Vorlesung Lineare Algebra II - SS 2023

Skript: Prof.Dr.A.Rosenschon
Dozent: Prof.Dr.W.Bley

1. Mengen und Abbildungen

Definition 1.1. Eine Menge M ist eine Zusammmenfassung von gewissen Objekten, den sogenannten Elementen von M. Die leere Menge $\emptyset$ ist die Menge, die kein Objekt enthält.

NB. Diese erste Definition ist nicht präzise; eine genaue Festlegung des Begriffs einer Menge, und der mit Mengen zulässigen Operationen, erfordert eine axiomatische Begründung der Mengenlehre, die für eine Einführung in die lineare Algebra nicht angebracht ist. Wir verwenden daher nur die obige, naive Definition.

- $m \in M$: m ist Element von M,
- $m \notin M$: m ist kein Element von M,
- $M = \{m_1, m_2, \dots\}$: M ist die Menge mit den Elemente $m_1, m_2 \dots$,
- $M = \{x \mid x \text{ erfüllt Eigenschaft } P\}$: Menge der x mit Eigenschaft P.

Beispiele 1.2. (a) $\mathbb{N} = \{1, 2, 3, \dots\}$ Menge der natürlichen Zahlen; $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ die Menge der natürlichen Zahlen einschliesslich 0.
(b) $\mathbb{Z} = \{\dots, -1, 0, 1, 2, \dots\}$ Menge der ganzen Zahlen.
(c) $\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z},\ q \neq 0\}$ Menge der rationalen Zahlen.
(d) $\mathbb{R}$ Menge der reellen Zahlen.

Definition 1.3. Sei M eine Menge.
(a) Eine Menge N ist eine Untermenge (oder Teilmenge) von M, $N \subseteq M$, falls jedes Element von N in M liegt. Ist $N \subseteq M$ und gibt es wenigstens ein $m \in M$ mit $m \notin N$, so schreibe $N \subsetneq M$. Zwei Mengen M, N sind gleich, $M = N$, wenn $N \subseteq M$ und $M \subseteq N$ gilt. Die Potenzmenge $P(M)$ ist die Menge aller Teilmengen von M; es ist $\emptyset \in P(M)$.
(b) Seien $N_j \subseteq M$, $j \in J$, J eine nichtleere Indexmenge (nicht unbedingt endlich). Die Vereinigung und der Durchschnitt der N_j sind definiert als die Mengen

$$\begin{aligned} \bigcup_{j\in J} N_j &= \{m \in M \mid m \in N_j \text{ für ein } j\}, \\ \bigcap_{j\in J} N_j &= \{m \in M \mid m \in N_j \text{ für alle } j\}. \end{aligned}$$

Ist $J = \emptyset$, so setze $\cup_j N_j = \emptyset$ und $\cap_j N_j = M$.
(c) Ist $N_j \subseteq M, j = 1, 2$, so ist die Differenz von N_1 und N_2 die Menge

$$N_1 \setminus N_2 = \{n_1 \in N_1 \mid n_1 \notin N_2\}.$$

(d) Seien $M_i \neq \emptyset, i = 1, \dots, k$ Mengen. Betrachte geordnete k-Tupel $(m_1, \dots, m_k), m_i \in M_i$, d.h. $(m_1, \dots, m_k) = (m'_1, \dots, m'_k) \Leftrightarrow m_i = m'_i$

für $i = 1, \ldots, k$. Das (kartesische) Produkt der M_i ist definiert als

$$M_1 \times \cdots \times M_k = \{(m_1, \ldots, m_k) \mid m_i \in M_i\}.$$

Seien A, B, C, N_j für $j \in J$ Teilmengen einer Menge M. Dann gilt:

- $A \cup B = B \cup A$ und $A \cap B = B \cap A$.
- $A \cup (B \cup C) = (A \cup B) \cup C$ und $A \cap (B \cap C) = (A \cap B) \cap C$.
- $A \cap (\cup_j N_j) = \cup_j (A \cap N_j)$ und $A \cup (\cap_j N_j) = \cap_j (A \cup N_j)$.

Beispiel 1.4. Sei $M = \mathbb{N}$, $N_1, N_2 \subseteq M$ die Mengen $N_1 = \{1\}$ und $N_2 = \{1, 2\}$. Dann ist $N_1 \cup N_2 = \{1, 2\}$, $N_1 \cap N_2 = \{1\}, N_1 \setminus N_2 = \emptyset, N_2 \setminus N_1 = \{2\}, N_1 \times N_2 = \{(1,1), (1,2)\}$ und $N_2 \times N_1 = \{(1,1), (2,1)\}$.

Definition 1.5. Sei M eine Menge und $N_j \subseteq M$, $j \in J$, Teilmengen. Die N_j bilden eine Partition von M, falls jedes $m \in M$ in genau einer der Teilmengen N_j liegt d.h. falls gilt

$$M = \cup_j N_j \text{ und } N_j \cap N_k = \emptyset \text{ für } j \neq k,\ j, k \in J.$$

Beispiel 1.6. Sei $N \subseteq M$ und $\overline{N} = M \setminus N$ (d.h. $\overline{N}$ ist das Komplement von N in M). Dann bilden N und $\overline{N}$ eine Partition von M.

Wir wollen Partitionen einer Menge charakterisieren. Ist $M = \cup_j N_j$ eine Partition, so nennen wir zwei Elemente $m, m' \in M$ äquivalent, $m \sim m'$, falls $m, m' \in N_j$ für ein j gilt. Für $m, m', m'' \in M$ folgt:

(i) $m \sim m$,
(ii) $m \sim m' \Rightarrow m' \sim m$,
(iii) $m \sim m'$ und $m' \sim m'' \Rightarrow m \sim m''$.

Wir zeigen, dass umgekehrt (i)-(iii) eine Partition bestimmen.

Definition 1.7. (a) Eine Relation R auf einer Menge M ist eine Teilmenge $R \subseteq M \times M$. Sind $m, m' \in M$ und gilt $(m, m') \in R$, so schreibe mRm' (m und m' stehen zueinander in Relation R).

(b) Eine Relation R auf einer Menge M ist eine Äquivalenzrelation, falls für alle Elemente $m, m', m'' \in M$ gilt:

(i) mRm (Reflexivität),
(ii) $mRm' \Rightarrow m'Rm$ (Symmetrie),
(iii) mRm' und $m'Rm'' \Rightarrow mRm''$ (Transitivität)

Ist R eine Äquivalenzrelation, so schreibe $\sim$ für R und $m \sim m'$ für mRm'; die Menge der zu einem $m \in M$ äquivalenten Elemente bildet die Äquivalenzklasse $[m]$ von m:

$$[m] = \{m' \in M \mid m \sim m'\} \subseteq M.$$

Lemma 1.8. *Sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Dann gilt für $m, m' \in M$ entweder $[m] \cap [m'] = \emptyset$ oder $[m] = [m']$;*

die verschiedenen Äquivalenzklassen bezüglich $\sim$ bilden eine Partition von M.

NB. Partition von $M \leftrightarrow$ Äquivalenzrelation auf M.

Beweis. Klar ist $\cup_m [m] \subseteq M$. Wegen (i) gilt für $m \in M$ stets $m \in [m]$, also ist auch $M \subseteq \cup_m [m]$ und somit $M = \cup_m [m]$. Sei $\emptyset \neq [m] \cap [m']$, zu zeigen ist $[m] = [m']$. Ist $m_0 \in [m] \cap [m']$, so gilt $m_0 \sim m$ und $m_0 \sim m'$. Sei $m_1 \in [m]$, d.h. $m_1 \sim m$. Wegen (ii) folgt aus $m_0 \sim m$ auch $m \sim m_0$, und (iii) angewandt auf $m_1 \sim m$ und $m \sim m_0$ liefert $m_1 \sim m_0$. Nochmalige Anwendung von (iii) auf $m_1 \sim m_0$ und $m_0 \sim m'$ zeigt $m_1 \sim m'$, also ist $m_1 \in [m']$ und $[m] \subseteq [m']$. Aus Symmetriegründen (vertauschen der Rollen von m und m') folgt genauso die umgekehrte Inklusion $[m'] \subseteq [m]$, d.h. es gilt $[m] = [m']$. $\square$

Beispiele 1.9. (a) Sei $M = \mathbb{R}^2$ und $L \subseteq M$ die Menge der Geraden in M. Für $l_1, l_2 \in L$ definiert

$$l_1 \sim l_2 \Leftrightarrow l_1 \mid\mid l_2 \quad \text{(d.h. die Geraden sind parallel)}$$

eine Äquivalenzrelation auf L.

(b) Sei $M = \mathbb{Z}$ und $m \geq 1$ eine ganze Zahl. Für $n_1, n_2 \in \mathbb{Z}$ definiere

$$n_1 \equiv n_2 (mod\ m) \Leftrightarrow m | (n_1 - n_2) \Leftrightarrow km = n_1 - n_2 \text{ für ein } k \in \mathbb{Z}.$$

Dann definiert $\equiv$ eine Äquivalenzrelation auf $\mathbb{Z}$:

(i) $n \equiv n (mod\ m)$, wegen $m | n - n = 0$,

(ii) $n_1 \equiv n_2 (mod\ m)$ heisst $n_1 - n_2 = km$ für ein $k \in \mathbb{Z}$; in diesem Fall folgt $-km = n_2 - n_1$, d.h. $n_2 \equiv n_1 (mod\ m)$.

(iii) Ist $n_1 \equiv n_2 (mod\ m)$, $km = n_1 - n_2$ und $n_2 \equiv n_3 (mod\ m)$, $lm = n_2 - n_3$, so folgt $(k+l)m = n_1 - n_3$, also $n_1 \equiv n_3 (mod\ m)$.

Sei $\mathbb{Z}/m\mathbb{Z}$ die Menge der Äquivalenzklassen dieser Äquivalenzrelation

$$\mathbb{Z}/m\mathbb{Z} = \{[r] \mid r \in \mathbb{Z}\}.$$

Nach Lemma 1.8 bilden die *verschiedenen* Äquivalenzklassen eine Partition von $\mathbb{Z}$; dies sind die Mengen $[r]$ für $0 \leq r < m$ (d.h. die verschiedenen Äquivalenzklassen entsprechen den möglichen Resten bei 'Division durch m' und werden daher auch 'Restklassen' genannt).

Ist $m = 2$, so besteht die Restklasse $[0]$ aus den geraden und die Restklasse $[1]$ aus den ungeraden ganzen Zahlen, d.h. $\mathbb{Z}/2\mathbb{Z} = \{[0], [1]\}$; weiter ist $\cdots [-2] = [0] = [2] = \cdots$ und $\cdots [-1] = [1] = [3] = \cdots$. Die entsprechende Partition von $\mathbb{Z}$ hat die Form

$$\mathbb{Z} = [0] \cup [1] = \{\text{gerade ganzen Zahlen}\} \cup \{\text{ungerade ganzen Zahlen}\}.$$

Definition 1.10. Seien $M \neq \emptyset \neq N$ Mengen.

(a) Eine Abbildung f von M nach N, $f : M \to N$, ordnet jedem $m \in M$ genau ein $n \in N$ zu; im Fall $f : m \mapsto n$ schreibe $f(m) = n$ (genauer: eine Abbildung ist eine Teilmenge $F \subseteq M \times N$, so dass es für jedes $m \in M$ genau ein $n \in N$ mit $(m, n) \in F$ gibt; in diesem Fall schreibe $f(m) = n$). Zwei Abbildungen $f, g : M \to N$ sind gleich, $f = g$, falls $f(m) = g(m)$ für alle $m \in M$ gilt. Setze

$$\mathrm{Abb}(M, N) = \{f \mid f : M \to N \text{ Abbildung}\}.$$

(b) Die Identitätsabbildung id_M auf M ist die Abbildung $\mathrm{id}_M : M \to M,\ \mathrm{id}_M(m) = m$ für alle $m \in M$.

(c) Seien M_i nichtleere Mengen, $i = 1, 2, 3$, $f \in \mathrm{Abb}(M_1, M_2)$ und $g \in \mathrm{Abb}(M_2, M_3)$. Das Kompositum $g \circ f$ von f und g ist die Abbildung

$$g \circ f : M_1 \to M_3,\ m_1 \mapsto g(f(m_1)),\ m_1 \in M_1.$$

• Für $f \in \mathrm{Abb}(M_1, M_2),\ g \in \mathrm{Abb}(M_2, M_3)$ und $h \in \mathrm{Abb}(M_3, M_4)$ gilt: $h \circ (g \circ f) = (h \circ g) \circ f$ (d.h. die Bildung von $\circ$ ist assoziativ).

• Ist $f \in \mathrm{Abb}(M, N)$, so gilt $f = \mathrm{id}_N \circ f$ und $f = f \circ \mathrm{id}_M$.

Definition 1.11. Seien $M \neq \emptyset \neq N$ Mengen und $f \in \mathrm{Abb}(M, N)$.

(a) Ist $U \subseteq M$, so ist das Bild von U unter f die Menge

$$f(U) = \{f(u) \mid u \in U\} \subseteq N.$$

(b) Ist $V \subseteq N$, so ist das Urbild von V unter f die Menge

$$f^{-1}(V) = \{m \mid f(m) \in V\} \subseteq M.$$

(c) f ist surjektiv, falls $f(M) = N$ gilt, d.h. zu jedem $n \in N$ gibt es ein $m \in M$ mit $f(m) = n$.

(d) f ist injektiv, falls aus $f(m_1) = f(m_2)$ mit $m_1, m_2 \in M$ stets $m_1 = m_2$ folgt; in diesem Fall besteht das Urbild eines jeden $n \in N$ aus maximal einem Element.

(e) f ist bijektiv, falls f injektiv und surjektiv ist.

Lemma 1.12. *Seien $M \neq \emptyset \neq N$ Mengen und sei $f \in \mathrm{Abb}(M, N)$.*

(a) f ist genau dann injektiv, wenn es ein $g \in \mathrm{Abb}(N, M)$ gibt, sodass gilt: $g \circ f = \mathrm{id}_M$.

(b) f ist genau dann surjektiv, wenn es ein $g \in \mathrm{Abb}(N, M)$ gibt, sodass gilt: $f \circ g = \mathrm{id}_N$.

(c) f ist genau dann bijektiv, wenn es ein $g \in \mathrm{Abb}(N, M)$ mit

$$g \circ f = \mathrm{id}_M \ \text{ und } f \circ g = \mathrm{id}_N$$

gibt; in diesem Fall ist g eindeutig bestimmt und ebenfalls bijektiv.

Beweis. Übung. □

Definition 1.13. Sei $f \in \mathrm{Abb}(M, N)$ bijektiv. Dann gibt es nach Lemma 1.12 eine eindeutig bestimmte Abbildung $g \in \mathrm{Abb}(N, M)$, sodass gilt: $g \circ f = \mathrm{id}_N$ und $f \circ g = \mathrm{id}_M$. In diesem Fall ist g ebenfalls bijektiv; $g = f^{-1}$ ist die zu f inverse Abbildung.

• Seien $f \in \mathrm{Abb}(M_1, M_2)$ und $g \in \mathrm{Abb}(M_2, M_3)$ bijektiv. Dann ist auch $g \circ f$ bijektiv und es gilt: $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. Genauer: Da die Bildung des Kompositums von Abbildungen assoziativ ist, gilt

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ (f^{-1} \circ g^{-1})) = g \circ ((f \circ f^{-1}) \circ g^{-1}) = \\ = g \circ (\mathrm{id}_{M_2} \circ g^{-1}) = g \circ g^{-1} = \mathrm{id}_{M_3};$$

ähnlich folgt $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \mathrm{id}_{M_1}$. Nach Lemma 1.12(c) ist daher $g \circ f$ bijektiv und $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Bemerkung 1.14. Zwei Mengen M, N sind *gleichmächtig*, falls es eine Bijektion von M auf N gibt, in diesem Fall schreibe $|M| = |N|$. Ist $M = \emptyset$, so setzte $|M| = 0$. Eine Menge M ist endlich, falls $M = \emptyset$ oder es eine Bijektion von M auf $\{1, \dots n\} \subseteq \mathbb{N}$ für ein geeignetes $n \in \mathbb{N}$ gibt. In diesem Fall ist n durch M eindeutig bestimmt, und $|M| = n$. Für endliche Mengen gilt: Ist $N \subseteq M$ und $|N| = |M|$, so ist $N = M$.

Für *unendliche* Mengen, d.h. nicht endliche Mengen, gilt dies nicht. Zum Beispiel, es ist $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q}$, aber $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$; Mengen mit der Eigenschaft, dass $|M| = |\mathbb{N}|$ gilt heissen *abzählbar* unendlich; die Menge $\mathbb{R}$ ist nicht abzählbar unendlich.

Die Frage, ob für eine unendliche Teilmenge $M \subseteq \mathbb{R}$ entweder $|M| = |\mathbb{N}|$ oder $|M| = |\mathbb{R}|$ gilt wurde von D. Hilbert 1900 als die *Kontinuumshypothese* formuliert. P. Cohen zeigte 1963, dass sich diese Frage mit den üblichen Axiomen der Mengenlehre weder beweisen noch widerlegen lässt.

Lemma 1.15. *Seien M, N nichtleere endliche Mengen mit $|M| = |N|$ und sei $f \in \mathrm{Abb}(M, N)$. Dann sind gleichwertig:*

(a) *f ist bijektiv,*
(b) *f ist injektiv,*
(c) *f ist surjektiv.*

NB. Das obige Lemma gilt nicht für unendlichen Mengen: Zum Beispiel, die Abbildung $f : \mathbb{Z} \to \mathbb{Z},\ f(n) = 2n$ ist eine Abbildung zwischen zwei (unendlichen) Mengen der gleichen Mächtigkeit, und ist injektiv, aber nicht surjektiv.

Beweis. (a)⇒(b): Trivial nach Definition. (b)⇒ (c): Da f injektiv ist, gilt $|M| = |f(M)|$, und wegen $|M| = |N|$ ist somit $|f(M)| = |N|$. Da

$f(M) \subseteq N$ und $|M| = |N|$ endlich ist, folgt $f(M) = N$, also ist f surjektiv. (c)$\Rightarrow$ (a): Für $n_1, n_2 \in N$, $n_1 \neq n_2$, ist $f^{-1}(n_1) \cap f^{-1}(n_2) = \emptyset$, d.h. die Urbilder $f^{-1}(n)$ der $n \in N$ definieren eine Partition von M

$$M = \bigcup_{n \in N} f^{-1}(n).$$

Da f surjektiv ist, gilt $|f^{-1}(n)| \geq 1$ für alle $n \in N$, damit folgt

$$|M| = |\bigcup_{n \in N} f^{-1}(n)| = \sum_{n \in N} |f^{-1}(n)| \geq \sum_{n \in N} 1 = |N|.$$

Da nach Annahme $|N| = |M|$ gilt, folgt $|f^{-1}(n)| = 1$ für $n \in N$, d.h. f ist injektiv. □

Lemma 1.16. *Sei $f : M \to N$ eine Abbildung und seien $M_1, M_2 \subseteq M$ und $N_1, N_2 \subseteq N$ Teilmengen. Dann gilt:*

(a) $f(M_1 \cup M_2) = f(M_1) \cup f(M_2)$,
(b) $f(M_1 \cap M_2) \subseteq f(M_1) \cap f(M_2)$,
(c) $f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2)$,
(d) $f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2)$.

Beweis. Übung. □

2. Gruppen I

Eine Menge G hat die Struktur einer (abelschen) Gruppe, wenn es eine 'Addition' mit den Eigenschaften der Addition in den ganzen Zahlen $\mathbb{Z}$ gibt; allgemein ist eine Gruppenstruktur auf einer Menge eine (nicht unbedingt kommutative) 'Verknüpfung' von Elementen mit den folgenden Eigenschaften:

Definition 2.1. Sei G eine nichtleere Menge. Eine Verknüpfung $\cdot$ auf G ist eine Abbildung $\cdot : G \times G \to G$, d.h. $\cdot$ ordnet jedem geordneten Paar $(a, b) \in G \times G$ ein Element $c \in G$ zu; schreibe $c = a \cdot b$. Eine Menge G, zusammen mit einer Verknüfung $\cdot$ ist eine Gruppe, falls gilt:

(1) $\cdot$ ist assoziativ: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in G$.
(2) es gibt ein (links)-neutrales Element $e \in G$ mit $e \cdot a = a$ für alle $a \in G$.
(3) zu jedem $a \in G$ gibt es ein (links)-inverses Element, d.h. ein $b \in G$ mit $b \cdot a = e$.

Die Gruppen G ist kommutativ oder abelsch, falls zusätzlich gilt:

(4) $a \cdot b = b \cdot a$ für $a, b \in G$.

• Aufgrund des Assoziativgesetzes (1) lassen sich Produkte von Elementen in einer Gruppe $(G, \cdot)$ beliebig klammern. Seien $a, b, c \in G$ mit $ba = b \cdot a = e$ und $cb = c \cdot b = e$. Dann gilt

$$ab = (ea)b = ((cb)a)b = (c(ba))b = (ce)b = c(eb) = cb = e,$$

d.h. $ba = e$ impliziert $ab = e$ (d.h. das links-inverse Element ist auch ein rechts-inverses Element). Weiter folgt damit auch

$$ae = a(ba) = (ab)a = ea = a,$$

also liefert $ea = a$ auch $ae = a$ (d.h. das links-neutrale Element e ist auch ein rechts-neutrales Element).

• Ist $(G, \cdot)$ eine Gruppe, so schreibe $e = 1$ (Einselement) und $b = a^{-1}$ für das zu a inverse Element. Sind $a_1, a_2, \ldots, a_n \in G$, so schreibe $\prod_{i=1}^{n} a_i = a_1 \cdot \cdots \cdot a_n$; nach Definition gilt $\prod_{i=1}^{0} a_i = 1$. Ist $(G, +)$ eine abelsche Gruppe, so schreiben wir oft '$+$' anstelle von '$\cdot$'; in diesem Fall setze $e = 0$ (Nullelement) und bezeichne das zu a inverse Element mit $-a$. Weiter ist dann $\sum_{i=1}^{n} a_i$ die Summe der endlich vielen Elemente $a_1, \ldots, a_n$; nach Definition ist $\sum_{i=1}^{0} a_i = 0$.

Beispiele 2.2. (a) $(\mathbb{Z}, +)$ ist abelsche Gruppe (übliche Addition).

(b) $(\mathbb{Q}, +)$ und $(\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}, \cdot)$ (übliche Addition und Multiplikation) sind abelsche Gruppen, genauso für $(\mathbb{R}, +)$ und $(\mathbb{R}^\times = \mathbb{R} \setminus \{0\}, \cdot)$.

(c) Die Menge $\mathbb{Z}[x]$ der Polynome in einer Variablen x mit ganzzahligen Koeffizienten bildet eine abelsche Gruppe mittels der Addition

$$f = \sum_{i=0}^{n} a_i x^i, g = \sum_{j=0}^{m} b_j x^j \Rightarrow f + g = \sum_{k=0}^{n+m} (a_k + b_k) x^k;$$

genauso ist die Menge solcher Polynome mit rationalen Koeffizienten $\mathbb{Q}[x]$ bzw. reellen Koeffizienten $\mathbb{R}[x]$ eine abelsche Gruppe.

(d) Sei M eine Menge und $\mathrm{Bij}(M)$ die Menge der bijektiven Abbildungen $M \to M$. Dann bildet $\mathrm{Bij}(M)$ mittels der Komposition $\circ$ von Abbildungen eine Gruppe: Sind $f, g \in \mathrm{Bij}(M)$, so ist auch $g \circ f \in \mathrm{Bij}(M)$, das neutrale Element ist die Identitätsabbildung id_M, und das zu einem $f \in \mathrm{Bij}(M)$ inverse Element ist die inverse Abbildung f^{-1}. Im Fall einer endlichen Menge $M = \{1, \ldots, n\} \subseteq \mathbb{N}$ schreibe $S_n = \mathrm{Bij}(M)$; für $n \geq 3$ ist die Gruppe S_n nicht abelsch.

(e) Sei $m \geq 1$ eine ganze Zahl. Für $a \in \mathbb{Z}$ betrachte die Äquivalenzklasse

$$[a] = \{a + mk \mid k \in \mathbb{Z}\} = a + m\mathbb{Z} \subseteq \mathbb{Z}$$

derjenigen Elemente von $\mathbb{Z}$, die zu a kongruent modulo m sind. Setze

$$[a] + [b] = [a + b],$$

d.h. definiere '+' auf den Äquivalenzklassen durch den Ausdruck auf der rechten Seite. Diese Addition von $[a]$ und $[b]$ ist wohl-definiert:

Ist $[a_1] = [a_2], a_1 - a_2 = km$ und $[b_1] = [b_2], b_1 - b_2 = lm$, so folgt $a_1+b_1-(a_2+b_2) = a_1-a_2+(b_1-b_2) = (k+l)m$, d.h. $[a_1+b_1] = [a_2+b_2]$.

Aus den Eigenschaften der Addition in $\mathbb{Z}$ ergibt sich, dass die Menge

$$\mathbb{Z}/m\mathbb{Z} = \{[a] = a + m\mathbb{Z} \mid a \in \mathbb{Z}\}$$

bzgl. der oben definierten Verknüpfung + die Struktur einer abelschen Gruppe mit neutralem Element $[0]$ hat; es ist $|\mathbb{Z}/m\mathbb{Z}| = m$.

Das nächste Lemma liefert elementare Rechenregeln in Gruppen:

Lemma 2.3. *Sei $(G, \cdot)$ eine Gruppe, $a, b, c \in G$.*

(a) $ab = ac \Rightarrow b = c$ *und* $ac = bc \Rightarrow a = b$,
(b) $(a^{-1})^{-1} = a$,
(c) $(ab)^{-1} = b^{-1}a^{-1}$.

Beweis. (a): Ist $ab = ac$, so liefert Multiplikation mit a^{-1} von links $a^{-1}(ab) = a^{-1}(ac)$ Wegen $a^{-1}(ab) = (a^{-1}a)b = eb = b$ und $a^{-1}(ac) = c$ folgt $b = c$; analog mit Multiplikation mit c^{-1} von rechts für den zweiten Fall. (b): Nach Definiton ist $(a^{-1})^{-1}a^{-1} = e$, Multiplikation mit a von rechts liefert $(a^{-1})^{-1} = a$. (c): Wegen $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}(eb) = b^{-1}b = e$ ist $(ab)^{-1} = b^{-1}a^{-1}$. □

Eine Untergruppe $H \subseteq G$ einer Gruppe G (bzgl. $\cdot$) ist eine Teilmenge, sodass die Einschränkung von $\cdot$ auf H eine Gruppenstruktur auf H definiert, insbesondere muss dazu das Produkt von zwei Elementen aus H in H liegen, und H alle Inversen und die 1 enthalten; genauer:

Definition 2.4. Sei G eine Gruppe. Eine Teilmenge $H \subseteq G$ ist eine Untergruppe von G, $H \leq G$, falls gilt

(a) $1 \in H$
(b) $a, b \in H \Rightarrow ab \in H$
(c) $a \in H \Rightarrow a^{-1} \in H$.

NB. Ist $\emptyset \neq H \subseteq G$ eine *nichtleere* Teilmenge, so lassen sich die Kriterien (a)-(c) der obigen Definition zu einer Bedingung vereinfachen:

$$\emptyset \neq H \subseteq G \textit{ ist Untergruppe, falls gilt: } a, b \in H \Rightarrow ab^{-1} \in H.$$

Konkret: Wegen $\emptyset \neq H$ gibt es ein $a \in H$ und die Bedingung impliziert $aa^{-1} = 1 \in H$, somit gilt (a). Ist $a \in H$ beliebig, so folgt aus $1 \in H$ nun $1a^{-1} = a^{-1} \in H$, also gilt (c). Da mit $a, b \in H$ auch $a, b^{-1} \in H$ ist, ist $a(b^{-1})^{-1} = ab \in H$, dies ist (b).

Beispiele 2.5. (a) In jeder Gruppe G gilt: $\{1\} \leq G$ und $G \leq G$ (die Untergruppen $\{1\}$ und G sind die trivalen Untergruppen von G).

(b) Als additiven Gruppen: $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R}$.

(c) Für die additiven Gruppen der Polynome mit ganzzahligen, rationalen und reellen Koeffizienten: $\mathbb{Z}[x] \leq \mathbb{Q}[x] \leq \mathbb{R}[x]$.

(d) Sei $m \geq 1$ eine ganze Zahl und $m\mathbb{Z} = \{mk \mid k \in \mathbb{Z}\} \subseteq \mathbb{Z}$. Dann ist $m\mathbb{Z} \leq \mathbb{Z}$ eine Untergruppe (für $m \geq 2$ ist $m\mathbb{Z} \subsetneq \mathbb{Z}$ und $|m\mathbb{Z}| = |\mathbb{Z}|$).

Nach Beispiel 2.5(d) bilden für $m \geq 1$ die m-Vielfachen $m\mathbb{Z} \subseteq \mathbb{Z}$ eine Untergruppe von $\mathbb{Z}$. Nach Beispiel 1.9(b) definiert

$$n_1 \equiv n_2 (mod\ m) \Leftrightarrow m|(n_1 - n_2) \Leftrightarrow (n_1 - n_2) \in m\mathbb{Z}$$

eine Äquivalenzrelation auf $\mathbb{Z}$.

Schreibt man die additive Gruppenoperation in $G = \mathbb{Z}$ multiplikativ und setzt man $U = m\mathbb{Z} \leq \mathbb{Z}$, so entspricht der additiven Relation $m|(n_1 - n_2)$ die multiplikative Relation $n_1 n_2^{-1} \in U$. Wir zeigen, dass diese multiplikative Relation bzgl. einer Untergruppe $U \leq G$ allgemein eine Äquivalenzrelation auf G definiert:

Lemma 2.6. *Sei G eine Gruppe und $U \leq G$ eine Untergruppe. Dann definiert $a \sim b \Leftrightarrow ab^{-1} \in U$ $(a, b \in G)$ eine Äquivalenzrelation auf G.*

Beweis. Wegen $aa^{-1} = 1 \in U$ gilt $a \sim a$. Ist $a \sim b$, also $ab^{-1} \in U$, so folgt $(ab^{-1})^{-1} = ba^{-1} \in U$, d.h. $b \sim a$. Ist $a \sim b$ und $b \sim c$, so gilt $ab^{-1} \in U$ und $bc^{-1} \in U$. Es folgt $ac^{-1} = (ab^{-1})(bc^{-1}) \in U$ und so $a \sim c$. □

Definition 2.7. Sei G eine Gruppe, $U \leq G$ eine Untergruppe und $\sim$ die durch U definierte Äquivalenzrelation auf G ($a \sim b \Leftrightarrow ab^{-1} \in U$). Ist $a \in G$, so ist die entsprechende Äquivalenzklasse die Menge

$$[a] = \{b \in G \mid a \sim b\} = \{b \in G \mid ab^{-1} \in U\} = \{ua \mid u \in U\} = Ua;$$

diese Mengen sind die Rechtsnebenklassen von U. Sind die Ua_j für $j \in J$ die verschiedenen Rechtsnebenklassen, so bilden diese eine Partition

$$G = \bigcup_{j \in J} Ua_j.$$

Ist $|J|$ endlich, so ist $|J|$ der Index von U in G; schreibe $|J| = |G : U|$.

• Genauso definiert $a \sim b \Leftrightarrow a^{-1}b \in$ eine Äquivalenzrelation auf G. Die Äquivalenzklasse von $a \in G$ ist die Linksnebenklasse

$$[a] = \{b \in G \mid a \sim nb\} = \{b \in G \mid a^{-1}b \in U\} = aU.$$

Ist G abelsch, so gilt $aU = Ua$; für eine nicht-abelsche Gruppe gilt dies im allgemeinen nicht.

• Der Versuch analog zur Definition der Addition auf $\mathbb{Z}/m\mathbb{Z}$ mittels der Addition auf $\mathbb{Z}$ eine Verknüpfung auf der Menge der Nebenklassen

$G/U = \{Ua \mid a \in G\}$ durch $Ua \cdot Ub = Uab$ zu definieren funktioniert für abelsche Gruppen, aber nicht für allgemeine Gruppen. Dies wird uns zu besonderen Untergruppen führen, den sogenannten Normalteilern.

Das folgende Resultat besagt, dass es einer endlichen Gruppe nicht Untergruppen beliebiger Kardinalität geben kann:

Theorem 2.8. *(Satz von Lagrange) Sei G eine endliche Gruppe (d.h. die Menge G ist endlich) und $U \leq G$ eine Untergruppe. Dann gilt:*

$$|G| = |G : U||U|.$$

NB. Das Theorem besagt: Gibt es eine Untergruppe $U \leq G$, so ist $|U|$ ein Teiler von $|G|$; dies ist eine notwendige Bedingung für die Existenz von Untergruppen; zum Beispiel kann eine Gruppe G mit Primzahlordnung $|G| = p$ nur die trivialen Untergruppen $\{1\}$ und G enthalten.

Beweis. Betrachte die Partition $G = \cup_j Ua_j$. Für $a \in G$ ist die Abbildung $U \to Ua,\ u \mapsto ua$ surjektiv (ist $ua \in Ua$, so gilt $u \mapsto ua$) und injektiv (ist $u_1 a = u_2 a$, so folgt $u_1 = u_2$), also eine Bijektion. Es folgt $|aU| = |U|$ und $|G| = |J||U| = |G : U||U|$. □

Bemerkung 2.9. Die Umkehrung des Satzes von Lagrange gilt nicht, d.h. im allgemeinen gibt es zu einem Teiler der Gruppenordnung $|G|$ einer endlichen Gruppe keine Untergruppe dieser Ordnung. Ein grundlegendes Resultat in diesem Kontext ist das folgende Theorem von Sylow: *Sei G eine endliche Gruppe, $|G| = n$ und p eine Primzahl die n teilt. Sei p^m die maximale p-Potenz in n. Dann gibt es eine Untergruppe $U \leq G$ mit $|U| = p^m$.* Zum Beispiel: Jede Gruppe G mit $|G| = 24 = 3 \cdot 2^3$ besitzt mindestens eine Untergruppe $U \leq G$ mit $|U| = 3$ und eine Untergruppe $V \leq G$ mit $|V| = 2^3 = 8$.

3. Körper

Ein Körper ist eine additiv geschriebene abelsche Gruppe, auf der zusätzlich eine Multiplikation definiert ist, die die Eigenschaften der Multiplikation von rationalen Zahlen erfüllt.

Definition 3.1. Ein Körper K ist eine Menge mit zwei Verknüpfungen $+$ und $\cdot$, sodass gilt:

(1) $(K, +)$ ist eine abelsche Gruppe mit Nullelement 0,
(2) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppen mit Einselement $1 \neq 0$,
(3) $a \cdot (b + c) = a \cdot b + a \cdot c$ und $(a + b) \cdot c = a \cdot c + b \cdot c$.

In Körpern gelten viele der 'üblichen' Rechenregeln. Für $a, b \in K$ ist:

- $0a = a0 = 0$,
- $(-1)a = -a$,
- $(-a)b = a(-b) = -ab$,
- $ab = 0 \Rightarrow a = 0$ oder $b = 0$.

NB. Nicht alle Eigenschaften der rationalen Zahlen gelten für allgemeine Körper. Zum Beispiel, in $\mathbb{Q}$ folgt für $n \in \mathbb{N}$ und $a \in \mathbb{Q}$ aus $n \cdot a = (1 + \cdots + 1) \cdot a = 0$ stets $a = 0$, aber es gibt Körper mit der Eigenschaft, dass $n \cdot a = 0$ für $n \neq 0$ und $a \neq 0$.

Beispiele 3.2. (a) $\mathbb{Q}$ und $\mathbb{R}$ sind Körper.

(b) Sei p eine Primzahl und $\mathbb{Z}/p\mathbb{Z}$ die Menge der Restklassen modulo p. Die Menge $\mathbb{Z}/p\mathbb{Z}$ ist bzgl. der Addition $[a] + [b] = [a + b]$ eine abelsche Gruppe mit Nullelement $[0]$. Analog definiert $[a] \cdot [b] = [ab]$ eine Multiplikation auf $\mathbb{Z}/p\mathbb{Z}$ mit Einselement $[1]$, sodass $\mathbb{Z}/p\mathbb{Z} \setminus \{[0]\}$ eine abelsche Gruppe ist. Aufgrund der Definition von $+$ und $\cdot$ in $\mathbb{Z}/p\mathbb{Z}$ (mittels $+$ und $\cdot$ in $\mathbb{Z}$) sind diese Operationen verträglich und $\mathbb{Z}/p\mathbb{Z}$ ist ein Körper mit p Elementen; vgl. Übung. In dem Körper $\mathbb{Z}/p\mathbb{Z}$ gilt $pa = 0$ für alle $a \in \mathbb{Z}/p\mathbb{Z}$.

Analog zur Definition einer Untergruppe einer Gruppe ist ein Unterkörper oder Teilkörper eines Körpers K eine Teilmenge $L \subseteq K$ mit der Eigenschaft, dass sich $+$ und $\cdot$ auf K zu Verknüpfungen $L \times L \to L$ auf L einschränken, und L bezüglich dieser Verküpfungen einen Körper bildet.

Definition 3.3. Sei K ein Körper. Ein Unterkörper $L \subseteq K$ ist eine Teilmenge, sodass gilt:

(a) $a, b \in L \Rightarrow a + b, a \cdot b \in L$,
(b) $0, 1 \in L$,
(c) $a \in L \Rightarrow -a \in L$,
(d) $0 \neq a \in L \Rightarrow a^{-1} \in L$.

Beispiele 3.4. (a) $\mathbb{Q}$ ist ein Unterkörper von $\mathbb{R}$.

(b) Betrachte die Menge $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subseteq \mathbb{R}$. Dann ist $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$ ein Unterkörper mit $\mathbb{Q} \subsetneq \mathbb{Q}(\sqrt{2}) \subsetneq \mathbb{R}$ (der Körper $\mathbb{Q}(\sqrt{2})$ ist der 'kleinste' Teilkörper von $\mathbb{R}$, der $\sqrt{2}$ enthält): Wir zeigen zunächst $\mathbb{Q} \subsetneq \mathbb{Q}(\sqrt{2}) \subsetneq \mathbb{R}$: Angenommen $\sqrt{2} = p/q \in \mathbb{Q}$, $p, q \in \mathbb{Z}, q \neq 0$, p/q gekürzt. Dann ist $p^2 = 2q^2$, also ist p^2 und damit p gerade (das Quadrat einer ungeraden Zahl ist ungerade). Sei $p = 2k$ für ein $k \in \mathbb{Z}$. Wegen $4k^2 = (2k)^2 = p^2 = 2q^2$ ist $2k^2 = q^2$, also ist q gerade: Widerspruch zu p/q ist gekürzt; also ist $\mathbb{Q} \subsetneq \mathbb{Q}(\sqrt{2})$. Wir nehmen nun an, dass

$\sqrt{3} \in \mathbb{Q}(\sqrt{2})$ ist, d.h. $\sqrt{3} = a + b\sqrt{2}$ für $a, b \in \mathbb{Q}$. Dann ist $a \neq 0$ (sonst $\sqrt{3} = b\sqrt{2}$, also $3 = 2b^2$) und $b \neq 0$ (sonst wäre $\sqrt{3} = a \in \mathbb{Q}$; ein ähnliches Argument wie für $\sqrt{2}$ zeigt, dass dies nicht gilt). Aus $\sqrt{3} = a + b\sqrt{2}$ folgt mit der binomischen Formel

$$3 = a^2 + 2ab\sqrt{2} + 2b^2,$$

und wegen $a \neq 0 \neq b$ dann $\sqrt{2} = (3 - a^2 - 2b^2)/2ab \in \mathbb{Q}$; Widerspruch zu $\sqrt{2} \notin \mathbb{Q}$. Also ist $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$ und damit auch $\mathbb{Q}(\sqrt{2}) \subsetneq \mathbb{R}$.

Einfaches Nachrechnen (in $\mathbb{R}$!) liefert die Formeln

$$\begin{aligned} (a + b\sqrt{2}) + (c + \sqrt{2}d) &= (a + c) + \sqrt{2}(b + d), \\ (a + b\sqrt{2}) \cdot (c + d\sqrt{2}) &= (ac + 2bd) + (ad + bc)\sqrt{2}, \end{aligned}$$

d.h. $\mathbb{Q}(\sqrt{2})$ ist abgeschlossen bzgl. $+$ und $\cdot$ und es gilt (a). Wegen $0 = 0 + 0\sqrt{2}$ und $1 = 1 + 0\sqrt{2}$ gilt (b). Da mit $a + b\sqrt{2}$ auch das additiv Inverse $-(a + b\sqrt{2}) = (-a) + (-b)\sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$ liegt, haben wir (c). Für (d) betrachte $0 \neq a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Dann ist $a \neq 0$ oder $b \neq 0$ und damit auch $0 \neq a - b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Es folgt $0 \neq (a - \sqrt{2}b)(a + \sqrt{2}b) = a^2 - 2b^2$ und das inverse $(a + b\sqrt{2})^{-1}$ ist durch folgende Formel gegeben

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} \in \mathbb{Q}(\sqrt{2}).$$

4. Vektorräume

Eine (abelsche) Gruppe ist eine algebraische Struktur, die die Eigenschaften der Addition in den ganzen Zahlen abstrahiert. Ähnlich ist die Definition eines Körpers eine abtrakte Formulierung der Eigenschaften der Addition und Multiplikation von rationalen Zahlen.

Die algebraische Struktur eines Vektorraums ist motiviert durch die reelle Ebene $\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\}$, zusammmen mit der Addition

$$v = (a, b),\ w = (c, d) \in \mathbb{R}^2 \Rightarrow v + w = (a + c, b + d),$$

und der Skalarmultiplikation

$$v = (a, b) \in \mathbb{R}^2, \alpha \in \mathbb{R} \ \Rightarrow \alpha \cdot v = (\alpha \cdot a,\ \alpha \cdot b).$$

In der abstrakten Formulierung werden die Vektoren Elemente einer Menge und die Skalare Elemente eines Körpers sein; zur Unterscheidung bezeichnen wir Vektoren mit lateinischen Buchstaben $a, b, c \ldots$ und Skalare mit griechischen Buchstaben $\alpha, \beta, \gamma, \ldots$.

Definition 4.1. Sei K ein Körper. Ein K-Vektorraum ist eine Menge V, zusammem mit einer (inneren) Verknüpfung $V \times V \to V,\ (v, w) \mapsto v + w$ (einer 'Addition' $+$) und einer (äusseren) Veknüpfung $K \times V \to$

V, $(\alpha, v) \mapsto \alpha \cdot v$ (einer 'Skalarmultiplikation' $\cdot$), sodass für $\alpha, \beta \in K$ und $v, w \in V$ gilt:

(1) V ist bzgl. $+$ eine abelsche Gruppe (insbesondere ist $V \neq \emptyset$),

(2) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ und $\alpha \cdot (v + w) = \alpha \cdot v + \beta \cdot w$,

(3) $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$,

(4) $1 \cdot v = v$.

Für K-Vektorräume gelten die folgenden Rechenregeln (hier ist 0_V das Nullelement in V und 0_K das Nullelement in K; im Weiteren werden diese Elemente nur mit 0 bezeichtnet, da es sich aus dem Kontext ergibt, welche 'Null' gemeint ist; weiter werden wir für $\alpha \cdot v$ oft einfach nur αv schreiben):

- $\alpha \cdot 0_V = 0_V$ für $\alpha \in K$,
- $0_K \cdot v = 0_V$ für $v \in V$,
- $(-\alpha) \cdot v = \alpha \cdot (-v)$ für $\alpha \in K$ und $v \in V$,
- $\alpha \cdot v = 0_V$ für $\alpha \in K$ und $v \in V$ impliziert $\alpha = 0_K$ oder $v = 0_V$,
- $\alpha \cdot (\sum_{i=1}^n v_i) = \sum_{i=1}^n (\alpha \cdot v_i)$ und $(\sum_{i=1}^n \alpha_i) \cdot v = \sum_{i=1}^n (\alpha_i \cdot v)$,

Beispiele 4.2. (a) Jede abelsche Gruppe enthält ein Nullelement 0 und ist daher eine nicht-leere Menge. Ist $V = \{0\}$ eine einelementige Menge, so ist V bzgl. $0 + 0 = 0$ eine abelsche Gruppe und für jeden Körper K und $\alpha \in K$ mittels $\alpha \cdot 0 = 0$ ein K-Vektorraum; V ist der triviale K-Vektorraum oder Nullraum.

(b) Sei K ein Körper und $n \in \mathbb{N}_0$. Dann ist das Produkt $K^n = \{(\alpha_1, \ldots, \alpha_n) \mid \alpha_i \in K\}$ bzgl. der komponentenweisen Addition und Skalarmultiplikation

$$\begin{aligned} (\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) &= (\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n), \\ \alpha \cdot (\alpha_1, \cdots, \alpha_n) &= (\alpha \cdot \alpha_1, \cdots, \alpha \cdot \alpha_n) \end{aligned}$$

ein K-Vektorraum; für $n = 0$ ist $K^0 = \{0\}$ der Nullraum. Die K-Vektorräume K^n sind die zentralen Beispiele in der linearen Algebra.

(c) Sei K ein Körper und M ein Menge. Dann ist die Menge $V = \mathrm{Abb}(M, K)$ der Abbildungen $M \to K$ ein K-Vektorraum bezüglich der 'punktweise' definierten Verknüpfungen: $f, g \in V$, $\alpha \in K$,

$$\begin{aligned} f + g : &\quad M \to K, \quad m \mapsto f(m) + g(m), \\ \alpha \cdot f : &\quad M \to K, \quad m \mapsto \alpha \cdot f(m) \end{aligned}$$

Diese Beispiele von K-Vektorräumen treten oft in der Analysis auf; zum Beispiel, ist $I = [0, 1] \subseteq \mathbb{R}$ das Einheitsintervall, und $K = \mathbb{R}$, so ist $V = \mathrm{Abb}(I, \mathbb{R})$ der $\mathbb{R}$-Vektorraum der reellwertigen Funktionen auf

dem Einheitsintervall.

(d) Sei K ein Körper und $K[x]$ die Menge der Polynome in x mit Koeffizienten in K. Dann ist $K[x]$ bzgl. der üblichen Addition von Polynomen (Addition der Koeffizienten) und der Skalarmultiplikation

$$\alpha \in K,\ f = \sum_{i=0}^{n} \alpha_i x^i \Rightarrow \alpha \cdot f = \sum_{i=0}^{n} (\alpha\alpha_i) x^i$$

ein K-Vektorraum.

(e) Sei K ein Körper und $k \subseteq K$ ein Unterkörper. Nach Definition ist $k \leq K$ eine abelsche Untergruppe und die Einschränkung der Multiplikation $K \times K \to K$ auf $k \times K \to K$ definiert ein Skalarprodukt, d.h. K ist ein k-Vektorraum. Insbesondere ist wegen der Inklusionen von Unterkörpern $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$, $\mathbb{R}$ nicht nur ein $\mathbb{R}$-Vektorraum ($\mathbb{R} = \mathbb{R}^1$), sondern auch ein $\mathbb{Q}$-Vektorraum bzw. $\mathbb{Q}(\sqrt{2})$-Vektorraum.

Ein K-linearer Unterraum eines K-Vektorraums V ist eine Teilmenge $U \subseteq V$, die bzgl. der Einschränkung der Addition und der Skalarmultiplikation von V auf U einen K-linearen Vektorraum definiert. Formal:

Definition 4.3. Sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ ist ein K-Untervektorraum oder K-linearer Unterraum von V, falls gilt:

(a) $\emptyset \neq U$,
(b) $a, b \in U \Rightarrow a + b \in U$,
(c) $\alpha \in K, a \in U \Rightarrow \alpha \cdot a \in U$ (insbesondere: $a \in U \Rightarrow -a \in U$).

NB. Ist V ein K-Vektorraum und $U \subseteq V$ ein K-linearer Unterraum, so bezeichnen wir U oft einfach als linearen Unterraum (d.h. ein linearer Unterraum eines K-Vektorraums ist stets ein Untervektorraum über demselben Körper K).

Beispiele 4.4. (a) Sei V ein K-Vektorraum. Dann ist $V \subseteq V$ stets ein Unterraum. Ist $v \in V$, so ist der von v erzeugte lineare Unterraum

$$\langle v \rangle = K \cdot v = \{\alpha \cdot v \mid \alpha \in K\} \subseteq V.$$

Insbesondere enthält jeder K-Vektorraum V die linearen Unterräume $\{0\}$ ($v = 0$) und V; dies sind die trivialen linearen Unterräume.

(b) Sei $m \leq n$, und sei $K^m \subseteq K^n$ die kanonische Inklusion, die ein m-Tupel $(\alpha_1, \ldots, \alpha_m) \in K^m$ mit dem n-Tupel $(\alpha_1, \ldots, \alpha_m, 0, \ldots 0) \in K^n$ identifiziert. Dann ist $K^m \subseteq K^n$ ein linearer Unterraum.

(c) Sei K ein Körper und M eine Menge. Ist $V = \mathrm{Abb}(M, K)$ der K-Vektorraum der K-wertigen Funktionen auf M, so bilden die stetigen (bzw. differenzierbaren, Polynomfunktionen) einen Unterraum von V.

(d) Wegen der Inklusionen $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$ sind $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$ und $\mathbb{R}$ $\mathbb{Q}$-lineare Unterräume des $\mathbb{Q}$-Vektorraums $\mathbb{R}$.

Wir betrachten Untervektorräume eines K-Vektorraums V.

Lemma 4.5. *Sei V ein K-Vektorraum und sei $\{U_i\}_{i\in I}$ eine Familie von linearen Unterräumen von V. Dann ist $U = \cap_{i\in I} U_i \subseteq V$ ebenfalls ein linearer Unterraum.*

Beweis. Aus $0 \in U_i$ für alle i folgt $0 \in U$, d.h. $U \neq \emptyset$. Sind $u, u' \in U$ und $\alpha \in K$, so ist $u + u' \in U_i$ und $\alpha \cdot u \in U_i$ für alle i, da die U_i lineare Unterräume sind. Damit folgt $u + u' \in U$ und $\alpha \cdot u \in U$. □

Wir wollen zu einer beliebigen Teilmenge $A \subseteq V$ eines K-Vektorraums den 'kleinsten' linearen Unterraum $\langle A\rangle \subseteq V$ bestimmen, der die gegebene Menge A enthält. Aus dem obigen Lemma ergibt sich, dass

$$\langle A\rangle = \cap\{U \mid U \subseteq V \text{ linearer Unteraum mit } A \subseteq U\}.$$

Klar ist, dass dieser lineare Unterraum $\langle A\rangle$ alle Elemente der Form

$$\sum_{i=1}^{n} \alpha_i a_i,\ n \in \mathbb{N}_0,\ \alpha_i \in K,\ a_i \in A$$

enthalten muss. Wir zeigen:

Lemma 4.6. *Sei V ein K-Vektorraum und $A \subseteq V$ eine Teilmenge. Der von der Teilmenge $A \subseteq V$ in V erzeugte lineare Unterraum ist*

$$\begin{aligned}\langle A\rangle &= \{\textstyle\sum_{i=1}^{n} \alpha_i a_i \mid n \in \mathbb{N}_0,\ \alpha_i \in K,\ a_i \in A\}\\ &= \cap\{U \mid U \subseteq V \textit{ linearer Unterraum mit } A \subseteq U\} \subseteq V.\end{aligned}$$

NB. Sei $\{a_i\}_{i\in I} \subseteq V$ eine Familie von Elementen von V und $A = \{a_i \mid i \in I\}$. Dann ist der von den a_i erzeugte lineare Unterraum

$$\langle a_i \mid i \in I\rangle = \langle A\rangle \subseteq V.$$

- $\langle\emptyset\rangle = \{0\}$,
- $A \subseteq \langle A\rangle$ für jede Teilmenge $A \subseteq V$,
- $U = \langle U\rangle$ für jeden linearen Unterraum $U \subseteq V$,
- Für Teilmengen $A, B \subseteq V$ gilt: $A \subseteq B \Rightarrow \langle A\rangle \subseteq \langle B\rangle$ und $A \subseteq \langle B\rangle \Rightarrow \langle A\rangle \subseteq \langle B\rangle$.

Beweis. Ist $A \neq \emptyset$, $a \in A$, so ist auch $0 = 0 \cdot a \in \langle A\rangle$; ist $A = \emptyset$, so ist nach Definition der leeren Summe $\sum_{i=1}^{0} \alpha_i \cdot a_i = 0$, also gilt $0 \in \langle A\rangle$ auch in diesem Fall. Sei $\alpha \in K$ ein Skalar und seien $a, b \in \langle A\rangle$, d.h. $a = \Sigma_{i=1}^{r} \alpha_i a_i$ und $b = \Sigma_{j=1}^{s} \beta_j b_j$. Dann sind auch $\alpha a = \sum_{i=1}^{r} (\alpha\alpha_i) a_i$ und

$a+b=\sum_{i=1}^{r}\alpha_i a_i+\sum_{j=1}^{s}\beta_j b_j$ in $\langle A\rangle$, d.h. $\langle A\rangle$ definiert einen linearen Unterraum.

Sei $U\subseteq V$ ein linearer Unterraum, der A enthält. Dann muss U auch die $\langle A\rangle$ definierenden Ausdrücke enthalten, also ist $\langle A\rangle\subseteq U$ und da dies für jeden solchen Unterraum U gilt, folgt $\langle A\rangle\subseteq\cap\{U\mid U\subseteq V \text{ linearer Unterraum mit } A\subseteq U\}$. Wegen $A\subseteq\langle A\rangle$ (da $1\cdot a=a$ für $a\in A$) ist auch $\langle A\rangle$ ein Unterraum, der A enthält. Also ist $\cap\{U\mid U\subseteq V \text{ linearer Unterraum mit } A\subseteq U\}\subseteq\langle A\rangle$ und damit gilt Gleichheit. □

Beispiele 4.7. (a) Sei $V=\mathbb{R}^2$. Ist $0\neq a\in\mathbb{R}^2$ ein beliebiger Vektor, so ist der von a erzeugte lineare Unterraum (vgl. Beispiel 4.4(a))

$$\langle a\rangle=\{\alpha\cdot a\mid\alpha\in\mathbb{R}\}\subseteq\mathbb{R}^2$$

genau die Gerade durch den Ursprung, die durch a erzeugt wird. Ist $0\neq b\in\mathbb{R}^2$ ein weiterer Vektor mit $a\neq\alpha\cdot b$ für alle $\alpha\in\mathbb{R}$ (d.h. die Vektoren a und b liegen nicht auf derselben Geraden), so ergibt sich

$$\langle a,b\rangle=\{\alpha\cdot a+\beta\cdot b\mid\alpha,\beta\in\mathbb{R}\}=\mathbb{R}^2.$$

Insbesondere gibt es für *jeden* Vektor $c\in\mathbb{R}^2$ Skalare $\alpha,\beta\in\mathbb{R}$, sodass

$$c=\alpha a+\beta b.$$

(b) Sei $V=\mathbb{Q}[x]$ der $\mathbb{Q}$-Vektorraum aller Polynome in x mit rationalen Koeffizienten. Dann gilt $\langle 1\rangle=\{\text{konstante Polynome }\}=\mathbb{Q}$, $\langle\{1,x\}\rangle=\{\text{Polynome vom Grad }\leq 1\}$, $\langle\{1,x,x^2\}\rangle=\{\text{Polynome vom Grad }\leq 2\}$, etc., d.h. keine *endliche* Teilmenge der Form $A=\{1,x,x^2,\dots,x^n\}$ hat die Eigenschaft, dass $\langle A\rangle=\mathbb{Q}[x]$ ist. Allerdings ist $\langle A\rangle=\mathbb{Q}[x]$ für $A=\{x^n\mid k\in\mathbb{N}_0\}$ (vgl. $x^0=1$).

Was betrachten Teilmengen $A\subseteq V$ für die $\langle A\rangle=V$ gilt:

Definition 4.8. Eine Menge $A=\{a_i\}_{i\in I}\subseteq V$ von Elementen eines K-Vektorraums V ist ein Erzeugendensystem von V, falls $\langle A\rangle=V$ gilt, d.h. falls jeder Vektor $v\in V$ eine Darstellung als endliche Summe

$$v=\sum_{i=1}^{n}\alpha_i a_i,\ \alpha_i\in K,\ a_i\in A$$

besitzt. Der K-Vektorraum V ist endlich erzeugt (über K), falls V eine endliches Erzeugensystem $A=\{a_1,\cdots,a_n\}$ besitzt.

