Lineare Algebra II

Sommersemester 2017

Universität Bayreuth

Michael Stoll

Inhaltsverzeichnis

Bildschirmversion vom 20. Juni 2017, 12:10 Uhr.

17. Summen von Untervektorräumen, Komplemente, Kodimension

Eines unserer Ziele in diesem Semester 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 (aber auch an anderen Stellen) 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 7.3), 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.

* **17.1. 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

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

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

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

statt $\sum_{i\in I} U_i$. ◊

Die Schreibweise erklärt sich durch die folgende Eigenschaft.

17.2. Lemma. *Sei V ein Vektorraum.* **LEMMA** Elemente der Summe

(1) *Sind $U_1, U_2, \ldots, U_n$ Untervektorräume von V, dann ist*
$U_1 + U_2 + \ldots + U_n = \{u_1 + u_2 + \ldots + u_n \mid u_1 \in U_1, u_2 \in U_2, \ldots, u_n \in U_n\}$.

(2) *Ist $(U_i)_{i\in I}$ eine Familie von Untervektorräumen von V, dann ist*
$$\sum_{i\in I} U_i = \Big\{ \sum_{i\in J} u_i \mid J \subset I \text{ endlich, } u_i \in U_i \text{ für alle } i \in J \Big\} .$$

Beweis. Es ist klar, dass die jeweils rechts stehende Menge in der links stehenden enthalten ist, denn ihre Elemente sind Linearkombinationen von Elementen von $U_1 \cup \ldots \cup U_n$ bzw. $\bigcup_{i\in I} U_i$ (Satz 7.9). Wir zeigen, dass die rechts stehende Menge ein Untervektorraum von V ist. Dann folgt analog zum Beweis von Satz 7.9, 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

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

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

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

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

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

zwei Elemente von U. Wenn wir $u_i = \mathbf{0}$ (bzw. $u'_i = \mathbf{0}$) 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 λ ein Skalar und $u = u_1 + u_2 + \ldots + u_n$ bzw. $u = \sum_{i \in J} u_i$ ein Element von U. Dann ist

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

wieder ein Element von U. ❑

17.3. Beispiele. **BSP** Summen von Unter-VR

- 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

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

(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.

* **17.4. Satz.** *Sei V ein Vektorraum mit Untervektorräumen U_1 und U_2. Dann gilt* **SATZ** Dimension der Summe

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

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 endlich-dimensional sind, etwa $\dim U_1 = n_1$ und $\dim U_2 = n_2$. Sei $m = \dim(U_1 \cap U_2) \le \min\{n_1, n_2\}$. Wir wählen eine Basis $(b_1, b_2, \ldots, b_m)$ von $U_1 \cap U_2$, die wir einerseits zu einer Basis $(b_1, \ldots, b_m, b'_{m+1}, b'_{m+2}, \ldots, b'_{n_1})$ von U_1 und andererseits zu einer Basis $(b_1, \ldots, b_m, b''_{m+1}, b''_{m+2}, \ldots, b''_{n_2})$ von U_2 ergänzen (Basisergänzungssatz mit Folgerung 9.7). Ich behaupte, dass

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

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 λ_i (für $i \in \{1, 2, \ldots, m\}$), λ'_i (für $i \in \{m+1, \ldots, n_1\}$) und λ''_i (für $i \in \{m+1, \ldots, n_2\}$) Skalare mit

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

Wir schreiben diese Gleichung als

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

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

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

sein; daraus folgt dann auch $\lambda_1 = \ldots = \lambda_m = 0$. ❑

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 $U_1 \cap U_2 = \{\mathbf{0}\}$.

17.5. 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 17.4. 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 $u = \mathbf{0}$. ❑

17.6. Definition. Wenn die Aussagen (1) und (2) in Lemma 17.5 gelten, dann heißt die Summe von U_1 und U_2 *direkt*. ♢ **DEF** direkte Summe zweier UVR

Die Eigenschaften (1) und (2) in Lemma 17.5 lassen sich auch so ausdrücken:

$$\forall u_1 \in U_1, u_2 \in U_2 \colon \bigl(u_1 + u_2 = \mathbf{0} \implies u_1 = u_2 = \mathbf{0}\bigr)\,.$$

In dieser Form lässt sich das verallgemeinern.

* **17.7. 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

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

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

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

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

Lemma 17.5 hat dann die folgende Verallgemeinerung.

17.8. 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

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

ist, wobei $-u$ als Element von U_i betrachtet wird. Definition 17.7 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 Basismenge 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 $B_i = \{b_{i1}, b_{i2}, \ldots, b_{im_i}\}$ mit $m_i = \dim U_i$ und seien λ_{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 $j \in \{1, 2, \ldots, m_i\}$, 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 für jedes $i \in I$

$$\begin{aligned} \sum_{j \in I} \dim U_j &\overset{(3)}{=} \dim \sum_{j \in I} U_j = \dim\Big(U_i + \sum_{j \in I \setminus \{i\}} U_j\Big) \\ &\overset{(*)}{\le} \dim U_i + \dim \sum_{j \in I \setminus \{i\}} U_j \le \dim U_i + \sum_{j \in I \setminus \{i\}} \dim U_j = \sum_{j \in I} \dim U_j\,, \end{aligned}$$

also muss bei $(*)$ Gleichheit herrschen. Aus Satz 17.4 folgt dann

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

❑

17.9. **Beispiel.** Hier ist ein Beispiel, das zeigt, dass die vielleicht erst einmal näher liegende Version „$\forall i, j \in I \colon i \neq j \Rightarrow U_i \cap U_j = \{\mathbf{0}\}$“ 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 $\pi_1 \colon V \to U_1$ und $\pi_2 \colon V \to U_2$. 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 π_1 und π_2 surjektiv, denn $\pi_1|_{U_1} = \mathrm{id}_{U_1}$ und $\pi_2|_{U_2} = \mathrm{id}_{U_2}$.

17.10. **Definition.** Die Abbildungen π_1 und π_2 heißen die *Projektionen* von V auf U_1 bzw. U_2 bezüglich der Zerlegung $V = U_1 \oplus U_2$. ♢ **DEF** Projektionen

Wenn wir mit $p_1, p_2 \colon V \to V$ die Abbildungen bezeichnen, die durch $p_i(v) = \pi_i(v)$ gegeben sind (sie unterscheiden sich von π_1 und π_2 nur durch den vergrößerten Wertebereich), dann gilt

$$p_1 \circ p_1 = p_1\,, \quad p_2 \circ p_2 = p_2\,, \quad \text{sowie} \quad p_1 \circ p_2 = p_2 \circ p_1 = \mathbf{0} \quad \text{und} \quad p_1 + p_2 = \mathrm{id}_V\,.$$

Umgekehrt gilt: Ist $p \colon V \to V$ ein *Projektor*, d.h. eine lineare Abbildung mit $p \circ p = p$, dann gilt $V = \mathrm{im}(p) \oplus \ker(p)$, wobei $\ker(p) = \mathrm{im}(\mathrm{id}_V - p)$ ist (Übung). Mit $p' = \mathrm{id}_V - p$ gilt dann auch $p' \circ p' = p'$, $p \circ p' = p' \circ p = \mathbf{0}$ und $p + p' = \mathrm{id}_V$. **DEF** Projektor

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

* 17.11. **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. ♢ **DEF** Komplement

Konkret bedeutet die Bedingung, dass $U \cap U' = \{\mathbf{0}\}$ und $U + U' = V$ gilt.

17.12. **Beispiele.** **BSP** Komplemente

- V ist das einzige Komplement von $\{\mathbf{0}\}$ in V und $\{\mathbf{0}\}$ 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. ♣

Gibt es immer ein Komplement?

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

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

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

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

weil $b_1, \ldots, b_n$ linear unabhängig sind. Die Dimensionsformel folgt aus Lemma 17.5. ❑

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 9.5.

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.

17.14. **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 $\phi = \pi \circ \iota \colon U_1' \to U_2'$; dabei sei $\iota \colon U_1' \to V$ die Inklusionsabbildung und $\pi \colon V \to U_2'$ die Projektion bezüglich der Zerlegung $V = U \oplus U_2'$. Dann gilt

$$\ker(\phi) = \ker(\pi) \cap U_1' = U \cap U_1' = \{\mathbf{0}\},$$

also ist ϕ injektiv. (Die erste Gleichheit folgt daraus, dass ι injektiv ist.) Wir müssen noch zeigen, dass ϕ auch surjektiv ist, dann ist ϕ 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. ❑

Damit ist folgende Definition sinnvoll.

∗ 17.15. **Definition.** Seien V ein Vektorraum und $U \subset V$ ein Untervektorraum, der ein Komplement U' in V hat. Dann heißt **DEF** Kodimension

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

die *Kodimension* von U in V. ♢

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

17.16. **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 $U_a = \{p \in P \mid p(a) = 0\} = \ker \mathrm{ev}_a$. 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 $\operatorname{codim}_P U_a = 1$. 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 15.16) sieht man analog: Sind $a_1, a_2, \ldots, a_n \in \mathbb{R}$ paarweise verschieden und ist $U_{a_1,\ldots,a_n} = U_{a_1} \cap U_{a_2} \cap \ldots \cap U_{a_n}$ der Untervektorraum der Polynomfunktionen, die in $a_1, \ldots, a_n$ verschwinden, dann ist der Untervektorraum $P_{<n}$ der Polynomfunktionen vom Grad $< n$ ein Komplement von $U_{a_1,\ldots,a_n}$ in P, also gilt $\operatorname{codim}_P U_{a_1,\ldots,a_n} = n$. 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 $a_1, \ldots, a_n$ verschwinden, sind von der Form $x \mapsto q(x)(x - a_1)(x - a_2) \cdots (x - a_n)$. ♣

Der Begriff der Kodimension erlaubt eine etwas genauere Formulierung des „Rangsatzes“ 10.17. Zur Erinnerung: Der Satz besagt, dass für eine lineare Abbildung $\phi\colon V \to W$ 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 ϕ endlich ist:

17.17. **Satz.** *Sei $\phi\colon V \to W$ eine lineare Abbildung mit $\operatorname{rk}(\phi) < \infty$. Dann gilt* **SATZ** Rangsatz mit Kodimension

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

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

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

folgt

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

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

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

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 $\lambda_1, \ldots, \lambda_m$ mit

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

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

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

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

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

  woraus $\lambda_1 = \ldots = \lambda_m = 0$ und damit $v = \mathbf{0}$ folgt. ❑

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

17.18. **Lemma.** *Sei $V = U_1 \oplus U_2 \oplus \ldots \oplus U_n$ eine Zerlegung eines Vektorraums V als direkte Summe von Untervektorräumen. Seien weiter* **LEMMA** direkte Summe von Endo-morphismen

$$f_1 \colon U_1 \to U_1, \quad f_2 \colon U_2 \to U_2, \quad \ldots, \quad f_n \colon U_n \to U_n$$

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

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

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

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

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

wobei $\pi_i \colon V \to U_i$ die Projektion auf U_i und $\iota_i \colon U_i \to V$ die Inklusionsabbildung ist. (Alternativ kann man das auch wie vor Definition 17.10 nachrechnen.) ❑

17.19. **Definition.** Die Abbildung f in Lemma 17.18 heißt die *direkte Summe* von $f_1, f_2, \ldots, f_n$; wir schreiben **DEF** direkte Summe von Endo-morphismen

$$f = f_1 \oplus f_2 \oplus \ldots \oplus f_n\,. \qquad \diamondsuit$$

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:

17.20. **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. $\diamondsuit$ **DEF** invarianter Untervektor-raum

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

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 $i \in \{1, 2, \ldots, n\}$. Dann können wir $f_i \in \operatorname{End}(U_i)$ definieren durch $f_i(u) = f(u)$ für $u \in U_i$; damit gilt $f = f_1 \oplus f_2 \oplus \ldots \oplus f_n$. ❑

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

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

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

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

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

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

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

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

$$\mathrm{Mat}_{B,B}(f) = \left(\begin{array}{c|c|c|c} M_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & M_2 & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & M_n \end{array}\right)$$

eine Block-Diagonalmatrix, wobei $M_i = \mathrm{Mat}_{B_i,B_i}(f_i)$ 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 $\mathrm{Mat}_{B,B}(f)$ 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. ❑

18. Äquivalenzrelationen, Quotientenräume und affine Unterräume

Wir erinnern uns daran, dass der Kern jeder linearen Abbildung $f\colon V \to V'$ ein Untervektorraum von V ist. In diesem Abschnitt wollen wir 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 17.13), es ist also $V = U \oplus U'$. Die zu dieser Zerlegung gehörende Projektion $\pi\colon V \to U'$ 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 $f\colon V \to V'$ gibt mit $\ker(f) = U$, dann gibt es auch eine surjektive solche Abbildung, denn wir können einfach die im Wertebereich eingeschränkte Abbildung $f\colon V \to \operatorname{im}(f)$ betrachten. Wir nehmen jetzt also an, dass wir so eine surjektive lineare Abbildung $f\colon V \to V'$ 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 λv_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 $f\colon V \to V'$, $v \mapsto [v]$, 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*.

18.1. **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*. ♢

18.2. **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 $\{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \le y\}$ und analog für $<$, $\ge$, $>$.

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

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

* 18.3. **Definition.** Seien X eine Menge und R eine Relation auf X. Dann heißt R eine *Äquivalenzrelation*, wenn folgende Bedingungen erfüllt sind:

DEF Äquivalenzrelation

(1) $\forall x \in X\colon x\,R\,x$ (Reflexivität).

(2) $\forall x, y \in X\colon (x\,R\,y \implies y\,R\,x)$ (Symmetrie).

(3) $\forall x, y, z \in X\colon (x\,R\,y \wedge y\,R\,z \implies x\,R\,z)$ (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 $a \le b$ folgt nicht unbedingt $b \le a$ und aus $a \mid b$ folgt nicht unbedingt $b \mid a$.

18.4. **Beispiele.** Beispiele für Äquivalenzrelationen, die wir schon kennen gelernt haben, sind die „Äquivalenz" und die „Ähnlichkeit" von Matrizen.

BSP Äquivalenzrelationen

Zur Erinnerung: Zwei Matrizen $A, B \in \mathrm{Mat}(m \times n, K)$ sind *äquivalent*, wenn sie dieselbe lineare Abbildung $K^n \to K^m$ bezüglich i.A. verschiedener Basen (von K^n und K^m) darstellen. Das ist gleichbedeutend damit, dass es Matrizen $P \in \mathrm{GL}(m, K)$ und $Q \in \mathrm{GL}(n, K)$ gibt mit $B = PAQ$. Wir hatten in der *Linearen Algebra I* gezeigt, dass das genau dann der Fall ist, wenn $\mathrm{rk}(A) = \mathrm{rk}(B)$ gilt.

Zwei Matrizen $A, B \in \mathrm{Mat}(n, K)$ sind *ähnlich*, wenn sie denselben Endomorphismus von K^n bezüglich i.A. verschiedener Basen von K^n darstellen. Das ist gleichbedeutend damit, dass es $P \in \mathrm{GL}(n, K)$ gibt mit $B = P^{-1}AP$. Über einem algebraisch abgeschlossenen Körper wie $\mathbb{C}$ wird die Klassifikation von Matrizen bis auf Ähnlichkeit durch die *Jordan-Normalform* geleistet, die wir später in dieser Vorlesung studieren werden. ♣

Man kann eine Äquivalenzrelation als eine „vergröberte" 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:

18.5. **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 Äquivalenzklassen

(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.
„(1) $\Leftrightarrow$ (2)": Das ist 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 $y \sim x$. ❑

18.6. **Definition.** In der Situation von Satz 18.5 schreiben wir

DEF Quotientenmenge

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

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

kanonische Surjektion

Wir können das auf unser Problem anwenden.

* **18.7. 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 $v \equiv v' \bmod U$ (gesprochen „v ist kongruent zu v' modulo U").*

LEMMA Kongruenz modulo U ist Äquivalenzrelation

$$v \equiv v' \bmod U \iff v - v' \in U\,.$$

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

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

Wir wollen jetzt gerne $V' = V/U$ setzen, mit der kanonischen Surjektion als linearer Abbildung $V \to V'$. Dazu müssen wir auf V/U eine K-Vektorraum-Struktur definieren. Wenn die kanonische Surjektion linear sein soll, dann müssen $[v] + [v'] = [v + v']$ und $\lambda[v] = [\lambda v]$ gelten (für $v, v' \in V$ und $\lambda \in K$). Wir müssen also nachweisen, dass diese Definitionen der Addition und der Skalarmultiplikation sinnvoll sind. Das wird durch zusätzliche Eigenschaften von $\equiv_U$ sichergestellt.

18.8. Lemma. *Die Relation $\equiv_U$ aus Lemma 18.7 hat folgende Eigenschaften:*

LEMMA Kongruenz modulo U ist Kongruenzrelation

(1) *Für $v_1, v_2, v_1', v_2' \in V$ gilt:*
Aus $v_1 \equiv v_1' \bmod U$ und $v_2 \equiv v_2' \bmod U$ folgt $v_1 + v_2 \equiv v_1' + v_2' \bmod U$.

(2) *Für $v, v' \in V$ und $\lambda \in K$ gilt: Aus $v \equiv v' \bmod U$ folgt $\lambda v \equiv \lambda v' \bmod U$.*

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

DEF Kongruenzrelation

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 $\lambda v - \lambda v' = \lambda(v - v') \in U$. ❑

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

SATZ Quotientenraum

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

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

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

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 $[v_2] = [v_2']$, d.h. $v_1 \equiv v_1' \bmod U$ und $v_2 \equiv v_2' \bmod U$. Nach Lemma 18.8 folgt dann $v_1 + v_2 \equiv v_1' + v_2' \bmod U$, 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 π linear ist, folgt schließlich direkt aus der Definition von Addition und Skalarmultiplikation in V/U. Tatsächlich ist die Definition gerade so gemacht, *damit* π linear wird! Es gilt dann (man beachte $\mathbf{0}_{V/U} = [\mathbf{0}_V]$, also $[v] = \mathbf{0}_{V/U} \iff v \in U$)

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

18.10. Definition. Der Vektorraum V/U („V modulo U“) heißt der *Quotientenraum* von V modulo U; die lineare Abbildung $\pi\colon V \to V/U$ heißt der zugehörige *kanonische Epimorphismus*. ♢

DEF Quotientenraum

kanonischer Epimorphismus

Damit ist die eingangs gestellte Frage positiv beantwortet.

