildungen zwischen zwei endlich-dimensionalen euklidischen Vektorräumen.

*Beweis.* Sei  $A = (a_{ij}) = \text{Mat}_B(f)$ . Es gilt nach Satz 25.11

$$f(e_j) = \langle e_1, f(e_j) \rangle e_1 + \langle e_2, f(e_j) \rangle e_2 + \dots + \langle e_n, f(e_j) \rangle e_n,$$

also ist

$$a_{ij} = \langle e_i, f(e_j) \rangle = \langle f(e_j), e_i \rangle = \langle e_j, f^*(e_i) \rangle$$

und eine analoge Überlegung zeigt, dass das  $a'_{ji}$  ist für  $A' = \text{Mat}_B(f^*)$ . Also gilt  $A' = A^\top$  wie behauptet. Die Aussage über selbst-adjungierte  $f$  folgt; die Aussage über Isometrien folgt mit Lemma 26.5. □

**26.8. Satz.** Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum und sei  $f \in \text{End}(V)$ . Wenn  $f$  orthogonal diagonalisierbar ist, dann ist  $f$  selbst-adjungiert.

**SATZ**  
orthog. diag.  
 $\Rightarrow$  selbst-adj.

*Beweis.* Nach Voraussetzung gibt es eine ONB  $B$  von  $V$ , sodass  $A = \text{Mat}_B(f)$  eine Diagonalmatrix ist. Dann gilt auch  $A = A^\top$ , also ist  $f$  nach Lemma 26.7 selbst-adjungiert. □

Zum Beweis der Gegenrichtung machen wir eine Vorüberlegung. Wir wissen, dass  $V$  zum Standardraum  $\mathbb{R}^n$  isometrisch ist, also können wir ohne Einschränkung  $V = \mathbb{R}^n$  (mit  $n > 0$ ) betrachten. Die Menge  $S = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$  (also die Oberfläche der  $n$ -dimensionalen Einheitskugel) ist eine abgeschlossene und beschränkte Teilmenge von  $\mathbb{R}^n$ , also ist  $S$  kompakt. Die Abbildung

$$h : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \mathbf{x} \mapsto \langle f(\mathbf{x}), \mathbf{x} \rangle$$

ist stetig, denn  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ist stetig und das Skalarprodukt ist ebenfalls stetig (vgl. Analysis). Als stetige Funktion nimmt die Abbildung  $h$  auf der kompakten Menge  $S$  ihr Maximum an, etwa in  $\mathbf{x}_0 \in S$ .

**26.9. Lemma.** *In der eben diskutierten Situation (mit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  selbst-adjungiert) ist  $\mathbf{x}_0$  ein Eigenvektor von  $f$  zum Eigenwert  $\lambda = h(\mathbf{x}_0)$ .*

**LEMMA**  
Maximum ist  
Eigenwert

*Beweis.* Wir bemerken zunächst, dass für alle  $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$  gilt

$$\langle f(\mathbf{x}), \mathbf{x} \rangle = \|\mathbf{x}\|^2 \left\langle f\left(\frac{1}{\|\mathbf{x}\|}\mathbf{x}\right), \frac{1}{\|\mathbf{x}\|}\mathbf{x} \right\rangle \leq \lambda \|\mathbf{x}\|^2,$$

denn  $\|\mathbf{x}\|^{-1}\mathbf{x} \in S$ . Für  $\mathbf{x} = \mathbf{0}$  gilt die Ungleichung  $\langle f(\mathbf{x}), \mathbf{x} \rangle \leq \lambda \|\mathbf{x}\|^2$  ebenfalls.

Wir zeigen jetzt, dass  $\mathbf{x}_0$  ein Eigenvektor ist. Wir können  $f(\mathbf{x}_0) = \mu\mathbf{x}_0 + \mathbf{y}$  schreiben mit  $\mu \in \mathbb{R}$  und  $\mathbf{y} \perp \mathbf{x}_0$  (nach Lemma 25.10 mit  $U = \langle \mathbf{x}_0 \rangle_{\mathbb{R}}$ ). Für  $t \in \mathbb{R}$  betrachten wir den Vektor  $\mathbf{x}_0 + t\mathbf{y}$ . Es gilt

$$\langle f(\mathbf{x}_0 + t\mathbf{y}), \mathbf{x}_0 + t\mathbf{y} \rangle \leq \lambda \|\mathbf{x}_0 + t\mathbf{y}\|^2 = \lambda(1 + t^2\|\mathbf{y}\|^2) = \lambda + t^2\lambda\|\mathbf{y}\|^2$$

(dabei haben wir  $\mathbf{x}_0 \perp \mathbf{y}$  und den „Pythagoras“ benutzt). Auf der anderen Seite ist

$$\begin{aligned} \langle f(\mathbf{x}_0 + t\mathbf{y}), \mathbf{x}_0 + t\mathbf{y} \rangle &= \langle f(\mathbf{x}_0) + tf(\mathbf{y}), \mathbf{x}_0 + t\mathbf{y} \rangle \\ &= \langle f(\mathbf{x}_0), \mathbf{x}_0 \rangle + t(\langle f(\mathbf{x}_0), \mathbf{y} \rangle + \langle f(\mathbf{y}), \mathbf{x}_0 \rangle) + t^2\langle f(\mathbf{y}), \mathbf{y} \rangle \\ &= \lambda + t(\langle f(\mathbf{x}_0), \mathbf{y} \rangle + \langle \mathbf{y}, f(\mathbf{x}_0) \rangle) + t^2\langle f(\mathbf{y}), \mathbf{y} \rangle \\ &= \lambda + 2t\langle f(\mathbf{x}_0), \mathbf{y} \rangle + t^2\langle f(\mathbf{y}), \mathbf{y} \rangle \\ &= \lambda + 2t\langle \mu\mathbf{x}_0 + \mathbf{y}, \mathbf{y} \rangle + t^2\langle f(\mathbf{y}), \mathbf{y} \rangle \\ &= \lambda + 2t\|\mathbf{y}\|^2 + t^2\langle f(\mathbf{y}), \mathbf{y} \rangle. \end{aligned}$$

Dabei haben wir verwendet, dass  $f$  selbst-adjungiert und  $\mathbf{x}_0 \perp \mathbf{y}$  ist. Für  $t > 0$  ergibt sich daraus die Ungleichung (nach Subtraktion von  $\lambda$  und Division durch  $t$ )

$$2\|\mathbf{y}\|^2 \leq t(\lambda\|\mathbf{y}\|^2 - \langle f(\mathbf{y}), \mathbf{y} \rangle);$$

wenn wir  $t$  von oben gegen null gehen lassen, folgt daraus  $\|\mathbf{y}\|^2 = 0$ , also  $\mathbf{y} = \mathbf{0}$  und damit  $f(\mathbf{x}_0) = \mu\mathbf{x}_0$ . Außerdem gilt

$$\lambda = \langle f(\mathbf{x}_0), \mathbf{x}_0 \rangle = \langle \mu\mathbf{x}_0, \mathbf{x}_0 \rangle = \mu\|\mathbf{x}_0\|^2 = \mu,$$

also ist  $\lambda$  der zu  $\mathbf{x}_0$  gehörende Eigenwert. □

**26.10. Folgerung.** *Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum mit  $\dim V > 0$  und sei  $f \in \text{End}(V)$  selbst-adjungiert. Dann hat  $f$  einen (reellen) Eigenwert.*

**FOLG**  
selbst-adj.  
Abb. haben  
Eigenwert

*Beweis.* Wir wählen eine ONB von  $V$ ; dann gibt es eine Isometrie  $\phi : \mathbb{R}^n \rightarrow V$  (mit  $n = \dim V$ ). Die Abbildung  $\tilde{f} = \phi^{-1} \circ f \circ \phi \in \text{End}(\mathbb{R}^n)$  ist ebenfalls selbst-adjungiert, hat also nach Lemma 26.9 einen Eigenwert  $\lambda \in \mathbb{R}$ . Da  $\tilde{f}$  und  $f$  dieselben Eigenwerte haben (ist  $\mathbf{x} \in \mathbb{R}^n$  Eigenvektor von  $\tilde{f}$  zum Eigenwert  $\lambda$ , dann ist  $\phi(\mathbf{x})$  Eigenvektor von  $f$  zum selben Eigenwert), gilt das auch für  $f$ . □

\* 26.11. **Satz.** Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum und sei  $f \in \text{End}(V)$  selbst-adjungiert. Dann ist  $f$  orthogonal diagonalisierbar, d.h.,  $V$  besitzt eine Orthonormalbasis, die aus Eigenvektoren von  $f$  besteht.

**SATZ**  
selbst-adj.  
Abb. sind  
orthogonal  
diag.bar

*Beweis.* Durch Induktion über  $n = \dim V$ . Für  $n = 0$  ist nichts zu beweisen (die leere Familie ist eine Basis aus Eigenvektoren). Sei also jetzt  $n > 0$  und die Aussage für  $\dim V = n - 1$  richtig. Nach Folgerung 26.10 hat  $f$  einen Eigenwert  $\lambda \in \mathbb{R}$ ; sei  $e_n \in V$  ein zugehöriger Eigenvektor mit  $\|e_n\| = 1$ . Sei  $U \subset V$  das orthogonale Komplement von  $\langle e_n \rangle_{\mathbb{R}}$ . Dann ist  $U$  ein  $f$ -invarianter Untervektorraum, denn für  $u \in U$  gilt

$$\langle f(u), e_n \rangle = \langle u, f(e_n) \rangle = \langle u, \lambda e_n \rangle = \lambda \langle u, e_n \rangle = 0,$$

also ist  $f(u) \in U$ .  $U$  ist (mit dem auf  $U \times U$  eingeschränkten Skalarprodukt von  $V$ ) ein euklidischer Vektorraum mit  $\dim U = n - 1$ , und  $f|_U$  ist ein selbst-adjungierter Endomorphismus von  $U$ . Nach der Induktionsannahme hat also  $U$  eine ONB  $(e_1, e_2, \dots, e_{n-1})$ , die aus Eigenvektoren von  $f$  besteht. Dann ist  $(e_1, \dots, e_{n-1}, e_n)$  eine ONB von  $V$  aus Eigenvektoren von  $f$ .  $\square$

Für Matrizen lässt sich das Ergebnis auch so formulieren:

26.12. **Folgerung.** Sei  $A \in \text{Mat}(n, \mathbb{R})$  eine symmetrische Matrix. Dann gibt es eine orthogonale Matrix  $P \in O(n)$ , sodass  $P^{\top}AP = P^{-1}AP$  eine Diagonalmatrix ist.

**FOLG**  
Spektralsatz

Das ist Satz 16.9, den wir im ersten Semester benutzt haben, um das Determinantenkriterium für positive Definitheit (Satz 16.13) zu beweisen.

*Beweis.* Sei  $f : \mathbf{x} \mapsto A\mathbf{x}$ , dann ist  $A$  die Matrix von  $f \in \text{End}(\mathbb{R}^n)$  bezüglich der Standardbasis  $E$ ; da  $A$  symmetrisch ist, ist  $f$  selbst-adjungiert (Lemma 26.7). Nach Satz 26.11 hat  $\mathbb{R}^n$  eine Orthonormalbasis  $B$  aus Eigenvektoren von  $f$ , also ist  $D = \text{Mat}_B(f)$  eine Diagonalmatrix. Die Matrix  $P = \text{Mat}_{B,E}(\text{id}_{\mathbb{R}^n})$  hat als Spalten die Vektoren von  $B$  und ist damit orthogonal (vergleiche die Bemerkung nach Definition 26.6). Außerdem ist

$$P^{-1}AP = \text{Mat}_{E,B}(\text{id}_{\mathbb{R}^n}) \text{Mat}_E(f) \text{Mat}_{B,E}(\text{id}_{\mathbb{R}^n}) = \text{Mat}_B(f) = D. \quad \square$$

## 27. KLASSIFIKATION VON QUADRIKEN

Wir arbeiten in diesem Abschnitt im Standardraum  $\mathbb{R}^n$ .

Eine lineare Gleichung (mit  $a_1, a_2, \dots, a_n, c \in \mathbb{R}$  gegeben, nicht alle  $a_j = 0$ , und  $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  gesucht)

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = c$$

hat als Lösungsmenge eine *affine Hyperebene* (also einen affinen Unterraum der Dimension  $n - 1$ ). Viel mehr gibt es dazu nicht zu sagen. Daher befassen wir uns jetzt mit *quadratischen* Gleichungen. Sie haben die allgemeine Form

$$\sum_{i,j=1}^n a_{ij}x_ix_j + \sum_{j=1}^n b_jx_j = c$$

oder kurz

$$\langle \mathbf{x}, A\mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{x} \rangle = c \quad \text{bzw.} \quad \mathbf{x}^\top A\mathbf{x} + \mathbf{b}^\top \mathbf{x} = c;$$

dabei ist  $\mathbf{0} \neq A = (a_{ij}) \in \text{Mat}(n, \mathbb{R})$  eine symmetrische Matrix,  $\mathbf{b} = (b_1, \dots, b_n)^\top$  ein Vektor und  $c \in \mathbb{R}$ . (Ist  $A$  die Nullmatrix, dann ist die Gleichung nicht wirklich quadratisch.)

\*

**27.1. Definition.** Die Lösungsmenge einer quadratischen Gleichung wie oben heißt *Quadratik* im  $\mathbb{R}^n$ . Eine Quadratik im  $\mathbb{R}^2$  heißt auch *Kegelschnitt*.  $\diamond$

**DEF**  
Quadratik

Die Bezeichnung „Kegelschnitt“ für Quadriken im  $\mathbb{R}^2$  kommt daher, dass sich (fast) alle solchen Quadriken als Schnitt des Doppelkegels

$$x_1^2 + x_2^2 - x_3^2 = 0$$

im  $\mathbb{R}^3$  mit einer Ebene realisieren lassen.

Man kann analog Quadriken auch über anderen Körpern (zum Beispiel  $\mathbb{C}$ ) definieren und studieren.

**27.2. Beispiel.** Ein einfaches Beispiel für eine Quadratik im  $\mathbb{R}^2$ , also für einen Kegelschnitt, ist der *Einheitskreis*, der die Lösungsmenge der quadratischen Gleichung

$$x_1^2 + x_2^2 = 1$$

ist. (Hier ist  $A = I_2$  die Einheitsmatrix,  $\mathbf{b} = \mathbf{0}$  und  $c = 1$ .) Allgemeiner ist ein Kreis mit Mittelpunkt  $(m_1, m_2)$  und Radius  $r$  ebenfalls ein Kegelschnitt; hier lautet die Gleichung

$$(x_1 - m_1)^2 + (x_2 - m_2)^2 = r^2,$$

was zu

$$x_1^2 + x_2^2 - 2m_1x_1 - 2m_2x_2 = r^2 - m_1^2 - m_2^2$$

äquivalent ist (also  $A = I_2$ ,  $\mathbf{b} = -2(m_1, m_2)^\top$  und  $c = r^2 - m_1^2 - m_2^2$ ).  $\clubsuit$

**BSP**  
Kreis

Eine Drehung, Spiegelung (allgemeiner eine Isometrie) oder Verschiebung ändert die geometrische Form einer Quadratik nicht. Deshalb sind wir an einer Normalform bzw. Klassifikation bis auf solche Geometrie erhaltenden Abbildungen interessiert. Wir geben diesen Abbildungen zuerst einen Namen. Vorher erinnern wir uns daran, dass die Determinante einer orthogonalen Matrix stets  $\pm 1$  ist:

$$A^\top A = I_n \implies 1 = \det(I_n) = \det(A^\top) \det(A) = \det(A)^2.$$

Die zugehörigen Isometrien  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  sind also orientierungserhaltend, wenn  $\det(A) = 1$  ist, und orientierungsumkehrend, wenn  $\det(A) = -1$  ist (vergleiche Definition 17.2). Die orthogonalen Matrizen mit Determinante 1 bilden eine Untergruppe von  $O(n)$ , die *spezielle orthogonale Gruppe*  $SO(n)$ .

**27.3. Definition.** Sei  $V$  ein euklidischer Vektorraum. Eine Abbildung  $f : V \rightarrow V$  heißt eine (*euklidische*) *Bewegung* von  $V$ , wenn es eine Isometrie  $h : V \rightarrow V$  und einen Vektor  $v_0 \in V$  gibt mit  $f(v) = h(v) + v_0$  für alle  $v \in V$ . Im Fall  $h = \text{id}_V$  heißt  $f$  auch *Translation* um  $v_0$ .

**DEF**  
Bewegung

Die Bewegung ist *orientierungserhaltend* bzw. *-umkehrend*, wenn  $h$  orientierungserhaltend (also  $\det(h) > 0$ ) bzw. *-umkehrend* ( $\det(h) < 0$ ) ist.  $\diamond$

Die Menge aller Bewegungen von  $V$  bildet eine Gruppe (Übung), die *Bewegungsgruppe* von  $V$ .

**27.4. Beispiel.** Ist  $H = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{b}^\top \mathbf{x} = c\}$  eine affine Hyperebene, dann gibt es eine Bewegung  $T$  mit  $T(H) = \{\mathbf{x} \in \mathbb{R}^n \mid x_1 = 0\}$ . Dafür ergänzen wir  $\mathbf{b}$  zu einer Basis von  $\mathbb{R}^n$  und wenden das Gram-Schmidt-Verfahren an. Wir erhalten eine ONB, deren erster Vektor ein skalares Vielfaches  $\lambda \mathbf{b}$  ist. Sei  $P$  die zugehörige orthogonale Matrix (deren Spalten diese ONB bilden). Wir identifizieren  $P$  mit der Isometrie  $\mathbf{x} \mapsto P\mathbf{x}$ . Dann ist

**BSP**  
Klassif.  
affiner  
Hyperebenen

$$\begin{aligned} P^{-1}(H) &= \{P^{-1}\mathbf{x} \mid \mathbf{x} \in H\} = \{\mathbf{x} \mid P\mathbf{x} \in H\} = \{\mathbf{x} \mid \mathbf{b}^\top P\mathbf{x} = c\} \\ &= \{\mathbf{x} \mid \mathbf{b}^\top (P^{-1})^\top \mathbf{x} = c\} = \{\mathbf{x} \mid (P^{-1}\mathbf{b})^\top \mathbf{x} = c\} = \{\mathbf{x} \mid x_1 = \lambda c\}, \end{aligned}$$

denn  $P\mathbf{e}_1 = \lambda \mathbf{b}$ , also ist  $P^{-1}\mathbf{b} = \lambda^{-1}\mathbf{e}_1$ . Translation um den Vektor  $(\lambda c, 0, \dots, 0)^\top$  ergibt schließlich die Hyperebene  $x_1 = 0$ .  $\clubsuit$

Das Hauptergebnis in diesem Abschnitt lautet wie folgt.

**\* 27.5. Satz.** Sei  $Q = \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, A\mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{x} \rangle = c\}$  eine Quadrik im  $\mathbb{R}^n$  (mit  $n \geq 1$ ) und sei  $r$  der Rang von  $A$ . Dann gibt es eine orientierungserhaltende Bewegung  $T$  des  $\mathbb{R}^n$ , reelle Zahlen  $a_1, \dots, a_r > 0$  und Vorzeichen  $\varepsilon_1, \dots, \varepsilon_r \in \{\pm 1\}$  mit

**SATZ**  
euklidische  
Normalform  
von Quadriken

$$\begin{aligned} T(Q) &= \left\{ \mathbf{x} \in \mathbb{R}^n \mid \varepsilon_1 \left(\frac{x_1}{a_1}\right)^2 + \dots + \varepsilon_r \left(\frac{x_r}{a_r}\right)^2 = 0 \right\} \quad \text{oder} \\ T(Q) &= \left\{ \mathbf{x} \in \mathbb{R}^n \mid \varepsilon_1 \left(\frac{x_1}{a_1}\right)^2 + \dots + \varepsilon_r \left(\frac{x_r}{a_r}\right)^2 = 1 \right\} \quad \text{oder} \\ T(Q) &= \left\{ \mathbf{x} \in \mathbb{R}^n \mid \varepsilon_1 \left(\frac{x_1}{a_1}\right)^2 + \dots + \varepsilon_r \left(\frac{x_r}{a_r}\right)^2 = x_{r+1} \right\}. \end{aligned}$$

Im zweiten Fall sind die  $(a_j, \varepsilon_j)$  bis auf Permutation eindeutig bestimmt, im dritten Fall ist zusätzlich eine Umkehr aller Vorzeichen möglich, im ersten Fall ist auch noch eine zusätzliche gemeinsame Skalierung der  $a_j$  zugelassen.

*Beweis.* Nach dem Spektralsatz 26.12 gibt es  $P \in O(n)$  mit  $P^\top A P = D$  diagonal. Dabei können wir annehmen, dass die  $n - r$  Nullen auf der Diagonalen von  $D$  am Ende kommen. Falls  $\det(P) = -1$ , können wir  $P$  ersetzen durch  $P' = P \text{diag}(-1, 1, \dots, 1)$ , dann gilt auch

$$P'^\top A P' = \text{diag}(-1, 1, \dots, 1) D \text{diag}(-1, 1, \dots, 1) = D$$

und  $\det(P') = 1$ . Wir können also  $P \in SO(n)$  annehmen. Wenn wir  $P\mathbf{x}$  in die Gleichung von  $Q$  einsetzen, erhalten wir

$$c = \mathbf{x}^\top P^\top A P \mathbf{x} + \mathbf{b}^\top P \mathbf{x} = \mathbf{x}^\top D \mathbf{x} + (\mathbf{b}^\top P) \mathbf{x}.$$

Ausgeschrieben lautet das

$$\lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + b'_1 x_1 + \dots + b'_n x_n = c.$$

(Dabei sind  $\lambda_1, \dots, \lambda_r$  die von null verschiedenen Eigenwerte von  $A$ .) Die neue Quadrik ist das Bild der ursprünglichen unter der orientierungserhaltenden Isometrie  $P^{-1}$ .

Hier können wir in den ersten  $r$  Variablen quadratisch ergänzen: Wir ersetzen  $x_j$  durch  $x_j - \frac{b'_j}{2\lambda_j}$  und erhalten die neue Gleichung

$$\lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + b'_{r+1} x_{r+1} + \dots + b'_n x_n = c'$$

mit

$$c' = c + \frac{1}{4} \left( \frac{(b'_1)^2}{\lambda_1} + \dots + \frac{(b'_r)^2}{\lambda_r} \right).$$

Diese Quadrik entsteht aus der vorigen durch eine Translation. Jetzt gibt es drei mögliche Fälle:

- $b'_{r+1} = \dots = b'_n = c' = 0$ . Dann hat die Gleichung die gewünschte Form mit  $\varepsilon_j = \text{sign } \lambda_j$  und  $a_j = 1/\sqrt{|\lambda_j|}$  für  $j \in \{1, 2, \dots, r\}$ .
- $b'_{r+1} = \dots = b'_n = 0$  und  $c' \neq 0$ . Dann teilen wir die Gleichung durch  $c'$ ; mit  $\varepsilon_j = \text{sign}(\lambda_j/c')$  und  $a_j = \sqrt{|c'/\lambda_j|}$  für  $j \in \{1, 2, \dots, r\}$  haben wir die zweite Form.
- $b'_j \neq 0$  für ein  $j \in \{r+1, \dots, n\}$ . Wie in Beispiel 27.4 kann der „lineare Teil“ der Gleichung auf die Form  $-\mu x_{r+1} = 0$  gebracht werden mit  $\mu \in \mathbb{R}^\times$ . Analog zum zweiten Fall setzen wir  $\varepsilon_j = \text{sign}(\lambda_j/\mu)$  und  $a_j = \sqrt{|\mu/\lambda_j|}$  um die gewünschte Form zu erhalten.