NB. Ist A ein Erzeugendensystem von V, d.h. $\langle A\rangle=V$, so hat nach Definition *jeder* Vektor $v\in V$ eine Darstellung als eine endliche Summe $v=\sum_{i=1}^{n}\alpha_i a_i$, wobei $\alpha_i\in K$ und $a_i\in A$ sind. Aber: diese Darstellung ist nicht unbedingt eindeutig. Zum Beispiel, die Menge

$A = \{(1,0),(0,1),(1,1)\}$ ist ein Erzeugendensystem von $\mathbb{R}^2$, aber der Nullvektor hat die beiden Darstellungen

$$\begin{aligned}(0,0) &= 0\cdot(1,0)+0\cdot(0,1)\\ (0,0) &= 1\cdot(1,0)+1\cdot(0,1)+(-1)\cdot(1,1)\end{aligned}$$

Beispiele 4.9. (a) Ist V ein K-Vektorraum und $A = V$, so gilt $\langle V\rangle = V$; $\langle V\rangle$ ist das triviale Erzeugendensystem.

(b) Nach Beispiel 4.7(a) bilden in $\mathbb{R}^2$ je zwei nicht-triviale und nicht auf einer Graden liegende Vektoren $a, b \in \mathbb{R}^2$ ein Erzeugendensystem, also ist $\mathbb{R}^2$ endlich erzeugt; jede Teilmenge $A \subseteq \mathbb{R}^2$ die mindestens zwei solche nicht-triviale und nicht auf einer Geraden liegenden Vektoren enthält ist ein Erzeugendensystem, d.h. $\mathbb{R}^2$ lässt sich von je zwei solchen (nicht unbedingt 'orthogonalen') Elementen erzeugen. Allgemein besitzt $\mathbb{R}^n$ ein Erzeugendensystem aus n Elementen: Setze $e_i = (0,\ldots,1,\ldots,0)$ d.h. e_i ist der Vektor in $\mathbb{R}^n$ mit 1 in der i-ten Komponente und 0 in allen anderen. Für $v = (\alpha_1,\ldots,\alpha_n) \in \mathbb{R}^n$ ist dann

$$v = \sum_{i=1}^{n} \alpha_i e_i,$$

d.h. $\langle e_1,\ldots,e_n\rangle = \mathbb{R}^n$. Analog folgt für jeden Körper K, dass die n Vektoren $\{e_1,\ldots,e_n\}$ den Vektorraum K^n erzeugen.

(c) Der $\mathbb{Q}$-Vektorraum $\mathbb{Q}[x]$ ist nicht endlich erzeugt: ist $A \subseteq \mathbb{Q}[x]$ eine endliche Menge von Polynomen $f_1,\ldots,f_n$, so ist die maximale Potenz von x in endlichen Summen der Form $\sum_{i=1}^{n} q_i f_i,\ q_i \in \mathbb{Q}$, beschränkt und jedes Polynom, dass eine höheren Potenz von x enthält kann nicht als eine solche Summe dargestellt werden.

(d) Betrachte den Körper $V = \mathbb{Q}(\sqrt{2})$ von Beispiel 3.4(b). Der Körper V ist ein V-Vektorraum und als solcher von 1 erzeugt, $\langle 1\rangle = V$. Da $\mathbb{Q} \subseteq V$ ein Unterkörper ist, ist V auch ein $\mathbb{Q}$-Vektorraum; für V als $\mathbb{Q}$-Vektorraum gilt nach Definition $\langle 1,\sqrt{2}\rangle = \{a+b\sqrt{2} \mid a,b \in \mathbb{Q}\} = V$.

Uns interessieren Erzeugendensysteme $A \subseteq V$ mit der Eigenschaft, dass sich jedes $v \in V$ *eindeutig* als eine endliche Linearkombination

$$v = \sum_{i=1}^{n} \alpha_i a_i,\ \alpha_i \in K,\ a_i \in A$$

schreiben lässt. In diesem Fall sind die α_i eindeutig bestimmt, sodass

$$v = \sum_{i=1}^{n} \alpha_i a_i \leftrightarrow v = (\alpha_1,\ldots,\alpha_n),$$

d.h. die α_i lassen sich als die Koordinaten von v bzgl. A interpretieren.

Diese Bedingung einer eindeuitgen Darstellung lässt sich wie folgt formulieren: Sei $A = \{a_1, \ldots, a_n\}$ endlich. Lässt sich jeder Vektor eindeutig als eine endliche Linearkombination der a_i darstellen, so gilt dies insbesondere für den Nullvektor. Wegen $0 = 0 \cdot a_1 + 0 \cdot a_2 + \cdots + 0 \cdot a_n$ folgt dann

$$(\#) \qquad \sum_{i=0}^{n} \alpha_i a_i = 0 \Leftrightarrow \alpha_i = 0 \text{ für alle } i = 0, \ldots, n.$$

Gilt umgekehrt (#), und sind $v = \sum_{i=1}^{n} \alpha_i a_i = \sum_{i=1}^{n} \beta_i a_i$ zwei Darstellungen von einem Vektor v, so folgt wegen $0 = v - v = \sum_{i=1}^{n} (\alpha_i - \beta_i) a_i$ aus (#) die Gleichheit der Koeffizienten $\alpha_i = \beta_i$ für $i = 0, \ldots, n$, d.h. die Darstellung von v als $v = \sum_{i=1}^{n} \alpha_i a_i$ ist eindeutig.

Dies führt zu dem Begriff der linearen Unabhängigkeit.

Definition 4.10. Sei V ein K-Vektorraum und seien $\{a_i\}_{i \in I}$ Vektoren in V. Die Menge $\{a_i\}_{i \in I}$ ist linear unabhängig (oder auch: die a_i sind linear unabhängig), falls für jede endliche Teilmenge $J \subseteq I$ gilt

$$\sum_{j \in J} \alpha_j a_j = 0 \Rightarrow \alpha_j = 0 \text{ für alle } j \in J.$$

Sind die a_i nicht linear unabhängig, so sind sie linear abhängig.

NB. Vektoren $\{a_1, \ldots, a_n\}$ sind linear unabhängig genau dann, wenn $\langle a_1, \ldots, a_n \rangle$ ein minimales Erzeugendensystem ist (d.h. kein a_i lässt sich durch die anderen a_j als $a_i = \sum_{j \neq i} \alpha_j a_j,\ \alpha_j \in K$ darstellen): Sind $\{a_1, \ldots, a_n\}$ linear abhängig, so gibt einen Ausdruck $\sum_{i=1}^{n} \alpha_i a_i = 0$ mit nicht alle $\alpha_i = 0$. Ist $\alpha_i \neq 0$, so folgt $a_i = \sum_{j \neq i} -\alpha_i^{-1} \alpha_j a_j$, d.h. $\langle a_1, \ldots, a_n \rangle$ ist nicht minimal. Ist umgekehrt $\langle a_1, \ldots, a_n \rangle$ nicht minimal, so gibt es ein a_i mit $a_i = \sum_{j \neq i} \alpha_j a_j$ und daher $(-1)a_i + \sum_{j \neq i} \alpha_j a_j = 0$; da in einem Körper $-1 \neq 0$ ist sind $\{a_1, \ldots, a_n\}$ linear abhängig.

- Die aus dem Nullvektor $\{0\}$ bestehende Menge ist linear abhängig.
- Eine Menge die einen Vektor und ein nicht-triviales Skalarvielfaches dieses Vektors enthält ist linear abhängig (vgl. $\alpha v + (-\alpha)v = 0$).
- Die aus einem Nichtnullvektor bestehende Menge $\{v \neq 0\}$ ist linear unabhängig.
- Die leere Menge $\emptyset$ ist linear unabhängig.

Beispiele 4.11. (a) Ist $V = \mathbb{R}^2$, so sind die Vektoren $e_1 = (1, 0)$ und $e_2 = (0, 1)$ linear unabhängig. Dies folgt formal, da die Gleichung

$$(0, 0) = \alpha \cdot (1, 0) + \beta \cdot (0, 1) = (\alpha, \beta)$$

nur die Lösung $\alpha = 0 = \beta$ hat. Die lineare Unabhängigkeit von e_1 und e_2 ist geometrisch klar: Jeder Vektor $a \in \mathbb{R}^2$ lässt sich auf genau eine Weise als Summe ('Parallelogramm') von Vielfachen von e_1 und e_2 darstellen. Aus dem gleichen Grund sind je zwei nicht-triviale Vektoren $a, b \in \mathbb{R}^2$, die nicht auf einer Graden liegen linear unabhängig.

(b) Sei $V = K^n$ und sei $e_i = (0, \ldots, 1, \ldots, 0)$ der Vektor mit 1 in der i-ten Komponente, $i = 1, \ldots, n$. Analog wie in (a) gilt für alle $\alpha_i \in K$

$$(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^{n} \alpha_i e_i.$$

Also ist $\sum_{i=1}^n \alpha_i e_i = 0$ genau dann, wenn $\alpha_i = 0$ für $i = 1, \ldots., n$, d.h. die $\{e_1, \ldots, e_n\}$ sind linear unabhängig.

(c) Ist $V = \mathbb{Q}[x]$ der $\mathbb{Q}$-Vektorraum der Polynome mit rationalen Koeffizienten, so die unendliche Menge $A = \{1, x, \ldots, x^n, \ldots \mid n \in \mathbb{N}_0\} \subseteq V$ linear unabhängig: Aus $\sum_{i=0}^n \alpha_i x^i = 0 = \sum_{i=0}^n 0x^i$ folgt $\alpha_i = 0$ für $i = 0, \ldots, n$.

Definition 4.12. Sei V ein K-Vektorraum. Eine Basis von V ist eine Erzeugendensystem $B = \{b_i \mid i \in I\} \subseteq V$ von V (d.h. $\langle B \rangle = V$), welches aus linear unabhängigen Vektoren besteht.

NB. Ist $B \subseteq V$ eine Basis, so besagt die erste Bedingung, dass jeder Vektor in V sich als (endliche) Linearkombination von Elementen aus B mit Koeffizienten aus K darstellen lässt. Die zweite Bedingung impliziert, dass diese Darstellung eindeutig ist. Inbesondere: Ist $B = \{b_i \mid i \in I\} \subseteq V$ eine Basis und $J \subseteq I$ eine endliche Teilmenge so gilt: $\sum_{j \in J} \alpha_j b_j = \sum_{j \in J} \beta_j b_j \;\Rightarrow\; \alpha_j = \beta_j$ für $j \in J$.

Theorem 4.13. *Sei V ein endlich erzeugter K-Vektorraum, $V = \langle a_1, \ldots, a_n \rangle$. Sei $1 \leq k \leq n$ und seien $c_1, \ldots, c_k$ linear unabhängige Vektoren in V. Dann gibt es eine Basis $\{b_1, \ldots, b_m\}$ von V, sodass gilt*

$$\{c_1, \ldots, c_k\} \subseteq \{b_1, \ldots, b_m\} \subseteq \{a_1, \ldots, a_n\},$$

d.h. jede linear unabhängige Menge $\{c_1, \ldots, c_k\}$ lässt sich durch Hinzunahme geeigneter Vektoren aus einem Erzeugendensystem zu einer Basis von V ergänzen.

NB. 1. Das Theorem besagt, dass jeder endlich erzeugte K-Vektorraum eine Basis besitzt: Ist $V = \{0\}$, so ist $\{\emptyset\}$ eine Basis. Ist $V \neq \{0\}$, so gibt es einen Vektor $0 \neq c \in V$, der linear unabhängig ist, und $\{c\}$ lässt sich durch Hinzunahme geeigneter Vektoren aus einem Erzeugendensystem von V zu einer Basis von V ergänzen.

2. Der Beweis impliziert, dass folgende Aussagen gleichwertig sind:

(a) $\{b_1, \ldots, b_m\}$ ist eine Basis von V,
(b) $\{b_1, \ldots, b_m\}$ ist maximale linear unabhängige Teilmenge in V,
(c) $\langle b_1, \ldots b_m \rangle$ ist minimales Erzeugendensystem von V.

Beweis. Wir können oBdA annehmen, dass $\{c_1, \ldots, c_k\} \subseteq \{a_1, \ldots, a_n\}$ ist; sei also $a_i = c_i$ für $i = 1, \ldots, k$. Betrachte die Teilmengen

$$\{a_1, \ldots, a_k\} \subseteq \{a_1, \ldots, a_m\} \subseteq \{a_1, \ldots, a_n\},$$

wobei $1 \leq k \leq m \leq n$ und $\{a_1, \ldots, a_m\}$ eine maximale linear unabhängige Teilmenge von $\{a_1, \ldots, a_n\}$ ist. Ist $m < j \leq n$ so sind die Vektoren $\{a_1, \ldots, a_m, a_j\}$ linear abhängig, d.h. es gibt eine Relation

$$\sum_{i=1}^{m} \alpha_i a_i + \alpha_j a_j = 0,$$

wobei ein $\alpha_i \neq 0$ oder $\alpha_j \neq 0$ ist. Angenommen $\alpha_j = 0$. Da die $a_1, \ldots a_m$ linear unabhängig folgt dann $\alpha_1 = \cdots = \alpha_m = 0$, Widerspruch. Somit ist $\alpha_j \neq 0$ und damit auch $a_j \in \langle a_1, \ldots a_m \rangle$, da

$$a_j = -\sum_{i=1}^{n} (\alpha_j^{-1} \alpha_i) a_i \in \langle a_1, \ldots, a_m \rangle.$$

Also ist $\langle a_1, \ldots, a_m \rangle = V$, d.h. $\{a_1, \ldots a_m\}$ ist eine Basis von V. □

Nach Theorem 4.13 hat jeder endlich erzeugte K-Vektorraum eine Basis. Klar ist, dass eine solche Basis nicht eindeutig bestimmt ist (zum Beispiel ist für $V = \mathbb{R}^2$ sowohl $\{(1,0),(0,1)\}$, als auch $\{(1,0),(1,1)\}$ eine Basis). Wir zeigen, dass die Anzahl der Basiselemente eine Invariante des Vektorraums und unabhängig von der Wahl der Basis ist.

Lemma 4.14. *(Austauschlemma) Sei V ein K-Vektorraum und $B = \{b_1, \ldots, b_n\} \subseteq V$ eine Basis von V. Ist $b = \sum_{1=1}^{n} \alpha_i b_i$ mit $\alpha_i \in K$ und $\alpha_i \neq 0$, so ist auch $B' = \{b_1, b_2, \ldots, b_{i-1}, b, b_{i+1}, \ldots, b_n\} \subseteq V$ eine Basis von V.*

Beweis. Sei $b = \sum_{i=1}^{n} \alpha_i b_i$ mit $\alpha_i \neq 0$. Durch Umnumerieren können wir ohne Einschränkung annehmen, dass $i = 1$ ist. Dann ist

$$b_1 = \alpha_1^{-1}(b - \sum_{i=2}^{n} \alpha_i b_i) \in \langle b, b_2, \ldots, b_n \rangle$$

und es folgt $\langle b, b_2, \ldots b_n \rangle = \langle b_1, \ldots, b_n \rangle = V$. Es bleibt zu zeigen: Die Vektoren $\{b, b_2, \ldots b_n\}$ sind linear unabhängig. Angenommen

$$\beta b + \sum_{i=2}^{n} \beta_i b_i = 0, \quad \beta, \beta_i \in K.$$

Dann ist

$$\beta(\sum_{i=1}^{n} \alpha_i b_i) + \sum_{i=2}^{n} \beta_i b_i = (\beta\alpha_1)b_1 + \sum_{i=2}^{n} (\beta\alpha_i + \beta_i)b_i = 0.$$

Da die $\{b_1, \ldots, b_n\}$ linear unabhängig sind folgt dann

$$\beta\alpha_1 = 0 \text{ und } \beta_i + \beta\alpha_i = 0 \text{ für } i = 2, \ldots, n.$$

Wegen $\alpha_1 \neq 0$ ist $\beta = 0$ und damit auch $\beta_i = 0$ für $i = 2, \ldots, n$, d.h. die $\{b, b_2, \ldots, b_n\}$ sind linear unabhängig und daher eine Basis. □

Theorem 4.15. *(Austauschsatz von Steinitz) Sei V ein K-Vektorraum und $\{b_1, \ldots, b_n\} \subseteq V$ eine Basis von V. Ist $\{a_1, \ldots a_m\} \subseteq V$ eine linear unabhängige Teilmenge, so ist $m \leq n$ und mit geeigneter Numerierung der b_i ist $\{a_1, \ldots, a_m, b_{m+1}, \ldots, b_n\}$ ebenfalls eine Basis von V (d.h. jede linear unabhängige Menge von Vektoren aus V lässt sich durch Hinzunahme geeigneter Vektoren aus einer Basis zu einer Basis erweitern).*

Beweis. Induktion nach m. Ist $m = 1$, so ist $a_1 \neq 0$ und $a_1 = \sum_{i=1}^{n} \alpha_i b_i$ mit $\alpha_i \neq 0$ für ein i; ohne Einschränkung ist $i = 1$. Nach Lemma 4.14 ist $\{a_1, b_2, \ldots, b_n\}$ ebenfalls eine Basis von V. Sei nun $1 < m \leq n$. Da die $\{a_1, \ldots, a_{m-1}\}$ linear unabhängig sind gibt es nach Induktion eine Basis $\{a_1, \ldots, a_{m-1}, b_m, b_{m+1}, \ldots b_n\}$ von V. Angenommen

$$a_m = \sum_{j=1}^{m-1} \beta_j a_j + \sum_{i=m}^{n} \gamma_i b_i, \ \beta_j, \gamma_i \in K.$$

Da die $a_1, \ldots, a_m$ linear unabhängig sind ist a_m keine Linearkombination von $a_1, \ldots, a_{m-1}$ und es gibt ein $\gamma_i \neq 0$; wir können durch Umnumerierung annehmen, dass $\gamma_m \neq 0$ ist. Nach Lemma 4.14 können wir dann b_m durch a_m ersetzen, und erhalten eine Basis $\{a_1, \ldots, a_m, b_{m+1}, \ldots, b_n\}$.

Angenommen $\{a_1, \ldots, a_m\}$ ist eine linear unabhängige Teilmenge von V mit $m > n$. Dann folgt wie oben, dass $\{a_1, \ldots, a_m\}$ eine Basis von V bildet. Wegen $m > n$ ist $a_n \in \langle a_1, \ldots, a_m \rangle$, was der linearen Unabhängigkeit von $\{a_1, \ldots, a_m\}$ widerspricht, also ist $m \leq n$. □

Theorem 4.16. *Sei V ein endlich erzeugter K-Vektorraum. Dann hat V eine Basis $\{b_1, \ldots, b_n\}$ und jede Basis hat genau n Elemente.*

Beweis. Nach Theorem 4.13 hat V eine Basis $B = \{b_1, \ldots, b_n\}$. Sei $B' = \{b'_i \mid i \in I\} \subseteq V$ eine weitere Basis von V. Ist $|I| > n$, so gibt es in V eine linear unabhängige Menge $\{b'_1, \ldots, b'_{n+1}\}$ mit mehr als n Elementen; Widerspruch zu Theorem 4.15. Also ist $|I| \leq n$. Vertauschen der Rollen von B und B' liefert $n \leq |I|$, also ist $n = |I|$. □

Definition 4.17. Sei $\{0\} \neq V$ ein endlich erzeugter K-Vektorraum. Die Dimension (oder K-Dimension) $\dim_K V$ ist die Anzahl der Elemente einer (und damit jeder) Basis von V. Ist $V = \{0\}$, so setze $\dim_K V = 0$.

NB. Mit Hilfe des 'Zornschen Lemmas' kann man die Existenz von Basen in beliebigen K-Vektorräumen (d.h. nicht unbedingt endlich erzeugten K-Vektorräumen) beweisen, d.h. für einen beliebigen K-Vektorraum gibt es eine Basis $B = \{b_i \mid i \in I\} \subseteq V$, sodass sich jedes Element von V eindeutig als eine *endliche* Linearkombination dieser Basiselemente darstellen lässt. Der Beweis liefert die Existenz, ist aber nicht konstruktiv. Je zwei Basen haben die gleiche Mächtigkeit. Ist V endlich erzeugt, so schreibe $\dim V < \infty$; ist V nicht endlich erzeugt, so setze $\dim V = \infty$.

Beispiele 4.18. (a) Für jeden Körper K und $n \geq 1$ ist $\{e_1, \ldots, e_n\} \subseteq K^n$ eine Basis; die sogenannte Standardbasis. Also ist $\dim K^n = n$.

(b) Betrachte die Inklusionen von Körpern $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$. Der $\mathbb{Q}$-Vektorraum $\mathbb{Q} = \mathbb{Q}^1$ hat Dimension 1. Der $\mathbb{Q}$-Vektorraum $\mathbb{Q}(\sqrt{2})$ wird von $\{1, \sqrt{2}\}$ erzeugt. Da $\{1, \sqrt{2}\}$ auch linear unabhängig sind, ist $\dim_{\mathbb{Q}} \mathbb{Q}(\sqrt{2}) = 2$. Für die reellen Zahlen $\mathbb{R}$, aufgefasst als $\mathbb{Q}$-Vektorraum, gilt $\dim_{\mathbb{Q}} \mathbb{R} = \infty$: Hätte $\mathbb{R}$ eine endliche $\mathbb{Q}$-Basis, so wäre $\mathbb{R}$ abzählbar, Widerspruch.

(c) Für den $\mathbb{Q}$-Vektorraum $\mathbb{Q}[x]$ der Polynome mit rationalen Koeffizienten gilt ebenfalls $\dim_{\mathbb{Q}} \mathbb{Q}[x] = \infty$; eine (unendliche) Basis ist durch die Menge $\{x^n \mid n \in \mathbb{N}_0\}$ gegeben.

(d) Sei K ein Körper, M eine Menge, und $V = \mathrm{Abb}(M, K)$ der K-Vektorraum der Abbildungen $M \to K$. Für $m \in M$ sei $f_m : M \to K$ die Abbildung $f_m(m) = 1$ und $f_m(m') = 0$ für $m' \neq m$. Seien $m_1, \ldots, m_n$ paarweise veschiedene Elemente von M. Angenommen es gilt

$$\sum_{i=1}^{n} \alpha_i f_{m_i} = 0, \quad \alpha_i \in K.$$

Dann ist

$$0 = (\sum_{i=1}^{n} \alpha_i f_{m_i})(m_j) = \alpha_j \text{ für } j = 1, \ldots, n.$$

Also sind die $f_{m_1}, \ldots, f_{m_n}$ linear unabhängig. Ist $|M| = \infty$, so liefert dies beliebig grosse Mengen von linear unabhängigen Elementen in V, also ist $\dim_K V = \infty$. Ist $|M| = n < \infty$ endlich, so hat jedes Element

$f \in V$ eine Darstellung als eine endliche Linearkombination

$$f = \sum_{m \in M} f(m) f_m,$$

d.h. die $\{f_m \mid m \in M\}$ bilden eine Basis von V und $\dim_K V = |M| = n$.

Sei V ein K-Vektorraum und seien $U_i \subseteq V$ lineare Unterräume, $i = 1, \ldots, k$. Die Summe der U_i ist definiert als die Menge

$$U_1 + \cdots + U_k = \{u_1 + \cdots + u_k \mid u_i \in U_i\} \subseteq V.$$

Es ist $U_1 + \cdots + U_k = \langle \cup_{i=1}^k U_i \rangle \subseteq V$, insbesondere ist die Summe $U_1 + \cdots + U_k \subseteq V$ ein linearer Unterraum von V; siehe Übung.

Lemma 4.19. *Sei V ein K-Vektorraum endlicher Dimension $\dim_K V = n < \infty$ und sei $U \subseteq V$ ein linearer Unterraum. Dann gilt*

(a) $\dim_K U \leq \dim_K V$,

(b) *Jede Basis $\{b_1, \ldots, b_k\} \subseteq U$ von U lässt sich zu einer Basis $\{b_1, \ldots, b_k, b_{k+1}, \ldots, b_n\} \subseteq V$ von V ergänzen.*

(c) *Es gibt einen linearen Unterraum $W \subseteq V$, sodass gilt: $V = U + W$ und $U \cap W = \{0\}$ (der Unterraum W ist ein Komplement von U in V),*

(d) $\dim_K U = \dim_K V$ *genau dann, wenn $U = V$.*

Beweis. (a): Ist $U = \{0\}$, so ist $\dim_K U = 0 \leq n$. Ist $U \neq \{0\}$, und ist $B \subseteq U$ eine linear unabhängige Teilmenge, so ist nach Theorem 4.15 $|B| \leq n$. Insbesondere hat eine Basis von U maximal n Elemente und $\dim_K U \leq \dim_K V$.

(b): Folgt aus Theorem 4.15.

(c): Ist $\{b_1, \ldots, b_k\}$ eine Basis von U, so lässt sich diese nach (b) zu einer Basis $\{b_1, \ldots, b_k, b_{k+1}, \ldots, b_n\}$ von V ergänzen. Für den linearen Unterraum $W = \langle b_{k+1}, \ldots, b_n \rangle$ gilt $U + W = V$ und $U \cap W = \{0\}$.

(d): Sei $\dim_K U = \dim_K V = n$. Ist $\{b_1, \ldots, b_n\}$ eine Basis von U, so ist nach (b) $\{b_1, \ldots, b_n\}$ auch eine Basis von V und $U = \langle b_1, \ldots, b_n \rangle = V$. Die Umkehrung ist trivial. □

Beispiel 4.20. Ist M eine Menge und $N \subseteq M$ eine Teilmenge, so ist das mengentheoretische Komplement $\overline{N} = M \setminus N = \{m \in M \mid m \notin N\}$ eindeutig bestimmt. Dies gilt im allgemeinen nicht für Komplemente von linearen Unterräumen: Ist V ein endlich erzeugter K-Vektorraum und $\{0\} \subsetneq U \subsetneq V$ ein linearer Unterraum, so hat U stets mehrere Komplemente. Ist $\{b_1, \ldots, b_k\}$ eine Basis von V und $\{b_1, \ldots, b_n\}$ eine Basis von V, so definiert für jedes $\alpha \in K$ der lineare Unterraum

$$W_\alpha = \langle b_{k+1}, \ldots, b_{n-1}, b_n + \alpha b_1 \rangle$$

ein Komplement von U in V. Offensichtlich ist $U + W_\alpha = V$. Angenommen $v \in U \cap W_\alpha$. Dann gibt es Linearkombinationen

$$v = \sum_{j=k+1}^{n-1} \alpha_j b_j + \alpha_n(b_n + \alpha b_1) = \sum_{i=1}^{k} \alpha_i b_i \in U \cap W_\alpha.$$

Es ergibt sich eine Darstellung der 0 und Koeffizientenvergleich zeigt $\alpha_1 = \cdots = \alpha_n = 0$, also ist $U \cap W_\alpha = \{0\}$. Für verschiedene $\alpha \in K$ sind die W_α verschieden; ist der Körper K unendlich, so gibt es unendlich viele Komplemente von U in V. Konkret: Sei $V = \mathbb{R}^2$, $U = \langle(1,0)\rangle$ und $\{(1,0),(0,1)\}$ die Standardbasis von V. Für $\alpha \in \mathbb{R}$ ist $W_\alpha + \langle(0,1) = \alpha(1,0)\rangle = \langle(\alpha,1)\rangle$. Für $\alpha, \alpha' \in \mathbb{R}$, $\alpha \neq \alpha'$ ist $(\alpha,1) \neq (\alpha',1)$, d.h. die durch diese Vektoren erzeugten Unterräume W_α und $W_{\alpha'}$ sind verschiedene Geraden durch den Ursprung; jede von der x-Achse $U = \langle(1,0)\rangle$ verschiedende Gerade durch den Ursprung liefert ein Komplement von U in V.

5. Lineare Abbildungen und Faktorräume

Wir wollen Vektorräume mittels Abbildungen vergleichen. Ein Vektorraum ist eine Menge, zusammen mit einer 'algebraischen' Struktur ('Addition' und 'Skalarmultiplikation') und für uns wichtig sind diejenigen Abbildungen, die mit dieser Struktur 'verträglich' sind.

Abstrakt sollte eine strukturerhaltende Abbildung die folgende Eigenschaft haben: Ist M eine Menge mit einer Verküpfung $*_M$ und N eine Menge mit einer Verknüpfung $*_N$, so ist eine Abbildung von Mengen $f : M \to N$ mit diesen Verknüpfungen verträglich, falls gilt

$$f(m *_M m') = f(m) *_N f(m'),\ m, m' \in M,$$

d.h. es ist egal, ob man zuerst in M verknüpft und dann abbildet oder zuerst abbildet und dann in N verknüpft. Weiter sollte eine solche Abbildung das neutrale Element e_M bzgl. $*_M$ auf das neutrale Element e_N bzgl. $*_N$ abbilden. Solche strukturerhaltenden Abbildungen werden Homomorphismen genannt.

Zum Beispiel: Sind G, H Gruppen, so ist ein Gruppenhomomorphismus eine Abbildung $f : G \to H$, sodass für alle $g, g' \in G$ gilt

$$f(g \cdot g') = f(g) \cdot f(g'),$$

(wegen $1_H \cdot f(1_G) = f(1_G \cdot 1_G) = f(1_G) \cdot f(1_G)$ gilt $f(1_G) = 1_H$).

Sind K, L Körper, so ist ein Körperhomomorphismus eine Abbildung

$f : K \to L$, sodass für alle $a, b \in K$ die folgenden Formeln gelten

$$\begin{aligned} f(a+b) &= f(a)+f(b), \\ f(a\cdot b) &= f(a)\cdot f(b), \end{aligned}$$

(wie oben folgt dann auch $f(1_K) = 1_L$).

Mittels strukturerhaltender Abbildung ergibt sich ein evidenter Begriff von 'gleichwertigen' oder 'isomorphen' algebraischen Strukturen: gibt es eine bijektive Abbildung (die beiden Mengen haben 'gleichviele' Elemente) die strukturerhaltend ist (es ist egal wo man verknüpft), so sind die Strukturen isomorph (aber nicht unbedingt identisch).

Wir betrachten strukturerhaltende Abbildungen von Vektorräume:

Definition 5.1. Sei K ein Körper und seien V, W K-Vektorräume.

(1) Eine Abbildung $f : V \to W$ ist linear (genauer: K-linear oder ein Homomorphismus von K-Vektorräumen), falls für alle $a_1, a_2, a \in V$ und $\alpha \in K$ gilt:

$$f(a_1 + a_2) = f(a_1) + f(a_2) \text{ und } f(\alpha \cdot a) = \alpha \cdot f(a).$$

Sei $Hom(V, W) = Hom_K(V, W)$ die Menge aller K-linearen Abbildungen von V nach W; ist $V = W$, so schreibe $End_K(V) = Hom_K(V, V)$, die Elemente von $End_K(V)$ sind die Endomorphismen von V.

(2) Ist $f \in Hom_K(V, W)$, so definiere Kern und Bild von f als

$$\begin{aligned} \ker(f) &= \{a \in V \mid f(a) = 0\} \subseteq V, \\ \mathrm{im}(f) &= \{f(a) \mid a \in V\} \subseteq W. \end{aligned}$$

(3) Eine K-lineare Abbildung $f \in Hom_K(V, W)$ ist ein Monomorphismus (bzw. Epimorphismus, Isomorphismus), falls f injektiv (bzw. surjektiv, bijektiv) ist. Gibt es einen Isomorphismus $f : V \to W$, so sind V und W isomorph, $V \cong W$.

• Sei $f \in Hom_K(V, W)$. Dann ist $f(0) = 0$ und $f(-a) = -f(a)$.

Lemma 5.2. *Seien V, W K-Vektorräume und sei $f \in Hom_K(V, W)$.*

(a) $\ker(f) \subseteq V$ *und* $\mathrm{im}(f) \subseteq W$ *sind lineare Unterräume.*

(b) f *ist ein Monomorphismus* $\Leftrightarrow \ker(f) = \{0\}$.

Beweis. (a): Wegen $0 = f(0)$ ist $0 \in \ker(f)$ und $\ker(f) \neq \emptyset$. Sind $a_1, a_2 \in \ker(f)$, so ist wegen $f(a_1 + a_2) = f(a_1) + f(a_2) = 0 + 0 = 0$ auch $a_1 + a_2 \in \ker(f)$. Für $a \in \ker(f)$ und $\alpha \in K$ ist $f(\alpha a) = \alpha f(a) = \alpha \cdot 0 = 0$, d.h. $\alpha a \in \ker(f)$. Der Beweis für $\mathrm{im}(f)$ ist ähnlich einfach.

(b): Ist f ein Monomorphismus und $a \in \ker(f)$, so ist $f(a) = 0 = f(0)$ und da f injektiv ist folgt $a = 0$, also ist $\ker(f) = \{0\}$. Sei

umgekehrt $\ker(f) = \{0\}$. Sind $a_1, a_2 \in V$ mit $f(a_1) = f(a_2)$, so ist $0 = f(a_1) - f(a_2) = f(a_1 - a_2)$, d.h. $a_1 - a_2 \in \ker(f) = \{0\}$ und somit $a_1 = a_2$, d.h. f ist ein Monomorphismus. $\square$

Beispiele 5.3. (a) Einfache Beispiele von linearen Abbildungen sind die Identität $\mathrm{id} : V \to V, a \mapsto a$ und die Nullabbildung $f : V \to W, a \mapsto 0$. Insbesondere gilt: Ist $U \subseteq V$ ein Untervektorraum, so ist die Inklusion $i = \mathrm{id}_U : U \to V,\ u \mapsto u$ linear.

(b) Sei $V = \mathbb{R}^3$ und $W = \mathbb{R}^2$. Seien $\alpha_{ij} \in \mathbb{R}, i = 1, 2, j = 1, 2, 3$ gegeben. Wir ordnen diese Skalare als ein formales Schema an und betrachten die Elemente von V und W als Spaltenvektoren (x_j) und (y_i). Setze

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 \end{pmatrix}$$

Dann definiert das obige Zuordnungsschema eine lineare Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Konkret ergibt sich, zum Beispiel für die folgende Wahl der Skalare

$$\begin{array}{lll} \alpha_{11} = 1 & \alpha_{12} = 2 & \alpha_{13} = -4 \\ \alpha_{21} = -7 & \alpha_{22} = 8 & \alpha_{23} = -10 \end{array}$$

die lineare Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & -4 \\ -7 & 8 & -10 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 - 4x_3 \\ -7x_1 + 8x_2 - 10x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Allgemein lassen sich so lineare Abbildungen $f : V = K^n \to K^m = W$ definieren: Sind $\alpha_{ij} \in K,\ i = 1, \ldots, m,\ j = 1, \ldots n$ fest gewählte Skalare, so definiert die Zuordnung

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_m \end{pmatrix}$$

wobei

$$y_i = \sum_{j=1}^{n} \alpha_{ij} x_j,\ i = 1, \ldots m$$

eine lineare Abbildung $f : K^n \to K^m$. Der Kern von f besteht aus allen Vektoren $(x_1, \ldots, x_n)$ mit der Eigenschaft, dass

$$\sum_{j=1}^{n} \alpha_{ij} x_j = 0, \ i = 1, \ldots, m;$$

d.h. den gemeinsamen Lösungen der obigen m linearen Gleichungen (in den Variablen $x_1, \ldots, x_n$). Das Bild von f besteht aus den Vektoren $(y_1, \ldots, y_m)$, sodass die m Gleichungen (in den Variablen $x_1, \ldots, x_n$)

$$\sum_{j=1}^{n} \alpha_{ij} x_j = y_i, \ i = 1, \ldots m,$$

eine Lösung haben. Wir werden zeigen, dass jede lineare Abbildung $K^n \to K^m$ diese Form hat und Techniken zur Lösung solcher Gleichungssysteme entwickeln.

Das nächste Lemma ist fundamental: Es zeigt, dass eine lineare Abbildung auf einer Basis festgelegt ist und man dabei beliebige Werte auf einer Basis vorgeben kann.

Lemma 5.4. *Seien V und W K-Vektorräume, $\{a_j \mid j \in J\}$ eine Basis von V und $\{b_i \mid i \in I\}$ eine Basis von W.*

(a) *Seien $c_j \in W$, $j \in J$ beliebig vorgegeben. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit $f(a_j) = c_j$ für $j \in J$.*

(b) *Seien $\alpha_{ij} \in K, i \in I, j \in J$, sodass für $j \in J$ nur endlich viele $\alpha_{ij} \neq 0$ sind. Dann gibt es genau ein $f \in Hom_K(V, W)$ mit*

$$f(a_j) = \sum_{j \in J} \alpha_{ij} b_i, \ j \in J.$$

NB. Seien $\{a_1, \ldots, a_n\}$ und $\{b_1, \ldots, b_m\}$ Basen von V bzw. W. Nach (a) ist eine lineare Abbildung $f : V \to W$ durch die Bilder der Basisvektoren eindeutig bestimmt. Jedes der Bilder $f(a_j)$ besitzt eine eindeutige Darstellung als Linearkombination der b_i; d.h. für $j = 1, \ldots n$ ist somit

$$f(a_j) = \sum_{i=1}^{m} \alpha_{ij} b_i.$$

Wir ordnen $\alpha_{1j}, \dots \alpha_{mj}$ $(j = 1, \dots n)$ die für diese Darstellung des j-ten Basisvektoren auftreten als die Spalten einer sogenannten Matrix an

$$(\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1j} & \cdots & \alpha_{1m} \\ \alpha_{21} & \cdots & \alpha_{2j} & \cdots & \alpha_{2m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \cdots & \alpha_{mj} & \cdots & \alpha_{mn} \end{pmatrix}$$

Damit lässt sich die lineare Abbildung f bzgl. dieser Basen durch eine solche Matrix darstellen; wir werden die Frage, wie man durch eine geschickte Wahl dieser Basen eine 'einfache' Matrix bekommt später studieren. Umgekehrt besagt (b), dass ein Schema der Form (α_{ij}) eine eindeutige lineare Abbildung bestimmt.

Beweis. (a): Jedes $a \in V$ hat eine eindeutige Darstellung $a = \sum_{j=1}^{n} \alpha_j a_j$. Ist $f : V \to W$ eine lineare Abbildung mit $f(a_j) = c_j$, so folgt sofort

$$f(a) = f(\sum_{j=1}^{n} \alpha_j a_j) = \sum_{j=1}^{n} \alpha_j f(a_j) = \sum_{j=1}^{n} \alpha_j c_j.$$

Also ist f durch die Werte $f(a_j)$ auf den Basiselementen a_j eindeutig festgelegt, d.h. die Zuordnung $f(a_j) = c_j$ bestimmt eine eindeutige Abbildung $f : V \to W$. Weiter ist die oben definierte Abbildung

$$a = \sum_{j=1}^{n} \alpha_j a_j \mapsto f(a) = \sum_{j=1}^{n} \alpha_j f(a_j) = \sum_{j=1}^{n} \alpha_j c_j$$

linear: Sind $a = \sum_{j=1}^{n} \alpha_j a_j,\ a' = \sum_{j=1}^{m} \alpha'_j a_j$ Vektoren in V, so ist

$$a + a' = \sum_{j=1}^{n} \alpha_j a_j + \sum_{j=1}^{m} \alpha'_j a_j = \sum\nolimits_{j=1}^{n+m} (\alpha_j + \alpha'_j) a_j$$

die eindeutige Darstellung von $a+a'$ als endliche Linearkombination der Basiselemente a_i ist. Nach Definition von f folgt dann sofort $f(a+a') = f(a) + f(a')$. Ein ähnliches Argument zeigt $\alpha f(a) = f(\alpha a)$.
(b): Folgt aus (a) mit $c_j = \sum_{j \in J} \alpha_{ij} b_i$ (dies ist eine endliche Summe). □

Beispiel 5.5. Betrachte die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2,\ v \mapsto 2v$; offensichtlich ist f linear. Bzgl. der Basis $\{(1,0),(0,1)\}$ von $\mathbb{R}^2$ auf (beiden Seiten der Abbildung) ergibt sich

$$\begin{aligned} (1,0) &\mapsto (2,0) = 2 \cdot (1,0) + 0 \cdot (0,1), \\ (0,1) &\mapsto (0,2) = 0 \cdot (1,0) + 2 \cdot (0,1), \end{aligned}$$

also ist die Matrixdarstellung von f bzgl. der Basis $\{(1,0),(0,1)\}$

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Wählt man $\{(1,0),(0,1)\}$ als Basis für den 'Definitionsbereich' $\mathbb{R}^2$ und $\{(1,0),(1,1)\}$ als Basis für den 'Zielbereich' $\mathbb{R}^2$, so ist wegen

$$\begin{aligned} (1,0) &\mapsto (2,0) = 2\cdot(1,0) + 0\cdot(0,1), \\ (0,1) &\mapsto (0,2) = -2\cdot(1,0) + 2\cdot(1,1), \end{aligned}$$

die Matrixdarstellung von f bzgl. dieser beiden Basen

$$\begin{pmatrix} 2 & -2 \\ 0 & 2 \end{pmatrix}.$$

Theorem 5.6. *Seien V und W K-Vektorräume mit $\dim_K V = n < \infty$. Dann sind gleichwertig*

(a) $\dim_K W = n$,

(b) *Es gibt einen Isomorphismus $f : V \to W$, d.h. $V \cong W$.*

NB. Das Theorem besagt: Ist V ein n-dimensionale K-Vektorraum, so ist $V \cong K^n$. Der Beweis zeigt, dass dieser Isomorphismus von der Wahl von Basen von V und K^n abhängt, d.h. jede Wahl von solchen Basen liefert einen Isomorphismus (aber es es gibt keinen eindeutigen Isomorphismus).

Beweis. (a)⇒(b): Sei $\dim_K W = n$ und sei $\{a_1,\dots,a_n\}$ eine Basis von V und $\{b_1,\dots,b_n\}$ eine Basis Basen von W. Nach Lemma 5.4(a) gibt es eine eindeutige lineare Abbildung $f : V \to W$ mit $f(a_j) = b_j$ für $j = 1,\dots,n$; genauer, für $a = \sum_{j=1}^n \alpha_j a_j$ ist $f(a) = \sum_{j=1}^n \alpha_j f(a_j) = \sum_{j=1}^n \alpha_j b_j$. Diese Abbildung f ist ein Epimorphismus: Angenommen $b = \sum_{j=1}^n \beta_j b_j \in W$. Dann ist $a = \sum_{j=1}^n \beta_j a_j \in V$ und für dieses Element a gilt

$$f(a) = \sum_{j=1}^n \beta_j f(a_j) = \sum_{j=1}^n \beta_j b_j = b.$$

Die Abbildung f ist ein Monomorphismus: Nach Lemma 5.2(b) genügt es zu zeigen, dass $\ker(f) = \{0\}$ ist. Sei $a = \sum_{j=1}^n \alpha_j a_j \in \ker(f)$, sodass

$$0 = f(a) = \sum_{j=1}^n \alpha_j f(a_j) = \sum_{j=1}^n \alpha_j b_j;$$

da die b_i linear unabhängig sind folgt $\alpha_j = 0$ für alle j, d.h. $a = 0$.

(b)⇒(a): Sei $f : V \to W$ ein Isomorphismus und sei $\{a_1,\dots,a_n\}$ eine Basis von V. Wir zeigen $\{f(a_1),\dots,f(a_n)\}$ ist eine Basis von W. Sei

$b \in W$. Da f ein Epimorphismus ist gibt es ein $a \in V$ mit $f(a) = b$. Ist $a = \sum_{j=1}^{n} \alpha_j a_j$, so folgt $b = f(a) = \sum_{j=1}^{n} \alpha_j f(a_j)$, also ist $b \in \langle f(a_1), \ldots, f(a_n)\rangle$ und da $b \in W$ beliebig war folgt $\langle f(a_1), \ldots, f(a_n)\rangle = W$. Angenommen

$$0 = \sum_{j=1}^{n} \alpha_j f(a_j) = f(\sum_{j=1}^{n} \alpha_j a_j).$$

Dann ist $\sum_{j=1}^{n} \alpha_j a_j \in \ker(f)$. Da f ein Monomorphismus ist zeigt nochmalige Anwendung von Lemma 5.2(b), dass $\ker(f) = \{0\}$ ist, also ist $\sum_{j=1}^{n} \alpha_j a_j = 0$ und da die a_j linear unabhängig sind folgt $\alpha_j = 0$ für alle j, d.h. die $\{f(a_1), \ldots, f(a_n)\}$ sind linear unabhängig. □

Eine lineare Abbildung $f : V \to W$ ist durch die linearen Unterräume $\mathrm{im}(f) \subseteq W$ und $\ker(f) \subseteq V$ charakterisiert; das Bild $\mathrm{im}(f)$ sind die in W 'sichtbaren' Elemente, der Kern $\ker(f)$ die Elemente in V, die in W 'verlorengehen' (d.h. kein nicht-triviales Bild haben). Um diese linearen Räume studieren zu können führen wir Faktorräume ein.

Die Idee hier ist die, das Bild $\mathrm{im}(f)$ mit einem Quotienten- oder Faktorraum $V/\ker(f)$ zu identifizieren. Jeder lineare Unterraum $U \subseteq V$ ist insbesondere eine abelsche Untergruppe, sodass nach Lemma 2.6 $a_1 \sim a_2 \Leftrightarrow a_1 - a_2 \in U$ eine Äquivalenzrelation auf V definiert. Betrachte die Menge der (verschiedenen) Äquivalenzklassen

$$V/U = \{a + U \mid a \in V\},$$

zusammen mit der *surjektive* Abbildung $q : V \to V/U,\ a \mapsto a + U$.

Wir zeigen, dass die Addition und Skalarmultiplikation auf V analog eine Addition und Skalarmultiplikation auf V/U induziert, sodass V/U ein K-Vektorraum und $f : V \to V/U$ eine lineare Abbildung ist.

Definition 5.7. Sei V ein K-Vektorraum und $U \subseteq V$ ein linearer Unterraum. Für $a \in V$ setze $a + U = \{a + u \mid u \in U\} \subseteq V$. Der Quotienten- oder Faktorraum von V nach U ist die Menge

$$V/U = \{a + U \mid a \in V\}.$$

Lemma 5.8. *Sei V ein K-Vektorraum und $U \subseteq V$ ein linearer Unterraum. Dann ist der Faktorraum V/U ein K-Vektorraum mittels*

$$(a_1 + U) + (a_2 + U) = (a_1 + a_2) + U \text{ und } \alpha(a + U) = \alpha a + U.$$

Insbesondere gilt in V/U : $0 = 0 + U = U$.

Die Abbildung $q : V \to V/U,\ a \mapsto a + U$ ist ein Epimorphismus.

Beweis. Zu zeigen ist zunächst, dass die Operationen auf V/U wohldefiniert sind. Ist $a_1 + U = a_1' + U$ und $a_2 + U = a_2' + U$, so ist nach Definition $a_1 - a_1' = u_1 \in U$ und $a_2 - a_2' = u_2 \in U$. Damit folgt

$$(a_1 + a_2) + U - [(a_1' + a_2') + U] = (u_1 + U) - (u_2 + U) = U.$$

Ähnlich für die Skalarmultiplikation: Ist $a_1 + U = a_2 + U$, also $a_1 - a_2 = u \in U$, und ist $\alpha \in K$, so ist auch $\alpha(a_1 - a_2) = \alpha u \in U$ und somit $\alpha a_1 + U = \alpha a_2 + U$.

Weiter ist V/U mit dieser Addition und Skalarmultplikation ein K-Vektorraum; dies folgt, da diese Operationen von den entsprechenden Operationen auf dem K-Vektorraum V induziert sind. Zum Beispiel,

$$\alpha((a_1 + U) + (a_2 + U)) = \alpha(a_1 + U) + \alpha(a_2 + U)$$

da in V gilt $\alpha(a_1 + a_2) = \alpha a_1 + \alpha a_2$. Aus dem gleichen Grund ist die (surjektive) Abbildung $q : V \to V/U,\ a \mapsto a + U$ K-linear, d.h. ein Epimorphismus. □

Lemma 5.9. *Sei V ein K-Vektorraum und seien $U \subseteq W \subseteq V$ lineare Unterräume. Für die K-Vektorräume $W/U, V/W$ und V/U gilt:*

(a) *Sei $\{w_i + U \mid i \in I\}$ eine Basis von W/U und $\{v_j + W \mid j \in J\}$ eine Basis von V/W. Dann ist $\{w_i + U, v_j + U \mid i \in I, j \in J\}$ eine Basis von V/U.*

(b) *Ist $\dim V = n < \infty$, so ist $\dim V/U = \dim V - \dim U$.*

Beweis. (a): Für $a \in V$ betrachte das Element $a + W \in V/W$. Da die $\{v_j + W \mid j \in J\}$ eine Basis von V/W bilden gibt es Skalare $\alpha_j \in K$, sodass $a + W = \sum_{j=1}^{n} \alpha_j (v_j + W)$ und weiter $a - \sum_{j=1}^{n} \alpha_j v_j \in W$. Betrachte $(a - \sum_{j=1}^{n} \alpha_j v_j) + U \in W/U$. Da die Menge $\{w_i + U \mid i \in I\}$ eine Basis von W/U bildet gibt es Skalare $\beta_i \in K$, sodass

$$(a - \sum_{j=1}^{n} \alpha_j v_j) + U = \sum_{i=1}^{m} \beta_i (w_i + U),$$

und somit

$$a + U = \sum_{j=1}^{n} \alpha_j (v_j + U) + \sum_{i=1}^{m} \beta_i (w_i + U),$$

d.h. $V/U = \langle v_j + U, w_i + U \mid j \in J, i \in I \rangle$. Wir zeigen die $\{v_j + U, u_i + U \mid j \in J, i \in I\}$ sind linear unabhängig. Angenommen in V/U gilt

$$\sum_{j=1}^{n} \alpha_j (v_j + U) + \sum_{i=1}^{m} \beta_i (w_i + U) = U \quad (\text{in } V/U \text{ ist } 0 = U)$$

mit $\alpha_j, \beta_i \in K$. Es folgt $\sum_{j=1}^{n} \alpha_j v_j + \sum_{i=1}^{m} \beta_i w_i \in U$. Da $W \subseteq V$ ein linearer Unterraum ist folgt aus $w_i \in W$ dann auch $\sum_{i=1}^{m} \beta_i w_i \in W$ und in V/W gilt $\sum_{j=1}^{n} \alpha_j (v_j + W) = W$. Da die Menge $\{v_j + W \mid j \in J\}$ eine Basis von V/W bilden liefert dies $\alpha_j = 0$ für $j = 1, \dots n$. Wegen $w_i + U \in W/U$ ergibt sich jetzt in W/U die Identität

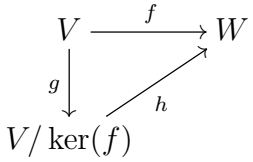
$$\sum_{i=1}^{m} \beta_i (w_i + U) = 0$$

und da die $\{w_i + U\}$ eine Basis von W/U bilden gilt $\beta_i = 0$ für $i = 1, \dots m$; dies beweist die Behauptung.

(b): Folgt aus (a) mit $U = W$. □

Theorem 5.10. *(Homomorphiesatz) Seien V, W K-Vektorräume und sei $f : V \to W$ eine K-lineare Abbildung.*

(a) *Es gibt einen Epimorphismus $g : V \to V/\ker(f)$ und einen Monomorphismus $h : V/\ker(f) \to W$, sodass $f = h \circ g$ und $\mathrm{im}(f) = \mathrm{im}(h)$ ist, d.h. das folgende Diagramm kommutiert*



und h induziert einen Isomorphismus $V/\ker(F) \cong \mathrm{im}(f)$.

(b) *Ist $\dim_K V = n < \infty$, so gilt die Formel*

$$\dim_K V = \dim_K \ker(f) + \dim_K \mathrm{im}(f).$$

Beweis. (a): Die Abbildung $g : V \to V/\ker(f)$ ist der evidente Epimorphismus $a \mapsto a + \ker(f)$. Die einzige Abbildung $h : V/\ker(f) \to W$, die die gewünschten Eigenschaften haben könnte ist definert durch

$$h : V/\ker(f) \to W,\ a + \ker(f) \mapsto f(a).$$

Zu zeigen ist: h ist ein wohldefinierter Monomorphismus mit $\mathrm{im}(f) = \mathrm{im}(h)$. Sei $a_1 + \ker(f) = a_2 + \ker(f)$. Dann ist $a_1 - a_2 \in \ker(f)$ und

$$h(a_1 + \ker(f)) - h(a_2 + \ker(f)) = f(a_1) - f(a_2) = f(a_1 - a_2) = 0,$$

d.h. h ist wohldefiniert. Weiter ist h linear, da f linear ist; es ist

$$\begin{aligned} h(\alpha(a + \ker(f)) &= h(\alpha a + \ker(f)) = f(\alpha a) = \alpha f(a) = \\ &= \alpha h(a + \ker(f)); \end{aligned}$$

und ähnlich für die Addition. Nach Konstruktion gilt für jedes $a \in V$

$$(h \circ g)(a) = h(a + \ker(f)) = f(a),$$

also ist $f = h \circ g$ und $\mathrm{im}(f) = \mathrm{im}(h)$. Ist $a + \ker(f) \in \ker(h)$, so ist

$$0 = h(a + \ker(f)) = f(a),$$

d.h. $a + \ker(f) = \ker(f)$ und h ist ein Monomorphismus.
(b): Nach (a) ist $V/\ker(f) \cong \mathrm{im}(f)$, d.h. diese beiden K-Vektorräume haben die gleiche Dimension $\dim_K(V/\ker(f)) = \dim_K \mathrm{im}(f)$. Im Fall $\dim_K V = n$ folgt mit Lemma 5.9(b) weiter

$$\dim_K \mathrm{im}(f) = \dim_K(V/\ker(f)) = \dim_K V - \dim_K \ker(f).$$

□

Lemma 5.11. *Seien V, W K-Vektorräume mit $\dim_K V = \dim_K W = n < \infty$ und $f : V \to W$ eine K-lineare Abbildung. Gleichwertig sind*

(a) *f ist ein Isomorphismus,*
(b) *f ist ein Monomorphismus,*
(c) *f ist ein Epimorphismus.*

Beweis. (a) ⇒ (b): Trivial. (b)⇒ (c): Ist f ein Monomorphismus, so ist $\ker(f) = \{0\}$. Weiter ist $\mathrm{im}(f) \subseteq W$ und f ist ein Epimorphismus genau dann, wenn $\mathrm{im}(f) = W$ ist; also nach Lemma 4.19(d) genau dann, wenn $\dim_K \mathrm{im}(f) = \dim_K W$ ist. Dies folgt aus dem Homorphiesatz 5.10(b): Wegen $\ker(f) = \{0\}$ ist $\dim_K \mathrm{im}(f) = \dim_K V = \dim_K W$.
(c)⇒ (a): Ist f ein Epimorphismus, so ist $\dim_K \mathrm{im}(f) = \dim_K W = \dim_K V$ und der Homomorphiesatz 5.10(b) liefert $\dim \ker(f) = 0$, d.h. $\ker(f) = \{0\}$ und f ist ein Monomorphismus. □

Wir geben eine geometrische Interpretation von Faktorräumen und dem Homomorphiesatz.

Beispiel 5.12. Sei $V = \mathbb{R}^2$, $W = \mathbb{R}^1$ und $f : V \to W$ die lineare Abbildung, die durch die Zuordnungen $(1, 0) \mapsto (1, 0)$ und $(0, 1) \mapsto (0, 0)$ bestimmt ist, d.h. $f(a, b) = a$; offensichtlich ist f ein Epimorphismus. Für $c \in W$ definiert das Urbild $f^{-1}(c) \subseteq V$ eine Teilmenge und nach Definition einer Abbildung ist für verschiedene $c, d \in W$ der Schnitt $f^{-1}(c) \cap f^{-1}(d) = \emptyset$, d.h. die Urbilder $\{f^{-1}(c) \mid c \in \mathbb{R}\}$ bilden eine Partition von V. Dabei ist $f^{-1}(0) = \ker(f) = \{(0, b) \mid b \in \mathbb{R}\} \subseteq V$ ein linearer Unterraum (die y-Achse) und für ein $c \in W$ ist das Urbild

$$\begin{aligned} f^{-1}(c) &= \{(c, b) \mid b \in \mathbb{R}\} = (c, 0) + \{(0, b) \mid b \in \mathbb{R}\} \\ &= (c, 0) + \ker(f) \subseteq V \end{aligned}$$

der ‘um $(c, 0)$ verschobene’ lineare Unterraum $\ker(f)$. Damit folgt

$$V/\ker(f) = \{(c, 0) + \ker(f) \mid c \in \mathbb{R}\}$$

d.h. die Elemente des Faktorraums sind die Urbilder von f und der Isomorphismus h aus dem Homomorphiesatz 5.10(a) ist die Abbildung

$$(c, 0) + \ker(f) \mapsto c;$$

insbesondere ist $c \mapsto f^{-1}(c)$ die dazu inverse Abbildung.