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

DEF Kodimension allgemein

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

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

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

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

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/U \arrow[ur, dashed, "\phi"'] & \end{tikzcd}$$

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

SATZ lin. Abb. auf V/U

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

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

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

das zeigt schon einmal die Eindeutigkeit; die Frage ist nur, ob wir ϕ 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

$$f(v) = f\bigl((v - v') + v'\bigr) = f(v - v') + f(v') = f(v')\,,$$

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

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

und

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

ϕ ist genau dann injektiv, wenn $\ker(\phi)$ trivial ist. Aus der Definition von ϕ folgt $\phi([v]) = \mathbf{0} \iff v \in \ker(f)$; außerdem gilt natürlich $[v] = \mathbf{0} \iff v \in U$. Beides zusammen zeigt $\ker(\phi) = \{\mathbf{0}\} \iff \ker(f) = U$. ❑

Satz 18.11 erlaubt uns, lineare Abbildungen mit Definitionsbereich V/U zu konstruieren.

Der folgende Satz ist wichtig. In den Algebra-Vorlesungen werden Sie ähnliche Resultate für Gruppen und Ringe (statt wie hier Vektorräume) kennen lernen.

SATZ Homomorphiesatz für lineare Abb.

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

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\
V/\ker(f) \arrow[r, "\cong", "\phi"'] & \mathrm{im}(f) \arrow[u, "\iota"']
\end{tikzcd}$$

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

Eine wichtige Aussage des Homomorphiesatzes ist, dass Bilder von Homomorphismen mit Quelle V (bis auf Isomorphie) dasselbe sind wie Quotientenräume von V.

Beweis. Nach Satz 18.11 gibt es eine eindeutig bestimmte lineare Abbildung $\tilde{\phi}\colon V/\ker(f) \to W$ mit $f = \tilde{\phi} \circ \pi$. Es gilt $\operatorname{im}(\tilde{\phi}) = \operatorname{im}(f)$, also können wir $\tilde{\phi}$ im Wertebereich einschränken zu $\phi\colon V/\ker(f) \to \operatorname{im}(f)$; es folgt $f = \iota \circ \phi \circ \pi$. Es bleibt zu zeigen, dass ϕ ein Isomorphismus ist: ϕ ist injektiv nach Satz 18.11 und surjektiv wegen $\operatorname{im}(\phi) = \operatorname{im}(\tilde{\phi}) = \operatorname{im}(f)$. ❑

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

18.13. **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

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

mit Kern N („Nullfolge“ heißt ja gerade „Grenzwert null“). Aus dem Homomorphiesatz 18.12 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. ♣

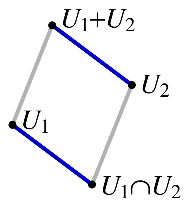
Weitere Anwendungen des Homomorphiesatzes sind durch die folgenden „Isomorphiesätze“ gegeben.

SATZ
Erster Isomorphiesatz

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

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

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 ϕ ist u ein Element von U_1.)

Beweis.

$$\begin{tikzcd}
U_1 \arrow[r] \arrow[d] \arrow[dr, "f"] & U_1 + U_2 \arrow[d] \\
U_1/(U_1 \cap U_2) \arrow[r, "\cong", "\phi"'] & (U_1 + U_2)/U_2
\end{tikzcd}$$

Wir betrachten die Verknüpfung $f\colon U_1 \to (U_1 + U_2)/U_2$ der Inklusionsabbildung $U_1 \to U_1 + U_2$ mit dem kanonischen Epimorphismus $U_1 + U_2 \to (U_1 + U_2)/U_2$. 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 18.12 existiert der Isomorphismus ϕ wie angegeben. ❑

SATZ Zweiter Isomorphiesatz

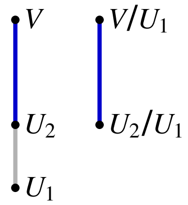
18.15. **Satz.** *Seien V ein Vektorraum und $U_1 \subset U_2 \subset V$ Untervektorräume. Dann ist U_2/U_1 ein Untervektorraum von V/U_1 und die Abbildung*

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

ist ein Isomorphismus.

Beweis.

$$\begin{tikzcd} V \arrow[r, "\pi"] \arrow[d] \arrow[dr, "f"] & V/U_1 \arrow[d] \\ V/U_2 \arrow[r, "\cong", "\phi"'] & (V/U_1)/(U_2/U_1) \end{tikzcd}$$



Sei $\pi\colon V \to V/U_1$ der kanonische Epimorphismus, dann ist $U_2/U_1 = \pi(U_2)$ ein Untervektorraum von V/U_1. Wir betrachten die Verknüpfung f von π mit dem kanonischen Epimorphismus $V/U_1 \to (V/U_1)/(U_2/U_1)$. 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 dann wieder aus dem Homomorphiesatz 18.12. ❑

DEF Nebenklasse Restklasse

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.

DEF Affiner Unterraum

* 18.16. **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. ♢

Wir kennen affine Unterräume bereits als Lösungsmengen von linearen Gleichungen (siehe Satz 12.10): Die Lösungsmenge jeder linearen Gleichung $f(x) = b$ (wobei $f\colon V \to W$ 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 $f = \mathbf{0}\colon V \to K$ und $b = 1$); für $A = v + U$ ist $A = \pi^{-1}([v])$ für den kanonischen Epimorphismus $\pi\colon V \to V/U$.

Das liefert uns zwei verschiedene Möglichkeiten, einen nicht-leeren affinen Unterraum A des endlich-dimensionalen Standard-Vektorraums K^n zu beschreiben (wir betrachten die Elemente von K^n als Spaltenvektoren):

(1) Als verschobener Untervektorraum in der Form $A = \boldsymbol{x}_0 + \langle \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \rangle$, wobei $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)$ eine Basis des linearen Unterraums $A - A$ ist.

(2) Als Lösungsmenge eines Systems von linearen Gleichungen $M\boldsymbol{x} = \boldsymbol{b}$.

Von einer Beschreibung der Form (2) kommen wir zu einer Beschreibung der Form (1), indem wir den Lösungsalgorithmus für lineare Gleichungssysteme verwenden. Für die umgekehrte Richtung müssen wir $U = \langle \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \rangle$ als Kern einer Matrix M schreiben. Sei dafür M' die $(m \times n)$-Matrix, deren Zeilen durch $\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_m^\top$ gegeben sind. Dann folgt, dass der Zeilenraum der gesuchten Matrix M genau der Kern von M' ist (Übung). Indem wir eine Basis des Kerns von M' als Zeilen in eine Matrix schreiben, bekommen wir also einen Kandidaten für M; die rechte Seite $\boldsymbol{b}$ ergibt sich dann als $\boldsymbol{b} = M\boldsymbol{x}_0$.

Welche Art der Beschreibung nützlicher ist, hängt davon ab, was wir mit dem affinen Unterraum anstellen wollen. Wenn wir zum Beispiel der Durchschnitt von zwei (oder mehr) affinen Unterräumen bestimmen wollen (der, wie wir bald sehen werden, wieder ein affiner Unterraum ist), dann ist Beschreibung (2) hilfreicher.

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 im Fall $U = \{\mathbf{0}\}$), denn jedes $v \in A$ erfüllt $A = v + U$.

Wir können affine Unterräume durch eine Abgeschlossenheitseigenschaft charakterisieren.

* 18.17. **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 alle $a_1, a_2, \ldots, a_n \in A$ und $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ mit $\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$ gilt $\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n \in A$.*

Beweis. „(1)⇒(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 $j \in \{1, 2, \ldots, n\}$, also erhalten wir

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

„(2)⇒(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 (2) 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 (2) gilt $\lambda a - \lambda v + v \in A$, also ist $\lambda u = (\lambda a - \lambda v + v) - v \in U$. ❑

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

18.18. **Folgerung.** *Sei V ein Vektorraum und sei $(A_i)_{i \in I}$ eine Familie von affinen Unterräumen von V mit $I \neq \emptyset$. Dann ist $\bigcap_{i \in I} A_i$ ebenfalls ein affiner Unterraum von V.* **FOLG** Durchschnitte von affinen Unterräumen

Beweis. Nach Satz 18.17 sind alle A_i abgeschlossen unter affinen Linearkombinationen. Seien jetzt $a_1, a_2, \ldots, a_n \in A = \bigcap_{i \in I} A_i$ und seien $\lambda_1, \lambda_2, \ldots, \lambda_n$ Skalare mit $\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$. Dann ist $a = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n \in A_i$ für alle $i \in I$, also ist $a \in A$. Damit ist A unter affinen Linearkombinationen abgeschlossen, also ist A wiederum nach Satz 18.17 ein affiner Unterraum von V. ❑

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).

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

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

- Die leere Menge ist ein affiner Unterraum.
- Jede einelementige Menge $\{\boldsymbol{x}\}$ 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.

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 $g = \boldsymbol{x}_0 + \langle \boldsymbol{y} \rangle$ beschrieben werden (mit „Aufpunkt“ $\boldsymbol{x}_0$ und „Richtungsvektor“ $\boldsymbol{y} \neq \boldsymbol{0}$) 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(affine Hülle von T) $\le \#T - 1$. 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 $A = \boldsymbol{x}_0 + \langle \boldsymbol{y}_1, \boldsymbol{y}_2 \rangle$ hat, wobei $g_1 \cap g_2 = \{\boldsymbol{x}_0\}$ ist und $\boldsymbol{y}_1, \boldsymbol{y}_2$ 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:

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

19. Der Dualraum

Wir hatten im ersten Semester schon gesehen, dass die Menge aller Homomorphismen $f: V \to W$ zwischen zwei K-Vektorräumen V und W selbst wieder die Struktur eines K-Vektorraums $\mathrm{Hom}(V, W)$ hat (Satz 10.20). Ein besonders wichtiger Spezialfall tritt auf, wenn $W = K$ ist.

* **19.1. Definition.** Sei V ein K-Vektorraum. Eine lineare Abbildung $\phi: V \to K$ heißt auch eine *Linearform* auf V. Der Vektorraum $V^* = \mathrm{Hom}(V, K)$ 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 $v \mapsto \lambda\phi(v)$.

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

$$\mathrm{pr}_j: (x_1, x_2, \dots, x_n) \longmapsto x_j, \qquad j \in \{1, 2, \dots, n\},$$

Linearformen.

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

$$\mathrm{ev}_x: f \longmapsto f(x)$$

eine Linearform auf V.

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

$$\mathrm{ev}_v: \phi \longmapsto \phi(v)$$

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

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 10.11). Daraus ergibt sich der folgende wichtige Satz.

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

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

* **19.4. Definition.** In der Situation von Satz 19.3 heißt die Basis $(v_1^*, \ldots, v_n^*)$ die *zur Basis* $(v_1, \ldots, v_n)$ *duale Basis* von V^*. ♢

DEF duale Basis

Man beachte, dass jedes Element v_i^* der dualen Basis von *allen* Elementen $v_1, \ldots, v_n$ 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 $v_1^*, v_2^*, \ldots, v_n^* \in V^*$ eine Basis bilden.

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

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

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

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

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

$$\bigl(\phi(v_1)v_1^* + \phi(v_2)v_2^* + \ldots + \phi(v_n)v_n^*\bigr)(v_j) = \phi(v_j) \,.$$

Das zeigt, dass ϕ eine Linearkombination von $v_1^*, v_2^*, \ldots, v_n^*$ ist.

(Alternativ könnten wir argumentieren, dass $\dim V^* = \dim \operatorname{Hom}(V, K) = \dim V = n$ ist, sodass die lineare Unabhängigkeit bereits ausreicht.) ❑

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 $\phi \notin \langle \{b_i^* \mid i \in I\} \rangle$.

19.5. Beispiel. Die duale Basis zur Standardbasis $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ von K^n besteht gerade aus den Koordinatenabbildungen $(\mathrm{pr}_1, \mathrm{pr}_2, \ldots, \mathrm{pr}_n)$. Wir bekommen einen Isomorphismus $f\colon K^n \to (K^n)^*$, indem wir $\mathbf{e}_j$ auf pr_j abbilden. Wenn wir die Elemente von K^n als Spaltenvektoren betrachten, dann können wir die Elemente des dualen K^n als Zeilenvektoren auffassen. Sind $\boldsymbol{x}, \boldsymbol{y} \in K^n$, dann ist die Auswertung von $\boldsymbol{x}^\top$ auf $\boldsymbol{y}$ gegeben durch die Multiplikation einer $(1 \times n)$-Matrix mit einer $(n \times 1)$-Matrix oder auch durch das Skalarprodukt:

BSP duale Basis der Standardbasis

$$\boldsymbol{x}^\top(\boldsymbol{y}) = f(\boldsymbol{x})(\boldsymbol{y}) = \boldsymbol{x}^\top \boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + \ldots + x_n y_n \,.$$

Es ist nämlich

$$f(\boldsymbol{x})(\boldsymbol{y}) = (x_1 \,\mathrm{pr}_1 + \ldots + x_n \,\mathrm{pr}_n)(\boldsymbol{y}) = x_1 y_1 + \ldots x_n y_n \,.$$ ♣

Aus dem Satz ergibt sich unmittelbar:

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

$$\dim V = \dim V^* .$$

Insbesondere sind V und V^* isomorph.

FOLG $V \cong V^*$ für V endl.-dim.

Die Aussage von Folgerung 19.6 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 $X \to \mathcal{P}(X)$. Zum Beweis sei $f\colon X \to \mathcal{P}(X)$ 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 ϕ_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 ϕ_T enthält). Es folgt, dass jede Basis von V^* echt mächtiger sein muss als B.

Ein Isomorphismus $V \to V^*$ ist — nach Wahl einer Basis $(v_1, v_2, \ldots, v_n)$ 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^{**}.

Bevor wir das zeigen, formulieren wir noch eine Aussage, die wir später brauchen.

19.7. Lemma.

LEMMA Fortsetzung linearer Abbildungen

(1) *Seien V und W Vektorräume und $U \subset V$ ein Untervektorraum. Ist außerdem $f\colon U \to W$ eine lineare Abbildung, dann kann man f zu einer linearen Abbildung $F\colon V \to W$ 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 $\phi(v) = 1$.*

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 Teil (1) an auf $U = \langle v \rangle$ und $f\colon U \to K$, $\lambda v \mapsto \lambda$. Für die allgemeinere Aussage betrachten wir $\mathbf{0} \neq [v] \in V/U$; dann gibt es eine Linearform $\bar{\phi}\colon V/U \to K$ mit $\bar{\phi}([v]) = 1$. Damit hat $\phi = \bar{\phi} \circ \pi$ die gewünschten Eigenschaften, wobei $\pi\colon V \to V/U$ der kanonische Epimorphismus ist. ❑

* **19.8. Satz.** *Sei V ein K-Vektorraum. Dann ist die folgende Abbildung ein injektiver Homomorphismus:* **SATZ** kanon. Abb. in den Bidualraum

$$\alpha_V\colon V \longrightarrow V^{**}, \quad v \longmapsto \bigl(\mathrm{ev}_v\colon \phi \mapsto \phi(v)\bigr)\,.$$

Ist V endlich-dimensional, dann ist α_V ein Isomorphismus.

Beweis. Es sind verschiedene Aussagen zu zeigen.

- $\mathrm{ev}_v \in V^{**}$ (siehe Beispiel 19.2): ev_v ist eine Abbildung $V^* \to K$; wir müssen zeigen, dass ev_v linear ist:
$$\mathrm{ev}_v(\phi + \phi') = (\phi + \phi')(v) = \phi(v) + \phi'(v) = \mathrm{ev}_v(\phi) + \mathrm{ev}_v(\phi')$$
und analog für die Skalarmultiplikation. (Hier benutzen wir die Definition der Vektorraumstruktur von V^*.)
- α_V ist linear: $\alpha_V(v + v') = \mathrm{ev}_{v+v'}$ bildet $\phi \in V^*$ auf $\phi(v + v') = \phi(v) + \phi(v') = \mathrm{ev}_v(\phi) + \mathrm{ev}_{v'}(\phi)$ ab, hat also denselben Effekt wie $\alpha_V(v) + \alpha_V(v')$. Analog sehen wir, dass $\alpha_V(\lambda v) = \mathrm{ev}_{\lambda v}$ 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^{**}.)
- α_V ist injektiv: Wir zeigen $\ker(\alpha_V) = \{\mathbf{0}\}$. Sei $v \neq \mathbf{0}$. Nach Lemma 19.7 gibt es $\phi \in V^*$ mit $\bigl(\alpha_V(v)\bigr)(\phi) = \phi(v) = 1 \neq 0$, 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 $\ker(\alpha_V)$.
- α_V ist Isomorphismus, falls $\dim V < \infty$: In diesem Fall gilt nach Folgerung 19.6 $\dim V = \dim V^* = \dim V^{**}$. Als injektive lineare Abbildung zwischen endlich-dimensionalen Vektorräumen derselben Dimension muss α_V dann ein Isomorphismus sein (Folgerung 10.14). ❑

Ist V endlich-dimensional, dann kann man also V und V^{**} durch den kanonischen Isomorphismus α_V miteinander identifizieren und damit V als den Dualraum von V^* betrachten. Das bedeutet, dass die Beziehung zwischen V und V^* symmetrisch ist. Das wird zum Beispiel durch die nächste Aussage illustriert.

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

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

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

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

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

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

Die erste Aussage haben wir im Beweis von Satz 19.3 verwendet, die zweite folgt durch Vertauschen der Rollen von V und V^*. Diese zweite Relation zeigt, dass man die Elemente v_i^* der dualen Basis als Koordinatenabbildungen bezüglich der Basis $(v_1, \ldots, v_n)$ von V interpretieren kann.