Die Eindeutigkeit ergibt sich daraus, dass die Koeffizienten der  $x_j^2$  zueinander im Verhältnis der Eigenwerte  $\neq 0$  von  $A$  stehen müssen; im zweiten Fall wird die Skalierung dadurch fixiert, dass die Konstante auf der rechten Seite 1 ist; im dritten Fall wird die Skalierung bis auf ein Vorzeichen durch den Koeffizienten von  $x_{r+1}$  festgelegt (denn das Vorzeichen von  $x_{r+1}$  in der Gleichung kann durch die orientierungserhaltende Isometrie, die nur die Vorzeichen von  $x_1$  und  $x_{r+1}$  ändert, umgedreht werden). □

**27.6. Definition.** In der Situation des zweiten Falls von Satz 27.5 mit  $r = n$  heißen die Zahlen  $a_1, a_2, \dots, a_n$  die *Halbachsen* von  $Q$ . Die Geraden  $T^{-1}(\langle \mathbf{e}_j \rangle_{\mathbb{R}})$  heißen die *Hauptachsen* von  $Q$ , der Punkt  $T^{-1}(\mathbf{0})$  der *Mittelpunkt* oder das *Zentrum* von  $Q$ . ◇

**DEF**  
Halbachsen  
Hauptachsen  
Mittelpunkt

Das erklärt auch die Bezeichnung *Hauptachsentransformation* für die Bewegung, die eine Quadrik in Normalform bringt.

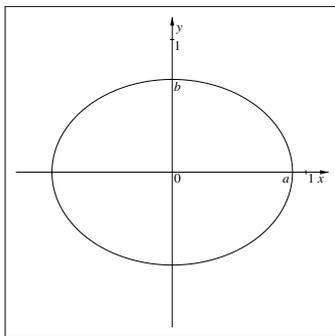
Je nach Verteilung der Vorzeichen erhalten wir verschiedene Typen von Quadriken. Für Kegelschnitte mit  $\text{rk}(A) = 2$  gibt es drei Möglichkeiten:

Vorz.		rechte Seite	
$\varepsilon_1$	$\varepsilon_2$	0	1
+	+	Punkt	Ellipse
+	-	Geradenpaar	Hyperbel
-	-	Punkt	leer

Bei Quadriken im  $\mathbb{R}^3$  mit Rang 3 sind es entsprechend vier:

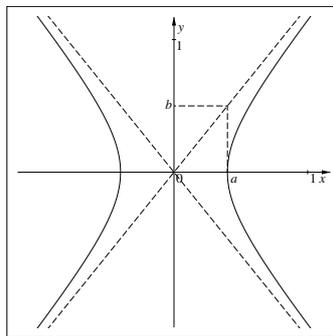
Vorz.			rechte Seite	
$\varepsilon_1$	$\varepsilon_2$	$\varepsilon_3$	0	1
+	+	+	Punkt	Ellipsoid
+	+	-	Doppelkegel	einschaliges Hyperboloid
+	-	-	Doppelkegel	zweischaliges Hyperboloid
-	-	-	Punkt	leer

### Einige Quadriken im $\mathbb{R}^2$



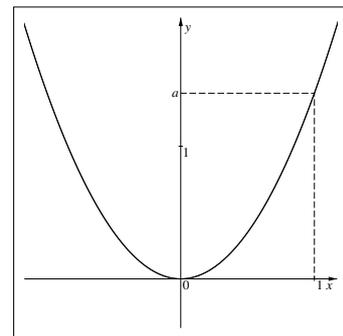
Ellipse:  

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$



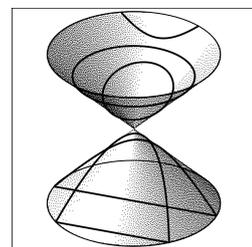
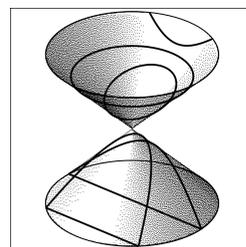
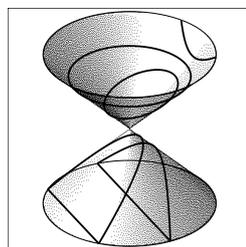
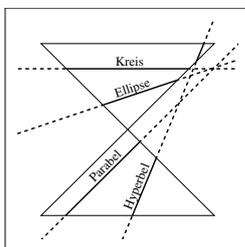
Hyperbel:  

$$\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1$$



Parabel:  

$$ax^2 = y$$

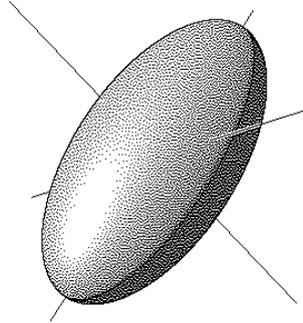


Man sieht, dass der Typ in der Form  $\mathbf{x}^T A \mathbf{x} = 1$  durch die Vorzeichen der Eigenwerte von  $A$  bestimmt ist; die Halbachsen sind durch  $1/\sqrt{|\lambda_j|}$  gegeben.

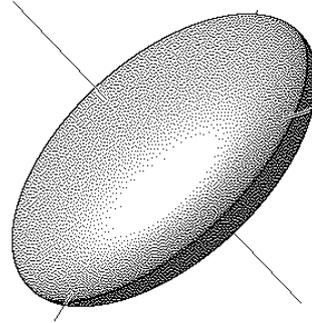
Kegelschnitte vom Rang 1 sind

Vorz.	rechte Seite		
$\varepsilon_1$	0	1	$x_2$
+	Doppelgerade	parallele Geraden	Parabel
-	Doppelgerade	leer	Parabel

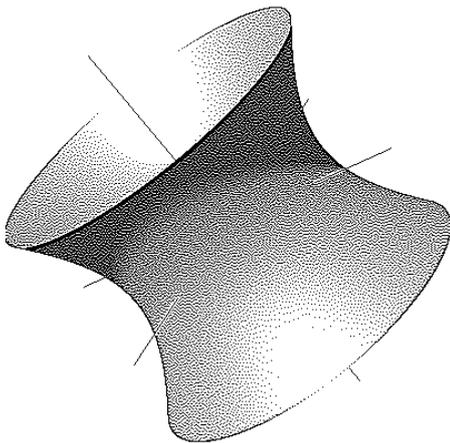
## Einige Quadriken im $\mathbb{R}^3$



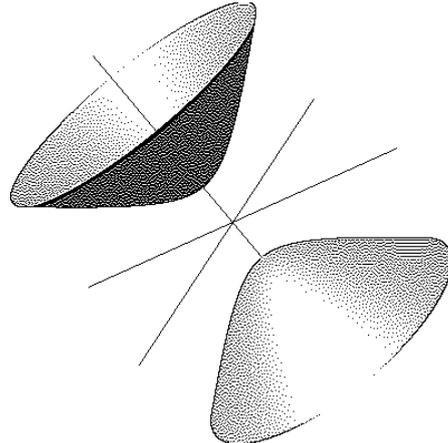
Prolates Rotationsellipsoid:  
 $5x^2 + y^2 + 5z^2 = 70$



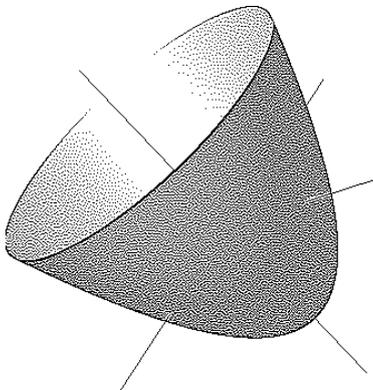
Oblates Rotationsellipsoid:  
 $x^2 + y^2 + 5z^2 = 70$



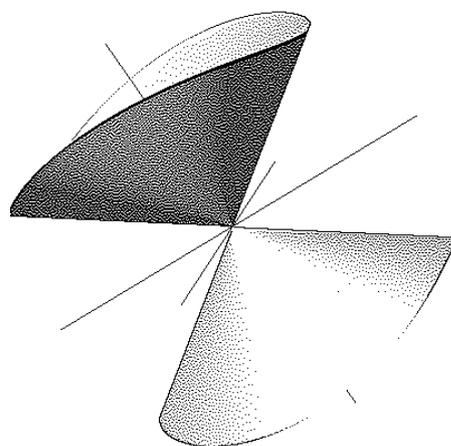
Einschaliges Rotationshyperboloid:  
 $x^2 + y^2 - z^2 = 30$



Zweischaliges Rotationshyperboloid:  
 $-x^2 - y^2 + z^2 = 5$



Rotationsparaboloid:  
 $x^2 + y^2 + 5z = 30$



Kegel:  
 $x^2 + 3y^2 - 2z^2 = 0$

Quadriken im  $\mathbb{R}^3$  vom Rang 2 sind

Vorz.		rechte Seite		
$\varepsilon_1$	$\varepsilon_2$	0	1	$x_3$
+	+	Doppelgerade	ellipt. Zylinder	ellipt. Paraboloid
+	-	Ebenenpaar	hyperbol. Zylinder	hyperbol. Paraboloid
-	-	Doppelgerade	leer	ellipt. Paraboloid

und vom Rang 1:

Vorz.		rechte Seite		
$\varepsilon_1$		0	1	$x_2$
+		doppelte Ebene	parallele Ebenen	parabol. Zylinder
-		doppelte Ebene	leer	parabol. Zylinder

27.7. **Beispiel.** Als einfaches Beispiel bestimmen wir die euklidische Normalform des Kegelschnitts (wir schreiben  $x, y$  statt  $x_1, x_2$  für die Koordinaten)

$$5x^2 + 4xy + 2y^2 = 1.$$

**BSP**  
Hauptachsen-  
transformation

Die zugehörige Matrix ist

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix};$$

der Vektor  $\mathbf{b}$  ist der Nullvektor und  $c = 1$ . Das charakteristische Polynom von  $A$  ist

$$\chi_A = \begin{vmatrix} X - 5 & -2 \\ -2 & X - 2 \end{vmatrix} = (X - 5)(X - 2) - 4 = X^2 - 7X + 6 = (X - 1)(X - 6),$$

also sind die Eigenwerte 1 und 6 mit zugehörigen Eigenvektoren

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Wie es sein muss, sind diese Vektoren zueinander orthogonal. Um eine ONB zu erhalten, müssen wir noch skalieren; das liefert die Transformationsmatrix

$$P = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \in \text{SO}(2)$$

mit

$$P^{-1}AP = P^TAP = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}.$$

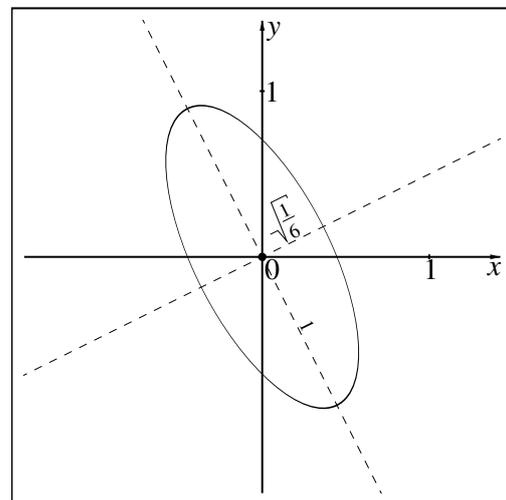
Die transformierte Gleichung lautet also

$$x^2 + 6y^2 = 1$$

oder

$$\left(\frac{x}{1}\right)^2 + \left(\frac{y}{1/\sqrt{6}}\right)^2 = 1;$$

das ist eine Ellipse mit Halbachsen 1 und  $1/\sqrt{6}$ .



Neben der euklidischen Normalform gibt auch die *affine Normalform*. Dabei sind statt Bewegungen *Affinitäten* erlaubt; eine Affinität ist eine Abbildung der Form

$$\mathbf{x} \mapsto A\mathbf{x} + \mathbf{b} \quad \text{mit} \quad A \in \text{GL}(n, \mathbb{R})$$

(also ein Automorphismus, gefolgt von einer Translation). Der Unterschied zu Bewegungen ist, dass zusätzlich eine Skalierung der Koordinaten möglich ist. Das hat den Effekt, dass in der Normalform aus Satz 27.5 alle  $a_j = 1$  gewählt werden können. Die affine Normalform legt bereits den Typ der Quadrik fest, da dieser nur von der Form der Gleichung und den Vorzeichen der quadratischen Terme abhängt. Für Kegelschnitte hat man also die folgenden affinen Normalformen:

Typ	Gleichung
Ellipse	$x^2 + y^2 = 1$
Hyperbel	$x^2 - y^2 = 1$
leere Menge	$-x^2 - y^2 = 1$
Punkt	$x^2 + y^2 = 0$
sich schneidende Geraden	$x^2 - y^2 = 0$
Parabel	$x^2 = y$
parallele Geraden	$x^2 = 1$
leere Menge	$-x^2 = 1$
Doppelgerade	$x^2 = 0$

Für Quadriken im  $\mathbb{R}^3$  sieht der „Zoo“ so aus:

Typ	Gleichung
Ellipsoid	$x^2 + y^2 + z^2 = 1$
einschaliges Hyperboloid	$x^2 + y^2 - z^2 = 1$
zweischaliges Hyperboloid	$x^2 - y^2 - z^2 = 1$
leere Menge	$-x^2 - y^2 - z^2 = 1$
Punkt	$x^2 + y^2 + z^2 = 0$
Doppelkegel	$x^2 + y^2 - z^2 = 0$
elliptisches Paraboloid	$x^2 + y^2 = z$
hyperbolisches Paraboloid	$x^2 - y^2 = z$
elliptischer Zylinder	$x^2 + y^2 = 1$
hyperbolischer Zylinder	$x^2 - y^2 = 1$
leere Menge	$-x^2 - y^2 = 1$
Gerade	$x^2 + y^2 = 0$
sich schneidende Ebenen	$x^2 - y^2 = 0$
parabolischer Zylinder	$x^2 = y$
parallele Ebenen	$x^2 = 1$
leere Menge	$-x^2 = 1$
Doppelebene	$x^2 = 0$

---



---

**Ende des Stoffs für die Lehramts-Klausur**

---



---

28. UNITÄRE VEKTORRÄUME

Wir wollen jetzt eine zu den euklidischen Vektorräumen analoge Theorie für komplexe (statt reelle) Vektorräume aufbauen. Eine symmetrische Bilinearform (wie im reellen Fall) können wir nicht verwenden, denn eine symmetrische Bilinearform  $\beta$  auf einem komplexen Vektorraum kann nicht positiv definit sein (man kann nicht einmal erreichen, dass  $\beta(v, v)$  stets reell ist), denn

$$\beta(\mathbf{i}v, \mathbf{i}v) = \mathbf{i}^2\beta(v, v) = -\beta(v, v).$$

Um das zu verhindern, modifizieren wir die Eigenschaften, die wir fordern.

\*

**28.1. Definition.** Sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Eine Abbildung  $\beta : V \times V \rightarrow \mathbb{C}$  heißt eine *Sesquilinearform* auf  $V$ , wenn sie linear im ersten und konjugiert-linear im zweiten Argument ist: Für alle  $v_1, v'_1, v_2, v'_2 \in V$  und alle  $\lambda \in \mathbb{C}$  gilt

$$\begin{aligned} \beta(v_1 + v'_1, v_2) &= \beta(v_1, v_2) + \beta(v'_1, v_2), & \beta(\lambda v_1, v_2) &= \lambda\beta(v_1, v_2); \\ \beta(v_1, v_2 + v'_2) &= \beta(v_1, v_2) + \beta(v_1, v'_2), & \beta(v_1, \lambda v_2) &= \bar{\lambda}\beta(v_1, v_2). \end{aligned}$$

**DEF**  
Sesqui-  
linearform  
hermitesch

Eine Sesquilinearform  $\beta$  auf  $V$  heißt *hermitesch*, wenn zusätzlich für alle  $v_1, v_2 \in V$  gilt

$$\beta(v_2, v_1) = \overline{\beta(v_1, v_2)}. \quad \diamond$$

„Sesqui-“ bedeutet „ $1\frac{1}{2}$ -fach“ (so wie „bi-“ „zweifach“ heißt); die konjugierte Linearität wird sozusagen halb gezählt (entsprechend heißt eine konjugiert-lineare Abbildung auch *semilinear*). Häufig wird in der Definition einer Sesquilinearform Linearität im zweiten und Semilinearität im ersten Argument gefordert (also umgekehrt wie in der Definition oben). Das macht keinen wesentlichen Unterschied; man muss nur beim Rechnen aufpassen, wann man Skalare konjugiert herausziehen muss.

Wir erinnern uns an die komplexe Konjugation: Für  $z = x + y\mathbf{i} \in \mathbb{C}$  mit  $x, y \in \mathbb{R}$  ist  $\bar{z} = x - y\mathbf{i}$ . Dann gilt für  $z, z_1, z_2 \in \mathbb{C}$

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2 \quad \text{und} \quad z\bar{z} = |z|^2.$$

$z$  ist genau dann reell, wenn  $z = \bar{z}$  ist.

Für eine hermitesche Sesquilinearform  $\beta$  auf  $V$  gilt dann  $\beta(v, v) \in \mathbb{R}$  für alle  $v \in V$ , denn

$$\beta(v, v) = \overline{\beta(v, v)}.$$

Daher ist die folgende Definition sinnvoll.

\*

**28.2. Definition.** Sei  $V$  ein komplexer Vektorraum und  $\beta$  eine hermitesche Sesquilinearform auf  $V$ .  $\beta$  heißt *positiv definit*, wenn für alle  $\mathbf{0} \neq v \in V$  gilt  $\beta(v, v) > 0$ . Eine positiv definite hermitesche Sesquilinearform auf  $V$  heißt auch ein *unitäres Skalarprodukt* auf  $V$ .

**DEF**  
unitäres  
Skalarprod.  
unitärer  
Vektorraum

Ein komplexer Vektorraum  $V$  zusammen mit einem unitären Skalarprodukt auf  $V$  heißt *unitärer Vektorraum*. Wie im reellen Fall schreiben wir das Skalarprodukt in einem unitären Vektorraum meistens in der Form  $\langle v_1, v_2 \rangle$ .  $\diamond$

28.3. Beispiele.

**BSP**  
unitäre  
Vektorräume

- Das Standardbeispiel ist  $\mathbb{C}^n$  mit dem Standard-Skalarprodukt

$$\langle (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \rangle = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n.$$

- Ein wichtiges Beispiel aus der Analysis ist der Raum  $V = \mathcal{C}([a, b], \mathbb{C})$  der stetigen Funktionen  $[a, b] \rightarrow \mathbb{C}$  mit dem Skalarprodukt

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

- Ein weiteres Beispiel aus der Analysis ist der Raum  $\ell^2(\mathbb{C})$  der komplexwertigen Folgen  $(a_n)_{n \in \mathbb{Z}}$  mit  $\sum_{n=-\infty}^{\infty} |a_n|^2 < \infty$ . Auf diesem Raum hat man das unitäre Skalarprodukt

$$\langle (a_n), (b_n) \rangle = \sum_{n=-\infty}^{\infty} a_n \bar{b}_n.$$

(Dass das wohldefiniert ist, folgt aus der Cauchy-Schwarzschen Ungleichung in Satz 25.4, denn man hat für alle  $N \in \mathbb{N}$ , dass

$$\begin{aligned} \left| \sum_{n=-N}^N a_n \bar{b}_n \right| &\leq \sum_{n=-N}^N |a_n| |b_n| \leq \sqrt{\sum_{n=-N}^N |a_n|^2} \sqrt{\sum_{n=-N}^N |b_n|^2} \\ &\leq \sqrt{\sum_{n=-\infty}^{\infty} |a_n|^2} \sqrt{\sum_{n=-\infty}^{\infty} |b_n|^2} < \infty \end{aligned}$$

ist.)



Wie im reellen Fall definieren wir die *Länge* eines Vektors  $v$  in einem unitären Vektorraum  $V$  durch

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Die grundlegenden Ungleichungen gelten weiterhin, mit im Wesentlichen denselben Beweisen.

\* 28.4. **Satz.** Sei  $V$  ein unitärer Vektorraum.

- (1) Für  $v \in V$  und  $\lambda \in \mathbb{C}$  gilt  $\|\lambda v\| = |\lambda| \|v\|$ .
- (2) (Cauchy-Schwarzsche Ungleichung) Für  $v, w \in V$  gilt  $|\langle v, w \rangle| \leq \|v\| \|w\|$  mit Gleichheit genau dann, wenn  $v$  und  $w$  linear abhängig sind.
- (3) (Dreiecksungleichung) Für  $v, w \in V$  gilt  $\|v+w\| \leq \|v\| + \|w\|$  mit Gleichheit genau dann, wenn  $v = \lambda w$  oder  $w = \lambda v$  ist mit  $\lambda \in \mathbb{R}_{\geq 0}$ .

**SATZ**  
Cauchy-  
Schwarz  
Dreiecksungl.

Wir übernehmen die Definitionen von *orthogonal*, *orthonormal* und *Orthonormalbasis* wörtlich von Definition 25.5. Das Gram-Schmidtsche Orthonormalisierungsverfahren aus Satz 25.8 funktioniert auch im unitären Fall; insbesondere hat ein endlich-dimensionaler unitärer Vektorraum stets eine ONB und jede orthonormale Teilmenge eines solchen unitären Vektorraums lässt sich zu einer ONB ergänzen.

Auch die Aussagen von Satz 25.11 und Satz 25.12 gelten, hier in der Form

$$(28.1) \quad v = \langle v, e_1 \rangle e_1 + \langle v, e_2 \rangle e_2 + \dots + \langle v, e_n \rangle e_n$$

für alle  $v \in V$ , wenn  $(e_1, e_2, \dots, e_n)$  eine Orthonormalbasis von  $V$  ist, und damit dann

$$\langle v, w \rangle = \langle v, e_1 \rangle \overline{\langle w, e_1 \rangle} + \langle v, e_2 \rangle \overline{\langle w, e_2 \rangle} + \dots + \langle v, e_n \rangle \overline{\langle w, e_n \rangle}$$

und speziell

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + |\langle v, e_2 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2.$$

(Da das Skalarprodukt nicht symmetrisch (sondern hermitesch) ist, kommt es jetzt auf die Reihenfolge an. Die richtige Anordnung ergibt sich daraus, dass die rechte Seite in Gleichung (28.1) linear (und nicht semilinear) in  $v$  sein muss.)

Wie im euklidischen Fall definiert man eine *Isometrie* zwischen unitären Vektorräumen als einen Isomorphismus, der mit dem Skalarprodukt verträglich ist. Es gilt wieder, dass ein Isomorphismus bereits dann eine Isometrie ist, wenn er Längen erhält, denn man kann auch ein unitäres Skalarprodukt durch Längen ausdrücken:

$$4\langle v, w \rangle = \|v + w\|^2 + \mathbf{i}\|v + \mathbf{i}w\|^2 - \|v - w\|^2 - \mathbf{i}\|v - \mathbf{i}w\|^2.$$

Es gilt auch die folgende Aussage analog zu Lemma 26.1:

**28.5. Lemma.** *Sei  $V$  ein endlich-dimensionaler unitärer Vektorraum und sei  $\phi \in V^*$ . Dann gibt es einen eindeutig bestimmten Vektor  $w \in V$  mit*

$$\phi(v) = \langle v, w \rangle \quad \text{für alle } v \in V.$$

**LEMMA**  
Linearformen  
via  $\langle \cdot, \cdot \rangle$

Die Version mit Fixierung des ersten Arguments gilt so nicht, denn  $v \mapsto \langle w, v \rangle$  ist semilinear, nicht linear. Man kann das aber reparieren, indem man die komplexe Konjugation nachschaltet: Wegen  $\langle v, w \rangle = \overline{\langle w, v \rangle}$  kann man  $\phi$  auch in der Form

$$\phi(v) = \overline{\langle w, v \rangle}$$

darstellen.

Wir übertragen die Definition der adjungierten Abbildung auf den unitären Fall.