Wir verallgemeinern dieses Beispiel.

Definition 5.13. Sei V ein K-Vektorraum, $U \subseteq V$ ein linearer Unterraum und $a \in V$. Ein affiner Unterraum $A \subseteq V$ ist eine Teilmenge

$$A = a + U = \{a + u \mid u \in U\} \subseteq V;$$

insbesondere ist jedes Element $a + U$ des Faktorraums V/U ein affiner Unterraum.

• Ist $A = a + U \subseteq V$ ein affiner Unterraum, so ist U eindeutig durch A bestimmt: Sei $A = a + U = a' + U'$, wir zeigen $U = U'$. Nach Annahme ist $U' = a - a' + U$, also gibt es ein $u \in U$ mit $a - a' + u = 0$. Da U ein linearer Unterraum ist folgt $a - a' = -u \in U$, und damit $U' = a - a' + U = -u + U = U$.
• Da für $A = a + U$ der lineare Unterraum U eindeutig durch A bestimmt ist, lässt sich dem affinen Unterraum A eine Dimension zuordnen: $\dim_K A = \dim_K U$. Ein affiner Unterraum der Dimension 1 ist eine affine Gerade; im Fall $\dim_K V = n < \infty$ ist ein affiner Unterraum der Dimension $n - 1$ eine affine Hyperebene.
• Sei $f : V \to W$ eine lineare Abbildung. Das Urbild von $f(0) = 0$

$$f^{-1}(f(0)) = f^{-1}(0) = \ker(f) \subseteq V$$

ist ein linearer Unterraum. Ist $a \in V$ ein beliebiger Vektor, so ist

$$f^{-1}(f(a)) = \{a + u \mid u \in \ker(f)\} = a + \ker(f)$$

ein affiner Unterraum, dies ist die Faser von f über a.

NB. Ist $f : V \to W$ linear, so ist $f : V \to \operatorname{im}(f)$ surjektiv und zu jedem $b \in \operatorname{im}(f)$ gibt es ein $a \in V$ mit $f(a) = b$, d.h. über jedem $b \in \operatorname{im}(f)$ liegt genau eine Faser $f^{-1}(b) = a + \ker(f)$. Die Aussage des Homomorphiesatzes ist, dass die Abbildung $b \mapsto f^{-1}(b)$ einen Isomorphismus $\operatorname{im}(f) \cong V/\ker(f)$ von Vektorräumen liefert.

Wir zeigen jeder affine Unterraum ist die Faser einer linearen Abbildung.

Lemma 5.14. *Sei V ein K-Vektorraum und $\emptyset \neq A \subseteq V$ eine Teilmenge. Dann sind gleichwertig:*

(a) *A ist ein affiner Unterraum, d.h. es gibt einen linearen Unterraum $U \subseteq V$ und ein $a \in V$ mit $A = a + U$,*

(b) *Es gibt einen K-Vektorraum W und eine lineare Abbildung $f : V \to W$, sodass A eine Faser von f ist, d.h. $A = f^{-1}(b)$ für ein $b \in W$,*

(c) *Seien $a_0, \ldots, a_k \in A$ und $\alpha_0, \ldots, \alpha_k \in K$ mit $\sum_{i=0}^{k} \alpha_i = 1$. Dann ist $\sum_{i=0}^{k} \alpha_i a_i \in A$.*

Beweis. (a)⇒(b): Sei $A = a + U$. Für den Epimorphismus $f : V \to V/U,\ a \mapsto a + U$ gilt $f^{-1}(f(a)) = a + \ker(f) = a + U = A$.
(b)⇒(c): Sei $A = f^{-1}(b)$ für ein $b \in W$. Seien $a_0, \ldots, a_k \in A$ und $\alpha_0, \ldots, \alpha_k \in K$ mit $\sum_{i=0}^{k} \alpha_i = 1$ gegeben. Die Linearität von f liefert

$$f(\sum_{i=0}^{k} \alpha_i a_i) = \sum_{i=0}^{k} \alpha_i f(a_i) = (\sum_{i=0}^{k} \alpha_i) b = 1 \cdot b = b,$$

d.h. $\sum_{i=0}^{k} \alpha_i a_i \in f^{-1}(b) = A$.
(c)⇒(a): Wähle fest ein $a_0 \in A$ und betrachte die Menge $\Delta A = \{a - a_0 \mid a \in A\} \subseteq V$. Es ist $A = a_0 + \Delta A$; wir zeigen $\Delta A \subseteq V$ ist ein linearer Unterraum. Wegen $a_0 \in A$ ist $0 \in \Delta A$, also ist $\Delta A \neq \emptyset$. Für $a_1 - a_0, a_2 - a_0 \in \Delta A$ ist nach (c) $a_1 + (-1)a_0 + a_2 \in A$, sodass

$$(a_1 - a_0) + (a_2 - a_0) = (a_1 - a_0 + a_2) - a_0 \in \Delta A.$$

Für $\alpha \in K$ folgt aus (c) weiter $\alpha a_1 + (1 - \alpha)a_0 \in A$; somit gilt auch

$$\alpha(a_1 - a_0) = (\alpha a_1 + (1 - \alpha)a_0) - a_0 \in \Delta A.$$

□

Beispiel 5.15. Sei V ein K-Vektorraum und $a_0, \ldots, a_k \in V$. Dann ist der kleinste affine Unterraum $A \subseteq V$, der die $\{a_0, \ldots, a_k\}$ enthält

$$A = \{a_0 + \sum_{i=1}^{k} \alpha_i (a_i - a_0) \mid \alpha_1, \ldots, \alpha_k \in K\} \subseteq V.$$

Nach Definition ist $A = a_0 + U$, wobei U der von den Vektoren $a_1 - a_0, \ldots, a_k - a_0$ erzeugte lineare Unterraum ist, d.h. A ist affiner Unterraum. Da $a_0 + \sum_{i=1}^{k} \alpha_i (a_i - a_0) = (1 - \sum_{i=1}^{k} \alpha_i) a_0 + \sum_{i=1}^{k} \alpha_i a_i$ folgt

$$A = \{\sum_{i=0}^{k} \alpha_i a_i \mid \alpha_0, \ldots, \alpha_k \in K \text{ mit } \sum_{i=0}^{k} \alpha_i = 1\}.$$

Nach Lemma 5.14(c) ist dies der kleinste affine Unterraum, der die Vektoren $a_0, \ldots, a_k$ enthält. Konkret: Sei $V = \mathbb{R}^2$ und seien $a_0, a_1 \in \mathbb{R}^2$ zwei Punkte. Der kleinste affine Unterraum der a_0 und a_1 enthält ist

$$A = \{a_0 + \alpha(a_1 - a_0) \mid \alpha \in \mathbb{R}\} = \{\alpha a_0 + \beta a_1 \mid \alpha,\ \beta \in \mathbb{R}, \alpha + \beta = 1\},$$

d.h. die Gerade durch die beiden Punkte a_0 und a_1.

6. Lineare Abbildungen und Matrizen

Wir studieren lineare Abildungen und zeigen dazu zunächst, dass die Menge der K-linearen Abbildungen $Hom_K(V,W)$ selbst ein K-Vektorraum ist. Damit hat $Hom_K(V,W)$ eine Basis und jede lineare Abbildung $V \to W$ hat eine eindeutige Darstellung als eine endliche Linearkombination von Elementen einer solchen Basis.

Lemma 6.1. *Seien V, W K-Vektorräume.*

(a) *Für $f, g \in Hom_K(V,W)$, $\alpha \in K$ und $a \in V$ setze*

$$(f+g)(a) = f(a) + g(a) \text{ und } (\alpha f)(a) = \alpha f(a);$$

mit diesen Operationen ist $Hom_K(V,W)$ ein K-Vektorraum.

(b) *Seien $\{a_1, \ldots, a_n\} \subseteq V$ und $\{b_1, \ldots, b_m\} \subseteq W$ Basen. Für $j = 1, \ldots, n$ und $i = 1, \ldots, m$ definiere $e_{ij} \in Hom_K(V,W)$ durch*

$$e_{ij}(a_k) = \begin{cases} 0 & j \neq k \\ b_i & j = k \end{cases}$$

Dann ist $\{e_{11}, \ldots, e_{mn}\}$ eine Basis von $Hom_K(V,W)$; insbesondere ist $\dim_K Hom_K(V,W) = \dim_K V \cdot \dim_K W$.

Beweis. (a): Nachrechnen.

(b): Ist $f \in Hom_K(V,W)$ so gilt bezüglich der gegebenen Basen

$$f(a_j) = \sum_{i=1}^{m} \alpha_{ij} b_i, \ j = 1, \ldots n.$$

Für diese Koeffizienten $\alpha_{ij}, i = 1, \ldots m, j = 1, \ldots n$ bilde

$$g = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} e_{ij}.$$

Nach Definition hat die lineare Abbildung g die Eigenschaft, dass

$$g(a_j) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} e_{ij}(a_j) = \sum_{i=1}^{m} \alpha_{ij} b_i = f(a_j),$$

d.h. $f = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} e_{ij}$ und die e_{ij} erzeugen $Hom_K(V,W)$. Sei

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} e_{ij} = 0, \ \beta_{ij} \in K.$$

Evaluierung dieser Abbildung auf a_j liefert für $j = 1, \ldots n$ die Identität

$$0 = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} e_{ij}(a_j) = \sum_{i=1}^{m} \beta_{ij} b_i.$$

Da die b_i linear unabhängig sind folgt $\beta_{ij} = 0$ für alle i, j, also sind $\{e_{11}, \ldots, e_{mn}\}$ linear unabhängig und bilden eine Basis. □

Sind $f : V_1 \to V_2$ und $g : V_2 \to V_3$ lineare Abbildungen, so schreibe $gf = g \circ f : V_1 \to V_3$ für die Komposition. Diese Komposition ist allgemein für Abbildungen definiert; wir zeigen, das die Verknüpfung von linearen Abbildungen wieder linear, mit der Addition verträglich und assoziativ ist.

Lemma 6.2. *Seien V_i K-Vektorräume, $i = 1, 2, 3, 4$.*

(a) *Sind $g \in Hom_K(V_2, V_3)$ und $f \in Hom_K(V_1, V_2)$ so definiert*
$$(gf)(a_1) = g(f(a_1)),\ a_1 \in V_1$$
eine lineare Abbildung $fg \in Hom_K(V_1, V_3)$.

(b) *Ist $g \in Hom_k(V_2, V_3)$ und sind $f_1, f_2 \in Hom_K(V_1, V_2)$, so gilt*
$$g(f_1 + f_2) = gf_1 + gf_2.$$

(c) *Sind $g_1, g_2 \in Hom_K(V_2, V_3)$ und $f \in Hom_K(V_1, V_2)$, so gilt*
$$(g_1 + g_2)f = g_1 f + g_2 f.$$

(d) *Sei $h \in Hom_K(V_3, V_4)$, $g \in Hom_K(V_2, V_3)$, $f \in Hom_K(V_1, V_2)$. Dann ist*
$$h(gf) = (hg)f.$$

Beweis. Nachrechnen. □

Wir betrachten Isomorphismen in $Hom_K(V, W)$. Ist $f : V \to W$ ein Isomorphismus, so ist f eine bijektive Abbildung und hat damit eine inverse bijektive Abbildung $g = f^{-1} : W \to V$. Wir zeigen, dass diese Abbildung $g = f^{-1}$ ebenfalls linear ist. Damit gelten für Isomorphismen diesselben Beziehungen wie für bijektive Abbildungen von Mengen.

Lemma 6.3. *Seien V_i K-Vektorräume, $i = 1, 2, 3$.*

(a) *Sei $f \in Hom_K(V_1, V_2)$ ein Isomorphismus. Dann gibt es genau ein $g \in Hom_K(V_2, V_1)$ mit $gf = \mathrm{id}_{V_1}$ und $fg = \mathrm{id}_{V_2}$; setze $g = f^{-1}$.*

(b) *Sind $f \in Hom_K(V_1, V_2)$ und $g \in Hom_K(V_2, V_3)$ Isomorphismen, so ist auch $gf \in Hom_K(V_1, V_3)$ ein Isomorphismus; es gilt: $(gf)^{-1} = f^{-1}g^{-1}$.*

Beweis. (a): Da $f : V_1 \to V_2$ eine Bijektion ist, gibt es nach Lemma 1.12 genau eine Bijektion $g : V_2 \to V_1$ mit $gf = \mathrm{id}_{V_1}$ und $fg = \mathrm{id}_{V_2}$. Wir zeigen g, dass linear ist. Für $a_2, a_2' \in V_2$ gilt

$$\begin{aligned} f(g(a_2 + a_2')) &= (fg)(a_2 + a_2') = \mathrm{id}_{V_2}(a_2 + a_2') = \mathrm{id}_{V_2}(a_2) + \mathrm{id}_{V_2}(a_2') \\ &= f(g(a_2)) + f(g(a_2')) = f(g(a_2) + g(a_2')). \end{aligned}$$

Da f injektiv ist, folgt damit $g(a_2 + a_2') = g(a_2) + g(a_2')$. Ähnlich zeigt man $\alpha g(a_2) = g(\alpha a_2)$.
(b): Folgt aus der Bemerkung nach Definition 1.13. □

Beispiel 6.4. Im Fall $V = W$ ist $Hom_K(V, W) = End_K(V)$. Nach Lemma 6.1(a) ist $End_K(V)$ ein K-Vektorraum. Die Verknüpfung von Endomorphismen $f : V \to V$ liefert eine 'Multiplikation' auf $End_K(V)$

$$f, g \in End_K(V) \Rightarrow fg = f \circ g \in End_K(V)$$

mit Einselement $\mathrm{id}_V : V \to V,\ \mathrm{id}_V(a) = a$. Für diese Multiplikation gelten die beiden Distributivgesetzen und das Assoziativgesetz. Weiter ist diese Multiplikation mit der Skalarmultiplikation verträglich, d.h.

$$\alpha(fg) = (\alpha f)g = f(\alpha g);\ f, g \in End_K(V),\ \alpha \in K.$$

Ein K-Vektorraum, zusammen mit einer Multiplikation, welche die obigen Verträglichkeitsbedingungen erfüllt ist eine K-Algebra.

Das Kroneckersymbol δ_{jk} ist definiert als $\delta_{jk} = 1$ falls $j = k$ und $\delta_{jk} = 0$ falls $j \neq k$. Ist $\{a_1, \ldots, a_n\}$ eine Basis von V, so bilden nach Lemma 6.1(b) die Endomorphismen $e_{ij} \in End_K(V)$ mit

$$e_{ij}(a_k) = \delta_{jk} a_i$$

eine Basis $\{e_{11}, \ldots, e_{nn}\}$ von $End_K(V)$; es ist $\dim_K End_K(V) = n^2$.
Für die Basiselemente $\{e_{ij}\}$ von $End_K(V)$ gelten die Formeln

$$e_{ij} e_{kl} = \delta_{jk} e_{il} \quad \text{und} \quad \sum_{i=1}^{n} e_{ii} = \mathrm{id}_V\,.$$

Ist $\dim_K V > 1$, so ist die Multiplikation via Verknüpfung von Abbildungen in $End_K(V)$ *nicht* kommutativ: Die obige Formel liefert $e_{12}e_{22} = \delta_{22}e_{12} = e_{12} \neq 0$ und $e_{22}e_{12} = \delta_{21}e_{22} = 0$, d.h. $e_{12}e_{22} \neq e_{22}e_{12}$.

Definition 6.5. Sei V ein K-Vektorraum. Ist $f \in End_K(V)$ ein Isomorphismus, so ist f regulär (auch 'invertierbar' bzw. 'Automorphismus'); ist f nicht regulär, so ist f singulär. Die regulären Abbildungen aus $End_K(V)$ bilden bzgl. der Verknüpfung von Endomorphismen eine multiplikative Gruppe mit neutralem Element id_V (vgl. Lemma 6.3); diese Gruppe bezeichnen wir mit $GL(V)$ (General Linear group).

Beispiel 6.6. Sei $K = \mathbb{Z}/p\mathbb{Z}$ der Körper mit p Elementen und sei V ein K-Vektorraum der Dimension n, d.h. $V \cong (\mathbb{Z}/p\mathbb{Z})^n$. Für zwei (beliebige) endlich-dimensionale K-Vektorräume V, W und $f \in Hom_K(V, W)$ gilt: f ist ein Isomorphismus genau dann, wenn f jede Basis von V auf eine Basis von W abbildet. Also ist die Anzahl der Elemente von

$GL(V)$ genau die Anzahl der verschiedenen Basen von V. Jede Basis $\{a_1, \ldots, a_n\}$ von V entsteht durch Wahl der a_i wie folgt:

$$\begin{array}{lll} 0 \neq a_1 \in V & p^n - 1 & \text{Möglichkeiten,} \\ a_1 \in V \setminus \langle a_1 \rangle & p^n - p & \text{Möglichkeiten,} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ a_n \in V \setminus \langle a_1, \ldots, a_{n-1} \rangle & p^n - p^{n-1} & \text{Möglichkeiten.} \end{array}$$

Damit ist $|GL(V)| = (p^n - 1)(p^n - p) \cdots (p^n - p^{n-1})$.

Definition 6.7. Seien V, W K-Vektorräume und sei $f \in Hom_K(V, W)$. Ist $\dim_K \operatorname{im}(f) = n < \infty$, so ist der Rang $r(f)$ von f definiert als

$$r(f) = \dim_K \operatorname{im}(f).$$

- Wegen $\operatorname{im}(f) \subseteq W$ ist stets $r(f) \leq \dim_K W$.
- Aus dem Homomorphiesatz 5.10(b) folgt für $\dim_K V = n < \infty$:

$$r(f) = \dim_K \operatorname{im}(f) = \dim_K V - \dim_K \ker(f)$$

Definition 6.8. Sei K ein Körper. Eine Matrix vom Typ (m, n) über K ist ein Schema von Skalaren $\alpha_{ij} \in K$ der folgenden Form

$$A = (\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \alpha_{m1} & \alpha_{m2}, & \ldots & \alpha_{mn} \end{pmatrix},$$

d.h. eine Matrix $A = (\alpha_{ij})$ vom Typ (m, n) besteht aus m Zeilen

$$z_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}),\ i = 1, \ldots, m$$

und n Spalten

$$s_j = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \cdot \\ \cdot \\ \alpha_{mj} \end{pmatrix},\ j = 1, \ldots n.$$

Sei $K^{m \times n}$ die Menge aller Matrizen vom Typ (m, n) über K.

Beispiel 6.9. Sei $K = \mathbb{R}$ und $m = 2 = n$. Seien $A = (\alpha_{ij}), B = (\beta_{ij})$ Matrizen vom Typ $(2, 2)$ über $\mathbb{R}$ und sei $\alpha \in \mathbb{R}$ ein Skalar. Setze

$$\begin{aligned} A + B &= \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} = \begin{pmatrix} \alpha_{11} + \beta_{11} & \alpha_{12} + \beta_{12} \\ \alpha_{21} + \beta_{21} & \alpha_{22} + \beta_{22} \end{pmatrix}, \\ \alpha A &= \alpha \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = \begin{pmatrix} \alpha\alpha_{11} & \alpha\alpha_{12} \\ \alpha\alpha_{21} & \alpha\alpha_{22} \end{pmatrix} \end{aligned}$$

Die Menge $\mathbb{R}^{2\times 2}$ der $(2,2)$-Matrizen über $\mathbb{R}$ ist mittels dieser Addition und Skalarmultplikation ein K-Vektorraum. Die Abbildung

$$\Theta : \mathbb{R}^{2\times 2} \to \mathbb{R}^4 : \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \mapsto (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})$$

ist ein Isomorphismus, also ist $\dim_{\mathbb{R}} \mathbb{R}^{2\times 2} = 4$. Ist $\{e_1, \ldots, e_4\}$ die Standardbasis von $\mathbb{R}^4$, so sind die Urbilder $\Theta^{-1}(e_1), \ldots \Theta^{-1}(e_4)$ genau die Matrizen E_{ij} mit einer 1 in der Stelle (i,j) und 0 sonst. Da Θ ein Isomorphismus ist bilden die $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ eine Basis von $\mathbb{R}^{2\times 2}$.

Das obige Beispiel hängt weder von $K = \mathbb{R}$ noch von $m = 2 = n$ ab, d.h. die Eigenschaften aus diesem Beispiel gelten allgemein für die Matrizen $K^{m\times n}$ vom Typ (m,n); genauer:

• Die Menge $K^{m\times n}$ aller Matrizen vom Typ (m,n) ist mittels

$$(\alpha_{ij}) + (\beta_{ij}) = (\alpha_{ij} + \beta_{ij}) \text{ und } \alpha(\alpha_{ij}) = (\alpha\alpha_{ij})$$

ein K-Vektorraum. Die Abbildung $\Theta : K^{m\times n} \to K^{mn}$

$$(\alpha_{ij}) \mapsto (\alpha_{11}, \cdots, \alpha_{1n}, \alpha_{21}, \cdots, \alpha_{2n}, \cdots, \alpha_{m1}, \ldots, \alpha_{mn})$$

definiert einen Isomorphismus von K-Vektorräumen $K^{m\times n} \cong K^{mn}$. Insbesondere ist $\dim_K K^{m\times n} = mn$.

• Sei $E_{ij} \in K^{m\times n}$ die Matrix mit 1 an der Stelle (i,j) und 0 sonst. Dann bilden die E_{ij} $(i = 1, \ldots, m; j = 1, \ldots, n)$ eine Basis von $K^{m\times n}$.

Seien V, W K-Vektorräume mid $\dim_K V = n$ und $\dim_K W = m$. Wir ordnen einer linearen Abbildung $f : V \to W$ eine Matrix in $K^{m\times n}$ zu (abhängig von der Wahl von Basen von V und W), um dann lineare Abbildungen mittels dieser zugeordneten Matrizen studieren.

Definition 6.10. Seien V und W K-Vektorräume, $X = \{v_1, \ldots, v_n\}$ eine Basis von V und $Y = \{w_1, \ldots, w_m\}$ eine Basis von W. Für eine lineare Abbildung $f \in Hom_K(V,W)$ haben die Werte $f(v_j)$ eine eindeutige Darstellung als endliche Linearkombiniation der w_i, d.h.

$$f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i, \ j = 1, \ldots n.$$

Setze $A = A_{f,X,Y} = (\alpha_{ij}) \in K^{m\times n}$; die Matrix $A_{f,X,Y}$ ist die Matrix von f bezüglich der Basen X und Y. Ist $V = W$ und $X = Y$, so schreibe $A_{f,X}$ für $A_{f,X,Y}$.

NB. Die Matrix $A_{f,X,Y}$ hängt von den Basen X, Y und von der Anordnung der Vektoren in diesen Basen ab.

Proposition 6.11. *Seien U, V, W K-Vektorräume mit Basen $X = \{u_1, \ldots, u_k\}, Y = \{v_1, \ldots, v_n\}$ und $Z = \{w_1, \ldots, w_m\}$. Dann gilt:*

(a) *Die Abbildung* $\kappa : Hom_K(U,V) \to K^{n\times k},\ f \mapsto A_{f,X,Y}$ *ist ein Isomorphismus.*

(b) *Seien* $g \in Hom_K(U,V)$ *und* $f \in Hom_K(V,W)$*. Sind* $A_{g,X,Y} = (\alpha_{ij})$ *und* $A_{f,X,Y} = (\beta_{rs})$*, so ist* $A_{fg,X,Z} = (c_{is})$ *mit*

$$c_{is} = \sum_{j=1}^{n} \alpha_{ij} b_{js}.$$

Beweis. (a): Da $Hom_K(U,V)$ und $K^{n\times k}$ die gleiche Dimension kn haben, genügt es nach Lemma 5.11 zu zeigen, dass κ linear und ein Monomorphismus ist. Dabei ist die zweite Aussage trivial: Nach Definition ist $A_{f,X,Y} = 0$ die Nullmatrix genau dann, wenn $f = 0$ ist. Wir zeigen die Abbildung κ ist linear: Seien $f_1, f_2 \in Hom_K(U,V)$ gegeben durch

$$f_1(u_j) = \sum_{i=1}^{n} \alpha_{ij} v_i \text{ und } f_2(u_j) = \sum_{i=1}^{n} \alpha'_{ij} v_i,\ j = 1, \dots n,$$

d.h. f_1 und f_2 entsprechen bzgl. der Basen X und Y den Matrizen

$$A_{f_1,X,Y} = (\alpha_{ij}) \text{ und } A_{f_2,X,Y} = (\alpha'_{ij}).$$

Dann gilt

$$(f_1 + f_2)(u_j) = \sum_{i=1}^{n} (\alpha_{ij} + \alpha'_{ij}) v_i,\ j = 1, \dots, n$$

also hat die lineare Abbildung $f_1 + f_2$ bzgl. X und Y die Matrix

$$A_{f_1+f_2,X,Y} = (\alpha_{ij} + \alpha'_{ij}) = (\alpha_{ij}) + (\alpha'_{ij}) = A_{f_1,X,Y} + A_{f_2,X,Y}$$

und es gilt $\kappa(f_1+f_2) = \kappa(f_1)+\kappa(f_2)$. Der Beweis von $\kappa(\alpha f_1) = \alpha\kappa(f_1)$, d.h. $A_{\ \alpha f_1,X,Y} = \alpha(A_{f_1,X,Y})$, ist ähnlich einfach.

(b): Aus $f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i$ und $g(u_s) = \sum_{j=1}^{n} \beta_{js} v_j$ folgt direkt

$$\begin{aligned}(fg)(u_s) &= f(g(u_s)) = f(\sum_{j=1}^{n} \beta_{js} v_j) = \sum_{j=1}^{n} \beta_{js} (f(v_j)) = \\ &= \sum_{j=1}^{n} \beta_{js} (\sum_{i=1}^{m} \alpha_{ij} w_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} (\alpha_{ij}\beta_{js}) w_i.\end{aligned}$$

Also ist $A_{fg,X,Z} = (c_{is})$ mit $c_{is} = \sum_{j=1}^{n} \alpha_{ij}\beta_{js}$. □

Der Beweis von (b) zeigt, wie Matrizen (von kompatiblem Typ) zu multiplizieren sind, sodass das entsprechende Produkt der Komposiition der zugrundeliegenden linearen Abbildungen entspricht. Dies motiviert folgende Definition des allgemeinen Matrizenprodukts.

Definition 6.12. Sei K ein Körper und seien $A = (\alpha_{ij}) \in K^{m\times n}$, sowie $B = (\beta_{jl}) \in K^{n\times k}$. Definere das Produkt $AB \in K^{m\times k}$ als die Matrix

$$AB = (c_{il}) \text{ mit } c_{il} = \sum_{j=1}^{n} \alpha_{ij}\beta_{jl}$$

• Sind $A = (\alpha_{ij}), A' = (\alpha'_{ij}) \in K^{m\times n}$ und $B = (\beta_{jl}), B' = (\beta'_{jl}) \in K^{n\times k}$, so gelten für die Matrixmultiplikation die beiden Distributivgesetze

$$(A + A')B = AB + A'B \text{ und } A(B + B') = AB + AB'.$$

• Für $A = (\alpha_{ij}) \in K^{m\times n}, B = (\beta_{jl}) \in K^{n\times k}$ und $C = (\gamma_{lr}) \in K^{k\times s}$ ist

$$A(BC) = (AB)C.$$

Beispiele 6.13. (a) Rein formal gilt

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

oder

$$\begin{pmatrix} 3 & 0 \\ 2 & 0 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -3 \\ 2 & -2 \\ 9 & 7 \end{pmatrix}.$$

(b) Betrachte die linearen Abbildungen

$$g : \mathbb{R}^3 \to \mathbb{R}^2, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

$$f : \mathbb{R}^2 \to \mathbb{R}^3, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Bezüglich der Standardbasen sind die g und f zugeordneten Matrizen

$$A_g = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \end{pmatrix} \text{ und } A_f = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Für das Kompositum $fg : \mathbb{R}^3 \to \mathbb{R}^3$ gilt nach Definition von f und g

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix},$$

also ist

$$A_{fg} = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \end{pmatrix} = A_f \cdot A_g.$$

Ist $\dim_K V = n < \infty$, so besagen die obigen Verträglichkeitsaussagen für Matrizen, dass $K^{n\times n}$ nicht nur ein K-Vektorraum, sondern auch eine K-Algebra ist, vgl. Beispiel 6.4. Dabei ist das Einselement in $K^{n\times n}$

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

die Einheitsmatrix vom Typ (n, n). Ist X eine Basis von V, so ist der Isomorphismus von K-Vektorräumen von Proposition 6.11(a)

$$\kappa : End_K(V) \to K^{n\times n},\ f \mapsto A_{f,X}$$

in Fakt ein Isomorphismus von K-Algebren (d.h. mit Produkten verträglich, d.h. $\kappa(gf) = \kappa(g)\kappa(f)$). Im folgenden verwenden wir diesen Isomorphismus um Aussagen über Endomorphismen in Aussagen über Matrizen zu übersetzen.

Nach Lemma 6.3 ist $f \in End_K(V)$ ein Automorphismus genau dann, wenn es ein $g \in End_K(V)$ mit $gf = \mathrm{id}_V$ und $fg = \mathrm{id}_V$ gibt. Ist X eine Basis von V, so folgt aus $gf = \mathrm{id}_V$ und $fg = \mathrm{id}_V$ durch Anwendung des Isomorphismus $\kappa : End_K(V) \to K^{n\times n}$ wegen $\kappa(\mathrm{id}_V) = E_n$ dann

$$A_{f,X}A_{g,X} = E_n = A_{g,X}A_{f,X}.$$

Dies motiviert folgende Definition für Matrizen:

Definition 6.14. Sei $A \in K^{n\times n}$. Gibt es ein $B \in K^{n\times n}$ mit $AB = E_n = BA$, so ist B eindeutig durch A bestimmt; B ist die zu A inverse Matrix, schreibe $B = A^{-1}$. Hat A eine inverse Matrix, so ist A invertierbar (oder regulär).

• Gibt es ein B mit $AB = E_n$ oder $BA = E_n$, so ist $B = A^{-1}$.

Beispiel 6.15. Sei K ein Körper und sei $A \in K^{2\times 2}$, genauer

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}.$$

Setze $d = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \in K$. Ist $d \neq 0$, so rechnet man nach

$$\frac{1}{d}\begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix}\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2$$

d.h. A ist invertierbar. Sei $d = 0$. Ist $\alpha_{12} \neq 0$ oder $\alpha_{22} \neq 0$, so ist

$$\begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ 0 & 0 \end{pmatrix} A = 0.$$

Wäre A invertierbar, so liefert Rechtsmultiplikation mit A^{-1} dann

$$\begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ 0 & 0 \end{pmatrix} = 0;$$

dies ist ein Widerspruch, also ist A nicht invertierbar. In Fall $\alpha_{12} = 0 = \alpha_{23}$ folgt wegen

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

ähnlich wie vorher, dass A nicht invertierbar ist. Wir haben gezeigt:

$$A \text{ ist invertierbar } \Leftrightarrow d \neq 0.$$

Die Determinantentheorie wird uns ein ähnliches Kriterium für Matrizen in $K^{n\times n}$ liefern.

Sei $f \in End_K(V)$ und sei $A_{f,X}$ die Matrix von f bezüglich einer Basis X von V. Dann ist f ein Automorphismus genau dann, wenn $A_{f,X}$ für jede Wahl einer solchen Basis X invertierbar ist:

Lemma 6.16. *Sei V ein K-Vektorraum der Dimension $n < \infty$ und sei $f \in End_K(V)$ ein Endomorphismus. Dann sind gleichwertig:*

(a) *f ist Automorphismus,*

(b) *Für jede Basis X von V ist $A_{f,X}$ invertierbar; weiter gilt*

$$A_{f^{-1},X} = A_{f,X}^{-1},$$

(c) *Für wenigstens eine Basis X von V ist $A_{f,X}$ invertierbar.*

NB. Das Lemma impliziert die folgenden Ausagen: Sei $X = \{v_1, \ldots, v_n\}$ eine Basis von V und seien $\alpha_{ij} \in K$ $(i, j = 1, \ldots, n)$ Skalare. Setze

$$w_j = \sum_{i=1}^{n} \alpha_{ij} v_i, \ j = 1, \ldots, n.$$

1) Dann ist $\{w_1, \ldots, w_n\}$ genau dann eine Basis von V, wenn die Matrix (α_{ij}) invertierbar ist: Die Abbildung $f(v_j) = w_j$ ist ein Endomorphismus mit Basis $A_{f,X} = (\alpha_{ij})$. Dabei ist f genau dann ein Automorphismus, wenn $\{w_1, \ldots, w_n\}$ eine Basis ist; nach (b) gilt dies genau dann, wenn (α_{ij}) invertierbar ist.

2) Ist (α_{ij}) invertierbar und $(\beta_{ij}) = (\alpha_{ij})^{-1}$, so ist

$$v_j = \sum_{k=1}^{n} \beta_{kj} w_k, \ j = 1, \ldots, n$$

Folgt wegen

$$\begin{aligned} v_j &= \textstyle\sum_{i=1}^{n} \delta_{ij} v_i = \sum_{i=1}^{n} (\sum_{k=1}^{n} \alpha_{ik}\beta_{kj}) v_i = \\ &= \textstyle\sum_{k=1}^{n} \beta_{kj} \sum_{i=1}^{n} \alpha_{ik} v_i = \sum_{k=1}^{n} \beta_{kj} w_k. \end{aligned}$$

Beweis. (a)⇒(b): Sei $f^{-1} \in End_K(V)$ die zu f inverse Abbildung, sodass $f^{-1}f = ff^{-1} = \mathrm{id}_V$. Ist X eine Basis von V, so liefert der Isomorphismus $f \mapsto A_{f,X}$ die Identitäten $A_{f^{-1},X}A_{f,X} = A_{f,X}A_{f^{-1},X} = E_n$, also ist die Matrix $A_{f,X}$ invertierbar und es gilt $A_{f^{-1},X} = A_{f,X}^{-1}$.

(b)⇒(c): Trivial.

(c)⇒(a): Sei $A_{f,X}$ invertierbar bzgl. X. Setze $g = \kappa^{-1}((A_{f,X})^{-1}) \in End_K(V)$. Da κ (und so κ^{-1}) ein Isomorphismus von K-Algebren ist, erhält die Abbildung κ^{-1} Produkte und Einselemente. Also ist

$$gf = \kappa^{-1}(A_{f,X}^{-1})\kappa^{-1}(A_{f,X}) = \kappa^{-1}(A_{f,X}^{-1}A_{f,X}) = \kappa^{-1}(E_n) = \mathrm{id}_V\,.$$

Genauso folgt $fg = \mathrm{id}_V$, d.h. f ist ein Automorphismus. □

Proposition 6.17. *(Basiswechsel) (a) Seien V, W K-Vektorräume. Ferner seien $X = \{v_1, \ldots, v_n\}$, $X' = \{v_1', \ldots, v_n'\}$ Basen von V und $Y = \{w_1, \ldots, w_m\}$, $Y' = \{w_1', \ldots, w_m'\}$ Basen von W, sodass*

$$\begin{aligned} v_j' &= \textstyle\sum_{i=1}^n \beta_{ij} v_i, \quad j = 1, \ldots, n, \\ w_l' &= \textstyle\sum_{k=1}^m \gamma_{kl} w_k, \quad l = 1, \ldots m. \end{aligned}$$

Dann sind (β_{ij}) und (γ_{kl}) invertierbar und für $f \in Hom_K(V, W)$ gilt

$$A_{f,X',Y'} = (\gamma_{kl})^{-1} A_{f,X,Y}(\beta_{ij}).$$

(b) Ist $V = W$, $X = Y$ und $X' = Y'$ so liefert dies die Identität

$$A_{f,X'} = (\beta_{ij})^{-1} A_{f,X}(\beta_{ij}).$$

Beweis. (a): Seien id_V und id_W die Identitäten auf V und V. Wegen

$$\mathrm{id}_V(v_j') = v_j' = \sum_{i=1}^n \beta_{ij} v_i \text{ und } \mathrm{id}_W(w_l') = w_l' = \sum_{k=1}^m \gamma_{kl} w_k$$

ist $A_{\mathrm{id}_V,X',X} = (\beta_{ij})$ und $A_{\mathrm{id}_W,Y,Y'} = (\gamma_{kl})^{-1}$. Nach Proposition 6.11(b) ist die zu einem Kompositum assoziierte Matrix das Produkt der zu den einzelnen Abbildungen assoziierten Matrizen, also folgt

$$\begin{aligned} A_{f,X',Y'} &= A_{(\mathrm{id}_W f \mathrm{id}_V),X',Y'} = A_{\mathrm{id}_W,Y,Y'} A_{f,X,Y} A_{\mathrm{id}_V,X',X} = \\ &= (\gamma_{kl})^{-1} A_{f,X,Y}(\beta_{ij}). \end{aligned}$$

(b): Ist ein Spezialfall von (a). □

Beispiel 6.18. Für ein konkretes Beispiel eines Basiswechsels wie in Proposition 6.17(a) betrachte die lineare Abbildung

$$f : \mathbb{R}^3 \to \mathbb{R}^2, \ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Seien zunächst $X = \{(1,0,0),(0,1,0),(0,0,1)\}$ und $Y = \{(1,0),(0,1)\}$ die Standardbasen. Die Koeffizienten der Darstellungen der Bilder der Basisvektoren sind die Spalten der Matrix $A_{f,X,Y}$. Wegen

$$\begin{array}{rcl} f(1,0,0) = (1,0) & = & 1\cdot(1,0) + 0\cdot(0,1) \\ f(0,1,0) = (0,1) & = & 0\cdot(1,0) + 1\cdot(0,1) \\ f(0,0,1) = (0,0) & = & 0\cdot(1,0) + 0\cdot(0,1) \end{array}$$

ist somit die Matrix von f Byzgl. der Basen X, Y gegeben durch

$$A_{f,X,Y} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Die Wahl von anderen Basen X' von $\mathbb{R}^3$ bzw. Y' von $\mathbb{R}^2$ ergibt eine andere Matrix. Zum Beispiel, sind $X' = \{(1,0,0),(1,1,0),(1,1,1)\}$ und $Y' = \{(1,1),(0,1)\}$ diese Basen, so liefert die analoge Rechnung

$$\begin{array}{rcl} f(1,0,0) = (1,0) & = & 1\cdot(1,1) - 1\cdot(0,1) \\ f(1,1,0) = (1,1) & = & 1\cdot(1,1) + 0\cdot(0,1) \\ f(1,1,1) = (1,1) & = & 1\cdot(1,1) + 0\cdot(0,1) \end{array}$$

die Matrix

$$A_{f,X',Y'} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \end{pmatrix}$$

Nach Proposition 6.17(a) lässt sich der Übergang von $A_{f,X,Y}$ nach $A_{f,X',Y'}$ durch invertierbare Matrizen beschreiben, die einem Basiswechsel entsprechen. Betrachte zunächst die Identitätsabbildung $\mathrm{id}_{\mathbb{R}^3}$ auf $\mathbb{R}^3$ bezüglich der Basen X' und X. Wegen

$$\begin{array}{rcl} (1,0,0) & = & 1\cdot(1,0,0) + 0\cdot(0,1,0) + 0\cdot(0,0,1) \\ (1,1,0) & = & 1\cdot(1,0,0) + 1\cdot(0,1,0) + 0\cdot(0,0,1) \\ (1,1,1) & = & 1\cdot(1,0,0) + 1\cdot(0,1,0) + 1\cdot(0,0,1) \end{array}$$

ist dann

$$A_{\mathrm{id}_{\mathbb{R}^3},X',X} = (\beta_{ij}) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

wobei die Matrix (β_{ij}) invertierbar ist, da die Identität ein offensichtlicher Isomorphismus ist. Genauso hat die Identität auf $\mathbb{R}^2$ bzgl. der Basen Y' und Y aufgrund von den Identitäten

$$\begin{array}{rl} (1,1) & = 1\cdot(1,0) + 1\cdot(0,1) \\ (0,1) & = 0\cdot(1,0) + 1\cdot(0,1) \end{array}$$

die Matrix

$$A_{\mathrm{id}_{\mathbb{R}^2},Y',Y} = (\gamma_{kl}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Auch (γ_{kl}) ist invertierbar, und nach Beispiel 6.15 gilt

$$A_{\mathrm{id}_{\mathbb{R}^2},Y,Y'} = (\gamma_{kl})^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Gleichwertig lässt sich $(\gamma_{kl})^{-1}$ als die Matrix der Identitätsabbildung $\mathrm{id}_{\mathbb{R}^2}$ bezüglich der Basen Y und Y' direkt bestimmen: Aufgrund von

$$\begin{aligned} (1,0) &= 1 \cdot (1,1) - 1 \cdot (0,1) \\ (0,1) &= 0 \cdot (1,1) + 1 \cdot (0,1) \end{aligned}$$

ist

$$A_{\mathrm{id}_{\mathbb{R}^2},Y,Y'} = (\gamma_{kl})^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Nach Proposition 6.17(a) gilt dabei die Beziehung

$$A_{f,X',Y'} = (\gamma_{kl})^{-1} A_{f,X,Y}(\beta_{ij}),$$

d.h. das folgende Diagramm kommutiert

$$\begin{array}{ccc} \mathbb{R}^3 \text{ mit Basis } X' & \xrightarrow{A_{f,X',Y'}} & \mathbb{R}^2 \text{ mit Basis } Y' \\ {\scriptstyle A_{\mathrm{id},X',X}=(\beta_{ij})}\big\downarrow \cong & & \cong \big\uparrow {\scriptstyle A_{\mathrm{id},Y,Y'}=(\gamma_{kl})^{-1}} \\ \mathbb{R}^3 \text{ mit Basis } X & \xrightarrow{A_{f,X,Y}} & \mathbb{R}^2 \text{ mit Basis } Y \end{array}$$

d.h. wir haben

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

wie man durch direktes Nachrechnen bestätigt.

Definition 6.19. Sei K ein Körper und sei $A = (\alpha_{ij}) \in K^{m \times n}$ eine Matrix. Betrachte die Zeilenvektoren $z_i = (\alpha_{i1}, \ldots, \alpha_{in}), i = 1, \ldots m$ (bzw. die Spaltenvektoren $s_j = (\alpha_{1j}, \ldots, \alpha_{mj}), j = 1, \ldots, n$) von A als Elemente von K^n (bzw. K^m). Dann ist der Zeilenrang $r_z(A)$ (bzw. Spaltenrang $r_s(A)$) die Anzahl der linear unabhängigen Zeilenvektoren (bzw. Spaltenvektoren), d.h.

$$r_z(A) = \dim_K \langle z_1, \ldots, z_m \rangle \text{ und } r_s(A) = \dim_K \langle s_1, \ldots, s_n \rangle.$$

• Offensichtlich ist $r_z(A) \leq min\{m, n\}$ und $r_s(A) \leq min\{m, n\}$.

Sei $f \in Hom_K(V, W)$ eine lineare Abbildung und sei $A_{f,X,Y} \in K^{m \times n}$ die Matrix von f bzgl. Basen X von V und Y von W. Dann ist der Rang der linearen Abbildung f der Spaltenrang der Matrix $A_{f,X,Y}$:

Lemma 6.20. *(a) Seien V, W K-Vektorräume, X eine Basis von V und Y eine Basis von W. Ist $f : V \to W$ eine K-lineare Abbildung, so gilt $r(f) = r_s(A_{f,X,Y})$*
(b) Sei $A \in K^{m\times n}$ und seien $B \in K^{n\times n}$ und $C \in K^{m\times m}$ invertierbar. Dann ist $r_s(A) = r_s(CAB)$.
(c) Sei $A \in K^{m\times n}$ und $B \in K^{n\times r}$. Dann ist $r_s(AB) \leq min\{r_s(A), r_s(B)\}$.

Beweis. (a): Sei $X = \{v_1, \ldots, v_n\}$ und $Y = \{w_1, \ldots, w_m\}$. Nach Definition von $A_{f,X,Y} = (\alpha_{ij})$ ist $f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i,\ j = 1, \ldots n$.
Sei $\{e_1, \ldots, e_n\}$ die Standardbasis von K^m. Die Zuordnung $w_i \mapsto e_i$ bestimmt eine eindeutige K-lineare Abbildung $g : W \to K^m$, genauer

$$g : W \to K^m, \sum_{i=1}^{m} \beta_i w_i \mapsto \begin{pmatrix} \beta_1 \\ \cdot \\ \cdot \\ \beta_m \end{pmatrix};$$

da g eine Basis auf eine Basis abbildet ist diese Abbildung ein Isomorphismus. Nach Übungblatt 7, Aufgabe 4 gilt $r(f) = r(gf) = \dim \mathrm{im}(gf)$, wobei

$$(gf)(v_j) = g(\sum_{i=1}^{m} \alpha_{ij} w_i) = \begin{pmatrix} \alpha_{1j} \\ \cdot \\ \cdot \\ \alpha_{mj} \end{pmatrix} = s_j,\ j = 1, \ldots n,$$

d.h. das Bild des j-ten Basisvektors unter gf ist die j-te Spalte von $A_{f,X,Y}$. Also ist $r(f) = \dim_K \langle s_1, \ldots s_n \rangle = r_s(A_{f,X,Y})$.
(b), (c): Siehe Übungsblatt 7, Aufgabe 4. □

Definition 6.21. Sei K ein Körper. Für eine Matrix $A \in K^{m\times n}$

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdot & \cdot & \cdot & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdot & \cdot & \cdot & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{m1} & \alpha_{m2} & \cdot & \cdot & \cdot & \alpha_{mn} \end{pmatrix}$$

definiere die zu A transponierte Matrix $A^t \in K^{n\times m}$ durch

$$A^t = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdot & \cdot & \cdot & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \cdot & \cdot & \cdot & \alpha_{m2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{1n} & \alpha_{2n} & \cdot & \cdot & \cdot & \alpha_{nm} \end{pmatrix},$$

d.h. A^t entsteht aus A durch Vertauschen der Zeilen und Spalten.

• Die Abbildung $K^{m\times n} \to K^{n\times m},\ A \mapsto A^t$ ist ein Isomorphismus.

• Für $A \in K^{m\times n}$ und $B \in K^{n\times r}$ gilt: $(AB)^t = B^t A^t$.

• Wir zeigen später: Jede lineare Abbildung $f \in Hom_K(V,W)$ induziert ein lineare Abbildung (die duale Abbildung) $f^* \in Hom_K(W^*, V^*)$, wobei $W^* = Hom_K(W,K)$ und $V^* = Hom_K(V,K)$ die entsprechenden Dualräume sind. Ist $A = A_{f,X,Y}$ die Matrix von f bgzl. der Basen X, Y, so ist $A^t = A_{f^*,Y^*,X^*}$ die Matrix von f^* bzgl. der dualen Basen X^*, Y^*.

Theorem 6.22. *Für jede Matrix $A \in K^{m\times n}$ gilt*

$$r_z(A) = r_s(A) \leq min\{m,n\}\ \textit{(d.h. Zeilenrang=Spaltenrang).}$$

• Schreibe $r(A) = r_z(A) = r_s(A)$; $r(A)$ ist der Rang der Matrix A.

Beweis. Sei $r = r_s(A)$. Nach Übungsblatt 7, Aufgabe 5 gibt es invertierbare Matrizen $B \in K^{m\times m}$ und $C \in K^{n\times n}$, sodass gilt

$$BAC = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n},$$

wobei E_r die Einheitsmatrix vom Typ r ist. Transponieren liefert

$$C^t A^t B^t = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times m}.$$

Wegen $E_m = (BB^{-1})^t = (B^{-1})^t B^t$ ist B^t invertierbar; ebenso ist C^t invertierbar. Mit Lemma 6.20(b) folgt aus den obigen Identitäten

$$r_s(A) = r_s(BAC) = r_s(C^t A^t B^t) = r_s(A^t) = r_z(A).$$

□

Sei $A \in K^{n\times n}$ und $V = K^n$ mit der Standardbasis $X = \{e_1, \ldots, e_n\}$. Nach Proposition 6.11 ist die Abbildung

$$\kappa : \mathrm{End}_K(V) \to K^{n\times n},\ f \mapsto A_{f,X}$$

ein Isomorphismus von K-Vektorräumen, d.h. es gibt eine eindeutige Abbildung $f \in \mathrm{End}_K(V)$ mit $A = A_{f,X}$. Da κ mit der Verknüpfung von Abbildungen und der Multiplikation von Matrizen verträglich ist (siehe Bemerkungen von Seite 44) ist κ ein Isomorphismus von K-Algebren. Also ist $\mathrm{id}_V = gf$ genau dann, wenn $E_n = A_{g,X} \cdot A_{f,X}$ und somit

$$f \text{ ist Isomorphismus } \Leftrightarrow A = A_{f,X} \text{ ist invertierbar .}$$

Für $f \in \mathrm{End}_K(V)$ ist $r(f) \leq \dim_K V = n$ und f ist ein Epimorphismus genau dann, wenn $r(f) = n$ ist. Nach Lemma 6.20(a) ist $r(f) = r(A) = r(A_{f,X})$, also gilt

$$f \text{ ist Epimorphismus } \Leftrightarrow r(A) = r(A_{f,X}) = n.$$

Da nach Lemma 5.11 (mit $V = W$) f genau dann ein Epimorphismus ist, wenn f ein Isomorphismus ist, ergibt sich aus den obigen Äquivalenzen folgendes Resultat:

Proposition 6.23. *Für $A \in K^{n\times n}$ sind gleichwertig:*

(a) $r(A) = n$

(b) *A ist invertierbar.*

7. Elementare Umformungen

Für jede Matrix $A \in K^{m\times n}$ ist nach Theorem 6.22 der Zeilenrang gleich dem Spaltenrang, d.h. $r_z(A) = r_s(A)$. In einfachen Beispielen lässt sich der Rang einer Matrix sofort ablesen. Zum Beispiel ist

$$r\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix} = 2 \text{ und } r\begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix} = 1$$

Eine allgemeine Matrix lässt sich durch 'elementare Umformungen' wie Addition und Subtraktion von Vielfachen von Zeilen auf eine solche einfachere Form bringen. Zum Beispiel,

$$A = \begin{pmatrix}2 & 0 & 1\\ 1 & 2 & 0\\ 1 & 1 & 1\end{pmatrix} \overset{z_3-z_2}{\to} \begin{pmatrix}2 & 0 & 1\\ 1 & 2 & 0\\ 0 & -1 & 1\end{pmatrix} \overset{z_2-1/2z_1}{\to} \begin{pmatrix}2 & 0 & 1\\ 0 & 2 & -1/2\\ 0 & -1 & 1\end{pmatrix}$$
$$\overset{z_3+1/2z_2}{\to} \begin{pmatrix}2 & 0 & 1\\ 0 & 2 & -1/2\\ 0 & 0 & -3/4\end{pmatrix} = A'$$

Diese Umformungen lassen sich als Linksmultiplikation mit Matrizen beschreiben:

$$z_3-z_2 \leftrightarrow T_1 = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{pmatrix} : \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix}2 & 0 & 1\\ 1 & 2 & 0\\ 1 & 1 & 1\end{pmatrix} = \begin{pmatrix}2 & 0 & 1\\ 1 & 2 & 0\\ 0 & -1 & 1\end{pmatrix}$$

und genauso

$$z_2-1/2z_1 \leftrightarrow T_2 = \begin{pmatrix}1 & 0 & 0\\ -1/2 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \text{ und } z_3+1/2z_1 \leftrightarrow T_3 = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1/2 & 0 & 1\end{pmatrix}$$

Also führt Linksmultiplikation mit diesen 'Transformationsmatrizen' T_1, T_2, T_3 die gegebene Matrix A in die einfachere Form A' über, d.h.

$$T_3T_2T_1A = A',$$

wobei die T_i invertierbar sind (sie haben offensichtlich Rang 3). Nach Lemma 6.20(b) verändert die Multiplikation mit einer invertierbaren

Matrix den Rang nicht, somit haben A und A' den gleichen Rang

$$r(A) = r(T_3T_2T_1A) = r(A') = 3.$$

Wir bestimmen allgemein 'Elementarmatrizen' mit der Eigenschaft, dass Multiplikation mit diesen Matrizen den Rang einer Matrix erhält und verallgemeinern das obige Beispiel zu einen Algorithmus, der es uns erlaubt den Rang einer Matrix zu berechnen.

Definition 7.1. Sei K ein Körper und $i \neq j$. Eine Elementarmatrix $T_{ij}(\alpha) = (\alpha_{ij}) \in K^{n\times n}$ hat die Form:

(a) $\alpha \in K$ an der Stelle (i, j) (wobei $i \neq j$ ist),
(b) 1 auf der Diagonalen (d.h. $\alpha_{ii} = 1$ für alle i),
(c) 0 an allen anderen Stellen.

Für $\alpha = 1$ setze $T_{ij} = T_{ij}(1)$. Sei $\mathbb{E}_n \subseteq K^{n\times n}$ die Menge der Elementarmatrizen vom Typ (n, n).

• Für Elementarmatrizen gelten die Beziehungen

$$\begin{aligned} T_{ij}(\alpha) \cdot T_{ij}(\beta) &= T_{ij}(\alpha + \beta), \\ T_{ij}(\alpha)T_{ij}(-\alpha) &= E_n, \end{aligned}$$

insbesondere ist jede Elementarmatrix invertierbar, d.h. regulär.

Lemma 7.2. *Sei $A \in K^{m\times n}$ mit Zeilen $z_1, .., z_m$ und Spalten $s_1, .., s_n$.*

(a) *Ist $T_{ij}(\alpha) \in K^{m\times m}$ (d.h. das Produkt $T_{ij}(\alpha)A$ existiert), so ist*

$$T_{ij}(\alpha)A = \begin{pmatrix} z_1 \\ \cdot \\ \cdot \\ z_i + \alpha z_j \\ \cdot \\ \cdot \\ z_m \end{pmatrix}.$$

(b) *Ist $T_{ij}(\beta) \in K^{n\times n}$ (d.h. das Produkt $AT_{ij}(\beta)$ existiert), so gilt*

$$AT_{ij}(\beta) = (s_1, \cdots, s_j + \beta s_i, \ldots, s_n).$$

Beweis. Nachrechnen. □

Definition 7.3. Sei $A \in K^{m\times n}$. Eine elementare Umformung von A ist eine Umformung, die durch die Multiplikation von A mit einer Elementarmatrix (von links oder von rechts) ensteht. Explizit hat jede elementare Umformung die Form

(a) Ersetzen der Zeile z_i von A durch $z_i + \alpha z_j$, wobei $\alpha \in K$ und $i \neq j$; die Zeilen z_k für $k \neq i$ bleiben unverändert. Gleichwertig: Multiplikation mit $T_{ij}(\alpha) \in \mathbb{E}_m$ von links.

(b) Ersetzen der Spalte s_j von A durch $s_j + \beta s_i$, wobei $\beta \in K$ und $i \neq j$; die Spalten s_k für $k \neq i$ bleiben unverändert. Gleichwertig: Multiplikation mit $T_{ij}(\beta) \in \mathbb{E}_n$ von rechts.

NB. Ist $A \in K^{m\times n}$, so entspricht die Anwendung endlich vieler elementarer Umformungen $T_i \in \mathbb{E}_m$ und $S_j \in \mathbb{E}_n$ auf A einer Umformung

$$A \quad \mapsto \quad T_k \cdots T_1 A S_1 \cdot S_l = A'.$$

Setze $C = T_k \cdots T_1$ und $B = S_1 \cdots S_l$. Dann sind C und B als Produkt von invertierbaren Matrizen invertierbar und nach Lemma 6.20(b) gilt

$$r(A) = r(CAB) = r(A').$$

Also verändert die Anwendung von (endlich viele) beliebigen elementare Umformungen auf A nicht den Rang von A; man kann somit eine gegebene Matrix A durch Links- bzw. Rechtsmutiplikation mit Elementarmatrizen in eine Matrix überführen, deren Rang leicht zu bestimmen ist. Dabei kann man entweder diese Operationen ad hoc auf eine gegebene Matrix anwenden, oder einen von diversen Algorithmen verwenden, die die gegebene Matrix in eine einfachere Form bringen.