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

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

ist — wir erhalten wieder die Lagrangesche Interpolationsformel, vergleiche Beispiel 10.15. ♣

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

$$\begin{tikzcd} V \arrow[d, "f"'] \arrow[dr, "f^\top(\phi)"] & \\ W \arrow[r, "\phi"] & K \end{tikzcd}$$

Wir erhalten eine Abbildung $f^\top \colon W^* \to V^*$.

* 19.11. **Definition.** Ist $f \colon V \to W$ linear, dann heißt $f^\top \colon W^* \to V^*$, $\phi \mapsto \phi \circ f$, 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.

19.12. 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 $f\colon V \to W$ linear und $A = \mathrm{Mat}_{B,B'}(f)$ die f bezüglich der Basen B und B' darstellende Matrix. Dann gilt*

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

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

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

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

also ist

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

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

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

Anwenden auf b_j ergibt

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

also sind die beiden Matrizen zueinander transponiert. ❑

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

LEMMA $f \mapsto f^\top$

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

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

Beweis. Φ ist linear, denn für $\phi \in W^*$ und $f, g \in \mathrm{Hom}(V,W)$ 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 $\Phi(\lambda f) = (\lambda f)^\top = \lambda f^\top = \lambda\Phi(f)$.

Φ ist injektiv, denn für $f \in \mathrm{Hom}(V,W)$ 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 19.7), folgt $f(v) = \mathbf{0}$ für alle $v \in V$, also $f = \mathbf{0}$.

Sind V und W beide endlich-dimensional, dann gilt (Satz 10.22)

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

also ist Φ als injektive lineare Abbildung zwischen zwei endlich-dimensionalen Vektorräumen derselben Dimension ein Isomorphismus. ❑

Wir zeigen noch einige weitere einfache Eigenschaften der transponierten Abbildung.

LEMMA
Eigenschaften von $f^\top$

19.14. **Lemma.**

(1) *Ist V ein Vektorraum, dann gilt* $\mathrm{id}_V^\top = \mathrm{id}_{V^*}$.

(2) *Sind V, V', V'' Vektorräume und $f: V \to V'$ und $g: V' \to V''$ lineare Abbildungen, dann gilt* $(g \circ f)^\top = f^\top \circ g^\top$.

(3) *Ist $f: V \to W$ ein Isomorphismus, dann ist auch $f^\top$ ein Isomorphismus und es gilt* $(f^\top)^{-1} = (f^{-1})^\top$.

Beweis.

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

(2) Für $\phi \in (V'')^*$ ist

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

(3) Nach den beiden ersten Teilen gilt

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

und

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

woraus die Behauptungen folgen. ❑

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

LEMMA
$f^\top$ und α_V

19.15. **Lemma.**

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

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

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\alpha_V"'] & W \arrow[d, "\alpha_W"] \\ V^{**} \arrow[r, "f^{\top\top}"] & W^{**} \end{tikzcd}$$

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

20. Bilinearformen

Nachdem wir im letzten Abschnitt Linearformen besprochen haben, kommen wir jetzt zu bilinearen Abbildungen und Bilinearformen.

∗ **20.1. Definition.** Seien K ein Körper und V_1, V_2, W drei K-Vektorräume. Eine Abbildung $\beta\colon V_1 \times V_2 \to W$ heißt *(K-)bilinear*, wenn β 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

$$\beta(v_1 + v_1', v_2) = \beta(v_1, v_2) + \beta(v_1', v_2)\,, \qquad \beta(\lambda v_1, v_2) = \lambda\beta(v_1, v_2)$$
$$\beta(v_1, v_2 + v_2') = \beta(v_1, v_2) + \beta(v_1, v_2')\,, \qquad \beta(v_1, \lambda v_2) = \lambda\beta(v_1, v_2)\,.$$

Ist $W = K$, dann heißt β eine *(K-)Bilinearform* oder auch *Paarung*. Gilt außerdem $V_1 = V_2 = V$, dann heißt β eine *(K-)Bilinearform auf V*. Wir bezeichnen den K-Vektorraum aller Bilinearformen $V_1 \times V_2 \to K$ mit $\operatorname{Bil}(V_1, V_2)$.

Ist $\beta\colon V \times V \to K$ eine Bilinearform auf V, dann heißt β *symmetrisch*, wenn für alle $v_1, v_2 \in V$ gilt, dass $\beta(v_2, v_1) = \beta(v_1, v_2)$ ist. β heißt *alternierend*, wenn für alle $v \in V$ gilt, dass $\beta(v, v) = 0$ ist. ♢

Ist $\beta\colon V \times V \to K$ eine alternierende Bilinearform, dann gilt $\beta(v_2, v_1) = -\beta(v_1, v_2)$. Das sieht man so:

$$\begin{aligned} 0 &= \beta(v_1 + v_2, v_1 + v_2) \\ &= \beta(v_1, v_1) + \beta(v_1, v_2) + \beta(v_2, v_1) + \beta(v_2, v_2) \\ &= \beta(v_1, v_2) + \beta(v_2, v_1)\,. \end{aligned}$$

Umgekehrt folgt aus $\beta(v_2, v_1) = -\beta(v_1, v_2)$ für alle $v_1, v_2 \in V$ die Gleichung $\beta(v, v) = -\beta(v, v)$, also $2\beta(v, v) = 0$ für alle $v \in V$. Kann man in K durch 2 teilen (im Körper $\mathbb{F}_2$ mit zwei Elementen ist $2 = 0$, dort geht das nicht, aber sonst praktisch immer), dann folgt, dass β alternierend ist. Im Normalfall sind alternierende Bilinearformen also dasselbe wie schief-symmetrische.

Bilineare Abbildungen treten häufig in Gestalt einer Multiplikationsabbildung auf.

20.2. Beispiele. Die Matrixmultiplikation

BSP
bilineare Abb.

$$\operatorname{Mat}(l \times m, K) \times \operatorname{Mat}(m \times n, K) \longrightarrow \operatorname{Mat}(l \times n, K), \quad (A, B) \longmapsto AB$$

ist eine bilineare Abbildung (das folgt aus den Rechenregeln für Matrizen). Genauso ist die Multiplikation von Polynomen

$$K[X] \times K[X] \longrightarrow K[X], \quad (p, q) \longmapsto pq$$

eine bilineare Abbildung.

Das *Standard-Skalarprodukt*

DEF
Standard-Skalarprodukt

$$K^n \times K^n \longrightarrow K, \quad \bigl((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)\bigr) \longmapsto x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

ist eine symmetrische Bilinearform auf K^n.

Die Abbildung

$$K^2 \times K^2 \longrightarrow K, \quad \bigl((x_1, x_2), (y_1, y_2)\bigr) \longmapsto x_1y_2 - x_2y_1$$

ist eine alternierende Bilinearform auf K^2.

Die *Spurform*

$$\operatorname{Mat}(m \times n, K) \times \operatorname{Mat}(m \times n, K) \longrightarrow K, \quad (A, B) \longmapsto \operatorname{Tr}(A^\top B) = \operatorname{Tr}(AB^\top)$$

ist eine symmetrische Bilinearform auf $\operatorname{Mat}(m \times n, K)$. ♣

Allgemein gilt (leichte Übung): Ist $\beta\colon V_1 \times V_2 \to W$ bilinear und $f\colon W \to W'$ linear, dann ist $f \circ \beta$ bilinear. Ähnlich wie wir linearen Abbildungen zwischen endlich-dimensionalen Vektorräumen mit festgelegten Basen Matrizen zuordnen können, können wir auch Bilinearformen durch Matrizen beschreiben.

20.3. **Definition.** Seien V und W endlich-dimensionale K-Vektorräume. Wir setzen $\dim V = m$ und $\dim W = n$. Sei $\beta\colon V \times W \to K$ eine Bilinearform. Seien weiter $B = (b_1, b_2, \ldots, b_m)$ eine Basis von V und $B' = (b'_1, b'_2, \ldots, b'_n)$ eine Basis von W. Dann heißt

DEF Matrix einer Bilinearform

$$\operatorname{Mat}_{B,B'}(\beta) = \bigl(\beta(b_i, b'_j)\bigr)_{1\le i\le m, 1\le j\le n} = \begin{pmatrix} \beta(b_1,b'_1) & \beta(b_1,b'_2) & \cdots & \beta(b_1,b'_n) \\ \beta(b_2,b'_1) & \beta(b_2,b'_2) & \cdots & \beta(b_2,b'_n) \\ \vdots & \vdots & & \vdots \\ \beta(b_m,b'_1) & \beta(b_m,b'_2) & \cdots & \beta(b_m,b'_n) \end{pmatrix}$$

die *Matrix von β* bezüglich B und B'. Im Fall $V = W$ und $B = B'$ schreiben wir auch $\operatorname{Mat}_B(\beta)$ statt $\operatorname{Mat}_{B,B}(\beta)$. ♢

Sind $v = x_1b_1 + x_2b_2 + \ldots + x_mb_m \in V$ und $v' = y_1b'_1 + y_2b'_2 + \ldots + y_nb'_n \in W$, dann ist $\beta(v, v') = \sum_{i=1}^m \sum_{j=1}^n x_iy_j\beta(b_i, b'_j)$, was sich in folgende Matrixmultiplikation übersetzen lässt (rechts steht eine 1×1-Matrix, die wir mit ihrem einzigen Eintrag identifizieren):

$$\beta(v, v') = (x_1\, x_2\, \ldots\, x_m) \operatorname{Mat}_{B,B'}(\beta) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} .$$

Seien B_V, B'_V Basen von V und B_W, B'_W Basen von W, sowie $\beta\colon V \times W \to K$ eine Bilinearform. Mit $A = \operatorname{Mat}_{B_V,B_W}(\beta)$, $A' = \operatorname{Mat}_{B'_V,B'_W}(\beta)$ und den Basiswechselmatrizen $P = \operatorname{Mat}_{B'_V,B_V}(\mathrm{id}_V)$ und $Q = \operatorname{Mat}_{B'_W,B_W}(\mathrm{id}_W)$ gilt dann (ähnlich wie für lineare Abbildungen)

$$A' = P^\top AQ .$$

Für Bilinearformen auf einem Vektorraum V, wo man nur eine Basis (von V) wählen kann, muss dabei $Q = P$ sein. In diesem Fall heißen zwei Matrizen A und A', die dieselbe Bilinearform (bezüglich zweier i.A. verschiedener Basen) beschreiben, auch *kongruent*. Das bedeutet also, dass es $P \in \mathrm{GL}(\dim V, K)$ gibt mit $A' = P^\top AP$; wir erhalten eine Äquivalenzrelation auf $\operatorname{Mat}(\dim V, K)$.

DEF kongruente Matrizen

Es gibt einen Zusammenhang zwischen Bilinearformen einerseits und Linearformen und Dualräumen andererseits.

20.4. **Lemma.**

LEMMA Linearformen aus einer Bilinearform

(1) *Seien V und W zwei K-Vektorräume und $\beta\colon V \times W \to K$ eine Bilinearform. Dann ist für jedes $v \in V$ die Abbildung $W \to K$, $w \mapsto \beta(v, w)$, eine Linearform auf W, und für jedes $w \in W$ ist die Abbildung $V \to K$, $v \mapsto \beta(v, w)$, eine Linearform auf V.*

(2) *Die durch (1) gegebenen Abbildungen*

$$\beta_L\colon V \longrightarrow W^*, \quad v \longmapsto \bigl(w \mapsto \beta(v, w)\bigr)$$

und

$$\beta_R\colon W \longrightarrow V^*, \quad w \longmapsto \bigl(v \mapsto \beta(v, w)\bigr)$$

sind linear.

(3) *Die sich aus* (2) *ergebenden Abbildungen*

$$\mathrm{Bil}(V,W) \longrightarrow \mathrm{Hom}(V,W^*), \quad \beta \longmapsto \beta_L$$

und

$$\mathrm{Bil}(V,W) \longrightarrow \mathrm{Hom}(W,V^*), \quad \beta \longmapsto \beta_R$$

sind Isomorphismen. Insbesondere sind $\mathrm{Hom}(V,W^*)$ *und* $\mathrm{Hom}(W,V^*)$ *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 $f \mapsto \big((v,w) \mapsto (f(v))(w)\big)$: Einerseits wird $\beta \in \mathrm{Bil}(V,W)$ wie folgt abgebildet:

$$\beta \longmapsto \beta_L \longmapsto \big((v,w) \mapsto (\beta_L(v))(w) = \beta(v,w)\big) = \beta\,;$$

andererseits haben wir für $f \in \mathrm{Hom}(V,W^*)$

$$f \longmapsto \big((v,w) \mapsto (f(v))(w)\big) \longmapsto \big(v \mapsto \big(w \mapsto (f(v))(w)\big) = f(v)\big) = f\,.$$

Die Bijektivität der zweiten Abbildung zeigt man analog. (Die Inverse ist $f \mapsto \big((v,w) \mapsto (f(w))(v)\big)$.) ❑

Ist B eine endliche Basis von V und B' eine endliche Basis von W, dann gilt

$$\mathrm{Mat}_{B,B'^*}(\beta_L) = \mathrm{Mat}_{B,B'}(\beta)^\top \qquad \text{und} \qquad \mathrm{Mat}_{B',B^*}(\beta_R) = \mathrm{Mat}_{B,B'}(\beta)\,.$$

20.5. **Definition.** Eine Bilinearform $\beta\colon V \times W \to K$ heißt *nicht-ausgeartet*, wenn $\beta_L\colon V \to W^*$ und $\beta_R\colon W \to V^*$ Isomorphismen sind. Anderenfalls heißt β *ausgeartet*. ♢ **DEF** Bilinearform nicht-ausgeartet

Wenn man eine solche nicht-ausgeartete Bilinearform hat, dann kann man (via β_L und β_R) V als Dualraum von W und umgekehrt betrachten: Zu jeder Linearform ϕ 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 ψ 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 α_V ein Isomorphismus ist. Das ist aber nur für endlich-dimensionale Vektorräume V der Fall.

Dann müssen V und W dieselbe Dimension haben: $\dim W = \dim W^* = \dim V$.

20.6. **Beispiel.** Sei V ein endlich-dimensionaler K-Vektorraum. Dann ist die *Auswertungspaarung* **BSP** nicht-ausg. Bilinearform

$$\mathrm{ev}\colon V \times V^* \longrightarrow K, \quad (v,\phi) \longmapsto \phi(v)$$

nicht-ausgeartet. Für beliebiges V gilt (Übung)

$$\mathrm{ev}_L = \alpha_V\colon V \longrightarrow V^{**} \qquad \text{und} \qquad \mathrm{ev}_R = \mathrm{id}_{V^*}\colon V^* \longrightarrow V^*\,. \qquad ♣$$

20.7. **Lemma.** *Seien V und W zwei K-Vektorräume derselben endlichen Dimension n, sei B eine Basis von V, B' eine Basis von W und $\beta \in \mathrm{Bil}(V, W)$. Wir setzen $A = \mathrm{Mat}_{B,B'}(\beta)$. Dann sind folgende Aussagen äquivalent:*

LEMMA Kriterium für nicht-ausgeartet

(1) *β ist nicht-ausgeartet.*

(2) $\ker(\beta_L) = \{\mathbf{0}\}$.

(3) $\ker(\beta_R) = \{\mathbf{0}\}$.

(4) $\det(A) \neq 0$.

Beweis. Dass aus (1) die Aussagen (2) und (3) folgen, ist klar nach Definition 20.5. Umgekehrt folgt aus (2) zunächst, dass β_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 (denn $\beta_L^\top$ ist ein Isomorphismus nach Lemma 19.14). Genauso zeigt man „(3)⇒(1)“. Schließlich ist (3) äquivalent dazu, dass $\mathrm{Mat}_{B',B^*}(\beta_R)$ invertierbar ist. Diese Matrix ist aber genau A, und „A invertierbar“ ist äquivalent zu „$\det(A) \neq 0$“. ❑

Wir wollen jetzt symmetrische Bilinearformen auf einem Vektorraum V genauer betrachten.

20.8. **Lemma.** *Seien K ein Körper, $n \in \mathbb{N}$, V ein n-dimensionaler K-Vektorraum mit Basis B und $\beta \colon V \times V \to K$ eine Bilinearform auf V. Sei weiter $A = \mathrm{Mat}_B(\beta)$. Dann gilt:*

LEMMA Matrix einer symm. Bilinearform

$$\beta \text{ ist symmetrisch} \quad \iff \quad A^\top = A$$

Eine Matrix $A \in \mathrm{Mat}(n, K)$ heißt *symmetrisch*, wenn $A^\top = A$ ist. Damit gilt also, dass eine Bilinearform genau dann symmetrisch ist, wenn ihre Matrix (bezüglich einer beliebigen Basis) symmetrisch ist.

DEF symmetrische Matrix

Beweis. Sei $B = (b_1, b_2, \ldots, b_n)$.

„⇒“: Ist β symmetrisch, dann ist $\beta(b_i, b_j) = \beta(b_j, b_i)$; das bedeutet gerade $A^\top = A$.

„⇐“: Sei $A^\top = A$. Dann gilt für Spaltenvektoren $\boldsymbol{x}, \boldsymbol{y} \in K^n$:

$$\boldsymbol{x}^\top A \boldsymbol{y} = (\boldsymbol{x}^\top A \boldsymbol{y})^\top = \boldsymbol{y}^\top A^\top \boldsymbol{x} = \boldsymbol{y}^\top A \boldsymbol{x} \,.$$

(Beachte: $\boldsymbol{x}^\top A \boldsymbol{y}$ ist eine 1×1-Matrix und damit gleich ihrer Transponierten.) Daraus folgt $\beta(v, v') = \beta(v', v)$ für alle $v, v' \in V$. ❑

Zum Beispiel sieht man so sehr leicht, dass das Standard-Skalarprodukt eine nicht-ausgeartete symmetrische Bilinearform auf K^n ist, denn die zugehörige Matrix bezüglich der Standardbasis ist die Einheitsmatrix I_n, die symmetrisch ist und $\det(I_n) = 1$ erfüllt.

Wir betrachten im Folgenden den Fall $K = \mathbb{R}$. Dann können wir zwischen positiven und negativen Elementen von $\mathbb{R}$ unterscheiden. Das führt zu folgender Definition.