**\* 28.6. Definition.** Seien  $V$  und  $W$  unitäre Vektorräume und sei  $f : V \rightarrow W$  linear.

**DEF**  
adjungierte  
Abbildung  
normal

- (1) Gibt es eine lineare Abbildung  $f^* : W \rightarrow V$ , sodass für alle  $v \in V$  und  $w \in W$  gilt  $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$ , dann heißt  $f^*$  die zu  $f$  *adjungierte Abbildung*. (Es gilt dann auch  $\langle f^*(w), v \rangle = \langle w, f(v) \rangle$ .)
- (2) Hat  $f \in \text{End}(V)$  eine adjungierte Abbildung  $f^*$  und gilt  $f = f^*$ , dann heißt  $f$  *selbst-adjungiert*. Das bedeutet also  $\langle f(v), v' \rangle = \langle v, f(v') \rangle$  für alle  $v, v' \in V$ .
- (3) Hat  $f \in \text{End}(V)$  eine adjungierte Abbildung  $f^*$  und gilt  $f \circ f^* = f^* \circ f$ , dann heißt  $f$  *normal*. ◇

Ähnlich wie im euklidischen Fall sieht man, dass

$$\text{Mat}_{B',B}(f^*) = \overline{\text{Mat}_{B,B'}(f)}^\top$$

gilt, wenn  $B$  und  $B'$  Orthonormalbasen von  $V$  bzw.  $W$  sind. Dabei steht  $\bar{A}$  für die Matrix, die aus  $A$  entsteht, indem man jeden Eintrag komplex konjugiert. Daran sieht man, dass zwar nach wie vor

$$(f + g)^* = f^* + g^*, \quad (f \circ g)^* = g^* \circ f^* \quad \text{und} \quad (f^*)^* = f$$

gelten (für geeignete  $f$  und  $g$ ), aber für die Skalarmultiplikation haben wir jetzt

$$(\lambda f)^* = \bar{\lambda} f^*, \quad \text{insbesondere } (\lambda \text{id}_V)^* = \bar{\lambda} \text{id}_V.$$

(Das bedeutet, dass die Abbildung  $\text{Hom}(V, W) \rightarrow \text{Hom}(W, V)$ ,  $f \mapsto f^*$ , jetzt semilinear ist statt linear wie im euklidischen Fall.)

Wenn wir den Standardraum  $\mathbb{C}^n$  mit der Standardbasis (die eine ONB ist) betrachten und eine Matrix  $A \in \text{Mat}(n, \mathbb{C})$  mit dem Endomorphismus  $x \mapsto Ax$  identifizieren, dann ist  $A^* = \bar{A}^\top$ .

\*

**28.7. Definition.** Eine Matrix  $A \in \text{Mat}(n, \mathbb{C})$  heißt *hermitesch*, wenn  $A = A^*$  ist.  $A$  heißt *unitär*, wenn  $AA^* = I_n$  ist, und *normal*, wenn  $AA^* = A^*A$  gilt.

**DEF**  
hermitesche,  
unitäre  
Matrix

Die Gruppe (!) der unitären  $n \times n$ -Matrizen heißt die *unitäre Gruppe* und wird mit  $U(n)$  bezeichnet. Die unitären Matrizen mit Determinante 1 bilden ebenfalls eine Gruppe, die *spezielle unitäre Gruppe*  $SU(n)$ .  $\diamond$

Eine Matrix ist genau dann unitär, wenn ihre Spalten (oder Zeilen) eine ONB bilden. Das ist genau dann der Fall, wenn die zugehörige lineare Abbildung eine Isometrie ist. Hermitesche und unitäre Matrizen sind normal.

Für die Determinante einer unitären Matrix  $A$  gilt

$$\begin{aligned} 1 &= \det(I_n) = \det(AA^*) = \det(A) \det(\bar{A}^\top) \\ &= \det(A) \det(\bar{A}) = \det(A) \overline{\det(A)} \\ &= |\det(A)|^2, \end{aligned}$$

also  $|\det(A)| = 1$ .

Wir wenden uns jetzt der Frage nach der unitären Diagonalisierbarkeit zu: Ein Endomorphismus  $f$  eines endlich-dimensionalen unitären Vektorraums  $V$  ist *unitär diagonalisierbar*, wenn  $V$  eine ONB hat, die aus Eigenvektoren von  $f$  besteht. Es wird sich herausstellen, dass das genau für normale Endomorphismen der Fall ist. Wir beginnen mit einem Lemma.

**28.8. Lemma.** Sei  $V$  ein unitärer Vektorraum und sei  $f \in \text{End}(V)$  normal.

**LEMMA**

- (1) Es gilt  $\|f^*(v)\| = \|f(v)\|$  für alle  $v \in V$ .
- (2) Ist  $v \in V$  ein Eigenvektor von  $f$  zum Eigenwert  $\lambda$ , dann ist  $v$  auch ein Eigenvektor von  $f^*$  zum Eigenwert  $\bar{\lambda}$ .

*Beweis.* Die erste Aussage sieht man so:

$$\|f^*(v)\|^2 = \langle f^*(v), f^*(v) \rangle = \langle f(f^*(v)), v \rangle = \langle f^*(f(v)), v \rangle = \langle f(v), f(v) \rangle = \|f(v)\|^2.$$

Die zweite Aussage folgt daraus: Zunächst einmal ist mit  $f$  auch  $f - \lambda \text{id}_V$  normal, denn

$$\begin{aligned} (f - \lambda \text{id}_V) \circ (f - \lambda \text{id}_V)^* &= (f - \lambda \text{id}_V) \circ (f^* - \bar{\lambda} \text{id}_V) \\ &= f \circ f^* - \lambda f^* - \bar{\lambda} f + |\lambda|^2 \text{id}_V \\ &= f^* \circ f - \lambda f^* - \bar{\lambda} f + |\lambda|^2 \text{id}_V \\ &= (f^* - \bar{\lambda} \text{id}_V) \circ (f - \lambda \text{id}_V) \\ &= (f - \lambda \text{id}_V)^* \circ (f - \lambda \text{id}_V). \end{aligned}$$

Aus  $f(v) = \lambda v$  ergibt sich dann

$$0 = \|(f - \lambda \text{id}_V)(v)\| = \|(f - \lambda \text{id}_V)^*(v)\| = \|(f^* - \bar{\lambda} \text{id}_V)(v)\| = \|f^*(v) - \bar{\lambda} v\|;$$

damit ist  $f^*(v) = \bar{\lambda}v$ . □

\* **28.9. Satz.** *Sei  $f$  ein Endomorphismus eines endlich-dimensionalen unitären Vektorraums  $V$ . Dann ist  $f$  unitär diagonalisierbar genau dann, wenn  $f$  normal ist.* **SATZ**  
Spektralsatz

*Beweis.* Sei zunächst  $f$  unitär diagonalisierbar. Dann ist die Matrix von  $f$  bezüglich einer ONB von  $V$ , die aus Eigenvektoren von  $f$  besteht, diagonal und damit normal. Es folgt, dass  $f$  ebenfalls normal ist.

Die umgekehrte Implikation beweisen wir durch Induktion über die Dimension  $n$  des Vektorraums  $V$ . Für  $n = 0$  (oder  $n = 1$ ) ist nichts zu zeigen. Sei also  $n \geq 1$ . Weil  $\mathbb{C}$  algebraisch abgeschlossen ist, hat das charakteristische Polynom von  $f$  eine Nullstelle, also hat  $f$  einen Eigenwert  $\lambda$  mit zugehörigem Eigenvektor  $v_n$ . Nach Skalieren können wir annehmen, dass  $\|v_n\| = 1$  ist. Nach Lemma 28.8 ist  $f^*(v_n) = \bar{\lambda}v_n$ . Wir betrachten das orthogonale Komplement von  $\langle v_n \rangle_{\mathbb{C}}$ :

$$U = \{u \in V \mid \langle u, v_n \rangle = 0\}.$$

Dann ist  $U$  ein  $f$ -invarianter Untervektorraum von  $V$ , denn für  $u \in U$  gilt

$$\langle f(u), v_n \rangle = \langle u, f^*(v_n) \rangle = \langle u, \bar{\lambda}v_n \rangle = \lambda \langle u, v_n \rangle = 0$$

und damit  $f(u) \in U$ . Analog sieht man, dass  $U$  ein  $f^*$ -invarianter Untervektorraum ist. Damit ist  $f|_U$  ein normaler Endomorphismus von  $U$  ( $U$  ist ein unitärer Vektorraum mit dem eingeschränkten Skalarprodukt); außerdem gilt wie im euklidischen Fall  $V = \langle v_n \rangle_{\mathbb{C}} \oplus U$ . Nach Induktionsannahme hat  $U$  eine ONB  $(v_1, \dots, v_{n-1})$  aus Eigenvektoren von  $f$ ; dann ist  $(v_1, \dots, v_{n-1}, v_n)$  eine ONB von  $V$  aus Eigenvektoren von  $f$ . □

Für Matrizen lautet die interessante Richtung dieser Aussage wie folgt (der Beweis ist analog zum euklidischen Fall):

\* **28.10. Folgerung.** *Ist  $A \in \text{Mat}(n, \mathbb{C})$  normal, dann gibt es eine unitäre Matrix  $P \in U(n)$ , sodass  $P^{-1}AP = P^*AP$  eine Diagonalmatrix ist.* **FOLG**  
Spektralsatz  
für Matrizen

Ein normaler Endomorphismus eines endlich-dimensionalen unitären Vektorraums hat genau dann nur reelle Eigenwerte, wenn er selbst-adjungiert ist (Übung).

## 29. ORTHOGONALE GRUPPEN UND QUATERNIONEN

Wir wollen uns jetzt die (speziellen) orthogonalen und unitären Gruppen  $O(n)$ ,  $SO(n)$ ,  $U(n)$  und  $SU(n)$  in kleinen Dimensionen  $n$  genauer ansehen. Unter anderem werden wir einen Zusammenhang kennenlernen zwischen den Gruppen  $SU(2)$  und  $SO(3)$ , bei dem der Schiefkörper der *Quaternionen* eine Rolle spielt.

Ein ganz trivialer Spezialfall ist  $n = 0$ , dann sind alle Matrixgruppen triviale Gruppen (sie bestehen nur aus dem neutralen Element, das hier die leere Matrix ist).

Im Fall  $n = 1$  haben wir es mit  $1 \times 1$ -Matrizen zu tun, die wir mit Elementen von  $\mathbb{R}$  oder  $\mathbb{C}$  identifizieren können. Transposition ist hier die Identität, ebenso die Determinante, also erhalten wir

$$\begin{aligned} O(1) &= \{\lambda \in \mathbb{R} \mid \lambda^2 = 1\} = \{\pm 1\} \subset \mathbb{R}^\times, \\ SO(1) &= \{\lambda \in \mathbb{R} \mid \lambda = 1\} = \{1\} \subset \mathbb{R}^\times, \\ U(1) &= \{\lambda \in \mathbb{C} \mid |\lambda|^2 = 1\} = S^1 \subset \mathbb{C}^\times, \\ SU(1) &= \{\lambda \in \mathbb{C} \mid \lambda = 1\} = \{1\} \subset \mathbb{C}^\times. \end{aligned}$$

Dabei ist

$$S^1 = \{z \in \mathbb{C} \mid |z| = 1\} \subset \mathbb{C}^\times$$

die *Kreisgruppe*. Für  $z = x + yi \in S^1$  gilt  $1 = |z|^2 = x^2 + y^2$ , also gibt es  $\alpha \in \mathbb{R}$  mit  $x = \cos \alpha$  und  $y = \sin \alpha$ , also  $z = \cos \alpha + i \sin \alpha$ . Dieser Winkel  $\alpha$  ist bis auf Addition eines ganzzahligen Vielfachen von  $2\pi$  eindeutig bestimmt. Den Ausdruck für  $z$  kann man auch so schreiben:

$$e^{i\alpha} = \sum_{n=0}^{\infty} \frac{(i\alpha)^n}{n!} = \sum_{m=0}^{\infty} (-1)^m \frac{\alpha^{2m}}{(2m)!} + i \sum_{m=0}^{\infty} (-1)^m \frac{\alpha^{2m+1}}{(2m+1)!} = \cos \alpha + i \sin \alpha;$$

es gilt

$$(\cos \alpha + i \sin \alpha)(\cos \beta + i \sin \beta) = e^{i\alpha} e^{i\beta} = e^{i(\alpha+\beta)} = \cos(\alpha + \beta) + i \sin(\alpha + \beta)$$

(durch Ausmultiplizieren der linken Seite und Vergleich von Real- und Imaginärteil erhält man die Additionstheoreme für Sinus und Cosinus).

Das führt zur *Polarkoordinatendarstellung* der komplexen Zahlen: Jede komplexe Zahl  $z \in \mathbb{C}$  kann geschrieben werden als  $z = re^{i\alpha}$  mit  $r = |z| \in \mathbb{R}_{\geq 0}$  eindeutig bestimmt; für  $z \neq 0$  ist  $\alpha$  wie oben eindeutig bestimmt bis auf Addition eines ganzzahligen Vielfachen von  $2\pi$ , für  $z = 0$  ist  $\alpha$  beliebig. Diese Darstellung eignet sich besonders gut zum Multiplizieren:

$$re^{i\alpha} \cdot r'e^{i\alpha'} = (rr')e^{i(\alpha+\alpha')}.$$

Wir betrachten als nächstes  $SO(2)$  und  $O(2)$ . Wir erinnern uns daran, dass eine orthogonale Matrix die Eigenschaft hat, dass ihre Spalten (oder Zeilen) eine Orthonormalbasis bilden. Die erste Spalte einer Matrix  $A \in O(2)$  hat also die Form  $(\cos \alpha, \sin \alpha)^\top$ , denn ihre Länge muss 1 sein. Die zweite Spalte muss ebenfalls Länge 1 haben und auf der ersten senkrecht stehen; das lässt genau die beiden Möglichkeiten

$$A_\alpha^+ = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{und} \quad A_\alpha^- = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}.$$

Die erste Matrix hat Determinante 1, ist also in  $SO(2)$ , die zweite hat Determinante  $-1$ , ist also in  $O(2) \setminus SO(2)$ .  $A_\alpha^+$  beschreibt eine Drehung um den Winkel  $\alpha$

gegen den Uhrzeigersinn (denn die beiden Standard-Basisvektoren  $\mathbf{e}_1$  und  $\mathbf{e}_2$  werden auf entsprechend gedrehte Vektoren abgebildet), während  $A_\alpha^-$  eine Spiegelung ist: Das charakteristische Polynom ist

$$\begin{vmatrix} X - \cos \alpha & -\sin \alpha \\ -\sin \alpha & X + \cos \alpha \end{vmatrix} = X^2 - \cos^2 \alpha - \sin^2 \alpha = X^2 - 1 = (X - 1)(X + 1);$$

also gibt es einen Eigenvektor, der fest bleibt (er spannt die Spiegelungsgerade auf) und senkrecht dazu einen, der das Vorzeichen wechselt. Ist  $\alpha = 2\beta$ , dann wird die Gerade, an der gespiegelt wird, erzeugt von  $(\cos \beta, \sin \beta)^\top$ . Daraus, dass  $A_\alpha^+$  eine Drehung um den Winkel  $\alpha$  beschreibt, folgt auch  $A_\alpha^+ A_\beta^+ = A_{\alpha+\beta}^+$  (das kann man mit den Additionstheoremen auch direkt nachrechnen). Es folgt:

29.1. **Satz.** *Die Abbildung*

$$\Phi : \mathrm{U}(1) \longrightarrow \mathrm{SO}(2), \quad e^{i\alpha} \longmapsto A_\alpha^+, \quad \text{bzw.} \quad x + y\mathbf{i} \longmapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

**SATZ**  
 $\mathrm{U}(1) \cong \mathrm{SO}(2)$

*ist ein Gruppenisomorphismus.*

Man definiert *Gruppenhomomorphismen* (analog zu Homomorphismen von Vektorräumen) als Abbildungen, die mit der Gruppenstruktur verträglich sind. Konkret ist ein Gruppenhomomorphismus von einer Gruppe  $(G, 1_G, *_G, i_G)$  in eine weitere Gruppe  $(H, 1_H, *_H, i_H)$  (zu Gruppen siehe Definition 3.6) eine Abbildung  $f : G \rightarrow H$  mit der Eigenschaft

$$f(g *_G g') = f(g) *_H f(g') \quad \text{für alle } g, g' \in G.$$

Es folgt dann (im Wesentlichen genauso wie bei linearen Abbildungen hinsichtlich der additiven Gruppenstruktur)  $f(1_G) = 1_H$  und  $f(i_G(g)) = i_H(f(g))$  für die Inversen. Ein *Gruppenisomorphismus* ist ein bijektiver Gruppenhomomorphismus; in diesem Fall ist die Umkehrabbildung ebenfalls ein Gruppenhomomorphismus.

*Beweis.* Dass beide angegebenen Abbildungsvorschriften dieselbe Abbildung definieren, ergibt sich aus

$$A_\alpha^+ = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{und} \quad e^{i\alpha} = \cos \alpha + \mathbf{i} \sin \alpha.$$

Die zweite Abbildungsvorschrift zeigt, dass  $\Phi$  wohldefiniert ist;  $\Phi$  ist bijektiv, denn auf beiden Seiten ist  $\alpha$  durch das Gruppenelement genau bis auf Addition von ganzzahligen Vielfachen von  $2\pi$  eindeutig bestimmt. Die Abbildung ist auch ein Gruppenhomomorphismus, da gilt

$$\Phi(e^{i\alpha})\Phi(e^{i\beta}) = A_\alpha^+ A_\beta^+ = A_{\alpha+\beta}^+ = \Phi(e^{i(\alpha+\beta)}) = \Phi(e^{i\alpha} \cdot e^{i\beta}). \quad \square$$

Ein weiterer Zugang zu  $\Phi$  geht über die Struktur von  $\mathbb{C}$  als reeller Vektorraum mit der kanonischen Basis  $(1, \mathbf{i})$ . Für  $z = x + y\mathbf{i} \in \mathbb{C}$  ist die Multiplikation mit  $z$  eine  $\mathbb{R}$ -lineare Abbildung  $\mathbb{C} \rightarrow \mathbb{C}$ ,  $w \mapsto zw$  (diese Abbildung ist natürlich tatsächlich sogar  $\mathbb{C}$ -linear). Die Matrix dieses Endomorphismus bezüglich der Basis  $(1, \mathbf{i})$  ist gerade

$$\Phi(z) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

Dies definiert  $\Phi : \mathbb{C} \rightarrow \mathrm{Mat}(2, \mathbb{R})$  mit der Eigenschaft  $\Phi(wz) = \Phi(w)\Phi(z)$ . Der Isomorphismus  $\Phi$  im Satz oben ist dann gerade die Einschränkung auf  $S^1$ .

Es gilt auch  $\Phi(w+z) = \Phi(w) + \Phi(z)$  und  $\Phi(1) = I_2$ :  $\Phi$  ist ein injektiver Ringhomomorphismus, der  $\mathbb{C}$  isomorph auf den Unterring von  $\text{Mat}(2, \mathbb{R})$  abbildet, dessen Elemente alle Matrizen der Form  $\begin{pmatrix} x & -y \\ y & x \end{pmatrix}$  sind.

Da  $\Phi(re^{i\alpha}) = r\Phi(e^{i\alpha})$  ist, sieht man, dass Multiplikation mit  $z = re^{i\alpha}$  eine *Drehstreckung* der komplexen Ebene  $\mathbb{C} \cong \mathbb{R}^2$  bewirkt: eine Drehung um den Winkel  $\alpha$  zusammen mit einer Streckung um den Faktor  $r$ .

Sie erinnern sich vielleicht aus dem Schulunterricht, dass sich aus der Verknüpfung zweier Spiegelungen der Ebene eine Drehung ergibt (um den Schnittpunkt der Spiegelachsen; der Drehwinkel ist das Doppelte des orientierten Winkels zwischen den Achsen). Wir wollen das jetzt präzisieren und verallgemeinern.

**29.2. Definition.** Zwei Matrizen  $A, B \in \text{Mat}(n, \mathbb{R})$  heißen *orthogonal ähnlich*, wenn es  $P \in O(n)$  gibt mit  $B = P^{-1}AP$ . DEF  
orthogonal  
ähnlich

Satz 26.11 kann dann so ausgedrückt werden:

*Jede symmetrische reelle Matrix ist orthogonal ähnlich zu einer Diagonalmatrix.*

Zwei Matrizen sind orthogonal ähnlich genau dann, wenn sie denselben Endomorphismus bezüglich zweier (möglicherweise) verschiedener ONBen beschreiben.

**29.3. Definition.** Sei  $V$  ein euklidischer (oder unitärer) Vektorraum. Eine Zerlegung  $V = \bigoplus_{i \in I} U_i$  als direkte Summe von Untervektorräumen heißt *orthogonal*, wenn die  $U_i$  paarweise orthogonal sind: DEF  
orthogonale  
direkte  
Summe

$$\forall i, j \in I, i \neq j \quad \forall u_i \in U_i, u_j \in U_j : u_i \perp u_j \quad \diamond$$

Haben wir eine orthogonale direkte Summe  $V = U \oplus U'$  und sind  $B$  und  $B'$  Orthonormalbasen von  $U$  und  $U'$ , dann ist  $B \cup B'$  eine Orthonormalbasis von  $V$ . Außerdem gilt  $U' = U^\perp$  und  $U = U'^\perp$ .

**29.4. Lemma.** *Seien  $V$  ein endlich-dimensionaler euklidischer Vektorraum und  $V = U \oplus U'$  eine orthogonale Zerlegung. Ist  $f : V \rightarrow V$  eine Isometrie und  $U$  unter  $f$  invariant, dann ist auch  $U'$  unter  $f$  invariant; insbesondere zerlegt sich  $f$  als  $f = f|_U \oplus f|_{U'}$ .* LEMMA

*Beweis.* Aus  $f(u) \in U$  für alle  $u \in U$  folgt für  $u' \in U'$  und  $u \in U$

$$\langle f(u'), f(u) \rangle = \langle u', u \rangle = 0.$$

Da  $f$  bijektiv ist, gilt  $f(U) = U$  (das folgt aus  $f(U) \subset U$  und  $\dim f(U) = \dim U$ ), damit folgt  $f(u') \in U^\perp = U'$ . □

**29.5. Satz.** *Jede orthogonale Matrix  $A \in O(n)$  ist orthogonal ähnlich zu einer Block-Diagonalmatrix, deren Blöcke die Form  $(1)$ ,  $(-1)$  oder  $A_\varphi^+ \in \text{SO}(2)$  haben. Die Blöcke sind bis auf ihre Reihenfolge und Ersetzen von  $A_\varphi^+$  durch  $A_{-\varphi}^+$  eindeutig bestimmt.* SATZ  
Normalform  
von  
orthogonalen  
Matrizen

*Beweis.* Wir zeigen die äquivalente Aussage, dass jede lineare Isometrie eines  $n$ -dimensionalen euklidischen Vektorraums bezüglich einer geeigneten ONB durch eine Blockmatrix der angegebenen Gestalt beschrieben wird, und zwar durch Induktion über  $n$ . Im Fall  $n = 0$  ist nichts zu zeigen. Sei also  $n > 0$  und  $f \in \text{End}(V)$  eine Isometrie,  $\dim V = n$ . Wir erinnern uns daran, dass alle komplexen Eigenwerte von  $f$  den Betrag 1 haben und dass die nicht-reellen Eigenwerte in konjugiert-komplexen Paaren auftreten.

Hat  $f$  einen reellen Eigenwert  $\lambda$  mit Eigenvektor  $e$ , den wir auf Länge 1 skalieren können, dann zerlegt sich  $f$  als  $f = \lambda \text{id}_U \oplus f|_{U'}$  mit  $U = \langle e \rangle_{\mathbb{R}}$  und  $U' = U^{\perp}$ . Nach Induktionsvoraussetzung gibt es eine ONB  $B$  von  $U'$ , sodass  $f|_{U'}$  bezüglich  $B$  durch eine Matrix der gewünschten Form beschrieben wird. Dann ist  $(e, B)$  eine ONB von  $V$  und die Matrix von  $f$  bezüglich dieser Basis entsteht durch Ergänzen eines  $1 \times 1$ -Blocks der Form  $(\lambda) = (\pm 1)$ ; sie hat damit ebenfalls die gewünschte Form.