Wir geben ein Beispiel eines solchen Algorithmus:

Theorem 7.4. *Sei $A = (\alpha_{ij}) \in K^{m\times n}$. Ist $A \neq 0$, so gibt es Elementarmatrizen $T_i \in \mathbb{E}_m$ $(i = 1, \ldots, k)$ und $S_j \in \mathbb{E}_n$ $(j = 1, \ldots, l)$, sodass die resultierende Matrix $A' = T_k \cdots T_1 A S_1 \cdots S_l$ die folgende Form hat*

$$A' = \begin{pmatrix} \alpha'_{11} & * & * & & & & \cdot & \cdot & * \\ 0 & \alpha'_{22} & * & & & & \cdot & \cdot & * \\ 0 & 0 & \alpha'_{33} & * & & & \cdot & \cdot & * \\ \cdot & & \cdot & & & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \alpha'_{rr} & * & \cdot & \cdot & * \\ 0 & 0 & \cdot & 0 & 0 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & & & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & & & & \cdot & \cdot & 0 \end{pmatrix},$$

wobei $\alpha'_{11} \cdots \alpha'_{rr} \neq 0$ ist. Es ist $r(A) = r(A') = r$.

Beweis. Schritt 1: Ist $\alpha_{11} \neq 0$, so gehe zu Schritt 4.

Schritt 2: Sei $\alpha_{11} = 0$, aber sei z_1 nicht die Nullzeile. Dann gibt es in z_1 einen nicht-trivialen Eintrag α_{1j} mit $j > 1$. Rechtsmultipikation mit der Elementarmatrix $T_{1j} \in \mathbb{E}_n$ ersetzt die erste Spalte s_1 durch $s_1 + s_j$, d.h. die Matrix AT_{1j} hat an der Stelle $(1, 1)$ den Eintrag $\alpha_{j1} \neq 0$. Gehe jetzt zu Schritt 4.

Schritt 3: Sei $\alpha_{11} = 0$ und sei z_1 eine Nullzeile. Da $A \neq 0$ ist, gibt es eine Zeile $z_i,\ i > 1$ mit einem nicht-trivalen Eintrag. Linksmultiplikation mit der Elementarmatrix $T_{1i} \in \mathbb{E}_m$ ersetzt z_1 durch $z_1 + z_i$; also hat die erste Zeile von $T_{1i}A$ die Form $(\beta_{i1}, \ldots, \beta_{in})$ mit $\beta_{ij} \neq 0$ für ein j. Ist $\beta_{i1} \neq 0$, so gehe zu Schritt 4, ist $\beta_{i1} = 0$, so gehe zu Schritt 2.

Schritt 4: Wir haben eine Matrix A mit $\alpha_{11} \neq 0$. Also sind die Quotienten $\frac{\alpha_{i1}}{\alpha_{11}} \in K$ und damit die Matrizen $T_{i1}(\frac{\alpha_{i1}}{\alpha_{11}}) \in \mathbb{E}_m$ für $i = 2, \ldots, m$ definiert. Die Matrix $T_{m1}(-\frac{\alpha_{m1}}{\alpha_{11}}) \cdots T_{21}(-\frac{\alpha_{21}}{\alpha_{11}})A$ hat die Form

$$\begin{pmatrix} z_1 \\ z_1 - \frac{\alpha_{21}}{\alpha_{11}} z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_m - \frac{\alpha_{m1}}{\alpha_{11}} z_1 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdot & \cdot & \alpha_{1n} \\ 0 & & & & \\ \cdot & & & & \\ \cdot & & A_1 & & \\ \cdot & & & & \\ 0 & & & & \end{pmatrix}$$

Schritt 5: Ist die verbleibende $(m-1) \times (n-1)$-Matrix A_1 die Nullmatrix, so sind wir fertig. Falls $A_1 \neq 0$ ist, wende die Schritte 1-4 auf A' an (jede elementare Umformung von A_1 lässt sich als elementare Umformung von A interpretieren). Nach endlich vielen Schritten liefert dies eine Matrix der gewünschten Form. □

Beispiele 7.5. (a) Sei $A \in \mathbb{R}^{3\times 3}$ die Matrix

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

Dem Algorithmus folgend beginnen wir mit Schritt 4. Dies liefert

$$\begin{pmatrix} 1 & -1 & 2 \\ 4 & 4 & -2 \\ 2 & 0 & 2 \end{pmatrix} \overset{T_{21}(-4)}{\to} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 8 & -10 \\ 2 & 0 & 2 \end{pmatrix} \overset{T_{31}(-2)}{\to} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 8 & -10 \\ 0 & 2 & -2 \end{pmatrix}$$

Iteration dieses Verfahrens auf die verbleibende Matrix $A_1 \in \mathbb{R}^{2\times 2}$ zeigt

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 8 & 10 \\ 0 & 2 & -2 \end{pmatrix} \overset{T_{32}(-1/4)}{\to} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 8 & 10 \\ 0 & 0 & 1/2 \end{pmatrix}$$

Also ist $r = 3$ und $r(A) = 3$.

(b) Betrachte die Matrix (mit reellen Koeffizienten)

$$A = \begin{pmatrix} 3 & 6 & 2 & 10 \\ 10 & 16 & 6 & 30 \\ 5 & 14 & 4 & 14 \end{pmatrix}$$

Elementare Umformungen (ohne Verwendung des Algorithmus) zeigen

$$\begin{pmatrix} 3 & 6 & 2 & 10 \\ 10 & 16 & 6 & 30 \\ 5 & 14 & 4 & 14 \end{pmatrix} \overset{z_2-3z_1}{\to} \begin{pmatrix} 3 & 6 & 2 & 10 \\ 1 & -2 & 0 & 0 \\ 5 & 14 & 4 & 14 \end{pmatrix}$$
$$\overset{z_3-2z_1}{\to} \begin{pmatrix} 3 & 6 & 2 & 10 \\ 1 & -2 & 0 & 0 \\ -1 & 2 & 0 & -6 \end{pmatrix}$$
$$\overset{z_3+z_2}{\to} \begin{pmatrix} 3 & 6 & 2 & 10 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & -6 \end{pmatrix}$$

Die Zeilen z_1, z_2, z_3 in der letzten Matrix sind offensichtlich linear unabhängig, also ist $r(A) = 3$.

Ist $A \in K^{n\times n}$ und $A = T_1 \cdots T_l$ das Produkt von Elementarmatrizen $E_i \in \mathbb{E}_n$, so ist A invertierbar. Aber nicht jede invertierbare Matrix ist das endliche Produkt von Elementarmatrizen. Zum Beispiel ist

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

invertierbar, aber nicht Produkt von Elementarmatrizen: Da $d = d(A) = 2 \neq 0$ ist A invertierbar, siehe Beispiel 6.15. Jede Elementarmatrix vom Typ $(2, 2)$ ist eine Matrix der folgenden Form

$$T = \begin{pmatrix} 1 & 0 \\ \gamma & 1 \end{pmatrix} \text{ oder } S = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix},$$

also ist $d(T) = 1 = d(S)$. Für $A, B \in \mathbb{R}^{2\times 2}$ gilt $d(AB) = d(A)d(B)$; also hat jedes endliche Produkt von Elementen aus $\mathbb{E}_2$ Determinante 1 und A kann keine Darstellung als ein solches Produkt haben.

Wir werden zeigen: Allgemein sind invertierbare Diagonalmatrizen $D = (d_{ij})$ mit $d_{11} \cdots d_{nn} \neq 1$ nicht als ein Produkt von Elementarmatrizen darstellbar.

Definition 7.6. Eine elementare Diagonalmatrix $D_i(\alpha) \in K^{n\times n}$ ist eine Matrix $D_i(\alpha) = (\alpha_{ij})$ mit den Einträgen

(a) $\alpha_{ii} = \alpha$ für ein $\alpha \in K \setminus \{0\}$,
(b) $\alpha_{jj} = 1$ für $j \neq i$,
(c) $\alpha_{ij} = 0$ für $i \neq j$.

Sei $\mathbb{D}_n \subseteq K^{n\times n}$ die Menge der Matrizen von diesem Typ.

• $D_i(\alpha) \cdot D_i(\alpha^{-1}) = E_n$, d.h. $D_i(\alpha)$ ist invertierbar.

Lemma 7.7. *Jede invertierbare Matrix $A \in K^{n\times n}$ hat eine Darstellung als ein endliches Produkt von Matrizen aus $\mathbb{E}_n^\times = \mathbb{E}_n \cup \mathbb{D}_n$.*

• Sei $GL_n(K)$ die Gruppe der invertierbaren Matrizen in $K^{n\times n}$ (bzgl. der Multiplikation von Matrizen). Das obige Lemma besagt, dass sich jedes Element von $GL_n(K)$ als ein endliches Produkt von Elementen von $\mathbb{E}_n^\times$ schreiben lässt; man sagt $GL_n(K)$ wird von $\mathbb{E}_n^\times$ erzeugt.

• Die Darstellung eines Elements von $GL_n(K)$ als ein endliches Produkt von Elementen aus $\mathbb{E}_n^\times$ ist nicht eindeutig. Zum Beispiel gilt in $GL_2(\mathbb{R})$:

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

Beweis. (Skizze) Sei $A \in K^{n\times n}$ invertierbar. Nach Proposition 6.23 ist $r(A) = n$, insbesondere hat A keine triviale Spalte. Ist $\alpha_{11} = 0$, so gibt es ein $i > 1$ mit $\alpha_{i1} \neq 0$ und $T_{1i}A$ hat α_{i1} an der Stelle $(1,1)$; wir können also annehmen, dass $\alpha_{11} \neq 0$ ist. Wie im Schritt 4 im Beweis von Theorem 7.4 lässt sich A durch Zeilenumformungen und Iteration (d.h. Anwendung auf A_1) in eine Matrix $B = T_k \cdots T_1 A$ der Form

$$B = \begin{pmatrix} \beta_{11} & * & * & * & * \\ 0 & \beta_{22} & * & * & * \\ 0 & 0 & \beta_{33} & * & * \\ \cdot & & \cdot & & \\ \cdot & & \cdot & \cdot & * \\ 0 & 0 & 0 & 0 & \beta_{nn} \end{pmatrix}$$

mit $\beta_{11}\cdots\beta_{nn} \neq 0$ überführen. Diese Matrix B lässt sich durch Linksmultiplikation mit Elementarmatrizen und Matrizen der Form $D_{ii}(\alpha)$ zu einer Einheitsmatrix machen: Man beseitigt die Einträge in der letzten Spalte $\beta_{1,n}, \beta_{2,n}, \cdots, \beta_{n-1,n}$ mit Hilfe der letzten Zeile, dann die Einträge in der vorletzten Spalte $\beta_{1,n-1}, \cdots, \beta_{n-2,n-1}$ mit Hilfe der vorletzten Zeile, etc. um so eine Diagonalmatrix mit nicht-trivalen Diagonaleinträgen zu erhalten. Multiplikation mit geeigneten Matrizen der Form $D_{ii}(\alpha)$ macht dann diese Diagonalmatrix zur Einheitsmatrix.

Also gibt es Elementarmatrizen $T_i \in \mathbb{E}_n$ $(i = 1, \ldots, r)$ sowie elementare Diagonalmatrizen $D_j \in \mathbb{D}_n$ $(j = 1, \ldots, l)$ mit $D_s \cdots D_1 T_r \cdots T_1 A = E_n$; es folgt $A = (D_1 \cdots D_1 T_r \cdots T_1)^{-1} = T_1^{-1} \cdots D_s^{-1}$. □

Beispiele 7.8. Der Beweis von Lemma 7.7 ist konstruktiv und liefert ein Verfahren die inverse Matrix A^{-1} zu berechnen: Ist $A \in K^{n\times n}$ invertierbar und sind $D_i \in \mathbb{D}_n$ bzw. $T_j \in \mathbb{E}_n$ mit $D_s \cdots D_1 T_r \ldots T_1 A = E_n$, so ist $A^{-1} = D_s \cdots D_1 T_r \cdots T_1 = D_s \cdots D_1 T_r \cdots T_1 E_n$. Um A^{-1} konkret zu berechnen schreibt man A und E_n nebeneinander, und formt A durch Linksmultiplikation mit Elementen aus $\mathbb{E}_n$ bzw. $\mathbb{D}_n$ zur Einheitsmatrix E_n um. Die analogen Umformungen angewandt auf E_n liefern die inverse Matrix A^{-1}.

(a) Wir betrachten $A \in \mathbb{R}^{3\times 3}$ und wenden das obige Verfahren zur Berechnung von A^{-1} an *ohne* vorher bestimmt zu haben, ob A^{-1} existiert:

$$
\begin{array}{rcl}
A & = & \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & -1 \\ 0 & 1 & -4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad = \quad E_3 \\
\overset{z_2-z_1}{\rightarrow} & & \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \\ 0 & 1 & -4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\overset{z_3-z_2}{\rightarrow} & & \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \\
\overset{z_1-z_2}{\rightarrow} & & \begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & -3 \\ 0 & 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \\
\overset{z_2-3z_3}{\rightarrow} & & \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 2 & -1 & 0 \\ -4 & 4 & -3 \\ 1 & -1 & 1 \end{pmatrix} \\
\overset{z_1+5z_3}{\rightarrow} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 7 & -6 & 5 \\ -4 & 4 & -3 \\ 1 & -1 & 1 \end{pmatrix} \\
\overset{D_3(-1)}{\rightarrow} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 7 & -6 & 5 \\ -4 & 4 & -3 \\ -1 & 1 & -1 \end{pmatrix} \quad = \quad A^{-1}
\end{array}
$$

Zur Kontrolle rechnet man nach:

$$
A^{-1} \cdot A = \begin{pmatrix} 7 & -6 & 5 \\ -4 & 4 & -3 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & -1 \\ 0 & 1 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = E_3
$$

(b) Wendet man dieses Verfahren auf eine Matrix A an, die nicht invertierbar ist, so zeigt sich das unterwegs:

$$
\begin{array}{rcl}
A & = & \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad = \quad E_3 \\
\overset{z_3-z_1}{\rightarrow} & & \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}
\end{array}
$$

Der Rang der letzten Matrix auf der A-Seite ist 2 und gleich dem Rang von A, d.h. $r(A) = 2 < 3$ und die Matrix A ist nicht invertierbar.

Sei $A \in K^{m\times n}$ mit $r(A) = r$. Nach Übungsblatt 7, Aufgabe 5 gibt es invertierbare Matrizen $C \in K^{m\times m}$ und $B \in K^{n\times n}$, sodass gilt

$$CAB = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in K^{m\times n}$$

Wir zeigen, wie sich mit Hilfe von Elementarmatrizen diese Matrizen C und B explizit berechnen lassen.

Wir benötigen dazu die folgende Definition.

Definition 7.9. Eine Matrix $A = (\alpha_{ij}) \in K^{m\times n}$ hat Zeilenstufenform, falls gilt:

(a) Es gibt ein r mit $0 \leq r \leq m$, sodass in den Zeilen mit Index 1 bis r jeweils nicht nur Nullen stehen, und in den Zeilen mit Index $r+1$ bis m nur Nullen stehen,

(b) Für jedes i mit $1 \leq i \leq r$ sei j_i der kleinste Index der Spalte, in der ein Eintrag ungleich Null steht, d.h. $j_i = \min\{j \mid \alpha_{ij} \neq 0\}$; hier ist $j_1 < j_2 < \cdots < j_r$.

Dabei ist $r = 0$ zulässig, in diesem Fall sind alle Einträge von A Null. Die ersten nichttrivialen Einträge in den nichttrivialen Zeilen $\alpha_{1j_1}, \ldots, \alpha_{rj_r}$ sind die Pivots (oder auch Angelpunkte) von A.

Beispiele 7.10. (a) Die Matrix

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

ist in Zeilenstufenform. Hier $m = 3, n = 4, r = 3$; weiter ist $j_1 = 2, j_2 = 3, j_3 = 4$, die Pivots sind die Einträge $\alpha_{12} = 2, \alpha_{23} = 1$ und $\alpha_{34} = 2$.
(b) Die folgende Matrix ist in Zeilenstufenform

$$A = \begin{pmatrix} 1 & 3 & 0 & 5 & 6 & 0 & 5 \\ 0 & 0 & 1 & 3 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Das Rechenverfahren zur Bestimmung von C und B basiert auf folgende Beobachtung: Jede Matrix $A \in K^{m\times n}$ lässt sich durch geeigenete Zeilenumformungen (d.h. Linksmultiplikation mit Elementarmatrizen) in Zeilenstufenform überführen. Also gibt es $T_1, \ldots T_k \in \mathbb{E}_m$, sodass

$$T_k \cdots T_1 A$$

in Zeilenstufenform ist. Weiter lässt sich jede Matrix in Zeilenstufenform (mit entsprechendem r) durch Spaltenumformungen (d.h. Rechtsmultiplikation mit Elementarmatrizen) in eine Matrix der Form

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

überführen, wobei C_r eine Diagonalmatrix vom Typ (r, r) mit nichttrivialen Einträgen ist. Durch Linksmultiplikation mit geeigneten elementaren Diagonalmatrizen $D_s \cdots D_1$ ergibt sich die Matrix

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

Also ist

$$D_s \cdots D_1 T_k \cdots T_1 A S_1 \ldots S_l = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

und so $C = D_s \cdots D_1 T_k \cdots T_1$ und $B = S_1 \cdots S_l$.

Beispiel 7.11. Betrachte die folgende Matrix mit reellen Einträgen

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

Offensichtlich hat A den Rang 2, d.h. es gibt invertierbare Matrizen C und B mit

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

Umd B und C zu bestimmen, schreiben wir im ersten Schritt E_2 und A nebeneinander und formen diese Matrizen durch Zeilenumformungen parallel so um, dass A in Zeilenstufenform übergeführt wird:

$$\begin{array}{rcll} E_2 & = & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \end{pmatrix} = A \\ & \overset{z_2 - 2z_1}{\to} & \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{pmatrix} = T_1 A \end{array}$$

Im zweiten Schritt schreiben wir T_1A und E_3 nebeneinander und bringen T_1A durch Spaltenumformungen auf die gewünschte Form:

$$\begin{array}{rcll} T_1A & = & \begin{pmatrix}1&2&0\\0&-2&1\end{pmatrix} & \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = E_3 \\ & \overset{s_2-2s_1}{\to} & \begin{pmatrix}1&0&0\\0&-2&1\end{pmatrix} & \begin{pmatrix}1&-2&0\\0&1&0\\0&0&1\end{pmatrix} = E_3S_1 \\ & \overset{s_2+3s_3}{\to} & \begin{pmatrix}1&0&0\\0&1&1\end{pmatrix} & \begin{pmatrix}1&-2&0\\0&1&0\\0&3&1\end{pmatrix} = E_3S_1S_2 \\ & \overset{s_3-s_2}{\to} & \begin{pmatrix}1&0&0\\0&1&0\end{pmatrix} & \begin{pmatrix}1&-2&2\\0&1&-1\\0&3&-2\end{pmatrix} = E_3S_1S_2S_3 \end{array}$$

Also ist $C = T_1 = \begin{pmatrix}1&0\\-2&1\end{pmatrix}$ und $B = S_1S_2S_3 = \begin{pmatrix}1&-2&2\\0&1&-1\\0&3&-2\end{pmatrix}$.

8. Lineare Gleichungssysteme

Wir betrachten Systeme von m linearen Gleichungen in n Variablen mit Koeffizienten in einem Körper K, d.h. ein System von Gleichungen

$$\begin{array}{ccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & \cdot & \cdot & \alpha_{1n}x_n & = & \beta_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & \cdot & \cdot & \alpha_{2n}x_n & = & \beta_2 \\ \cdot & & \cdot & & & & & \cdot \\ \cdot & & \cdot & & & & & \cdot \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & \cdot & \cdot & \alpha_{mn}x_n & = & \beta_m, \end{array}$$

mit Variablen x_j und Skalaren $\alpha_{ij}, \beta_i \in K$; dafür schreiben wir auch

$$\text{(L)} \quad \sum_{j=1}^{n} \alpha_{ij}x_j = \beta_i, \ i = 1, \ldots, m.$$

Eine Lösung von (L) ist eine gemeinsame Lösung der m Gleichungen.

Das System (L) lässt sich als ein Matrizenprodukt auffassen

$$Ax = \begin{pmatrix}\alpha_{11}&\cdot&\cdot&\cdot&\alpha_{1n}\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \alpha_{m1}&\cdot&\cdot&\cdot&\alpha_{mn}\end{pmatrix} \cdot \begin{pmatrix}x_1\\x_2\\\cdot\\\cdot\\x_n\end{pmatrix} = \begin{pmatrix}\beta_1\\\beta_2\\\cdot\\\cdot\\\beta_m\end{pmatrix} = b$$

Der Matrix $A = (\alpha_{ij})$ definiert eine lineare Abbildung $A : K^n \to K^m$ durch $x \mapsto Ax$. Damit lässt sich das System (L) wie folgt interpretieren: Ist $b \in K^m$ fest gewählt, so gilt $A^{-1}(b) = \{x \in K^n \mid Ax = b\}$, d.h. die Elemente von $A^{-1}(b)$ sind genau die Lösungen von (L).

Die wesentlichen Fragestellungen nach der Existenz und Eindeutigkeit von Lösungen von (L) lassen sich somit wie folgt formulieren:

(1) Ist $A^{-1}(b) \neq \emptyset$, d.h. gibt es eine Lösung?

(2) Ist $|A^{-1}(b)| = 1$, d.h. gibt es eine eindeutige Lösung?

Das System (L) lässt sich wie folgt umschreiben

$$\begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \cdot \\ \cdot \\ \alpha_{m1} \end{pmatrix} \cdot x_1 + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \cdot \\ \cdot \\ \alpha_{m2} \end{pmatrix} \cdot x_2 + \cdots + \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \cdot \\ \cdot \\ \alpha_{mn} \end{pmatrix} \cdot x_n = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_m \end{pmatrix}$$

Dabei sind die auf der linken Seite auftretenden Vektoren genau die Spalten $s_1, \ldots, s_n$ der Matrix A, und falls eine Lösung x exisitiert, dann muss b eine Linearkombination der s_i sein und es gilt

$$x_1 s_1 + \cdots x_n s_n = b.$$

Die Fragen nach Existenz und Eindeutgkeit von Lösungen lassen sich somit auch wie folgt ausdrücken:

(1') Ist $b \in \langle s_1, \ldots, s_n \rangle$?

(2') Falls $b \in \langle s_1, \ldots, s_n \rangle$, sind $s_1, \ldots, s_n$ linear unabhängig?

Wir formalisieren diese Überlegungen und geben zunächst einfach zu berechnende Existenz- und Eindeutigkeitskriterien für die Lösbarkeit eines solchen Systems von linearen Gleichungen.

Definition 8.1. Sei $A = (\alpha_{ij}) \in K^{m \times n}$ eine Matrix und $b = (\beta_i) \in K^m$ ein Spaltenvektor. Ein lineares Gleichungssystem (L) ist ein System von Gleichungen der Form $Ax = b,\ x = (x_j) \in K^n$, d.h. die m Gleichungen

$$(\mathrm{L}): \quad \sum_{j=1}^{n} \alpha_{ij} x_j = \beta_i,\ i = 1, \ldots, m.$$

Ist $b = 0$, so ist (L) homogen; ist $b \neq 0$, so ist (L) inhomogen. Die erweiterte Koeffizientenmatrix des linearen Gleichungssystems (L) ist

$$B = [A, b] = \begin{pmatrix} \alpha_{11} & \cdot & \cdot & \cdot & \alpha_{1n} & \beta_1 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \alpha_{m1} & \cdot & & & \alpha_{mn} & \beta_n \end{pmatrix}$$

Lemma 8.2. *Sei $A \in K^{m\times n}$ eine Matrix.*

(a) *Die Lösungen des homogenen Systems $Ax = 0$ ist genau* $\ker(A)$*; insbesondere bilden die Lösungen von $Ax = 0$ einen linearen Unterraum von K^n der Dimension $n - r(A)$.*

(b) *Ist x_0 eine Lösung von $Ax = b$, so ist $x_0 + \ker(A)$ die Menge aller Lösungen von $Ax = b$.*

• Nach (b) bilden die Lösungen von $Ax = b$ einen affinen Unterraum von K^n; nach (a) hat dieser affine Unterraum Dimension $n - r(A)$.

• Aus (a) folgt: Ist $n > m$, so hat $Ax = 0$ nicht-triviale Lösungen $x \neq 0$.

Beweis. (a): Sei $A : K^n \to K^m, x \mapsto Ax$ die durch die Matrix A definierte lineare Abbildung. Die Menge der Lösungen von $Ax = 0$ ist $A^{-1}(0) = \ker(A) \subseteq K^n$ und damit ein linearer Unterraum. Da $\dim \operatorname{im}(A) = r(A)$ ergibt sich mit dem Homomorphiesatzes 5.10(b) $\dim \ker(A) = \dim K^n - \dim \operatorname{im}(A) = n - r(A)$.

(b): Sei x_0 eine Lösung von $Ax = b$. Ist $y \in K^n$ mit $Ay = 0$, so folgt $A(x_0 + y) = b$, sodass $x_0 + \ker(A) \subseteq A^{-1}(b)$. Ist umgekehrt $x \in A^{-1}(b)$ eine Lösung von $Ax = b$, so setze $y = x - x_0$. Dann ist $Ay = 0$ und $x = x_0 + y \in x_0 + \ker(A)$; dies zeigt $A^{-1}(b) \subseteq x_0 + \ker(A)$. □

Proposition 8.3. *(Existenz) Sei $A \in K^{m\times n}$ und sei $b \in K^m$. Betrachte das System (L): $Ax = b$ mit erweiterter Koeffizientenmatrix $B = [A, b]$. Dann gilt: (L) ist lösbar genau dann, wenn $r(A) = r(B)$ ist.*

Beweis. Sei $b = (\beta_i)$ und seien $s_1, \dots s_n$ die Spalten von A. Dann ist (L) genau dann lösbar, wenn $b \in \langle s_1, \dots, s_n\rangle$ ist. Dies gilt wegen

$$r(A) = \dim\langle s_1, \dots, s_n\rangle \leq \dim\langle s_1, \dots, s_k, b\rangle = r(B)$$

genau dann, wenn $r(A) = r(B)$ ist. □

Proposition 8.4. *(Eindeutigkeit) Sei $A \in K^{m\times n}$ und $b \in K^m$ mit $A^{-1}(b) \neq \emptyset$ (d.h. (L): $Ax = b$ hat eine Lösung). Dann hat (L) genau dann eine eindeutige Lösung, wenn $Ax = 0$ nur die triviale Lösung $x = 0$ hat; dies gilt genau dann, wenn $r(A) = n$ ist.*

• Sei $A \in K^{m\times n}$, sodass $Ax = b$ für alle $b \in K^m$ lösbar ist. Nach Proposition 8.3 ist dann $r(A) = m$. Sind diese Lösungen eindeutig, so folgt mit Proposition 8.4 $r(A) = n$. Also ist in diesem Fall A vom Typ (n, n) und wegen $r(A) = n$ invertierbar. Ist A^{-1} die inverse Matrix, so sind die eindeutigen Lösungen von $Ax = b$ genau die $x = A^{-1}b$.

Beweis. Ist $x_0 \in A^{-1}(b)$, so sind nach Lemma 8.2(b) die Lösungen von $Ax = b$ genau die Elemente von $x_0 + \ker(A)$ und x_0 ist die einzige

Lösung genau dann, wenn $\ker(A) = \{0\}$ ist. Da $\dim \ker(A) = n - r(A)$ gilt $\ker(A) = \{0\}$ genau dann, wenn $r(A) = n$ ist. □

Beispiel 8.5. Betrachte das lineare Gleichungssystem mit reellen Koeffizienten

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & 2x_3 & = & 0 \\ 2x_1 & + & 3x_2 & & & = & 9 \\ & & 2x_2 & + & x_3 & = & -1 \end{array}$$

Wir haben die Matrizen

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix} \text{ und } B = [A, b] = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 2 & 3 & 0 & 9 \\ 0 & 2 & 1 & -1 \end{pmatrix}$$

Für A berechnet man $r(A)$ mittels der elementaren Zeilenumformungen

$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix} \overset{z_2 - 2z_1}{\to} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -4 \\ 0 & 2 & 1 \end{pmatrix} \overset{z_3 - 2z_1}{\to} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 9 \end{pmatrix}$$

Es ist $r(A) = 3$. Aus $3 = r(A) \leq r(B) \leq 3$ folgt $r(A) = r(B) = 3$ und nach Proposition 8.3 hat $Ax = b$ eine Lösung. Da $r(A) = 3$ ist nach Proposition 8.4 diese Lösung eindeutig; es ist $x = A^{-1}b$.

Die Bestimmung von $r(A)$ und $r(B)$ kann durch elementare Umformungen durchgeführt werden und liefert nach den obigen Kriterien Existenz- und Eindeutigkeitsaussagen. Eine Variante solcher Umformungen liefert einen Lösungsalgorithmus für ein Gleichungssystem (L): $Ax = b$ mit $B = [A, b]$. Die zulässige Umformungen von (L) sind:

(a) Vertauschen der Zeilen von B (Permutation der Gleichungen),

(b) Vertauschen der Spalten von A (Permutation der $x_1, \ldots x_n$),

(c) Zeilenübergänge in B der Form $z_i \to z_i + \alpha z_j, i \neq j, \alpha \in K$.

Diese Operationen ändern die Lösungsmenge von (L) nicht (abgesehen von einer eventuellen Umnumerierung der x_i).

Gauss-Algorithmus. Wir können annehmen, dass A keine Nullspalte hat (d.h. jede Variable x_j kommt in dem System (L) nicht-trivial vor).

1. Nach Anwendung von (a) und (b) ist $A = (\alpha_{ij})$ mit $\alpha_{11} \neq 0$.

2. Mittels $z_i \to z_i - \alpha_{i1}/\alpha_{11} z_1$ für $i \geq 2$ ergibt sich ein System der Form

$$\begin{array}{rcl} \alpha_{11}x_1 + \sum_{k=2}^{n} \alpha_{1k}x_k & = & \beta_1 \\ \sum_{k=2}^{n} \alpha'_{jk}x_k & = & \beta'_j,\ j = 2, \ldots n. \end{array}$$

3. Betrachte das reduzierte System von $m-1$ Gleichungen

$$\sum_{k=2}^{n} \alpha'_{jk} x_k = \beta'_j,\ j = 2, \dots n.$$

Ist $(\alpha'_{ij}) = 0$ ist, so sind wir fertig. Ist $(\alpha'_{ij}) \neq 0$, so liefert Anwendung von **1.** und **2.** auf dieses System und dann weitere Iteration ein Gleichungsssystem von m Gleichungen in n Variablen y_i der Form

$$\begin{array}{rrcl} & \beta_{11}y_1 + \beta_{12}y_2 + \cdots + \beta_{1k}y_k + \cdots + \beta_{1n}y_n &=& \beta'_1 \\ & \beta_{22}y_2 + \cdots + \beta_{2k}y_k + \cdots + \beta_{2n}y_n &=& \beta'_2 \\ & & \cdot & \cdot \\ & & \cdot & \cdot \\ \text{(L')}: & \beta_{kk}y_k + \cdots + \beta_{kn}y_n &=& \beta'_k \\ & 0 &=& \beta'_{k+1} \\ & & \cdot & \\ & & \cdot & \\ & 0 &=& \beta'_m \end{array}$$

mit $\beta_{rr} \neq 0$ für $r = 1, \dots k$ (d.h. die Matrix (β_{ij}) ist in Zeilenstufenform mit Pivots $\beta_{11}, \dots \beta_{kk}$). Für das System (L') gilt:

- Ist $\beta'_r \neq 0$ für ein $r = k+1, \dots m$, so ist (L) nicht lösbar.
- Ist $\beta'_{k+1} = \cdots = \beta'_m = 0$ so ist (L') lösbar (und die $y_{k+1}, \dots, y_n$ können beliebig gewählt werden). Wegen $\beta'_{jj} \neq 0$ für $j = 1, \dots k$ lässt sich (L') sukzessive nach $y_k, y_{k-1}, \dots, y_1$ auslösen; eine eindeutige Lösung von (L') liegt dabei nur dann vor, wenn $k = n$ ist.

Beispiele 8.6. (a) Betrachte das System (mit reellen Koeffizenten)

$$\text{(L)}:\ B = [A, b] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 2 & 0 & 2 & 2 \end{pmatrix}$$

Es ist $\alpha_{11} \neq 0$. Die Operation $z_3 - 2z_1$ liefert ein neues System

$$\text{(L')}:\ B' = [A', b'] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & -2 & 0 & 0 \end{pmatrix}$$

Das reduzierte System hat $\alpha'_{22} = 0$. Vertauschen der 2. und 3. Spalte von A' ergibt eine Matrix A'' mit $\alpha''_{22} \neq 0$ und demselben b, d.h.

$$\text{(L")}:\ B'' = [A'', b'] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -2 & 0 \end{pmatrix}$$

Aus dieser Darstellung lässt sich durch sukzessive Auflösung die eindeutige Lösung $y = (0, 1, 0)$ bestimmen. Da wir die 2. und 3. Spalte vertauscht haben, ist die eindeutige Lösung von $B = [A, b]$ dann $x = (0, 0, 1)$.

(b) Die erweitere Koeffizientenmatrix (in $\mathbb{R}^{3\times 5}$)

$$B = [A, b] = \begin{pmatrix} 4 & 0 & 0 & 0 & 4 \\ 0 & 2 & 2 & 4 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

ist bereits in Zeilenstufenform. Für die Lösungen ergibt sich

$$x_4 = 1,\ 2x_2 + 2x_3 + 4 = 2,\ x_1 = 1.$$

Auflösen der 2. Gleichung nach x_2 liefert $x_2 = -x_3 - 1$. Also lässt sich jede Lösung x von $Ax = b$ in folgender Form schreiben

$$x = \begin{pmatrix} 1 \\ -x_3 - 1 \\ x_3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Dabei ist $x_0 = (1, -1, 0, 1)^t$ eine spezielle Lösung von $Ax = b$. Weiter liegt $(0, -1, 1, 0)^t \in \ker(A)$; wegen $r(A) = 3$ folgt $\dim \ker(A) = 4 - 3 = 1$, d.h. $(0, -1, -1, 0)$ ist eine Basis von $\ker(A)$. Die obige Beschreibung der Lösungsmenge von $Ax = b$ entspricht damit genau der Darstellung als affiner Unterraum $x_0 + \ker(A)$ von Lemma 8.2(b).

(c) Hat die erweiterte Koeffizientenmatrix die Form

$$B = [A, b] = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

so folgt aus der letzten Gleichung $0x_3 = 1$; das diese Gleichung keine Lösung hat ist das System $Ax = b$ nicht lösbar.

(d) Betrachte die folgende Matrix mit reellen Koeffizienten

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 7 \\ 0 & 2 & 5 \end{pmatrix}$$

Es ist $r(A) = 3$: Dies folgt, zum Beispiel, da allgemein gilt

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} A_1 B_1 & 0 \\ 0 & A_2 B_2 \end{pmatrix}$$

und in diesem Beispiel für die entsprechende 2×2-Matrix A_2

$$d(A_2) = 3 \cdot 5 - 2 \cdot 7 = 1 \neq 0$$

ist, d.h. $r(A_2) = 2$ und damit $r(A) = 3$. Damit hat $Ax = b$ für *jedes* $b \in \mathbb{R}^3$ die eindeutige Lösung $x = A^{-1}b$. Wegen $d = 1$ folgt weiter

$$A_2^{-1} = \begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix} \text{ und } A^{-1} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 5 & -7 \\ 0 & -2 & 3 \end{pmatrix}$$

Konkret ist zum Beispiel für $b = (1, 1, 1)^t$ die Lösung von $Ax = b$ somit

$$x = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 5 & -7 \\ 0 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -2 \\ 1 \end{pmatrix}$$

wie man durch direktes Nachrechnen leicht bestätigt.

9. Gruppen II

Um Determinanten von Matrizen definieren zu können benötigen wir einige elementare Aussagen der Gruppentheorie:

Definition 9.1. Seien G und H (multiplikativ geschriebene) Gruppen.

(1) Ein Homomorphismus (oder Gruppenhomomorphismus) ist eine Abbildung $f : G \to H$, sodass für alle $g_1, g_2 \in G$ gilt

$$f(g_1 g_2) = f(g_1) f(g_2).$$

(2) Ist $f : G \to H$ ein Gruppenhomomorphismus, so der Kern von f (bzw. das Bild von f) $\ker(f) = \{g \in G \mid f(g) = 1\}$ (bzw. $\mathrm{im}(f) = \{f(g) \mid g \in G\}$).

(3) Ein Homomorphismus $f : G \to H$ heisst Monomorphismus (bzw. Epimorphismus, Isomorphismus) falls f injektiv (bzw. surjektiv, bijektiv) ist. Gibt es einen Isomorphismus $f : G \to H$, so sind G und H isomorph, $G \cong H$. Die Isomorphismen $G \to G$ sind die Automorphismen von G.

• Für einen Gruppenhomomorphismus f gilt stets: $f(1) = 1$ und $f(g^{-1}) = f(g)^{-1}$.

• Kern und Bild eines Homomorphismus $f : G \to H$ sind Untergruppen; $\ker(f) \leq G$ und $\mathrm{im}(f) \leq H$.

• Ein Homomorphismus $f : G \to H$ ist genau dann ein Monomorphismus, wenn $\ker(f) = \{1\}$ ist.

• Ein Homomorphismus $f : G \to H$ ist genau dann ein Isomorphismus, wenn es einen Homomorphismus $g : H \to G$ mit $g \circ f = \mathrm{id}_G$ und $f \circ g = \mathrm{id}_H$ gibt. In diesem Fall ist $g = f^{-1}$.

Beispiele 9.2. (a) Sei G eine Gruppen und $a \in G$. Dann ist die Abbildung 'Konjugation mit a', d.h. $f_a : G \to G,\ g \mapsto a^{-1}ga$ ein Automorphismus: f_a ist Homomorphismus, da für $g_1, g_2 \in G$ gilt

$$f_a(g_1)f_a(g_2) = a^{-1}g_1aa^{-1}g_2a = a^{-1}g_1g_2a = f_a(g_1g_2).$$

Ist $f_a(g) = a^{-1}ga = 1$, so folgt $ga = a$ und $g = 1$, also ist $\ker(f_a) = \{1\}$ und f_a ist ein Monomorphismus. Für $g \in G$ ist $f_a(aga^{-1}) = g$, somit ist f_a auch surjektiv.

(b) Seien V, W K-Vektorräume und sei $f \in Hom_K(V, W)$. Dann besagt

$$f(v + v') = f(v) + f(v') \text{ für alle } v, v' \in V,$$

dass f ein Homomorphismus zwischen den V und W zugrundeliegenden additiven Gruppen $(V, +)$ und $(W, +)$ ist. In diesem Fall sind die abelschen Gruppen $\ker(f) \subseteq V$ und $\mathrm{im}(f) \subseteq W$ die den entsprechenden linearen Unterräumen zugrundeliegenden abelschen Gruppen.

(c) Sei K ein Körper und $A = (\alpha_{ij}) \in K^{2\times 2}$. Setze

$$\det(A) = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \in K.$$

Nach Beispiel 6.15 ist A invertierbar genau dann, wenn $\det(A) \neq 0$ ist. Sind $A, B \in K^{2\times 2}$, so zeigt eine einfache direkte Rechnung $\det(AB) = \det(A)\det(B)$. Also definiert det einen Homomorphismus

$$\det : GL_2(K) \to K^\times.$$

Ist $\alpha \in K^\times$, so ist

$$\det \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} = \alpha \in K^\times,$$

d.h. der Homomorphismus det ist ein Epimorphismus.

Definition 9.3. Sei G eine Gruppe. Eine Untergruppe $U \leq G$ ist ein Normalteiler (oder eine normale Untergruppe), $U \trianglelefteq G$, falls gilt

$$u \in U,\ g \in G \Rightarrow g^{-1}ug \in U.$$

Ist $U < G$ (d.h. $U \neq G$), so schreibe $U \lhd G$.

Beispiele 9.4. (a) Die trivialen Untergruppen $\{1\} < G$ und $G \leq G$ einer jeden Gruppe G sind Normalteiler, die trivialen Normalteiler.

(b) Ist G abelsch, so folgt aus $g^{-1}ug = g^{-1}gu = 1u = u \in U$, dass jede Untergruppe ein Normalteiler ist.

(c) Sei $f : G \to H$ ein Gruppenhomomorphismus und $U = \ker(f) \leq G$. Für $u \in U$ und $g \in G$ ist $f(g^{-1}ug) = f(g^{-1})f(u)f(g) = f(g)^{-1}f(g) = 1$, also ist $U = \ker(f) \trianglelefteq G$.

(d) Sei $G = S_3$ die Gruppe der bijektiven Abbildungen von $\{1, 2, 3\}$.

Für verschiedene Ziffern $i, j, k \in \{1, 2, 3\}$ sei (i, j) die bijektive Abbildung $i \mapsto j, j \mapsto i, k \mapsto k$. Wegen $(i, j)(i, j) = \mathrm{id}$ definiert jede solche Transposition (i, j) eine Untergruppe $\{\mathrm{id}, (i, j)\} < G$ der Ordnung 2. Sei nun $U = \{\mathrm{id}, (1, 2)\} < G$, $u = (1, 2)$ und $g = (1, 3) \in G$. Dann ist

$$g^{-1}ug = (13)(12)(13) = (23) \notin U,$$

d.h. die Untergruppe $U = \{\mathrm{id}, (1, 2)\} < S_3$ ist kein Normalteiler.

Ist V ein K-Vektorraum und $W \subseteq V$ ein linearer Unterraum, so ist auf dem Faktorraum $V/W = \{a + W \mid a \in V\}$ die Verküpfung $(a_1 + W) + (a_2 + W) = a_1 + a_2 + W$ wohldefiniert. Der naive Versuch analog für eine Gruppe G und eine Untergruppe $U \leq G$ eine Faktorgruppe zu definieren scheitert: In multiplikativer Notation ist die entsprechende Menge $G/U = \{gU \mid g \in G\}$ und die 'evidente' Mutiplikation

$$g_1U \cdot g_2U = g_1g_2U$$

ist im allgemeinen *nicht* wohl-definiert. Ist $g_1U = g_1'U$ und $g_2U = g_2'U$, so ist $g_1' = g_1u_1$ und $g_2' = g_2u_2$ für geeignete $u_1, u_2 \in U$. Es folgt

$$g_1'g_2' = g_1u_1g_2u_2 = g_1(g_2g_2^{-1})u_1g_2u_2 = g_1g_2(g_2^{-1}u_1g_2)u_2,$$

d.h. $g_1g_2U = g_1'g_2'U$ gilt nur dann, wenn $g_2^{-1}u_1g_2 \in U$ ist; dies ist gerade die Normalteilerbedingung an U. Insbesondere induziert die Multiplikation auf G nur dann eine wohl-definierte Multiplikation auf G/U, wenn $U \subseteq G$ nicht nur eine Untergruppe, sondern ein Normalteiler ist.

Lemma 9.5. *Sei $N \trianglelefteq G$ ein Normalteiler und $G/N = \{gN \mid g \in G\}$.*

(a) *Die Menge G/N ist mittels der Verknüpfung*

$$g_1N \cdot g_2N = g_1g_2N, \quad g_1, g_2 \in G$$

eine Gruppe mit neutralem Element N

(b) *Die Abbildung $\pi : G \to G/N$, $g \mapsto gN$ ist ein Epimorphismus mit $\ker(\pi) = N$.*

Beweis. (a): Da $N \trianglelefteq G$ ist, ist die Gruppenoperation auf G/N wohldefiniert, siehe oben; wegen $N = 1N$ und $1N \cdot gN = 1gN = gN$ ist N das neutrale Element.

(b): Die Abbildung π ist offensichtlich surjektiv und nach Definition der Gruppenoperation auf G/N ein Homomorphismus. Die letzte Aussage $N = \ker(\pi)$ gilt da $g \in \ker(\pi) \Leftrightarrow gN = N \Leftrightarrow g \in N$. □

Theorem 9.6. *(Homomorphiesatz) Sei $f : G \to H$ ein Gruppenhomomorphismus. Dann gibt es einen Epimorphismus $\pi : G \to G/\ker(f)$ und einen Monomorphismus $h : G/\ker(f) \to H$ mit $f = h \circ \pi$ und $\mathrm{im}(f) = \mathrm{im}(h)$.*

• Der Monomorphismus $h : G/\ker(f) \to H$ induziert einen Isomorphismus $h : G/\ker(f) \xrightarrow{\cong} \mathrm{im}(f)$.

Beweis. Da $\ker(f) \trianglelefteq G$ ein Normalteiler ist hat $G/\ker(f)$ eine Gruppenstruktur. Die Abbildungen π und h sind die evidenten Abbildungen

$$\begin{aligned} \pi : &\quad G \to G/\ker(f), \quad g \mapsto g\ker(f), \\ h : &\quad G/\ker(f) \to H, \quad g\ker(f) \mapsto f(g) \end{aligned}$$

Dabei ist π nach Konstruktion ein Epimorphismus. Man rechnet nach, dass h wohl-definiert ist; die restlichen Eigenschaften sind dann klar. □

Wir betrachten nun die symmetrische Gruppe S_n der bijektiven Abbildungen der Menge $\{1, \ldots, n\}$ aus Beipiel 2.2(d). Die Gruppenoperation auf S_n ist die Verknüpfung von Abildungen, S_n ist eine endliche Gruppe mit $|S_n| = n!$. Für $n \geq 3$ ist die Gruppe S_n *nicht* abelsch.

Die Elemente von S_n heissen Permutationen und wir bezeichnen diese mit kleinen griechischen Buchstaben; dabei sei ι das neutrale Element von S_n, d.h. die Identitätsabbildung. Für $\tau \in S_n$ schreibe

$$\tau = \begin{pmatrix} 1 & 2 & \cdot & \cdot & \cdot & n \\ \tau(1) & \tau(2) & \cdot & \cdot & \cdot & \tau(n) \end{pmatrix}$$

Bei feststehendem n lassen wir zur Vereinfachung der Notation oft die Ziffern mit $\tau(j) = j$ weg, zum Beispiel schreiben wir für $n = 6$ so

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

Definition 9.7. Seien $a_1, \cdots, a_k$ paarweise verschiedene Ziffern aus der Menge $\{1, .., n\}$. Ein k-Zykel in S_n ist eine Permutation der Form

$$\psi = \begin{pmatrix} a_1 & a_2 & \cdot & \cdot & a_{k-1} & a_k \\ a_2 & a_3 & \cdot & \cdot & a_k & a_1 \end{pmatrix};$$

wir verwenden für einen solchen k-Zykel auch die Notation

$$\psi = (a_1, a_2, \cdots, a_{k-1}, a_k) = (a_2, a_3, \cdots, a_k, a_1).$$

Zwei Zyklen $(a_1, a_2, \cdots, a_k)$ und $(b_1, b_2, \cdots, b_l)$ heissen disjunkt, falls $\{a_1, \ldots a_k\} \cap \{b_1, \ldots b_l\} = \emptyset$ ist. Eine Transposition ist ein 2-Zykel (i, j) (nach Definition ist dabei $i \neq j$).

Lemma 9.8. *Sei S_n die symmetrische Gruppe mit $n > 1$.*

(a) *Jede Permutation $\tau \in S_n$ hat eine Darstellung als ein Produkt von disjunkten Zyklen.*

(b) *Es gilt* $(a_1, a_2, \cdots, a_k) = (a_1, a_k)(a_1, a_{k-1}) \cdots (a_1, a_2)$, *d.h. jede Permutation lässt sich als ein Produkt von Transpositionen schreiben.*

• Ist $n = 5$, so gilt nach (b) $(1, 2, 3) = (1, 3)(1, 2)$. Wegen $(1, 2, 3) = (1, 3)(1, 2) = (1, 3)(1, 2)(4, 5)(4, 5)$ ist die Darstellung einer Permutation als ein Produkt von Transpositionen *nicht* eindeutig.

Beweis. (a): Jede Permutation lässt sich sukzessive in disjunkte Zyklen aufteilen, zum Beispiel ist

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{pmatrix} = (1, 2)(4, 5, 6)$$

Der Beweis von (a) ist eine Formalisierung dieses Prozesses und eine einfache Übung.

(b): Die erste Aussage folgt durch Nachrechnen (Permutationen sind von rechts zu lesen), die zweite aus (a). □

Theorem 9.9. *Sei* $n > 1$ *und sei* $\{-1, 1\}$ *die multiplikative Gruppe.*

(a) *Es gibt einen Epimorphismus*

$$\operatorname{sgn} : S_n \to \{-1, +1\}$$

mit $\operatorname{sgn}(\tau) = -1$ *für alle Transpositonen* $\tau \in S_n$.

(b) *Sei* K *ein Körper und* $f : S_n \to K^\times$ *ein Homomorphismus. Dann ist entweder* $f(\rho) = 1$ *für alle* $\rho \in S_n$, *oder es ist* $char(K) \neq 2$ *und* $f = \operatorname{sgn}$.

NB. Ist $\pi \in S_n$ und $\pi = \tau_1 \cdots \tau_k$ eine Zerlegung in Transpositionen, so ist nach (a) sgn ein Homomorphismus und $\operatorname{sgn}(\tau_i) = -1$ für alle i, also ist $\operatorname{sgn}(\pi) = (-1)^k$ (d.h. $\operatorname{sgn}(\pi) = 1$, falls π eine Darstellung durch eine gerade Anzahl von Transpositionen hat, und $\operatorname{sgn}(\pi) = -1$ sonst). Dies zeigt weiter: Die Zerlegung von π in ein Produkt vonTranspositionen ist nicht eindeutig, aber für jede solche Zerlegung gilt, dass die Parität (gerade oder ungerade) der Anzahl der Faktoren eindeutig ist.

Beweis. (a): Sei $T = \{i, j\} \subseteq \{1, \ldots, n\}$ mit $i < j$. Für $\tau \in S_n$ setze

$$Z_\tau(T) = \begin{cases} 1 & \text{falls } \tau(i) \leq \tau(j), \\ -1 & \text{falls } \tau(i) > \tau(j) \end{cases}$$

und definiere

$$\operatorname{sgn}(\tau) = \prod_T Z_\tau(T) \in \{-1, 1\},$$

wobei das Produkt über alle $T = \{i, j\} \subseteq \{1, \dots n\}$ mit $i \neq j$ läuft. Für $\rho \in S_n$ und $T = \{i, j\}$ setze $\rho T = \{\rho(i), \rho(j)\}$. Wir zeigen

$$(\#) \qquad Z_{\tau\rho}(T) = Z_\tau(\rho T) Z_\rho(T).$$

Gilt dies, so folgt

$$\mathrm{sgn}(\tau\rho) = \prod_T Z_{\tau\rho}(T) = \prod_T Z_\tau(\rho T) \prod_T Z_\rho(T) = \mathrm{sgn}(\tau)\mathrm{sgn}(\rho),$$

da mit T auch ρT alle 2-elementigen Teilmengen von $\{1, \dots n\}$ durchläuft; insbesondere definiert sgn einen Homomorphismus. Die Behauptung (#) ergibt sich aus der folgenden Tabelle, die die $Z_*(T)$ bzgl. der relative Lage von $\{\rho(i), \rho(j)\}$ und $\{\tau(\rho(i)), \tau(\rho(j))\}$ beschreibt. Für $T = \{i, j\}$ ist

	$Z_\rho(T)$	$Z_\tau(\rho T)$	$Z_{\tau\rho}(T)$
$\rho(i) \leq \rho(j)$ und $\tau(\rho(i)) \leq \tau(\rho(j))$	1	1	1
$\rho(i) \leq \rho(j)$ und $\tau(\rho(i)) > \tau(\rho(j))$	1	-1	-1
$\rho(i) > \rho(j)$ und $\tau(\rho(i)) \leq \tau(\rho(j))$	-1	-1	1
$\rho(i) > \rho(j)$ und $\tau(\rho(i)) > \tau(\rho(j))$	-1	1	-1

Es bleibt zu zeigen: $\mathrm{sgn}(\tau) = -1$ für jede Transposition $\tau \in S_n$. Ist $\tau = (1, 2)$, so ist $Z_\tau(T) = -1$ für $T = \{1, 2\}$ und $Z_\tau(T) = 1$ sonst, d.h. $\mathrm{sgn}(\tau) = -1$. Ist $\tau' = (i, j)$ beliebig, so gibt es ein

$$\rho = \begin{pmatrix} 1 & 2 & \cdots \\ i & j & \cdots \end{pmatrix} \in S_n$$

mit $\rho(1) = i$ und $\rho(2) = j$. Dann ist $\tau' = \rho\tau\rho^{-1}$, und es folgt

$$\mathrm{sgn}(\tau') = \mathrm{sgn}(\rho\tau\rho^{-1}) = \mathrm{sgn}(\rho)\mathrm{sgn}(\tau)\mathrm{sgn}(\rho)^{-1} = sgn(\tau) = -1,$$

wobei die vorletzte Identität verwendet, dass $\{-1, +1\}$ eine abelsche Gruppe ist.

(b): Sei $f : S_n \to K^\times$ ein Homomorphismus. Ist τ eine Transposition, so ist $\tau^2 = \iota$ und $1 = f(\tau^2) = f(\tau)^2$, also $f(\tau) \in \{-1, 1\}$. Da sich jede Permutation nach Lemma 9.8(b) als Produkt von Transpositonen schreiben lässt, folgt $f(\tau) \in \{-1, 1\}$ für jedes $\tau \in S_n$. Ist $\tau = (1, 2)$ und ist $\tau' = (i, j)$ eine beliebige Transposition, so liefert der Beweis von (a) ein $\rho \in S_n$ mit $\tau' = \rho\tau\rho^{-1}$. Also ist (die Gruppe $\{-1, 1\}$ ist abelsch)

$$f(\tau') = f(\rho\tau\rho^{-1}) = f(\rho)f(\tau)f(\rho)^{-1} = f(\tau)$$

d.h. für jede Transposition τ' gilt $f(\tau') = 1$ falls $f(\tau) = 1$ und $f(\tau') = -1$ falls $f(\tau) = -1$ ist. Ist $\rho \in S_n$ und ist $\rho = \tau_1 \cdots \tau_k$ eine Darstellung

als ein Produkt von Transpositionen, so gilt nach Teil (a)

$$\operatorname{sgn}(\rho) = \prod_{j=1}^{k} \operatorname{sgn}(\tau_j) = (-1)^k.$$

Andererseits ist

$$f(\rho) = \prod_{j=1}^{k} f(\tau_j) = \begin{cases} (-1)^k & \text{falls } f(\tau) = -1 \\ 1 & \text{falls } f(\tau) = 1 \end{cases}$$

Dies zeigt die Behauptung: Ist $f(\rho) \neq 1$ für ein $\rho \in S_n$, so ist $char(K) \neq 2$ (sonst ist $-1 = 1$). Weiter gibt es eine Transposition τ_j mit $f(\tau_j) = -1$ und damit gilt auch $f(\tau) = -1$, sodass $f = \operatorname{sgn}$. □

Definition 9.10. Für $n \geq 2$ ist $A_n = \ker\{\operatorname{sgn} : S_n \to \{-1, 1\}\}$ die alternierende Gruppe auf n Ziffern.

- Es ist $A_n \trianglelefteq S_n$ und $|S_n : A_n| = 2$.
- Für jedes $\tau \in S_n$ mit $\tau(\pi) = -1$ ist $S_n = A_n \cup \tau A_n = \tau A_n \cup A_n$, wobei die Vereinigung jeweils disjunkt ist.

Beispiel 9.11. Sei $U < S_n$ eine Untergruppe mit $|S_n : U| = 2$. Dann ist $U = A_n$: Betrachte die Abbildung

$$f : S_n \to \{-1, 1\},\ f(\tau) = \begin{cases} 1 & \text{falls } \tau \in U, \\ -1 & \text{falls } \tau \notin U. \end{cases}$$

Man rechnet nach, dass f ein Homormorphismus ist; die Behauptung folgt dann mit Theorem 9.9(b).