* **20.9. Definition.** Seien V ein $\mathbb{R}$-Vektorraum und $\beta\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf V. **DEF** positiv/negativ (semi)definit indefinit

(1) β heißt *positiv semidefinit*, wenn $\beta(v, v) \geq 0$ ist für alle $v \in V$.

(2) β heißt *positiv definit*, wenn $\beta(v, v) > 0$ ist für alle $\mathbf{0} \neq v \in V$.

(3) β heißt *negativ semidefinit*, wenn $\beta(v, v) \leq 0$ ist für alle $v \in V$.

(4) β heißt *negativ definit*, wenn $\beta(v, v) < 0$ ist für alle $\mathbf{0} \neq v \in V$.

(5) β heißt *indefinit*, wenn es $v, v' \in V$ gibt mit $\beta(v, v) > 0$ und $\beta(v', v') < 0$.

Seien $n \in \mathbb{N}$ und $A \in \operatorname{Mat}(n, \mathbb{R})$ symmetrisch, also $A^\top = A$. Im Folgenden betrachten wir $\boldsymbol{x} \in \mathbb{R}^n$ stets als Spaltenvektor.

(1) A heißt *positiv semidefinit*, wenn $\boldsymbol{x}^\top A \boldsymbol{x} \geq 0$ ist für alle $\boldsymbol{x} \in \mathbb{R}^n$.

(2) A heißt *positiv definit*, wenn $\boldsymbol{x}^\top A \boldsymbol{x} > 0$ ist für alle $\mathbf{0} \neq \boldsymbol{x} \in \mathbb{R}^n$.

(3) A heißt *negativ semidefinit*, wenn $\boldsymbol{x}^\top A \boldsymbol{x} \leq 0$ ist für alle $\boldsymbol{x} \in \mathbb{R}^n$.

(4) A heißt *negativ definit*, wenn $\boldsymbol{x}^\top A \boldsymbol{x} < 0$ ist für alle $\mathbf{0} \neq \boldsymbol{x} \in \mathbb{R}^n$.

(5) A heißt *indefinit*, wenn es $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ gibt mit $\boldsymbol{x}^\top A \boldsymbol{x} > 0$ und $\boldsymbol{y}^\top A \boldsymbol{y} < 0$. ♢

Daraus folgt im Fall $\dim V < \infty$, dass β genau dann positiv/negativ (semi-)definit bzw. indefinit ist, wenn das für $\operatorname{Mat}_B(\beta)$ mit irgendeiner Basis B von V gilt.

20.10. Beispiele. Das Standard-Skalarprodukt auf $\mathbb{R}^n$ ist positiv definit, denn $x_1^2 + x_2^2 + \ldots + x_n^2 > 0$, wenn nicht alle x_j null sind. **BSP** positiv definite Bilinearformen

Die Spurform auf $\operatorname{Mat}(m \times n, \mathbb{R})$ ist ebenfalls positiv definit, denn für eine Matrix $A = (a_{ij}) \in \operatorname{Mat}(m \times n, \mathbb{R})$ ist

$$\operatorname{Tr}(A^\top A) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \,.$$

(Wenn man $\operatorname{Mat}(m \times n, \mathbb{R})$ mit $\mathbb{R}^{mn}$ in der üblichen Weise identifiziert, dann ist die Spurform einfach das Standard-Skalarprodukt.) ♣

Die Matrix einer positiv definiten symmetrischen Bilinearform kann auch negative Einträge haben und eine symmetrische Matrix mit lauter positiven Einträgen braucht nicht positiv definit zu sein.

20.11. Beispiele. Seien **BSP**

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} .$$

Dann ist A positiv definit, denn

$$\begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 - 2xy + 2y^2 = 2(x - \tfrac{1}{2}y)^2 + \tfrac{3}{2}y^2 \,,$$

und B ist nicht positiv semidefinit, denn

$$\begin{pmatrix} 1 & -1 \end{pmatrix} B \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1^2 + 4 \cdot 1 \cdot (-1) + (-1)^2 = -2 \,.$$

Tatsächlich ist B indefinit, denn $\mathbf{e}_1^\top B \mathbf{e}_1 = 1$. ♣

20.12. Beispiel. Ist V ein endlich-dimensionaler reeller Vektorraum mit einer positiv definiten symmetrischen Bilinearform $\beta\colon V \times V \to \mathbb{R}$, dann ist β nicht-ausgeartet: Wir zeigen $\ker(\beta_L) = \{\mathbf{0}\}$. Sei also $v \in \ker(\beta_L)$. Dann ist

BSP
pos. def. $\Rightarrow$ nicht-ausg.

$$0 = \mathbf{0}(v) = \bigl(\beta_L(v)\bigr)(v) = \beta(v, v)\,.$$

Wäre $v \neq \mathbf{0}$, dann hätten wir $\beta(v, v) > 0$, also muss $v = \mathbf{0}$ sein. ♣

Unser nächstes Ziel wird es sein, ein relativ einfaches Kriterium herzuleiten, mit dem man entscheiden kann, ob eine symmetrische Matrix positiv (oder negativ) definit ist. Dies geschieht im Hinblick auf Anwendungen in der Analysis II (dort wird es um Kriterien gehen, wann eine Funktion mehrerer Variabler ein Maximum oder Minimum hat).

Dafür werden wir folgende Aussage verwenden, die wir allerdings jetzt noch nicht beweisen können. Das werden wir bald nachholen. Zuerst noch eine Definition.

20.13. Definition. Sei $n \in \mathbb{N}$. Eine Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ heißt *orthogonal*, wenn $A^\top A = I_n$ ist. Wir schreiben $\mathrm{O}(n)$ für die Menge der orthogonalen Matrizen in $\mathrm{Mat}(n, \mathbb{R})$. ♢

DEF
orthogonale Matrix

Dann ist insbesondere A invertierbar (mit $A^{-1} = A^\top$). Man prüft ohne große Schwierigkeiten nach, dass $\mathrm{O}(n)$ eine Gruppe (mit der Matrixmultiplikation als Verknüpfung) ist.

20.14. Satz. *Ist $A \in \mathrm{Mat}(n, \mathbb{R})$ symmetrisch, dann ist A (über $\mathbb{R}$) orthogonal diagonalisierbar: Es gibt eine Matrix $P \in \mathrm{O}(n)$, sodass $P^\top AP = P^{-1}AP = D$ eine Diagonalmatrix ist.*

SATZ
Spektral-satz

Daraus folgt leicht:

20.15. Lemma. *Sei $A \in \mathrm{Mat}(n, \mathbb{R})$ symmetrisch. Dann gilt:*

LEMMA
Definitheit über Eigen-werte

(1) *A ist genau dann positiv semidefinit, wenn A keinen negativen Eigenwert hat.*
(2) *A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.*
(3) *A ist genau dann negativ semidefinit, wenn A keinen positiven Eigenwert hat.*
(4) *A ist genau dann negativ definit, wenn alle Eigenwerte von A negativ sind.*
(5) *A ist genau dann indefinit, wenn A positive und negative Eigenwerte hat.*

Beweis. Nach Satz 20.14 gibt es $P \in \mathrm{O}(n)$, sodass

$$P^\top AP = P^{-1}AP = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

eine Diagonalmatrix ist; ihre Diagonaleinträge sind gerade die Eigenwerte von A. Nach Lemma 20.8 ist D die Matrix der A entsprechenden symmetrischen Bilinearform auf $\mathbb{R}^n$ bezüglich einer anderen Basis (gegeben durch die Spalten von P), also ist A genau dann positiv definit, wenn D positiv definit ist. Für einen Spaltenvektor $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ gilt

$$\boldsymbol{x}^\top D\boldsymbol{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2\,.$$

Sind alle $\lambda_j > 0$, dann ist das positiv für alle $\boldsymbol{x} \neq \mathbf{0}$, also ist D (und damit A) positiv definit. Ist hingegen $\lambda_j \leq 0$ für ein j, dann ist $\boldsymbol{x}^\top D\boldsymbol{x} \leq 0$ für $\boldsymbol{x} = \mathbf{e}_j$, und

D (und damit A) ist nicht positiv definit. Die anderen Aussagen sieht man auf die gleiche Weise. ❑

Das Definitheitskriterium wird mit Hilfe von Determinanten geeigneter Untermatrizen formuliert, sogenannten Minoren.

20.16. **Definition.** Seien K ein Körper, $m, n \in \mathbb{N}$, $A = (a_{ij}) \in \operatorname{Mat}(m \times n, K)$ und $0 \leq r \leq \min\{m, n\}$. Eine $r \times r$-*Untermatrix* von A ist eine Matrix der Form $(a_{i_k, j_l})_{1 \leq k,l \leq r}$, wobei $1 \leq i_1 < i_2 < \ldots < i_r \leq m$ und $1 < j_1 < j_2 < \ldots < j_r \leq n$. Man wählt also r Zeilen und r Spalten von A aus und bildet die Matrix aus den Einträgen in diesen Zeilen und Spalten. **DEF** Untermatrix Minor Hauptminor

Ein r-*Minor* von A ist die Determinante einer $r \times r$-Untermatrix von A. Im Fall $m = n$ ist ein r-*Hauptminor* von A ein r-Minor von A, sodass in der obigen Notation $i_1 = j_1$, $i_2 = j_2$, $\ldots$, $i_r = j_r$ gilt (man wählt also dieselben Zeilen- und Spaltenindizes aus). Der *führende* r-*Hauptminor* von A ist die Determinante der Untermatrix $(a_{ij})_{1 \leq i,j \leq r}$, die aus den ersten r Zeilen und Spalten von A gebildet wird. ♢

Minoren sind manchmal nützlich, um den Rang einer Matrix zu beschreiben.

20.17. **Lemma.** *Seien K ein Körper, $m, n \in \mathbb{N}$, $A \in \operatorname{Mat}(m \times n, K)$ und sei $1 \leq r \leq \min\{m, n\}$. Dann sind folgende Aussagen äquivalent:* **LEMMA** Rang über Minoren

(1) $\operatorname{rk}(A) < r$.

(2) *Alle r-Minoren von A verschwinden.*

Beweis. „(1) ⇒ (2)“: Sei A' eine $r \times r$-Untermatrix von A. Da je r Spalten von A linear abhängig sind, gilt das auch für die Spalten von A', also ist $\det A' = 0$.

„(2) ⇒ (1)“: Wir nehmen an, dass $\operatorname{rk}(A) \geq r$ ist und zeigen, dass es einen nicht verschwindenden r-Minor gibt. Nach Voraussetzung gibt es r linear unabhängige Spalten in A; sei B die $m \times r$-Matrix, die aus diesen r Spalten besteht. Dann ist $\operatorname{rk}(B) = r$, also hat B auch r linear unabhängige Zeilen. Sei A' die Matrix, die aus diesen r Zeilen von B besteht; dann ist A' eine $r \times r$-Untermatrix von A. Außerdem ist $\operatorname{rk}(A') = r$, also ist der r-Minor $\det(A')$ von A nicht null. ❑

Mit Hilfe der Minoren lassen sich auch die weiteren Koeffizienten des charakteristischen Polynoms ausdrücken, denn es gilt für eine Matrix $A \in \operatorname{Mat}(n, K)$ mit charakteristischem Polynom $p \in K[X]$:

$$p = \sum_{k=0}^{n} (-1)^k s_k(A) X^{n-k},$$

wobei $s_k(A)$ die Summe der k-Hauptminoren von A ist. Für $k = 1$ ist das gerade die Spur von A, denn die 1-Hauptminoren sind genau die Einträge auf der Diagonalen; für $k = n$ ist das die Determinante von A (der einzige n-Hauptminor). Eine Möglichkeit das einzusehen besteht darin, die Multilinearität der Determinante als Funktion (z.B.) der Zeilen einer Matrix zu verwenden (vergleiche das Kleingedruckte auf Seite 106 (LA I)). Für eine Teilmenge T von $\{1, 2, \ldots, n\}$ sei A_T die $n \times n$-Matrix, deren j-te Zeile für $j \in T$ mit der j-ten Zeile von A und für $j \notin T$ mit der j-ten Zeile von I_n übereinstimmt. Dann ist

$$p = \det(XI_n - A) = \sum_{T \subset \{1,2,\ldots,n\}} \det(-A_T) X^{n-\#T} = \sum_{T \subset \{1,2,\ldots,n\}} (-1)^{\#T} \det(-A_T) X^{n-\#T}$$

und $\det(A_T)$ ist gerade der $\#T$-Minor von A, der zu den Zeilen und Spalten mit Nummern in T gehört (wie man durch Entwicklung nach den anderen Zeilen sieht).

Wir wollen die Minoren jetzt aber benutzen, um nachzuweisen, dass eine symmetrische Matrix positiv (oder negativ) definit ist. Dafür zunächst noch ein einfaches Lemma.

20.18. **Lemma.** *Sei $A = (a_{ij}) \in \mathrm{Mat}(n, \mathbb{R})$ eine positiv definite symmetrische Matrix. Dann ist für $r \in \{1, 2, \ldots, n\}$ die Untermatrix $A' = (a_{ij})_{1\le i,j\le r}$ von A ebenfalls positiv definit.*

LEMMA
Untermatrizen erben positive Definitheit

Beweis. Sei $\mathbf{0} \neq \boldsymbol{x}' = (x_1, \ldots, x_r) \in \mathbb{R}^r$. Wir müssen zeigen, dass $(\boldsymbol{x}')^\top A' \boldsymbol{x}' > 0$ ist. Sei $\boldsymbol{x} = (x_1, x_2, \ldots, x_r, 0, \ldots, 0) \in \mathbb{R}^n$ (wir fügen also $n - r$ Nullen an). Dann ist $\boldsymbol{x} \neq \mathbf{0}$, also nach Voraussetzung $\boldsymbol{x}^\top A \boldsymbol{x} > 0$. Es genügt also zu zeigen, dass $(\boldsymbol{x}')^\top A' \boldsymbol{x}' = \boldsymbol{x}^\top A \boldsymbol{x}$ ist. Mit $x_j = 0$ für $j \in \{r+1, r+2, \ldots, n\}$ gilt

$$\boldsymbol{x}^\top A \boldsymbol{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = \sum_{i=1}^{r} \sum_{j=1}^{r} a_{ij} x_i x_j = (\boldsymbol{x}')^\top A' \boldsymbol{x}' . \qquad \square$$

* 20.19. **Satz.** *Seien $n \in \mathbb{N}$ und $A = (a_{ij}) \in \mathrm{Mat}(n, \mathbb{R})$ eine symmetrische Matrix. Für $r \in \{1, 2, \ldots, n\}$ sei $d_r(A) = \det(a_{ij})_{1\le i,j\le r}$ der führende r-Hauptminor von A. Dann gilt:*

(1) *A ist positiv definit $\iff$ $d_r(A) > 0$ für alle $r \in \{1, 2, \ldots, n\}$.*

(2) *A ist negativ definit $\iff$ $(-1)^r d_r(A) > 0$ für alle $r \in \{1, 2, \ldots, n\}$.*

SATZ
Determinantenkriterium für positiv definit

Die Bedingung für „negativ definit" heißt also $d_1(A) < 0$, $d_2(A) > 0$, $d_3(A) < 0$ usw.: Die führenden Hauptminoren alternieren im Vorzeichen. Man merkt sich das am besten an den Vorzeichen der führenden Hauptminoren von $-I_n$.

Beweis. Wir beweisen zunächst Aussage (1). Die Richtung „$\Rightarrow$" folgt aus Lemma 20.15, denn mit A sind nach Lemma 20.18 auch die Matrizen $A_r = (a_{ij})_{1\le i,j\le r}$ positiv definit, und eine positiv definite Matrix hat positive Determinante (denn die ist das Produkt der (positiven) Eigenwerte). Die Richtung „$\Leftarrow$" zeigen wir durch Induktion über n. Der Fall $n = 0$ (oder $n = 1$) ist klar. Für den Schritt von n auf $n+1$ sei $A \in \mathrm{Mat}(n+1, \mathbb{R})$ symmetrisch mit positiven führenden Hauptminoren $d_r(A)$ für alle $r \in \{1, 2, \ldots, n+1\}$. Das gilt dann entsprechend auch für die Matrix $A_n \in \mathrm{Mat}(n, \mathbb{R})$ (denn $d_r(A_n) = d_r(A)$ für $r \le n$). Nach Induktionsvoraussetzung ist A_n positiv definit. Das heißt, dass für Spaltenvektoren $\mathbf{0} \neq \boldsymbol{x} \in \langle \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \rangle \subset \mathbb{R}^{n+1}$ stets $\boldsymbol{x}^\top A \boldsymbol{x} > 0$ ist. Wir zeigen jetzt, dass A höchstens einen negativen Eigenwert haben kann: Nach Satz 20.14 gibt es $P \in \mathrm{O}(n+1)$ mit

$$P^\top A P = P^{-1} A P = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n+1})$$

diagonal. Wären wenigstens zwei Eigenwerte negativ, etwa λ_i und λ_j, mit zugehörigen Eigenvektoren $\boldsymbol{y}_i = P\mathbf{e}_i$ und $\boldsymbol{y}_j = P\mathbf{e}_j$ (als Spaltenvektoren), dann hätten wir für $(0,0) \neq (\alpha, \beta) \in \mathbb{R}^2$

$$(\alpha \boldsymbol{y}_i + \beta \boldsymbol{y}_j)^\top A (\alpha \boldsymbol{y}_i + \beta \boldsymbol{y}_j) = (\alpha \mathbf{e}_i + \beta \mathbf{e}_j)^\top P^\top A P (\alpha \mathbf{e}_i + \beta \mathbf{e}_j) = \lambda_i \alpha^2 + \lambda_j \beta^2 < 0 .$$

Da die $n+2$ Vektoren $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \boldsymbol{y}_i, \boldsymbol{y}_j \in \mathbb{R}^{n+1}$ nicht linear unabhängig sein können, gibt es $(0,0) \neq (\alpha, \beta) \in \mathbb{R}^2$ mit $\mathbf{0} \neq \boldsymbol{x} = \alpha \boldsymbol{y}_i + \beta \boldsymbol{y}_j \in \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle$. Dann müsste aber sowohl $\boldsymbol{x}^\top A \boldsymbol{x} < 0$ als auch $\boldsymbol{x}^\top A \boldsymbol{x} > 0$ gelten, ein Widerspruch. Es kann also keine zwei negativen Eigenwerte geben. Da das Produkt aller Eigenwerte $d_{n+1}(A) = \det(A)$ positiv ist, kann es auch nicht genau einen negativen Eigenwert geben (und natürlich kann null kein Eigenwert sein), also sind alle Eigenwerte von A positiv; nach Lemma 20.15 ist A also positiv definit.