Hat  $f$  keinen reellen Eigenwert, dann hat  $f$  ein Paar konjugiert-komplexer Eigenwerte  $\lambda = e^{i\varphi}$ ,  $\bar{\lambda} = e^{-i\varphi}$ . Wir identifizieren für einen Moment  $V$  mit  $\mathbb{R}^n$ . Ist  $e \in \mathbb{C}^n$  ein komplexer Eigenvektor zum Eigenwert  $\bar{\lambda}$ , dann haben wir

$$\begin{aligned} f(\text{Re}(e)) + \mathbf{i}f(\text{Im}(e)) &= f(e) = \bar{\lambda}e = (\cos \varphi - \mathbf{i} \sin \varphi)(\text{Re}(e) + \mathbf{i} \text{Im}(e)) \\ &= (\cos \varphi \cdot \text{Re}(e) + \sin \varphi \cdot \text{Im}(e)) + \mathbf{i}(-\sin \varphi \cdot \text{Re}(e) + \cos \varphi \cdot \text{Im}(e)). \end{aligned}$$

Außerdem sind  $\text{Re}(e)$  und  $\text{Im}(e)$  orthogonal und von gleicher Länge, denn im unitären Vektorraum  $\mathbb{C}^n$  gilt  $e \perp \bar{e}$  (denn  $e$  und  $\bar{e}$  sind Eigenvektoren zu verschiedenen Eigenwerten des normalen Endomorphismus  $f$ , also nach Satz 28.9 bis auf Skalierung Teil einer ONB); daraus folgt

$$\begin{aligned} 0 &= \langle e, \bar{e} \rangle = \langle \text{Re}(e) + \mathbf{i} \text{Im}(e), \text{Re}(e) - \mathbf{i} \text{Im}(e) \rangle \\ &= \langle \text{Re}(e), \text{Re}(e) \rangle + \mathbf{i} \langle \text{Im}(e), \text{Re}(e) \rangle + \mathbf{i} \langle \text{Re}(e), \text{Im}(e) \rangle - \langle \text{Im}(e), \text{Im}(e) \rangle \\ &= \|\text{Re}(e)\|^2 - \|\text{Im}(e)\|^2 + 2\mathbf{i} \langle \text{Re}(e), \text{Im}(e) \rangle \end{aligned}$$

(beachte, dass  $\langle \text{Re}(e), \text{Im}(e) \rangle$  als Skalarprodukt zweier reeller Vektoren reell ist); es folgt  $\|\text{Re}(e)\| = \|\text{Im}(e)\|$  und  $\langle \text{Re}(e), \text{Im}(e) \rangle = 0$ . Bei geeigneter Skalierung bilden also  $\text{Re}(e)$  und  $\text{Im}(e)$  eine ONB eines zweidimensionalen Untervektorraums  $U$  von  $V$  und bezüglich dieser Basis ist  $f|_U$  gegeben durch die Drehmatrix

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} = A_{\varphi}^+ \in \text{SO}(2).$$

Wir zerlegen  $V = U \oplus U^{\perp}$  und wenden wie im ersten Fall die Induktionsvoraussetzung auf  $f|_{U^{\perp}}$  an. In diesem Fall wird die Matrix durch den Block  $A_{\varphi}^+$  ergänzt.

Die Eindeutigkeit folgt aus dem Vergleich der Eigenwerte.  $\square$

Als Spezialfall erhalten wir folgende Aussage:

**29.6. Folgerung.** Sei  $A \in \text{SO}(3)$ . Dann hat  $A$  einen Eigenvektor  $e_1$  zum Eigenwert 1 mit  $\|e_1\| = 1$ . Ist  $(e_1, e_2, e_3)$  eine positiv orientierte Orthonormalbasis mit erstem Element  $e_1$  und  $P \in \text{SO}(3)$  die Matrix mit Spalten  $e_1, e_2, e_3$ , dann ist

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}$$

**FOLG**  
Elemente  
von  $\text{SO}(3)$   
sind  
Drehungen

mit einem  $\varphi \in \mathbb{R}$ . Ist  $A \in O(3) \setminus SO(3)$ , dann hat  $A$  den Eigenwert  $-1$ , und mit einer geeigneten Matrix  $P \in SO(3)$  hat man

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}$$

mit  $\varphi \in \mathbb{R}$ .

*Beweis.* Das folgt aus Satz 29.5 und seinem Beweis. Man beachte, dass für  $\varphi = 0$  man zwei Diagonalblöcke  $(1)$  und für  $\varphi = \pi$  zwei Diagonalblöcke  $(-1)$  erhält.  $\square$

Im Fall  $A \in SO(3)$  beschreibt  $A$  also eine Drehung um die Achse  $\mathbb{R}e_1$  mit dem Drehwinkel  $\alpha$  (gemessen in der von  $e_2$  und  $e_3$  aufgespannten Ebene in der Orientierung von  $e_2$  nach  $e_3$ ). Ist  $A \in O(3) \setminus SO(3)$ , dann kommt zur Drehung noch eine Spiegelung an der Ebene  $\langle e_2, e_3 \rangle_{\mathbb{R}}$  hinzu.

Wir verallgemeinern die Begriffe „Spiegelung“ und „Drehung“ auf höhere Dimensionen.

**29.7. Definition.** Sei  $n \in \mathbb{N}$ . Ein Element  $A \in O(n)$  heißt *Spiegelung*, wenn  $\dim E_{-1}(A) = 1$  und  $\dim E_1(A) = n - 1$  ist.

**DEF**  
Spiegelung  
Drehung

$A$  heißt *Drehung*, wenn  $\dim E_1(A) = n - 2$  und  $A \in SO(n)$  ist.  $\diamond$

Die Definition von „Drehung“ wurde korrigiert, um den Fall einer Drehung um  $\pi$  ( $180^\circ$ ) mit einzuschließen. Anderenfalls stimmt die Aussage von Lemma 29.8 unten nicht, wenn die Eigenräume zum Eigenwert  $-1$  der Spiegelungen orthogonal zueinander sind.

Ist  $A$  eine Spiegelung, dann kann man  $\mathbb{R}^n$  als orthogonale direkte Summe zerlegen in  $\mathbb{R}^n = \langle e_{-1} \rangle_{\mathbb{R}} \oplus E_1(A)$  (eine direkte Summe heißt *orthogonal*, wenn die Summanden paarweise orthogonal sind; Untervektorräume  $U_1$  und  $U_2$  heißen *orthogonal*, wenn  $\forall u_1 \in U_1, u_2 \in U_2 : u_1 \perp u_2$  gilt); dabei sei  $e_{-1}$  ein Eigenvektor der Länge 1 zum Eigenwert  $-1$ . Man kann also jedes  $\mathbf{x} \in \mathbb{R}^n$  schreiben als  $\mathbf{x} = \lambda e_{-1} + \mathbf{y}$  mit  $\mathbf{y} \perp e_{-1}$ , und dann ist  $A\mathbf{x} = -\lambda e_{-1} + \mathbf{y}$ , was genau eine Spiegelung an der Hyperebene  $E_1(A)$  beschreibt. Wir haben oben schon gesehen, dass jedes Element von  $O(2) \setminus SO(2)$  eine Spiegelung ist.

Ist  $A$  eine Drehung, dann hat die Normalform von  $A$  eine Drehmatrix als Block (oder zwei Blöcke  $(-1)$ ) und sonst nur Blöcke  $(1)$ .

**29.8. Lemma.** *Das Produkt zweier verschiedener Spiegelungen ist eine Drehung. Jede Drehung lässt sich als Produkt zweier Spiegelungen schreiben.*

**LEMMA**  
Sp.  $\circ$  Sp.  
= Drehung

*Beweis.* Seien  $A$  und  $B$  zwei verschiedene Spiegelungen in  $O(n)$ . Wegen  $A \neq B$  ist  $\dim(E_1(A) \cap E_1(B)) = n - 2$  (aus der Dimensionsformel für Summen und Durchschnitte folgt, dass diese Dimension  $n - 2$  oder  $n - 1$  sein muss; wäre sie  $n - 1$ , dann wäre  $E_1(A) = E_1(B)$  also auch  $E_{-1}(A) = E_1(A)^\perp = E_1(B)^\perp = E_{-1}(B)$ , was  $A = B$  bedeuten würde). Wir wählen eine ONB von  $\mathbb{R}^n$ , deren letzte  $n - 2$  Elemente eine Basis von  $U = E_1(A) \cap E_1(B)$  bilden. Bezüglich dieser ONB haben die Spiegelungen die Form  $A' \oplus I_{n-2}$  und  $B' \oplus I_{n-2}$  mit  $A', B' \in O(2) \setminus SO(2)$ . Dann ist aber  $A'B' \in SO(2)$  (denn  $\det(A'B') = \det(A') \det(B') = (-1)^2 = 1$ ), und es folgt, dass  $AB$  eine Drehung ist.

Sei jetzt umgekehrt  $A$  eine Drehung. Durch eine geeignete Zerlegung von  $\mathbb{R}^n$  als orthogonale direkte Summe zerlegt sich  $A$  in der Form  $A' \oplus I_{n-2}$  mit  $A' \in \text{SO}(2)$ . Es genügt also, die Behauptung für  $A'$  zu zeigen. Ist  $2\varphi$  der Drehwinkel, dann gilt

$$\begin{aligned} A' &= \begin{pmatrix} \cos 2\varphi & -\sin 2\varphi \\ \sin 2\varphi & \cos 2\varphi \end{pmatrix} = \begin{pmatrix} \cos^2 \varphi - \sin^2 \varphi & -2 \sin \varphi \cos \varphi \\ 2 \sin \varphi \cos \varphi & \cos^2 \varphi - \sin^2 \varphi \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

und alle Matrizen im letzten Produkt sind in  $\text{O}(2)$ , wobei das Produkt der ersten drei und auch die letzte Matrix in  $\text{O}(2) \setminus \text{SO}(2)$ , also Spiegelungen sind.  $\square$

\* **29.9. Folgerung.** *Jedes Element von  $\text{O}(n)$  ist ein Produkt von höchstens  $n$  Spiegelungen. Für Elemente von  $\text{SO}(n)$  ist die Anzahl der Spiegelungen gerade, sonst ungerade.*

**FOLG**  
Spiegelungen erzeugen  $\text{O}(n)$

*Beweis.* Sei  $A \in \text{O}(n)$ . Nach Satz 29.5 gibt es  $P \in \text{O}(n)$  mit

$$P^{-1}AP = \text{diag}(\underbrace{(1), \dots, (1)}_r, \underbrace{(-1), \dots, (-1)}_s, A_{\varphi_1}^+, \dots, A_{\varphi_t}^+)$$

mit  $r + s + 2t = n$ . Es ist klar, dass die rechte Seite geschrieben werden kann als ein Produkt von  $s$  Spiegelungen und  $t$  Drehungen (man ersetze jeweils alle Blöcke bis auf einen durch die entsprechende Einheitsmatrix). Jede Drehung ist ein Produkt von zwei Spiegelungen; insgesamt hat man ein Produkt von  $s + 2t \leq n$  Spiegelungen  $S_1, \dots, S_{s+2t}$ . Dann ist  $A = (PS_1P^{-1})(PS_2P^{-1}) \dots (PS_{s+2t}P^{-1})$  ebenfalls ein Produkt von  $s + 2t$  Spiegelungen.

Die zweite Aussage folgt aus einem Vergleich der Determinanten, denn eine Spiegelung hat Determinante  $-1$ .  $\square$

Um eine Drehung im Raum zu beschreiben, braucht man eine Matrix mit neun reellen Einträgen. Auf der anderen Seite zeigt Folgerung 29.6, dass so eine Drehung durch den Einheitsvektor  $e_1$  (der die Drehachse beschreibt) und den Winkel  $\alpha$  eindeutig beschrieben werden kann. Da  $e_1 \in S^2 = \{\mathbf{x} \in \mathbb{R}^3 \mid \|\mathbf{x}\| = 1\}$  sich in etwas Zweidimensionalem und  $\alpha$  oder äquivalent  $e^{i\alpha} \in S^1$  sich in etwas Eindimensionalem bewegt, gibt es eigentlich nur drei „Freiheitsgrade“ und nicht neun. Wir werden jetzt eine algebraische Struktur einführen, die eine kompaktere Beschreibung erlaubt.

\* **29.10. Definition.** Sei  $\mathbb{H} = \mathbb{R}^4$ , wobei wir die Elemente der Standardbasis mit  $1, \mathbf{i}, \mathbf{j}, \mathbf{k}$  bezeichnen. Dann wird  $\mathbb{H}$  zu einem Schiefkörper durch die Multiplikation, die die Skalarmultiplikation fortsetzt, die Distributivgesetze erfüllt und auf der Basis durch

**DEF**  
Schiefkörper der Quaternionen

$$\begin{aligned} 1\mathbf{i} &= \mathbf{i} = \mathbf{i}1, & 1\mathbf{j} &= \mathbf{j} = \mathbf{j}1, & 1\mathbf{k} &= \mathbf{k} = \mathbf{k}1, & \mathbf{i}^2 &= \mathbf{j}^2 = \mathbf{k}^2 = -1 \\ \mathbf{i}\mathbf{j} &= \mathbf{k}, & \mathbf{j}\mathbf{i} &= -\mathbf{k}, & \mathbf{j}\mathbf{k} &= \mathbf{i}, & \mathbf{k}\mathbf{j} &= -\mathbf{i}, & \mathbf{k}\mathbf{i} &= \mathbf{j}, & \mathbf{i}\mathbf{k} &= -\mathbf{j} \end{aligned}$$

festgelegt ist. Mit dieser Struktur heißt  $\mathbb{H}$  der Schiefkörper der *Quaternionen*, ein Element von  $\mathbb{H}$  heißt eine *Quaternion*.

Ist  $\alpha = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$  mit  $a, b, c, d \in \mathbb{R}$ , dann ist  $\bar{\alpha} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$  die zu  $\alpha$  *konjugierte* Quaternion. Ist  $\bar{\alpha} = \alpha$ , dann heißt  $\alpha$  *reell*; ist  $\bar{\alpha} = -\alpha$ , dann heißt  $\alpha$  eine *reine Quaternion*. Der dreidimensionale Untervektorraum der reinen

Quaternionen wird mit  $\text{Im } \mathbb{H}$  bezeichnet.  $\text{Re}(\alpha) = a \in \mathbb{R}$  heißt der *Skalarteil* von  $\alpha$ ,  $\text{Im}(\alpha) = b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \text{Im } \mathbb{H}$  der *Vektorteil* von  $\alpha$ .  $\diamond$

Die Bezeichnung  $\mathbb{H}$  ehrt Sir William Rowan Hamilton, der die Quaternionen (wieder-)entdeckte, ihnen ihren Namen gab und sie intensiv studierte.

Hier ist natürlich noch Einiges zu zeigen.

- $\mathbb{H}$  ist ein Ring:

Das geht wohl am einfachsten dadurch, dass man die Elemente von  $\mathbb{H}$  mit gewissen komplexen  $2 \times 2$ -Matrizen identifiziert:

$$\Psi : \mathbb{H} \longrightarrow \text{Mat}(2, \mathbb{C}), \quad a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \longmapsto \begin{pmatrix} a + b\mathbf{i} & c + d\mathbf{i} \\ -c + d\mathbf{i} & a - b\mathbf{i} \end{pmatrix};$$

das Bild von  $\Psi$  besteht aus allen Matrizen der Form  $\begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$  mit  $z, w \in \mathbb{C}$ . Es ist klar, dass  $\Psi(\alpha + \beta) = \Psi(\alpha) + \Psi(\beta)$  ist, und man rechnet nach, dass auch  $\Psi(\alpha\beta) = \Psi(\alpha)\Psi(\beta)$  gilt (es genügt, das auf der Basis zu prüfen). Da die Ring-Axiome im Matrizenring gelten, ist auch  $\mathbb{H}$  ein Ring (den wir mit einem Unterring von  $\text{Mat}(2, \mathbb{C})$  identifizieren können).

- $\mathbb{H}$  ist ein Schiefkörper:

Es ist klar, dass die Multiplikation in  $\mathbb{H}$  nicht kommutativ ist. Ist  $\alpha \in \mathbb{H}$  nicht null, dann ist (mit  $\alpha$  wie oben und  $z = a + b\mathbf{i}$ ,  $w = c + d\mathbf{i}$ )

$$\det(\Psi(\alpha)) = z\bar{z} + w\bar{w} = |z|^2 + |w|^2 = a^2 + b^2 + c^2 + d^2 > 0,$$

also ist  $\Psi(\alpha)$  invertierbar, und weil

$$\begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}^{-1} = \frac{1}{|z|^2 + |w|^2} \begin{pmatrix} \bar{z} & -w \\ \bar{w} & z \end{pmatrix}$$

wieder im Bild von  $\Psi$  liegt, ist auch  $\alpha$  invertierbar mit  $\alpha^{-1} = \Psi^{-1}(\Psi(\alpha)^{-1})$ .

Man sieht leicht, dass  $\Psi(\bar{\alpha}) = \Psi(\alpha)^*$  ist; daraus folgt

$$\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta} \quad \text{und} \quad \overline{\alpha\beta} = \bar{\beta} \cdot \bar{\alpha}.$$

Man beachte die Vertauschung der Faktoren! Außerdem gilt für  $\alpha = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$

$$\alpha\bar{\alpha} = \bar{\alpha}\alpha = a^2 + b^2 + c^2 + d^2 = |\alpha|^2;$$

so definieren wir  $|\alpha| \in \mathbb{R}_{\geq 0}$ . Aus den Eigenschaften der Konjugation oder aus  $\det(\Psi(\alpha)) = |\alpha|^2$  folgt  $|\alpha\beta| = |\alpha| \cdot |\beta|$ . Man kann zum Beispiel so argumentieren:

$$|\alpha\beta|^2 = (\alpha\beta)\overline{(\alpha\beta)} = \alpha\beta\bar{\beta}\bar{\alpha} = \alpha|\beta|^2\bar{\alpha} = \alpha\bar{\alpha}|\beta|^2 = |\alpha|^2|\beta|^2.$$

Wenn man für  $\xi = x_1 + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}$  und  $\eta = y_1 + y_2\mathbf{i} + y_3\mathbf{j} + y_4\mathbf{k}$  die Gleichung  $|\xi|^2|\eta|^2 = |\bar{\xi}\eta|^2$  ausschreibt (beachte  $|\xi| = |\bar{\xi}|$ ), erhält man eine Formel, die ein Produkt von Summen von vier Quadraten wieder als Summe von vier Quadraten darstellt:

$$\begin{aligned} & (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) \\ &= (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 - x_3y_4 + x_4y_3)^2 \\ &+ (x_1y_3 + x_2y_4 - x_3y_1 - x_4y_2)^2 + (x_1y_4 - x_2y_3 + x_3y_2 - x_4y_1)^2 \end{aligned}$$

Daraus folgt zum Beispiel, dass das Produkt zweier natürlicher Zahlen, die Summen von vier Quadratzahlen sind, wieder eine Summe von vier Quadratzahlen ist. Dies ist ein wichtiger Schritt im Beweis des *Vier-Quadrate-Satzes* von Lagrange. Der Satz besagt, dass *jede* natürliche Zahl  $n$  Summe von vier Quadratzahlen ist (dabei ist null als Summand erlaubt); die eben gemachte Beobachtung erlaubt es, sich auf den Fall zu beschränken, dass  $n$  eine Primzahl ist.

Die obige Gleichung ist analog zur entsprechenden Gleichung für zwei Quadrate, die man aus der Multiplikativität des komplexen Absolutbetrags erhält:

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2.$$

Wenn wir reine Quaternionen mit Vektoren im  $\mathbb{R}^3$  identifizieren, dann lässt sich das Produkt zweier reiner Quaternionen recht elegant schreiben als

$$\xi \cdot \eta = -\langle \xi, \eta \rangle + \xi \times \eta;$$

der Skalarteil des Produkts ist also bis auf das Vorzeichen das Skalarprodukt und der Vektorteil ist das Vektorprodukt der beiden Vektoren. Aus der Multiplikativität des Betrags folgt dann

$$|\xi|^2|\eta|^2 = \langle \xi, \eta \rangle^2 + |\xi \times \eta|^2 = |\xi|^2|\eta|^2 \cos^2 \angle(\xi, \eta) + |\xi \times \eta|^2$$

und damit  $|\xi \times \eta| = |\xi||\eta| \sin \angle(\xi, \eta)$ .

Aus der Multiplikativität des Absolutbetrags folgt auch, dass

$$S^3 = \{ \alpha \in \mathbb{H} \mid |\alpha| = 1 \}$$

eine (nicht-kommutative) Gruppe unter der Multiplikation von  $\mathbb{H}$  ist. Die Matrizen im Bild von  $\Psi$  haben die Eigenschaft, dass ihre beiden Spalten (oder auch Zeilen) dieselbe Länge haben und zueinander orthogonal sind (bezüglich des unitären Skalarprodukts auf  $\mathbb{C}^2$ ). Die Länge der Spalten von  $\Psi(\alpha)$  ist gerade  $|\alpha|$ . Daraus folgt, dass  $\Psi(S^3) \subset \text{SU}(2)$  ist (denn für  $|\alpha| = 1$  ist  $\Psi(\alpha)$  unitär und die Determinante ist  $\det(\Psi(\alpha)) = |\alpha|^2 = 1$ ). Umgekehrt liegt jedes Element von  $\text{SU}(2)$  im Bild von  $\Psi$  (denn die erste Zeile hat die Form  $(z, w)$  mit  $|z|^2 + |w|^2 = 1$ , dann muss die zweite Zeile die Form  $\lambda(-\bar{w}, \bar{z})$  haben mit  $|\lambda| = 1$ , und da die Determinante dann  $\lambda$  ist, muss  $\lambda = 1$  sein). Es folgt:

29.11. **Satz.** Die Einschränkung von  $\Psi$  liefert einen Gruppenisomorphismus

$$S^3 \longrightarrow \text{SU}(2).$$

**SATZ**  
 $S^3 \cong \text{SU}(2)$

Multiplikation mit einer Quaternion von links oder von rechts ergibt einen Endomorphismus von  $\mathbb{H}$  als reeller Vektorraum. Wir können auch von links *und* rechts mit jeweils einer fest gewählten Quaternion multiplizieren.

29.12. **Lemma.** Sei  $\alpha \in \mathbb{H}$  und  $m_\alpha : \mathbb{H} \rightarrow \mathbb{H}$ ,  $\xi \mapsto \alpha\xi\bar{\alpha}$ . Dann ist  $\text{Im } \mathbb{H}$  ein unter  $m_\alpha$  invarianter reeller Untervektorraum von  $\mathbb{H}$ .

**LEMMA**  
Invarianz  
von  $\text{Im } \mathbb{H}$

Ist  $\alpha \in S^3$ , dann ist die Einschränkung von  $\alpha$  auf  $\text{Im } \mathbb{H}$  eine orientierungserhaltende Isometrie (also eine Drehung), und alle Drehungen von  $\text{Im } \mathbb{H}$  haben diese Form.

*Beweis.* Sei  $\xi \in \text{Im } \mathbb{H}$ , also  $\bar{\xi} = -\xi$ . Dann gilt

$$\overline{m_\alpha(\xi)} = \overline{\alpha\xi\bar{\alpha}} = \bar{\alpha}\bar{\xi}\bar{\alpha} = \alpha(-\xi)\bar{\alpha} = -\alpha\xi\bar{\alpha} = -m_\alpha(\xi),$$

also ist  $m_\alpha(\xi) \in \text{Im } \mathbb{H}$ . Weiter gilt für  $\alpha \in S^3$

$$|m_\alpha(\xi)| = |\alpha\xi\bar{\alpha}| = |\alpha||\xi||\bar{\alpha}| = |\alpha|^2|\xi| = |\xi|,$$

also ist  $m_\alpha$  eine Isometrie.

Alle Links- oder Rechts-Multiplikationen mit festen Quaternionen  $\beta \neq 0$  haben Determinante  $|\beta|^4 > 0$ , sind also orientierungserhaltend. Damit ist  $m_\alpha$  als Automorphismus von  $\mathbb{H}$  orientierungserhaltend. Wegen  $m_\alpha(1) = 1$  (für  $\alpha \in S^3$ ) hat die

Einschränkung von  $m_\alpha$  auf  $\text{Im } \mathbb{H}$  dieselbe Determinante 1 wie  $m_\alpha$ . Also ist auch die Einschränkung auf  $\text{Im } \mathbb{H}$  orientierungserhaltend.

Eine Drehung um die vom Einheitsvektor  $\varepsilon \in \text{Im } \mathbb{H}$  erzeugte Gerade mit dem Winkel  $2\varphi$  bekommt man als  $m_\alpha$  mit  $\alpha = \cos \varphi + \varepsilon \sin \varphi$ : Es gilt dann  $\alpha\varepsilon = \varepsilon\alpha$ , also

$$m_\alpha(\varepsilon) = \alpha\varepsilon\bar{\alpha} = \alpha\varepsilon\alpha^{-1} = \varepsilon\alpha\alpha^{-1} = \varepsilon$$

und für  $\xi \in \text{Im } \mathbb{H}$  mit  $\langle \varepsilon, \xi \rangle = 0$  gilt

$$\varepsilon\xi = -\xi\varepsilon = \varepsilon \times \xi = \xi \text{ um } \pi/2 \text{ um die Achse } \mathbb{R}\varepsilon \text{ gedreht}$$

und  $\varepsilon\xi\varepsilon = \xi$ . Es folgt

$$m_\alpha(\xi) = (\cos^2 \varphi - \sin^2 \varphi)\xi + 2 \cos \varphi \sin \varphi(\varepsilon \times \xi) = \cos(2\varphi)\xi + \sin(2\varphi)(\varepsilon \times \xi),$$

was genau eine Drehung um den Winkel  $2\varphi$  in der zu  $\varepsilon$  senkrechten Ebene in  $\text{Im } \mathbb{H}$  beschreibt.  $\square$

Analog zu linearen Abbildungen definiert man den *Kern* eines Gruppenhomomorphismus  $f : G \rightarrow H$  als  $\ker(f) = \{g \in G \mid f(g) = 1_H\}$ .