Bemerkung 9.12. Für $n = 3$ und $n \geq 5$ sind $\{1\}, A_n$ und S_n die einzigen Normalteiler von S_n, und A_n besitzt nur die trivialen Normalteiler $\{1\}$ und A_n (man sagt A_n ist eine *einfache* Gruppe). Für $n = 4$ ist A_4 nicht-einfach, da $V = \{\iota, (12)(34), (13)(24), (14)(23)\}$ einen Normalteiler mit $|V| = 4$ definiert. Die Existenz dieser Untergruppe ist der Grund dafür, dass Lösungen von Polynomgleichungen der Form

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0,\ a_i \in \mathbb{C},$$

für $n \leq 4$ stets durch Wurzeln ausgedrückt werden können (der Fall $n = 2$ ist die 'Mitternachtsformel'), dies für $n \geq 5$ aber nicht gilt. Die systematische Analyse der Lösungen solcher Polynomgleichungen erfolgt im Rahmen der Galoistheorie und ist ein Thema der Vorlesung 'Algebra'.

10. Determinanten

Nach Beispiel 6.15 gilt lässt sich für eine Matrix $A = (\alpha_{ij}) \in K^{2\times 2}$ aus den Koeffizienten α_{ij} bestimmen, ob A invertierbar ist, genauer

$$A^{-1} \text{ existiert } \Leftrightarrow d = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \neq 0.$$

Ist A invertierbar, so ist die inverse Matrix A^{-1} durch die Formel

$$A^{-1} = \frac{1}{d}\begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix}$$

gegeben, die ebenfalls d involviert. Der Ausdruck $d \in K$ ist ein Spezialfall einer sogenannten Determinante. Wir ordnen im folgenden jeder Matrix $A \in K^{n\times n}$ eine Determinante $\det(A) \in K$ zu. Diese Invariante enkodiert Information über die Invertierbarkeit von A, lässt sich zur Berechnung von A^{-1} benüzten, und hat geometrische Bedeutung.

Wir definieren die Determinante einer quadratischen Matrix allgemeiner für Matrizen $A = (\alpha_{ij})$, deren Einträge α_{ij} nicht Elemente eines Körpers, sondern allgemeiner Elemente eines kommutativen Rings sind. Ein kommutativer Ring ist dabei eine Menge mit zwei Verknüpfungen, die alle Bedingungen an einen Körper erfüllt, mit Ausnahme der Existenz von mutiplikativ inversen Elementen, genauer:

Definition 10.1. Ein Ring R ist eine Menge, zusammen mit zwei Verknüpfungen $+$ und $\cdot$, sodass gilt:

(1) R ist bzgl. $+$ eine abelsche Gruppe (mit neutralem Element 0),
(2) Es gibt ein $1 \in R$ mit $1r = r = r1$ für $r \in R$; es gilt das Assoziativgesetz $r_1(r_2r_3) = (r_1r_2)r_3$ für $r_1, r_2, r_3 \in R$.
(3) Es gelten die Distributivgesetze, d.h. für $r_1, r_2, r_3 \in R$ ist

$$r_1(r_2 + r_3) = r_1r_2 + r_1r_3 \text{ und } (r_1 + r_2)r_3 = r_1r_3 + r_2r_3.$$

Ein Ring R ist kommutativ, falls zusätzlich gilt

(4) $r_1r_2 = r_2r_1$ für $r_1, r_2 \in R$.

• Die Definition besagt nicht, dass wie in einem Körper $0 \neq 1$ sein muss, inbesondere ist der Nullring $R = \{0\}$ definiert.

Beispiele 10.2. (a) Die ganzen Zahlen $\mathbb{Z}$ formen bzgl. der üblichen Addition und Multiplikation einen kommutativen Ring.

(b) Ist R ein (kommutativer) Ring, so bildet die Menge der n-Tupel

$$R^n = \{(r_1, \ldots, r_n) \mid r_i \in R\}$$

bzgl. der komponentenweisen Addition und Multiplikation wieder einen (kommutativen) Ring.

(c) Die Menge der Polynome $K[x]$ (bzw. $R[x]$) über einem Körper (bzw.

einem kommutativen Ring R) bilden bzgl. der Addition und Mutiplikation von Polynomen einen kommutativen Ring.

(d) Sei $R^{n\times n}$ die Menge der quadratischen Matrizen $A = (\alpha_{ij})$ vom Typ (n, n) mit Einträgen α_{ij} aus einem kommutativen Ring R. Dann ist $R^{n\times n}$ bzgl. der Addition und Multipliaktion von Matrizen ein Ring. Der Ring $R^{n\times n}$ ist für $n \geq 2$ in der Regel (z.B. falls $0 \neq 1$ in R) nicht kommutativ.

Definition 10.3. Sei R ein kommutativer Ring und $A = (\alpha_{ij}) \in R^{n\times n}$ eine quadratische Matrix vom Typ (n, n). Die Determinante von A ist

$$\det(A) = \sum_{\tau\in S_n} \operatorname{sgn}(\tau)\alpha_{1\tau(1)}\alpha_{2\tau(2)}\cdots\alpha_{n\tau(n)} \in R$$

• Ist $n = 6$, so hat die die Determinante definierende Formel 720 Terme, d.h. obige Definition liefert ein Ringelement, welches a priori nur schwer berechenbar ist.

Beispiele 10.4. (a) Sei $A = (\alpha_{ij}) \in R^{n\times n}$ eine Dreiecksmatrix mit $\alpha_{ij} = 0$ für $j > i$. Dann sind die in der Determinante $\det(A)$ auftretenden Summanden $\alpha_{1\tau(1)}\alpha_{2\tau(2)}\cdots\alpha_{n\tau(n)}$ nur dann nicht-trivial, wenn $\tau(1) \leq 1, \tau(2) \leq 2, \cdots, \tau(n) \leq n$ ist. Dies gilt nur für die Identitätsabbildung ι, sodass $\det(A)$ das Produkt der Diagonaleinträge ist

$$\det(A) = \alpha_{11}\alpha_{22}\cdots\alpha_{nn}.$$

(b) Sei $A = (\alpha_{ij}) \in R^{2\times 2}$. Es gilt $S_2 = \{\iota, \tau\}$, wobei $\tau = (1, 2)$ ist. Nach Definition ist die Determinante von A dann (vgl. Beispiel 6.15)

$$\begin{aligned}\det(A) &= \operatorname{sgn}(\iota)\alpha_{1\iota(1)}\alpha_{2\iota(2)} + \operatorname{sgn}(1,2)\alpha_{1\tau(1)}\alpha_{2\tau(2)}\\ &= 1\cdot\alpha_{11}\alpha_{22} + (-1)\cdot\alpha_{12}\alpha_{21}\\ &= \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \in R.\end{aligned}$$

(c) In $\mathbb{R}^2$ seien Vektoren $s_1 = (x_1, y_1)^t$ und $s_2 = (x_2, y_2)^t$ gegeben. Wir schreiben diese Vektoren in Polarkoordinaten, d.h. in der Form

$$x_j = r_j\cos(\alpha_j) \text{ und } y_j = r_j\sin(\alpha_j),$$

wobei $0 \leq \alpha_1 \leq \alpha_2 \leq \pi/2$ sei. Sei $\gamma = \alpha_2 - \alpha_1$ der Winkel zwischen (x_2, y_2) und (x_1, y_1). Elementar-Geometrische Überlegungen zeigen, dass die Fläche F des von den Spaltenvektoren s_1 und s_2 aufgespannten Parallelogramms durch die folgende Formel gegeben ist

$$F = r_1 r_2\sin(\gamma).$$

Betrachte den Winkel α_1 zwischen s_1 und der x-Achse. Die Matrix

$$D = \begin{pmatrix}\cos(-\alpha_1) & -\sin(-\alpha_1)\\ \sin(-\alpha_1) & \cos(-\alpha_1)\end{pmatrix}$$

beschreibt die Drehung um den Winkel $-\alpha_1$. Es gilt $\det(D) = 1$. Für

$$A = (s_1, s_2) = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$$

gilt also

$$DA = \begin{pmatrix} r_1 & * \\ 0 & r_2 \sin(\gamma) \end{pmatrix}$$

wobei (siehe Übung) $\det(A) = \det(D)\det(A) = \det(DA) = r_1 r_2 \sin(\gamma)$. Es folgt $F = \det(A)$, d.h. die Determinante entspricht genau dem Volumen des durch die Vektoren aufgespannten Parallelogramms.

Das obige Beispiel (c) suggeriert, dass die Determinante allgemein eine 'Volumenfunktion' ist. Wir zeigen im folgenden, dass jede 'abstrakte Volumenfunktion' bis auf eine Konstante durch die Determinanten gegeben ist. Wir beginnen mit elementaren Eigenschaften.

Für $A \in R^{n\times n}$ mit Zeilen $z_1, \dots z_n$ und Spalten $s_1, \dots, s_n$ betrachten wir im folgenden $\det(A)$ als eine Funktion der Zeilen bzw. Spalten

$$\det(A) = f_{\det}(z_1, \dots, z_n) = g_{\det}(s_1, \dots, s_n).$$

Lemma 10.5. *Sei $A \in R^{n\times n}$. Dann gilt*

(a) $\det(A) = \det(A^t)$,

(b) *Für $r, r' \in R$ und $z_j, z_j' \in R^n$ gilt die Formel*

$$f_{\det}(*, rz_j + r'z_j', *) = r f_{\det}(*, z_j, *) + r' f_{\det}(*, z_j', *)$$

(d.h. für $R = K$ ein Körper und j fest ist die Abbildung $z_j \mapsto f_{\det}(z_1, \dots, z_{j-1}, z_j, z_{j+1}, \dots, z_n)$ linear).

(c) *Ist $z_i = z_j$ für ein $i \neq j$, so ist $f_{\det}(z_1, \dots, z_n) = 0$.*

(d) *Die zu (b) und (c) analogen Aussagen für $g_{\det}(s_1, \dots, s_n)$ gelten.*

Beweis. (a): Sei $A^t = (\beta_{ij})$ mit $\beta_{ij} = \alpha_{ji}$. Nach Definition von A^t ist

$$\begin{aligned} \det(A^t) &= \textstyle\sum_{\tau\in S_n} \operatorname{sgn}(\tau)\beta_{1\tau(1)} \cdots \beta_{n\tau(n)} \\ &= \textstyle\sum_{\tau\in S_n} \operatorname{sgn}(\tau)\alpha_{\tau(1)1} \cdots \alpha_{\tau(n)n} \end{aligned}$$

Wegen $\operatorname{sgn}(\tau)^2 = 1$ ist $\operatorname{sgn}(\tau) = \operatorname{sgn}(\tau)^{-1} = \operatorname{sgn}(\tau^{-1})$. Da τ eine Bijektion ist gilt $\tau(i) = j$ genau dann, wenn $i = \tau^{-1}(j)$ ist. Umordnung der Faktoren (dies verwendet, dass der Ring R kommutativ ist) liefert

$$\alpha_{\tau(1)1} \cdots \alpha_{\tau(n)n} = \alpha_{1\tau^{-1}(1)} \cdots \alpha_{n\tau^{-1}(n)}.$$

Also ist

$$\det(A^t) = \sum_{\tau\in S_n} \operatorname{sgn}(\tau^{-1})\alpha_{1\tau^{-1}(1)} \cdots \alpha_{n\tau^{-1}(n)} = \det(A).$$

(b): Sei $z_j = (\alpha_{j1}, \ldots, \alpha_{jn})$ und $z'_j = (\alpha'_{j1}, \ldots, \alpha'_{jn})$. Dann gilt wegen

$$\begin{aligned}\textstyle\sum_{\tau\in S_n} \operatorname{sgn}(\tau)\alpha_{1\tau(1)}\ldots(r\alpha_{j\tau(j)} + r'\alpha'_{j\tau(j)})\ldots\alpha_{n\tau(n)} = \\ = r\textstyle\sum_{\tau\in S_n} \operatorname{sgn}(\tau)\alpha_{1\tau(1)}\ldots\alpha_{j\tau(j)}\ldots\alpha_{n\tau(n)} \\ + r'\textstyle\sum_{\tau\in S_n} \operatorname{sgn}(\tau)\alpha_{1\tau(1)}\ldots\alpha'_{j\tau(j)}\ldots\alpha_{n\tau(n)}\end{aligned}$$

die Formel $f_{\det}(*, rz_j + r'z'_j, *) = rf_{\det}(*, z_j, *) + r'f_{\det}(*, z'_j, *)$.
(c): Sei $z_i = z_j$ mit $i < j$ und sei $\sigma = (i, j)$. Dann ist $S_n = A_n \cup A_n\sigma$ eine disjunkte Zerlegung und $\sum_{\tau\in S_n} = \sum_{\pi\in A_n} + \sum_{\pi\sigma\in A_n\sigma}$. Es folgt

$$\begin{aligned} & \operatorname{sgn}(\pi)(\alpha_{1\pi(1)}\ldots\alpha_{n\pi(n)}) + \operatorname{sgn}(\pi\sigma)(\alpha_{1\pi\sigma(1)}\ldots\alpha_{n\pi\sigma(n)}) \\ = \; & \operatorname{sgn}(\pi)(\alpha_{1\pi(1)}\ldots\alpha_{n\pi(n)}) - \operatorname{sgn}(\pi)(\alpha_{1\pi\sigma(1)}\ldots\alpha_{n\pi\sigma(n)}) \\ = \; & \operatorname{sgn}(\pi)(\alpha_{1\pi(1)}\ldots\alpha_{n\pi(n)} - \alpha_{1\pi\sigma(1)}\ldots\alpha_{n\pi\sigma(n)}) = 0\end{aligned}$$

da wegen $z_i = z_j$ auch $\alpha_{i\pi\sigma(i)} = \alpha_{i\pi(j)} = \alpha_{j\pi(j)}$.
(d): Folgt mit (a) aus (b) und (c). □

Definition 10.6. Sei R ein kommutativer Ring. Eine Abbildung

$$V : (R^n)^n \to R$$

ist eine abstrakte Volumenfunktion auf R^n falls gilt:

(1) Für $r, r' \in R$ und $z_j, z'_j \in R^n$ ist

$$V(*, rz_j + r'z'_j, *) = rV(*, z_j, *) + r'V(*, z'_j, *)$$

(2) Ist $z_i = z_j \in R^n$, $i \neq j$, so ist $V(z_1, \ldots, z_n) = 0$.

• Nach Lemma 10.5(b)(c) definiert die Determinante eine abstrakte Volumenfunktion $f_{\det} : (R^n)^n \to R,\ (z_1, \ldots, z_n) \mapsto f_{\det}(z_1, \ldots, z_n)$.

Proposition 10.7. *Für eine abstrakte Volumenfunktion V auf R^n gilt*

(a) *Für $i \neq j$ und $r \in R$ ist*

$$V(z_1, \ldots, z_i + rz_j, \ldots, z_n) = V(z_1, \ldots, z_i, \ldots, z_n)$$

(b) *Für $\tau \in S_n$ ist*

$$V(z_{\tau(1)}, \ldots, z_{\tau(n)}) = \operatorname{sgn}(\tau)V(z_1, \ldots z_n).$$

(c) *Ist $z_i = (\alpha_{i1}, \ldots, \alpha_{in})$ und sind $e_1, \ldots, e_n$ die Vektoren in R^n mit 1 an der Stelle i und 0 sonst, so ist*

$$V(z_1, \ldots, z_n) = \det(\alpha_{ij})V(e_1, \ldots e_n).$$

NB. Die letzte der obigen Aussagen besagt, dass jede abstrakte Volumenfunktion auf R^n bis auf eine Konstante die Determinante ist:

$$V(z_1, \ldots, z_n) = f_{det}(z_1, \ldots, z_n) \cdot V(e_1, \ldots, e_n) = f_{det}(z_1, \ldots, z_n) \cdot c,$$

wobei $c = V(e_1, \ldots, e_n) \in R$ eine Konstante ist.

Beweis. (a): Folgt direkt aus den Eigenschaften (1) und (2) von V.
(b): Sei $\tau = (i, j)$ mit $i < j$. Da V linear in jeder Komponente ist folgt

$$V(*, z_i, *, z_j, *) + V(*, z_j, *, z_i, *) = V(*, z_i + z_j, *, z_j + z_i, *) = 0$$

und die Behauptung gilt für eine Transposition. Eine beliebige Permutation $\tau \in S_n$ lässt sich τ nach Lemma 9.8(b) als ein Produkt geeigneter Transpositionen $\tau = \tau_1 \cdots \tau_k$ schrieben. Setze $\rho = \tau_2 \cdots \tau_k$. Induktion (nach der Anzahl der Faktoren in einer solchen Zerlegung) liefert

$$\begin{aligned} V(z_{\tau(1)}, \ldots, z_{\tau(n)}) &= V(z_{\tau_1\rho(1)}, \ldots, z_{\tau_1\rho(n)}) \\ &= \operatorname{sgn}(\tau_1) V(z_{\rho(1)}, \ldots, z_{\rho(n)}) \\ &= \operatorname{sgn}(\tau_1)\operatorname{sgn}(\rho) V(z_1, \ldots, z_n) \\ &= \operatorname{sgn}(\tau) V(z_1, \ldots, z_n) \end{aligned}$$

(c): Sei $z_i = \sum_{j=1}^{n} \alpha_{ij} e_j$ für $i = 1, \ldots, n$. Da nach (1) V in jeder Komponente linear ist folgt

$$V(z_1, \ldots, z_n) = \sum_{(j_1, \ldots, j_n)} \alpha_{1j_1} \ldots \alpha_{nj_n} V(e_{j_1}, \ldots, e_{j_n}).$$

Taucht in einem solchen n-Tupel $(j_1, \ldots, j_n)$ eine Zahl wiederholt auf, so gilt nach (2) $V(e_{j_1}, \ldots, e_{j_n}) = 0$. Damit sind die nicht-trivialen Summanden in der obigen Summe genau diejenigen mit $\{j_1, \ldots, j_n\} = \{1, \ldots, n\}$, und zu jedem solchen Summanden gibt es genau eine Permutation $\tau \in S_n$ mit $j_i = \tau(i)$. Mit (b) folgt so

$$\begin{aligned} V(z_1, \ldots, z_n) &= \textstyle\sum_{\tau \in S_n} \alpha_{1\tau(1)} \ldots \alpha_{n\tau(n)} \operatorname{sgn}(\tau) V(e_1, \ldots, e_n) \\ &= \det(\alpha_{ij}) V(e_1, \ldots, e_n) \end{aligned}$$

□

Lemma 10.8. *Für $A, B \in R^{n\times n}$ gilt* $\det(AB) = \det(A)\det(B)$.

• Sei K ein Körper und $A \in K^{n\times n}$. Dann ist A genau dann invertierbar, wenn die Spaltenvektoren von A linear unabhängig sind, also genau dann, wenn $\det(A) \neq 0$ ist. Insbesondere definiert det eine Abbildung $\det : GL_n(K) \to K^\times$; dies ist ein Epimorphismus.

• Seien $z_1, \ldots, z_n \in R^n$ und V eine beliebige abstrakte Volumenfunktion. Dann gibt es eine Konstante $c \in R$, sodass

$$V(z_1, \ldots, z_n) = c \cdot f_{det}(z_1, \ldots, z_n).$$

Ist $A \in R^{n\times n}$ mit Zeilen $z_1, \ldots, z_n$, und ist $B \in R^{n\times n}$ eine weitere Matrix, so hat AB die Zeilen $z_1B, \ldots, z_nB$ und das obige Lemma besagt

$$V(z_1B, \ldots, z_nB) = c{\cdot}\det(AB) = c{\cdot}\det(A)\det(B) = V(z_1, \ldots z_n)\det(B),$$

d.h. $\det(B)$ ist der 'Verzerrungsfaktor' der Volumenfunktion V bei Anwendung von B.

Beweis. Seien $z_1, \ldots, z_n$ die Zeilen von A. Betrachte die Abbildung

$$f_B : (R^n)^n \to R,\ (z_1, \ldots, z_n) \mapsto f_{\det}(z_1B, \ldots, z_nB) = \det(AB).$$

Wir zeigen, dass f_B eine abstrakte Volumenfunktion auf R^n definiert. Gilt dies, so gibt es nach Proposition 10.7 eine Konstante $c(B)$ mit

$$f_B(z_1, \ldots, z_n) = f_{\det}(z_1, \ldots, z_n)c(B) = \det(A)c(B).$$

Ist speziell $A = E_n$ die Einheitsmatrix mit den Zeilen $e_1, \ldots, e_n$, so ist

$$\det(B) = \det(E_nB) = f_B(e_1, \ldots, e_n) = \det(E_n)c(B) = c(B),$$

also ist $c(B) = \det(B)$ und $\det(AB) = f_B(z_1, \ldots, z_n) = \det(A)\det(B)$.

Wir verifizieren die Eigenschaften einer Volumenfunktion: Sei $z_i = z_j$ für ein $i \neq j$. Dann stimmen in AB die Zeilen z_iB und z_jB überein, und nach Lemma 10.5(c) ist $\det(AB) = 0$, also gilt

$$f_B(z_1, \ldots, z_n) = \det(AB) = 0.$$

Sei $A = (\alpha_{ij}) \in R^{n\times n}$ mit Zeilen $z_1, \ldots, z_n$, und sei $z'_j = (\alpha'_{j1}, \ldots, \alpha'_{jn})$. Nach Lemma 10.5(b) gilt für $B \in R^{n\times n}$ und $r, r' \in R$ die Formel

$$\det \begin{pmatrix} z_1B \\ \cdot \\ \cdot \\ (rz_j + r'z'_j)B \\ \cdot \\ \cdot \\ z_nB \end{pmatrix} = r \det \begin{pmatrix} z_1B \\ \cdot \\ \cdot \\ z_jB \\ \cdot \\ \cdot \\ z_nB \end{pmatrix} + r' \det \begin{pmatrix} z_1B \\ \cdot \\ \cdot \\ z'_jB \\ \cdot \\ \cdot \\ z_nB, \end{pmatrix}$$

also ist

$$f_B(*, rz_j + r'z'_j, *) = rf_B(*, z_j, *) + r'f_B(*, z'_j, *)$$

und f_B definiert eine abstrakte Volumenfunktion. □

Wir kommen zur Berechnung von Determinanten.

Lemma 10.9. *(Kästchensatz) Seien $B \in R^{m\times m}$, $C \in R^{n\times n}$ und weiter $D \in R^{n\times m}$. Setze $k = m + n$ und betrachte die $k \times k$-Matrix*

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

Dann gilt: $\det(A) = \det(B)\det(C)$.

Beweis. Sei $E_l \in R^{l\times l}$ die Einheitsmatrix. Aufgrund der Darstellung

$$\begin{pmatrix} B & 0 \\ D & C \end{pmatrix} = \begin{pmatrix} B & 0 \\ D & E_n \end{pmatrix} \begin{pmatrix} E_m & 0 \\ 0 & C \end{pmatrix},$$

genügt es nach Lemma 10.8 die folgenden Identitäten zu beweisen

$$\det \begin{pmatrix} B & 0 \\ D & E_n \end{pmatrix} = \det(B) \text{ und } \det \begin{pmatrix} E_m & 0 \\ 0 & C \end{pmatrix} = \det(C)$$

Seien $z_1, \ldots, z_n \in R^n$ und sei $C(z_i) \in R^{n \times n}$ die durch die Zeilen $z_1, \ldots, z_n$ definierte Matrix. Dann definiert die Abbildung

$$h : (R^n)^n \to R,\ (z_1, \ldots, z_n) \mapsto \det \begin{pmatrix} E_m & 0 \\ 0 & C(z_i) \end{pmatrix}$$

eine abstrakte Volumenfunktion. Nach Lemma 10.7(b) unterscheidet sich h von der Determinatenfunktion nur um eine Konstante, d.h. es gibt ein $c \in R$, sodass für jede Matrix $C(z_i)$ die folgende Identität gilt

$$h(z_1, \ldots, z_n) = c \cdot f_{\det}(z_1, \ldots, z_n) = c \cdot \det(C(z_i)).$$

Ist $C(z_i) = E_n$ die Einheitsmatrix mit den Zeilen $e_1, \ldots, e_n$, so folgt

$$1 = \det(E_k) = h(e_1, \ldots, e_n) = c \cdot f_{\det}(e_1, \ldots, e_n) = c \cdot \det(E_n) = c,$$

also ist $c = 1$ und somit $h(z_1, \ldots, z_n) = f_{\det}(z_1, \ldots, z_n)$. Dies zeigt

$$\det \begin{pmatrix} E_m & 0 \\ 0 & C \end{pmatrix} = \det(C).$$

Die verbleibende Behauptung folgt analog: Für $z_1, \ldots, z_m \in R^m$ und $B(z_i) \in R^{m \times m}$ die Matrix mit Zeilen $z_1, \ldots, z_m$ definiert die Abbildung

$$k : (R^m)^m \to R,\ (z_1, \ldots, z_m) \mapsto \det \begin{pmatrix} B(z_i) & 0 \\ D & E_n \end{pmatrix}$$

eine abstrakte Volumenfunktion. Also gibt es ein $c \in R$, sodass

$$k(z_1, \ldots, z_m) = c \cdot f_{\det}(z_1, \ldots, z_m) = c \cdot \det(B(z_i)).$$

Der Fall $B(z_i) = E_m$ liefert $c = 1$ und es folgt

$$\det \begin{pmatrix} B & 0 \\ D & E_n \end{pmatrix} = \det(B).$$

□

Beispiele 10.10. (a) Sei K ein Körper und $A, B, C, D \in K^{n \times n}$ mit $AC = CA$. Wegen $AC = CA$ gilt dann die Identiät

$$\begin{pmatrix} E_n & 0 \\ -C & A \end{pmatrix} \cdot \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & AD - CB \end{pmatrix}.$$

Mit Lemma 10.9 folgt

$$\det(A) \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \det(AD - CB)$$

Gilt weiter $\det(A) \neq 0$, so kann man durch $\det(A)$ teilen und erhält

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB),$$

d.h. die Determinante der $2n \times 2n$-Matrix auf der linken Seite dieser Gleichung ist gleich der Determinante der $n \times n$-Matrix auf der rechten Seite. (d.h. für $n = 4$ lässt sich eine Summe mit $8! = 40320$ Termen durch eine Summe mit $4! = 24$ Termen berechnen).
(b) Seien $A, B, C, D \in K^{n \times n}$ wie in (a). Gilt $AC = CA$, so kann man zeigen

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB),$$

auch wenn $\det(A) = 0$ ist. Für $AC \neq CA$ gilt dies nicht: Betrachte

$$A = D = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ und } B = -C = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.$$

Es gilt $\det(A) = 1$ und $AC \neq CA$. Direktes Nachrechnen zeigt weiter

$$\det \begin{pmatrix} A & -C \\ C & A \end{pmatrix} = 2 \neq 1 = \det(A^2 + C^2).$$

Wir geben abschliessend zwei allgemeine Methoden zur Berechnung von Determinaten an.

I. Zeilenstufenform. Sei $R = K$ ein Körper und $A \in K^{n \times n}$. Nach Theorem 7.4 gibt es Elementarmatrizen $T_1, \ldots, T_k, S_1, \ldots, S_l \in K^{n \times n}$, sodass $A' = T_k \cdots T_1 A S_1 \cdots S_l$ eine obere Dreiecksmatrix ist. Nach Definition haben alle Elementarmatrizen die Determinante 1 und Lemma 10.8 besagt, dass $\det(A') = \det(A)$ ist. Für die obere Dreiecksmatrix A' ist $\det(A')$ nach Beispiel 10.4(a) und Lemma 10.5(a) das Produkt der Diagonaleinträge und leicht berechenbar.

Beispiel 10.11. Sei $A \in \mathbb{R}^{3 \times 3}$ die Matrix

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

Wie in Beispiel 7.5(a) lässt sich A mittels elementarer Umformungen in die Matrix

$$A' = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 8 & 10 \\ 0 & 0 & 1/2 \end{pmatrix}$$

überführen. Es folgt $\det(A) = 1 \cdot 8 \cdot (1/2) = 4$.

II. Laplace Entwicklung. Betrachte $A = (\alpha_{ij}) \in R^{3\times 3}$. Sei $A_{ij} \in R$ die Determinante der Matrix, die aus A durch Ersetzen der i-ten Zeile durch die Zeile e_j (mit 1 an der Stelle j und 0 sonst) entsteht, z.B.

$$A_{12} = \det \begin{pmatrix} 0 & 1 & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}.$$

Dies liefert 9 Elemente A_{ij} von R. Sei $\widetilde{A} = (A_{ij})^t \in R^{3\times 3}$. Dann ist

$$A \cdot \widetilde{A} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \cdot \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}$$

Dieses Produkt hat an der Stelle $(1,1)$ den Ausdruck

$$\alpha_{11} \cdot \det \begin{pmatrix} 1 & 0 & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} + \alpha_{12} \cdot \det \begin{pmatrix} 0 & 1 & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} + \\ + \alpha_{13} \cdot \det \begin{pmatrix} 0 & 0 & 1 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$$

Da die Determinante in einer Zeile linear ist, ist dies genau $\det(A)$. Analog steht in dem Produkt $A \cdot \widetilde{A}$ an den Stellen $(2,2)$ und $(3,3)$ das Ringelement $\det(A)$. An den Stellen (i,j) mit $i \neq j$ ist der Eintrag 0, da nach Definition von $\widetilde{A}$ in diesem Fall der Eintrag die Determinante einer Matrix mit zwei identischen Zeilen ist. Also ist $A \cdot \widetilde{A} = \det(A) \cdot E_3$. Insbesondere ist im Fall $R = K$ ein Körper die Matrix A invertierbar genau dann, wenn $\det(A) \neq 0$, und in diesem Fall ist $A^{-1} = \det(A)^{-1}\widetilde{A}$. Weiter ergibt sich eine explizite Möglichkeit für die Berechnung von $\det(A)$: Die in der obigen Summe auftretenden Determinanten lassen sich nach einer Vertauschung von Spalten mittels dem Kästchensatz 10.9 leicht berechnen. Da eine solche Vertauschung τ nach Lemma 10.7(c) die Determinante nur um $\operatorname{sgn}(\tau)$ ändert, folgt

$$\det(A) = \alpha_{11} \det \begin{pmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} - \alpha_{12} \det \begin{pmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{pmatrix} + \alpha_{13} \det \begin{pmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix}$$

Dies ist die Berechnung von $\det(A)$ mittels 'Entwicklung nach der 1.Zeile'. Analog lässt sich die Determinante mittels Entwicklung nach einer beliebigen Zeile oder Spalte berechnen.

Sei konkret K ein Körper und betrachte die Matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 2 & 0 & 3 \end{pmatrix} \in K^{3\times 3}$$

Für die Determinante von A ergibt sich durch Entwicklung nach der 1. Zeile die Formel

$$\begin{aligned} \det(A) &= 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} -1 & 2 \\ 2 & 3 \end{pmatrix} + 1 \cdot \det \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \\ &= 15 \end{aligned}$$

Für die Adjunkte $\widetilde{A} = (A_{ij})^t$ ergibt sich

$$A_{11} = \det \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 2 \\ 2 & 0 & 3 \end{pmatrix} = 3$$

und

$$A_{12} = \det \begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & 2 \\ 2 & 0 & 3 \end{pmatrix} = 7,$$

sowie nach Berechung von weiteren 7 Determinanten die Adjunkte

$$\widetilde{A} = \begin{pmatrix} 3 & -6 & 3 \\ 7 & 1 & -3 \\ -2 & 4 & 3 \end{pmatrix}$$

sodass

$$A^{-1} = \frac{1}{\det(A)} \widetilde{A} = \frac{1}{15} \begin{pmatrix} 3 & -6 & 3 \\ 7 & 1 & -3 \\ -2 & 4 & 3 \end{pmatrix}.$$

Allgemein gilt:

Definition 10.12. Sei $A = (\alpha_{ij}) \in R^{n\times n}$, und sei $A_{ij} \in R$ die Determinante der Matrix, die aus A durch Ersetzen der i-ten Zeile durch $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit 1 an der Stelle j entsteht, d.h.

$$A_{ij} = \det \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & \cdots & 1 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdots & \cdot \\ \alpha_{n1} & \cdots & \alpha_{nj} & \cdots & \alpha_{nn} \end{pmatrix}$$

Die Adjunkte $\widetilde{A}$ von A ist die Matrix $(A_{ij})^t$.

• Die explizite Berechnung von $\widetilde{A}$ für $A \in R^{n\times n}$ erfordert die Berechnung von n^2 Determinanten.

• Sei $A \setminus \{ij\}$ die Matrix, die aus A durch Streichen der i-ten Zeile und der j-ten Spalte entsteht. Aus Lemma 10.7(b) und Lemma 10.9 folgt

$$A_{ij} = (-1)^{i+j} \det(A \setminus \{ij\}).$$

• Die Einträge A_{ij} lassen sich gleichwertig als die Determinante derjenigen Matrix definieren, die aus A durch Ersetzen der j-ten Spalte durch $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit 1 an der Stelle i entsteht.

Theorem 10.13. *Sei $A = (\alpha_{ij}) \in R^{n\times n}$. Dann gilt:*

(a) $A\widetilde{A} = det(A)E_n$, *d.h.* $\sum_{j=1}^n \alpha_{ij}A_{kj} = \delta_{ik}\det(A)$; *ist* $k = i$, *so ist* $\sum_{j=1}^n \alpha_{ij}A_{ij} = det(A)$ *(Entwicklung nach der i-ten Zeile).*

(b) $\widetilde{A}A = det(A)E_n$, *d.h.* $\sum_{j=1}^n A_{ji}\alpha_{jk} = \delta_{ik}\det(A)$; *ist* $k = i$, *so ist* $\sum_{j=1}^n A_{ji}\alpha_{ji} = det(A)$ *(Entwicklung nach der i-ten Spalte).*

(c) *Ist $R = K$ ein Körper, so ist A genau dann invertierbar, wenn $det(A) \neq 0$ ist. In diesem Fall ist $A^{-1} = det(A)^{-1}\widetilde{A}$.*

Beweis. (a): Das Produkt $A\widetilde{A}$ hat an der Stelle (i, k) den Eintrag

$$\sum_{j=1}^n \alpha_{ij}A_{kj} = \delta_{ik}\det(A).$$

Somit sind die einizgen nicht-trivialen Einträge von $A\widetilde{A}$ die Einträge $\det(A)$ auf der Diagonalen ($i = k$), und $A\widetilde{A} = \det(A)E_n$.

(b): Analog mittels A_{ij} als Determinante nach Spaltenvertauschung.

(c): Existiert A^{-1}, so ist $AA^{-1} = E_n$, und $\det(A) \neq 0$ folgt aus

$$1 = \det(E_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}).$$

Ist umgekehrt $\det(A) \neq 0$, so gilt nach (a) $A^{-1} = \det(A)^{-1}\widetilde{A}$. □

Beispiele 10.14. (a) Betrachte die Matrix mit reellen Einträgen

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

Nach Theorem 10.13(a) ist $\det(A) = \sum_{j=1}^3 \alpha_{ij}A_{ij}$. Im Fall $i = 1$ (Entwicklung nach der 1. Zeile) ergibt sich die Formel

$$\det(A) = \alpha_{11}A_{11} + \alpha_{12}A_{12} + \alpha_{13}A_{13}.$$

Wegen $\alpha_{11} = 1$ folgt mit dem Kästchensatz 10.9

$$A_{11} = \det(1) \cdot \det \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix} = 1 \cdot (4-1) = 3,$$

also ist $\alpha_{11}A_{11} = 1 \cdot 3 = 3$. Nachrechnen liefert $\alpha_{12}A_{12} = 2 \cdot 2 = 4$ und $\alpha_{13}A_{13} = 3 \cdot (-8) = -24$. Also ist $\det(A) = 3 + 4 - 24 = -17$.

Effizienter ist hier die Entwicklung nach der 1. Spalte (da der Eintrag $\alpha_{21} = 0$ ist). Explizit:

$$\det(A) = 1 \cdot \det \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix} = 1 \cdot (4-1) + 2 \cdot (2-12) = -17.$$

(b) Für die Matrix

$$A = \begin{pmatrix} 5 & 0 & 3 & -1 \\ 3 & 0 & 0 & 4 \\ -1 & 2 & 4 & -2 \\ 1 & 0 & 0 & 5 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

liefert Entwicklung nach den jeweiligen Spalten mit vielen Nullen

$$\det(A) = -2 \cdot \det \begin{pmatrix} 5 & 3 & -1 \\ 3 & 0 & 4 \\ 1 & 0 & 5 \end{pmatrix} = -2 \cdot (-3) \det \begin{pmatrix} 3 & 4 \\ 1 & 5 \end{pmatrix} = 6 \cdot 11 = 66.$$

11. Dualräume

Wir ordnen einem K-Vektorraum V einen dualen K-Vektorraum V^* zu und studieren die Beziehungen zwischen V und V^*. Als Anwendung beweisen wir mit Hilfe von Dualräumen, dass sich affine Unterräume als Schnitte von affinen Hyperebenen interpretieren lassen.

Sind V und W K-Vektorräume, so ist die Menge der linearen Abbildungen $Hom_K(V, W)$ nach Lemma 6.1(a) bzgl. der punktweisen Addition und Skalarmultiplikation ein K-Vektorraum. Ist $\{a_i \mid i \in I\}$ eine Basis von V und sind $\{w_i \mid i \in I\}$ beliebige Elemente von W, so bestimmt nach Lemma 5.4(a) die Zuordnung $f(a_i) = w_i$ eine eindeutige lineare Abbildung $f \in \mathrm{Hom}_K(V, W)$.

Wir betrachten im folgenden den Spezialfall $W = K^1 = K$.

Definition 11.1. Sei V ein K-Vektorraum. Dann ist der K-Vektorraum

$$V^* = Hom_K(V, K)$$

der zu V duale Vektorraum (oder Dualraum zu V); die Elemente von V^* heissen Linearformen auf V.

• Nach Lemma 5.4(a) ist V^* ein K-Vektorraum. Ist $\{a_1, \ldots, a_n\}$ eine Basis von V, so folgt wie in Beispiel 4.18(d), dass die $f_i \in V^*$ mit

$$f_i(a_j) = \delta_{ij},\ i = 1, \ldots, n$$

eine Basis von V^* bilden; die Basis $\{f_1, \ldots, f_n\}$ ist die zu $\{a_1, \ldots, a_n\}$ duale Basis (und umgekehrt); insbesondere ist $\dim_K V = \dim_K V^* = n$.

Beispiele 11.2. (a) Sei $V = \mathbb{R}^2$ mit Basis $\{a_1, a_2\}$, $a_1 = (1, 0)$ und $a_2 = (1, 1)$. Dann ist die zu $\{a_1, a_2\}$ duale Basis $\{f_1, f_2\}$ gegeben durch

$$f_1(a_1) = 1,\ f_1(a_2) = 0,\ f_2(a_1) = 0 \text{ und } f_2(a_2) = 1.$$

Die zu der Standardbasis $\{e_1, e_2\}$ von V duale Basis hat die Form $\{f_1', f_2'\}$ mit $f_1'(e_1) = 1,\ f_1'(e_2) = 0,\ f_2'(e_1) = 0$ und $f_2'(e_2) = 1$. Hier ist $f_2'(a_1) = f_2'(e_1) = 0$ und $f_2'(a_2) = f_2'(e_1 + e_2) = 0 + 1 = 1$, d.h. f_2 und f_2' bestimmen dieselbe Abbildung auf V, aber $f_1 \neq f_1'$, da $f_1'(e_2) = 0$ und $f_1(e_2) = f_1(a_2) - f_1(a_1) = -1$ (d.h. $a_1 = e_1$ impliziert nicht $f_1 = f_1'$ und $a_2 \neq e_2$ impliziert nicht $f_2 \neq f_2'$).

(b) Sei $V = \mathbb{R}[x]$ der $\mathbb{R}$-Vektorraum der Polynome mit reellen Koeffizienten und Basis $\{x^i \mid i \in \mathbb{N}_0\}$. Betrache die $f_i \in V^*$ mit $f_i(x^j) = \delta_{ij}$. Da sich die Linearform $f \in V^*$ mit $f(x^i) = 1$ für alle i nicht als endliche Linearkombination der f_i darstellen lässt ist $f \notin \langle f_i |\ i \in \mathbb{N}_0 \rangle$; insbesondere bilden die zu der Basis $\{x^i \mid i \in \mathbb{N}_0\}$ von V dualen Elemente $\{f_i \mid i \in \mathbb{N}_0\}$ keine Basis von V^*.

Lemma 11.3. *(Trennungslemma) Sei V ein K-Vektorraum und sei $U \subseteq V$ ein linearer Unterraum endlicher Dimension $\dim_K U = n$. Ist $v \in V \setminus U$, so gibt es ein $f \in V^*$ mit $f(u) = 0$ für $u \in U$ und $f(v) = 1$.*

Beweis. Sei $\{u_1, \ldots, u_n\}$ eine Basis von U und $v \in V \setminus U$. Dann ist $\{u_1, \ldots, u_n, v\}$ linear unabhängig und es gibt ein $g \in \langle u_1, \ldots, u_n, v \rangle^*$ mit $g(u_i) = 0$ und $g(v) = 1$. Die Linearform $f \in V^*$ definiert durch $f(a) = 0$ für $a \in V \setminus \langle u_1, \ldots, u_n, v \rangle$ und $f = g$ sonst hat die geforderten Eigenschaften. □

Proposition 11.4. *Sei V ein K-Vektorraum und $V^{**} = (V^*)^*$. Setze*

$$T : V \to V^{**},\ (Tv)(f) = f(v),\ v \in V,\ f \in V^*.$$

(a) *T ist ein Monomorphismus,*

(b) *Ist $\dim_K V = n < \infty$, so ist T ein Isomorphismus.*

Beweis. (a): Nachrechnen zeigt, dass $Tv \in V^{**}$ und T linear ist; zum Beispiel: Die Abbildung Tv ist additiv, da für $f_1, f_2 \in V^*$ gilt

$$(Tv)(f_1 + f_2) = (f_1 + f_2)(v) = f_1(v) + f_2(v) = (Tv)(f_1) + (Tv)(f_2).$$

Sei $0 \neq v \in V$. Nach Lemma 24.3 (angewandt mit $U = \{0\}$) gibt es ein $f \in V^*$ mit $f(v) = 1$. Es folgt $(Tv)(f) = f(v) = 1$, d.h. $Tv \neq 0$ und T ist ein Monomorphismus.
(b): Ist $\dim_K V = n$, so gilt $\dim_K V^* = \dim_K V^{**} = n$ und der Monomorphismus T ist nach Lemma 5.11 ein Isomorphismus. □

Sei V ein K-Vektorraum und $U \subseteq V$ ein linearer Unterraum. Ist $f \in V^*$, d.h. $f : V \to K$ ist linear, so gibt es eine K-lineare Restriktionsabbildung $R : V^* \to U^*, f \mapsto Rf = f|_U$. Setze $U^\perp = \ker(R) = \{f \in V^* \mid f(U) = 0\} \subseteq V^*$. Mit dem Homomorphiesatz folgt $V^*/U^\perp \cong \mathrm{im}(R)$.

Weiter definiert $I : (V/U)^* \to V^*, g \mapsto Ig$ mit $(Ig)(v) = g(v+U)$ eine K-lineare Abbildung. Für $u \in U$ ist $(Ig)(u) = g(u+U) = g(U) = 0$, sodass $I : (V/U)^* \to U^\perp$.

Wir definieren allgemeiner:

Definition 11.5. Sei V ein K-Vektorraum und $U \subseteq V$ ein linearer Unterraum.

(a) Die Restriktion $R : V^* \to U^*$ ist die K-lineare Abbildung
$$(Rf)(u) = f(u),\ u \in U,\ f \in V^*.$$

(b) Die Inflation $I : (V/U)^* \to V^*$ ist die K-lineare Abbildung
$$(Ig)(v) = g(v + U),\ v \in V,\ g \in (V/U)^*.$$

(c) Ist $M \subseteq V$ eine Teilmenge, so ist $M^\perp$ ($\perp$ 'senkrecht')
$$M^\perp = \{f \in V^* \mid f(m) = 0 \text{ für alle } m \in M\}$$
die Menge der Linearformen, die M annullieren.

(d) Ist $S \subseteq V^*$ eine Teilmenge, so ist $S^\top$ ($\top$ 'umgedreht senkrecht')
$$S^\top = \{v \in V \mid s(v) = 0 \text{ für alle } s \in S\}$$
die Menge der Vektoren, die von Linearformen in S annulliert werden.

• $M^\perp \subseteq V^*$ ist für jede Menge M ein linearer Unterraum.

Beispiel 11.6. Sei $V = \mathbb{R}^2$ mit der Standardbasis $\{e_1, e_2\}$ und sei $\{f_1, f_2\}$ die entsprechende duale Basis von V^*. Sind $v = \alpha_1 e_1 + \alpha_2 e_2 \in V$ und $f = \beta_1 f_1 + \beta_2 f_2 \in V^*$, so folgt
$$f(v) = (\beta_1 f_1 + \beta_2 f_2)(\alpha_1 e_1 + \alpha_2 e_2) = \beta_1\alpha_1 + \beta_2\alpha_2.$$
Wir verwenden diese Relation um für $U = \langle e_2 \rangle \subseteq V$ das Objekt $(U^\perp)^\top \subseteq V$ zu berechnen: Nach Definition ist $U^\perp = \{f \in V^* \mid f(u) = 0 \text{ für alle } u \in U\} = \{f \in V^* \mid f(e_2) = 0\}$. Für $v = e_2$ (d.h. $\alpha_1 = 0$

und $\alpha_2 = 1$) ergibt sich aus der obigen Relation $f(e_2) = 0$ genau dann, wenn $\beta_2 = 0$, d.h. $f(e_2) = 0$ genau dann, wenn $f = \beta_1 f_1$ und somit $U^\perp = \langle f_1 \rangle \subseteq V^*$. Betrachte $(U^\perp)^\top = \langle f_1 \rangle^\top \subseteq V$. Nach Definition ist $\langle f_1 \rangle^\top = \{v \in V \mid \beta_1 f_1(v) = 0\}$. Ist $v = \alpha_1 e_1 + \alpha_2 e_2$, so ist $f_1(v) = \alpha_1$, d.h. $\langle f_1 \rangle^\top = \langle e_2 \rangle = U \subseteq V$. Wir haben gezeigt $(U^\perp)^\top = U$, dies ist ein Spezialfall eines allgemeinen Dualitätssatzes.

Lemma 11.7. *Sei V ein K-Vektorraum und $U \subseteq V$ ein linearer Unterraum. Dann gilt*

(a) *$R : V^* \to U^*, f \mapsto Rf$, ist linear mit $\ker(R) = U^\perp$.*

(b) *Ist $\dim_K V = n < \infty$, so ist R ein Epimorphismus und es gilt*

$$\dim_K U^\perp = n - \dim_K U.$$

(c) *$I : (V/U)^* \to U^\perp$ ist ein Isomorphismus.*

Beweis. (a): Folgt aus der Diskussion vor Definition 24.5.

(b): Sei $\{u_1, \ldots, u_r\}$ eine Basis von U und $\{u_1, \ldots, u_r, v_{r+1}, \ldots, v_n\}$ eine Basis von V. Erweitere $f \in U^*$ auf $g \in V^*$ durch $g(u_i) = f(u_i)$ und $g(v_i) = 0$. Damit ist $g(u) = f(u)$ für $u \in U$, d.h. $f = Rg$ und R ist surjektiv. Wegen $\mathrm{im}(R) = U^*$, $\ker(R) = U^\perp$ und $\dim_K U = \dim_K U^*$ liefert der Homomorphiesatz die Identität $\dim_K U^\perp = n - \dim_K U$.

(c): Betrachte die lineare Abbildung $I : (V/U)^* \to U^\perp$. Sei $f \in U^\perp$. Dann definiert $\bar{f}(v + U) = f(v)$ ein Element von $(V/U)^*$ mit $I\bar{f} = f$, d.h. I ist surjektiv. Sei $g \in \ker(I)$. Für $v \in V$ ist $0 = (Ig)(v) = g(v + U)$, also ist $g = 0$ und I ist injektiv. □

Theorem 11.8. *(Dualitätssatz) Sei V ein K-Vektorraum endlicher Dimension $\dim_K V = n$. Dann gilt*

(a) *Ist $U \subseteq V$ ein linearer Unterraum, so ist $U^{\perp\top} = U$.*

(b) *Ist $S \subseteq V^*$ ein linearer Unterraum, so ist $S^{\top\perp} = S$.*

(c) *Sind U_1 und U_2 lineare Unterräume von V, so ist*

$$(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp \text{ und } (U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp.$$

Beweis. (a): Es ist $U^{\perp\top} = \{v \in V \mid f(v) = 0 \text{ für } f \in U^\perp\} \supseteq U$. Ist $v \in U^{\perp\top} \setminus U$, so gibt es nach dem Trennungslemma 24.3 ein $f \in V^*$ mit $f(v) = 1$ und $f(u) = 0$ für $u \in U$; insbesondere ist $f \in U^\perp$. Da $v \in U^{\perp\top}$ folgt $0 = f(v) = 1$, Widerspruch. Also gilt Gleichheit $U = U^{\perp\top}$.

(b): Nach Definition von $S^{\top\perp}$ gilt $S^{\top\perp} \supseteq S$. Da S und $S^{\top\perp}$ lineare Unterräume endlicher Dimension sind genügt zu zeigen $\dim_K S^{\top\perp} = \dim_K S$. Betrachte die Abbildung $T : V \to V^{**}$, $(Tv)(f) = f(v)$, $v \in$

$V,\ f \in V^*$. Für $s \in S \subseteq V^*$ ist $(Tv)(s) = s(v)$. Dies liefert die Identität

$$\begin{aligned} S^\top &= \{v \in V \mid s(v) = 0 \text{ für alle } s \in S\} \\ &= \{v \in V \mid (Tv)(s) = 0 \text{ für alle } s \in S\}. \end{aligned}$$

Nach Proposition 24.4(b) ist T ein Isomorphismus ist, also ist $\dim_K S^\top = \dim_K T(S^\top)$ und es folgt

$$\begin{aligned} \dim_K S^\top &= \dim_K T(\{v \in V \mid (Tv)(s) = 0 \text{ für alle } s \in S\}) \\ &= \dim_K \{f \in V^{**} \mid f(s) = 0 \text{ für alle } s \in S\} \\ &= \dim_K S^\perp \\ &= \dim_K V^* - \dim_K S, \end{aligned}$$

wobei die letzten Gleichung aus Lemma 24.8(b) (angewandt auf $U = S$ und $V = V^*$) folgt. Nach nochmaliger Anwendung von Lemma 24.8(b) (mit $U = S^\top$ und $V = V$) folgt wegen $\dim_K V = \dim_K V^*$ dann

$$\dim_K S^{\top\perp} = \dim_K V - \dim_K S^\top = \dim_K S,$$

und somit $S = S^{\top\perp}$.

(c): Übung. □

Als Anwendung des Dualitätssatz betrachten wir affine Unterräume eines n-dimensionalen K-Vektorraums V. Nach Definition 5.13 hat ein m-dimensionaler affiner Unterraum von V die Form $H = a + U$, wobei $a \in V$ und $U \subseteq V$ ein m-dimensionaler linearer Unterraum ist. Ist $\dim_K U = n - 1$, so ist $H = a + U$ eine affine Hyperebene.

Ist $V = \mathbb{R}^2$, so sind die affinen Hyperebenen genau die Geraden in V und jeder Punkt ist der Schnitt von 2 Geraden. Im Fall $V = \mathbb{R}^3$ sind die affinen Hyperebenen die Flächen, jeder Punkt ist Schnitt von 3 Flächen und jede Gerade ist Schnitt von zwei Flächen. Dies suggeriert, dass allgemein ein m-dimensionaler affiner Unterraum einens n-dimensionalen Vektorraums der Schnitt von $n-m$ vielen affinen Hyperebenen ist. Wir beweisen dies unter Verwendung des Dualitätssatz 24.9.

Die affinen Hyperebenen lassen sich mittels Linearformen wie folgt beschreiben: Sei $\dim_K V = n$ und $0 \neq f \in V^*$. Dann ist $\mathrm{im}(f) = K$ und $\dim \ker(f) = n - 1$. Wähle ein $w \in V$ mit $f(w) \neq 0$. Für $\alpha \in K$ setze $a = \alpha f(w)^{-1} w \in V$, sodass $f(a) = \alpha$ ist. Dann definiert

$$H_{f,\alpha} = a + \ker(f) = \{v \in V \mid f(v) = \alpha\}$$

eine affine Hyperebene in V. Jede affine Hyperebene $H = a+U \subseteq V$ hat diese Form: Wegen $\dim_K U = n-1$ ist nach Lemma 24.8(b) $\dim U^\perp = 1$ und $U^\perp$ wird von einem $0 \neq f \in V^*$ erzeugt. Für $\alpha = f(a)$ folgt

$$H = H_{f,\alpha}.$$

Theorem 11.9. *Sei V ein n-dimensionaler K-Vektorraum. Dann ist jeder m-dimensionale affine Unterraum der Durchschnitt von $n-m$ affinen Hyperebenen.*

Beweis. Sei $H = a + U \subseteq V$, $a \in V$, $U \subseteq V$ ein m-dimensionaler linearer Unterraum. Nach Lemma 24.8(b) ist $\dim_K U^\perp = n - m$; sei $\{f_1, \ldots, f_{n-m}\}$ eine Basis von $U^\perp$. Evaluierung der f_i bei a liefert Skalare α_i und affine Hyperebenen H_{f_i,α_i}. Wir zeigen $H = \cap_{i=1}^{n-m} H_{f_i,\alpha_i}$:

Ist $v = a + u \in H$, so gilt für $i = 1, \ldots, n-m$ nach Definition

$$f_i(v) = f_i(a) + f_i(u) = f_i(a) + 0 = \alpha_i.$$

Also ist $v \in \cap_{i=1}^{n-m} H_{f_i,\alpha_i}$. Ist umgekehrt $v \in \cap_{i=1}^{n-m} H_{f_i,\alpha_i}$, so folgt wegen $f_i(v-a) = 0$ dann $v - a \in f_i^\top$ und somit $v - a \in \langle f_1, \ldots, f_{n-m}\rangle^\top$. Der Dualitätssatz 24.9(a) zeigt $\langle f_1, \ldots, f_{n-m}\rangle^\top = U^{\perp\top} = U$, also ist $v - a \in U$ und $v \in a + U = H$. □

Bemerkung 11.10. Für lineare Gleichungssysteme ergibt sich folgende geometrische Interpretation: Für ein lineares Gleichnungssystem

$$(\mathrm{L}): \qquad \sum_{j=1}^{n} \alpha_{ij} x_j = \beta_i, \quad i = 1, \ldots, m,$$

betrachte die m Gleichungen separat. Die i-te Gleichung hat die Form

$$(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})(x_1, \ldots, x_n)^t = \beta_i.$$

Interpretiert man $x = (x_1, \ldots, x_n)^t$ als Vektor in $V = K^n$ und ist $(\alpha_{i1}, \ldots, \alpha_{in})$ nicht der Nullvektor, so definiert

$$f_i : V \to K, \; x \mapsto (\alpha_{i1}, \ldots, \alpha_{in})x.$$

eine nicht-triviale Linearform und die Lösung der i-ten Gleichung entspricht den Elementen der affinen Hyperebene H_{f_i,β_i}. Die Lösungsmenge von (L) ist genau der Schnitt der m affinen Hyperebenen H_{f_i,β_i}, insbesondere hat (L) eine Lösung falls dieser Schnitt nicht leer ist, und eine eindeutige Lösung, falls dieser Schnitt aus genau einem Punkt besteht.

Die zulässigen Spaltenumformungen, d.h. die Operationen

(a)	$f_i \mapsto \alpha f_i$,	$\alpha \in K^\times$,
(b)	$f_i \mapsto f_i + f_j$,	$i \neq j$,
(c)	Vertauschung von f_i und f_j,	$i \neq j$,

sind Operationen in V^*; man kann zeigen, dass diese Operationen den Schnitt der H_{f_i,β_i} nicht ändern. Genauer: Anwendung der Spaltenumformungen (a)-(c) liefert lineare Abbildungen $f_i' : V \to K$, $i = 1, \ldots, m$, zusammen mit Skalaren β_i', $i = 1, \ldots, m$, sodass gilt

$$\cap_{i=1}^{m} H_{f_i,\beta_i} = \cap_{i=1}^{m} H_{f_i',\beta_i'}.$$

Definition 11.11. Seien V, W K-Vektorräume und $A \in \mathrm{Hom}_K(V, W)$. Ist $f \in W^*$ und $v \in V$, so definiert $f(Av)$ ein Element von K. Setze

$$A^* : W^* \to V^*,\ f \mapsto \{v \mapsto f(Av)\};$$

die Abbildung A^* ist die zu A duale Abbildung.