Aussage (2) folgt aus Aussage (1): A ist genau dann negativ definit, wenn $-A$ positiv definit ist, und für die führenden Hauptminoren gilt $d_r(-A) = (-1)^r d_r(A)$. ❑

20.20. **Beispiele.** Wir betrachten wieder **BSP**

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Die führenden Hauptminoren von A sind $d_1(A) = 2$, $d_2(A) = 2^2 - 1^2 = 3$, was bestätigt, dass A positiv definit ist. Hingegen sind die führenden Hauptminoren von B gegeben durch $d_1(B) = 1$ und $d_2(B) = 1^2 - 2^2 = -3$, was bestätigt, dass B nicht positiv definit ist (und auch nicht negativ definit, denn dafür haben beide Minoren das falsche Vorzeichen). ♣

Ist die symmetrische Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ nur positiv semidefinit, dann folgt wie im Beweis von „$\Rightarrow$“, dass die führenden Hauptminoren von A alle ≥ 0 sein müssen. Die Umkehrung gilt dann aber im Allgemeinen nicht.

20.21. **Beispiel.** Die Matrix $A = \left(\begin{smallmatrix} 0 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ hat nichtnegative führende Hauptminoren (beide sind null), ist aber nicht positiv semidefinit (denn $\mathbf{e}_2^\top A \mathbf{e}_2 = -1$). ♣ **BSP**

Es gibt auch ein Determinanten-Kriterium für positive (oder negative) Semidefinitheit. Es lautet wie folgt.

Satz. *Seien $n \in \mathbb{N}$ und $A = (a_{ij}) \in \mathrm{Mat}(n, \mathbb{R})$ eine symmetrische Matrix. Dann gilt:* **SATZ** Determinantenkriterium für semidefinit

(1) *A ist positiv semidefinit $\iff$ $d \geq 0$ für alle Hauptminoren d von A.*

(2) *A ist negativ semidefinit $\iff$ $\forall r \in \{1, 2, \ldots, n\} \colon (-1)^r d \geq 0$ für alle r-Hauptminoren d von A.*

(3) *A ist indefinit $\iff$ es gibt einen $2r$-Hauptminor $d < 0$ von A, oder es gibt einen $(2r+1)$-Hauptminor $d > 0$ und einen $(2r'+1)$-Hauptminor $d' < 0$ von A.*

Beweis. Aussage (3) folgt formal-logisch aus (1) und (2) (A ist genau dann indefinit, wenn A weder positiv noch negativ semidefinit ist). Aussage (2) folgt aus (1) durch Anwendung von (1) auf $-A$. Es genügt also, die erste Aussage zu zeigen. Die Richtung „$\Rightarrow$“ ist wieder klar: Jede Haupt-Untermatrix von A ist positiv semidefinit, hat also nichtnegative Eigenwerte und damit nichtnegative Determinante.

Zum Beweis von „$\Leftarrow$“ nehmen wir an, dass alle Hauptminoren von A nichtnegativ sind. Wir bemerken zunächst, dass aus Lemma 20.15 folgt, dass eine symmetrische Matrix mit nicht verschwindender Determinante positiv oder negativ definit oder indefinit sein muss. Sei $K = \ker(A) \subset \mathbb{R}^n$ und $k = \dim K$. Wir wählen eine Basis $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k)$ von K. Wir können diese Basis durch Hinzunahme von $n-k$ Standard-Basisvektoren $\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \ldots, \mathbf{e}_{j_{n-k}}$ (mit $1 \leq j_1 < j_2 < \ldots < j_{n-k} \leq n$) zu einer Basis von $\mathbb{R}^n$ ergänzen (Basisergänzungssatz 9.5). Sei $V = \langle \mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \ldots, \mathbf{e}_{j_{n-k}} \rangle$; dann ist $V \cap K = \{\mathbf{0}\}$ und jeder Vektor $\boldsymbol{x} \in \mathbb{R}^n$ kann eindeutig geschrieben werden als $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_1$ mit $\boldsymbol{x}_0 \in K$ und $\boldsymbol{x}_1 \in V$. Sei $\beta \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ die symmetrische Bilinearform, deren Matrix bezüglich der Standard-Basis A ist. Dann gilt $\beta(\boldsymbol{x}, \boldsymbol{x}_0) = \beta(\boldsymbol{x}_0, \boldsymbol{x}) = 0$ für alle $\boldsymbol{x} \in \mathbb{R}^n$ und $\boldsymbol{x}_0 \in K$.

Für $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_1$ wie oben gilt also $\beta(\boldsymbol{x}, \boldsymbol{x}) = \beta(\boldsymbol{x}_1, \boldsymbol{x}_1)$. Sei A' die $(n-k) \times (n-k)$-Untermatrix von A zu den Zeilen- und Spaltenindizes $j_1, j_2, \ldots, j_{n-k}$. Dann ist A' eine Matrix der Bilinearform $\beta' = \beta|_{V \times V}$, und nach der obigen Überlegung ist A genau dann positiv semidefinit, wenn das für A' gilt. Außerdem ist $\ker(A') = \{\mathbf{0}\}$ (wegen $K \cap V = \{\mathbf{0}\}$), also ist $\det(A') \neq 0$. Damit ist A' positiv oder negativ definit oder indefinit. Wie im Beweis von Satz 20.19 zeigt man induktiv, dass A' keine zwei negativen Eigenwerte haben kann. Wegen $\det(A') > 0$ (hier verwenden wir die Voraussetzung) müssen alle Eigenwerte von A' positiv sein. Damit ist A' positiv definit, also ist A positiv semidefinit. ❑

Dieses Kriterium ist sehr viel weniger nützlich als Satz 20.19: Es gibt 2^n Hauptminoren, aber nur n führende Hauptminoren. Der Aufwand dafür, *alle* Hauptminoren zu testen, wird also schon für relativ kleine n zu groß, um praktikabel zu sein. Zum Glück gibt es bessere Möglichkeiten. Wir werden darauf noch genauer eingehen.

21. Euklidische und unitäre Vektorräume

Wir haben am Ende des letzten Abschnitts schon damit begonnen, von der „allgemeinen“ linearen Algebra über beliebigen Grundkörpern etwas wegzugehen und Resultate für den speziellen Körper $\mathbb{R}$ 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.

∗ 21.1. **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.
euklidischer Vektorraum

„Euklidisch“ nach Euklid von Alexandria, da man in euklidischen Vektorräumen euklidische Geometrie betreiben kann.

Um Verwechslungen zu vermeiden, notieren wir den von einer Menge A erzeugten Untervektorraum als $\langle A\rangle_{\mathbb{R}}$.

21.2. **Beispiele.**

BSP
eukl. VR

- Das *Standard-Skalarprodukt* $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^\top \boldsymbol{y}$ auf $\mathbb{R}^n$ (mit Spaltenvektoren $\boldsymbol{x}$, $\boldsymbol{y}$) 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. ♣


C. Hermite
1822–1901

Wir wollen jetzt eine analoge Definition für komplexe (statt reelle) Vektorräume formulieren. Eine symmetrische Bilinearform (wie im reellen Fall) können wir nicht verwenden, denn eine symmetrische Bilinearform β auf einem komplexen Vektorraum kann nicht positiv definit sein (man kann nicht einmal erreichen, dass $\beta(v, v)$ stets reell ist), denn
$$\beta(\boldsymbol{i}v, \boldsymbol{i}v) = \boldsymbol{i}^2\beta(v, v) = -\beta(v, v)\,.$$
Um das zu verhindern, modifizieren wir die Eigenschaften, die wir fordern.

∗ 21.3. **Definition.** Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $\beta\colon V \times V \to \mathbb{C}$ 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
$$\beta(v_1 + v_1', v_2) = \beta(v_1, v_2) + \beta(v_1', v_2), \quad \beta(\lambda v_1, v_2) = \lambda\beta(v_1, v_2);$$
$$\beta(v_1, v_2 + v_2') = \beta(v_1, v_2) + \beta(v_1, v_2'), \quad \beta(v_1, \lambda v_2) = \bar{\lambda}\beta(v_1, v_2)\,.$$
Eine Sesquilinearform β auf V heißt *hermitesch*, wenn zusätzlich für alle $v_1, v_2 \in V$ gilt
$$\beta(v_2, v_1) = \overline{\beta(v_1, v_2)}\,.$$
♢

DEF
Sesquilinearform
hermitesch

„Hermitesch" nach Charles Hermite.

„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 $z = x + y\boldsymbol{i} \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ ist $\bar{z} = x - y\boldsymbol{i}$. Dann gilt für $z, z_1, z_2 \in \mathbb{C}$

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2\,, \qquad \overline{z_1 z_2} = \bar{z}_1\,\bar{z}_2 \quad \text{und} \quad z\bar{z} = |z|^2\,.$$

z ist genau dann reell, wenn $z = \bar{z}$ ist.

Für eine hermitesche Sesquilinearform β 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.

∗ 21.4. **Definition.** Sei V ein komplexer Vektorraum und β eine hermitesche Sesquilinearform auf V. β 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 $\langle v_1, v_2 \rangle$. ♢

21.5. **Beispiele.**

BSP unitäre Vektorräume

- Das Standardbeispiel ist $\mathbb{C}^n$ mit dem Standard-Skalarprodukt
$$\big\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \big\rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n\,.$$
- Auch das Beispiel aus der Analysis lässt sich übertragen: Der Raum $\mathcal{C}([a, b], \mathbb{C})$ der stetigen Funktionen $[a, b] \to \mathbb{C}$ mit dem Skalarprodukt
$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\,dx\,.$$
ist ein unitärer Vektorraum. ♣

Damit können wir jetzt Längen und Winkel einführen.

∗ 21.6. **Definition.** Sei V ein euklidischer oder unitärer 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 $\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \ldots + x_n^2}$ die übliche euklidische Länge eines Vektors. Im $\mathbb{C}^n$ gilt analog $\|(x_1, \ldots, x_n)\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$. Die Standardbasis besteht jeweils aus Einheitsvektoren.

Wir beweisen einige Eigenschaften der Länge.

SATZ
Cauchy-Schwarz
Dreiecksungl.

∗ **21.7. Satz.** *Sei V ein euklidischer oder unitärer Vektorraum.*

(1) *Für $v \in V$ und $\lambda \in \mathbb{R}$ bzw. $\mathbb{C}$ gilt $\|\lambda v\| = |\lambda|\|v\|$.*

(2) *(Cauchy-Schwarzsche Ungleichung) Für $v, w \in V$ gilt $|\langle v, w\rangle| \le \|v\|\|w\|$ mit Gleichheit genau dann, wenn v und w linear abhängig sind.*

(3) *(Dreiecksungleichung) Für $v, w \in V$ gilt $\|v+w\| \le \|v\|+\|w\|$ mit Gleichheit genau dann, wenn $v = \lambda w$ oder $w = \lambda v$ ist mit $\lambda \in \mathbb{R}_{\ge 0}$.*

A.-L. Cauchy
1789–1857

H.A. Schwarz
1843–1921

Beweis.

(1) $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda\bar{\lambda}\langle v, v\rangle} = \sqrt{|\lambda|^2\langle v, v\rangle} = |\lambda|\sqrt{\langle v, v\rangle} = |\lambda|\|v\|$.

(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

$$0 \le \langle v', v'\rangle = \langle v', v\rangle = \langle v, v\rangle - \frac{\langle v, w\rangle}{\|w\|^2}\langle w, v\rangle = \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2}\,,$$

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

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&= \langle v, v\rangle + \langle v, w\rangle + \overline{\langle v, w\rangle} + \langle w, w\rangle = \|v\|^2 + 2\,\mathrm{Re}(\langle v, w\rangle) + \|w\|^2 \\
&\le \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \le \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2\,.
\end{aligned}$$

Die Ungleichung folgt. Gleichheit ist äquivalent zu $\langle v, w\rangle = \|v\|\|w\|$; dafür müssen v und w linear abhängig sein und damit das Skalarprodukt links positiv ist (beachte $\langle \lambda w, w\rangle = \lambda\|w\|^2$) muss der Skalarfaktor reell und ≥ 0 sein. ❑

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 (für $v, w \neq \mathbf{0}$)

$$(v, w) \longmapsto d(v, w) = \|v - w\|$$

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.

* 21.8. **Definition.** Sei V ein euklidischer oder unitärer Vektorraum. **DEF** Winkel orthogonal orthonormal ONB

(1) Im euklidischen Fall ist für zwei Vektoren $v, w \in V$ mit $v, w \neq \mathbf{0}$ 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

$$U^\perp = \{v \in V \mid \forall w \in U \colon v \perp w\}$$

das *orthogonale Komplement* von U in V.

(4) Eine Teilmenge $A \subset V$ heißt *orthogonal*, wenn ihre Elemente paarweise orthogonal sind ($\forall v, w \in A \colon v \neq w \Rightarrow v \perp w$). A heißt *orthonormal*, wenn zusätzlich alle Elemente von A Einheitsvektoren sind.

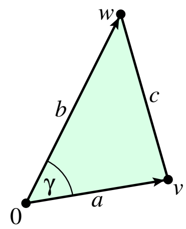
(5) Eine Basis B von V heißt eine *Orthonormalbasis* oder kurz *ONB* von V, wenn sie aus paarweise orthogonalen Einheitsvektoren besteht (wenn also die Menge der Vektoren in B orthonormal ist). ♢

Die Cauchy-Schwarzsche Ungleichung stellt sicher, dass die Definition des Winkels sinnvoll ist, denn es gilt ja für jedes euklidische Skalarprodukt

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|\|w\|} \leq 1\,.$$

Zwei Vektoren in einem euklidischen Vektorraum sind genau dann orthogonal, wenn wenigstens einer der Nullvektor ist oder der Winkel zwischen ihnen $\pi/2$ (entsprechend $90°$) ist.

21.9. **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 γ. Dann gilt $\langle v, w\rangle = \|v\|\|w\| \cos\gamma$, also **BSP** Cosinussatz Pythagoras



$$c^2 = \|v - w\|^2 = \|v\|^2 - 2\langle v, w\rangle + \|w\|^2 = a^2 - 2ab\cos\gamma + b^2\,;$$

das ist der *Cosinussatz*. Gilt $v \perp w$ (dann ist γ ein rechter Winkel), dann vereinfacht sich das zum *Satz des Pythagoras*

$$c^2 = a^2 + b^2\,.$$ ♣

Die Standardbasis ist eine Orthonormalbasis des Standardraums $\mathbb{R}^n$ oder $\mathbb{C}^n$.

21.10. **Lemma.** *Sei V ein euklidischer oder unitärer Vektorraum und seien die Vektoren $v_1, v_2, \ldots, v_n \in V$ paarweise orthogonal und von $\mathbf{0}$ verschieden. Dann ist $(v_1, v_2, \ldots, v_n)$ linear unabhängig.* **LEMMA** orthogonal ⇒ lin.unabh.

Beweis. Seien $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ bzw. $\mathbb{C}$ mit $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = \mathbf{0}$. Es folgt für alle $j \in \{1, 2, \ldots, n\}$:

$$0 = \langle \mathbf{0}, v_j\rangle = \Big\langle \sum_{i=1}^n \lambda_i v_i, v_j \Big\rangle = \sum_{i=1}^n \lambda_i \langle v_i, v_j\rangle = \lambda_j \|v_j\|^2\,,$$

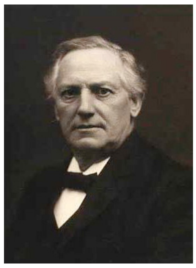
und weil $v_j \neq \mathbf{0}$ ist, muss $\lambda_j = 0$ sein. ❑

Gibt es immer eine Orthonormalbasis? Der folgende wichtige Satz zeigt, dass man aus jeder endlichen Basis eine Orthonormalbasis konstruieren kann.

∗ **21.11. Satz.** *Sei V ein euklidischer oder unitärer Vektorraum mit gegebener Basis $B = (b_1, b_2, \ldots, b_n)$. Dann bilden die wie folgt sukzessive definierten Vektoren e_j eine ONB von V; dabei gilt $\langle e_1, e_2, \ldots, e_j\rangle_{\mathbb{K}} = \langle b_1, b_2, \ldots, b_j\rangle_{\mathbb{K}}$ für alle $j \in \{0, 1, \ldots, n\}$ (mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$).*

SATZ Gram-Schmidt-Orthonormalisierung

$$\begin{aligned}
e_1 &= \frac{1}{\|v_1\|} v_1 \quad \textit{mit} \quad v_1 = b_1 \\
e_2 &= \frac{1}{\|v_2\|} v_2 \quad \textit{mit} \quad v_2 = b_2 - \langle b_2, e_1\rangle e_1 \\
e_3 &= \frac{1}{\|v_3\|} v_3 \quad \textit{mit} \quad v_3 = b_3 - \langle b_3, e_1\rangle e_1 - \langle b_3, e_2\rangle e_2 \\
&\vdots \qquad \vdots \\
e_n &= \frac{1}{\|v_n\|} v_n \quad \textit{mit} \quad v_n = b_n - \sum_{j=1}^{n-1} \langle b_n, e_j\rangle e_j
\end{aligned}$$



J.P. Gram
1850–1916

E. Schmidt
1876–1959
(© Konrad Jacobs)

Dieses Verfahren ist nach Jørgen Pedersen Gram und Erhard Schmidt benannt.