\* 29.13. **Satz.** *Die Abbildung*

$$S^3 \longrightarrow \text{SO}(3), \quad \alpha \longmapsto m_\alpha|_{\text{Im } \mathbb{H}}$$

*ist ein surjektiver Gruppenhomomorphismus mit Kern  $\{\pm 1\}$ . Vorschalten von  $\Psi^{-1} : \text{SU}(2) \rightarrow S^3$  liefert demnach einen surjektiven Gruppenhomomorphismus  $\text{SU}(2) \rightarrow \text{SO}(3)$  mit Kern  $\{\pm I_2\}$ .*

**SATZ**  
 $\text{SO}(3) \cong$   
 $\text{SU}(2)/\{\pm I\}$

Ähnlich wie wir für einen Vektorraum  $V$  und einen Untervektorraum  $U$  den Quotientenvektorraum  $V/U$  definiert haben, kann man für eine Gruppe  $G$  und eine Untergruppe  $H$  (die eine zusätzliche Eigenschaft haben muss — sie muss ein sogenannter *Normalteiler* sein) die Quotientengruppe  $G/H$  definieren. Der Kern eines Gruppenhomomorphismus  $\varphi : G \rightarrow G'$  ist stets ein Normalteiler, und man hat wieder einen *Homomorphiesatz*  $G/\ker(\varphi) \cong \text{im}(\varphi)$ .

*Beweis.* Dass die Abbildung wohldefiniert und surjektiv ist, wurde in Lemma 29.12 gezeigt. Dass es sich um einen Gruppenhomomorphismus handelt, folgt aus der Beziehung  $m_{\alpha\beta} = m_\alpha \circ m_\beta$ :

$$m_{\alpha\beta}(\xi) = (\alpha\beta)\xi(\overline{\alpha\beta}) = \alpha\beta\xi\bar{\beta}\bar{\alpha} = \alpha m_\beta(\xi)\bar{\alpha} = m_\alpha(m_\beta(\xi)) = (m_\alpha \circ m_\beta)(\xi).$$

Ist  $\alpha \in S^3$  im Kern, dann gilt  $\alpha\xi\bar{\alpha} = \xi$ , oder äquivalent (wegen  $\bar{\alpha}\alpha = 1$ )  $\alpha\xi = \xi\alpha$  für alle  $\xi \in \text{Im } \mathbb{H}$  und damit auch für alle  $\xi \in \mathbb{H}$ . Schreibt man  $\alpha = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$  und setzt  $\xi = \mathbf{i}, \mathbf{j}$  ein, dann sieht man, dass  $b = c = d = 0$  sein müssen. Es folgt  $\alpha = a = \pm 1$ . Umgekehrt ist klar, dass diese beiden Elemente im Kern liegen.  $\square$

Wenn man es vermeiden möchte, die Quaternionen, mit denen man rechnet, auf Länge 1 zu bringen, dann kann man auch für beliebiges  $\alpha \in \mathbb{H}^\times$  die Abbildung

$$\xi \longmapsto \frac{1}{|\alpha|^2} \alpha\xi\bar{\alpha} = \alpha\xi\alpha^{-1}$$

(denn es ist  $\alpha\bar{\alpha} = |\alpha|^2$ , also  $\bar{\alpha}/|\alpha|^2 = \alpha^{-1}$ ) betrachten. Das ist gleichbedeutend mit  $m_{\alpha/|\alpha|}$ , hat aber den Vorteil, dass man Quadratwurzeln vermeidet. Das ergibt dann einen surjektiven Gruppenhomomorphismus  $\mathbb{H}^\times \rightarrow \text{SO}(3)$  mit Kern  $\mathbb{R}^\times$ .

In jedem Fall sieht man, dass eine Drehung im  $\mathbb{R}^3 \cong \text{Im } \mathbb{H}$  durch eine Quaternion (bis auf reelle Skalierung), also durch ein Quadrupel reeller Zahlen, beschrieben werden kann. Verknüpfung von Drehungen entspricht der Multiplikation von

Quaternionen. Das bedeutet 16 reelle Multiplikationen, während die Multiplikation zweier reeller  $3 \times 3$ -Matrizen 27 reelle Multiplikationen benötigt. (In beiden Fällen lässt sich die Anzahl der Multiplikationen durch geschicktes Umformen auf Kosten von zusätzlichen Additionen verringern; trotzdem bleibt die Version mit Quaternionen vorteilhaft.) Wegen der effizienteren Darstellung und Verknüpfung werden Quaternionen daher in Anwendungen wie zum Beispiel in der Computergrafik eingesetzt.

## 30. ÄUSSERE DIREKTE SUMME UND TENSORPRODUKT

Wir hatten zu Beginn dieses Semesters direkte Summen von Untervektorräumen kennengelernt. Die direkte Summe  $U_1 \oplus U_2$  hat die Eigenschaft, dass sie  $U_1$  und  $U_2$  enthält, von beiden zusammen erzeugt wird und zwischen  $U_1$  und  $U_2$  keine Relationen bestehen, was sich im trivialen Durchschnitt  $U_1 \cap U_2 = \{\mathbf{0}\}$  äußert. Man kann diese Eigenschaften auch etwas anders formulieren:

**30.1. Lemma.** *Seien  $V$  ein  $K$ -Vektorraum und  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen von  $V$ , deren Summe direkt ist, und sei  $U = \bigoplus_{i \in I} U_i$ . Sei  $W$  ein weiterer  $K$ -Vektorraum.*

*Dann gilt: Ist  $(f_i : U_i \rightarrow W)_{i \in I}$  eine Familie linearer Abbildungen, dann gibt es **genau eine** lineare Abbildung  $f : U \rightarrow W$ , sodass  $f|_{U_i} = f_i$  gilt für alle  $i \in I$ .*

**LEMMA**  
Charakterisierung von direkten Summen

*Beweis.* Jedes Element  $u$  von  $U$  kann eindeutig geschrieben werden in der Form

$$u = \sum_{i \in I} u_i$$

mit  $u_i \in U_i$  für alle  $i \in I$  und  $u_i = \mathbf{0}$  für alle bis auf endlich viele  $i \in I$  (die formal möglicherweise unendliche Summe ist dann definiert als die Summe über die endlich vielen Terme  $\neq \mathbf{0}$ ). Das folgt aus der Definition einer direkten Summe. Wenn  $f$  existiert, dann muss gelten

$$f(u) = \sum_{i \in I} f(u_i) = \sum_{i \in I} f_i(u_i),$$

also definieren wir  $f$  in dieser Weise. Damit existiert  $f$  als Abbildung und ist eindeutig bestimmt. Es ist leicht zu sehen, dass  $f$  linear ist, was die Existenz (und Eindeutigkeit) von  $f$  als lineare Abbildung zeigt.  $\square$

Wir wollen jetzt für beliebige Vektorräume  $V_i$  einen neuen Vektorraum  $V$  konstruieren, der sich wie eine direkte Summe der  $V_i$  verhält. Wir können nicht mehr davon ausgehen, dass die  $V_i$  in  $V$  enthalten sind, darum ersetzen wir die Inklusion durch eine lineare Abbildung. Das führt auf die folgende Definition.

**\* 30.2. Definition.** Sei  $(V_i)_{i \in I}$  eine Familie von  $K$ -Vektorräumen. Eine (*äußere*) direkte Summe der  $V_i$  ist ein  $K$ -Vektorraum  $V$  zusammen mit einer Familie von linearen Abbildungen  $(\iota_i : V_i \rightarrow V)_{i \in I}$  mit der folgenden „universellen Eigenschaft“: Zu jedem  $K$ -Vektorraum  $W$  und jeder Familie  $(f_i : V_i \rightarrow W)_{i \in I}$  von linearen Abbildungen gibt es *genau eine* lineare Abbildung  $f : V \rightarrow W$  mit  $f \circ \iota_i = f_i$  für alle  $i \in I$ .  $\diamond$

**DEF**  
(äußere) direkte Summe

Wir nennen diese direkte Summe die „äußere“, um sie von der „inneren“ direkten Summe von Untervektorräumen zu unterscheiden, die sich innerhalb eines festen Vektorraums abspielt.

Wir betrachten den Fall von zwei Vektorräumen  $V_1$  und  $V_2$ . Für jeden Vektorraum  $W$  mit linearen Abbildungen  $\iota_1 : V_1 \rightarrow W$  und  $\iota_2 : V_2 \rightarrow W$  bekommen wir für jeden weiteren Vektorraum  $W$  eine Abbildung

$$\text{Hom}(V, W) \longrightarrow \text{Hom}(V_1, W) \times \text{Hom}(V_2, W), \quad f \longmapsto (f \circ \iota_1, f \circ \iota_2).$$

Die Definition oben lässt sich dann so ausdrücken:  $(V, (\iota_1, \iota_2))$  ist genau dann eine direkte Summe von  $V_1$  und  $V_2$ , wenn diese Abbildung stets bijektiv ist. (In diesem Fall ist sie sogar ein Isomorphismus, denn die Abbildung ist linear, wobei die

Vektorraumstruktur rechts komponentenweise definiert ist.) Entsprechendes gilt für beliebige Familien von Vektorräumen.

**30.3. Beispiel.** Ist  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen eines Vektorraums  $V$ , deren Summe  $U$  direkt ist, dann ist  $U$  zusammen mit den Inklusionsabbildungen  $U_i \hookrightarrow U$  eine äußere direkte Summe der  $U_i$ . Das ist gerade der Inhalt von Lemma 30.1. ♣

**BSP**  
direkte  
Summe

Objekte, die durch eine universelle Eigenschaft definiert sind, erfreuen sich einer sehr starken Eindeutigkeit („eindeutig bis auf eindeutigen Isomorphismus“). Im Fall der direkten Summe sieht das so aus:

**30.4. Satz.** Sei  $(V_i)_{i \in I}$  eine Familie von  $K$ -Vektorräumen und seien  $(V, (\iota_i)_{i \in I})$  und  $(V', (\iota'_i)_{i \in I})$  zwei direkte Summen der  $V_i$ . Dann gibt es einen eindeutig bestimmten Isomorphismus  $\varphi : V \rightarrow V'$  mit  $\iota'_i = \varphi \circ \iota_i$  für alle  $i \in I$ .

**SATZ**  
Eindeutigkeit  
der direkten  
Summe

*Beweis.* Wir wenden die universelle Eigenschaft der direkten Summe  $V$  an auf  $W = V'$  und die Abbildungen  $\iota'_i$ . Das liefert eine eindeutig bestimmte lineare Abbildung  $\varphi : V \rightarrow V'$  mit  $\iota'_i = \varphi \circ \iota_i$  für alle  $i \in I$ . Genauso können wir die universelle Eigenschaft der direkten Summe  $V'$  anwenden auf  $W = V$  und die Abbildungen  $\iota_i$ . Das liefert eine ebenfalls eindeutig bestimmte lineare Abbildung  $\varphi' : V' \rightarrow V$  mit  $\iota_i = \varphi' \circ \iota'_i$  für alle  $i \in I$ . Die Verknüpfung  $f = \varphi' \circ \varphi : V \rightarrow V$  erfüllt

$$f \circ \iota_i = \varphi' \circ (\varphi \circ \iota_i) = \varphi' \circ \iota'_i = \iota_i$$

für alle  $i \in I$ . Dies gilt auch für  $\text{id}_V$ ; wegen der Eindeutigkeit in der universellen Eigenschaft (von  $V$ , mit  $W = V$  und  $f_i = \iota_i$ ) folgt also  $\varphi' \circ \varphi = f = \text{id}_V$ . Dasselbe Argument mit vertauschten Rollen zeigt  $\varphi \circ \varphi' = \text{id}_{V'}$ . Das zeigt, dass  $\varphi$  ein Isomorphismus ist; die Eindeutigkeit hatten wir bereits festgestellt. □

Dieser Satz besagt, dass es nicht darauf ankommt, wie man eine direkte Summe konstruiert, denn alles, was in der einen direkten Summe passiert, hat eine eindeutige Entsprechung in der anderen.

Es bleibt aber die Frage, ob so eine direkte Summe immer existiert.

**\* 30.5. Satz.** Jede Familie  $(V_i)_{i \in I}$  von  $K$ -Vektorräumen hat eine äußere direkte Summe.

**SATZ**  
Existenz  
der direkten  
Summe

*Beweis.* Sei  $V \subset \prod_{i \in I} V_i$  die Teilmenge aller Familien  $(v_i)_{i \in I}$  mit  $v_i = \mathbf{0}$  für alle bis auf endlich viele  $i \in I$ , und sei für  $j \in I$  die Abbildung  $\iota_j : V_j \rightarrow V$  gegeben durch  $v \mapsto (v_i)_{i \in I}$  mit  $v_j = v$  und  $v_i = \mathbf{0} \in V_i$  für  $i \neq j$ . Es ist leicht nachzuprüfen, dass  $V$  mit komponentenweiser Addition und Skalarmultiplikation ein  $K$ -Vektorraum ist und dass die  $\iota_i$  dann lineare Abbildungen sind. Wir müssen noch die universelle Eigenschaft nachweisen. Sei dazu  $W$  ein  $K$ -Vektorraum und sei  $(f_i : V_i \rightarrow W)_{i \in I}$  eine Familie linearer Abbildungen. Wir definieren  $f : V \rightarrow W$  durch

$$f((v_i)_{i \in I}) = \sum_{i \in I} f_i(v_i)$$

(die Summe ist definiert, weil alle bis auf endlich viele Summanden null sind). Dann gilt offenbar  $f \circ \iota_i = f_i$  für alle  $i \in I$ ; es ist auch leicht zu sehen, dass  $f$  linear ist. Auf der anderen Seite gilt für  $v = (v_i)_{i \in I} \in V$ , dass  $v = \sum_{i \in I} \iota_i(v_i)$  ist; das zeigt, dass  $f$  nicht anders definiert werden kann. Damit ist  $f$  auch eindeutig bestimmt. □

Wir schreiben

$$\bigoplus_{i \in I} V_i$$

für die äußere direkte Summe. In der Notation wird nicht zwischen innerer und äußerer direkter Summe unterschieden; was gemeint ist, sollte jeweils aus dem Kontext klar sein. Im Fall  $I = \{1, 2, \dots, n\}$  schreiben wir wie üblich häufig auch  $V_1 \oplus V_2 \oplus \dots \oplus V_n$ .

Nun wollen wir eine ähnliche Konstruktion betrachten, bei der nicht Familien von linearen Abbildungen  $V_i \rightarrow W$  durch eine lineare Abbildung  $V \rightarrow W$  ersetzt werden, sondern wir wollen *bilineare* Abbildungen  $V_1 \times V_2 \rightarrow W$  durch *lineare* Abbildungen  $V \rightarrow W$  ersetzen. Das führt auf folgende Definition:

\* **30.6. Definition.** Seien  $V_1$  und  $V_2$  zwei  $K$ -Vektorräume. Ein  $K$ -Vektorraum  $V$  zusammen mit einer bilinearen Abbildung  $\beta : V_1 \times V_2 \rightarrow V$  heißt ein *Tensorprodukt* von  $V_1$  und  $V_2$ , wenn es für jeden  $K$ -Vektorraum  $W$  und jede bilineare Abbildung  $b : V_1 \times V_2 \rightarrow W$  genau eine lineare Abbildung  $f : V \rightarrow W$  gibt mit  $b = f \circ \beta$ .  $\diamond$

**DEF**  
Tensor-  
produkt

Wir schreiben  $\text{Bil}(V_1, V_2; W)$  für den Vektorraum der bilinearen Abbildungen von  $V_1 \times V_2$  nach  $W$ ; dann besagt die Definition, dass  $(V, \beta)$  genau dann ein Tensorprodukt von  $V_1$  und  $V_2$  ist, wenn die Abbildung

$$\text{Hom}(V, W) \longrightarrow \text{Bil}(V_1, V_2; W), \quad f \longmapsto f \circ \beta$$

für alle  $W$  bijektiv (und damit ein Isomorphismus) ist.

Wie für die direkte Summe gilt, dass Tensorprodukte bis auf eindeutigen Isomorphismus eindeutig bestimmt sind:

**30.7. Satz.** Seien  $V_1$  und  $V_2$  zwei  $K$ -Vektorräume und seien  $(V, \beta)$  und  $(V', \beta')$  zwei Tensorprodukte von  $V_1$  und  $V_2$ . Dann gibt es einen eindeutig bestimmten Isomorphismus  $\varphi : V \rightarrow V'$  mit  $\beta' = \varphi \circ \beta$ .

**SATZ**  
Eindeutigkeit  
des Tensor-  
produkts

*Beweis.* Der Beweis ist völlig analog zum Beweis von Satz 30.4. Wir wenden die universelle Eigenschaft von  $V$  an auf die bilineare Abbildung  $\beta'$ , das liefert eine eindeutig bestimmte lineare Abbildung  $\varphi : V \rightarrow V'$  mit  $\beta' = \varphi \circ \beta$ . Analog gibt es  $\varphi' : V' \rightarrow V$  mit  $\beta = \varphi' \circ \beta'$ . Die Eindeutigkeit in der universellen Eigenschaft liefert  $\varphi' \circ \varphi = \text{id}_V$  und  $\varphi \circ \varphi' = \text{id}_{V'}$ ; damit ist  $\varphi$  ein Isomorphismus.  $\square$

Wir schreiben  $V_1 \otimes V_2$  (oder  $V_1 \otimes_K V_2$ , wenn es auf den Körper ankommt) für ein Tensorprodukt von  $V_1$  und  $V_2$ ; die bilineare Abbildung  $\beta$  wird dann in der Form  $\beta(v_1, v_2) = v_1 \otimes v_2$  notiert.

Was ist eine „universelle Eigenschaft“? Die Definitionen 30.2 und 30.6 haben Folgendes gemeinsam: Wir haben gewisse „Objekte“ (im Fall der direkten Summen sind das Paare  $(W, (f_i)_{i \in I})$  aus einem Vektorraum  $W$  und einer Familie linearer Abbildungen  $f_i : V_i \rightarrow W$ , im Fall des Tensorprodukts sind es Paare  $(W, b)$  aus einem Vektorraum  $W$  und einer bilinearen Abbildung  $b : V_1 \times V_2 \rightarrow W$ ), zwischen denen es „Morphismen“ gibt (bei der direkten Summe ist ein Morphismus  $(W, (f_i)_{i \in I}) \rightarrow (W', (f'_i)_{i \in I})$  eine lineare Abbildung  $\varphi : W \rightarrow W'$  mit  $\varphi \circ f_i = f'_i$  für alle  $i \in I$ , beim Tensorprodukt ist ein Morphismus  $(W, b) \rightarrow (W', b')$  eine lineare Abbildung  $\varphi : W \rightarrow W'$  mit  $\varphi \circ b = b'$ ). Die universelle Eigenschaft besagt dann, dass es von dem betreffenden „universellen“ Objekt genau einen Morphismus zu jedem anderen Objekt gibt (dann hat man ein *initiales* Objekt) oder auch, dass es von jedem Objekt genau einen Morphismus zum universellen Objekt gibt (dann hat man ein *finale*s Objekt). Man kann sehr abstrakt formulieren,

welche Eigenschaften die Objekte und Morphismen haben müssen (sie bilden dann eine sogenannte *Kategorie*); der Teil der Mathematik, der sich damit beschäftigt, heißt *Kategorientheorie* und wird gerne liebevoll als „abstract nonsense“ bezeichnet.

Ein einfaches, aber triviales Beispiel erhalten wir, wenn wir als Objekte  $K$ -Vektorräume und als Morphismen lineare Abbildungen betrachten. Ein universelles Objekt  $U$  hat dann die Eigenschaft, dass es immer *genau eine* lineare Abbildung  $U \rightarrow V$  gibt, für jeden  $K$ -Vektorraum  $V$ . Es ist dann leicht zu sehen, dass hier der Null-Vektorraum ein universelles (initiales) Objekt ist. Er ist übrigens auch ein finales Objekt in dieser Kategorie.

Die Eindeutigkeitsaussage führt auf die Frage nach der Existenz des Tensorprodukts. Zuerst noch eine Definition.

**30.8. Definition.** Sei  $X$  eine Menge und  $K$  ein Körper. Sei

$$K^{(X)} = \{(\lambda_x)_{x \in X} \mid \lambda_x = 0 \text{ für alle bis auf endlich viele } x \in X\}.$$

Dann ist  $K^{(X)}$  ein Untervektorraum von  $K^X$  mit Basis  $(\mathbf{e}_x)_{x \in X}$ , wobei (analog zum Standardvektorraum  $K^n$ )  $\mathbf{e}_x = (\delta_{x,y})_{y \in X}$  die Familie ist, deren Komponenten alle null sind bis auf die  $x$ -te Komponente, die den Wert 1 hat.  $\diamond$

Der Beweis der letzten beiden Aussagen ist eine Übungsaufgabe.

**DEF**  
Vektorraum  
mit gegebener  
Basis

**\* 30.9. Satz.** Seien  $B_1$  und  $B_2$  Basen der  $K$ -Vektorräume  $V_1$  und  $V_2$ . Dann ist  $V = K^{(B_1 \times B_2)}$  zusammen mit

$$\beta : V_1 \times V_2 \longrightarrow V, \quad \left( \sum_{v \in B_1} \lambda_v v, \sum_{v' \in B_2} \mu_{v'} v' \right) \longmapsto (\lambda_v \mu_{v'})_{(v,v') \in B_1 \times B_2}$$

ein Tensorprodukt von  $V_1$  und  $V_2$ . (In den Summen sind alle bis auf endlich viele Koeffizienten  $\lambda_v$  bzw.  $\mu_{v'}$  null.)

**SATZ**  
Existenz  
des Tensor-  
produkts

*Beweis.* Wir schreiben  $\mathbf{e}_{(v,v')}$  wie in Definition 30.8 für die Elemente der „Standard-Basis“ von  $V$ . Sei  $W$  ein weiterer  $K$ -Vektorraum und  $b : V_1 \times V_2 \rightarrow W$  bilinear. Eine lineare Abbildung  $f : V \rightarrow W$  mit  $f \circ \beta = b$  muss dann für alle  $v \in B_1$  und  $v' \in B_2$  die Gleichung

$$f(\mathbf{e}_{(v,v')}) = f(\beta(v, v')) = b(v, v')$$

erfüllen. Es gibt genau eine lineare Abbildung  $f$ , die auf der Standard-Basis von  $V$  diese Werte annimmt (Satz 9.11). Es bleibt zu zeigen, dass für diese lineare Abbildung tatsächlich  $f \circ \beta = b$  gilt: Seien  $v_1 = \sum_{v \in B_1} \lambda_v v \in V_1$  und  $v_2 = \sum_{v' \in B_2} \mu_{v'} v' \in V_2$ . Dann gilt

$$\begin{aligned} f(\beta(v_1, v_2)) &= f((\lambda_v \mu_{v'})_{(v,v') \in B_1 \times B_2}) = f\left( \sum_{(v,v') \in B_1 \times B_2} \lambda_v \mu_{v'} \mathbf{e}_{(v,v')} \right) \\ &= \sum_{(v,v') \in B_1 \times B_2} \lambda_v \mu_{v'} f(\mathbf{e}_{(v,v')}) = \sum_{(v,v') \in B_1 \times B_2} \lambda_v \mu_{v'} b(v, v') \\ &= b\left( \sum_{v \in B_1} \lambda_v v, \sum_{v' \in B_2} \mu_{v'} v' \right) = b(v_1, v_2). \quad \square \end{aligned}$$

Damit ist die Existenz des Tensorprodukts jedenfalls für endlich-dimensionale Vektorräume gezeigt. Da (unter Verwendung des Auswahlaxioms) jeder Vektorraum eine Basis hat, gilt die Existenzaussage auch allgemein.