• Nach Definition ist $(A^*f)(v) = (f \circ A)(v) = f(Av) : V \to K$ die Komposition von zwei K-linearen Abbildungen und somit ein Element von V^*. Seien $f_1, f_2 \in W^*$. Für $v \in V$ gilt

$$\begin{aligned}(A^*(f_1 + f_2))(v) &= (f_1 + f_2)(Av) = f_1(Av) + f_2(Av) = \\ &= (A^*f_1)(v) + (A^*f_2)(v) = (A^*f_1 + A^*f_2)(v),\end{aligned}$$

also ist $A^*(f_1 + f_2) = A^*f_1 + A^*f_2$. Ähnlich folgt $A^*(\alpha f) = \alpha(A^*f)$ für $f \in W^*$ und $\alpha \in K$; damit ist $A^* \in \mathrm{Hom}_K(W^*, V^*)$.

• Sind V, W K-Vektorräume, $A, B \in \mathrm{Hom}_K(V, W)$ und $\alpha \in K$, so gilt

$$(A + B)^* = A^* + B^* \text{ und } (\alpha A)^* = \alpha A^*.$$

• Seien V_i K-Vektorräume, $i = 1, 2, 3$, $A \in \mathrm{Hom}_K(V_1, V_2)$ und $B \in \mathrm{Hom}_K(V_2, V_3)$. Dann gilt für das Kompositum der dualen Abbildungen

$$(BA)^* = A^*B^*.$$

Wir zeigen: Hat $A \in \mathrm{Hom}_K(V, W)$ bzgl. Basen von V und W die Matrix (α_{ij}), so hat A^* bzgl. der dualen Basis die Matrix $(\alpha_{ij})^t$, d.h.

$$(A \leftrightarrow (\alpha_{ij})) \Leftrightarrow (A^* \leftrightarrow (\alpha_{ij})^t).$$

Lemma 11.12. *Seien V, W K-Vektorräume, $\{v_1, \ldots, v_n\}$ eine Basis von V, $\{w_1, \ldots, w_m\}$ eine Basis von W, und $A \in \mathrm{Hom}_K(V, W)$ mit*

$$Av_i = \sum_{j=1}^{m} \alpha_{ji} w_j, \quad i = 1, \ldots, n.$$

Ist $\{f_1, \ldots, f_n\}$ die duale Basis zu $\{v_1, \ldots, v_n\}$ und $\{g_1, \ldots, g_m\}$ die duale Basis zu $\{w_1, \ldots, w_m\}$, so gilt für die duale Abbildung A^*

$$A^*g_j = \sum_{i=1}^{n} \alpha_{ji} f_i, \quad j = 1, \ldots, m.$$

Beweis. Nach Definition der dualen Abbildung gilt

$$(A^*g_j)(v_k) = g_j(Av_k) = g_j(\sum_{l=1}^{m} \alpha_{lk} w_l) = \sum_{l=1}^{m} \alpha_{lk}(g_j w_l).$$

Nach Definition $g_j(w_l) = \delta_{jl}$, also ist $g_j(w_l) = 0$ für $j \neq l$ und es folgt

$$(A^* g_j)v_k = \alpha_{jk}.$$

Wegen $f_i(v_k) = \delta_{ik}$ folgt genauso $f_i(v_k) = 0$ für $i \neq k$, und damit

$$(A^* g_j)(v_k) = \alpha_{jk} = (\sum_{i=1}^{n} \alpha_{ji} f_i)(v_k).$$

□

Ist V ein K-Vektorraum, so gilt nach Definition 24.5(3) und (4)

$$\begin{aligned} M \subseteq V &\Rightarrow M^\perp = \{f \in V^* \mid f(m) = 0 \text{ für alle } m \in M\}, \\ S \subseteq V^* &\Rightarrow S^\top = \{v \in V \mid s(v) = 0 \text{ für alle } s \in S\}. \end{aligned}$$

Proposition 11.13. *Seien V, W K-Vektorräume und sei $A \in \mathrm{Hom}_K(V, W)$. Dann gilt*

(a) $\ker(A^*) = \mathrm{im}(A)^\perp$.

(b) $\ker(A) = \mathrm{im}(A^*)^\top$.

Ist $\dim_K W < \infty$, so gilt weiter

(c) *A ist Epimorphismus $\Leftrightarrow$ A^* ist Monomorphismus.*

(d) *A ist Monomorphismus $\Leftrightarrow$ A^* ist Epimorphismus.*

(e) *A ist Isomorphismus $\Leftrightarrow$ A^* ist Isomorphismus.*

Beweis. (a): Aus den Definitionen ergibt sich direkt

$$\begin{aligned} \ker(A^*) &= \{g \in W^* \mid (A^* g)(v) = g(Av) = 0 \text{ für alle } v \in V\} \\ &= \{g \in W^* \mid g(w) = 0 \text{ für alle } w \in \mathrm{im}(A)\} \\ &= \mathrm{im}(A)^\perp. \end{aligned}$$

(b): Eine Richtung ist offensichtlich, da

$$\begin{aligned} \mathrm{im}(A^*)^\top &= \{v \in V \mid g(v) = 0 \text{ für alle } g \in \mathrm{im}(A^*)\} \\ &= \{v \in V \mid (A^* f)(v) = f(Av) = 0 \text{ für alle } f \in W^*\} \\ &\supseteq \ker(A). \end{aligned}$$

Sei $v \in \mathrm{im}(A^*)^\top$. Dann ist (vgl.oben) $(A^* f)(v) = f(Av) = 0$ für alle $f \in W^*$. Angenommen $Av \neq 0$. Das Trennungslemma 24.3 (angewandt mit $U = \{0\}$ und $V = W$) liefert ein $f \in W^*$ mit $f(Av) = 1$, Widerspruch. Somit ist $Av = 0$ und $v \in \ker(A)$.

(c): Ist A ein Epimorphismus, so ist $\mathrm{im}(A) = W$ und $\ker(A^*) = \mathrm{im}(A)^\perp = W^\perp = \{0\}$ nach (a), also ist A^* ein Monomorphismus. Ist umgekehrt A^* ein Monomorphismus, so gilt $\{0\} = \ker(A^*) = \mathrm{im}(A)^\perp$ nach (a), und der Dualitätssatz 24.9 liefert $\mathrm{im}(A) = \mathrm{im}(A)^{\perp\top} = \{0\}^\top = W$, d.h. A ist ein Epimorphimus.

(d): Sei A ein Monomorphismus. Dann ist $\{0\} = \ker(A) = \mathrm{im}(A^*)^\top$ nach (b), und Dualität 24.9 zeigt $\mathrm{im}(A^*) = \mathrm{im}(A^*)^{\top\perp} = \{0\}^\perp = V^*$, d.h. A^* ist ein Epimorphismus. Ist A^* ein Epimorphismus, so folgt direkt nach (b) $\ker(A) = \mathrm{im}(A^*)^\top = (V^*)^\top = \{0\}$ und A ist ein Monomorphismus.

(e): Folgt aus (c) und (d). □

12. Polynome und ihre Nullstellen

Wir werden im folgenden lineare Abbildungen genauer studieren, insbesondere die Frage, wann ein Endomorphismus diagionalisierbar ist, d.h. unter welchen Bedingungen es eine Basis gibt, so dass die Matrix eines Endomorphismus bzgl. dieser Basis eine Diagonalmatrix ist. Ist eine lineare Abbildung diagonalisierbar, so sind die in der Diagonalen auftretenden Skalare die sogenannten Eigenwerte der Abbildung; diese Eigenwerte sind die Nullstellen eines Polynoms. Wir betrachten daher zunächst elementare Eigenschaften von Polynomringen.

Definition 12.1. Seien R, S Ringe (vgl. Definition 10.1).

(1) Eine Abbildung $f : R \to S$ ist ein Ringhomomoprhimus, falls
   (a) $f(r_1 + r_2) = f(r_1) + f(r_2),\ \ r_1, r_2 \in R$,
   (b) $f(r_1 r_2) = f(r_1) f(r_2),\ \ r_1, r_2 \in R$,
   (c) $f(1_R) = 1_S$.

(2) Ein Monomorphismus (bzw. Epimorphismus, Isomorphismus) ist ein injektiver (bzw. surjektiver, bijektiver) Ringhomomorphimus.

In der Vorlesung sind wir bei der Definition des Polynomrings rein intuitiv vorgegangen. Hier nun eine formale Definition des Polynomrings.

Definition 12.2. Sei R ein Ring. Der Polynomring $R[x]$ über R ist

$$R[x] = \{(a_0, a_1, \dots) \mid a_j \in R, \text{ nur endlich viele } a_j \neq 0\},$$

mit Addition und Multiplikation definiert durch

$$(a_j) + (b_j) = (a_j + b_j) \text{ und } (a_j)(b_j) = (c_j) \text{ mit } c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Lemma 12.3. *Sei R ein Ring mit Einselement 1.*

(a) *$R[x]$ ist ein Ring mit Einselement $1 = (1, 0, 0, \dots)$; der Ring $R[x]$ ist genau dann kommutativ, wenn R kommutativ ist.*

(b) *Die Abbildung $R \to R[x],\ a \mapsto (a, 0, 0, \dots)$ ist ein Monomorphismus von Ringen.*

(c) *Ist K ein Körper, so ist $K[x]$ eine kommutative K-Algebra. Ist $x = (0, 1, 0, \dots)$, so ist $\{x^j \mid j = 0, 1, 2, \dots\}$ eine K-Basis von $K[x]$.*

NB. Für x wie in (c) folgt $x^j = (0, \dots, 1, 0, \dots)$ (1 an der Stelle j). Da die x^j eine Basis bilden, lässt sich jedes Element von $K[x]$ als

$$\sum_{j=0}^{n} a_j x^j = (a_0, a_1, a_2, \dots, a_n, 0, \dots)$$

schreiben, d.h. als ein Polynom in x mit Koeffizienten in K. Dabei ist

$$(\sum_{i=0}^{m} a_i x^i)(\sum_{j=0}^{n} b_j x^j) = \sum_{k=0}^{n+m} (\sum_{i+j=k} a_i b_j) x^k,$$

d.h. die Multiplikation ist die übliche Multiplikation von Polynomen.

Beweis. Einfaches Nachrechnen. □

Definition 12.4. Sei K ein Körper. Ist $0 \neq f(x) = \sum_{j=0}^{n} a_j x^j \in K[x]$ mit $a_n \neq 0$, so ist der Grad von f definiert als $\operatorname{Grad} f = n$. Ist $f = 0$, so setze $\operatorname{Grad} f = -\infty$; ist $\operatorname{Grad} f = n$ und $a_n = 1$, so ist f normiert. Oftmals schreiben wir auch $\deg(f)$ anstelle von $\operatorname{Grad}(f)$.

- $\operatorname{Grad}(f + g) \leq max\{\operatorname{Grad}(f), \operatorname{Grad}(g)\}$ (wobei $max\{-\infty, n\} = n$),
- $\operatorname{Grad}(fg) = \operatorname{Grad}(f) + \operatorname{Grad}(g)$ (wobei $-\infty = -\infty + n$).

Definition 12.5. Sei R ein Ring. Dann ist R nullteilerfrei, falls für alle $a, b \in R$ gilt:

$$ab = 0 \implies a = 0 \text{ oder } b = 0.$$

Einen nullteilerfreien kommutativen Ring nennt man auch einen Integritätsbereich oder Integritätsring.

Bemerkung 12.6. Falls R ein nullteilerfreier Ring ist, so ist auch $R[x]$ nullteilerfrei.

Proposition 12.7. *(Division mit Rest) Seien $f, g \in K[x]$ mit $g \neq 0$. Dann gibt es eindeutig bestimmte $h, r \in K[x]$, so dass gilt*

$$f = gh + r \text{ mit } \operatorname{Grad}(r) < \operatorname{Grad}(g).$$

Beweis. Ist $\operatorname{Grad}(f) < \operatorname{Grad}(g)$, so setze $h = 0$ und $r = f$. Sei $f = \sum_{j=0}^{m} a_j x^j$ und $g = \sum_{k=0}^{n} b_k x^k$ mit $m \geq n$ und $a_m b_n \neq 0$. Betrachte das Polynom

$$f_1 = f - \frac{a_m}{b_n} x^{m-n} g = c_0 + c_1 x + \cdots c_{m-1} x^{m-1}$$

mit $\mathrm{Grad}(f_1) \leq m-1$. Mittels Induktion nach m gilt $f_1 = h_1 g + r_1$ mit $h_1, r_1 \in K[x]$ und $\mathrm{Grad}(r_1) < n$. Also ist

$$f = g(\frac{a_m}{b_n} x^{m-n} + h_1) + r_1$$

mit $\mathrm{Grad}(r_1) < \mathrm{Grad}(g)$. Es bleibt zu zeigen, dass h und r eindeutig bestimmt sind. Ist $f = h_1 g + r_1$ mit $h_1, r_1 \in K[x]$ und $\mathrm{Grad}(r_1) < n$ eine weitere Darstellung dieser Form, so folgt

$$r - r_1 = (h_1 - h)g.$$

Die Annahme $h \neq h_1$ impliziert den Widerspruch

$$n = \mathrm{Grad}(g) \leq \mathrm{grad}(g) + \mathrm{Grad}(h_1 - h) = \mathrm{Grad}(r - r_1) < n.$$

Also ist $h_1 = h$ und wegen $hg + r = h_1 g + r_1$ dann auch $r_1 = r$. □

Definition 12.8. Sei K ein Körper, $\mathcal{A}$ eine K-Algebra und $c \in \mathcal{A}$. Ist $f(x) = \sum_{j=0}^{n} a_j x^j \in K[x]$, so setze $f(c) = \sum_{j=0}^{n} a_j c^j \in \mathcal{A}$. Die Abbildung

$$\alpha = \alpha_c : K[x] \to \mathcal{A},\ f \mapsto f(c)$$

ist der Einsetzungshomomorphismus (bzgl. c); α_c ist ein Homomorphismus von K-Algebren, d.h. α_c ist K-linear und ein Ringhomomorphismus.

Beispiele 12.9. (a) Seien $K \subseteq L$ Körper. Dann ist L eine K-Algebra und für $f \in K[x]$ und $c \in L$ ist $f(c) \in L$ definiert.
(b) Sei V ein K-Vektorraum und $\mathcal{A} = \mathrm{End}_K(V)$. Ist $f = \sum_{j=0}^{n} a_j x^j$ ein Polynom in $K[x]$ und $\phi \in \mathrm{End}_K(V)$, so ist $f(\phi)$ der Endomorphismus

$$f(\phi) = \sum_{j=0}^{n} a_j \phi^j = a_0 id_V + a_1 \phi + a_2 \phi^2 + \cdots + a_n \phi^n.$$

(c) Ist K endlich mit $|K| = q = p^n$, so gilt $c^q = c$ für alle $c \in K$. Ist $f(x) = x^q - x \in K[x]$, so ist $f \neq 0$, aber $f(c) = 0$ für alle $c \in K$, d.h. das Polynom $f \in K[x]$ ist von der durch f induzierten Abbildung $K \to K,\ c \mapsto f(c)$, zu unterscheiden. Ist $\phi : K^2 \to K^2,\ (x_1, x_2) \mapsto (0, x_1)$ so folgt $f(\phi) = \phi^q - \phi = -\phi \neq 0$, da $\phi^2 = 0$ ist.

Lemma 12.10. *Seien $K \subseteq L$ Körper, $f \in K[x]$ und $c \in L$.*

(a) *Ist $f(c) = 0$, so ist $f = (x - c)h$ für ein eindeutig bestimmtes Polynom $h \in L[x]$.*
(b) *Ist $f \neq 0$ und $f(c) = 0$, so gibt es ein eindeutiges bestimmtes $m \in \mathbb{N}$ und ein eindeutig bestimmtes Polynom $h \in L[x]$ mit $f = (x - c)^m h$ und $h(c) \neq 0$.*

Beweis. Teil (a) folgt einfach durch Teilen von f durch $(x-c)$ mit Rest. Es ist $f = (x-c)g+r$ mit $\deg(r) < 1$. Also ist $r(x) = r_0$ ein konstantes Polynom und Einsetzen von c liefert $0 = f(c) = (c-c)g(c) + r_0 = r_0$. Die Existenzaussage in Teil (b) folgt durch sukzessives Anwenden von (a). Zur Eindeutigkeit: Es sei

$$f(x) = (x-c)^{m_1}h_1 = (x-c)^{m_1}h_2 \text{ mit } h_1(c) \neq 0, h_2(c) \neq 0.$$

oE sei $m_1 \leq m_2$. Dann folgt aus der Annahme $m_1 < m_2$

$$(x-c)^{m_1}(h_1 - (x-c)^{m_2-m_1}h_2) = 0.$$

Da $K[x]$ nullteilerfrei ist, folgt $(h_1 - (x-c)^{m_2-m_1}h_2) = 0$ und damit $h_1(c) = 0$. Widerspruch! Also ist $m_1 = m_2$ und die Nullteilerfreiheit impliziert sodann $h_1 = h_2$. □

Definition 12.11. Seien $K \subseteq L$ Körper, $f \in K[x]$ und $c \in L$.

(1) Ist $f \in K[x]$ und $f(c) = 0$, so ist c eine Nullstelle von f.
(2) Die Zahl m aus Lemma 12.10 (b) nennt mant die Vielfachheit der Nullstelle c von f.

Lemma 12.12. *Seien $K \subseteq L$ Körper und sei $0 \neq f \in K[x]$. Seien $c_1, \ldots, c_r$ die paarweise verschiedenen Nullstellen von f in L mit Vielfachheiten $m_1, \ldots, m_r$. Dann gibt es ein $g \in L[x]$, so dass gilt:*

$$f = \prod_{j=1}^{r}(x-c_j)^{m_j}g \text{ und } g(c_j) \neq 0 \text{ für } j = 1, \ldots, r.$$

Weiter ist

$$r \leq \sum_{j=1}^{r} m_j \leq \mathrm{Grad}(f).$$

Insbesondere hat f höchstens $\mathrm{Grad}(f)$ *viele verschiedene Nullstellen.*

• Sind $f, g \in K[x]$ und gibt es unendlich viele $c \in K$ mit $f(c) = g(c)$, so hat $f - g$ unendlich viele Nullstellen, also ist $f = g$.

Beweis. Nach Annahme ist $f = (x-c_1)^{m_1}h$ mit $h \in L[x]$ und $h(c_1) \neq 0$, d.h. die Behauptung gilt für $r = 1$. Da die c_j Nullstellen von f sind, folgt $h(c_j) = 0$ für $j = 2, \ldots, r$, und Induktion nach r liefert

$$h = \prod_{j=2}^{r}(x-x_j)^{s_j}g,$$

wobei $s_j \geq 1$ und $g(c_j) \neq 0$ für $j = 2, \ldots, r$ ist. Für diese j setze

$$k_j = (x-c_1)^{m_1} \prod_{i=2, i\neq j}^{r} (x-c_i)^{s_i}g.$$

Also ist $f = (x - c_j)^{s_j} k_j$ mit $k_j(c_j) \neq 0$ für $j = 2, \ldots, r$. Nach Lemma 12.10(b) ist die Vielfachheit m_j einer Nullstelle c_j von f eindeutig bestimmt, somit folgt $s_j = m_j$ für $j = 2, \ldots r$. Dies liefert die Darstellung

$$f = \prod_{j=1}^{r} (x - c_j)^{m_j} g,$$

und $\mathrm{Grad}(f) = \sum_{j=1}^{r} \mathrm{Grad}(x - c_j)^{m_j} + \mathrm{Grad}(g) \geq \sum_{j=1}^{r} m_j \geq r.$ □

Die Vielfachheit einer Nullstelle lässt sich leicht mittels der (formalen) Ableitung bestimmen:

Definition 12.13. Die 1. Ableitung von $f = \sum_{j=0}^{n} a_j x^j \in K[x]$ ist

$$f' = f^{(1)} = \sum_{j=1}^{n} j a_j x^{j-1};$$

die höheren Ableitungen $f^{(k)}$ für $k \geq 2$ sind rekursiv definiert durch

$$f^{(k)} = (f^{(k-1)})'.$$

- $(f + g)' = f' + g'$ und $(fg)' = f'g + fg'$.
- Ist $\mathrm{Grad}(f) = n$, so ist $\mathrm{Grad}(f') = n - 1$ falls $char(K) \nmid n$, und $\mathrm{Grad}(f') \leq n - 1$ falls $char(K) | n$.

Lemma 12.14. *Sei $f \in K[x]$ mit $\mathrm{Grad}(f) \geq 1$, und sei $c \in K$. Ist $char(K) = 0$ oder $char(K) > m$, so ist c eine m-fache Nullstelle von f genau dann, wenn $f(c) = f'(c) = \cdots = f^{(m-1)}(c) = 0 \neq f^{(m)}(c)$ ist.*

Beweis. Übung. □

Beispiele 12.15. (a) Sei K ein Körper und $f = x^n - a$ mit $0 \neq a \in K$. Es sei $char(K) = 0$ oder $char(K) \nmid n$. Sei L ein Erweiterungskörper von K, d.h. $K \subseteq L$, und $c \in L$ eine Nullstelle von f. Dann ist $f'(c) = nc^{n-1} \neq 0$. Also ist c eine einfache Nullstelle (d.h. mit Vielfachheit $m = 1$) und f hat in keinem Erweiterungskörper von K mehrfache Nullstellen.

(b) Sei $char(K) = p > 0$, $a \in K$ und $f = x^p - a$, d.h. $\mathrm{Grad}(f) = char(K)$. Es ist $f' = px^{p-1} = 0$. Wegen $char(K) = p$ gilt $x^p - c^p = (x - c)^p$. Ist also c eine Nullstelle von f in einem Körper L, $K \subseteq L$, so ist

$$0 = x^p - c^p = (x - c)^p,$$

d.h. c ist eine p-fache Nullstelle von f.

Definition 12.16. Seien $K \subseteq L$ Körper und sei $f \in K[x]$. Dann zerfällt f über L, falls es $a, c_1, \cdots, c_n \in L$ gibt, so dass in $L[x]$ gilt

$$f = a \prod_{j=1}^{n} (x - c_j).$$

Ein Körper K ist algebraisch abgeschlossen, falls jedes $f \in K[x]$ mit $\text{Grad}(f) \geq 1$ in K einen Nullstelle hat (also über K zerfällt).

Bemerkungen 12.17. (a) Der Fundamentalsatz der Algebra besagt, dass jedes $f \in \mathbb{C}[x]$ mit $\text{Grad}(f) \geq 1$ in $\mathbb{C}$ einen Nullstelle besitzt. Also ist $\mathbb{C}$ algebraisch abgeschlossen. Inbesondere gilt: Sind $K \subseteq \mathbb{C}$ Körper und ist $f \in K[x]$, so liegen alle Nullstellen von f in $\mathbb{C}$.

(b) Da $x^2 + 1 \in \mathbb{R}[x]$ keine reelle Nullstelle hat, ist $\mathbb{R}$ nicht algebraisch abgeschlossen.

(c) Sei K ein endlicher Körper mit $|K| = q = p^n$. Dann gilt $c^q = c$ für alle $c \in K$. Also hat $f = x^q - x + 1 \in K[x]$ keine Nullstelle in K und K ist nicht algebraisch abgeschlossen. Endliche Körper sind also nie algebraisch abgeschlossen.

(d) Ein Satz der Algebra besagt, dass es zu jedem Körper K einen algebraisch abgeschlossenen Körper L mit $K \subseteq L$ gibt.

13. Charakteristisches Polynom und Eigenwerte

Sei V ein K-Vektorraum endlicher Dimension und $f : V \to V$ ein Endomorphismus. Gibt es eine Basis $B = \{v_1, \ldots, v_n\}$ von V, so dass die Matrix $A = A_{f,B}$ von f bzgl. B Diagonalform hat (d.h. die einzigen nicht-trivialen Einträge liegen auf der Diagonalen), so werden wir f diagonalisierbar nennen. In diesem Fall gilt $f(v_i) = \alpha_{ii} v_i,\ i = 1, \ldots, n$. Die $\alpha_{ii} \in K$ sind die sogenannten Eigenwerte von f, und die v_i die entsprechenden Eigenvektoren. Die Frage nach der Diagonalisierbarkeit von f lässt sich auf die Existenz von Eigenwerten mit bestimmten Eigenschaften zurückführen.

Definition 13.1. (a) Sei V ein K-Vektorraum, $f \in End_K(V)$ ein Endomorphismus. Ein Skalar $\alpha \in K$ ist ein Eigenwert von f, falls es einen nicht-trivialen Vektor $0 \neq v \in V$ mit $f(v) = \alpha v$ gibt; in diesem Fall ist v ein Eigenvektor zum Eigenwert α. Das Spektrum $\sigma(f)$ von f ist die Menge aller Eigenwerte.

(b) Sei $A \in K^{n \times n}$ eine quadratische Matrix. Ein Skalar $\alpha \in K$ ist ein Eigenwert von f, falls es einen nicht-trivialen Vektor $0 \neq v \in K^n$ mit $Av = \alpha v$ gibt; in diesem Fall ist v ein Eigenvektor zum Eigenwert α. Das Spektrum $\sigma(A)$ von f ist die Menge aller Eigenwerte der Matrix A.

• (b) ist ein Spezialfall von (a), nämlich für $V = K^n$ und

$$f = f_A : K^n \longrightarrow K^n, \quad v \mapsto Av.$$

Lemma 13.2. *Sei V ein K-Vektorraum und $f \in End_K(V)$. Sind $v_1, \ldots, v_r \in V$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\alpha_1, \ldots, \alpha_r$, so sind die $v_1, \ldots, v_r$ linear unabhängig.*

• Ist $\dim_K V = n$, so hat f höchstens n verschiedene Eigenwerte.

• Sei $\dim_K V = n$. Hat $f \in End_K(V)$ genau n verschiedene Eigenwerte $\alpha_1, \cdots, \alpha_n$, so bilden die entsprechenden Eigenvektoren $v_1, \ldots, v_n$ eine Basis von V. Die Matrix von f bzgl. dieser Basis ist eine Diagonalmatrix mit Diagonaleinträgen $\alpha_1, \ldots, \alpha_n$.

Beweis. Induktion nach r. Ist $r = 1$, so ist der Eigenvektor v_1 zum Eigenwert α_1 linear unabhängig (da nach Definition $0 \neq v_1$ ist). Sei also $r > 1$. Seien $\beta_1, \ldots, \beta_r \in K$ mit $\sum_{i=1}^r \beta_i v_i = 0$. Dann ist

$$\sum_{i=1}^{r} \alpha_i \beta_i v_i = \sum_{i=1}^{n} \beta_i f(v_i) = f(\sum_{i=1} \beta_i v_i) = f(0) = 0.$$

Wegen $\alpha_1(\sum_{i=1}^r \beta_i v_i) = \sum_{i=1}^r \alpha_1 \beta_i v_i = 0$ folgt für die Differenz

$$\sum_{i=2}^{r} (\alpha_i - \alpha_1)\beta_i v_i = 0.$$

Nach Induktion sind die $v_2, \ldots, v_r$ linear unabhängig. Wegen $(\alpha_i - \alpha_1)\beta_i = 0$ und $\alpha_i - \alpha_1 \neq 0$ für $i = 2, \ldots, r$ folgt zunächst $\beta_2 = \cdots = \beta_r = 0$, und dann auch $\beta_1 = 0$ (da $\beta_1 v_1 = 0$ mit $0 \neq v_1$ ist). □

Definition 13.3. Sei V ein K-Vektorraum und $f \in End_K(V)$.

(1) Sei $\alpha \in \sigma(f)$. Dann ist der Eigenraum von f zu α der lineare Unterraum

$$V_f(\alpha) = V(\alpha) = Kern(\alpha id_V - f) = \{v \in V \mid f(v) = \alpha v\} \subseteq V.$$

(2) Ist $\dim_K V = n < \infty$, so nennt man f diagonalisierbar (über K), falls es es eine Basis $v_1, \ldots, v_n$ bestehend aus Eigenvektoren gibt. Falls $\alpha_1, \ldots, \alpha_n$ die zugehörigen Eigenwerte sind (im Allgemeinen natürlich nicht paarweise verschieden), so ist die Matrix von f bezüglich dieser Basis eine Diagonalmatrix; die Diagonaleinträgen sind genau die zugehörigen Eigenwerte α_i.

Beispiele 13.4. (a) Sei $K = \mathbb{R}$ oder $\mathbb{C}$ und sei $V = K^2$ mit Basis $\{v_1, v_2\}$. Definiere $f \in End_K(V)$ durch $f(v_1) = v_2$ und $f(v_2) = -v_1$. Bzgl. der gegebenen Basis hat f die Matrix

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

Sei $\alpha \in K$ ein Eigenwert von f mit Eigenvektor $0 \neq v \in V$. Wegen $f^2 = -id_V$ ist dann $-v = f^2(v) = f(\alpha v) = \alpha^2 v$. Also ist $\alpha^2 = -1$. Im Fall $K = \mathbb{R}$ existieren also keine Eigenwerte. Im Fall $K = \mathbb{C}$ ist $\sigma(f) = \{-i, i\}$. Direktes Nachrechnen zeigt, dass

$$f(v_1 + iv_2) = -i(v_1 + iv_2) \text{ und } f(v_1 - iv_2) = i(v_1 - iv_2).$$

Also bilden $\{v_1 + iv_2, v_1 - iv_2\}$ eine Basis von V zu den verschiedenen Eigenwerten $-i, i$, und A ist diagonalisierbar über $\mathbb{C}$.

(b) Sei $K = \mathbb{R}$ und $V = K^2$. Falls v_1, v_2 die Standardbasis ist, so entspricht die Abbildung f in Beispiel (a) geometrisch einer Drehung um $\pi/2$ (da f die Basisvektoren $v_1 = (1, 0)$ und $v_2 = (0, 1)$ auf $(0, 1)$ und $(-1, 0)$ abbildet). Sei allgemein $f = D(\phi)$ eine Drehung in V um den Winkel ϕ, $0 \leq \phi < 2\pi$. Ein Eigenvektor von f ist ein Vektor, der auf ein Vielfaches von sich selbst abgebildet wird. Da eine Drehung offenbar die Länge eines Vektors erhält, kann dieses Vielfaches nur ± 1 sein, und für $\phi \neq 0, \pi$ hat f keine Eigenwerte.

(c) Sei $V = K[x]$ und sei $f \in End_K(V)$ definiert durch $f : p \mapsto xp$. Angenommen es gibt ein $\alpha \in K$ sowie ein $0 \neq p \in V = K[x]$ mit $f(p) = \alpha p$. Dann liefert Gradvergleich den Widerspruch

$$1 + Grad(p) = Grad(xp) = Grad(f(p)) = Grad(\alpha p) \leq Grad(p),$$

also hat f keine Eigenwerte.

Wir zeigen, dass die Eigenwerte eines Endomorphismus genau die Nullstellen des sogenannten charakteristischen Polynoms sind.

Definition 13.5. Sei V ein n-dimensionaler K-Vektorraum, und $f \in End_K(V)$ ein Endomorphismus. Sei $A = A_{f,B} \in K^{n \times n}$ die Matrix von f bzgl. einer gewählten Basis B von V. Sei $E = E_n \in K^{n \times n}$ die Einheitsmatrix. Dann ist

$$\chi_f(x) = det(xE - A) \in K[x]$$

das charakteristische Polynom von f.

• Zur Wohldefiniertheit: Seien $A = A_{f,B}$ und $A' = A_{f,B'}$ Matrizen von f bzgl. zweier Basen B und B' von V. Dann gibt es eine invertierbare

Matrix T mit $A' = T^{-1}AT$. Wegen

$$\begin{aligned} det(xE - A') &= det(T^{-1}(xE - A)T) \\ &= det(T)^{-1}det(xE - A)det(T) \\ &= det(xE - A) \end{aligned}$$

ist das charakteristische Polynom wohldefiniert, d.h. unabhängig von der Wahl der darstellenden Matrix. Insbesondere haben ähnliche Matrizen (d.h. $A, A' \in K^{n\times n}$ mit $A' = T^{-1}AT$) dieselben Eigenwerte.

• Sei $A = (\alpha_{ij}) \in K^{n\times n}$. Dann hat das Polynom $det(xE - A)$ die Form

$$x^n - Tr(A)x^{n-1} + \cdots + (-1)^n det(A),$$

wobei $Tr(A) = \sum_{i=1}^{n} \alpha_{ii}$ die Spur von A ist. Damit gilt: Ist $\dim_K V = n$ und $f \in End_K(V)$, so ist das charakteristische Polynom $\chi_f(x)$ ein normiertes Polynom vom Grad n.

• Ähnliche Matrizen haben die gleiche Spur und die gleiche Determinante. Es gilt $Tr(f) = Tr(A)$ und $\det(f) = \det(A)$ für jede darstellende Matrix A von f.

Lemma 13.6. *Sei* $\dim_K V < \infty$ *und* $f \in End_K(V)$. *Für* $\alpha \in K$ *gilt dann:*

$$\alpha \textit{ ist Eigenwert von } f \Leftrightarrow \chi_f(\alpha) = 0.$$

Beweis. Nach Definition ist $\alpha \in K$ genau dann ein Eigenwert von f, wenn $Kern(\alpha id_V - f) \neq 0$ ist. Wegen $\dim_K V < \infty$ gilt dies genau dann, wenn $\alpha id_V - f$ kein Isomorphismus ist. Ist $A = A_{f,B}$ die darstellende Matrix von f (bzgl. einer beliebigen Basis B), so ist $\alpha id_V - f$ genau dann kein Isomorphismus, wenn $\alpha E - A$ nicht invertierbar ist. Nach Theorem 10.13(c) ist dies gleichwertig zu $det(\alpha E - A) = 0$, d.h. $\chi_f(\alpha) = 0$. □

Beispiele 13.7. (a) Sei $K = \mathbb{R}$ oder $\mathbb{C}$ und $V = K^2$ mit Basis $\{v_1, v_2\}$. Sei $f \in End_K(V)$ mit $f(v_1) = v_2$ und $f(v_2) = -v_1$ wie in Beispiel 13.4(a). Sei A die Matrix von f bzgl. $\{v_1, v_2\}$. Dann ist

$$\chi_f(x) = det(xE - A) = det \begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix} = x^2 + 1.$$

Für $K = \mathbb{R}$ hat f keine Eigenwerte; ist $K = \mathbb{C}$, so ist $\sigma(A) = \{-i, i\}$.
(b) Sei $V = \mathbb{R}^3$ und $f \in End_K(V)$ bzgl. der Standardbasis durch

$$A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$$

gegeben. Dann ist das charakteristische Polynom $\chi_f(x) = det(xE - A)$ gegeben durch

$$\chi_f(x) = det \begin{pmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{pmatrix} = x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2;$$

die Eigenwerte von f sind $1, 2$.

(c) Ist $K = \mathbb{R}$ und ist $n = \dim_K V$ ungerade, so hat $f \in End_K(V)$ Eigenwerte: Nach Annahme ist $n = Grad(\chi_f(x))$ ungerade, somit gilt

$$\lim_{x \to \infty} \chi_f(x) = \infty \text{ und } \lim_{x \to -\infty} \chi_f(x) = -\infty.$$

Nach dem Zwischenwertsatz gibt es ein $\alpha \in \mathbb{R}$ mit $\chi_f(\alpha) = 0$.

(d) Ist K ein algebraisch abgeschlossener Körper (wie zum Beispiel $K = \mathbb{C}$), so hat $\chi_f(x)$ Nullstellen, d.h. es gibt Eigenwerte.

Definition 13.8. Sei $\dim_K V = n < \infty$ und $f \in End_K(V)$. Die Vielfachheit $v(f, \alpha)$ eines Eigenwerts α von f ist die Vielfachheit von α als Nullstelle von $\chi_f(x)$, d.h. ist $\chi_f = (x - \alpha)^m g$ mit $g \in K[x]$ und $g(\alpha) \neq 0$, so ist $v(f, \alpha) = m$.

Lemma 13.9. *Sei $\dim_K V = n < \infty$ und $f \in End_K(V)$. Ist $\alpha \in \sigma(f)$ ein Eigenwert von f mit Vielfachheit $v(f, \alpha)$, so gilt*

$$1 \leq \dim_K V(\alpha) \leq v(f, \alpha).$$

Beweis. Sei $m = \dim_K V(\alpha)$. Sei $v_1, \ldots, v_m$ eine Basis des Eigenraums $V(\alpha)$. Wir ergänzen $v_1, \ldots, v_m$ zu einer Basis

$$B = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$$

von V. Bezüglich dieser Basis ist die darstellende Matrix $A = A_{f,B}$ eine Kästchenmatrix

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

mit einer $m \times m$-Diagonalmatrix A, deren Diagonaleinträge allesamt gleich α sind, und Matrizen $B \in K^{m \times (n-m)}, D \in K^{(m-n) \times (m-n)}$. Für das charakteriszische Polynom erhalten wir also $\chi_f(x) = (x - \alpha)^m h$ mit einem Polynom $h \in K[x]$. Es folgt: $m \leq v(\alpha, f)$. □

Proposition 13.10. *Sei $\dim_K V = n < \infty$ und $f \in End_K(V)$. Dann sind gleichwertig:*

(a) *f ist diagonalisierbar,*

(b) *$\chi_f(x) = \prod_{i=1}^{r} (x - \alpha_i)^{n_i} \in K[x]$ (wobei die α_i die paarweise verschiedenen Nullstellen sind), und $\dim_K V(\alpha_i) = v(f, \alpha_i) = n_i$.*

Beweis. $\Rightarrow$: Ist A diagionalisierbar, so gibt es paarweise verschiedene Eigenwerte $\alpha_1, \dots \alpha_r$ und eine Basis bestehend aus Eigenwerten

$$v_{1,1}, \dots, v_{1,n_1}, v_{2,1}, \dots, v_{2,n_2}, \dots, v_{r,1}, \dots, v_{r,n_r} \tag{1}$$

wobei $v_{i,1}, \dots, v_{i,n_i}$ jeweils linear unabhängige Vektoren in $V(\alpha_i)$ sind. Insbesondere gilt also $n_i \leq \dim V(\alpha_i)$. Sei A die Matrix von f bezüglich dieser Basis. Es folgt $\chi_f(x) = \det(xE - A) = \prod_{i=1}^{r}(x - \alpha_i)^{n_i}$, d.h. $v(\alpha_i, f) = n_i$. Lemma 13.9 impliziert nun $\dim(V(\alpha_i)) \leq n_i$, also insgesamt $\dim(V(\alpha_i)) = n_i = v(\alpha_i, f)$.

$\Leftarrow$: Nach Annahme ist $\chi_f(x) = \prod_{i=1}^{r}(x - \alpha_i)^{n_i}$ und $\dim_K V(\alpha_i) = n_i$ für $i = 1, \dots, r$. Sei für $i = 1, \dots, r$ die Menge $v_{i,1}, \dots, v_{i,n_i}$ eine Basis von $V(\alpha_i)$. Wegen

$$n = \deg(\chi_f(x)) = \sum_{i=1}^{r} n_i,$$

genügt es zu zeigen, dass die Vektoren

$$v_{1,1}, \dots, v_{1,n_1}, v_{2,1}, \dots, v_{2,n_2}, \dots, v_{r,1}, \dots, v_{r,n_r}$$

linear unabhängig sind. Sei dazu

$$\sum_{i=1}^{r} \sum_{j=1}^{n_i} \beta_{i,j} v_{i,j} = 0.$$

Setze $w_i := \sum_{j=1}^{n_i} \beta_{i,j} v_{i,j}$. Dann ist w_i entweder 0 oder ein Eigenvektor zu α_i. Da die α_i paarweise verschieden sind, impliziert Lemma 13.2, dass $w_i = 0$ für $i = 1, \dots, r$ gilt. Die lineare Unabhängigkeit der $v_{i,1}, \dots, v_{i,n_i}$ impliziert $\beta_{i,1} = \dots = \beta_{i,n_i} = 0$ für $i = 1, \dots, r$. $\square$

Der obige Beweis zeigt insbesondere auch

• Sei $\{\alpha_1, \dots, \alpha_r\}$ die Menge der Eigenwerte von f. Dann ist f genau dann diagonalisierbar, wenn $V = \bigoplus_{i=1}^{r} V(\alpha_i)$.

Beispiele 13.11. (a) In Beispiel 13.7(b) ist $\chi_f(x) = (x - 1)(x - 2)^2$. Für den Eigenwert 2 ist $\dim_K V(2) = \dim_K Kern(2E - A)$ genau die Dimension des Kerns der linearen Abbildung, die bzgl. der Standardbasis durch

$$\begin{pmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{pmatrix}$$

beschrieben ist. Durch elementare Zeilenumformungen folgt, dass diese Matrix Zeilenrang 1 hat. Also ist $\dim_K V(2) = 2$. Weiter ist $\dim_K V(1) = 1$ (da 1 ein Eigenwert ist, ist $\dim_K V(1) \geq 1$; die Umkehrung folgt aus Lemma 13.9). Nach Proposition 13.10 ist A diagonalisierbar.

(b) Sei $\dim(V) = 2$ und $f \in End_K(V)$ die lineare Abbildung, die auf

der Basis $\{v_1, v_2\}$ durch $f(v_1) = 0$ und $f(v_2) = v_1$ definiert ist. Die Matrix von f bzgl. der Basis v_1, v_2 ist

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

Weiter ist $\chi_f(x) = x^2 \in K[x]$, also ist 0 der einzige Eigenwert von f mit Vielfachheit 2. Der entsprechende Eigenraum $V(0) = Kern(0E - A) = Kern(A)$ hat offensichtlich Dimension 1, also ist wegen $\dim_K V(0) = 1 < 2 = v(0, f)$ die Abbildung f nach Proposition 13.10 *nicht* diagonalisierbar. Gleichwertig: Wäre f diagonalisierbar, so wäre A ähnlich zur Null-Matrix 0, was aber wegen $1 = r(A) \neq r(0) = 0$ nicht gelten kann.

Sei V ein endlich-dimensionaler K-Vektorraum, $\dim(V) = n < \infty$, und $f \in End_K(V)$. Dann stiftet jede Wahl einer Basis von V einen Isomorphismus $\mathrm{End}(V) \simeq K^{n \times n}$. Also ist $\dim(\mathrm{End}(V)) = n^2$ und die Endomorphismen $1, f, f^2, \ldots, f^{n^2}$ sind linear abhängig. Also gibt es Skalare $\alpha_0, \ldots, \alpha_{n^2} \in K$, nicht alle gleich 0, so dass

$$\alpha_0 + \alpha_1 f + \ldots + \alpha_{n^2} f^{n^2} = 0.$$

Es gibt also ein Polynom $h \in K[x], h \neq 0$, so dass $h(f) = 0$.

Definition 13.12. Sei V ein endlich-dimensionaler K-Vektorraum und $f \in End_K(V)$. Das Minimalpolynom μ_f von f ist das normierte Polynom kleinsten Grades mit $\mu_f(f) = 0$.

• Dies ist wohldefiniert, denn: Seien $p, q \in K[x]$ zwei normierte Polynome gleichen (minimalen) Grades mit der Eigenschaft $p(f) = 0 = q(f)$. Falls $p \neq q$, so wäre $0 \neq d := p - q$ ein Polynom kleineren Grades mit $d(f) = 0$, im Widerspruch zur Minimalität des Grades von p und q.

Theorem 13.13. *(Caley-Hamilton) Sei V ein n-dimensionaler K-Vektorraum, und sei $f \in \mathrm{End}_K(V)$ mit charakteristischem Polynom χ_f und Minimalpolynom μ_f. Dann gilt*

(a) $\chi_f(f) = 0$,
(b) *μ_f ist ein Teiler von χ_f (also ist $Grad(\mu_f) \leq Grad(\chi_f) = n$),*
(c) $\alpha \in \sigma(f) \Leftrightarrow \mu_f(\alpha) = 0$,
(d) *Hat f genau n verschiedene Eigenwerte, so ist $\mu_f = \chi_f$.*

Bemerkung 13.14. Der entscheidende Punkt im Beweis von Theorem 13.13 ist die Aussage, dass $\chi_f(f) = 0$ gilt. Sei A eine darstellende Matrix von f. Da $\mathrm{End}(V) \simeq K^{n \times n}$ ein Isomorphismus von K-Algebren ist, gilt

$$\chi_f(f) = 0 \iff \chi_f(A) = 0.$$

Der offensichtliche "Beweis"

$$\chi_f(A) = det(AE - \overline{A}) = det(0) = 0$$

ist falsch, da das Einsetzen in $\chi_f(x)$ und in $det(xE - A)$ verschiedene Dinge sind. Man beachte, dass $\chi_f(A) \in K^{n\times n}$ und $det(AE - \overline{A}) \in K$ gilt. Hier werden völlig verschiedene mathematische Objekte verglichen.

Beweis. (a): Sei $A \in K^{n\times n}$ die Matrix von f bzgl. einer Basis; wir zeigen $\chi_f(A) = 0$. Setze $B = xE - A \in K[x]^{n\times n}$, so dass $\chi_f(x) = det(xE-A) = det(B)$ ist. Es sei $\widetilde{B}$ die zu B adjunkte Matrix von 10.12, deren Einträge Unterdeterminanten von B vom Typ $(n-1, n-1)$ sind. Nach Theorem 10.13(a) ist

$$B\widetilde{B} = det(B)E = det(xE - A)E = \chi_f E,$$

wobei $\chi_f E$ die $n \times n$-Diagonalmatrix mit den Diagonaleinträgen χ_f bezeichnet. Da die Einträge in $\widetilde{B}$ Polynome in x vom Grad $\leq n - 1$ sind, hat die Matrix $\widetilde{B}$ eine Darstellung als formale Linearkombination

$$\widetilde{B} = \sum_{i=0}^{n-1} C_i x^i,\ C_i \in K^{n\times n}.$$

Sei $\chi_f(x) = \sum_{i=0}^{n} a_i x^i$ (mit $a_n = 1$). Wegen $\chi_f E = B\widetilde{B}$ gilt die Identität

$$(a_0 + a_1 x + \cdots + x^n)E = (xE - A)(C_0 + C_1 x + \cdots + C_{n-1}x^{n-1}),$$

die wir als eine Gleicheit von zwei Polynomen mit Koeffizienten in $K^{n\times n}$ interpretieren. Koeffizientenvergleich ergibt die Matrizenrelationen

$$\begin{array}{rcl} a_0 E &=& -AC_0 \\ a_1 E &=& C_0 - AC_1 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n-1}E &=& C_{n-2} - AC_{n-1} \\ E &=& C_{n-1} \end{array}$$

Es ist

$$\chi_f(A) = \sum_{i=0}^{n} a_i A^i = \sum_{i=0}^{n} A^i(a_i E),$$

und Ersetzen der $\alpha_i E$ durch die obigen Ausdrücke liefert

$$\begin{array}{rl} \chi_f(A) & = -AC_0 + A(C_0 - AC_1) + \ldots + A^{n-1}(C_{n-2} - AC_{n-1}) + A^n C_{n-1} \\ & = 0. \end{array}$$

(b): Teile χ_f durch μ_f mit Rest, also $\chi_f = q\mu_f + r$ mit $q, r \in K[x]$ und $\deg(r) < \deg(\mu_f)$. Wegen a) folgt $0 = \chi_f(A) = q(A)\mu_f(A) + r(A) = r(A)$. Aus der Minimalität des Grades von μ_f folgt $r = 0$.

(c): Sei $\alpha \in \sigma(f)$ und sei $\mu_f(x) = b_0 + b_1 x + \cdots + x^k,\ b_i \in K$. Ist $0 \neq v \in V$ ein Eigenvektor zum Eigenwert α, so gilt (wegen $\mu_f(A) = 0$)

$$0 = \mu_f(A)v = (\sum_{i=0}^{n} b_i A^i)v = \sum_{i=0}^{k} b_i(\alpha^i v) = (\sum_{i=0}^{k} b_i \alpha^i)v = \mu_f(\alpha)v.$$

Da $0 \neq v$ ist, folgt $\mu_f(\alpha) = 0$.

(d): Nach (b) ist $Grad(\mu_f) \leq \mathrm{Grad}(\chi_f) = n$. Hat f genau n verschiedene Eigenwerte, so folgt aus (c) $Grad(\mu_f) = n$, und somit $\mu_f = \chi_f$. □

Beispiele 13.15. (a) Sei $V = \mathbb{R}^2$ und $f \in End_{\mathbb{R}}(V)$ bzgl. der Standardbasen durch

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

gegeben. Eine direkte Rechnung zeigt

$$\chi_f(x) = det \begin{pmatrix} x-2 & -1 \\ -1 & x-2 \end{pmatrix} = (x-2)^2 - 1 = x^2 - 4x + 3 = (x-1)(x-3).$$

Wegen $\mu_f | \chi_f$ ist das Minimalpolynom χ_f von f eines der Polynome

$$x - 1,\ x - 3,\ \text{oder } x^2 - 4x + 3.$$

Da nach Theorem 13.13(c) sowohl 1 als auch 3 Nullstellen von μ_f sind, ist $\mu_f = \chi_f$.

Alternativ kann man natürlich direkt Theroem 13.13 (d) anwenden.

(b) Sei $V = \mathbb{R}^3$ und sei $f \in \mathrm{End}_{\mathbb{R}}(V)$ bzgl. der Standardbasen durch

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

gegeben. Dann ist das charakteristische Polynom $\chi_f(x) = x(x+3)^2$. Wegen

$$A^2 = \begin{pmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{pmatrix} = -3A$$

ist $p(x) = x^2 + 3x = x(x+3)$ in ein Polynom mit $p(A) = 0$. Teilen mit Rest wie im Beweis von Teil (b) zeigt $\mu_f | p$. Da jeder Eigenwert von f eine Nullstelle von μ_f ist, kann μ_f kein Polynom vom Grad 1 sein; damit ist $\mu_f(x) = x(x+3) \neq x(x+3)^2 = \chi_f(x)$.

Erinnerung: Sei V ein K-Vektorraum und seien $V_1, \ldots, V_r \subseteq V$ Unterräume. Dann nennen wir

$$V_1 + \ldots + V_r := \{v_1 + \ldots + v_r \mid v_i \in V_i\}$$

die Summe der V_i. Falls jedes $v \in V_1 + \ldots + V_r$ eine eindeutige Dartellung in dieser Form hat, so nennt man die Summe direkt und schreibt $V_1 \oplus \ldots \oplus V_r$. Falls $V = V_1 + \ldots + V_r$, so sagt man, V ist die Summe der V_i. Falls $V = V_1 \oplus \ldots \oplus V_r$, so sagt man, V ist die direkte Summe der V_i.

Es gilt:

$$V = V_1 \oplus \ldots \oplus V_r \iff V_i \cap \left(\sum_{j=1, j\neq i}^{r} V_i \right) = \{0\} \text{ für alle } i = 1, \ldots, r.$$

Lemma 13.16. *Sei V ein endlich-dimensionaler K-Vektorraum und sei $f \in End_K(V)$. Ist $\mu_f = g_1 \cdots g_r$ mit paarweise teilerfremden Faktoren $g_i \in K[x]$, $i = 1, \ldots, r$, so ist*

$$V = \bigoplus_{i=1}^{r} \ker(g_i(f)).$$

Beweis. Sei $h_i = \prod_{j\neq i} g_j$. Nach Annahme ist $ggT(h_1, \ldots, h_r) = 1$ in $K[x]$, und nach Lemma 15.9 gibt es $k_i \in K[x]$, $i = 1, \ldots, r$ mit

$$1 = \sum_{i=1}^{r} k_i h_i.$$

Ist $v \in V$, so folgt $v = \mathrm{id}_V v = \sum_{i=1}^{r} k_i(f)h_i(f)v$. Weiter ist

$$g_i(f)[k_i(f)h_i(f)]v = k_i(f)[g_i(f)h_i(f)]v = k_i(f)[\mu_f(f)]v = 0,$$

d.h. $h_i(f)k_i(f)v \in \ker(g_i(f))$, und somit

$$v = \sum_{i=1}^{r} k_i(f)h_i(f)v \in \sum_{i=1}^{r} \ker(g_i(f)).$$

Dies zeigt $V \subseteq \sum_{i=1}^{r} \ker(g_i(f))$; es folgt $V = \sum_{i=1}^{r} \ker(g_i(f))$.

Sei $U = \ker(g_i(f)) \cap \sum_{j\neq i} \ker(g_j(f))$. Zu zeigen ist: $U = 0$. Sei dazu $u \in U$. Dann gilt $g_i(f)u = 0 = h_i(u)u$. Da die Polynome g_i und h_i teilerfremd sind, gibt es nach Lemma 15.9 $s_i, t_i \in K[x]$ mit $1 = t_i g_i + s_i h_i$. Es folgt

$$0 = (t_i(f)g_i(f))u + (s_i(f)h_i(f))u = \mathrm{id}_V u = u.$$

□

Bemerkung 13.17. Die im Beweis benutzten Tatsachen über Teilerfremdheit und größte gemeinsame Teiler werden im nächsten Kapitel bewiesen.

Theorem 13.18. *Sei V ein n-dimensionaler K-Vektorraum und sei $f \in End_K(V)$ ein Endomorphismus. Dann sind gleichwertig:*

(a) *f ist diagonalisierbar,*

(b) *$\mu_f(x) = \prod_{i=1}^{r}(x - \alpha_j)$ mit paarweise verschiedenen $\alpha_i \in K$ (diese α_i sind offenbar genau die Eigenwerte von f).*

Beweis. $\Rightarrow$: Ist f diagonalisierbar, so gibt es eine Zerlegung

$$V = \bigoplus_{i=1}^{r} V_f(\alpha_i) = \bigoplus_{i=1}^{r} \ker(\alpha_i \,\mathrm{id}_V - f),$$

wobei die $V_f(\alpha_i)$ genau die Eigenräume zu den paarweise verschiedenen Eigenwerten sind. Also ist $\prod_{i=1}^{r}(\alpha_i \,\mathrm{id}_V - f)V = 0$, so dass

$$\mu_f \mid \prod_{i=1}^{r}(x - \alpha_i)$$

gilt. Nach Theorem 13.13(c) ist jeder Eigenwert von f eine Nullstelle von μ_f, also ist $\mu_f(x) = \prod_{i=1}^{r}(x - \alpha_i)$.
$\Leftarrow$: Sei $\mu_f(x) = \prod_{i=1}^{r}(x - \alpha_i)$ mit paarweise verschiedenen α_i. Setze $g_i(x) := x - \alpha_i$, so dass $\ker(g_i(f)) = \ker(f - \alpha_i \,\mathrm{id}_V) = V_f(\alpha_i)$ ist. Nach Annahme sind die g_i teilerfremd, so dass mit Lemma 13.16 folgt

$$V = \bigoplus_{i=1}^{r} V_f(\alpha_i) = \bigoplus_{i=1}^{r} \ker(\alpha_i \,\mathrm{id}_V - f).$$

□

Beispiele 13.19. (a) Die Endomorphismen f in den Beispielen 13.15(a) und (b) haben Minimalpolynom $\mu_f(x) = (x-1)(x-3)$ bzw. $\mu_f(x) = x(x+3)$. Nach Theorem 13.18 sind diese f diagonalisierbar.
(b) Sei f der Endomorphismus von Beispiel 13.11(b). Wegen $\chi_f(x) = x^2$ ist das Minimalpolynom μ_f eines der Polynome x oder x^2. Da $f \neq 0$ ist, folgt $\mu_f(x) = x^2$, also ist f nicht diagonalisierbar.

14. Ringe und Ideale

Wir betrachten Ideale in Ringen. Ideale sind Kerne von Ringhomomorphismen, und daher das Analogon von Normalteilern in Gruppen.

Wir erinnern an die Definition des Kerns und des Bildes eines Ringhomomorphismus.

Definition 14.1. Seien R, S Ring. Ist $f : R \to S$ ein Ringhomomorphismus, so sind Kern und Bild von f definiert als

$$\ker(f) = \{r \in R \mid f(r) = 0\} \text{ und } \mathrm{im}(f) = \{f(r) \mid r \in R\}.$$

• Ist $f : R \to S$ ein Ringhomomorphismus, so gilt trivialerweise

$$x_1, x_2 \in \ker(f) \Rightarrow x_1 + x_2 \in \ker(f).$$

Ist $x \in \ker(f)$ und sind $r_1, r_2 \in R$, so folgt wegen

$$f(r_1 x r_2) = f(r_1)f(x)f(r_2) = f(r_1) \cdot 0 \cdot f(r_2) = 0$$

weiter $r_1 x r_2 \in \ker(f)$. Dies führt uns zu dem Begriff eines Ideals.