Beweis. Zunächst einmal gilt $\langle e_1, \ldots, e_j\rangle = \langle b_1, \ldots, b_j\rangle$ für alle $j \in \{0, 1, \ldots, n\}$. Wir beweisen das durch Induktion über j (bis $j = n$). Die Aussage ist klar für $j = 0$ (auf beiden Seiten steht der Null-Vektorraum). Sie sei jetzt gezeigt für $j - 1$. Es ist

$$v_j = b_j - \sum_{i=1}^{j-1} \langle b_j, e_i\rangle e_i \in \langle e_1, \ldots, e_{j-1}, b_j\rangle = \langle b_1, \ldots, b_j\rangle$$

und wegen $b_j \notin \langle b_1, \ldots, b_{j-1}\rangle$ (denn die b_i bilden eine Basis, sind also linear unabhängig) ist $v_j \neq \mathbf{0}$. Damit ist e_j wohldefiniert und $e_j \in \langle b_1, \ldots, b_j\rangle$, woraus $\langle e_1, \ldots, e_j\rangle \subset \langle b_1, \ldots, b_j\rangle$ folgt. Umgekehrt ist

$$b_j = \sum_{i=1}^{j-1} \langle b_j, e_i\rangle e_i + \|v_j\| e_j \in \langle e_1, \ldots, e_j\rangle$$

und damit $\langle b_1, \ldots, b_j\rangle \subset \langle e_1, \ldots, e_j\rangle$.

Wir sehen insbesondere, dass $(e_1, \ldots, e_n)$ ein Erzeugendensystem von V ist.

Wir zeigen jetzt durch Induktion, dass $\{e_1, \ldots, e_n\}$ orthonormal ist. Nach Lemma 21.10 sind die e_j dann auch linear unabhängig, bilden also eine Basis von V.

Sei $1 \leq j \leq n$. Wir nehmen an, dass $\{e_1, \ldots, e_{j-1}\}$ orthonormal ist (das gilt trivialerweise für $j = 1$). Dann ist $v_j \perp e_i$ für alle $i < j$, denn

$$\langle v_j, e_i\rangle = \langle b_j, e_i\rangle - \sum_{k=1}^{j-1} \langle b_j, e_k\rangle\langle e_k, e_i\rangle = \langle b_j, e_i\rangle - \langle b_j, e_i\rangle = 0\,.$$

Wir haben schon gesehen, dass $v_j \neq \mathbf{0}$ ist. Damit ist e_j ein Einheitsvektor und (als skalares Vielfaches von v_j) ebenfalls orthogonal zu $e_1, \ldots, e_{j-1}$. ❑

Man kann das von endlichen auf abzählbar unendliche Basen verallgemeinern (Übung).

Wenn man das Verfahren auf eine Basis des $\mathbb{R}^n$ anwendet, dann lässt sich die Beziehung zwischen der ursprünglichen Basis und der daraus konstruierten Orthonormalbasis in der Form $B = PT$ schreiben, wobei in den Spalten von B die gegebenen Basisvektoren $b_1, b_2, \ldots, b_n$ und in den Spalten von P die neuen Basisvektoren $e_1, e_2, \ldots, e_n$ stehen. Weil die neue Basis eine ONB ist, ist P eine orthogonale Matrix. Die Matrix T

ist eine obere Dreiecksmatrix, deren Diagonaleinträge $\|v_1\|, \|v_2\|, \ldots, \|v_n\|$ positiv sind. Satz 21.11 impliziert also, dass jede invertierbare Matrix $B \in \mathrm{GL}(n, \mathbb{R})$ als Produkt PT geschrieben werden kann mit $P \in O(n)$ und einer oberen Dreiecksmatrix T mit positiven Diagonaleinträgen. Diese Zerlegung ist sogar eindeutig bestimmt; sie ist unter dem Namen QR-Zerlegung bekannt und wird in verschiedenen Algorithmen der numerischen Mathematik benötigt. (Wenn man die Basen in umgekehrter Reihenfolge in die Matrizen einträgt oder die Zeilen statt der Spalten verwendet oder beides, dann erhält man Versionen mit unteren Dreiecksmatrizen oder/und der umgekehrten Reihenfolge TP der Faktoren.)

21.12. **Beispiel.** Wir erzeugen eine ONB aus der Basis $B = (b_1, b_2, b_3)$ von $\mathbb{R}^3$ mit $b_1 = (1, 1, 1)$, $b_2 = (1, -1, 1)$, $b_3 = (1, 0, 0)$. Wir schreiben $N(\boldsymbol{x})$ für $\|\boldsymbol{x}\|^{-1}\boldsymbol{x}$ (für Vektoren $\boldsymbol{x} \neq \mathbf{0}$). Wir erhalten **BSP ONB**

$$e_1 = N(b_1) = \frac{1}{\sqrt{3}}(1, 1, 1)$$

$$e_2 = N(b_2 - \langle e_1, b_2\rangle e_1) = N\big((\tfrac{2}{3}, -\tfrac{4}{3}, \tfrac{2}{3})\big) = \frac{1}{\sqrt{6}}(1, -2, 1)$$

$$e_3 = N(b_3 - \langle e_1, b_3\rangle e_1 - \langle e_2, b_3\rangle e_2) = N\big((\tfrac{1}{2}, 0, -\tfrac{1}{2})\big) = \frac{1}{\sqrt{2}}(1, 0, -1)$$ ♣

Wir rechtfertigen die Bezeichnung „orthogonales Komplement“ für $U^\perp$:

21.13. **Lemma.** *Seien V ein euklidischer oder unitärer 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. Also ist $U \cap U^\perp = \{\mathbf{0}\}$.

Sei jetzt $v \in V$ beliebig und $(e_1, \ldots, e_n)$ eine ONB von U (die nach Satz 21.11 existiert). Wir setzen

$$v_1 = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2 + \ldots + \langle v, e_n\rangle e_n \in U \quad \text{und} \quad v_2 = v - v_1\,.$$

Dann gilt jedenfalls $v = v_1 + v_2$. Es bleibt zu zeigen, dass $v_2 \in U^\perp$ ist. Sei dazu $u = \lambda_1 e_1 + \ldots + \lambda_n e_n \in U$. Dann gilt

$$\langle v_2, u\rangle = \sum_{j=1}^{n} \bar{\lambda}_j \langle v - v_1, e_j\rangle$$

und

$$\langle v - v_1, e_j\rangle = \langle v, e_j\rangle - \sum_{i=1}^{n} \langle v, e_i\rangle\langle e_i, e_j\rangle = \langle v, e_j\rangle - \langle v, e_j\rangle = 0\,.$$

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.

* **21.14. Satz.** *Sei V ein euklidischer oder unitärer Vektorraum mit Orthonormalbasis $(e_1, e_2, \ldots, e_n)$.* **SATZ** Parsevalsche Gleichung

(1) *Für alle $v \in V$ gilt $v = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2 + \ldots + \langle v, e_n\rangle e_n$.*

(2) *Für alle $v \in V$ gilt $\|v\|^2 = |\langle v, e_1\rangle|^2 + |\langle v, e_2\rangle|^2 + \ldots + |\langle v, e_n\rangle|^2$.*

(3) *Für alle $v, w \in V$ gilt $\langle v, w\rangle = \langle v, e_1\rangle\overline{\langle w, e_1\rangle} + \ldots + \langle v, e_n\rangle\overline{\langle w, e_n\rangle}$.*

Beweis. Sei die rechte Seite in der ersten Gleichung u, dann ist wie im Beweis von Lemma 21.13 (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 eigentliche Parsevalsche Gleichung ist eine analoge Aussage in sogenannten „Prä-Hilbert-Räumen“: Eine (auch unendliche) orthonormale Menge $A \subset V$ ist genau dann ein „vollständiges Orthonormalsystem“ (das bedeutet, dass der von A erzeugte Untervektorraum von V bezüglich der durch das Skalarprodukt definierten Topologie dicht in V ist), wenn Gleichung (2) oben in der Form $\|v\|^2 = \sum_{e\in A} |\langle v, e\rangle|^2$ für alle $v \in V$ gilt.

Die letzte Aussage im obigen Satz besagt also, dass die Matrix der Bilinearform (oder Sesquilinearform) $\langle\cdot,\cdot\rangle$ bezüglich einer ONB die Einheitsmatrix ist:

$$\langle x_1e_1 + x_2e_2 + \ldots + x_ne_n, y_1e_1 + y_2e_2 + \ldots + y_ne_n\rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n\,.$$

Für eine Sesquilinearform β ist die Matrix $A = \mathrm{Mat}_B(\beta)$ genauso definiert wie für Bilinearformen. Mit $B = (b_1, \ldots, b_n)$ gilt dann

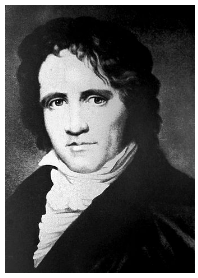
$$\beta(x_1b_1 + \ldots + x_nb_n, y_1b_1 + \ldots + y_nb_n) = (x_1\ x_2\ \cdots\ x_n)A\begin{pmatrix}\bar{y}_1\\ \bar{y}_2\\ \vdots\\ \bar{y}_n\end{pmatrix}.$$

Wenn $(e_1, \ldots, e_n)$ zwar orthonormal, aber nicht unbedingt eine Basis ist, dann gilt immerhin noch eine Ungleichung:

* **21.15. Satz.** *Sei V ein euklidischer oder unitärer Vektorraum und sei außerdem $\{e_1, e_2, \ldots, e_n\} \subset V$ eine orthonormale Menge. Dann gilt für alle $v \in V$* **SATZ** Besselsche Ungleichung

$$\|v\|^2 \geq |\langle v, e_1\rangle|^2 + |\langle v, e_2\rangle|^2 + \ldots + |\langle v, e_n\rangle|^2$$

mit Gleichheit genau für $v \in \langle e_1, e_2, \ldots, e_n\rangle_{\mathbb{K}}$ (mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$).



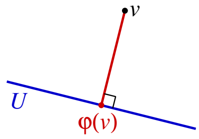
F.W. Bessel
1784–1846

Beweis. Sei $U = \langle e_1, e_2, \ldots, e_n\rangle_{\mathbb{K}}$. Wie im Beweis von Lemma 21.13 können wir $v \in V$ schreiben als $v = u + v'$ mit $u = \sum_{j=1}^n \langle v, e_j\rangle e_j \in U$ und $v' \in U^\perp$. Dann gilt $\|v\|^2 = \|u\|^2 + \|v'\|^2 = \sum_{j=1}^n |\langle v, e_j\rangle|^2 + \|v'\|^2$, wobei wir Satz 21.14 verwendet haben. Daraus folgt die Ungleichung; Gleichheit ist äquivalent mit $v' = \mathbf{0}$, also mit $v \in U$. ❑

Auch diese Ungleichung gilt allgemeiner für beliebige orthonormale Mengen oder Familien. Dies folgt aus dem oben formulierten endlichen Fall, indem man beliebige endliche Teilmengen oder -familien betrachtet.

Die Metrik auf einem euklidischen oder unitären Vektorraum V hat die schöne Eigenschaft, dass es zu einem beliebigen Vektor $v \in V$ und einem beliebigen endlich-dimensionalen Untervektorraum $U \subset V$ stets ein eindeutig bestimmtes Element $u \in U$ mit minimalem Abstand zu v gibt.

LEMMA orthogonale Projektion



21.16. **Lemma.** *Sei V ein euklidischer oder unitärer Vektorraum, sei $U \subset V$ ein endlich-dimensionaler Untervektorraum und sei $v \in V$. Dann gibt es ein eindeutig bestimmtes $u = \varphi(v) \in U$ mit*

$$d(v, u) = \|v - u\| = d(v, U) := \min\{d(v, u') \mid u' \in U\}\,.$$

Die Abbildung $\varphi\colon V \to U$ ist linear und erfüllt $\varphi|_U = \mathrm{id}_U$, also auch $\varphi|_U \circ \varphi = \varphi$.

DEF orthogonale Projektion

Die lineare Abbildung $\varphi\colon V \to U$ heißt die *orthogonale Projektion* auf U.

Beweis. Nach Lemma 21.13 können wir v eindeutig in der Form $v = u + v'$ schreiben mit $u \in U$ und $v' \in U^\perp$. Dann hat $\varphi(v) = u$ die geforderte Eigenschaft. Sei dazu $u' \in U$ beliebig. Dann gilt

$$\begin{aligned}\|v - u'\|^2 &= \|(v - u) + (u - u')\|^2 = \|v' + (u - u')\|^2 \\ &= \|v'\|^2 + \|u - u'\|^2 \ge \|v'\|^2 = \|v - u\|^2\end{aligned}$$

(beachte, dass $v' \perp u - u'$). Das zeigt, dass u den Abstand zu v minimiert und dass für jedes $u' \neq u$ der Abstand zu v größer ist.

Die Aussagen über φ sind klar, da φ die Projektion $V \to U$ bezüglich der Zerlegung $V = U \oplus U^\perp$ ist. ❑

BSP lineare Regression

21.17. **Beispiel.** Es seien n Punkte $(x_j, y_j) \in \mathbb{R}^2$ gegeben. Wir suchen eine Gerade $y = ax + b$, die möglichst nahe an diesen Punkten vorbei läuft, wobei wir „nahe“ durch den Unterschied in der y-Koordinate messen, also durch $|y_j - (ax_j + b)|$. Gauß hat dazu die *Methode der kleinsten Quadrate* eingeführt: Wir möchten $a, b \in \mathbb{R}$ so bestimmen, dass der „Fehler“

C.F. Gauß (1777–1855)

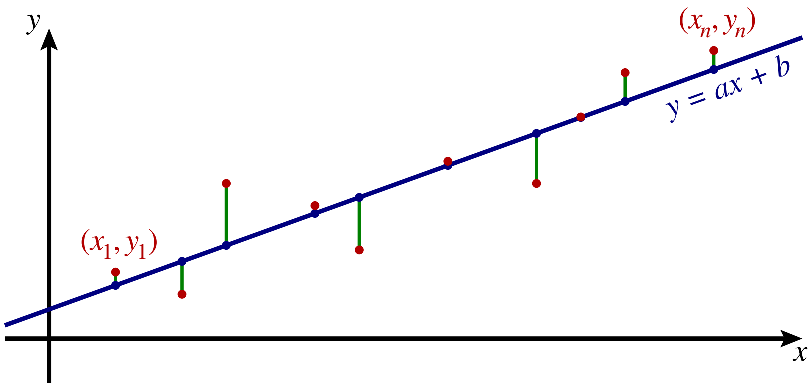
$$F(a, b) = |y_1 - (ax_1 + b)|^2 + |y_2 - (ax_2 + b)|^2 + \ldots + |y_n - (ax_n + b)|^2$$

minimal wird. Wir betrachten dazu die lineare Abbildung

$$\phi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^n\,, \quad (a, b) \longmapsto (ax_1 + b, ax_2 + b, \ldots, ax_n + b)\,;$$

wir setzen $U = \mathrm{im}(\phi)$. Dann ist, mit $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$,

$$F(a, b) = \|\phi(a, b) - \boldsymbol{y}\|^2\,.$$



Der Fehler wird also genau dann minimal, wenn $\phi(a, b)$ das Bild von $\boldsymbol{y}$ unter der orthogonalen Projektion auf U ist. Falls ϕ injektiv ist (das ist der Fall, sobald man mindestens zwei verschiedene x_j hat), dann ist (a, b) dadurch eindeutig bestimmt. ♣

Isomorphismen zwischen euklidischen oder unitären Vektorräumen, die zusätzlich das Skalarprodukt erhalten, haben einen besonderen Namen.

* **21.18. Definition.** Seien V und W zwei euklidische oder zwei unitäre Vektorräume. Eine Abbildung $f\colon V \to W$ 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 Skalarprodukt von W, rechts das von V.) Gibt es so eine Isometrie, dann heißen V und W *isometrisch.* ♢ **DEF** Isometrie

Das Skalarprodukt $\langle v, w \rangle$ lässt sich durch Längen ausdrücken:

$$4\langle v, w \rangle = \|v + w\|^2 - \|v - w\|^2$$

im euklidischen Fall und

$$4\langle v, w \rangle = \|v + w\|^2 + \boldsymbol{i}\|v + \boldsymbol{i}w\|^2 - \|v - w\|^2 - \boldsymbol{i}\|v - \boldsymbol{i}w\|^2$$

im unitären Fall. Daher genügt es, statt der zweiten Bedingung nur zu fordern, dass $\|f(v)\| = \|v\|$ ist für alle $v \in V$.

Man kann das im euklidischen Fall 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.

21.19. Beispiel. Ist V ein euklidischer Vektorraum mit ONB $(e_1, e_2, \ldots, e_n)$, dann ist die Abbildung **BSP** Isometrie

$$\mathbb{R}^n \longrightarrow V, \quad (x_1, x_2, \ldots, x_n) \longmapsto x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

eine Isometrie. Das ist gerade der Inhalt von Satz 21.14. Analog für einen unitären Vektorraum und die entsprechende Abbildung $\mathbb{C}^n \to V$.

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, und jeder n-dimensionale unitäre Vektorraum ist zum unitären Standard-Vektorraum $\mathbb{C}^n$ isometrisch.

Allgemein gilt: Ein Isomorphismus $V \to W$ zwischen endlich-dimensionalen euklidischen oder unitären Vektorräumen ist genau dann eine Isometrie, wenn er eine Orthonormalbasis auf eine Orthonormalbasis abbildet. ♣

22. Euklidische und unitäre Diagonalisierung

Sei V ein endlich-dimensionaler euklidischer Vektorraum. Wir hatten schon in Beispiel 20.12 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 (und funktioniert analog im unitären Fall):

22.1. **Lemma.** *Sei V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und sei $\phi \in V^*$ eine Linearform auf V. Dann gibt es ein eindeutig bestimmtes Element $w \in V$ mit $\phi(v) = \langle v, w\rangle = \overline{\langle w, v\rangle}$ 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. Im euklidischen Fall verwenden wir Beispiel 20.12: Da β nicht-ausgeartet ist, ist $\beta_R \colon V \to V^*$, $w \mapsto (v \mapsto \langle v, w\rangle)$, ein Isomorphismus. Dann ist klar, dass $w = \beta_R^{-1}(\phi)$ als einziges Element von V die gewünschte Eigenschaft hat.