Es gibt auch eine Basis-freie Konstruktion des Tensorprodukts, die allerdings ziemlich „brutal“ und „verschwenderisch“ anmutet. Wir setzen  $\mathcal{V} = K^{(V_1 \times V_2)}$  (das ist also ein Vektorraum, der für *jedes Element*  $(v_1, v_2) \in V_1 \times V_2$  ein Basiselement  $\mathbf{e}_{(v_1, v_2)}$  hat) und definieren  $\mathcal{U}$  als den Untervektorraum von  $\mathcal{V}$ , der von allen Elementen einer der Formen

$$\begin{aligned} & \mathbf{e}_{(\lambda v_1, v_2)} - \lambda \mathbf{e}_{(v_1, v_2)}, \quad \mathbf{e}_{(v_1, \lambda v_2)} - \lambda \mathbf{e}_{(v_1, v_2)}, \\ & \mathbf{e}_{(v_1 + v'_1, v_2)} - \mathbf{e}_{(v_1, v_2)} - \mathbf{e}_{(v'_1, v_2)}, \quad \mathbf{e}_{(v_1, v_2 + v'_2)} - \mathbf{e}_{(v_1, v_2)} - \mathbf{e}_{(v_1, v'_2)} \end{aligned}$$

mit  $v_1, v'_1 \in V_1$ ,  $v_2, v'_2 \in V_2$  und  $\lambda \in K$  erzeugt wird. Dann setzen wir  $V = \mathcal{V}/\mathcal{U}$  und  $\beta(v_1, v_2) = [\mathbf{e}_{(v_1, v_2)}]$ . Man rechnet nach, dass  $\beta$  bilinear ist (das kommt direkt aus der Definition von  $\mathcal{U}$ ). Ist  $b : V_1 \times V_2 \rightarrow W$  bilinear, dann definiert man zunächst eine lineare Abbildung  $F : \mathcal{V} \rightarrow W$  durch  $F(\mathbf{e}_{(v_1, v_2)}) = b(v_1, v_2)$  (eindeutige Festlegung durch Bild der Basis). Aus der Bilinearität von  $b$  folgt, dass  $\mathcal{U}$  im Kern von  $F$  enthalten ist; es gibt dann (das ist die universelle Eigenschaft des Quotientenraums, siehe Satz 22.10) eine eindeutig bestimmte lineare Abbildung  $f : V \rightarrow W$  mit  $F = f \circ \pi$ , wobei  $\pi : \mathcal{V} \rightarrow V$  der kanonische Epimorphismus ist. Dann gilt  $f \circ \beta = b$ . Die Eindeutigkeit von  $f$  ist auch leicht zu sehen — das Bild von  $\beta$  erzeugt  $V$ , also gibt es höchstens eine lineare Abbildung, die auf dem Bild von  $\beta$  gegebene Werte annimmt.

**30.10. Beispiel.** Seien  $K$  ein Körper und  $m, n \in \mathbb{N}$ . Dann ist der Vektorraum  $\text{Mat}(m \times n, K)$  isomorph zum Tensorprodukt  $K^m \otimes K^n$ . Dabei ist die bilineare Abbildung  $\beta : K^m \times K^n \rightarrow \text{Mat}(m \times n, K)$  gegeben durch  $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \cdot \mathbf{y}^\top$  („Spaltenvektor mal Zeilenvektor“); das Bild von  $\beta$  besteht genau aus der Nullmatrix und den Matrizen vom Rang 1. Daran sieht man sehr schön, dass **keineswegs alle Elemente von  $V_1 \otimes V_2$  die Form  $v_1 \otimes v_2$  haben!** Es gibt ja auch Matrizen von höherem Rang (jedenfalls, wenn  $m$  und  $n$  größer als 1 sind). ♣

**BSP**  
Matrizenraum  
als Tensor-  
produkt

Das Tensorprodukt  $V_1 \otimes V_2$  wird von den Elementen der Form  $v_1 \otimes v_2$  (also dem Bild von  $\beta$ ) erzeugt. (Das folgt aus der Konstruktion in Satz 30.9 oder auch direkt aus der universellen Eigenschaft: Wäre  $U = \langle \text{im}(\beta) \rangle$  nicht ganz  $V_1 \otimes V_2$ , dann könnte man ein Komplement  $U' \neq \{0\}$  von  $U$  wählen und darauf die lineare Abbildung aus der universellen Eigenschaft beliebig definieren, was der Eindeutigkeit widerspräche.) Man kann dann fragen, wie viele solche Elemente man höchstens braucht, um ein beliebiges Element darzustellen.

**30.11. Satz.** Seien  $V$  und  $V'$  zwei  $K$ -Vektorräume und sei  $(b_1, \dots, b_n)$  eine Basis von  $V'$ . Dann lässt sich jedes Element  $w$  von  $V \otimes V'$  eindeutig schreiben als

$$w = v_1 \otimes b_1 + v_2 \otimes b_2 + \dots + v_n \otimes b_n$$

mit  $v_1, v_2, \dots, v_n \in V$ .

**SATZ**  
Elemente  
des Tensor-  
produkts

Man kann das so interpretieren, dass man beim Übergang von  $V'$  zu  $V \otimes V'$  die skalaren Koeffizienten der Basis  $(b_1, \dots, b_n)$  durch „Koeffizienten“ aus  $V$  ersetzt.

*Beweis.* Wir betrachten folgende bilineare Abbildung  $b : V \times V' \rightarrow V^n$ :

$$b\left(v, \sum_{j=1}^n \lambda_j b_j\right) = (\lambda_1 v, \lambda_2 v, \dots, \lambda_n v).$$

Wegen der universellen Eigenschaft gibt es dann eine eindeutig bestimmte lineare Abbildung  $\varphi : V \otimes V' \rightarrow V^n$  mit  $\varphi(v \otimes b_j) = (\delta_{ij} v)_{1 \leq i \leq n}$ . Auf der anderen Seite haben wir die lineare Abbildung

$$\psi : V^n \rightarrow V \otimes V', \quad (v_1, v_2, \dots, v_n) \mapsto v_1 \otimes b_1 + v_2 \otimes b_2 + \dots + v_n \otimes b_n$$

und es gilt offensichtlich  $\varphi \circ \psi = \text{id}_{V^n}$  und  $(\psi \circ \varphi)(v \otimes b_j) = v \otimes b_j$  für alle  $v \in V$  und  $j \in \{1, 2, \dots, n\}$ . Da  $V \otimes V'$  von allen  $v \otimes b_j$  erzeugt wird, folgt

daraus  $\psi \circ \varphi = \text{id}_{V \otimes V'}$ , also ist  $\psi$  ein Isomorphismus, was genau die Behauptung ist. (Dass die  $v \otimes b_j$  Erzeuger von  $V \otimes V'$  sind, kommt daher, dass jedes  $v \otimes v'$  eine Linearkombination dieser spezielleren Elemente ist:  $v' = \sum_j \lambda_j b_j$  impliziert  $v \otimes v' = \sum_j \lambda_j (v \otimes b_j)$ .)  $\square$

**30.12. Beispiel.** Für eine  $m \times n$ -Matrix  $A$  bedeutet das, dass  $A$  Summe von höchstens  $n$  Matrizen vom Rang 1 ist. (Genauer gilt, dass  $A$  Summe von genau  $r = \text{rk}(A)$  Matrizen vom Rang 1 ist.)  $\clubsuit$

**BSP**  
Matrizen

Für direkte Summe und Tensorprodukt gelten Rechenregeln, die denen in einem kommutativen „Halbring“ wie den natürlichen Zahlen ähneln (ein Halbring ist wie ein Ring, nur dass die Existenz von additiven Inversen nicht verlangt wird).

**30.13. Satz.** Seien  $V_1, V_2, V_3$  drei  $K$ -Vektorräume. Dann gibt es kanonische Isomorphismen

**SATZ**  
„Rechen-  
regeln“  
für  $\oplus$  und  $\otimes$

$$\begin{aligned} (V_1 \oplus V_2) \oplus V_3 &\cong V_1 \oplus (V_2 \oplus V_3) \\ V_1 \oplus V_2 &\cong V_2 \oplus V_1 \\ \{0\} \oplus V_1 &\cong V_1 \\ (V_1 \otimes V_2) \otimes V_3 &\cong V_1 \otimes (V_2 \otimes V_3) \\ V_1 \otimes V_2 &\cong V_2 \otimes V_1 \\ K \otimes V_1 &\cong V_1 \\ \{0\} \otimes V_1 &\cong \{0\} \\ (V_1 \oplus V_2) \otimes V_3 &\cong (V_1 \otimes V_3) \oplus (V_2 \otimes V_3) \end{aligned}$$

Wegen der Assoziativität des Tensorprodukts schreibt man auch einfach  $V_1 \otimes V_2 \otimes V_3$  (und analog mit mehr als drei „Faktoren“).

*Beweis.* Wir zeigen hier exemplarisch nur eine der Aussagen; die übrigen Beweise sollten Sie als Übungsaufgaben betrachten. Die Beweis-Struktur ist immer dieselbe: Man konstruiert natürliche lineare Abbildungen in beiden Richtungen und zeigt unter Verwendung der universellen Eigenschaften, dass sie zueinander invers sind.

Wir beweisen das „Assoziativgesetz“  $(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes (V_2 \otimes V_3)$ . Dazu fixieren wir erst einmal  $v_3 \in V_3$ . Die Abbildung

$$V_1 \times V_2 \longrightarrow V_1 \otimes (V_2 \otimes V_3), \quad (v_1, v_2) \longmapsto v_1 \otimes (v_2 \otimes v_3)$$

ist offensichtlich bilinear und führt daher zu einer linearen Abbildung

$$f_{v_3} : V_1 \otimes V_2 \longrightarrow V_1 \otimes (V_2 \otimes V_3) \quad \text{mit} \quad f_{v_3}(v_1 \otimes v_2) = v_1 \otimes (v_2 \otimes v_3).$$

Da  $f_{v_3}$  linear von  $v_3$  abhängt (d.h.,  $V_3 \rightarrow \text{Hom}(V_1 \otimes V_2, V_1 \otimes (V_2 \otimes V_3))$ ,  $v_3 \mapsto f_{v_3}$ , ist linear), ist die Abbildung

$$b : (V_1 \otimes V_2) \times V_3 \longrightarrow V_1 \otimes (V_2 \otimes V_3), \quad b(v, v_3) = f_{v_3}(v)$$

bilinear. Deshalb gibt es eine lineare Abbildung

$$\varphi : (V_1 \otimes V_2) \otimes V_3 \longrightarrow V_1 \otimes (V_2 \otimes V_3) \quad \text{mit} \quad \varphi((v_1 \otimes v_2) \otimes v_3) = v_1 \otimes (v_2 \otimes v_3).$$

Analog gibt es eine lineare Abbildung

$$\psi : V_1 \otimes (V_2 \otimes V_3) \longrightarrow (V_1 \otimes V_2) \otimes V_3 \quad \text{mit} \quad \psi(v_1 \otimes (v_2 \otimes v_3)) = (v_1 \otimes v_2) \otimes v_3.$$

Da die angegebenen Elemente jeweils ein Erzeugendensystem bilden, folgt, dass  $\varphi$  und  $\psi$  zueinander inverse Isomorphismen sind.  $\square$

Die wichtigsten Prinzipien beim Umgang mit Tensorprodukten sind:

- Es gibt genau dann eine (dann auch eindeutig bestimmte) lineare Abbildung  $f : V_1 \otimes V_2 \rightarrow W$  mit  $f(v_1 \otimes v_2) = b(v_1, v_2)$  für alle  $v_1 \in V_1, v_2 \in V_2$ , wenn  $b : V_1 \times V_2 \rightarrow W$  bilinear ist.
- Die Abbildung  $(v_1, v_2) \mapsto v_1 \otimes v_2$  ist selbst bilinear.
- Die Elemente der Form  $v_1 \otimes v_2$  erzeugen  $V_1 \otimes V_2$ , aber im Allgemeinen hat nicht jedes Element von  $V_1 \otimes V_2$  diese Form.

Zum Beispiel ist die Auswertung von linearen Abbildungen

$$\text{Hom}(V, W) \times V \longrightarrow W, \quad (f, v) \longmapsto f(v)$$

bilinear und führt daher zu einer linearen Abbildung

$$\text{Hom}(V, W) \otimes V \longrightarrow W \quad \text{mit} \quad f \otimes v \longmapsto f(v).$$

Auch die Abbildung

$$V^* \times W \longrightarrow \text{Hom}(V, W), \quad (\phi, w) \longmapsto (v \mapsto \phi(v)w)$$

ist bilinear und führt zu einer kanonischen linearen Abbildung

$$V^* \otimes W \longrightarrow \text{Hom}(V, W).$$

**30.14. Satz.** *Seien  $V$  und  $W$  zwei  $K$ -Vektorräume. Ist  $V$  oder  $W$  endlich-dimensional, dann ist die kanonische lineare Abbildung  $\Phi : V^* \otimes W \rightarrow \text{Hom}(V, W)$  ein Isomorphismus.*

**SATZ**  
 $\text{Hom}(V, W)$   
 $\cong V^* \otimes W$

*Beweis.* Sei zunächst  $W$  endlich-dimensional und  $(b_1, \dots, b_m)$  eine Basis von  $W$ . Jede lineare Abbildung  $f : V \rightarrow W$  hat dann die Form

$$f(v) = f_1(v)b_1 + f_2(v)b_2 + \dots + f_m(v)b_m$$

mit eindeutig bestimmten Linearformen  $f_1, f_2, \dots, f_m \in V^*$ . Wir definieren

$$\Psi : \text{Hom}(V, W) \longrightarrow V^* \otimes W, \quad f \longmapsto f_1 \otimes b_1 + f_2 \otimes b_2 + \dots + f_m \otimes b_m;$$

dann gilt  $\Phi \circ \Psi = \text{id}_{\text{Hom}(V, W)}$  (klar) und  $\Psi \circ \Phi = \text{id}_{V^* \otimes W}$ : Sei  $w = \lambda_1 b_1 + \dots + \lambda_m b_m$ , dann ist

$$\begin{aligned} (\Psi \circ \Phi)(\phi \otimes w) &= (\Psi \circ \Phi)(\lambda_1 \phi \otimes b_1 + \dots + \lambda_m \phi \otimes b_m) \\ &= \Psi(v \mapsto \lambda_1 \phi(v)b_1 + \dots + \lambda_m \phi(v)b_m) \\ &= \lambda_1 \phi \otimes b_1 + \dots + \lambda_m \phi \otimes b_m \\ &= \phi \otimes (\lambda_1 b_1 + \dots + \lambda_m b_m) \\ &= \phi \otimes w. \end{aligned}$$

Damit gilt  $\Psi \circ \Phi = \text{id}$  auf einem Erzeugendensystem von  $V^* \otimes W$ , also ist  $\Psi \circ \Phi = \text{id}$ . Also ist  $\Phi$  ein Isomorphismus.

Sei jetzt  $V$  endlich-dimensional mit Basis  $(b_1, \dots, b_n)$  und dualer Basis  $(b_1^*, \dots, b_n^*)$  von  $V^*$ . Eine lineare Abbildung  $f : V \rightarrow W$  ist eindeutig festgelegt durch die beliebig wählbaren Bilder  $f(b_1), \dots, f(b_n)$ . Wir definieren

$$\Psi : \text{Hom}(V, W) \rightarrow V^* \otimes W, \quad f \longmapsto b_1^* \otimes f(b_1) + \dots + b_n^* \otimes f(b_n).$$

Dann sind  $\Phi$  und  $\Psi$  wieder invers zueinander, denn

$$\begin{aligned} (\Psi \circ \Phi)(\phi \otimes w) &= \Psi(v \mapsto \phi(v)w) \\ &= b_1^* \otimes (\phi(b_1)w) + \dots + b_n^* \otimes (\phi(b_n)w) \\ &= \phi(b_1)b_1^* \otimes w + \dots + \phi(b_n)b_n^* \otimes w \\ &= (\phi(b_1)b_1^* + \dots + \phi(b_n)b_n^*) \otimes w \\ &= \phi \otimes w \end{aligned}$$

(also  $\Psi \circ \Phi = \text{id}$  auf einem Erzeugendensystem, damit gilt  $\Psi \circ \Phi = \text{id}$ ) und

$$\begin{aligned} (\Phi \circ \Psi)(f) &= \Phi(b_1^* \otimes f(b_1) + \dots + b_n^* \otimes f(b_n)) \\ &= \Phi(b_1^* \otimes f(b_1)) + \dots + \Phi(b_n^* \otimes f(b_n)) \\ &= (v \mapsto b_1^*(v)f(b_1) + \dots + b_n^*(v)f(b_n)) \\ &= f, \end{aligned}$$

denn die Abbildung in der vorletzten Zeile bildet für alle  $j \in \{1, 2, \dots, n\}$  das Basiselement  $b_j$  auf  $f(b_j)$  ab.  $\square$

Sind  $V$  und  $W$  beide unendlich-dimensional, dann ist  $\Phi$  zwar noch injektiv, aber nicht mehr surjektiv — das Tensorprodukt ist „zu klein“, um alle Homomorphismen zu spezifizieren: Sei  $B$  eine (unendliche) Basis von  $W$ , dann hat jedes Element des Tensorprodukts  $V^* \otimes W$  die Form  $t = \sum_{b \in B'} \phi_b \otimes b$  mit einer endlichen Teilmenge  $B' \subset B$  (das beweist man ähnlich wie in Satz 30.11). Die lineare Abbildung  $\Phi(t)$  bildet  $v \in V$  auf  $\sum_{b \in B'} \phi_b(v)b$  ab, das Bild von  $\Phi(t)$  ist also im endlich-dimensionalen Untervektorraum  $\langle B' \rangle$  von  $W$  enthalten. Ähnlich wie in Satz 30.14 sieht man, dass jede lineare Abbildung mit endlich-dimensionalem Bild (also mit endlichem Rang) im Bild von  $\Phi$  liegt. Es gibt aber stets lineare Abbildungen  $V \rightarrow W$ , deren Bild unendlich-dimensional ist. Im Fall  $V = W$  liegt zum Beispiel  $\text{id}_V$  nicht im Bild von  $\Phi$ .

Im Fall  $V = W$  endlich-dimensional haben wir dann einen Isomorphismus

$$\text{End}(V) = \text{Hom}(V, V) \xrightarrow{\cong} V^* \otimes V$$

(die Umkehrabbildung von  $\Phi$  in Satz 30.14) und wir können folgende Komposition bilden:

$$\text{End}(V) \xrightarrow{\cong} V^* \otimes V = \text{Hom}(V, K) \otimes V \xrightarrow{\text{ev}} K,$$

wobei die letzte Abbildung die von der Auswertung induzierte Abbildung ist. Diese Abbildung  $V^* \otimes V \rightarrow K$  (oder entsprechend  $V \otimes V^* \rightarrow K$ ) heißt auch *Kontraktion*.

**30.15. Satz.** *Die so definierte Abbildung  $\text{End}(V) \rightarrow K$  ist die Spur  $f \mapsto \text{Tr}(f)$ .*

**SATZ**  
Spur  
Basis-frei

*Beweis.* Da die Spur über Matrizen definiert ist, müssen wir zuerst eine Basis  $B = (b_1, \dots, b_n)$  von  $V$  wählen; sei  $(b_1^*, \dots, b_n^*)$  die duale Basis von  $V^*$ . Dann ist für  $f \in \text{End}(V)$

$$\text{Mat}_B(f) = (b_i^*(f(b_j)))_{i,j},$$

also

$$\begin{aligned} \text{Tr}(f) &= \text{Tr}(\text{Mat}_B(f)) \\ &= b_1^*(f(b_1)) + \dots + b_n^*(f(b_n)) \\ &= \text{ev}(b_1^* \otimes f(b_1) + \dots + b_n^* \otimes f(b_n)) \\ &= \text{ev}(\Phi^{-1}(f)) \end{aligned}$$

(vergleiche den Beweis von Satz 30.14 für  $\dim V < \infty$ ). Das ist genau die Behauptung.  $\square$

Wir betrachten jetzt das Zusammenspiel von linearen Abbildungen mit dem Tensorprodukt.

**30.16. Lemma.** Seien  $V, V', W, W'$  vier  $K$ -Vektorräume und seien  $f : V \rightarrow W$  und  $f' : V' \rightarrow W'$  lineare Abbildungen. Dann gibt es genau eine lineare Abbildung  $f \otimes f' : V \otimes V' \rightarrow W \otimes W'$  mit  $(f \otimes f')(v \otimes v') = f(v) \otimes f'(v')$  für alle  $v \in V$  und  $v' \in V'$ .

**LEMMA**  
Tensor-  
produkt  
von Abb.

*Beweis.* Die Abbildung  $b : V \times V' \rightarrow W \otimes W', (v, v') \mapsto f(v) \otimes f'(v')$ , ist bilinear; nach der universellen Eigenschaft von  $V \otimes V'$  existiert also eine eindeutig bestimmte lineare Abbildung  $f \otimes f'$  wie angegeben.  $\square$

Es gilt dann  $\text{id}_V \otimes \text{id}_{V'} = \text{id}_{V \otimes V'}$ , und wenn  $V_1 \xrightarrow{f} V_2 \xrightarrow{g} V_3$  und  $V'_1 \xrightarrow{f'} V'_2 \xrightarrow{g'} V'_3$  lineare Abbildungen sind, dann gilt  $(g \circ f) \otimes (g' \circ f') = (g \otimes g') \circ (f \otimes f')$ , wie man leicht auf Elementen der Form  $v \otimes v'$  nachprüft.

**30.17. Beispiel.** Seien  $V_1, V_2$  und  $V_3$  drei endlich-dimensionale  $K$ -Vektorräume. Wir betrachten

**BSP**  
Komposition  
als  
Kontraktion

$$\begin{aligned} \text{Hom}(V_1, V_2) \times \text{Hom}(V_2, V_3) &\cong (V_1^* \otimes V_2) \times (V_2^* \otimes V_3) \\ &\longrightarrow V_1^* \otimes (V_2 \otimes V_2^*) \otimes V_3 \xrightarrow{\text{id} \otimes \text{ev} \otimes \text{id}} V_1^* \otimes K \otimes V_3 \\ &\cong V_1^* \otimes V_3 \cong \text{Hom}(V_1, V_3). \end{aligned}$$

Diese Abbildung ist dieselbe wie die Komposition  $(g, f) \mapsto f \circ g$ . Es genügt, das für  $g : v_1 \mapsto \phi(v_1)w$  und  $f : v_2 \mapsto \phi'(v_2)w'$  nachzuweisen, wobei  $\phi \in V_1^*, \phi' \in V_2^*$  und  $w \in V_2, w' \in V_3$ . Wir erhalten

$$\begin{aligned} (g, f) &\longmapsto (\phi \otimes w, \phi' \otimes w') \longmapsto \phi \otimes (w \otimes \phi') \otimes w' \\ &\longmapsto \phi \otimes \phi'(w) \otimes w' \longmapsto \phi'(w) \phi \otimes w' \\ &\longmapsto (v_1 \mapsto \phi'(w)\phi(v_1)w' = \phi'(\phi(v_1)w)w' = (f \circ g)(v_1)) \\ &= f \circ g \end{aligned}$$

Im Fall  $V_1 = V_3$  erhält man dann auch sehr leicht die Beziehung  $\text{Tr}(f \circ g) = \text{Tr}(g \circ f)$  (Übung).  $\clubsuit$

**30.18. Beispiel.** Ist  $B = (b_1, \dots, b_n)$  eine Basis von  $V$  und  $B' = (b'_1, \dots, b'_m)$  eine Basis von  $W$ , dann ist

**BSP**  
Kronecker-  
produkt von  
Matrizen

$$B'' = (b_1 \otimes b'_1, \dots, b_1 \otimes b'_m, b_2 \otimes b'_1, \dots, b_2 \otimes b'_m, \dots, b_n \otimes b'_1, \dots, b_n \otimes b'_m)$$

eine Basis von  $V \otimes W$ . Ist  $f$  ein Endomorphismus von  $V$  und  $g$  ein Endomorphismus von  $W$ , dann ist  $f \otimes g$  ein Endomorphismus von  $V \otimes W$ . Die zugehörige Matrix  $A'' = \text{Mat}_{B''}(f \otimes g)$  heißt das *Kronecker-Produkt* von  $A = (a_{ij}) = \text{Mat}_B(f)$  und  $A' = \text{Mat}_{B'}(g)$ . Es gilt dann

$$A'' = \begin{pmatrix} a_{11}A' & a_{12}A' & \cdots & a_{1n}A' \\ a_{21}A' & a_{22}A' & \cdots & a_{2n}A' \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}A' & a_{n2}A' & \cdots & a_{nn}A' \end{pmatrix}.$$

Man schreibt dafür auch  $A'' = A \otimes A'$ .