Definition 14.2. Sei R ein Ring. Eine Teilmenge $I \subseteq R$ ist ein (2-seitiges) Ideal, falls I eine Untergruppe der additiven Gruppe von R ist, und weiter gilt: $x \in I, r_1, r_2 \in R \Rightarrow r_1 x r_2 \in I$.

• Der Kern eines Ringhomomorphismus ist also stets ein Ideal.

• Ist $I \subseteq A$ ein Ideal mit $1 \in I$, so ist $I = R$. Allgemeiner: Falls I ein invertierbares Element u enthält, so ist $I = R$.

• Ist R kommutativ, so ist $r_1 x r_2 = r_1 r_2 x = x r_1 r_2$ und eine additive Untergruppe $I \subseteq R$ ist genau dann ein Ideal, wenn gilt

$$x \in I, r \in R \Rightarrow rx \in I \quad \text{(oder gleichwertig } xr \in I\text{)}.$$

• Ist R ein Ring und sind $I_j \subseteq R,\ j = 1, 2$, Ideale, so sind auch

$$\begin{aligned} I_1 \cap I_2 &= \{x \mid x \in I_1 \text{ und } x \in I_2\}, \\ I_1 + I_2 &= \{x_1 + x_2 \mid x_1 \in I_1,\ x_2 \in I_2\}, \\ I_1 I_2 &= \{\textstyle\sum_{i=1}^{k} x_i x_i' \mid x_i =\in I_1,\ x_i' \in I_2,\ k \in \mathbb{N}\} \end{aligned}$$

Ideale in R.

• Die Bildung von Summen und Produkten von Idealen ist assoziativ und distributiv, zum Beispiel ist $I_1(I_2 + I_3) = I_1 I_2 + I_1 I_3$. Weiter gilt

$$I_i + I_i = I_i,\ RI_i = I_i = I_i R, \text{ und } I_1 I_2 \subseteq I_1 \cap I_2.$$

Beispiele 14.3. (a) Sei $R = \mathbb{Z}$. Die Mengen $m\mathbb{Z} = \{mx \mid x \in \mathbb{Z}\} \subseteq \mathbb{Z}$ für $m \in \mathbb{Z}$ sind Ideale in $\mathbb{Z}$. Weiter ist jede additive Untergruppe $H \subseteq \mathbb{Z}$ von dieser Form: Ist $H = \{0\}$, so ist $H = 0\mathbb{Z}$. Ist $H \neq 0$, also enthält H positive Elemente (ist $0 \neq x \in H$, so ist auch $-x \in H$). Sei m das kleinste positive Element in H. Dann ist $m\mathbb{Z} \subseteq H$, da $H \subseteq \mathbb{Z}$ eine additive Untergruppe ist. Ist umgekehrt $a \in H$, so liefert Division mit Rest $a = qm + r$ mit $0 \leq r < m$. Somit ist $r = a - qm \in H$, und Minimalität von m liefert $r = 0$, also $a = qm \in m\mathbb{Z}$. Da jedes Ideal von $\mathbb{Z}$ eine additive Untergruppe ist, hat jedes Ideal $I \subseteq \mathbb{Z}$ die Form $I = m\mathbb{Z}$ für ein geeignetes $m \in \mathbb{Z}$ (d.h. $I = m\mathbb{Z}$ wird von einem einzigen Element m erzeugt).

(b) Sei $R = K[x]$ der Polynomring über einem Körper. Genauso wie in (a) folgt mittels Division mit Rest (vgl. 12.7), dass jedes Ideal $I \subseteq K[x]$ die Form $fK[x] = \{fg \mid g \in K[x]\}$ hat.

(c) Sei $R = \mathbb{Z}[x]$ der Polynomring über $\mathbb{Z}$, $(x) = xR \subseteq R$ das von x erzeugte Ideal (Polynome mit konstantem Term 0), und $(2) = 2R \subseteq R$ das von 2 erzeugte Ideal (Polynome mit geraden Koeffizienten). Sei

$$(2, x) = \{\sum_{j=0}^{n} a_j x^j \mid a_0 \text{ ist gerade}, n \in \mathbb{N}\} \subseteq R.$$

Dann definiert $(2, x)$ ein Ideal, das nicht aus den Vielfachen eines einzigen Elements besteht, siehe Übung 1.

Lemma 14.4. *Sei R ein Ring (mit $1 = 1_R$ und $0 = 0_R$), und sei $I \subseteq R$ ein Ideal.*

(a) *Die Menge $R/I = \{r + I \mid r \in R\}$ ist mittels der Operationen*

$$\begin{aligned}(r_1 + I) + (r_2 + I) &= r_1 + r_2 + I,\\ (r_1 + I) \cdot (r_2 + I) &= r_1 r_2 + I\end{aligned}$$

ein Ring mit Einselement $1 + I$ und Nullelement $0 + I = I$.

(b) *Die Abbildung $f : R \to R/I,\ r \mapsto r + I$ ist ein Epimorphismus von Ringen mit $\ker(f) = I$.*

• Ist I ein Ideal in R, und sind $r_1, r_2 \in R$, so schreibe

$$r_1 \equiv r_2 (mod\ I) \Leftrightarrow r_1 - r_2 \in I \iff r_1 + I = r_2 + I.$$

• Nach Konstruktion gilt: R kommutativ $\Rightarrow R/I$ kommutativ.

Beweis. (a): Da $I \subseteq R$ eine additive Untergruppe ist, ist die Addition genau die Addition der additiven Faktorgruppe, und damit wohldefiniert. Also genügt es zu zeigen, dass die Mutiplikation wohldefiniert ist: Sei $r_i + I = r_i' + I$ für $i = 1, 2$. Dann ist $r_i - r_i' \in I$ und wegen

$$r_1 r_2 - r_1' r_2' = r_1(r_2 - r_2') + (r_1 - r_1')r_2' \in I,$$

folgt $r_1 r_2 + I = r_1' r_2' + I$. Da R ein Ring ist, rechnet man leicht nach, dass dies auch für R/I gilt.

(b): Trivial. □

Damit folgt analog zu Vektorräume (vgl. Theorem 5.10) und Gruppen (vgl. Theorem 9.6) der Homomorphiesatz für Ringe:

Theorem 14.5. *(Homomorphiesatz) Sei $f : R \to S$ ein Homomorphismus von Ringen. Sei $I \subseteq R$ ein Ideal mit $I \subseteq \ker(f)$. Sei $\pi : R \to R/I,\ r \mapsto r + I$, der kanonische Epimorphismus. Dann gibt es einen eindeutig bestimmten Ringhomomorphismus $\bar{f} : R/I \to S$, so dass $f = \bar{f} \circ \pi$. Es gilt:*

(1) $\ker(\bar{f}) = \ker(f)/I$.

(2) *$\bar{f}$ ist injektiv* $\iff I = \ker(f)$.

(3) $\operatorname{im}(\bar{f}) = \operatorname{im}(f)$

Insbesondere induziert jeder Ringhomomorphismus einen Isomorphismus $R/\ker(f) \cong \operatorname{im}(f)$.

Beweis. Nachrechnen zeigt, dass die offensichtliche Abbildung

$$h : R/I \to S,\ r + I \mapsto f(s)$$

wohldefiniert ist, und die gewünschten Eigenschaften hat. □

Beispiele 14.6. (a) Ist $R = \mathbb{Z}$, und ist $I \subseteq R$ ein Ideal, so ist nach Beispiel 14.3(a) $I = m\mathbb{Z}$ für ein $m \in \mathbb{Z}$. Der Faktorring R/I ist die additive Gruppe $\mathbb{Z}/m\mathbb{Z}$, zusammen mit der von der Multiplikation auf $\mathbb{Z}$ induzierten Multiplikation.

(b) Sei $R = K[x]$ und sei $0 \neq f \in K[x]$ ein Polynom mit $\operatorname{Grad}(f) = n$. Dann definiert $I = fK[x] \subseteq R$ ein Ideal. Der Quotient $R/I = K[x]/fK[x]$ ist eine K-Algebra, also insbesondere ein K-Vektorraum.

Wir zeigen: Die Menge $\{x^j + I \mid j = 0, 1, \ldots, n-1\} \subseteq R/I$ ist eine K-Basis des K-Vektorraums $R/I = K[x]/fK[x]$; insbesondere ist $\dim_K R/I = \dim_K K[x]/fK[x] = \operatorname{Grad}(f) = n$.

Ist $g \in K[x]$, so ist nach Division mit Rest $g = fh + r$ mit $\operatorname{Grad}(r) < \operatorname{Grad}(f)$. Also gilt in R/I stets $g + I = r + I$, und R/I wird von den $\{x^j + I,\ j = 0, \ldots, n-1\}$ erzeugt. Lineare Unabhängigkeit folgt, da

$$0 + I = \sum_{j=0}^{n-1} c_j(x^j + I) = \left(\sum_{j=0}^{n-1} c_j x^j\right) + I$$

impliziert, dass

$$\sum_{j=0}^{n-1} c_j x^j = f(x)h(x) \text{ für ein } h \in K[x].$$

Durch Vergleich der Grade erhält man $h = 0$ und folglich $c_0 = \cdots = c_{n-1} = 0$.

(c) Sei $K = \mathbb{Q}$ und $f(x) = x^2 - 2 \in \mathbb{Q}[x]$ in Beispiel 14.6(b). Aus (c) folgt, dass der Faktorring $\mathbb{Q}[x]/f\mathbb{Q}[x]$ ein $\mathbb{Q}$-Vektorraum der Dimension 2 mit Basis $\{1 + f\mathbb{Q}[x], x + f\mathbb{Q}[x]\}$ ist. Betrachten $\sqrt{2} \in \mathbb{C}$ und

$$\alpha = \alpha_{\sqrt{2}} : \mathbb{Q}[x] \to \mathbb{C},\ g \mapsto g(\sqrt{2})$$

Dieser Einsetzungshomomorphismus α induziert einen Isomorphismus

$$\mathbb{Q}[x]/f\mathbb{Q}[x] \cong \mathbb{Q}(\sqrt{2}),$$

d.h. der linksstehende Faktorring ist ein Körper; siehe Übungsblatt.

(d) Sei $0 \neq I \subseteq K^{n\times n}$ ein Ideal im Matrixring der $n \times n$-Matrizen

mit Einträgen in einem Körper K. Sei $\{E_{ij} \mid i,j = 1,\ldots,n\} \subseteq K^{n\times n}$ die K-Standardbasis von $K^{n\times n}$, so dass $E_{ij}E_{kl} = \delta_{jk}E_{il}$ gilt, und sei $0 \neq A = \sum_{i,j=1}^{n} \alpha_{ij}E_{ij} \in I$ ein Element mit $\alpha_{st} \neq 0$. Dann ist

$$E_{ks}AE_{tl} = \alpha_{st}E_{ks}E_{st}E_{tl} = \alpha_{st}E_{kl} \in I.$$

Da I ein K-Vektorraum ist, folgt $E_{kl} \in I$ für alle k,l, d.h. $I = K^{n\times n}$.

Proposition 14.7. *(Chinesischer Restsatz) Sei R ein kommutativer Ring und seien $I_j \subseteq R,\ j = 1,\ldots,n$ Ideale mit $I_i + I_j = R$ für $i \neq j$.*

(a) *Für $s,t \in \mathbb{N}$ und $i \neq j$ gilt $I_i^t + I_j^s = R$,*

(b) *Es ist $R = I_1 \cdots I_{n-1} + I_n$ und $I_1 \cdots I_n = I_1 \cap \cdots \cap I_n$,*

(c) *Seien $r_j \in R,\ j = 1,\ldots,n$ beliebig vorgegeben. Dann gibt es ein $r \in R$ mit $r \equiv r_j (mod\ I_j)$ für $j = 1,2,\ldots,n$.*

Beweis. Die Aussagen (a) und (b) werden auf dem Übungsblatt bewiesen. Aussage (c) beweisen wir mittels Induktion nach n. Für den Induktionsanfang sei $n = 2$. Wegen $I_1 + I_2 = R$ gibt es $y \in I_1$ und $x \in I_2$, so dass $a_1 - a_2 = x - y$. Dann ist $a := a_1 + y = a_2 + x$ eine Lösung der simultanen Kongruenz $a \equiv a_i (\mathrm{mod}\ I_i), i = 1,2$. Zum Induktionsschritt $n \longrightarrow n+1$. Nach Induktion gibt es $y \in R$ mit

$$y \equiv a_i (\mathrm{mod}\ I_i), \quad , i = 1,\ldots,n.$$

Wegen $I_1 \cdots I_n + I_{n+1} = R$ liefert der Fall $n = 2$ ein Element $a \in R$ mit

$$a \equiv y (\mathrm{mod}\ I_1 \cdots I_n), \quad a \equiv a_{n+1} (\mathrm{mod}\ I_{n+1}).$$

Wegen $I_1 \cdots I_n \subseteq I_i$ für $i = 1,\ldots,n$ folgt $y \equiv a_i (\mathrm{mod}\ I_i)$ und somit

$$a \equiv a_i (\mathrm{mod}\ I_i) \text{ für } i = 1,\ldots,n+1.$$

□

Sind $R_1,\ldots,R_n$ (kommutative) Ringe, so ist das kartesische Produkt $\prod_{i=1}^{n} R_i = R_1 \times \cdots R_n$ mittels komponentenweiser Addition und Multiplikation wieder ein (kommutativer) Ring. Aus dem Chinesischen Restsatz folgt:

Korollar 14.8. *Sei R ein kommutativer Ring, und seien $I_1,\ldots,I_n$ Ideale in R mit $I_i + I_j = R$ für $i \neq j$. Dann ist die Abbildung*

$$f : R \to \prod_{i=1}^{n} R/I_i,\ r \mapsto (r + I_1,\ldots,r + I_n)$$

ein surjektiver Ringhomomorphismus mit $Kern(f) = \cap_{i=1}^n I_i = \prod_{i=1}^n I_i$. *Insbesondere ist*

$$R/\bigcap_{i=1}^n I_i = R/\left(\prod_{i=1}^n I_i\right) \cong \prod_{i=1}^n R/I_i.$$

Beweis. Der chinesische Restsatz liefert die Surjektivität der Abbildung f. Damit ist die induzierte Abbildung $\bar{f} : R/\ker(f) \longrightarrow \prod_{i=1}^n R/I_i$ nach dem Homomorphiesatz ein Isomorphismus und die Behauptung folgt aus $\ker(f) = \cap_{i=1}^n I_i = \prod_{i=1}^n I_i$. □

Beispiel 14.9. Sei $R = \mathbb{Z}$ und seien $m_1, \dots m_n \in \mathbb{Z}$ paarweise teilerfremde ganze Zahlen. Betrachte die Ideale $I_j = m_j\mathbb{Z} \subseteq \mathbb{Z}$. Der euklidische Algorithmus liefert für $i \neq j$ ganze Zahlen $x_i, x_j \in \mathbb{Z}$ mit $x_i m_i + x_j m_j = 1$. Also gilt $I_i + I_j = \mathbb{Z}$ für $i \neq j$. Der Chinesische Restsatz zeigt, dass es für vorgegebene $r_j \in \mathbb{Z},\ j = 1, \cdots, n$ ein $r \in \mathbb{Z}$ gibt, so dass

$$r \equiv r_j (mod\ m_j\mathbb{Z}),\ j = 1, \dots, n.$$

Die Zahl r ist dabei eindeutig bestimmt modulo $m_1 \cdots m_n$. Insbesondere ist für teilerfremde $m, n \in \mathbb{Z}$ dann $mn\mathbb{Z} = m\mathbb{Z} \cap n\mathbb{Z}$ und die kanonischen Abildungen induzieren einen Ringisomorphismus

$$\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

15. Arithmetik in Integritätsbereichen

Der Ring der ganzen Zahlen $\mathbb{Z}$ hat folgende Eigenschaften:

(a) $\mathbb{Z}$ hat keine nicht-trivialen Nullteiler, d.h. aus $mn = 0,\ m, n \in \mathbb{Z}$ folgt stets $m = 0$ oder $n = 0$.

(b) Jedes Element $0 \neq n \in \mathbb{Z}$ hat eine Primfaktorenzerlegung, d.h.

$$n = ep_1^{n_1} \cdots p_r^{n_r},\ p_i \text{ Primzahl },\ e \in \{-1, 1\},\ n_i \in \mathbb{N};$$

diese Zerlegung ist eindeutig (bis auf Reihenfolge).

(c) Jedes Ideal $I \subseteq \mathbb{Z}$ hat die Form $I = m\mathbb{Z}$ für ein $m \in \mathbb{Z}$, d.h. jedes Ideal wird von einem Element erzeugt.

(d) In $\mathbb{Z}$ gibt es eine 'Division mit Rest': Sind $n, m \in \mathbb{Z}, m \neq 0$, so gibt es $q, r \in \mathbb{Z}$ mit $n = qm + r$ und $r = 0$ oder $|r| < |m|$.

Wir studieren Verallgemeinerungen dieser Begriffe für allgemeine (kommutative) Ringe. Insbesondere entwickeln wir die grundlegenden Begriffe der Arithmetik in den Ringen $\mathbb{Z}$ und $K[x]$, die wir bereits im vorherigen Kapitel (siehe Bemerkung 13.17) benutzt haben und die wir im folgenden Kapitel zu einer genaueren Untersuchung von linearen Abbildungen benötigen werden.

Definition 15.1. Sei R ein kommutativer Ring. Sind $x, y \in R$ und gilt $x = ry$ für ein $r \in R$ (d.h. y teilt x), so schreibe $y|x$.

(1) Ein Element $x \in R$ ist ein Nullteiler, falls es ein $r \in R \setminus \{0\}$ mit $rx = 0$ gibt. Der Ring R ist ein Integritätsbereich, falls R keine nicht-trivialen Nullteiler hat, d.h. aus $xy = 0$ folgt stets $x = 0$ oder $y = 0$.

(2) Ein Element $x \in R$ ist eine Einheit, falls es ein $r \in R$ mit $rx = 1$ gibt. Die Menge der Einheiten von R bildet bzgl. der Multiplikation in R eine abelsche Gruppe, die wir mit $R^\times$ bezeichnen.

(3) Ein Element $0 \neq p \in R \setminus R^\times$ ist irreduzibel, falls aus $p = xy$ mit $x, y \in R$ stets $x \in R^\times$ oder $y \in R^\times$ folgt.

Definition 15.2. Sei R ein kommutativer Ring. Für $x, y \in R \setminus \{0\}$ definieren wir

$$x \sim y :\iff x \mid y \text{ und } y \mid x.$$

Wir sagen dann, x ist assoziiert zu y.

Lemma 15.3. *Sei R ein kommutativer und nullteilerfreier Ring. Seien $x, y \in R \setminus \{0\}$. Dann gilt:*

$$x \sim y \iff \exists u \in R^\times : x = uy \iff (x) = (y).$$

Weiter gilt für $x_1, x_2, , y_1, y_2 \in R \setminus \{0\}$ mit $x_1 \sim x_2, y_1 \sim y_2$ die Äquivalenz

$$x_1 \mid y_1 \iff x_2 \mid y_2.$$

Beweis. Leicht (Übung). □

Beispiele 15.4. (a) Der Ring $R = \mathbb{Z}$ ist ein Integritätsbereich; weiter ist $R^\times = \{-1, 1\}$ und die irreduziblen Elemente von R sind genau die Elemente der Form $\pm p$, wobei p eine Primzahl ist.

(b) Sei K ein Körper, und seien $x, y \in K$, $x \neq 0$ mit $xy = 0$. Dann ist $x^{-1}xy = y = 0$, so dass alle Körper automatisch Integritätsbereiche sind. Genauso sind alle Polynomringe $R[x]$ über einem Integritätsbereich R wieder Integritätsbereiche: Seien $f, g \in R[x]$ mit $fg = 0$ gegeben. Falls $f \neq 0 \neq g$, so ist

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0, \quad a_i \in R, a_n \neq 0, \\ g(x) &= b_m x^m + b_{m-1}x^{m-1} + \ldots + b_0, \quad b_j \in R, b_m \neq 0 \end{aligned}$$

und somit

$$f(x)g(x) = a_n b_m x^{n+m} + \ldots$$

mit $a_n b_m \neq 0$, da R nullteilerfrei ist. Insbesondere ist also $fg \neq 0$.

(c) Sei R ein Integritätsbereich und $R[x]$ der entsprechende Polynomring. Auf einem der Übungsblätter haben wir gezeigt:

$$R[x]^\times = R^\times.$$

(d) Der Faktorringe $\mathbb{Z}/m\mathbb{Z}$ von $\mathbb{Z}$ für $m \geq 2$ ist genau dann ein Integritätsbereiche, wenn m eine Primzahl ist. In diesem Fall ist $\mathbb{Z}/m\mathbb{Z}$ ein Körper.

Definition 15.5. Sei R ein Integritätsbereich.

(1) Sei $0 \neq x \in R$. Eine Zerlegung von x in irreduzible Elemente (oder Primzerlegung) ist eine Darstellung von x als

$$x = ep_1 \cdots p_n,$$

wobei die p_i irreduzibel (nicht unbedingt verschieden) und e eine Einheit ist. Diese Darstellung ist eindeutig, falls aus

$$x = ep_1 \dots p_n = e'q_1 \cdots q_m,$$

stets $n = m$ und (nach eventueller Umordnung) $u_i p_i = q_i$ mit $u_i \in R^\times,\ i = 1, \dots, n$ folgt. Der Ring R ist faktoriell, falls jedes Element $0 \neq x \in R$ eine eindeutige Primzerlegung hat.

(2) Ein Ideal $I \subseteq R$ ist ein Hauptideal, falls es von einem Element erzeugt wird, d.h. $I = xR = \{xr \mid r \in R\}$. Ist jedes Ideal von R ein Hauptideal, so ist R ein Hauptidealring.

(3) Der Ring R ist ein euklidischer Ring, falls es eine Abbildung $\phi : R \setminus \{0\} \to \mathbb{N}_0$ gibt, so dass gilt: Seien $x, y \in R$ mit $y \neq 0$. Dann gibt es $q, r \in R$, so dass $x = qy + r$ mit $r = 0$ oder $\phi(r) < \phi(y)$.

Wir zeigen:

$$R \text{ euklidischer Ring} \Rightarrow R \text{ Hauptidealring} \Rightarrow R \text{ faktoriell.}$$

Lemma 15.6. *Jeder euklidische Ring ist ein Hauptidealring.*

Beweis. (Vgl. Beispiel 14.3(a),(b)) Sei $0 \neq I \subseteq R$ ein Ideal. Wähle $0 \neq y \in I$ mit $\phi(y)$ minimal. Ist $x \in I$, so ist $x = qy + r$ für geeigente $q, r \in R$ mit $r = 0$ oder $\phi(r) < \phi(y)$. Wegen $r = x - qy \in I$ folgt $r = 0$, also ist $x = qy$ und $I = yR$. □

Beispiele 15.7. (a) Der Ring $R = \mathbb{Z}$ ist ein euklidischer Ring mittels der Abbildung $\phi(m) = |m|$ (Absolutbetrag), also auch ein Hauptidealring, vgl. Beispiel 14.3(a).

(b) Sei $R = K[x]$ der Polynomring über einem Körper. Der Ring R ist ein euklidischer Ring bzgl. $\phi(f) = \mathrm{Grad}(f)$, siehe Lemma 12.7, und

damit auch ein Hauptidealring, vgl. Beispiel 14.3(b).
(c) Der Ring $R = \mathbb{Z}[x]$ ist kein Hauptidealring, siehe Beispiel 14.3(c).

Definition 15.8. Sei R ein Integritätsbereich, und $r_1, \ldots, r_n \in R$. Ein Element $0 \neq d \in R$ ist ein grösster gemeinsamer Teiler von $r_1, \ldots, r_n$, $d = ggT(r_1, \ldots, r_n)$, falls die folgenden beiden Eigenschaften gelten:

(a) $d \mid r_i$ für $i = 1, \ldots, n$.
(b) Ist $0 \neq c \in R$ ein Element mit $c \mid r_i$ für $i = 1, \ldots, n$, so gilt $c \mid d$.

• Sind d, d' grösste gemeinsame Teiler von $r_1, \ldots, r_n$, so gilt $d|d'$ und $d'|d$, d.h. es gibt $s, t \in R$ mit $d = sd'$ und $d' = td$. Also ist $d = std$ und da R ein Integritätsbereich ist folgt aus $d(1 - st) = 0,\ d \neq 0$ dann $st = 1$; somit sind $s, t \in R^\times$. Umgekehrt gilt: Ist d ein grösster gemeinsamer Teiler von $r_1, \ldots r_n$ und ist $e \in R^\times$, so ist auch ed ein grösster gemeinsamer Teiler von $r_1, \ldots, r_n$. Insbesondere ist ein grösster gemeinsamer Teiler nur bis auf Einheiten, d.h. bis auf Assoziiertheit, eindeutig bestimmt.
• Im Allgemeinen existieren in einem beliebigen Integritätsbereich keine größten gemeinsamen Teiler.

Lemma 15.9. *Sei R ein Hauptidealring, und seien $r_1, \ldots r_n \in R$. Sei $(r_1, \ldots, r_n) \subseteq R$ das von den $r_1, \ldots, r_n$ erzeugte Ideal und $d \in R$, so dass $(d) = (r_1, \ldots, r_n)$ gilt. Dann ist d ein ggT von $r_1, \ldots r_n$. Umgekehrt gilt für jeden ggT d von $r_1, \ldots r_n$ die Beziehung $(d) = (r_1, \ldots, r_n)$.*
Insbesondere exisiert $ggT(r_1, \ldots, r_n) \in R$.

Beweis. Sei $(r_1, \ldots, r_n) = dR = (d)$. Dann gilt $r_i \in (d)$ für $i = 1, \ldots, n$, also $d|r_i$. Sei c ein gemeinsamer Teiler von $r_1, \ldots, r_n$. Dann folgt $(d) = (r_1, \ldots, r_n) \subseteq (c)$ und damit $c|d$.

Sei umgekehrt $d = ggT(r_1, \ldots, r_n)$. Dann folgt aus $d|r_i$ sofort $r_i \in (d)$, und damit $(r_1, \ldots, r_n) \subseteq (d)$. Sei $(r_1, \ldots, r_n) = (c)$. Dann folgt $c|r_i$ für $i = 1, \ldots, n$ und daher $c|d$, bzw. $(d) \subseteq (c) = (r_1, \ldots, r_n)$.

□

Bemerkung 15.10. In euklidischen Ringen kann man den größten gemeinsamen Teiler algorithmisch bestimmen. Dies geschieht mit dem sogenannten euklidischen Algorithmus. Seien dazu $a, b \in R$, $b \neq 0$. oE können wir $a \neq 0 \neq b$ voraussetzen. Setze $a_0 := a, a_1 := b$. Wir teilen

sukzessive mit Rest und erhalten

$$\begin{aligned} a_0 &= q_0 a_1 + a_2, \quad \phi(a_2) < \phi(a_1), \\ a_1 &= q_1 a_2 + a_3, \quad \phi(a_3) < \phi(a_2), \\ &\vdots \\ a_{n-1} &= q_{n-1} a_n + a_{n+1}, \quad \phi(a_{n+1}) < \phi(a_n) \\ a_n &= q_n a_{n+1} + 0, \end{aligned}$$

d.h. in der n-ten Stufe tritt erstmals der Rest 0 auf. Dann zeigt man leicht, dass a_{n+1} der ggT ist. Zudem kann man durch Auflösen von unten nach oben eine Darstellung des ggT als Linearkombination von a und b berechnen. Aus der Beziehung

$$\mathrm{ggT}(a_1, \ldots, a_n) = \mathrm{ggT}(\mathrm{ggT}(a_1, \ldots, a_{n-1}), a_n)$$

erhält man ein induktives Verfahren.

Lemma 15.11. *Sei R ein Hauptidealring und $p \in R$ ein irreduzibles Element. Sind $r_1, r_2 \in R$ mit $p|r_1 r_2$, so gilt $p|r_1$ oder $p|r_2$.*

Beweis. Sei $p \nmid r_1$. Angenommen $d = ggT(p, r_1) \notin R^\times$. Sind $s, t \in R$ mit $p = sd$ und $r_1 = td$, so folgt aus $d \notin R^\times$ dann $s \in R^\times$, da p irreduzibel ist. Dann ist $r_1 = ts^{-1}p$, ein Widerspruch zu $p \nmid r_1$. Also ist $d \in R^\times$ und wegen $R = (d) = (p, r_1)$ gibt es $a, b \in R$ mit $1 = ap + br_1$. Es folgt $r_2 = pr_2 a + r_1 r_2 b$, und wegen $p|r_1 r_2$ gilt dann auch $p|r_2$. □

Lemma 15.12. *Sei R ein Hauptidealring und*

$$(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \ldots$$

eine aufsteigende Kette von Idealen. Dann wird die Kette stationär, d.h. e gibt $n \in \mathbb{N}$, so dass für alle $k \geq n$ gilt: $(a_k) = (a_n)$.

Beweis. Sei $J := \cup_n (a_n)$. Man zeigt leicht, dass J ein Ideal in R ist: seien dazu $r \in R$ und $x_1, x_2 \in J$. Dann gibt es einen Index n, so dass $x_1, x_2 \in (a_n)$. Also ist auch $rx_1, x_1 + x_2 \in (a_n) \subseteq J$. Da R ein Hauptidealring ist, gibt es ein $a \in R$ mit $J = (a)$. Sei m, so dass $a \in (a_m)$. Dann gilt für alle $n \geq m$ die Gleichheit $(a_n) = (a_m)$, denn $(a_n) \subseteq J = (a) \subseteq (a_m) \subseteq (a_n)$. □

Proposition 15.13. *Jeder Hauptidealring ist faktoriell.*

Beweis. Sei R ein Hauptidealring. Wir beweisen zunächst die Existenz einer Primzerlegung. Angenommen es gäbe ein Element $0 \neq a \in R \backslash R^\times$, das keine Primzerlegung hat. Dann ist a reduzibel. Es gelte $a = a_1 a_1'$. Wenn a_1 und a_1' eine Primzerlegung hätten, so auch das Produkt. oE habe also a_1 keine Primzerlegung. Offensichtlich gilt $(a) \subseteq (a_1)$. Per

Induktion können wir also eine nicht-stationäre Kette $(a) \subseteq (a_1) \subseteq (a_2) \dots$ konstruieren, im Widerspruch zu Lemma 15.12.

Zur Eindeutigkeit: Sei $a \in R \setminus R^\times$. Sei

$$a = up_1 \cdots p_n = vq_1 \cdots q_m \text{ mit } u, v \in R^\times, p_i, q_j \text{ irreduzibel.}$$

Wegen Bemerkung 15.11 teilt p_1 einen der Faktoren in $vq_1 \cdots q_m$. Nach Umnumerierung gelte daher $p_1 \mid q_1$. Da q_1 irreduzibel ist, folgt $p_1 = u_1 q_1$ mit $u_1 \in R^\times$ und

$$uu_1 p_2 \cdots p_n = vq_2 \cdots q_m.$$

Mit absteigender Induktion erhält man $n = m$.

□

Abschließend notieren wir die folgende Proposition, die zusammen mit Lemma 15.9 die in Bemerkung 13.17 erwähnten Lücken schließt.

Proposition 15.14. *Seien R ein faktorieller Ring und $a_1, \dots, a_n \in R$. Sei*

$$a_i = u_i \prod_{j=1}^{r} p_j^{e_{ij}}, \quad e_{ij} \in \mathbb{N}_0,$$

mit paarweise verschiedenen irreduziblen Elementen p_j die Primzerlegung von $a_i, i = 1, \dots, n$. Dann gilt:

$$ggT(a_1, \dots, a_n) = \prod_{j=1}^{r} p_j^{min\{e_{ij} | i=1,\dots,n\}}.$$

Beweis. Siehe Übungsblatt. □

16. Moduln über Hauptidealringen

Ein K-Vektorraum ist eine abelsche Gruppe V, auf der zusätzlich eine Skalarmultiplikation mit Elementen aus dem Körper K definiert ist, so dass gewisse Verträglichkeitsbedingungen gelten. Eine abelsche Gruppe M, zusammen mit einer solchen Skalarmultiplikation durch Elemente aus einem Ring R (anstelle eines Körpers) ist ein R-Modul. Wir untersuchen elementare Eigenschaften solcher R-Moduln, und bestimmen die Struktur eines endlich erzeugten Moduls über einem Hauptidealring. Dieser Struktursatz erlaubt es uns, eine weitere kanonische Darstellung eines Endomorphismus als Matrix zu bestimmen, die sogenannte Jordan-Normalform.

Definition 16.1. Sei R ein Ring mit 1 (nicht unbedingt kommutativ). Eine Menge M ist ein R-(Links-)Modul, falls es eine Verknüpfung $+ : M \times M \to M,\ (m, m') \mapsto m + m'$ und eine Skalarmultiplikation $\cdot : R \times M \to M,\ (r, m) \mapsto rm$ gibt, so dass gilt:

(1) M ist bzgl. $+$ eine abelsche Gruppe mit neutralem Element 0,
(2) $(r_1 + r_2)m = r_1m + r_2m$ und $r(m_1 + m_2) = rm_1 + rm_2$,
(3) $(r_1r_2)m = r_1(r_2m)$,
(4) $1m = m$.

Beispiele 16.2. (a) Ist $R = K$ ein Körper, so sind die K-Moduln genau die K-Vektorräume; vgl. Definition 4.1.

(b) Seien $v_1, v_2 \in \mathbb{R}^2$ zwei linear unabhängige Vektoren. Dann ist $M := \{av_1 + bv_2 \mid a, b \in \mathbb{Z}\}$ ein $\mathbb{Z}$-Modul. Anschaulich ist M das von den Vektoren v_1, v_2 aufgespannte Gitter.

(c) Sei $R = \mathbb{Z}$. Ist M bzgl. $+$ eine abelsche Gruppe, so ist M ein $\mathbb{Z}$-Modul: Sei $n \in \mathbb{N}_0$ und $m \in M$, so definiert $n \cdot m := m + \ldots + m$ (n-fache Summe von m), und $(-n) \cdot m = n \cdot (-m)$ eine $\mathbb{Z}$-Skalarmultiplikation auf M. $\mathbb{Z}$-Moduln sind genau die abelschen Gruppen.

(d) Jeder Ring R ist ein R-Modul bzgl. der auf R definierten Multiplikation. Allgemeiner: Die R-Moduln M mit $M \subseteq R$ sind genau die Ideale von R.

(e) Sei V ein K-Vektorraum, und sei $A \in End_K(V)$ fest gewählt. Dann ist der K-Vektorraum V ein $K[x]$-Modul (bzgl. A) mittels

$$f \cdot v = f(A)v, \ f \in K[x].$$

Nach dem Satz von Cayley-Hamilton gilt für jedes $v \in V$, dass $\chi_A \cdot v = 0$.

Definition 16.3. Sei R ein Ring.

(1) Sei M ein R-Modul. Eine Menge $N \subseteq M$ ist ein R-Untermodul, falls $N \leq M$ eine abelsche Untergruppe ist, die bzgl. $\cdot$ abgeschlossen ist, d.h. für $r \in R$ und $n \in N$ gilt $rn \in N$.

(2) Sei M ein R-Modul und $N \subseteq M$ ein R-Untermodul. Die Faktorgruppe $M/N = \{m+N \mid m \in M\}$ mit der üblichen Addition $(m_1 + N) + (m_2 + N) = (m_1 + m_2) + N$ ist ein R-Modul mittels

$$r(m + N) = rm + N;$$

der R-Modul M/N ist der Faktor- oder Quotientenmodul.

(3) Seien M, N R-Moduln. Eine Abbildung $f : M \to N$ ist ein R-Modulhomomorphismus (oder R-Homomorphismus), falls gilt:
(a) $f(m + m') = f(m) + f(m')$,
(b) $f(rm) = rf(m)$.
Die Menge aller R-Homomorphismen $Hom_R(M, N)$ ist eine abelsche Gruppe mittels $(f_1 + f_2)(m) := f_1(m) + f_2(m)$. Ist R ein kommutativer Ring, so ist $Hom_R(M, N)$ ein R-Modul bzgl.

$$(rf)(m) := rf(m).$$

Ist $f \in Hom_R(M,N)$, so sind Kern und Bild von f, d.h.

$$\begin{aligned} \ker(f) &= \{m \in M \mid f(m) = 0\}, \\ \mathrm{im}(f) &= \{f(m) \mid m \in M\} \end{aligned}$$

wieder R-Moduln. Die Abbildung f ist ein Monomorphismus, falls $\ker(f) = \{0\}$, und ein Epimorphismus falls $Bild(f) = N$ ist. Gilt beides, so ist f ein Isomorphismus. Es gilt der Homomorphiesatz: Für einen R-Homomorphismus $f : M \to N$ ist

$$M/\ker(f) \cong Bild(f).$$

(4) Sei M ein R-Modul. Ist $N \subseteq M$ eine Teilmenge, so der von N erzeugte R-Modul $\langle N \rangle$ definiert durch $\{\sum_{j=1}^{r} r_j n_j \mid r_j \in R,\ n_j \in N\}$; dies ist der kleinste R-Untermodul von M, der die Menge N enthält. Gilt $M = \langle m_1, \ldots, m_n \rangle$, so nennt man M endlich erzeugt.

(5) Sei M ein R-Modul und seien $N_i \subseteq M,\ i \in I$, R-Untermoduln. Dann ist $\sum_{i\in I} N_i = \{\sum_{i\in I} n_i \mid n_i \in N_i,$ höchstens endlich viele $n_i \neq 0\}$ die Summe der N_i; dies ist ein R-Untermodul. Diese Summe ist eine direkte Summe, $\oplus_{i\in I} N_i$, falls jedes Element eine eindeutige Darstellung als eine endliche Summe $\sum_{i=1}^{k} n_i, n_i \in N_i$ hat.

(6) Sind $N_1, \ldots, N_n$ R-Moduln (die nicht unbedingt R-Untermoduln eines R-Moduls M sind), so ist die direkte Summe der N_i

$$\bigoplus_{i=1}^{n} N_i = N_1 \times \cdots \times N_n,$$

mengentheoretisch das kartesische Produkt; $\bigoplus_i N_i$ ist ein R-Modul mittels komponentenweiser Addition und Skalarmultiplikation. Der Spezialfall $N_i = R$ für $i = 1, \ldots, n$ liefert den freien R-Modul $R^n = \oplus_{i=1}^{n} R = R \times \cdots \times R$ (n-faches Produkt).

Für K-Vektorräume gilt: Jeder endlich erzeugte K-Vektorraum V ist isomorph zu K^n für ein geeignetes n, d.h. $V \cong K^n$. Weiter gilt für endlich-erzeugte (und damit endlich-dimensionale) K-Vektorräume: $V \cong W \Rightarrow V \cong K^n \cong W$ für ein geeignetes n, d.h. die Dimension eines endlich erzeugten K-Vektorraums bestimmt den Isomorphietyp.

Das Analogon zu einem K-Vektorraum der Form K^n für R-Moduln ist ein freier R-Modul $R^n = \oplus_{i=1}^{n} R$. Schreibt man die Elemente als n-Tupel $(r_1, \ldots, r_n),\ r_i \in R$, so wird R^n von den n 'Basisvektoren' $e_i = (0, \ldots, 1, \ldots, 0)$ mit 1 an der Stelle i erzeugt, und jedes Element $x =$

$(r_1, \ldots, r_n)$ hat eine eindeutige Darstellung als eine Linearkombination

$$x = \sum_{i=1}^{n} r_i e_i,$$

d.h. die $e_1, \ldots, e_n$ bilden eine R-Basis. Das Analogon zu den obigen Aussagen für K-Vektorräume für allgemeine R-Moduln wäre somit: Jeder endlich erzeugte R-Modul ist isomorph zu R^n für ein n, und dieses n bestimmt den Isomorphietyp. Die folgenden Beispiele zeigen, dass diese Aussagen für R-Moduln im Allgemeinen nicht gelten:

Beispiele 16.4. (a) Ist $M = \mathbb{Z}/3\mathbb{Z}$, so ist M nach 16.2(b) ein $\mathbb{Z}$-Modul. Dabei ist M trivialerweise endlich erzeugt, aber nicht isomorph zu einer direkten Summe der Form $\mathbb{Z} \oplus \mathbb{Z} \cdots \oplus \mathbb{Z}$.

(b) Es gibt (nicht-kommutative) Ringe mit der Eigenschaft, dass $R^m \cong R^n$ nicht unbedingt $n = m$ impliziert (vgl. Übung); inbesondere macht der evidente Begriff der Dimension für einen allgemeinen (freien) R-Modul keinen Sinn.

Wir zeigen im folgenden, dass für kommutative Ringe die zweite Pathologie nicht auftritt, d.h. für einen kommutativen Ring R folgt aus $R^n \cong R^m$ stets $n = m$. Wir benötigen dazu ein weiteres Resultat aus der Ringtheorie; der Beweis verwendet das mengentheoretische 'Zornsche Lemma' (siehe auch Bemerkung nach Definition 4.17), das ein Kriterium für die Existenz von 'maximalen Elementen' liefert. Genauer: Sei $\emptyset \neq S$ eine Menge, auf der eine Ordnungsrelation $\leq$ definiert ist, d.h. für s, s', s'' in S gelten

(i) $s \leq s$,
(ii) $s \leq s'$ und $s' \leq s \Rightarrow s = s'$,
(iii) $s \leq s'$ und $s' \leq s'' \Rightarrow s \leq s''$.

Beispiele 16.5. (a) Ist M eine Menge und $P(M)$ die Potenzmenge von M, so definiert die Inklusion $\subseteq$ eine solche Ordnungsrelation auf $P(M)$.

(b) Die Menge der natürlichen Zahlen ist bezüglich der Relation

$$a \leq b :\Longleftrightarrow a|b$$

geordnet.

Ist S eine Menge mit Ordnungsrelation $\leq$, so ist eine Teilmenge $K \subseteq S$ eine Kette von M, falls für $k, k' \in K$ stets $k \leq k'$ oder $k' \leq k$ gilt. Eine Kette K ist induktiv, falls es ein $m \in S$ gibt, so dass $k \leq m$ für alle $k \in K$ ist.

Lemma von Zorn. *Sei $S \neq \emptyset$ eine Menge mit Ordnungsrelation $\leq$. Ist*

jede Kette in S induktiv, so gibt es maximale Elemente $m \in S$, d.h. aus $m \leq s$ mit $s \in S$ folgt $m = s$.

Bemerkung 16.6. Wir beweisen das Lemma von Zorn nicht, da der Beweis eher in die Mengenlehre gehört. Da das Zornsche Lemma äquivalent zum sogenannten Auswahlaxiom ist, könnte man das Zornsche Lemma auch einfach als Axiom annehmen. Da es aber auf den ersten Blick nicht sehr intuitiv ist, führt man es meist auf das Auswahlaxiom zurück, das wiederum auf den ersten Blick offensichtlich zu sein scheint. Es lautet: Sei A eine Menge von nicht-leeren Mengen. Dann gibt es eine Funktion f mit Definitionsbereich A und der Eigenschaft $f(X) \in X$ für alle $X \in A$, d.h. man kann aus allen Mengen X in A gleichzeitig ein Element auswählen.

Lemma 16.7. *Sei R ein Ring.*

(a) *Ist $I \subseteq R$ ein Ideal, so gibt es ein maximales Ideal $M \subsetneq R$ mit $I \subseteq M$ (d.h. ist $J \subsetneq R$ ein Ideal mit $M \subseteq J \subsetneq R$, so ist $M = J$). Insbesondere hat R maximale Ideale.*

(b) *Sei R kommutativ. Dann ist ein Ideal $M \subsetneq R$ genau dann maximal, wenn R/M ein Körper ist.*

Beweis. (a): Sei $S = \{J \mid J \subsetneq R \text{ Ideal mit } I \subseteq J\}$. Ein maximales Ideal $M \subsetneq R$ mit $I \subseteq M$ ist ein maximales Element aus S bzgl. der durch die Inklusion $\subseteq$ definierten Ordnungsrelation, d.h. es genügt zu zeigen, dass S die Voraussetzungen für das Zornsche Lemma erfüllt. Wegen $I \in S$ ist $\emptyset \neq S$. Sei $K \subseteq S$ eine Kette in S. Setze $s(K) = \cup\{J \mid J \in K\} \subseteq R$. Es genügt zu zeigen: $s(K) \in S$; gilt dies, so folgt die Behauptung mit dem Zornschen Lemma. Offensichtlich ist $I \subseteq s(K)$. Sind $s, s' \in s(K)$, so gibt es Ideale $J, J' \in K$ mit $s \in J$ und $s' \in J'$. Da K eine Kette ist, gilt oBdA $J \subseteq J'$. Damit folgt

$$s + s' \in J + J' = J' \subseteq s(K).$$

Sind $r, r' \in r$, so sind offenbar $rsr', rs'r' \in s(K)$. Es ist noch $s(K) \neq R$ zu zeigen. Wäre $s(K) = R$, so wäre $1 \in s(K)$, also gäbe es ein $J \in K$ mit $1 \in J$. Widerspruch! Insgesamt haben wir also eine obere Schranke zu K gefunden, d.h. jede Kette ist induktiv.

(b): Sei R kommutativ und $M \subsetneq R$ ein maximales Ideal. Für $r \in R \setminus M$ betrachte das Ideal $M + (r)$. Wegen $M \subseteq M + (r) \subseteq R$ folgt aus der Maximalität von M dann $R = M + (r)$. Also gibt es $m \in M$ und $s \in R$ mit $1 = m + sr$, so dass $1 + M = (s + M)(r + M)$ in R/M. Damit hat jedes $0 \neq r + M \in R/M$ hat ein multiplikativ Inverses $s + M$, d.h. R/M ist ein Körper. Sei umgekehrt $M \subseteq R$ ein Ideal, so dass R/M ein Körper ist. Sei J ein Ideal mit $M \subsetneq J$. Es ist zu zeigen, dass $J = R$ gilt. Sei

dazu $a \in J \setminus M$. Dann ist $a + M \in R/M$ ungleich Null, hat also ein Inverses bezüglich der Multiplikation, da R/M nach Voraussetzung ein Körper ist. Also gibt es $b \in R$ mit $(a + M) \cdot (b + M) = (1 + M)$, was äquivalent zu $ab - 1 \in M$ ist. Wegen $ab \in M \subseteq J$ ist dann auch $1 \in J$, also $J = R$. □

Bemerkung 16.8. Sei R ein Hauptidealring. In der Übung wird gezeigt:

$$(p) \text{ ist maximal } \iff p \text{ ist irreduzibel.}$$

Lemma 16.9. *Sei R ein kommutativer Ring. Dann gilt:*

$$R^n \cong R^m \Leftrightarrow n = m.$$

Beweis. Sei $F = R^m$, $F' = R^n$, und sei $F \cong F'$ als R-Moduln. Nach Lemma 16.7(a) gibt es in R ein maximales Ideal $M \subsetneq R$, und nach Lemma 16.7(b) ist $k = R/M$ ein Körper. Die Teilmenge $MF = \oplus_{i=1}^m M \subseteq \oplus_{i=1}^m R = F$ ist ein R-Untermodul, und $F/MF \cong \oplus_{i=1}^n R/M \cong \oplus_{i=1}^n k = k^n$. Dasselbe Argument zeigt $F'/MF' \cong k^m$. Wegen $F \cong F'$ folgt $k^n \cong k^m$, also ist $n = m$. Die Umkehrung ist trivial. □

Definition 16.10. Sei R ein kommutativer Ring. Ist $F = \oplus_{i=1}^n R$ ein (endlich erzeugter) freier R-Modul, so ist $n = Rg(F)$ der Rang von F.

• Der Rang $Rg(F)$ ist nach Lemma 16.9(b) wohldefiniert, und verallgemeinert den Dimensionsbegriff von Vektorräumen.

Wir betrachten im folgenden beliebige freie R-Moduln (d.h. nicht unbedingt endlich erzeugt), dabei sei $F = \oplus_{i \in I} Re_i$ der freie R-Modul mit Erzeugenden $\{e_i \mid i \in I\}$, d.h. $R \to Re_j$, $r \mapsto re_j$ ist ein Isomorphismus von R-Moduln; inbesondere ist $\sum_{i \in I} r_i e_j = 0$, dann und genau dann wenn $r_i = 0$ für alle $i \in I$ gilt.

Lemma 16.11. *Sei R ein Ring.*

(a) *Sei $F = \oplus_{i \in I} Re_i$ ein freier R-Modul, M ein R-Modul, und $m_i \in M$, $i \in I$. Dann gibt es einen eindeutig bestimmten R-Homomorphismus $f \in Hom_R(F, M)$ mit $f(e_i) = m_i$ für $i \in I$.*

(b) *Sei F ein freier R-Modul. Ist M ein R-Modul und $f : M \to F$ ein surjektiver R-Homomorphismus, so gibt es einen Untermodul $N \subseteq M$ mit $M = \ker(f) \oplus N$ und $N \cong F$.*

Beweis. (a): Einfaches Nachrechnen zeigt, dass die Abbildung

$$f : F \to M, \ \sum_{i=1}^{k} r_i e_i \mapsto \sum_{i=1}^{k} r_i m_i$$

einen R-Homomorphismus definiert. Wohldefiniertheit und Eindeutigkeit sind leicht zu verifizieren.

(b): Sei $F = \oplus_{i\in I} Re_i$ und sei $f : M \to F$ surjektiv. Seien $m_i,\ i \in I$, so dass $f(m_i) = e_i$. Ist $m \in M$, so folgt mit geeigneten $r_i \in R$, dass

$$f(m) = \sum_{i\in I} r_i e_i = \sum_{i\in I} r_i f(m_i) = f(\sum_{i\in I} r_i m_i),$$

d.h. $m - \sum_{i\in I} r_i m_i \in \ker(f)$. Ist $N = \langle m_i \mid i \in I\rangle$, so folgt $M = \ker(f) + N$. Angenommen $\sum_{i\in I} r_i m_i \in N \cap \ker(f)$. Dann folgt

$$0 = f(\sum_{i\in I} r_i m_i) = \sum_{i\in I} r_i e_i,$$

also ist $r_i = 0$ für alle i, und $N \cap \ker(f) = \{0\}$. Daher gilt $M = N \oplus \ker(f)$ und $F \cong M/\ker(f) \cong N$. □

Definition 16.12. Sei R ein Integritätsbereich und sei M ein R-Modul. Ein Element $m \in M$ ist ein Torsionselement, falls es ein $0 \neq r \in R$ mit $rm = 0$ gibt; setze $T(M) = \{m \in M \mid m \text{ ist Torsionselement}\}$. Ein R-Modul M ist torsionsfrei, falls $T(M) = \{0\}$ ist.

• Man rechnet leicht nach, dass $T(M) \subseteq M$ ein R-Untermodul ist. Weiter ist der Faktormodul $M/T(M)$ torsionsfrei: Sei $0 \neq r \in R$ und $m \in M$, so dass $r(m+T(M)) = T(M)$ ist. Dann ist $rm \in T(M)$, d.h. es gibt ein $0 \neq r' \in R$ mit $r'rm = 0$. Wegen $0 \neq rr'$ folgt $m \in T(M)$.

Beispiel 16.13. Ist M eine (additiv geschriebene) abelsche Gruppe, so sind die Torsionselemente von M als $\mathbb{Z}$-Modul genau diejenigen Elemente $m \in M$, so dass $nm = 0$ für ein $0 \neq n \in \mathbb{Z}$ ist. Zum Beispiel, $T(\mathbb{Z}/m\mathbb{Z}) = \mathbb{Z}/m\mathbb{Z}$, $T(\mathbb{Z}) = \{0\}$, und $T(\mathbb{Z} \oplus \mathbb{Z}/m\mathbb{Z}) \simeq \mathbb{Z}/m\mathbb{Z}$.

Proposition 16.14. *Sei R ein Hauptidealring.*

(a) *Sei $M \subseteq F = R^n$ ein R-Untermodul des freien R-Moduls F vom Rang n. Dann ist M ein freier R-Modul vom Rang $\leq n$ (d.h. $M \cong R^k$ für ein $k \leq n$),*

(b) *Sei M ein endlich erzeugter R-Modul. Dann gilt:*

$$M \text{ ein freier } R\text{-Modul} \iff M \text{ ist torsionsfrei.}$$

Jeder endlich-erzeugte torsionfreie R-Modul ist also isomorph zu R^k für ein $k \geq 0$.

(c) *Ist M ein endlich erzeugter R-Modul, so ist $M \cong T(M) \oplus F$, wobei $F \cong R^k$ für ein geeignetes k ist.*

Beweis. (a): Induktion nach n. Ist $n = 1$, so ist ein R-Untermodul $M \subseteq F = R$ ein Ideal in R. Da R ein Hauptidealring ist, gibt es ein $a \in R$ mit $M = (a) = Ra$; also ist M ein freier R-Modul mit Erzeuger a vom Rang 0 (falls $a = 0$) oder 1 (falls $a \neq 0$), d.h. die

Behauptung gilt für $n = 1$. Sei nun $n > 1$. Seien $e_1, \ldots, e_n$ Erzeuger von F, d.h. $F = \oplus_{i=1}^n Re_i$. Setze $F' = R^{n-1} = \oplus_{i=2}^n Re_i$. Dann definiert

$$\pi : F \to F', \ \sum_{i=1}^n r_i e_i \mapsto \sum_{i=2}^n r_i e_i$$

einen surjekitven R-Homomorphismus mit $\ker(\pi) = Re_1$ frei vom Rang 1. Das Bild $\pi(M)$ von $M \subseteq F$ unter π ist ein R-Untermodul von F'. Nach Induktion ist $\pi(M)$ frei vom Rang $\leq n-1$. Lemma 16.11(b), angewandt auf $\pi|_M : M \to \pi(M)$, liefert eine Zerlegung $M = (M \cap \ker(\pi)) \oplus B$, mit $B \cong \pi(M)$. Da der R-Untermodul $M \cap \ker(\pi) \subseteq \ker(\pi) = Re_1$ frei vom Rang ≤ 1 ist, ist M frei vom Rang $\leq n$.

(b): Sei $M = \langle m_1, \ldots, m_n \rangle$ endlich erzeugt und torsionsfrei. Für $i = 1, \ldots, n$ ist der R-Untermodul $Rm_i \subseteq M$ frei. Sei $k \geq 1$ maximal, so dass mit geeigneter Numerierung der R-Untermodul von M

$$F = Rm_1 \oplus \cdots \oplus Rm_k$$

ein freier R-Modul ist. Für $j > k$ gilt somit $F \cap Rm_j \neq \{0\}$. Also gibt es für $j = k+1, \ldots, n$ jeweils ein $0 \neq r_j \in R$, so dass

$$r_j m_j = \sum_{i=1}^k r_{ji} m_i.$$

Ist $r = r_{k+1} \cdots r_n$, so ist $r \neq 0$ und $rm_j \in F$ für $j = k+1, \ldots, n$. Damit liegt das r-Vielfache aller Erzeuger von M in F, und es gilt $rM \subseteq F$. Da $rM \subseteq F$ ein R-Untermodul ist, ist rM nach (a) ein freier R-Modul. Die Abbildung $g : M \to rM, \ m \mapsto rm$, ist ein surjektiver R-Homomorphismus mit $\ker(g) = \{m \in M \mid rm = 0\} \subseteq T(M) = \{0\}$. Also gilt $M \cong rM \subseteq F$, und M ist frei (vom Rang $\leq k$).

(c): Ist M ein beliebiger R-Modul, so ist $M/T(M)$ torsionsfrei. Ist M endlich erzeugt ist, so ist auch der Quotient $M/T(M)$ endlich erzeugt. Nach (b) ist $M/T(M) \cong R^k$ ein freier R-Modul. Nach Lemma 16.11(b), angewandt auf die Quotientenabbildung $M \to M/T(M)$ mit Kern $T(M)$ ist $M = T(M) \oplus F$, wobei $F \cong M/T(M) \cong R^k$ frei ist. $\square$

Die Aussagen von Proposition 16.14 sind falsch, falls M nicht endlich erzeugt ist:

Beispiele 16.15. (a) Betrachte den $\mathbb{Z}$-Modul $\mathbb{Q}$. Man zeigt leicht, dass $\mathbb{Q}$ als $\mathbb{Z}$-Modul nicht endlich erzeugt ist. Offensichtlich ist $T(\mathbb{Q}) = \{0\}$, d.h. $\mathbb{Q}$ ist torsionsfrei. Da es zu jeden $q \in \mathbb{Q}$ ein $q' \in \mathbb{Q}$ mit $q = 2q'$ gibt, ist der $\mathbb{Z}$-Homomorphismus $q \mapsto 2q$ surjektiv, und $\mathbb{Q}$ ist kein freier $\mathbb{Z}$-Modul.