Im unitären Fall haben wir auch die Abbildung $\beta_R \colon V \to V^*$, die $w \in V$ abbildet auf $v \mapsto \beta(v, w)$; β_R ist jetzt allerdings nicht linear, sondern semilinear, da $\beta_R(\lambda w) = \bar{\lambda}\beta_R(w)$ gilt (das kommt von der Semilinearität von β im zweiten Argument). Man kann das reparieren, indem man statt V^* den „konjugierten“ Vektorraum $\overline{V^*}$ betrachtet, der dieselbe zu Grunde liegende Menge und dieselbe Addition hat wie V^*, aber die geänderte Skalarmultiplikation $\lambda \cdot \phi = \bar{\lambda}\phi$ (links die Skalarmultiplikation von $\overline{V^*}$, rechts die von V^*). Dann ist $\beta_R \colon V \to \overline{V^*}$ eine lineare Abbildung zwischen komplexen Vektorräumen derselben endlichen Dimension. Wie im euklidischen Fall sieht man, dass $\ker(\beta_R) = \{\mathbf{0}\}$ ist; damit ist β_R ein Isomorphismus und wir können wie im euklidischen Fall schließen, dass $w = \beta_R^{-1}(\phi)$ das einzige Element mit der gewünschten Eigenschaft ist. ❑

Für β_L haben wir im unitären Fall einen Isomorphismus $V \to \bar{V}^*$ von V auf den Dualraum des konjugierten Vektorraums $\bar{V}$.

Für einen endlich-dimensionalen euklidischen Vektorraum gibt es also einen *kanonischen* Isomorphismus $V \to V^*$. Ist $(e_1, e_2, \ldots, e_n)$ eine Orthonormalbasis von V und $(e_1^*, e_2^*, \ldots, e_n^*)$ die dazu duale Basis von V^*, dann identifiziert dieser Isomorphismus e_j mit e_j^*, denn

$$\langle e_i, e_j\rangle = \delta_{ij} = e_j^*(e_i)\,.$$

Im unitären Fall ergibt sich ein Isomorphismus $V \to \overline{V^*}$.

22.2. **Beispiel.** Sei $V = \mathbb{R}^3$ (Elemente als Spaltenvektoren) mit dem Standard-Skalarprodukt. Seien $v_1, v_2 \in V$. Dann ist $v \mapsto \det(v_1, v_2, v)$ (das bezeichne die Determinante der Matrix mit den Spalten v_1, v_2, v) eine Linearform auf V, also gibt es nach Lemma 22.1 einen eindeutig bestimmten Vektor $v_1 \times v_2 \in V$ mit

BSP Vektorprodukt

$$\langle v_1 \times v_2, v\rangle = \det(v_1, v_2, v) \qquad \text{für alle } v \in V\,.$$

Dieser Vektor $v_1 \times v_2$ heißt das *Vektorprodukt* oder *Kreuzprodukt* von v_1 und v_2. Das Vektorprodukt hat folgende Eigenschaften:

DEF Vektorprodukt

(1) Die Abbildung $V \times V \longrightarrow V$, $(v_1, v_2) \longmapsto v_1 \times v_2$, ist bilinear.

Das folgt aus der Linearität der Determinante in jeder Spalte der Matrix.

(2) Sind $v_1, v_2 \in V$ linear abhängig, dann ist $v_1 \times v_2 = \mathbf{0}$.

Das folgt daraus, dass dann $\det(v_1, v_2, v) = 0$ ist für alle v.

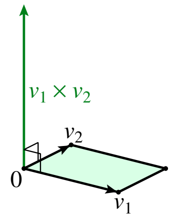
(3) Für alle $v_1, v_2 \in V$ gilt $(v_1 \times v_2) \perp v_1$ und $(v_1 \times v_2) \perp v_2$.

$\langle v_1 \times v_2, v_1 \rangle = \det(v_1, v_2, v_1) = 0$, ebenso für v_2.

(4) Es gilt die explizite Formel

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Beweis als Übung.



(5) Für alle $v_1, v_2 \in V \setminus \{\mathbf{0}\}$ gilt $\|v_1 \times v_2\| = \|v_1\|\|v_2\| \sin \angle(v_1, v_2)$.

Beweis als Übung. Zu zeigen ist $\|v_1 \times v_2\|^2 + \langle v_1, v_2\rangle^2 = \|v_1\|^2\|v_2\|^2$.

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 14.21), denn

$$\det(v_1, v_2, v_1 \times v_2) = \langle v_1 \times v_2, v_1 \times v_2 \rangle = \|v_1 \times v_2\|^2 > 0\,. \qquad \clubsuit$$

Der $\mathbb{R}^3$ zusammen mit dem Vektorprodukt ist ein Beispiel für eine sogenannte *Lie-Algebra* (nach Sophus Lie). Eine Lie-Algebra über einem Körper K ist ein K-Vektorraum V zusammen mit einer bilinearen Abbildung $V \times V \to V$, die üblicherweise $(v, w) \mapsto [v, w]$ geschrieben wird und die folgenden Bedingungen erfüllt:

- $\forall v \in V\colon [v, v] = \mathbf{0}$.
- $\forall v_1, v_2, v_3 \in V\colon [v_1, [v_2, v_3]] + [v_2, [v_3, v_1]] + [v_3, [v_1, v_2]] = \mathbf{0}$ (Jacobi-Identität).

Am einfachsten rechnet man die Jacobi-Identität für das Vektorprodukt nach, indem man sich überlegt, dass ihre Gültigkeit erhalten bleibt, wenn man einen Vektor skaliert oder ein skalares Vielfaches eines Vektors zu einem anderen addiert (also unter elementaren Spaltenumformungen an der aus den drei Vektoren als Spalten gebildeten Matrix). Damit kann man den Beweis auf die beiden trivialen Fälle $v_1 = \mathbf{0}$ und $(v_1, v_2, v_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ reduzieren.

Ein anderes Beispiel einer Lie-Algebra ist $\mathrm{Mat}(n, K)$ mit $[A, B] = AB - BA$.

Wir schreiben weiterhin $\mathbb{K}$ für $\mathbb{R}$ (im euklidischen Kontext) oder $\mathbb{C}$ (im unitären Kontext).

22.3. **Folgerung.** *Seien V und W euklidische oder unitäre Vektorräume mit $\dim V < \infty$ und sei $f\colon V \to W$ linear. Dann gibt es eine eindeutig bestimmte lineare Abbildung $f^*\colon W \to V$ mit* **FOLG** adjungierte Abbildung

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle \qquad \text{für alle } v \in V \text{ und } w \in W.$$

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 22.1 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 $f^*\colon W \to V$. Es bleibt zu zeigen, dass f^* linear ist. Seien $w, w' \in W$. Dann gilt für alle $v \in V$:

$$\begin{aligned} \langle v, f^*(w + w')\rangle &= \langle f(v), w + w'\rangle = \langle f(v), w\rangle + \langle f(v), w'\rangle \\ &= \langle v, f^*(w)\rangle + \langle v, f^*(w')\rangle = \langle v, f^*(w) + f^*(w')\rangle\,; \end{aligned}$$

die Eindeutigkeitsaussage in Lemma 22.1 zeigt dann $f^*(w+w') = f^*(w)+f^*(w')$. Analog für die Skalarmultiplikation: Seien $\lambda \in \mathbb{K}$ und $w \in W$. Dann gilt für alle $v \in V$:

$$\langle v, f^*(\lambda w)\rangle = \langle f(v), \lambda w\rangle = \bar{\lambda}\langle f(v), w\rangle = \bar{\lambda}\langle v, f^*(w)\rangle = \langle v, \lambda f^*(w)\rangle\,,$$

also wie eben $f^*(\lambda w) = \lambda f^*(w)$. ❑

Die führt auf folgende Begriffsbildung.

* **22.4. Definition.** Seien V und W euklidische oder unitäre Vektorräume und sei $f: V \to W$ linear. **DEF** (selbst-) adjungierte Abbildung normale Abbildung

(1) Gibt es eine lineare Abbildung $f^*: W \to V$, 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 $f \in \mathrm{End}(V)$ 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 $f \in \mathrm{End}(V)$ eine adjungierte Abbildung f^* und gilt $f \circ f^* = f^* \circ f$, dann heißt f *normal*. ♢

Folgerung 22.3 besagt, dass es für V endlich-dimensional stets adjungierte Abbildungen gibt.

22.5. Lemma. *Seien V_1, V_2 und V_3 euklidische oder unitäre Vektorräume und seien $f, g: V_1 \to V_2$ und $h: V_2 \to V_3$ linear, mit adjungierten Abbildungen f^*, g^* und h^* (das gilt z.B., wenn die Vektorräume endlich-dimensional sind), und sei $\lambda \in \mathbb{K}$. Dann gilt:* **LEMMA** Eigenschaften von f^*

(1) *$f+g$ und λf haben adjungierte Abbildungen; es ist $(f+g)^* = f^*+g^*$ und $(\lambda f)^* = \bar{\lambda} f^*$.*

(2) *$h \circ f$ hat eine adjungierte Abbildung; es gilt $(h\circ f)^* = f^* \circ h^*$.*

(3) *f^* hat eine adjungierte Abbildung; es ist $(f^*)^* = f$.*

(4) *f ist eine Isometrie $\iff$ f ist ein Isomorphismus mit $f^{-1} = f^*$.*

Insbesondere sind selbst-adjungierte Abbildungen und Isometrien normal.

Beweis.

(1) Das geht ähnlich wie beim Nachweis der Linearität von f^* im Beweis von Folgerung 22.3: Für $v \in V_1$, $w \in V_2$ ist

$$\begin{aligned}\langle (f+g)(v), w\rangle &= \langle f(v)+g(v), w\rangle = \langle f(v), w\rangle + \langle g(v), w\rangle \\ &= \langle v, f^*(w)\rangle + \langle v, g^*(w)\rangle = \langle v, f^*(w)+g^*(w)\rangle\,,\end{aligned}$$

woraus $(f+g)^* = f^*+g^*$ folgt, und für $v \in V_1$, $\lambda \in \mathbb{K}$, $w \in V_2$ ist

$$\langle \lambda f(v), w\rangle = \lambda\langle f(v), w\rangle = \lambda\langle v, f^*(w)\rangle = \langle v, \bar{\lambda} f^*(w)\rangle\,,$$

und damit $(\lambda f)^* = \bar{\lambda} f^*$.

(2) Für $v \in V_1$, $w \in V_3$ gilt

$$\begin{aligned}\langle (h \circ f)(v), w\rangle &= \langle h(f(v)), w\rangle = \langle f(v), h^*(w)\rangle \\ &= \langle v, f^*(h^*(w))\rangle = \langle v, (f^* \circ h^*)(w)\rangle\,,\end{aligned}$$

also ist $(h \circ f)^* = f^* \circ h^*$.

(3) Für $v \in V_2$, $w \in V_1$ ist

$$\langle f^*(v), w\rangle = \overline{\langle w, f^*(v)\rangle} = \overline{\langle f(w), v\rangle} = \langle v, f(w)\rangle\,;$$

das bedeutet $(f^*)^* = f$.

(4) Für alle $v, w \in V_1$ gilt

$$\langle f(v), f(w)\rangle = \langle v, f^*(f(w))\rangle\,,$$

und das ist genau dann dasselbe wie $\langle v, w\rangle$ für alle $v, w \in V_1$, wenn $f^* \circ f = \mathrm{id}_{V_1}$ ist. Ist f eine Isometrie, dann ist also f ein Isomorphismus und es gilt $f^* \circ f = \mathrm{id}_{V_1}$, also ist $f^{-1} = f^*$. Ist umgekehrt f ein Isomorphismus mit $f^{-1} = f^*$, dann folgt $f^* \circ f = \mathrm{id}_{V_1}$ und f ist eine Isometrie. ❑

Wir definieren analoge Begriffe für Matrizen. Die folgende Definition wiederholt im Wesentlichen Definition 20.13.

∗ 22.6. **Definition.** Eine Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ heißt *orthogonal*, wenn sie die Gleichung $A^{-1} = A^\top$ erfüllt. Wir schreiben $\mathrm{O}(n)$ für die Gruppe (!) der orthogonalen $n \times n$-Matrizen; $\mathrm{O}(n)$ heißt die *orthogonale Gruppe*.

DEF orthogonale Matrix

Die orthogonalen Matrizen mit Determinante 1 bilden ebenfalls eine Gruppe, die *spezielle orthogonale Gruppe* $\mathrm{SO}(n)$. ♢

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.

∗ 22.7. **Definition.** Für eine Matrix $A = (a_{ij}) \in \mathrm{Mat}(m \times n, \mathbb{C})$ sei $\bar{A} = (\bar{a}_{ij}) \in \mathrm{Mat}(m \times n, \mathbb{C})$ und $A^* = \bar{A}^\top$.

DEF hermitesche, unitäre, normale Matrix

Eine Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ heißt *hermitesch*, wenn $A = A^*$ ist. A heißt *unitär*, wenn $AA^* = I_n$ ist, und *normal*, wenn $AA^* = A^*A$ gilt.

Die Gruppe (!) der unitären $n \times n$-Matrizen heißt die *unitäre Gruppe* und wird mit $\mathrm{U}(n)$ bezeichnet. Die unitären Matrizen mit Determinante 1 bilden ebenfalls eine Gruppe, die *spezielle unitäre Gruppe* $\mathrm{SU}(n)$. ♢

Wie im euklidischen Fall ist eine quadratische Matrix genau dann unitär, wenn ihre Spalten (oder Zeilen) eine ONB von $\mathbb{C}^n$ bilden. 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 $|\det(A)| = 1$.

Die folgende Aussage zeigt, wie sich das Adjungieren auf die beschreibenden Matrizen auswirkt.

22.8. Lemma. *Seien V und W zwei endlich-dimensionale euklidische oder unitäre Vektorräume und seien $B = (e_1, e_2, \ldots, e_n)$ und $B' = (e'_1, e'_2, \ldots, e'_m)$ Orthonormalbasen von V bzw. W. Für eine lineare Abbildung $f\colon V \to W$ gilt dann*

$$\operatorname{Mat}_{B',B}(f^*) = \overline{\operatorname{Mat}_{B,B'}(f)}^\top \quad \textit{bzw.} \quad \operatorname{Mat}_{B,B'}(f)^* .$$

LEMMA Matrix der adjungierten Abbildung

Beweis. Seien $A = (a_{ij}) = \operatorname{Mat}_{B,B'}(f)$ und $A' = (a'_{ij}) = \operatorname{Mat}_{B',B}(f^*)$. Es gilt nach Satz 21.14

$$\begin{aligned} f(e_j) &= \langle f(e_j), e'_1\rangle e'_1 + \langle f(e_j), e'_2\rangle e'_2 + \ldots + \langle f(e_j), e'_m\rangle e'_m \quad \text{und} \\ f^*(e'_i) &= \langle f^*(e'_i), e_1\rangle e_1 + \langle f^*(e'_i), e_2\rangle e_2 + \ldots + \langle f^*(e'_i), e_n\rangle e_n , \end{aligned}$$

also ist $a_{ij} = \langle f(e_j), e'_i\rangle$ und $a'_{ji} = \langle f^*(e'_i), e_j\rangle$ und damit

$$a_{ij} = \langle f(e_j), e'_i\rangle = \langle e_j, f^*(e'_i)\rangle = \overline{\langle f^*(e'_i), e_j\rangle} = \bar{a}'_{ji} .$$

Es gilt also $A' = A^*$ ($= A^\top$ im euklidischen Fall) wie behauptet. ❑

22.9. Folgerung. *Sei V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum mit Orthonormalbasis $B = (e_1, e_2, \ldots, e_n)$ und sei $f\colon V \to V$ linear. Im euklidischen Fall gilt*

FOLG Matrix für selbst-adj. Endom.

$$f \textit{ ist selbst-adjungiert} \iff \operatorname{Mat}_B(f) \textit{ ist symmetrisch}$$

und

$$f \textit{ ist eine Isometrie} \iff \operatorname{Mat}_B(f) \textit{ ist orthogonal.}$$

Im unitären Fall gilt entsprechend

$$\begin{aligned} f \textit{ ist selbst-adjungiert} &\iff \operatorname{Mat}_B(f) \textit{ ist hermitesch,} \\ f \textit{ ist normal} &\iff \operatorname{Mat}_B(f) \textit{ ist normal} \end{aligned}$$

und

$$f \textit{ ist eine Isometrie} \iff \operatorname{Mat}_B(f) \textit{ ist unitär.}$$

Beweis. Die Aussagen über selbst-adjungierte f folgen direkt aus Lemma 22.8. Die Aussagen über Isometrien folgen mit Lemma 22.5. ❑

Wir wollen uns jetzt mit der Frage beschäftigen, wann es für einen Endomorphismus f eines euklidischen oder unitären Vektorraums V eine Orthonormalbasis von V gibt, die aus Eigenvektoren von f besteht. Man sagt dann, f sei *orthogonal diagonalisierbar* bzw. *unitär diagonalisierbar*.

DEF orthogonal/ unitär diagonalisierbar

Da sich die beiden Fälle hier doch stärker unterscheiden, behandeln wir zunächst den euklidischen Fall. Das Resultat wird sein, dass genau die selbst-adjungierten Endomorphismen orthogonal diagonalisierbar sind. Wir zeigen zuerst die einfachere Richtung.

22.10. Satz. *Sei V ein endlich-dimensionaler euklidischer Vektorraum und sei $f \in \mathrm{End}(V)$. 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 $A = \mathrm{Mat}_B(f)$ eine Diagonalmatrix ist. Dann gilt auch $A = A^\top$, also ist f nach Folgerung 22.9 selbst-adjungiert. ❑

Zum Beweis der Gegenrichtung machen wir eine Vorüberlegung, für die wir auf einige Grundtatsachen aus der Analysis zurückgreifen. 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 $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| = 1\}$ (also die Oberfläche der n-dimensionalen Einheitskugel) ist eine abgeschlossene und beschränkte Teilmenge von $\mathbb{R}^n$, also ist S kompakt (Satz von Heine-Borel). Die Abbildung

$$h\colon \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \boldsymbol{x} \longmapsto \langle f(\boldsymbol{x}), \boldsymbol{x}\rangle$$

ist stetig, denn $f\colon \mathbb{R}^n \to \mathbb{R}^n$ ist stetig und das Skalarprodukt ist ebenfalls stetig. Als stetige Funktion nimmt die Abbildung h auf der kompakten Menge S ihr Maximum an, etwa in $\boldsymbol{x}_0 \in S$.