Es gilt  $\text{Tr}(A \otimes A') = \text{Tr}(A) \text{Tr}(A')$  und  $\det(A \otimes A') = \det(A)^m \det(A')^n$  (Übung).  $\clubsuit$

Der folgende Satz gehört eigentlich in das Kapitel über euklidische Vektorräume.

**Satz.** Sei  $A = (\mathbf{x}_1 | \mathbf{x}_2 | \cdots | \mathbf{x}_n) \in \text{Mat}(n, \mathbb{R})$ . Dann gilt

$$|\det(A)| \leq \|\mathbf{x}_1\| \|\mathbf{x}_2\| \cdots \|\mathbf{x}_n\|$$

mit Gleichheit genau dann, wenn eine Spalte null ist oder die Spalten  $\mathbf{x}_j \in \mathbb{R}^n$  paarweise orthogonal sind.

**SATZ**  
Hadamardsche  
Ungleichung

*Beweis.* Wir beweisen die Aussage durch Induktion. Der Fall  $n = 1$  (und auch der Fall  $n = 0$ ) ist klar. Sei also  $n \geq 2$ . Dass Gleichheit gilt, wenn eine Spalte null ist, ist offensichtlich. Wir können also annehmen, dass  $\mathbf{x}_j \neq 0$  ist für alle  $j \in \{1, 2, \dots, n\}$ . Mit dem Gram-Schmidt-Verfahren konstruieren wir eine ONB  $B$  von  $\mathbb{R}^n$ , deren erstes Element ein skalares Vielfaches von  $\mathbf{x}_1$  ist; sei  $P$  die Matrix, deren Spalten die Vektoren in  $B$  sind. Dann ist

$$P^{-1}A = \left( \begin{array}{c|c} \|\mathbf{x}_1\| & \mathbf{y}^\top \\ \mathbf{0} & A' \end{array} \right)$$

mit  $\mathbf{y} = (y_2, \dots, y_n)^\top \in \mathbb{R}^{n-1}$ . Es gilt (beachte  $P \in \text{O}(n)$ , damit  $\det(P) = \pm 1$ )

$$|\det(A)| = |\det(P^{-1}A)| = \|\mathbf{x}_1\| |\det(A')|.$$

Wir schreiben  $A' = (\mathbf{x}'_2 | \cdots | \mathbf{x}'_n)$ . Da  $P^{-1}$  orthogonal ist, haben die Spalten von  $P^{-1}A$  dieselbe Länge wie die Spalten von  $A$ , also gilt  $\|\mathbf{x}_j\|^2 = y_j^2 + \|\mathbf{x}'_j\|^2$ . Aus der Induktionsvoraussetzung ergibt sich

$$|\det(A')| \leq \|\mathbf{x}'_2\| \cdots \|\mathbf{x}'_n\| \leq \|\mathbf{x}_2\| \cdots \|\mathbf{x}_n\|.$$

Daraus folgt die behauptete Ungleichung. In der zweiten Ungleichung oben gilt Gleichheit genau dann, wenn  $\mathbf{y} = \mathbf{0}$  ist, und das bedeutet gerade, dass  $\mathbf{x}_1 \perp \mathbf{x}_j$  ist für alle  $j \in \{2, \dots, n\}$ . Nach Induktionsvoraussetzung gilt Gleichheit in der ersten Ungleichung genau dann, wenn ein  $\mathbf{x}'_j = \mathbf{0}$  ist oder alle  $\mathbf{x}'_j$  paarweise orthogonal sind. Beide Bedingungen zusammen gelten genau dann, wenn alle  $\mathbf{x}_j$  paarweise orthogonal sind (den Fall  $\mathbf{x}_j = \mathbf{0}$  hatten wir ja ausgeschlossen).  $\square$

Eine  $n \times n$ -Matrix  $A$  mit Einträgen  $\pm 1$ , für die  $|\det(A)| = n^{n/2}$  gilt (das bedeutet gerade, dass die Spalten (oder Zeilen) paarweise orthogonal sind; äquivalent ist also die Bedingung  $A^\top A = nI_n$ ), heißt *Hadamard-Matrix*. Man überlegt sich relativ leicht, dass es solche Matrizen nur für  $n = 1$ ,  $n = 2$  und  $n = 4m$  mit  $m \geq 1$  geben kann. Es ist ein offenes Problem, ob es für *alle* diese  $n$  Hadamard-Matrizen gibt. Da man leicht zeigen kann, dass das Kronecker-Produkt zweier Hadamard-Matrizen wieder eine Hadamard-Matrix ist, folgt jedenfalls, dass die Menge der natürlichen Zahlen  $n$ , für die es eine  $n \times n$ -Hadamard-Matrix gibt, multiplikativ abgeschlossen ist. Zum Beispiel ist

$$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

eine Hadamard-Matrix, also gibt es für jedes  $n = 2^k$  Hadamard-Matrizen. Mit Hilfe von zahlentheoretischen Konstruktionen erhält man Hadamard-Matrizen für

$$n = 4m = q + 1 \quad \text{oder} \quad n = 8m + 4 = 2(q + 1),$$

wenn  $q$  die Potenz einer Primzahl ist (damit bekommt man

$$n = 4, 8, 12, 20, 24, 28, 32, 36, 44, 48, 52, 60, 68, 72, 76, 80, 84, \dots;$$

die weiteren Werte

$$n = 16, 40, 56, 64, 88, \dots$$

bekommt man aus der Multiplikativität; für  $n = 92$  muss man sich schon was anderes überlegen). Die bekannten Konstruktionen decken nicht alle Fälle ab. Anscheinend ist  $n = 668$  der kleinste ungelöste Fall.

31. SYMMETRISCHE UND ALTERNIERENDE POTENZEN

Wir erweitern die Definition von bilinearen Abbildungen auf Abbildungen mit (möglicherweise) mehr als zwei Argumenten.

\* **31.1. Definition.** Seien  $V_1, V_2, \dots, V_n$  und  $W$   $K$ -Vektorräume. Eine Abbildung  $m : V_1 \times V_2 \times \dots \times V_n \rightarrow W$  heißt *(K-)multilinear*, wenn sie in jedem Argument linear ist, d.h., für jedes  $j \in \{1, 2, \dots, n\}$  und für alle  $v_i \in V_i$  ist die Abbildung

**DEF**  
multilineare  
Abbildung  
symmetrisch  
alternierend

$$\begin{array}{ccc} V_j & \longrightarrow & V_1 \times V_2 \times \dots \times V_n & \xrightarrow{m} & W \\ v & \longmapsto & (v_1, \dots, v_{j-1}, v, v_{j+1}, \dots, v_n) & & \end{array}$$

linear. Im Fall  $W = K$  heißt  $m$  eine *Multilinearform*.

Eine multilineare Abbildung  $m : V^n \rightarrow W$  (also mit  $V_1 = V_2 = \dots = V_n = V$ ) heißt *symmetrisch*, wenn für alle  $v_1, \dots, v_n \in V$  und alle Permutationen  $\sigma \in S_n$  gilt

$$m(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = m(v_1, v_2, \dots, v_n).$$

(Es kommt also nicht auf die Reihenfolge der Argumente an.)

Eine multilineare Abbildung  $m : V^n \rightarrow W$  heißt *alternierend*, wenn

$$m(v_1, \dots, v_n) = \mathbf{0}$$

ist, sobald es  $i \neq j$  gibt mit  $v_i = v_j$ . Daraus folgt

$$m(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma)m(v_1, v_2, \dots, v_n)$$

für alle  $\sigma \in S_n$ . ◇

Um Letzteres zu sehen, genügt es eine Transposition  $\sigma$  zu betrachten (denn jede Permutation ist Produkt von Transpositionen und das Vorzeichen  $\varepsilon$  ist multiplikativ). Wenn  $\sigma$  zum Beispiel 1 und 2 vertauscht, dann betrachten wir

$$\begin{aligned} \mathbf{0} &= m(v_1 + v_2, v_1 + v_2, v_3, \dots, v_n) - m(v_1, v_1, v_3, \dots, v_n) - m(v_2, v_2, v_3, \dots, v_n) \\ &= m(v_1, v_2, v_3, \dots, v_n) + m(v_2, v_1, v_3, \dots, v_n) \\ &= m(v_1, v_2, v_3, \dots, v_n) + m(v_{\sigma(1)}, v_{\sigma(2)}, v_{\sigma(3)}, \dots, v_{\sigma(n)}), \end{aligned}$$

woraus mit  $\varepsilon(\sigma) = -1$

$$m(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma)m(v_1, v_2, \dots, v_n)$$

folgt.

**31.2. Beispiele.** Die Abbildung  $K[X]^n \rightarrow K[X], (p_1, \dots, p_n) \mapsto p_1 \cdots p_n$ , ist eine symmetrische multilineare Abbildung.

**BSP**  
multilineare  
Abbildungen

Die Determinante  $(K^n)^n \rightarrow K, (\mathbf{x}_1, \dots, \mathbf{x}_n) \mapsto \det(\mathbf{x}_1, \dots, \mathbf{x}_n)$ , ist eine alternierende Multilinearform.

Das Vektorprodukt  $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \times \mathbf{y}$ , ist eine alternierende bilineare Abbildung. ♣

Analog zum Tensorprodukt von zwei Vektorräumen kann man das Tensorprodukt von  $n$  Vektorräumen definieren.

**31.3. Definition.** Seien  $V_1, V_2, \dots, V_n$   $K$ -Vektorräume. Ein *Tensorprodukt* von  $V_1, V_2, \dots, V_n$  ist ein  $K$ -Vektorraum  $V$  zusammen mit einer multilinearen Abbildung  $\mu : V_1 \times \dots \times V_n \rightarrow V$ , sodass es für jeden  $K$ -Vektorraum  $W$  und jede multilineare Abbildung  $m : V_1 \times \dots \times V_n \rightarrow W$  genau eine lineare Abbildung  $f : V \rightarrow W$  gibt mit  $f \circ \mu = m$ .  $\diamond$

**DEF**  
Tensorprodukt

Wie üblich ist dieses Tensorprodukt eindeutig bis auf eindeutigen Isomorphismus. Die Existenz sieht man wie folgt:

**31.4. Lemma.** Ist  $(V, \mu)$  ein Tensorprodukt von  $V_1, \dots, V_n$ , dann ist  $V \otimes V_{n+1}$  mit der Abbildung  $(v_1, \dots, v_n, v_{n+1}) \mapsto \mu(v_1, \dots, v_n) \otimes v_{n+1}$  ein Tensorprodukt von  $V_1, \dots, V_{n+1}$ .

**LEMMA**  
Existenz  
des allg.  
Tensorprodukts

*Beweis.* Wir müssen die universelle Eigenschaft nachprüfen. Sei dazu  $W$  ein Vektorraum und  $m : V_1 \times \dots \times V_n \times V_{n+1} \rightarrow W$  multilinear. Für einen zunächst fest gewählten Vektor  $v_{n+1} \in V_{n+1}$  ist die Abbildung

$$m_{v_{n+1}} : V_1 \times \dots \times V_n \longrightarrow W, \quad (v_1, \dots, v_n) \longmapsto m(v_1, \dots, v_n, v_{n+1})$$

multilinear, also gibt es (weil  $(V, \mu)$  ein Tensorprodukt von  $V_1, \dots, V_n$  ist) eine eindeutig bestimmte lineare Abbildung  $f_{v_{n+1}} : V \rightarrow W$  mit  $f_{v_{n+1}} \circ \mu = m_{v_{n+1}}$ . Dann ist die Abbildung  $b : V \times V_{n+1}, (v, v_{n+1}) \mapsto f_{v_{n+1}}(v)$ , bilinear, also gibt es eine eindeutig bestimmte lineare Abbildung  $f : V \otimes V_{n+1} \rightarrow W$  mit

$$\begin{aligned} m(v_1, \dots, v_n, v_{n+1}) &= m_{v_{n+1}}(v_1, \dots, v_n) = f_{v_{n+1}}(\mu(v_1, \dots, v_n)) \\ &= b(\mu(v_1, \dots, v_n), v_{n+1}) = f(\mu(v_1, \dots, v_n) \otimes v_{n+1}). \end{aligned}$$

Damit ist die universelle Eigenschaft nachgewiesen.  $\square$

Induktion über die Anzahl  $n$  der zu verarbeitenden Vektorräume zeigt dann, dass immer ein Tensorprodukt existiert.

Man schreibt  $V_1 \otimes V_2 \otimes \dots \otimes V_n$  für „das“ Tensorprodukt von  $V_1, V_2, \dots, V_n$  und  $v_1 \otimes \dots \otimes v_n$  für  $\mu(v_1, \dots, v_n)$ . Das Lemma zeigt, dass diese Schreibweise mit der früher eingeführten (Weglassen von Klammern bei sukzessiven Tensorprodukten von je zwei Vektorräumen) kompatibel ist.

**31.5. Definition.** Ist  $n \geq 1$  und  $V$  ein  $K$ -Vektorraum, dann schreiben wir  $V^{\otimes n}$  für das Tensorprodukt  $\underbrace{V \otimes V \otimes \dots \otimes V}_{n \text{ Faktoren}}$ ; außerdem setzen wir  $V^{\otimes 0} = K$ .  $V^{\otimes n}$  heißt die  $n$ -te *Tensorpotenz* von  $V$ .  $\diamond$

**DEF**  
Tensor-  
potenz

Man kann dann die direkte Summe

$$T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n}$$

betrachten. Darauf erhält man eine natürliche Ringstruktur, indem man die Multiplikation auf den direkten Summanden als

$$(v_1 \otimes v_2 \otimes \dots \otimes v_n) \cdot (w_1 \otimes w_2 \otimes \dots \otimes w_m) = v_1 \otimes \dots \otimes v_n \otimes w_1 \otimes \dots \otimes w_m$$

definiert und dann „bilinear fortsetzt“ (also mit Hilfe der Distributivgesetze auf beliebige Summen von solchen Elementen ausdehnt). Diese Multiplikation ist dann auch  $K$ -bilinear. Allgemein nennt man einen Ring  $R$ , der gleichzeitig (mit derselben Addition) ein  $K$ -Vektorraum ist und dessen Multiplikation  $K$ -bilinear ist, eine  $K$ -Algebra. Der Ring  $T(V)$  ist also eine  $K$ -Algebra und heißt die *Tensoralgebra* von  $V$ . Falls  $\dim V \geq 2$  ist, dann ist  $T(V)$  nicht kommutativ (seien  $v, w \in V$  linear unabhängig, dann ist

$v \otimes w \neq w \otimes v$ ). Für  $V = \langle x \rangle$ , also  $\dim V = 1$ , ist  $T(V)$  isomorph zum Polynomring  $K[X]$ , denn  $V^{\otimes n} = K(x \otimes \cdots \otimes x) = Kx^n$ , die Elemente von  $T(V)$  sind also Polynome in  $x$ , und die Multiplikation ist durch  $x^n \cdot x^m = x^{n+m}$  gegeben. (Im Fall  $V = \{0\}$  ist  $T(V) = K$ , da  $V^{\otimes 0} = K$  und  $V^{\otimes n} = \{0\}$  ist für  $n \geq 1$ .)

Die Tensoralgebra ist auch durch eine universelle Eigenschaft charakterisiert. Ein *Homomorphismus von  $K$ -Algebren* zwischen  $K$ -Algebren  $A$  und  $A'$  ist eine Abbildung  $f : A \rightarrow A'$  mit  $f(a_1 + a_2) = f(a_1) + f(a_2)$ ,  $f(a_1 \cdot a_2) = f(a_1) \cdot f(a_2)$ ,  $f(1) = 1$  und  $f(\lambda a) = \lambda f(a)$  für alle  $a, a_1, a_2 \in A$  und  $\lambda \in K$  ( $f$  ist also sowohl mit der Ringstruktur als auch mit der Struktur als  $K$ -Vektorraum verträglich). Die universelle Eigenschaft der Tensoralgebra ist dann: Zu jeder  $K$ -Algebra  $A$  und jeder  $K$ -linearen Abbildung  $\phi : V \rightarrow A$  gibt es genau einen Homomorphismus von  $K$ -Algebren  $f : T(V) \rightarrow A$  mit  $\phi = f \circ \iota$ . Dabei ist  $\iota : V \rightarrow T(V)$  die Inklusion des direkten Summanden  $V = V^{\otimes 1}$  in  $T(V)$ .

Wir interessieren uns nun dafür, durch einen geeigneten Vektorraum die symmetrischen bzw. alternierenden multilinearen Abbildungen  $V^n \rightarrow W$  zu klassifizieren, analog dazu, wie das Tensorprodukt beliebige multilineare Abbildungen klassifiziert. Wegen der Eindeutigkeit von universellen Objekten verwenden wir im Folgenden den bestimmten Artikel („die“ statt „eine“).

**31.6. Definition.** Sei  $V$  ein  $K$ -Vektorraum und sei  $n \in \mathbb{N}$ . Die  *$n$ -te symmetrische Potenz* von  $V$  ist ein  $K$ -Vektorraum  $S^n V$  zusammen mit einer symmetrischen multilinearen Abbildung  $\sigma : V^n \rightarrow S^n V$  (oft  $\sigma(v_1, \dots, v_n) = v_1 \cdot v_2 \cdots v_n$  geschrieben), sodass die folgende universelle Eigenschaft gilt: Zu jedem  $K$ -Vektorraum  $W$  und jeder symmetrischen multilinearen Abbildung  $s : V^n \rightarrow W$  gibt es genau eine lineare Abbildung  $f : S^n V \rightarrow W$  mit  $f \circ \sigma = s$ .  $\diamond$

**DEF**  
symmetrische  
Potenz

**31.7. Definition.** Sei  $V$  ein  $K$ -Vektorraum und sei  $n \in \mathbb{N}$ . Die  *$n$ -te alternierende Potenz* (auch „äußere Potenz“) von  $V$  ist ein  $K$ -Vektorraum  $\bigwedge^n V$  zusammen mit einer alternierenden multilinearen Abbildung  $\alpha : V^n \rightarrow \bigwedge^n V$  (die meist  $\alpha(v_1, \dots, v_n) = v_1 \wedge v_2 \wedge \cdots \wedge v_n$  geschrieben wird), sodass die folgende universelle Eigenschaft gilt: Zu jedem  $K$ -Vektorraum  $W$  und jeder alternierenden multilinearen Abbildung  $a : V^n \rightarrow W$  gibt es genau eine lineare Abbildung  $f : \bigwedge^n V \rightarrow W$  mit  $f \circ \alpha = a$ .  $\diamond$

**DEF**  
alternierende  
Potenz

Da es um spezielle multilineare Abbildungen geht, sollten sich  $S^n V$  und  $\bigwedge^n V$  irgendwie aus der Tensorpotenz  $V^{\otimes n}$  konstruieren lassen. Das geht wie folgt:

**31.8. Satz.** Sei  $V$  ein  $K$ -Vektorraum und sei  $n \in \mathbb{N}$ . Sei  $U \subset V^{\otimes n}$  der Untervektorraum, der von allen Elementen der Form

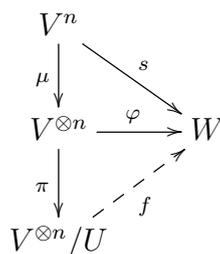
$$v_1 \otimes v_2 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n)}$$

erzeugt wird; dabei sind  $v_1, v_2, \dots, v_n \in V$  und  $\sigma \in S_n$ . Sei  $\pi : V^{\otimes n} \rightarrow V^{\otimes n}/U$  der kanonische Epimorphismus und  $\mu : V^n \rightarrow V^{\otimes n}$  die kanonische multilineare Abbildung. Dann ist  $(S^n V, \sigma) \cong (V^{\otimes n}/U, \pi \circ \mu)$ .

**SATZ**  
Konstruktion  
der symm.  
Potenz

*Beweis.* Wir müssen die universelle Eigenschaft nachprüfen. Sei also  $W$  ein  $K$ -Vektorraum und  $s : V^n \rightarrow W$  multilinear und symmetrisch. Aus der universellen

Eigenschaft von  $V^{\otimes n}$  ergibt sich, dass es eine eindeutig bestimmte lineare Abbildung  $\varphi : V^{\otimes n} \rightarrow W$  gibt mit  $\varphi \circ \mu = s$ .



Da  $s$  symmetrisch ist, gilt

$$\begin{aligned}
 \varphi(v_1 \otimes v_2 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n)}) \\
 = s(v_1, v_2, \dots, v_n) - s(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \mathbf{0}
 \end{aligned}$$

für alle Erzeuger von  $U$ , also ist  $U \subset \ker(\varphi)$ . Deshalb gibt es eine (dann auch eindeutig bestimmte) lineare Abbildung  $f : V^{\otimes n}/U \rightarrow W$  mit  $\varphi = f \circ \pi$ , also  $f \circ (\pi \circ \mu) = \varphi \circ \mu = s$ .  $\square$

Für die alternierende Potenz funktioniert das analog.

**31.9. Satz.** Sei  $V$  ein  $K$ -Vektorraum und sei  $n \in \mathbb{N}$ . Sei  $U \subset V^{\otimes n}$  der Untervektorraum, der von allen Elementen der Form

$$v_1 \otimes v_2 \otimes \cdots \otimes v_n \quad \text{mit } v_i = v_j \text{ für zwei Indizes } i \neq j$$

erzeugt wird; dabei sind  $v_1, v_2, \dots, v_n \in V$ . Sei  $\pi : V^{\otimes n} \rightarrow V^{\otimes n}/U$  der kanonische Epimorphismus und  $\mu : V^n \rightarrow V^{\otimes n}$  die kanonische multilineare Abbildung. Dann ist  $(\wedge^n V, \alpha) \cong (V^{\otimes n}/U, \pi \circ \mu)$ .

**SATZ**  
Konstruktion  
der alt.  
Potenz

*Beweis.* Wir müssen die universelle Eigenschaft nachprüfen. Sei also  $W$  ein  $K$ -Vektorraum und  $a : V^n \rightarrow W$  multilinear und alternierend. Aus der universellen Eigenschaft von  $V^{\otimes n}$  ergibt sich, dass es eine eindeutig bestimmte lineare Abbildung  $\varphi : V^{\otimes n} \rightarrow W$  gibt mit  $\varphi \circ \mu = a$ . Da  $a$  alternierend ist, gilt

$$\varphi(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = a(v_1, v_2, \dots, v_n) = \mathbf{0}$$

für alle Erzeuger von  $U$ , also ist  $U \subset \ker(\varphi)$ . Deshalb gibt es eine (dann auch eindeutig bestimmte) lineare Abbildung  $f : V^{\otimes n}/U \rightarrow W$  mit  $\varphi = f \circ \pi$ , also  $f \circ (\pi \circ \mu) = \varphi \circ \mu = a$ .  $\square$

Als nächstes überlegen wir uns, wie eine Basis von  $S^n V$  bzw.  $\wedge^n V$  aussieht, wenn wir eine Basis von  $V$  kennen. Als ersten Schritt leiten wir ein Kriterium her dafür, wann eine Multilinearform symmetrisch bzw. alternierend ist.

**31.10. Lemma.** Sei  $V$  ein  $K$ -Vektorraum und  $\phi : V^n \rightarrow K$  eine Multilinearform; sei weiter  $B = (b_1, \dots, b_m)$  eine Basis von  $V$ .

- (1)  $\phi$  ist symmetrisch genau dann, wenn für alle  $1 \leq i_1 \leq \dots \leq i_n \leq m$  und alle  $\sigma \in S_n$  gilt

$$\phi(b_{i_{\sigma(1)}}, b_{i_{\sigma(2)}}, \dots, b_{i_{\sigma(n)}}) = \phi(b_{i_1}, b_{i_2}, \dots, b_{i_n}).$$

**LEMMA**  
Kriterium  
für symm.  
bzw. alt.

- (2)  $\phi$  ist alternierend genau dann, wenn für alle  $1 \leq i_1 < \dots < i_n \leq m$  und alle  $\sigma \in S_n$  gilt

$$\phi(b_{i_{\sigma(1)}}, b_{i_{\sigma(2)}}, \dots, b_{i_{\sigma(n)}}) = \varepsilon(\sigma)\phi(b_{i_1}, b_{i_2}, \dots, b_{i_n})$$

und außerdem

$$\phi(b_{j_1}, b_{j_2}, \dots, b_{j_n}) = 0$$

ist, wenn zwei der Indizes übereinstimmen.