22.11. Lemma. *In der eben diskutierten Situation (mit $f\colon \mathbb{R}^n \to \mathbb{R}^n$ selbst-adjungiert) ist $\boldsymbol{x}_0$ ein Eigenvektor von f zum Eigenwert $\lambda = h(\boldsymbol{x}_0)$.*

LEMMA
Maximum ist
Eigenwert

Beweis. Wir bemerken zunächst, dass für alle $\mathbf{0} \neq \boldsymbol{x} \in \mathbb{R}^n$ gilt

$$\langle f(\boldsymbol{x}), \boldsymbol{x}\rangle = \|\boldsymbol{x}\|^2 \Big\langle f\Big(\frac{1}{\|\boldsymbol{x}\|}\boldsymbol{x}\Big), \frac{1}{\|\boldsymbol{x}\|}\boldsymbol{x}\Big\rangle \le \lambda\|\boldsymbol{x}\|^2,$$

denn $\|\boldsymbol{x}\|^{-1}\boldsymbol{x} \in S$. Für $\boldsymbol{x} = \mathbf{0}$ gilt die Ungleichung $\langle f(\boldsymbol{x}), \boldsymbol{x}\rangle \le \lambda\|\boldsymbol{x}\|^2$ ebenfalls.

Wir zeigen jetzt, dass $\boldsymbol{x}_0$ ein Eigenvektor ist. Wir können $f(\boldsymbol{x}_0) = \mu\boldsymbol{x}_0 + \boldsymbol{y}$ schreiben mit $\mu \in \mathbb{R}$ und $\boldsymbol{y} \perp \boldsymbol{x}_0$ (nach Lemma 21.13 mit $U = \langle \boldsymbol{x}_0\rangle_{\mathbb{R}}$). Für $t \in \mathbb{R}$ betrachten wir den Vektor $\boldsymbol{x}_0 + t\boldsymbol{y}$. Es gilt

$$\langle f(\boldsymbol{x}_0 + t\boldsymbol{y}), \boldsymbol{x}_0 + t\boldsymbol{y}\rangle \le \lambda\|\boldsymbol{x}_0 + t\boldsymbol{y}\|^2 = \lambda(1 + t^2\|\boldsymbol{y}\|^2) = \lambda + t^2\lambda\|\boldsymbol{y}\|^2$$

(dabei haben wir $\boldsymbol{x}_0 \perp \boldsymbol{y}$ und den „Pythagoras“ benutzt). Auf der anderen Seite ist

$$\begin{aligned}
\langle f(\boldsymbol{x}_0 + t\boldsymbol{y}), \boldsymbol{x}_0 + t\boldsymbol{y}\rangle &= \langle f(\boldsymbol{x}_0) + tf(\boldsymbol{y}), \boldsymbol{x}_0 + t\boldsymbol{y}\rangle \\
&= \langle f(\boldsymbol{x}_0), \boldsymbol{x}_0\rangle + t\big(\langle f(\boldsymbol{x}_0), \boldsymbol{y}\rangle + \langle f(\boldsymbol{y}), \boldsymbol{x}_0\rangle\big) + t^2\langle f(\boldsymbol{y}), \boldsymbol{y}\rangle \\
&= \lambda + t\big(\langle f(\boldsymbol{x}_0), \boldsymbol{y}\rangle + \langle \boldsymbol{y}, f(\boldsymbol{x}_0)\rangle\big) + t^2\langle f(\boldsymbol{y}), \boldsymbol{y}\rangle \\
&= \lambda + 2t\langle f(\boldsymbol{x}_0), \boldsymbol{y}\rangle + t^2\langle f(\boldsymbol{y}), \boldsymbol{y}\rangle \\
&= \lambda + 2t\langle \mu\boldsymbol{x}_0 + \boldsymbol{y}, \boldsymbol{y}\rangle + t^2\langle f(\boldsymbol{y}), \boldsymbol{y}\rangle \\
&= \lambda + 2t\|\boldsymbol{y}\|^2 + t^2\langle f(\boldsymbol{y}), \boldsymbol{y}\rangle.
\end{aligned}$$

Dabei haben wir verwendet, dass f selbst-adjungiert und $\boldsymbol{x}_0 \perp \boldsymbol{y}$ ist. Für $t > 0$ ergibt sich daraus die Ungleichung (nach Subtraktion von λ und Division durch t)

$$0 \le 2\|\boldsymbol{y}\|^2 \le t\big(\lambda\|\boldsymbol{y}\|^2 - \langle f(\boldsymbol{y}), \boldsymbol{y}\rangle\big);$$

wenn wir t von oben gegen null gehen lassen, folgt daraus $\|\boldsymbol{y}\|^2 = 0$, also $\boldsymbol{y} = \mathbf{0}$ und damit $f(\boldsymbol{x}_0) = \mu\boldsymbol{x}_0$. Außerdem gilt

$$\lambda = \langle f(\boldsymbol{x}_0), \boldsymbol{x}_0\rangle = \langle \mu\boldsymbol{x}_0, \boldsymbol{x}_0\rangle = \mu\|\boldsymbol{x}_0\|^2 = \mu,$$

also ist λ der zu $\boldsymbol{x}_0$ gehörende Eigenwert. ❑

22.12. Folgerung. *Sei V ein endlich-dimensionaler euklidischer Vektorraum mit $\dim V > 0$ und sei $f \in \mathrm{End}(V)$ 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 $\phi\colon \mathbb{R}^n \to V$ (mit $n = \dim V$). Die Abbildung $\tilde{f} = \phi^{-1} \circ f \circ \phi \in \mathrm{End}(\mathbb{R}^n)$ ist ebenfalls selbst-adjungiert, hat also nach Lemma 22.11 einen Eigenwert $\lambda \in \mathbb{R}$. Da $\tilde{f}$ und f dieselben Eigenwerte haben (ist $\boldsymbol{x} \in \mathbb{R}^n$ Eigenvektor von $\tilde{f}$ zum Eigenwert λ, dann ist $\phi(\boldsymbol{x})$ Eigenvektor von f zum selben Eigenwert), gilt das auch für f. ❑

* **22.13. Satz.** *Sei V ein endlich-dimensionaler euklidischer Vektorraum und sei $f \in \mathrm{End}(V)$. f ist genau dann orthogonal diagonalisierbar, d.h., V besitzt eine Orthonormalbasis, die aus Eigenvektoren von f besteht, wenn f selbst-adjungiert ist.*

SATZ Spektralsatz über $\mathbb{R}$

Beweis. Die eine Richtung ist Satz 22.10. Die andere Richtung beweisen wir durch Induktion über $n = \dim V$. Für $n = 0$ ist nichts zu beweisen (die leere Familie ist eine ONB aus Eigenvektoren). Sei $n > 0$ und die Aussage für $\dim V = n-1$ richtig. Nach Folgerung 22.12 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 $\dim U = n - 1$ (denn U ist ein Komplement des eindimensionalen Unterraums $\langle e_n \rangle_{\mathbb{R}}$), und $f|_U$ ist ein selbst-adjungierter Endomorphismus von U. Nach der Induktionsannahme hat also U eine ONB $(e_1, e_2, \ldots, e_{n-1})$, die aus Eigenvektoren von f besteht. Dann ist $(e_1, \ldots, e_{n-1}, e_n)$ eine ONB von V aus Eigenvektoren von f. ❑

Für Matrizen lässt sich die interessante Richtung dieses Satzes auch so formulieren:

* **22.14. Folgerung.** *Sei $A \in \mathrm{Mat}(n, \mathbb{R})$ eine symmetrische Matrix. Dann gibt es eine orthogonale Matrix $P \in \mathrm{O}(n)$, sodass $P^\top A P = P^{-1} A P$ eine Diagonalmatrix ist.*

FOLG Spektralsatz für Matrizen über $\mathbb{R}$

Das ist Satz 20.14, den wir bereits benutzt haben, um das Determinanten-Kriterium für positive Definitheit (Satz 20.19) zu beweisen.

Beweis. Sei $f\colon \boldsymbol{x} \mapsto A\boldsymbol{x}$, dann ist A die Matrix von $f \in \mathrm{End}(\mathbb{R}^n)$ bezüglich der Standardbasis E; da A symmetrisch ist, ist f selbst-adjungiert (Folgerung 22.9). Nach Satz 22.13 hat $\mathbb{R}^n$ eine Orthonormalbasis B aus Eigenvektoren von f, also ist $D = \mathrm{Mat}_B(f)$ eine Diagonalmatrix. Die Matrix $P = \mathrm{Mat}_{B,E}(\mathrm{id}_{\mathbb{R}^n})$ hat als Spalten die Vektoren von B und ist damit orthogonal (vergleiche die Bemerkung nach Definition 22.6). Außerdem ist

$$P^{-1}AP = \mathrm{Mat}_{E,B}(\mathrm{id}_{\mathbb{R}^n})\,\mathrm{Mat}_E(f)\,\mathrm{Mat}_{B,E}(\mathrm{id}_{\mathbb{R}^n}) = \mathrm{Mat}_B(f) = D\,. \qquad ❑$$

Wir wenden uns jetzt der Frage nach der unitären Diagonalisierbarkeit zu. Es wird sich herausstellen, dass genau die normalen Endomorphismen unitär diagonalisierbar sind. Wir beginnen mit einem Lemma.

22.15. Lemma. *Sei V ein unitärer Vektorraum und sei $f \in \operatorname{End}(V)$ normal.*

LEMMA Eigenwerte und -vektoren von f^*

(1) *Es gilt $\|f^*(v)\| = \|f(v)\|$ für alle $v \in V$.*

(2) *Ist $v \in V$ ein Eigenvektor von f zum Eigenwert λ, 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 $f - \lambda \operatorname{id}_V$ normal, denn

$$\begin{aligned}(f - \lambda \operatorname{id}_V) \circ (f - \lambda \operatorname{id}_V)^* &= (f - \lambda \operatorname{id}_V) \circ (f^* - \bar{\lambda} \operatorname{id}_V) \\ &= f \circ f^* - \lambda f^* - \bar{\lambda} f + |\lambda|^2 \operatorname{id}_V \\ &= f^* \circ f - \lambda f^* - \bar{\lambda} f + |\lambda|^2 \operatorname{id}_V \\ &= (f^* - \bar{\lambda} \operatorname{id}_V) \circ (f - \lambda \operatorname{id}_V) \\ &= (f - \lambda \operatorname{id}_V)^* \circ (f - \lambda \operatorname{id}_V) .\end{aligned}$$

Aus $f(v) = \lambda v$ und Aussage (1) ergibt sich dann

$$0 = \|(f - \lambda \operatorname{id}_V)(v)\| = \|(f - \lambda \operatorname{id}_V)^*(v)\| = \|(f^* - \bar{\lambda} \operatorname{id}_V)(v)\| = \|f^*(v) - \bar{\lambda} v\| ;$$

damit ist $f^*(v) = \bar{\lambda} v$. ❑

* **22.16. Satz.** *Sei f ein Endomorphismus eines endlich-dimensionalen unitären Vektorraums V. f ist genau dann unitär diagonalisierbar, wenn f normal ist.*

SATZ Spektralsatz über $\mathbb{C}$

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 λ mit zugehörigem Eigenvektor v_n. Nach Skalieren können wir annehmen, dass $\|v_n\| = 1$ ist. Nach Lemma 22.15 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 $V = \langle v_n\rangle_{\mathbb{C}} \oplus U$. Nach Induktionsannahme hat U eine ONB $(v_1, \ldots, v_{n-1})$ aus Eigenvektoren von f; dann ist $(v_1, \ldots, v_{n-1}, v_n)$ 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):

* **22.17. Folgerung.** *Ist $A \in \mathrm{Mat}(n, \mathbb{C})$ normal, dann gibt es eine unitäre Matrix $P \in \mathrm{U}(n)$, sodass $P^{-1}AP = P^*AP$ eine Diagonalmatrix ist.* **FOLG** Spektralsatz für Matrizen über $\mathbb{C}$

Die Verallgemeinerung des reellen Spektralsatzes auf den unitären Fall charakterisiert die selbst-adjungierten Endomorphismen auch in diesem Fall:

22.18. Satz. *Ein normaler Endomorphismus eines endlich-dimensionalen unitären Vektorraums ist genau dann selbst-adjungiert, wenn er nur reelle Eigenwerte hat.* **SATZ** selbst-adj. über $\mathbb{C}$

Beweis. Sei V ein endlich-dimensionaler unitärer Vektorraum und sei $f : V \to V$ normal. Aus Satz 22.16 folgt, dass f unitär diagonalisierbar ist; sei also B eine ONB von V, die aus Eigenvektoren von f besteht. Dann ist $D = \mathrm{Mat}_B(f)$ eine Diagonalmatrix, deren Diagonaleinträge gerade die Eigenwerte von f sind; außerdem ist f genau dann selbst-adjungiert, wenn D hermitesch ist (Folgerung 22.9). Das bedeutet hier $D = D^* = \bar{D}^\top = \bar{D}$, was damit äquivalent ist, dass die Eigenwerte reell sind. ❑

Das liefert eine Möglichkeit, den reellen Spektralsatz aus der komplexen Version abzuleiten; damit kann man das Kompaktheitsargument im Beweis von Satz 22.13 (d.h., die Aussage von Lemma 22.11) umgehen: Satz 22.18 garantiert, dass alle Eigenwerte eines selbst-adjungierter Endomorphismus eines endlich-dimensionalen euklidischen Vektorraums reell sind (man kann eine zugehörige Matrix als Matrix über $\mathbb{C}$ betrachten). Das liefert die Aussage von Folgerung 22.12, die wir für den Beweis gebraucht haben. Allerdings haben wir für den komplexen Fall verwendet, dass $\mathbb{C}$ algebraisch abgeschlossen ist, was in gewisser Weise eine schwieriger zu zeigende Aussage ist als die Sätze über kompakte Mengen im $\mathbb{R}^n$. (Üblicherweise sehen Sie den ersten Beweis in der „Einführung in die Funktionentheorie“ im vierten Semester.)

23. Quadratische Formen

Aus einer bilinearen Abbildung kann man eine „quadratische“ Abbildung machen: Ist $\beta\colon V \times V \to W$ bilinear, dann hat die Abbildung

$$q\colon V \longrightarrow W, \quad v \longmapsto \beta(v, v)$$

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.

* 23.1. **Definition.** Sei V ein K-Vektorraum. Eine *quadratische Form* auf V ist eine Abbildung $q\colon V \to K$, 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 $\mathrm{Qu}(V)$.

Zwei quadratische Formen q und q' auf V heißen *äquivalent*, wenn es einen Isomorphismus $f\colon V \to V$ gibt, sodass $q' = q \circ f$ ist. ♢

Die Parallelogramm-Gleichung folgt aus den beiden Eigenschaften in der Definition.

Analog zu symmetrischen Bilinearformen definiert man positive Definitheit usw. für quadratische Formen über $\mathbb{R}$.

23.2. **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 $\mathbf{0} \neq v \in V$ 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. ♢

Für das Folgende ist es wichtig, dass wir durch 2 teilen können. Deshalb noch eine Definition.

23.3. **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 $\mathrm{char}(K)$ für die Charakteristik von K. ♢ **DEF** Charakteristik

Die Charakteristik ist entweder null oder eine Primzahl: Wäre $\mathrm{char}(K) = n$ 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 $1 + 1 = 0$.

23.4. Lemma. *Sei V ein K-Vektorraum mit* $\operatorname{char}(K) \neq 2$. *Wir schreiben* $\operatorname{Sym}(V)$ *für den Vektorraum der symmetrischen Bilinearformen auf V. Dann ist* **LEMMA** symm. bil. = quadr.

$$\operatorname{Sym}(V) \longrightarrow \operatorname{Qu}(V), \quad \beta \longmapsto \bigl(v \mapsto \beta(v,v)\bigr)$$

ein Isomorphismus.

Man kann also quadratische Formen mit den zugehörigen symmetrischen Bilinearformen identifizieren; insbesondere sind auch quadratische Formen durch symmetrische Matrizen beschrieben. Wir schreiben $\operatorname{Mat}_B(q)$ für diese Matrix; es gilt für $v = x_1 b_1 + \ldots x_n b_n$ (wenn $B = (b_1, \ldots, b_n)$ ist)

$$q(v) = \boldsymbol{x}^\top \operatorname{Mat}_B(q) \boldsymbol{x}\,,$$

wobei $\boldsymbol{x}$ der Spaltenvektor $(x_1, \ldots, x_n)^\top$ ist.

Beweis. Dass die angegebene Abbildung wohldefiniert und linear ist, ist klar. Wir zeigen, dass sie bijektiv ist, indem wir die Umkehrabbildung angeben:

$$q \longmapsto \bigl((v,w) \mapsto \tfrac{1}{2}(q(v+w) - q(v) - q(w))\bigr)$$

(hier verwenden wir $\operatorname{char}(K) \neq 2$). Nach Definition 23.1 ist das Bild eine (symmetrische) Bilinearform, also haben wir eine Abbildung $\operatorname{Qu}(V) \to \operatorname{Sym}(V,V)$. Wir prüfen nach, dass das tatsächlich die Inverse ist:

$$\begin{aligned} q &\longmapsto \bigl((v,w) \mapsto \tfrac{1}{2}(q(v+w) - q(v) - q(w))\bigr) \\ &\longmapsto \bigl(v \mapsto \tfrac{1}{2}(q(2v) - 2q(v)) = q(v)\bigr) \\ &= q \end{aligned}$$

und

$$\begin{aligned} \beta &\longmapsto \bigl(v \mapsto \beta(v,v)\bigr) \\ &\longmapsto \bigl((v,w) \mapsto \tfrac{1}{2}(\beta(v+w,v+w) - \beta(v,v) - \beta(w,w)) \\ &\qquad\qquad = \tfrac{1}{2}(\beta(v,w) + \beta(w,v)) = \beta(v,w)\bigr) \\ &= \beta\,. \end{aligned}$$

❑