*Beweis.* Dass die Bedingungen notwendig sind, folgt unmittelbar aus der Definition. Es bleibt zu zeigen, dass sie auch hinreichend sind. Es ist erst einmal klar, dass aus den angegebenen Bedingungen dieselben Aussagen folgen ohne die Voraussetzung, dass die  $i_k$  monoton wachsend sind. Zur Vereinfachung schreiben wir  $(v_1, \dots, v_n)^\sigma$  für  $(v_{\sigma(1)}, \dots, v_{\sigma(n)})$ , wobei  $v_1, \dots, v_n \in V$  und  $\sigma \in S_n$  sind.

- (1) Wir müssen zeigen, dass  $\phi(\underline{v}^\sigma) = \phi(\underline{v})$  gilt für alle  $\underline{v} = (v_1, \dots, v_n) \in V^n$  und alle  $\sigma \in S_n$ . Wir schreiben  $v_j = \sum_{i=1}^m \lambda_{ij} b_i$ , dann ist

$$\phi(\underline{v}) = \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_1,1} \lambda_{i_2,2} \cdots \lambda_{i_n,n} \phi(b_{i_1}, b_{i_2}, \dots, b_{i_n}).$$

Es folgt

$$\begin{aligned} \phi(\underline{v}^\sigma) &= \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_1,\sigma(1)} \lambda_{i_2,\sigma(2)} \cdots \lambda_{i_n,\sigma(n)} \phi(b_{i_1}, b_{i_2}, \dots, b_{i_n}) \\ &\stackrel{(*)}{=} \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_{\sigma(1)},\sigma(1)} \lambda_{i_{\sigma(2)},\sigma(2)} \cdots \lambda_{i_{\sigma(n)},\sigma(n)} \phi(b_{i_{\sigma(1)}}, b_{i_{\sigma(2)}}, \dots, b_{i_{\sigma(n)}}) \\ &\stackrel{(**)}{=} \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_1,1} \lambda_{i_2,2} \cdots \lambda_{i_n,n} \phi(b_{i_{\sigma(1)}}, b_{i_{\sigma(2)}}, \dots, b_{i_{\sigma(n)}}) \\ &= \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_1,1} \lambda_{i_2,2} \cdots \lambda_{i_n,n} \phi(b_{i_1}, b_{i_2}, \dots, b_{i_n}) \\ &= \phi(\underline{v}). \end{aligned}$$

Bei (\*) haben wir die Indizes  $i_1, \dots, i_n$  mittels  $\sigma$  vertauscht, was nur einer Umordnung der Summe entspricht. Bei (\*\*) haben wir die Faktoren  $\lambda_{i_k,k}$  in die „richtige“ Reihenfolge gebracht, was ihr Produkt nicht ändert. Am Schluss haben wir die Voraussetzung verwendet.

- (2) Wir müssen zeigen, dass  $\phi(\underline{v}) = 0$  ist, wenn in  $\underline{v} = (v_1, \dots, v_n)$  zwei Komponenten übereinstimmen. Wir nehmen an, dass  $v_1 = v_2$  ist (der allgemeine Fall geht genauso). Wie oben schreiben wir die  $v_j$  als Linearkombination der Basis; es gilt dann  $\lambda_{i1} = \lambda_{i2} = \lambda_i$ . Sei  $\tau \in S_n$  die Transposition, die 1

und 2 vertauscht; es ist  $\varepsilon(\tau) = -1$ . Damit erhalten wir

$$\begin{aligned} \phi(\underline{v}) &= \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_1} \lambda_{i_2} \lambda_{i_3,3} \cdots \lambda_{i_n,n} \phi(b_{i_1}, b_{i_2}, \dots, b_{i_n}) \\ &= \sum_{i=1}^m \sum_{i_3=1}^m \cdots \sum_{i_n=1}^m \lambda_i^2 \lambda_{i_3,3} \cdots \lambda_{i_n,n} \underbrace{\phi(b_i, b_i, b_{i_3}, \dots, b_{i_n})}_{=0} \\ &\quad + \sum_{1 \leq i_1 < i_2 \leq m} \sum_{i_3=1}^m \cdots \sum_{i_n=1}^m \lambda_{i_1} \lambda_{i_2} \lambda_{i_3,3} \cdots \lambda_{i_n,n} \cdot \\ &\quad \cdot \underbrace{(\phi(b_{i_1}, b_{i_2}, b_{i_3}, \dots, b_{i_n}) + \phi(b_{i_2}, b_{i_1}, b_{i_3}, \dots, b_{i_n}))}_{=0} \\ &= 0. \end{aligned}$$

□

31.11. **Satz.** Sei  $V$  ein Vektorraum mit Basis  $B = (b_1, b_2, \dots, b_m)$ .

- (1)  $S^n B = (b_{i_1} \cdot b_{i_2} \cdots b_{i_n})_{1 \leq i_1 \leq i_2 \leq \dots \leq i_n \leq m}$  ist eine Basis von  $S^n V$ .
- (2)  $\wedge^n B = (b_{i_1} \wedge b_{i_2} \wedge \cdots \wedge b_{i_n})_{1 \leq i_1 < i_2 < \dots < i_n \leq m}$  ist eine Basis von  $\wedge^n V$ .

**SATZ**  
Basen von  
 $S^n V, \wedge^n V$

Insbesondere ist

$$\dim S^n V = \binom{m+n-1}{n} \quad \text{und} \quad \dim \wedge^n V = \binom{m}{n}.$$

*Beweis.* Wie in Satz 30.9 sieht man, dass

$$B^{\otimes n} = (b_{i_1} \otimes \cdots \otimes b_{i_n})_{1 \leq i_1, \dots, i_n \leq m}$$

eine Basis von  $V^{\otimes n}$  ist. Nach Satz 31.8 bzw. Satz 31.9 sind die kanonischen linearen Abbildungen

$$V^{\otimes n} \longrightarrow S^n V \quad \text{bzw.} \quad V^{\otimes n} \longrightarrow \wedge^n V$$

surjektiv. Da sie die Basis  $B^{\otimes n}$  auf  $S^n B$  bzw.  $\pm \wedge^n B \cup \{0\}$  abbilden, bilden  $S^n B$  bzw.  $\wedge^n B$  jedenfalls ein Erzeugendensystem von  $S^n V$  bzw.  $\wedge^n V$ . Es bleibt zu zeigen, dass sie auch linear unabhängig sind. Dafür überlegen wir uns, dass es zu jedem Basiselement  $b$  eine Linearform  $\phi_b$  auf  $S^n V$  bzw.  $\wedge^n V$  gibt, die auf  $b$  den Wert 1 und auf allen anderen Basiselementen den Wert 0 annimmt. Daraus folgt die lineare Unabhängigkeit: Seien nämlich  $\lambda_b$  Skalare mit  $\sum_{b \in S^n B} \lambda_b b = \mathbf{0}$  (bzw.  $\sum_{b \in \wedge^n B} \lambda_b b = \mathbf{0}$ ), dann folgt

$$0 = \phi_{b'}(\mathbf{0}) = \phi_{b'}\left(\sum_b \lambda_b b\right) = \sum_b \lambda_b \phi_{b'}(b) = \lambda_{b'}$$

für alle  $b' \in S^n B$  (bzw.  $b' \in \wedge^n B$ ).

Im Fall  $S^n V$  seien  $1 \leq i_1 \leq \dots \leq i_n \leq m$  und  $b \in S^n B$  das zugehörige Element. Wir definieren eine Linearform  $\phi$  auf  $V^{\otimes n}$ , indem wir die Bilder der Elemente von  $B^{\otimes n}$  festlegen:

$$\phi(b_{j_1} \otimes b_{j_2} \otimes \cdots \otimes b_{j_n}) = 1, \quad \text{falls es } \sigma \in S_n \text{ gibt mit } j_k = i_{\sigma(k)} \text{ für alle } k$$

und  $= 0$  sonst. Dann ist  $(v_1, \dots, v_n) \mapsto \phi(v_1 \otimes \cdots \otimes v_n)$  nach Lemma 31.10 eine symmetrische Multilinearform auf  $V$ , induziert also eine Linearform  $\phi_b$  auf  $S^n V$ , die die gewünschten Eigenschaften hat.

Im Fall  $\bigwedge^n V$  seien  $1 \leq i_1 < \dots < i_n \leq m$  und  $b \in \bigwedge^n B$  das zugehörige Element. Wir definieren wieder eine Linearform  $\phi$  auf  $V^{\otimes n}$ , indem wir die Bilder der Elemente von  $B^{\otimes n}$  festlegen:

$$\phi(b_{j_1} \otimes b_{j_2} \otimes \dots \otimes b_{j_n}) = \varepsilon(\sigma), \quad \text{falls es } \sigma \in S_n \text{ gibt mit } j_k = i_{\sigma(k)} \text{ für alle } k$$

und  $= 0$  sonst. Wieder nach Lemma 31.10 ist  $(v_1, \dots, v_n) \mapsto \phi(v_1 \otimes \dots \otimes v_n)$  eine alternierende Multilinearform auf  $V$ , induziert also eine Linearform  $\phi_b$  auf  $\bigwedge^n V$ , die die gewünschten Eigenschaften hat.

Die Formeln für die Dimensionen ergeben sich daraus, dass die Elemente von  $\bigwedge^n B$  genau den  $n$ -elementigen Teilmengen der Menge  $\{1, 2, \dots, m\}$  entsprechen, und aus der Überlegung, dass die schwach monoton wachsenden Tupel  $(i_1, \dots, i_n)$  mittels der Abbildung

$$(i_1, \dots, i_n) \mapsto \{i_1, i_2 + 1, i_3 + 2, \dots, i_n + n - 1\}$$

genau den  $n$ -elementigen Teilmengen von  $\{1, 2, \dots, m + n - 1\}$  entsprechen.  $\square$

Für  $n > \dim V$  ist also  $\bigwedge^n V = \{0\}$ , und  $\bigwedge^{\dim V} V$  ist eindimensional.

**31.12. Beispiele.** Das Vektorprodukt  $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \times \mathbf{y}$ , ist alternierend und bilinear und induziert deshalb eine lineare Abbildung  $\bigwedge^2 \mathbb{R}^3 \rightarrow \mathbb{R}^3$ . Diese bildet  $\mathbf{e}_1 \wedge \mathbf{e}_2$  auf  $\mathbf{e}_3$ ,  $\mathbf{e}_1 \wedge \mathbf{e}_3$  auf  $-\mathbf{e}_2$  und  $\mathbf{e}_2 \wedge \mathbf{e}_3$  auf  $\mathbf{e}_1$  ab, ist also ein Isomorphismus.

**BSP**  
Vektorprodukt  
Determinante

Die Determinante liefert entsprechend eine lineare Abbildung  $\bigwedge^n(K^n) \rightarrow K$ . Dabei wird der Erzeuger  $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n$  von  $\bigwedge^n(K^n)$  auf  $1 \in K$  abgebildet (denn  $\det(I_n) = 1$ ), also ist diese Abbildung ebenfalls ein Isomorphismus.  $\clubsuit$

Analog zur Tensoralgebra  $T(V)$  kann man die *symmetrische Algebra*

$$S(V) = \bigoplus_{n=0}^{\infty} S^n V$$

und die *alternierende Algebra*

$$\bigwedge(V) = \bigoplus_{n=0}^{\infty} \bigwedge^n V$$

betrachten. Die symmetrische Algebra ist kommutativ; ist  $V$  endlich-dimensional, dann ist  $S(V)$  isomorph zum Polynomring über  $K$  in  $\dim V$  Unbestimmten (die einer Basis von  $V$  entsprechen). Die symmetrische Algebra  $S(V^*)$  des Dualraums hat als Elemente gerade die Polynomfunktionen auf  $V$  (die Auswertung ist gegeben durch

$$(\phi_1 \cdot \phi_2 \cdot \dots \cdot \phi_n)(v) = \phi_1(v)\phi_2(v) \cdot \dots \cdot \phi_n(v).$$

Das liefert eine Möglichkeit, mit solchen Funktionen zu arbeiten, ohne dafür Koordinaten (also eine Basis) einführen zu müssen. Die symmetrische Algebra von  $V$  hat folgende universelle Eigenschaft: Für jede *kommutative*  $K$ -Algebra  $A$  und jede  $K$ -lineare Abbildung  $\phi : V \rightarrow A$  gibt es genau einen Homomorphismus von  $K$ -Algebren  $f : S(V) \rightarrow A$  mit  $\phi = f \circ \iota$  (mit  $\iota : V = S^1 V \hookrightarrow S(V)$ ).

Die alternierende Algebra eines  $n$ -dimensionalen Vektorraums  $V$  ist endlich-dimensional; die Dimension ist

$$\dim \bigwedge(V) = \sum_{k=0}^n \binom{n}{k} = 2^n.$$

Eine wichtige Eigenschaft des alternierenden Produkts ist, dass man damit lineare Unabhängigkeit testen kann.

**31.13. Satz.** *Seien  $V$  ein Vektorraum und  $v_1, \dots, v_n \in V$ . Dann sind  $v_1, \dots, v_n$  linear unabhängig genau dann, wenn  $v_1 \wedge \dots \wedge v_n \neq \mathbf{0}$  in  $\wedge^n V$  ist.*

**SATZ**  
lin. Unabh.  
über  $\wedge^n V$

*Beweis.* Sind  $v_1, \dots, v_n$  linear abhängig, dann ist einer der Vektoren eine Linearkombination der übrigen. Sei etwa  $v_n = \lambda_1 v_1 + \dots + \lambda_{n-1} v_{n-1}$ ; dann ist

$$v_1 \wedge \dots \wedge v_{n-1} \wedge v_n = \sum_{i=1}^{n-1} \lambda_i v_1 \wedge \dots \wedge v_{n-1} \wedge v_i = \mathbf{0},$$

da in jedem Summanden im alternierenden Produkt zwei gleiche Argumente stehen. (Das Argument funktioniert für jede alternierende multilineare Abbildung.)

Sind  $v_1, \dots, v_n$  linear unabhängig und ist  $V$  endlich-dimensional, dann können wir  $(v_1, \dots, v_n)$  zu einer Basis von  $V$  ergänzen; nach Satz 31.11 ist dann  $v_1 \wedge \dots \wedge v_n$  Element einer Basis von  $\wedge^n V$  und damit insbesondere nicht null.

Im allgemeinen Fall zeigt diese Überlegung, dass  $v_1 \wedge \dots \wedge v_n \neq \mathbf{0}$  ist in  $\wedge^n U$  mit  $U = \langle v_1, \dots, v_n \rangle$ . Dann gibt es eine alternierende Multilinearform  $\phi$  auf  $U^n$  mit  $\phi(v_1, \dots, v_n) \neq 0$ . Sei  $U'$  ein Komplement von  $U$  in  $V$  und  $p : V \rightarrow U$  die zugehörige Projektion; dann ist  $m : V^n \xrightarrow{p^n} U^n \xrightarrow{\phi} K$  eine alternierende Multilinearform auf  $V^n$  mit  $m(v_1, \dots, v_n) \neq 0$ . Daraus folgt  $v_1 \wedge \dots \wedge v_n \neq \mathbf{0}$  in  $\wedge^n V$ .  $\square$

Lineare Abbildungen  $V \rightarrow W$  induzieren lineare Abbildungen zwischen den symmetrischen und alternierenden Potenzen.

**31.14. Lemma.** *Seien  $V$  und  $W$  zwei  $K$ -Vektorräume und  $f : V \rightarrow W$  eine lineare Abbildung. Für jedes  $n \in \mathbb{N}$  gibt es dann eindeutig bestimmte lineare Abbildungen  $S^n f : S^n V \rightarrow S^n W$  und  $\wedge^n f : \wedge^n V \rightarrow \wedge^n W$  mit*

**LEMMA**  
 $S^n f, \wedge^n f$

$(S^n f)(v_1 \cdots v_n) = f(v_1) \cdots f(v_n)$  bzw.  $(\wedge^n f)(v_1 \wedge \dots \wedge v_n) = f(v_1) \wedge \dots \wedge f(v_n)$   
für alle  $v_1, \dots, v_n \in V$ .

*Beweis.* Wir beweisen die Aussage über  $\wedge^n f$ , die andere zeigt man analog. Da  $f$  linear ist, ist

$$V^n \longrightarrow \wedge^n W, \quad (v_1, \dots, v_n) \longmapsto f(v_1) \wedge \dots \wedge f(v_n)$$

eine alternierende multilineare Abbildung. Die universelle Eigenschaft von  $\wedge^n V$  liefert dann die gewünschte Abbildung  $\wedge^n f$  und zeigt, dass sie eindeutig bestimmt ist.  $\square$

Wie beim Tensorprodukt gilt dann natürlich auch  $S^n(g \circ f) = (S^n g) \circ (S^n f)$  und  $\wedge^n(g \circ f) = (\wedge^n g) \circ (\wedge^n f)$ .

Als (krönenden?) Abschluss dieses Kapitels zeigen wir, wie man eine Basis-freie Definition der Determinante bekommen kann.

**31.15. Satz.** *Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n < \infty$ . Für  $f \in \text{End}(V)$  gilt dann*

**SATZ**  
Determinante  
Basis-frei

$$\wedge^n f = \det(f) \text{id}_{\wedge^n V}.$$

*Beweis.* Wir erinnern uns daran, dass  $\dim \wedge^n V = 1$  ist. Also ist jeder Endomorphismus von  $\wedge^n V$  durch Multiplikation mit einem Skalar gegeben. Wir müssen zeigen, dass für  $\wedge^n f \in \text{End}(\wedge^n V)$  dieser Skalar  $\lambda(f)$  gerade  $\det(f)$  ist. Sei  $(b_1, \dots, b_n)$  eine Basis von  $V$ . Dann ist  $b_1 \wedge \dots \wedge b_n$  ein von null verschiedenes Element von  $\wedge^n V$  (Satz 31.11), also ist  $\lambda(f)$  durch

$$f(b_1) \wedge \dots \wedge f(b_n) = (\wedge^n f)(b_1 \wedge \dots \wedge b_n) = \lambda(f) b_1 \wedge \dots \wedge b_n.$$

eindeutig festgelegt. Daran sieht man, dass  $\lambda(f)$  linear in jedem  $f(b_j)$  ist und verschwindet, wenn  $f(b_i) = f(b_j)$  ist für  $i \neq j$ . Außerdem ist  $\lambda(\text{id}_V) = 1$ . Wenn man  $f(b_j)$  über die Basisdarstellung mit der  $j$ -ten Spalte von  $\text{Mat}_B(f)$  identifiziert, dann sind das gerade die Eigenschaften, die die Determinante charakterisieren (vergleiche Satz 13.3). Also muss  $\lambda(f) = \det(f)$  sein.  $\square$

Als Folgerung erhalten wir einen ganz schmerzlosen Beweis der Multiplikativität der Determinante.

**31.16. Folgerung.** Sei  $V$  ein endlich-dimensionaler Vektorraum und seien  $f$  und  $g$  Endomorphismen von  $V$ . Dann gilt  $\det(g \circ f) = \det(g) \det(f)$ .

**FOLG**  
det ist  
multiplikativ

*Beweis.* Sei  $\dim V = n$  und  $\mathbf{0} \neq w \in \wedge^n V$ . Es gilt nach Satz 31.15:

$$\begin{aligned} \det(g \circ f)w &= (\wedge^n(g \circ f))(w) = (\wedge^n g)((\wedge^n f)(w)) \\ &= (\wedge^n g)(\det(f)w) = \det(g) \det(f)w. \end{aligned}$$

Aus  $w \neq \mathbf{0}$  folgt  $\det(g \circ f) = \det(g) \det(f)$ .  $\square$

Die universelle Eigenschaft von  $\wedge^n V$  besagt, dass der Raum der alternierenden multilinearen Abbildungen  $V^n \rightarrow K$  kanonisch isomorph zu  $(\wedge^n V)^*$  ist. Im Fall  $V = \mathbb{R}^m$  kann man  $\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_n \in \wedge^n \mathbb{R}^m$  als ein das von  $\mathbf{x}_1, \dots, \mathbf{x}_n$  im  $\mathbb{R}^m$  aufgespannte Parallelotop repräsentierendes Objekt auffassen. Eine Linearform in  $(\wedge^n \mathbb{R}^m)^*$  weist dann jedem solchen (orientierten) Parallelotop eine Zahl zu, die man als zum Beispiel im Fall  $n = 2$  und  $m = 3$  als den darauf entfallenden Durchfluss einer strömenden Flüssigkeit interpretieren kann (ob man den positiv oder negativ zählt, hängt von der Orientierung des Parallelogramms und der Richtung des Flusses ab). Will man zum Beispiel den Gesamtfluss durch ein Flächenstück  $F$  im  $\mathbb{R}^3$  bestimmen, so muss man den Durchfluss durch viele kleine Flächenstücke zusammenzählen (und einen geeigneten Grenzübergang durchführen). An jedem Punkt  $\mathbf{x}$  von  $\mathbb{R}^3$  braucht man ein  $\phi_{\mathbf{x}} \in (\wedge^2 \mathbb{R}^3)^*$ , das sagt, welchen Durchfluss (kleine) Parallelogramme in der Nähe von  $\mathbf{x}$  haben sollen. So eine Zuordnung  $\phi : \mathbf{x} \mapsto \phi_{\mathbf{x}}$  heißt eine *Differentialform* (hier wäre es eine 2-Form ( $n = 2$ ) auf dem  $\mathbb{R}^3$ ). Den Fluss berechnet man, indem man „ $\phi$  über  $F$  integriert“.

Für endlich-dimensionale Vektorräume gilt  $(\wedge^n V)^* \cong \wedge^n(V^*)$  (Übung), sodass man diese beiden Räume identifizieren kann. Die zu  $(\mathbf{e}_1, \dots, \mathbf{e}_m)$  duale Basis von  $(\mathbb{R}^m)^*$  schreibt man in diesem Zusammenhang üblicherweise  $(dx_1, \dots, dx_m)$ . Eine 2-Form  $\phi$  wie oben hat dann etwa die Gestalt

$$\phi_{\mathbf{x}} = f_{12}(\mathbf{x}) dx_1 \wedge dx_2 + f_{13}(\mathbf{x}) dx_1 \wedge dx_3 + f_{23}(\mathbf{x}) dx_2 \wedge dx_3$$

mit Funktionen  $f_{12}, f_{13}, f_{23} : \mathbb{R}^3 \rightarrow \mathbb{R}$ . Man kann den Vektor  $(f_{23}(\mathbf{x}), -f_{13}(\mathbf{x}), f_{12}(\mathbf{x}))$  als Geschwindigkeit mal Massendichte der Strömung an der Stelle  $\mathbf{x}$  interpretieren: Die  $\mathbf{e}_1$ -Komponente der Strömung macht sich nicht auf den parallel dazu ausgerichteten Parallelogrammen  $\mathbf{e}_1 \wedge \mathbf{e}_2$  und  $\mathbf{e}_1 \wedge \mathbf{e}_3$  bemerkbar, sondern nur auf  $\mathbf{e}_2 \wedge \mathbf{e}_3$ ; deswegen entspricht sie  $f_{23}$ . Das andere Vorzeichen bei  $f_{13}$  hat mit der Orientierung zu tun: Sind  $\mathbf{x}_1, \mathbf{x}_2$  und der „Flussvektor“  $\mathbf{y}$  in dieser Reihenfolge positiv orientiert, dann soll der Fluss positiv sein (und zwar soll gerade  $\det(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$  herauskommen).

Eine 3-Form

$$\eta_{\mathbf{x}} = g(\mathbf{x}) dx_1 \wedge dx_2 \wedge dx_3$$

liefert eine Massen- oder Ladungsdichte. Zum Beispiel erhält man (für  $g = 1$ ) als Integral von  $dx_1 \wedge dx_2 \wedge dx_3$  über eine Teilmenge  $A \subset \mathbb{R}^3$  gerade das Volumen von  $A$ . Näheres dazu gibt es in der Vektoranalysis.