(b) Sei p eine Primzahl, und sei $A_i = \mathbb{Z}/p^{2i}\mathbb{Z}$. Dann ist die Menge

$$A = \{(b_1, b_2, \dots) \mid b_i \in A_i\}$$

ein $\mathbb{Z}$-Modul bzgl. der komponentenweisen Operationen; offensichtlich ist A nicht endlich erzeugt. Für dieses A gilt: $T(A)$ ist kein direkter Summand von A (siehe Übungsblatt).

Um die Struktur von Torsionsmoduln zu studieren, setzen wir ab jetzt voraus, dass R ein euklidischer Ring ist. Im folgenden sei $M = \bigoplus_{i=1}^{n} Ru_i$ ein freier R-Modul vom Rang n und $N \subseteq M$ ein Teilmodul gegeben durch Erzeugende $v_1, \dots, v_k$.

Theorem 16.16. *(Elementarteilersatz) Es gibt Basen $u_1, \dots, u_n$ von M und $v_1, \dots, v_m$ von N mit $m \leq n$ und*

$$v_i = \epsilon_i u_i, i = 1, \dots, m, \quad \epsilon_1 \mid \epsilon_2 \mid \dots \mid \epsilon_m.$$

Beweis. Der Beweis ist algorithmisch und kann als Verallgemeinerung des Gaußschen Algorithmus aufgefasst werden. Sei zunächst eine beliebige Basis $u_1, \dots, u_n$ von M gegeben und beliebige Erzeugende $v_1, \dots, v_k$ von N. Schreibe

$$v_k = \sum_{i=1}^{n} \alpha_{ik} u_i \text{ mit } \alpha_{ik} \in R.$$

Sei $A = (\alpha_{ik}) \in R^{n \times k}$. Wir werden die Matrix A durch schrittweises Abändern der Basis $u_1, \dots, u_n$ und des Erzeugendensystems $v_1, \dots, v_k$ in die Form

$$\left(\begin{array}{ccc|c} \epsilon_1 & & & \\ & \ddots & & 0 \\ & & \epsilon_m & \\ \hline & 0 & & 0 \end{array}\right)$$

transfromieren. Erlaubte Abänderungen sind dabei:

(1) Vertauschung zweier u oder v. Dies entspricht der Vertauschung zweier Zeilen oder Spalten.

(2) Ersetzung eines u_i durch $u_i + \lambda u_j, \lambda \in R, i \neq j$. Wegen

$$v_k = \sum_{l=1}^{n} \alpha_{lk} u_l = \sum_{l=1, l \neq i,j}^{n} \alpha_{lk} u_l + \alpha_{ik}(u_i + \lambda u_j) + (\alpha_{jk} - \alpha_{ik}\lambda) u_j$$

entspricht dies der Ersetzung der j-ten Zeile durch j-te Zeile minus λ mal i-te Zeile.

(3) Ersetzung eines v_i durch $v_i - \lambda v_j, \lambda \in R, i \neq j$. Dies entspricht der Ersetzung der i-ten Spalte durch i-te Spalte minus λ mal j-te Spalte.

Wir wenden nun den folgenden Algorithmus auf die Matrix A an:

Schritt 1: Durch Vertauschen von Zeilen und Spalten bringe man das Element ungleich Null von kleinster euklidischer Norm an die Stelle $(1,1)$.

Schritt 2: Durch Subtraktion geeigneter Vielfacher der ersten Zeile, kann man erreichen, dass A von der Form

$$A = \begin{pmatrix} \alpha_{11} & * & \dots & * \\ \gamma_2 & & & \\ \vdots & & * & \\ \gamma_n & & & \end{pmatrix}$$

ist, wobei jedes γ_i entweder gleich Null ist oder $\varphi(\gamma_i) < \varphi(\alpha_{11})$ gilt. Falls alle γ_i gleich Null sind, so gehe zu Schritt 3. Andernfalls gehe zu Schritt 1.

Schritt 3: Durch Subtraktion geeigneter Vielfacher der ersten Spalte, kann man sodann erreichen, dass A von der Form

$$A = \left(\begin{array}{c|ccc} \alpha_{11} & \gamma_1 & \dots & \gamma_k \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)$$

ist, wobei jedes γ_i entweder gleich Null ist oder $\varphi(\gamma_i) < \varphi(\alpha_{11})$ gilt. Falls alle γ_i gleich Null sind, so gehe zu Schritt 4. Andernfalls gehe zu Schritt 1.

Schritt 4: Falls $A' = 0$ so beende den Algorithmus.

Schritt 5: Falls alle Einträge von A' durch α_{11} teilbar sind, so gehe mit A' in Schritt 1.

Schritt 6: Sei α_{ik} ein Koeffizient in A', der nicht durch α_{11} teilbar ist. Teile mit Rest,

$$\alpha_{ik} = \alpha_{11}\beta + \gamma, \gamma \neq 0, \varphi(\gamma) < \varphi(\alpha_{11}).$$

Addiere nun die erste Zeile zur i-ten Zeile und subtrahiere dann β mal 1. Spalte von der k-ten Spalte. Dann kommt an der Stelle (i,k) gerade γ zu stehen. Gehe mit A in Schritt 1.

Der Algorithmus endet nach endlich vielen Schritten, da $\varphi\colon R \setminus \{0\} \longrightarrow \mathbb{N}_0$ und in den Schritten 2,3 und 5 die minimale euklidische Norm der Elemente von A verringert wird. □

Bemerkung 16.17. Die Aussage des Elementarteilersatzes ist gleichbedeutend mit der folgenden Aussage: Es gibt Matrizen $P \in \mathrm{Gl}_n(R)$,

$Q \in \mathrm{Gl}_k(R)$, so dass

$$PAQ = \left(\begin{array}{ccc|c} \epsilon_1 & & & \\ & \ddots & & 0 \\ & & \epsilon_m & \\ \hline & 0 & & 0 \end{array}\right)$$

Wir wenden nun den Elementarteilersatz an, um die Struktur von endlich-erzeugten Moduln über euklidischen Ringen zu studieren. Sei dazu $M = \langle m_1, \ldots, m_n \rangle$ ein endlich erzeugter R-Modul. Sei $F = \bigoplus_{i=1}^n Ru_i$ der freie Modul vom Rang n. Dann ist die Abbildung

$$f\colon F \longrightarrow M, \quad \sum_{i=1}^n r_i u_i \mapsto \sum_{i=1}^n r_i m_i,$$

ein surjektiver Modulhomomorphismus. Sei $N := \ker(f)$. Dann gilt nach dem Isomorphiesatz:

$$F/N \simeq M.$$

Nach dem Elementarteilersatz gibt Basen $u_1, \ldots, u_n$ von F und $v_1, \ldots, v_m$ von N mit $m \leq n$ sowie $\epsilon_1, \ldots, \epsilon_m$ mit

$$v_1 = \epsilon_1 u_1, \ldots, v_m = \epsilon_m u_m, \quad \epsilon_1 \mid \epsilon_2 \mid \ldots \mid \epsilon_m.$$

Es folgt dann:

$$\begin{aligned} F/N &\simeq \bigoplus_{i=1}^m Ru_i/R\epsilon_i u_i \oplus \bigoplus_{i=m+1}^n Ru_i \\ &\simeq \bigoplus_{i=1}^m R/R\epsilon_i \oplus R^{m-n}. \end{aligned} \tag{2}$$

Insbesondere gilt also wegen $R/R\epsilon = 0$ für $\epsilon \in R^\times$

$$T(M) \simeq \bigoplus_{i=1, \epsilon_i \notin R^\times}^m R/R\epsilon_i, \quad r = \mathrm{rk}(M) = n - m.$$

Bis auf die Eindeutigkeitsaussage haben wir damit den folgenden wichtigen Satz vollständig bewiesen.

Theorem 16.18. *Sei R ein euklidischer Ring und M ein endlich erzeugter R-Modul. Dann gibt es $r \in \mathbb{N}_0$ und $\epsilon_1, \ldots, \epsilon_s \in R \setminus R^\times$ mit $\epsilon_1 \mid \epsilon_2 \mid \ldots \mid \epsilon_s$, so dass*

$$M \simeq \bigoplus_{i=1}^s R/R\epsilon_i \oplus R^r.$$

Der Rang r ist eindeutig durch den Isomorphietyp von M bestimmt. Ebenso sind die $\epsilon_1, \ldots, \epsilon_s$ bis auf Assoziiertheit durch den Isomorphietyp von M eindeutig bestimmt.

Definition 16.19. Die durch M eindeutig bis auf Assoziiertheit bestimmten Elemente $\epsilon_1, \ldots, \epsilon_s$ aus Satz 16.18 nennt man die Elementarteiler von M. Falls $R = \mathbb{Z}$, so normieren wir die Elementarteiler durch dir Forderung $\epsilon_i > 0$. Falls $R = \mathbb{K}[x]$, so normieren wir die Elementarteiler durch die Forderung, dass der führende Koeffizient von ϵ_i gleich 1 ist. Damit werden die Elementarteiler in diesen Fällen eindeutig (und nicht nur eindeutig bis auf Assoziiertheit).

Bemerkung 16.20. Seien M, M' zwei endlich erzeugte R-Moduln mit Rang r und r' sowie Elementarteilern $\epsilon_1, \ldots, \epsilon_s$ und $\epsilon'_1, \ldots, \epsilon'_{s'}$. Dann sind M und M' genau dann isomorph, wenn $r = r'$ gilt und die Reihen der Elementarteiler übereinstimmen.

Im folgenden werden wir den Nachweis der noch fehlenden Eindeutigkeit erbringen. Da R als euklidischer Ring insbesondere faktoriell ist, können wir jedes ϵ_i eindeutig in der Form

$$\epsilon_i = u_i \prod_{j=1}^{t} \pi_j^{e_{ij}}$$

mit $e_{ij} \in \mathbb{N}_0$, $u_i \in R^\times$ und paarweise verschiedenen irreduziblen Elemente $\pi_j \in R$ schreiben. Da wir $e_{ij} = 0$ erlauben, können wir oE für alle i über die selbe Menge von Primelementen $\pi_1, \ldots, \pi_t$ laufen. Nach dem Chinesischen Restsatz erhalten wir

$$R/R\epsilon_i \simeq \bigoplus_{j=1}^{t} R/\left(\pi_j^{e_{ij}}\right), \quad i = 1, \ldots, s,$$

und zusammenfassend

$$M \simeq R^r \oplus \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{t} R/\left(\pi_j^{e_{ij}}\right) \simeq R^r \oplus \bigoplus_{j=1}^{t} \left(\bigoplus_{i=1}^{s} R/\left(\pi_j^{e_{ij}}\right) \right) \tag{3}$$

Definition 16.21. Sei R ein Hauptidealring und M ein R-Modul. Sei π ein irreduzibles Element von R. Dann nennt man

$$T_\pi(M) = \{m \in M \mid \pi^e m = 0 \text{ für ein geeignetes } e \in \mathbb{N}_0\}$$

den π-Torsionsmodul von M.

Man zeigt leicht, dass $T_\pi(M)$ ein Teilmodul von $T(M) \subseteq M$ ist. Aus (3) folgt

$$T_{\pi_j}(M) \simeq \bigoplus_{i=1}^{s} R/\left(\pi_j^{e_{ij}}\right).$$

Definition 16.22. Die Elemente

$$\pi_1^{e_{1,1}}, \ldots, \pi_1^{e_{s,1}}, \pi_2^{e_{1,2}}, \ldots, \pi_2^{e_{s,2}}, \ldots, \pi_t^{e_{1,t}}, \ldots, \pi_t^{e_{s,t}}$$

mit

$$\begin{aligned} &e_{1,1} \geq e_{2,1} \geq \ldots \geq e_{s,1} \geq 0, \\ &e_{1,2} \geq e_{2,2} \geq \ldots \geq e_{s,2} \geq 0, \\ &\ldots \\ &e_{1,t} \geq e_{2,t} \geq \ldots \geq e_{s,t} \geq 0 \end{aligned}$$

nennt man die Invariantenteiler von M.

Lemma 16.23. *Sei M ein endlich erzeugter R-Torsionsmodul. Dann entsprechen sich Elementarteiler und Invariantenteiler eineindeutig.*

Beweis. Leicht. □

Es genügt also zu zeigen, dass die Reihe der Invariantenteiler eindeutig durch den Isomorphietyp von M bestimmt ist.

Für den noch ausstehenden Eindeutigkeitsbeweis können wir uns oE auf die Betrachtung von endlich erzeugten R-Torsionsmoduln M beschränken. Für jedes irreduzible Element π ist $T_\pi(M)$ bis auf Isomorphie eindeutig durch den Isomorphietyp von M bestimmt. Es reicht also zu zeigen, dass die Invariantenteiler

$$\pi_j^{e_{1,j}}, \ldots, \pi_j^{e_{s,j}} \text{ mit } e_{1,j} \geq e_{2,j} \geq \ldots \geq e_{s,j}$$

eindeutig durch $T_{\pi_j}(M)$ bestimmt sind. Dazu setzen wir $\pi := \pi_j$ und entsprechend $T := T_{\pi_j}(M)$. Sei $k := R/(\pi)$ der Restklassenkörper. Dann ist $T/\pi T$ ein k-Vektorraum und es gilt

$$\dim_k(T/\pi T) = \text{ Anzahl der } e_{i,j} > 0.$$

Sodann betrachten wir den k-Vektorraum $\pi T/\pi^2 T$ und zeigen

$$\dim_k(\pi T/\pi^2 T) = \text{ Anzahl der } e_{i,j} > 1.$$

Also ist die Anzahl der $e_{i,j}$ mit $e_{i,j} = 1$ gegeben durch

$$\dim_k(T/\pi T) - \dim_k(\pi T/\pi^2 T),$$

was nur vom Isomorphietyp von T abhängt. Sukzessive zeigt man auf diese Weise, dass

$$|\{e_{i,j} \text{ mit } e_{i,j} = l\}| = \dim_k(\pi^{l-1}T/\pi^l T) - \dim_k(\pi^l T/\pi^{l+1}T).$$

Die rechte Seite hängt dabei nur vom Isomorphietyp von T ab. Der Isomorphietyp von T wiederum ist vom Isomorphietyp von M eindeutig bestimmt. Damit ist der Beweis der Eindeutigkeit vollständig erbracht.

Wir fassen unser Hauptresultat nochmals zusammen.

Theorem 16.24. *Sei R ein euklidischer Ring und M ein endlich erzeugter R-Modul. Dann gibt es $r \in \mathbb{N}_0$, irreduzible Elemente $\pi_1, \ldots, \pi_t$, sowie natürliche Zahlen $e(1,j) \geq \cdots \geq e(s,j) \geq 0$, $j = 1, \ldots, t$, so dass*

$$\begin{aligned} M &\cong R^r \oplus T_{\pi_1}(M) \oplus \cdots T_{\pi_t}(M), \text{ wobei} \\ T_{\pi_j}(M) &\cong R/(\pi_j^{e_{1,j}}) \oplus R/(\pi_j^{e_{2,j}}) \oplus \cdots \oplus R/(\pi_j^{e_{s,j}}). \end{aligned}$$

Die Zahlen r sowie die Invariantenteiler (und damit auch die Elementarteiler) sind eindeutig (bis auf Assoziiertheit) durch den Isomorphietyp von M bestimmt.

• Ein R-Modul M ist zyklisch, falls $M = Rm$ für ein $m \in M$ ist. Das obige Theorem besagt, dass über einem euklidischen Ring jeder endlich erzeugte R-Modul eine direkte Summe von zyklischen R-Modulen ist.
• Der obige Satz gilt allgemeiner für Hauptidealringe, wobei der Beweis dann etwas komplizierter und vor allem nicht mehr algorithmisch ist.

Abschließend notieren wir den sogenannten Hauptsatz über endlich erzeugte abelsche Gruppen. Man erinnere sich, dass die endlich erzeugten abelschen Gruppen genau die endlich erzeugten $\mathbb{Z}$-Moduln sind.

Korollar 16.25. *Für eine endlich erzeugte abelsche Gruppe A gibt es ein $r \in \mathbb{N}_0$, sowie Primzahlen $p_1, \ldots, p_t$ sowie natürliche Zahlen $e_{1,j} \geq \cdots \geq e_{s,j} \geq 0$, $j = 1, \ldots, t$, so dass*

$$\begin{aligned} A &\cong \mathbb{Z}^k \oplus T_{p_1}(A) \oplus \cdots T_{p_t}(A), \text{ wobei} \\ T_{p_j}(A) &\cong \mathbb{Z}/(p_j^{e_{1,j}}) \oplus \mathbb{Z}/(p_j^{e_{2,j}}) \oplus \cdots \oplus \mathbb{Z}/(p_j^{e_{s,j}}). \end{aligned}$$

17. Allgemeine und Jordansche Normalform

Definition 17.1. Sei R ein kommutativer Ring mit 1.

(a) Seien $A, B \in K^{n\times m}$. Dann heißt A äquivalent zu B, in Zeichen $A \sim B$, falls es $P \in \mathrm{Gl}_n(R)$ und $Q \in \mathrm{Gl}_m(R)$ gibt, so dass

$$B = P^{-1}AQ$$

gilt.

(b) Seien $A, B \in M_n(K) = K^{n\times n}$. Dann heißt A konjugiert oder ähnlich zu B, in Zeichen $A \approx B$, falls es $S \in \mathrm{Gl}_n(K)$ gibt, so dass

$$B = S^{-1}AS$$

gilt.

Bemerkung 17.2. Sowohl $\sim$ als auch $\approx$ sind Äquivalenzelationen. Falls $f\colon V \longrightarrow W$ eine lineare Abbildung ist zwischen endlich-erzeugten K-Vektorräumen mit $\dim(V) = m, \dim(W) = n$, so kann man nach

Wahl von Basen X und Y in V und W die Abbildung f durch eine Matrix $A = A_{f,X,Y}$ darstellen. Bei der Wahl anderer Basen X' und Y' erhält man eine zu A äquivalente Matrix $B = A_{f,X',Y'}$, d.h. $B = P^{-1}AQ$, wobei P und Q die Basisübergangsmatrizen sind.

Entsprechendes gilt für Endomorphismen $f\colon V \longrightarrow V$ und dem Begriff der Ähnlichkeit.

Wir werden im Folgenden in jeder Äquivalenzklasse bez. $\approx$ einen kanonischen Vertreter bestimmen. Diesen Vertreter nennen wir dann die allgemeine Normalform von A. Auf diese Weise können wir entscheiden, ob zwei gegebene Matrizen A und B konjugiert sind. Wir verwenden den Struktursatz 16.24 um die allgemeine bzw. Jordan Normalformen eines Endomorphismus zu bestimmen.

Definition 17.3. Sei $A \in M_n(K) = K^{n\times n}$. Dann heißt die Matrix

$$M_A(x) := xE - A \in M_n(K[x])$$

die charakteristische Matrix von A.

Sei V ein K-Vektorraum, und sei $A \in End_K(V)$ fest gewählt. Betrachte wie in Beispiel 16.2(d) V als $K[x]$-Modul mittels

$$K[x] \times V \to V,\ (f, v) \mapsto f \cdot v = f(A)v.$$

Sei $\dim_K V < \infty$. Das Minimalpolynom μ_A von A ist ein nicht-triviales Polynom, so dass $\mu_A(A) = 0$. Also ist $\mu_A \cdot V = 0$, d.h. V ist ein $K[x]$-Torsionsmodul. Als endlich-dimensionaler K-Vektorraum ist V endlich erzeugt, und somit ein endlich erzeugter $K[x]$-Torsionsmodul. Da $K[x]$ ein euklidischer Ring ist, können wir Theorem 16.24 auf V anwenden. Die resultierende Zerlegung des $K[x]$-Moduls V in eine endliche direkte Summe von zyklischen $K[x]$-Torsionsmoduln ist die Grundlage der allgemeinen Theorie der Normalformen.

Im Folgenden werden wir nicht mehr zwischen $A \in \mathrm{End}_K(V)$ und einer darstellenden Matrix $A \in K^{n\times n}$ unterscheiden. Dies stellt zwar einen kleinen Missbrauch der Notation dar, da die darstellende Matrix von der Wahl einer Basis in V abhängt, vereinfacht jedoch die Notationen erheblich. Sämtlich Definitonen und Resultate sind aber aus offensichtlichen Gründen davon unabhängig.

Wenn wir V als $K[x]$-Modul mittels einer Matrix A betrachten, dann schreiben wir im weiteren $V = V_A$, um anzudeuten, dass $f(x) \cdot v := f(A)v$ für alle Polynome $f \in K[x]$ und alle $v \in V$.

Theorem 17.4. *Seien $A, B \in M_n(K)$. Dann sind folgende Aussagen äquivalent:*

(i) A ist konjugiert zu B, in Zeichen, $A \approx B$.
(ii) $V_A \simeq V_B$ als $K[x]$-Moduln.
(iii) $K[x]^n/\langle M_A(x)\rangle \simeq K[x]^n/\langle M_B(x)\rangle$ als $K[x]$-Moduln.
(iv) $M_A(x)$ und $M_B(x)$ sind äquivalent, in Zeichen, $M_A(x) \sim M_B(x)$.
(v) $M_A(x)$ und $M_B(x)$ haben die selben Elementarteiler.
(vi) $M_A(x)$ und $M_B(x)$ haben die selben Invariantenteiler.

Beweis. "(i) $\Longrightarrow$ (ii)": Sei $B = S^{-1}AS$. Betrachte den Homomorphismus

$$g\colon V_A \longrightarrow V_B, \quad v \mapsto S^{-1}v.$$

Wegen $S \in \mathrm{Gl}_n(K)$ ist g ein Isomorphismus. Um zu zeigen, dass es ein Isomorphismus von $K[x]$-Moduln ist, genügt es zu zeigen, das $g(x\cdot v) = x \cdot g(v)$ gilt. Dies zeigt die folgende Rechnung:

$$g(x \cdot v) = g(Av) = S^{-1}Av = S^{-1}ASS^{-1}v = BS^{-1}v = x \cdot g(v).$$

"(ii) $\Longrightarrow$ (i)": Sei $g\colon V_A \longrightarrow V_B$ ein Isomorphismus von $K[x]$-Moduln. Dann ist g insbesondere eine K-lineare Abbildung und nach Wahl einer Basis von V wird g durch eine Matrix $S \in \mathrm{Gl}_n(K)$ dargestellt. Aus $g(x \cdot v) = x \cdot g(v)$ folgt $SA = BS$, also $A = S^{-1}BS$.

"(ii) $\Longleftrightarrow$ (iii)": Wir zeigen, dass $K[x]^n/\langle M_A(x)\rangle \simeq V_A$ als $K[x]$-Moduln. Sei dazu $v_1, \ldots, v_n$ eine Basis von $V = V_A$. Betrachte den surjektiven $K[x]$-Modulhomomorphismus

$$h\colon K[x]^n \longrightarrow V_A, \quad e_i \mapsto v_i,$$

wobei $e_1, \ldots, e_n$ die Standardbasis des $K[x]^n$ bezeichnet.Wir zeigen, dass $\langle M_A(x)\rangle$ im Kern von h liegt. Die j-te Spalte der charakteristischen Matrix $M_A(x)$ ist von der Form $(-\alpha_{1,j}, \ldots, x - \alpha_{jj}, \ldots, -\alpha_{n,j})$. Bezüglich der Basis $e_1, \ldots, e_n$ entspricht die j-te Spalte also dem Element $s_j = -\sum_{i=1, i\neq j}^{n} \alpha_{ij}e_i + (x - \alpha_{jj})e_j$ und es gilt

$$\begin{aligned} h(s_j) &= -\sum_{i=1, i\neq j}^{n} \alpha_{ij}v_i + (x - \alpha_{jj})v_j \\ &= -\sum_{i=1}^{n} \alpha_{ij}v_i + Av_j \\ &= 0 \end{aligned}$$

Damit induziert h eine wohldefinierte surjektive Abbildung

$$\bar{h}\colon K[x]^n/\langle M_A(x)\rangle \longrightarrow V_A.$$

Um zu zeigen, dass $\bar{h}$ ein Isomorphismus ist, genügt es zu beweisen, dass $\dim_K(K[x]^n/\langle M_A(x)\rangle) = n$ ist. Nach dem Elementarteilersatz ist

$K[x]^n/\langle M_A(x)\rangle)$ isomorph zu $\bigoplus_{i=1}^n K[x]/(c_i(x))$ mit Polynomen $c_i(x)$, so dass $c_1(x) \mid c_2(x) \mid \ldots \mid c_n(x)$. Es folgt:

$$\begin{aligned} dim_K(K[x]^n/\langle M_A(x)\rangle &= dim_K\left(K[x]^n/\left\langle \begin{pmatrix} c_1(x) & & \\ & \ddots & \\ & & c_n(x) \end{pmatrix} \right\rangle\right) \\ &= \sum_{i=1}^n \dim_K(K[x]/(c_i(x)) \\ &= \sum_{i=1}^n \deg(c_i(x)) \\ &= \deg(c_1(x)\cdots c_n(x)) \\ &\overset{(*)}{=} \deg(\det(M_A(x))) = n, \end{aligned}$$

wobei die Gleichheit $(*)$ aus folgendem Grund gilt. Die Matrizen $M_A(x)$ und $\mathrm{diag}(c_1(x),\ldots,c_n(x))$ sind äquivalent, d.h. es gibt $P, Q \in \mathrm{Gl}_n(K[x])$ mit $\mathrm{diag}(c_1(x),\ldots,c_n(x)) = P^{-1}M_A(x)Q$. Wegen $\det(P), \det(Q) \in K[x]^\times = K^\times$ folgt die Behauptung aus der Multiplikativität der Determinante.

Die Aussagen (iii), (iv), (v) und (vi) sind allesamt äquivalente Umformungen des Elementarteilersatzes. □

Im letzten Beweis haben wir insbesondere das folgende Lemma sowie Teil (a) von Lemma 17.6 bewiesen.

Lemma 17.5. *Seien $A \in M_n(K)$. Dann induziert die kanonische Abbildung $K[x]^n \longrightarrow V_A$ einen Isomorphismus*

$$K[x]^n/\langle M_A(x)\rangle \simeq V_A.$$

Lemma 17.6. *Seien $A \in M_n(K)$ und $c_1(x) \mid c_2(x) \mid \ldots \mid c_n(x)$ die Polynome aus dem Elementarteilersatz, so dass*

$$K[x]^n/\langle M_A(x)\rangle \simeq \bigoplus_{i=1}^n K[x]/(c_i(x))$$

gilt. Dann gilt:
(a) $\chi_A(x) = \prod_{i=1}^n c_i(x)$.
(b) $\mu_A(x) = c_n(x)$.

Beweis. (a) wurde im Beweis von Satz 17.4 bereits bewiesen. (b) folgt aus

$$V_A \simeq K[x]^n/\langle M_A(x)\rangle \simeq \bigoplus_{i=1}^n K[x]/(c_i(x))$$

zusammen mit $c_i(x) \mid c_n(x)$. □

Der Satz 17.4 rechtfertigt die folgende Definition.

Definition 17.7. Sei $A \in M_n(K)$. Seien $g_1(x), \ldots, g_r(x)$ die Elementarteiler von $M_A(x)$. oE seien die g_i normiert. Dadurch sind die g_i durch A eindeutig bestimmt (nicht nur bis auf Einheiten in $K[x]$). Wir nennen die g_i im weiteren auch die Elementarteiler der Matrix A. Analog bezeichnen wir die normierten Invariantenteiler von $M_A(x)$ auch als die Invariantenteiler von A.

Beispiel 17.8. Die Matrix $A = \begin{pmatrix} 1 & 2 \\ 5 & 13 \end{pmatrix} \in M_2(\mathbb{Q})$ hat die charakteristische Matrix

$$M_A(x) = \begin{pmatrix} x-1 & -2 \\ -5 & x-13 \end{pmatrix} \in M_2(\mathbb{Q}[x]).$$

Man berechnet $c_1(x) = 1, c_2(x) = (x-1)(x-13) - 10$. Die Matrix $B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \in M_2(\mathbb{Q})$ hat die charakteristische Matrix

$$M_A(x) = \begin{pmatrix} x-1 & -1 \\ 0 & x-2 \end{pmatrix} \in M_2(\mathbb{Q}[x]).$$

$c_1(x) = 1, c_2(x) = (x-1)(x-2)$.

Die Matrizen A und B sind also nicht ähnlich.

Zu $g(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in K[x], \deg(g) = n \geq 1$, betrachte

$$B_g := \begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & & & -a_2 \\ & & \cdots & & \cdots \\ & & & \cdots & \cdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix}$$

Für $n = 1$ ist $B_g = (-a_0)$. Wir nennen B_g die Begleitmatrix zu g.

Lemma 17.9. *Sei $g(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in K[x], \deg(g) = n \geq 1$ und $B = B_g$ die Begleitmatrix. Dann gilt:*

(i) Das charakteristische Polynom von B ist gegeben durch $\chi_B(x) = g(x)$.

(ii) Es gilt:

$$M_B(x) \sim \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & g \end{pmatrix}$$

Beweis. Übung. □

Korollar 17.10. *Es seien $g_1(x), \ldots, g_r(x) \in K[x]$ normierte Polynome vom Grad ≥ 1 mit*

$$g_i \mid g_{i+1}, \quad i = 1, \ldots, r-1.$$

Sei

$$B = B_{g_1,\ldots,g_r} = \begin{pmatrix} B_{g_1} & & \\ & \ddots & \\ & & B_{g_r} \end{pmatrix} \in M_n(K),$$

wobei $n = \sum_{i=1}^{r} \deg(g_i(x))$. Dann hat B die Elementarteiler $g_1, \ldots, g_r$.

Beweis. Es gilt

$$\begin{aligned} M_B(x) &= \begin{pmatrix} xE - B_{g_1} & & \\ & \ddots & \\ & & xE - B_{g_r} \end{pmatrix} \\ &\sim \left(\begin{array}{cccc|c|ccccc} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & g_1 & & & & & & \\ \hline & & & & & \ddots & & & & \\ \hline & & & & & & 1 & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & 1 & \\ & & & & & & & & & g_r \end{array}\right) \\ &\sim \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & g_1 & \\ & & & & \ddots \\ & & & & & g_r \end{pmatrix} \end{aligned}$$

Die erste Äquivalenz folgt hierbei aus dem Lemma 17.9, die zweite durch einfache Zeilen- und Spaltenvertauschungen. □

Theorem 17.11. *(Frobeniussche Normalform) Sei $A \in M_n(K)$. Dann ist A zu genau einer Matrix $B_{g_1,\ldots,g_r}$ ähnlich, wobei $g_1, \ldots, g_r$ normierte Polynome vom Grad ≥ 1 sind mit $g_1 \mid g_2 \mid \ldots \mid g_r$. Die Polynome $g_1, \ldots, g_r$ sind hierbei genau die Elementarteiler der charakteristischen Matrix $M_A(x)$.*

Beweis. Da $M_A(x)$ und $M_{B_{g_1,\ldots,g_r}}(x)$ die selben Elementarteiler haben, folgt dies sofort aus dem Satz 17.4. □

Beispiel 17.12. Sei

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{pmatrix} \in M_n(\mathbb{Q}).$$

Dann zeigt man mit dem Algorithmus aus dem Elementarteilersatz, dass

$$M_A(x) \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & (x-1)(x-2) \end{pmatrix}.$$

Also hat die Frobeniussche Normalform zwei Kästchen korrespondierend zu $x-1$ und $(x-1)(x-2) = x^2 - 3x + 2$ und ist von der Form

$$\left(\begin{array}{c|cc} 1 & & \\ \hline & 0 & -2 \\ & 1 & 3 \end{array}\right).$$

Lemma 17.13. *Sei $g = h_1 \cdots h_k$ ein Produkt von paarweise teilerfremden Polynomen $h_1, \ldots, h_k \in K[x]$ vom Grad ≥ 1. Dann gilt:*

$$B_g \approx \begin{pmatrix} B_{h_1} & & \\ & \ddots & \\ & & B_{h_k} \end{pmatrix}.$$

Beweis. Es sei

$$B = B_{h_1,\ldots,h_k} = \begin{pmatrix} B_{h_1} & & \\ & \ddots & \\ & & B_{h_k} \end{pmatrix}.$$

Wie im Beweis zu Lemma 17.10 zeigt man

$$M_B(x) \sim H(x) := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & h_1 & & \\ & & & & \ddots & \\ & & & & & h_k \end{pmatrix}.$$

Es ist also zu zeigen, dass $H(x)$ die selben Elementarteiler wie $M_{B_g}(x)$ hat, nämlich g. Nun sind die Elementarteiler von $H(x)$ per Definiton

genau die Elementarteiler von $K[x]^n/\langle H(x)\rangle$. Es gilt

$$\begin{aligned} K[x]^n/\langle H(x)\rangle &\simeq K[x]/(h_1(x)) \oplus \ldots \oplus K[x]/(h_r(x)) \\ &\simeq K[x]/(g(x)). \end{aligned}$$

Die zweite Isomorphie folgt hierbei aus dem Chinesischen Restsatz. An dieser Stelle geht die paarweise Teilerfremdheit der Polynome h_i ein. Also hat $H(x)$ nach Satz 16.16 den Elementarteiler g. □

Theorem 17.14. *(Weierstraßsche Normalform) Sei $A \in M_n(K)$. Dann gibt es ein bis auf Reihenfolge eindeutig bestimmtes System $h_1, \ldots, h_m$ von Potenzen normierter irreduzibler Polynome, so dass A zu der Matrix*

$$B_{h_1,\ldots,h_m} = \begin{pmatrix} B_{h_1} & & \\ & \ddots & \\ & & B_{h_m} \end{pmatrix}$$

ähnlich ist. Die Polynome $h_1, \ldots, h_r$ sind hierbei genau die Invariantenteiler der charakteristischen Matrix $M_A(x)$.

Beweis. Seien $g_1, \ldots, g_r$ die Elementarteiler von A. Nach dem Frobeniusschen Normalformensatz gilt

$$A \approx \begin{pmatrix} B_{g_1} & & \\ & \ddots & \\ & & B_{g_m} \end{pmatrix} \tag{4}$$

Schreibe nun wie beim Übergang von Elementarteilern zu Invariantenteilern (siehe Lemma 16.23) $g_j = h_{j,1} \cdots h_{j,s_j}$ mit Potenzen von paarweise verschiedenen irreduziblen Polynomen. Dann gilt wegen Lemma 17.13

$$B_{g_j} \approx \begin{pmatrix} B_{h_{j,1}} & & \\ & \ddots & \\ & & B_{h_{j,s_j}} \end{pmatrix}$$

Zusammensetzen gemäß (4) vervollständigt den Beweis. □

Beispiel 17.15. (1) In Beispiel 17.12 ist die Weierstraßsche Normalform gegeben durch

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

(2) Seien

$$\begin{aligned} g_1(x) &= x - 1, \\ g_2(x) &= (x-1)(x-2) = x^2 - 3x + 2, \\ g_3(x) &= (x-1)(x-2)(x-3) = x^3 - 6x^2 + 11x - 6 \end{aligned}$$

die Elementarteiler einer Matrix $A \in M_6(\mathbb{Q})$. Dann ist die Frobeniussche Normalform gegeben durch

$$\left(\begin{array}{c|cc|ccc} 1 & & & & & \\ \hline & 0 & -2 & & & \\ & 1 & 3 & & & \\ \hline & & & 0 & 0 & 6 \\ & & & 1 & 0 & -11 \\ & & & 0 & 1 & 6 \end{array}\right)$$

Die Weierstraßsche Normalform ist gegeben durch

$$\left(\begin{array}{c|cc|ccc} 1 & & & & & \\ \hline & 1 & 0 & & & \\ & 0 & 2 & & & \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 2 & 0 \\ & & & 0 & 0 & 3 \end{array}\right)$$

(3) Seien

$$\begin{aligned} g_1(x) &= (x+1)^2 = x^2 + 2x + 1, \\ g_2(x) &= x(x+1)^2 = x^3 + 2x^2 + x. \end{aligned}$$

die Elementarteiler einer Matrix $A \in M_6(\mathbb{Q})$. Dann ist die Frobeniussche Normalform gegeben durch

$$\left(\begin{array}{cc|ccc} 0 & -1 & & & \\ 1 & -2 & & & \\ \hline & & 0 & 0 & 0 \\ & & 1 & 0 & -1 \\ & & 0 & 1 & -2 \end{array}\right)$$

Die Weierstraßsche Normalfrom ist gegeben durch

$$\left(\begin{array}{cc|ccc} 0 & -1 & & & \\ 1 & -2 & & & \\ \hline & & 0 & & \\ & & \hline & 0 & -1 \\ & & & 1 & -2 \end{array}\right)$$

Falls K ein algebraisch abgeschlossener Körper ist (z.B. $K = \mathbb{C}$), so kommen als Invariantenteiler nur Potenzen von linearen Polynomen vor. Sei also

$$h(x) = (x - \alpha)^e, \quad e \geq 1.$$

Lemma 17.16. *Sei K beliebig und $h(x) = (x-\alpha)^e, e \geq 1$. Dann gilt:*

$$B_h \approx J(\alpha, e) := \begin{pmatrix} \alpha & & & & 0 \\ 1 & \alpha & & & 0 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \alpha & 0 \\ & & & 1 & \alpha \end{pmatrix}.$$

Für $e = 1$ ist $J(\alpha, e) = (\alpha)$.

Beweis. Mit dem Algorithmus aus dem Elementarteilersatz zeigt man

$$M_{J(\alpha,e)}(x) \sim \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & (x-\alpha)^e \end{pmatrix}$$

(siehe Übung). Mit Satz 17.4 folgt dann die Behauptung. □

Matrizen der Form $J(\alpha, e)$ nennen wir eine Jordanmatrix oder auch ein Jordankästchen.

Theorem 17.17. *(Jordansche Normalform) Sei $A \in M_n(K)$ und das charakteristische Polynom $\chi_A(x)$ zerfalle vollständig in Linearfaktoren. Dann gibt es ein bis auf die Reihenfolge eindeutig bestimmtes System von Jordanmatrizen $J_1, \ldots, J_m$, so dass*

$$A \approx \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}$$

Beweis. Wegen Lemma 17.6 ist

$$\chi_A(x) = \prod_{i=1}^{s} c_i(x).$$

Also zerfallen die Elementarteiler vollständig in Linearfaktoren. Die Invariantenteiler h_j von A sind Teiler der $c_i(x)$ und zerfallen daher ebenfalls vollständig in Linearfaktoren, d.h.

$$h_j(x) = (x-\alpha_j)^{e_j}. \tag{5}$$

Damit folgt die Behauptung aus Lemma 17.16. □

Bemerkung 17.18. (1) Wegen

$$\chi_A(x) = \prod_j h_j(x) = \prod_j (x-\alpha_j)^{e_j}$$

sind die α_j genau die Eigenwerte von A.

(2) Aus der Jordanschen Normalform lässt sich auch das Minimalpolynom direkt ablesen. Sortiere dazu die Jordankästchen nach Eigenwerten und Größe,

$$\begin{pmatrix} J(\alpha_1, e_{1,1}) & & & & & & & \\ & \ddots & & & & & & \\ & & J(\alpha_1, e_{1,n_1}) & & & & & \\ & & & \ddots & & & & \\ & & & & \ddots & & & \\ & & & & & J(\alpha_s, e_{s,1}) & & \\ & & & & & & \ddots & \\ & & & & & & & J(\alpha_s, e_{1,n_s}) \end{pmatrix}$$

mit $e_{i,1} \leq e_{i,2} \leq \ldots \leq e_{i,n_i}$. Dann gilt:

$$\mu_A(x) = \prod_{i=1}^{s} (x - \alpha_i)^{e_{i,n_i}}.$$

(3) Falls $\chi_A(x)$ vollständig in Linearfaktoren zerfällt, so sind die folgenden drei Aussagen äquivalent:

(i) A ist diagonalisierbar.
(ii) Die Jordansche Normalform ist eine Diagonalmatrix.
(iii) Jedes Jordankästchen hat Rahmengröße 1.

Beweis: Übung.

Beispiel 17.19. (1) In Beispiel 17.15(2) ist die Weierstraßsche Normalform gleich der Jordanschen Normalform.

(2) In Beispiel 17.15(3) ist die Jordansche Normalform gegeben durch

$$\left(\begin{array}{cc|ccc} -1 & 0 & & & \\ 1 & -1 & & & \\ \hline & & 0 & & \\ \hline & & & -1 & 0 \\ & & & 1 & -1 \end{array}\right)$$

18. Innere Produkträume

Wir betrachten K-Vektorräume V, wobei $K = \mathbb{R}$ oder $K = \mathbb{C}$ ist; unser Ziel ist, einen Begriff der Länge sowie des Winkels zwischen zwei Vektoren einzuführen. Wir verwenden dazu Funktionen $\phi : V \times V \to K$, die ein sogenanntes inneres Produkt auf V definieren. Zum Beispiel: Ist $K = \mathbb{R}^2$, $v_1 = (\alpha_1, \alpha_2)$ und $v_2 = (\beta_1, \beta_2)$, so definiert

$$\phi : \mathbb{R}^2 \times \mathbb{R}^2 \to K, \ \phi(v_1, v_2) := \alpha_1\beta_1 + \alpha_2\beta_2$$

eine solche Abbildung. Meist werden wir $(v_1, v_2) := \phi(v_1, v_2)$ schreiben, wobei die Bedeutung stets aus dem Kontext hervorgeht. Oftmals wird in der Literatur auch die Schreibweise $\langle v_1, v_2 \rangle := \phi(v_1, v_2)$ verwendet.

Ein offensichtlicher Begriff der Länge von v_1 ist

$$\| v_1 \| = \sqrt{\alpha_1^2 + \alpha_2^2} = \sqrt{(v_1, v_1)} \geq 0,$$

und der Winkel γ zwischen zwei Vektoren $v_1 \neq 0 \neq v_2$ ist durch

$$(v_1, v_2) = cos(\gamma) \| v_1 \| \| v_2 \|$$

bestimmt. Die analogen Formeln gelten für $\mathbb{R}^n$.

Im folgenden betrachten wir nur Vektorräume über den Körpern $\mathbb{R}$ oder $\mathbb{C}$; wir bezeichnen einen solchen Vektorraum als $\mathbb{K}$-Vektorraum.

Definition 18.1. Sei V ein $\mathbb{K}$-Vektorraum. Ein inneres Produkt (oder Skalarprodukt) auf V ist eine Abbildung $\phi\colon V \times V \longrightarrow \mathbb{K}$, die einem Paar von Vektoren $v_1, v_2 \in V$ einen Skalar $(v_1, v_2) := \phi(v_1, v_2)$ zuordnet, so dass folgende Regeln gelten

(1) $(v_1 + v_2, v_3) = (v_1, v_3) + (v_2, v_3)$,
(2) $(\alpha v_1, v_2) = \alpha(v_1, v_2),\ \alpha \in K$,
(3) $(v_2, v_1) = \overline{(v_1, v_2)}$ (komplexe Konjugation),
(4) $(v_1, v_1) \geq 0$ und $(v_1, v_1) = 0 \Leftrightarrow v_1 = 0$.

Ein endlich-dimensionaler $\mathbb{R}$-Vektorraum (bzw. $\mathbb{C}$-Vektorraum) V, zusammen mit einem fest gewählten inneren Produkt ist ein euklidischer (bzw. unitärer) Vektorraum.

• Aus (1)-(3) folgt, dass für jedes innere Produkt gilt

(5) $(v_1, \alpha v_2 + v_3) = \overline{\alpha}(v_1, v_2) + (v_1, v_3)$.

• Die Eigenschaften (1), (2) und (5) bestimmen eine sogenannte Sesquilinearform (d.h., additiv in beiden Variablen, linear in der ersten und semilinear in der zweiten Variable). Man nennt eine Sesquilinearform hermitesch, wenn (3) gilt. Eine hermitesche Sesquilinearform nennt man positiv definit, wenn (4) gilt.

• Für alle $v \in V$ gilt stets $(0, v) = (v, 0) = 0$. Dies folgt aus $(0, v) = (w - w, v) = (w, v) - (w, v) = 0$, wobei hier $w \in V$ beliebig ist.

• Wegen der hermiteschen Eigenschaft (3) folgt für alle $v \in V$ die Gleichung $(v, v) = \overline{(v, v)}$ und damit $(v, v) \in \mathbb{R}$. Die Bedingung in (4) ergibt also Sinn.

• Sei $K = \mathbb{C}$, so schreibe für $v_1, v_2 \in V$ den Skalar (v_1, v_2) als

$$(v_1, v_2) = Re(v_1, v_2) + iIm(v_1, v_2).$$

Für eine komplexe Zahl z gilt $Im(z) = Re(-iz)$, also ist $Im(v_1, v_2) = Re(-i(v_1, v_2)) = Re(v_1, iv_2)$, und es folgt

$$(v_1, v_2) = Re(v_1, v_2) + Re(v_1, iv_2),$$

d.h. das innere Produkt ist durch seinen 'Realteil' $Re(\ ,\ )$ bestimmt.

Beispiele 18.2. (a) Sei $\mathbb{K} = \mathbb{C}$ und $V = \mathbb{C}^n$. Sind $v_1 = (\alpha_1, \ldots, \alpha_n)$ und $v_1 = (\beta_1, \ldots, \beta_n)$ in V, so definiert

$$(v_1, v_2) = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$$

ein inneres Produkt. Im Fall $\mathbb{K} = \mathbb{R}$ liefert die entsprechende Formel

$$(v_1, v_2) = \sum_{i=1}^{n} \alpha_i \beta_i$$

das übliche euklidische Skalarprodukt. Wir nennen diese inneren Produkte auch das Standardskalarprodukt auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.

(b) Sei V der Vektorraum der stetigen komplexwertigen (oder reellwertigen) Funktionen auf dem Einheitsintervall $[0, 1]$. Dann definiert

$$(f, g) = \int_0^1 f(t)\overline{g(t)}dt$$

ein inneres Produkt auf V. Die Stetigkeit geht beim Nachweis von (4) ein.

(c) Ist $V = \mathbb{R}^2$, $v_1 = (\alpha_1, \alpha_2)$ und $v_1 = (\beta_1, \beta_2)$ und $B = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$. Dann definiert

$$(v_1, v_2) := v_1^t B v_2 = \alpha_1\beta_1 - \alpha_2\beta_1 - \alpha_1\beta_2 + 4\alpha_2\beta_2$$

ein inneres Produkt auf V. Wegen $(v_1, v_1) = (\alpha_1 - \alpha_2)^2 + 3\alpha_2^2$ gilt $(v_1, v_1) \geq 0$ und $(v_1, v_1) = 0 \Leftrightarrow v_1 = 0$; die Bedingungen (1)-(3) sind offensichtlich.

(d) Sei $V = M_n(K)$ der $\mathbb{K}$-Vektorraum aller $n \times n$-Matrizen. Für Matrizen $A, B \in V$ definieren wir

$$(A, B) := \mathrm{Tr}(AB^*),$$

wobei $B^* := \overline{B}^t$. Die Eigenschaften eines inneren Produkt lassen sich leicht nachrechnen. Zum Beispiel erhält man die positive Definitheit aus

$$(A, A) = \mathrm{Tr}(AA^*) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}\bar{\alpha}_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} |\alpha_{ij}|^2,$$

wobei hier $A = (\alpha_{ij})$ ist.

Lemma 18.3. *Seien V, W $\mathbb{K}$-Vektorräume, und sei $(\ ,\ )$ ein inneres Produkt auf W. Ist $A : V \to W$ eine injektive lineare Abbildung, so definert $p_A(v_1, v_2) = (Av_1, Av_2)$ ein inneres Produkt auf V.*

Beweis. Leichtes Nachrechnen; die Injektivität geht beim Nachweis von (4) ein. □

Beispiele 18.4. (a) Sei V ein $\mathbb{K}$-Vektorraum mit Basis $a_1, \ldots, a_n$. Ist $e_1, \ldots, e_n$ die Standardbasis von $\mathbb{K}^n$ und ist $A : V \to K^n$ die lineare Abbildung mit $a_i \mapsto e_i,\ i = 1, \ldots, n$ (d.h. A ist der 'natürliche' Isomorphismus $V \cong K^n$ bzgl. der gegebenen Basis), so ist

$$p_A(\sum_j \alpha_j a_j, \sum_k \beta_k a_k) = \sum_{j=1}^{n} \alpha_j \overline{\beta}_j.$$

Insbesondere gibt es zu jeder Basis von V ein inneres Produkt auf V mit $(a_i, a_j) = \delta_{ij}$; man kann leicht sehen, dass es genau ein solches inneres Produkt gibt; wir werden zeigen, dass jedes innere Produkt auf V das innere Produkt einer Basis ist.

(b) Sei V der Vektorraum von Beispiel 18.2(d). Ist $W = V$, und ist $A : V \to W$ die Abbildung 'Multiplikation mit t', d.h. $(Af)(t) = tf(t)$, $0 \le t \le 1$, so ist A linear (Nachrechnen) und injektiv (ist $Af = 0$, so ist $tf(t) = 0$ für $t > 0$; da f stetig ist, folgt $f(0) = 0$, und somit $f = 0$). Mit Lemma 18.3 ergibt sich ein inneres Produkt

$$p_A(f, g) = \int_0^1 tf(t)\overline{tg(t)}dt = \int_0^1 f(t)\overline{g(t)}t^2 dt.$$

Definition 18.5. Sei V ein $\mathbb{K}$-Vektorraum mit einem inneren Produkt $(\ ,\ )$. Für $v \in V$ definiere die Norm (oder 'Länge') von v als

$$\| v \| = \sqrt{(v, v)} \ge 0 \text{ (positive Quadratwurzel)} .$$

Lemma 18.6. *Sei V ein $\mathbb{K}$-Vektorraum mit einem inneren Produkt $(\ ,\ )$. Sind $v_1, v_2 \in V$ und $\alpha \in \mathbb{K}$, so gilt*

(a) $\| \alpha v_1 \| = |\alpha| \, \| v_1 \|$,

(b) $\| v_1 \| > 0$ *für* $0 \ne v_1$,

(c) $|(v_1, v_2)| \ \le \| v_1 \| \| v_2 \|$ *(Cauchy-Schwarz Ungleichung),*

(d) $\| v_1 + v_2 \| \le \| v_1 \| + \| v_2 \|$ *(Dreiecksungleichung).*

Beweis. (a) und (b) folgen leicht aus den Definitionen. Betrachte (c): Ist $v_1 = 0$, so ist $(0, v_2) = 0$, d.h. (c) gilt in diesem Fall (für jedes v_2). Ist $v_1 \ne 0$, so setze

$$v_3 = v_2 - \frac{(v_2, v_1)}{\| v_1 \|^2} v_1.$$

Dann ist $(v_3, v_1) = 0$ und (c) folgt unter Verwendung von $z\overline{z} = |z|^2$ aus der Ungleichung

$$\begin{aligned} 0 &\leq \| v_3 \|^2 = (v_3, v_3) = (v_3, v_2) \\ &= (v_2, v_2) - \frac{\overline{(v_2, v_1)}}{\| v_1 \|^2}(v_2, v_1) = \| v_2 \|^2 - \frac{|(v_1, v_2)|^2}{\| v_1 \|^2}. \end{aligned}$$

Mittels (c) folgt mit $z + \overline{z} = 2Re(z)$ und $Re(z) \leq |z|$

$$\begin{aligned} \| v_1 + v_2 \|^2 &= \| v_1 \|^2 + (v_1, v_2) + (v_2, v_1) + \| v_2 \|^2 \\ &= \| v_1 \|^2 + 2Re(v_1, v_2) + \| v_2 \|^2 \\ &\leq \| v_1 \|^2 + 2|(v_1, v_2)| + \| v_2 \|^2 \\ &\overset{(c)}{\leq} \| v_1 \|^2 + 2 \| v_1 \| \| v_2 \| + \| v_2 \|^2 \\ &= (\| v_1 \| + \| v_2 \|)^2. \end{aligned}$$

Dies zeigt $\| v_1 + v_2 \| \leq \| v_1 \| + \| v_1 \|$ und beweist (d). □

Definition 18.7. Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt $(\, , \,)$. Zwei Vektoren $v_1, v_2 \in V$ sind orthogonal, falls $(v_1, v_2) = 0$ ist. Eine Menge $M \subseteq V$ von Vektoren ist eine orthogonale Menge, falls je zwei verschiedene Vektoren in M orthogonal sind. Eine orthogonale Menge M ist orthonormal, falls $\| v \| = 1$ für alle $v \in M$.

• Der Nullvektor 0 ist orthogonal zu jedem Vektor in V, und der einzige Vektor mit dieser Eigenschaft.

• Eine orthonormale Menge besteht aus paarweise aufeinander senkrechten Vektoren der Länge 1.

Beispiele 18.8. (a) Die Standardbasen von $\mathbb{R}^n$ und $\mathbb{C}^n$ sind orthonormale Mengen bzgl. dem Standardskalarprodukt.

(b) Sei $V = \mathbb{C}^{n \times n}$ und sei $E_{ij} \in V$ die Matrix mit 1 an der Stelle (i, j) und 0 sonst. Die Menge $M = \{E_{ij} \mid i, j = 1, \ldots, n\}$ ist eine orthonormale Menge bzgl. des inneren Produkts von Beispiel 18.2(c), da

$$(E_{ij}, E_{rs}) = tr(E_{ij} E_{sr}) = \delta_{js} \delta_{ir}.$$

Lemma 18.9. *Sei V ein $\mathbb{K}$-Vektorraum mit inneren Produkt $(\,)$. Eine orthogonale Menge $M \subseteq V$ von nicht-trivialen Vektoren ist linear unabhängig.*

NB. Der Beweis zeigt: Sind $v_1, \ldots, v_n$ nicht-triviale orthogonale Vektoren und v eine Linearkombination der v_i, so gilt

$$v = \sum_{i=1}^{n} \frac{(v, v_i)}{\| v_i \|^2} v_i.$$

Beweis. Sei $M \subseteq V$ eine orthogonale Menge von nicht-trivialen Vektoren, und seien $v_1, \ldots, v_n$ verschiedene Vektoren in M. Angenommen $v = \sum_{i=1}^{n} \alpha_i v_i$. Da die $v_1, \ldots, v_n$ orthogonal sind, folgt

$$(v, v_k) = (\sum_i \alpha_i v_i, v_k) = \sum_i \alpha_i (v_i, v_k) = \alpha_k (v_k, v_k).$$

Nach Annahme ist $v_k \neq 0$, also ist $(v_k, v_k) = \| v_k \|^2 \neq 0$ und deshalb

$$\alpha_k = \frac{(v, v_k)}{\| v_k \|^2}, \quad \text{für } 1 \leq k \leq n.$$

Ist $v = 0$, so ist $(v, v_k) = (0, v_k) = 0$ und $\alpha_k = 0$ für alle $k = 1, \ldots, n$. Also ist M eine linear unabhängige Menge. □

Theorem 18.10. *(Gram-Schmidt) Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit einem inneren Produkt. Dann hat V eine orthonormale Basis.*

NB. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $a_1, \ldots, a_n$ eine Basis von V. Das mittels $a_1, \ldots, a_n$ definierte innere Produkt (vgl. 18.4)

$$(\sum_j \alpha_j a_j, \sum_k \beta_k a_k) = \sum_{j=1}^{n} \alpha_j \overline{\beta}_j$$

hat die Eigenschaft, dass die gegebene Basis eine Orthonormalbasis bzgl. dieses inneren Produkts ist. Theorem 18.10 besagt, dass jedes innere Produkt auf V von dieser Form ist: Ist $(\ ,\ )$ ein inneres Produkt auf V, so gibt es eine Basis $b_1, \ldots, b_n$, die bzgl. $(\ ,\ )$ orthonormal ist. Wegen $(b_i, b_j) = \delta_{ij}$ gilt dann $(\sum_j \alpha_j b_j, \sum_k \beta_k b_k) = \sum_{j=1}^{n} \alpha_j \overline{\beta}_j$.

Beweis. Der Beweis ist algorithmisch. Sei $b_1, \ldots, b_n$ eine Basis von V. Wir konstruieren zunächst aus $b_1, \ldots, b_n$ eine orthogonale Basis $a_1, \ldots, a_n$. Setze dazu $a_1 := b_1$. Wir nehmen an, dass wir bereits nicht-triviale $a_1, \ldots, a_k$ für $k < n$ konstruiert haben mit den Eigenschaften

(i) $(a_i, a_j) = 0, 1 \leq i, j \leq k, i \neq j$
(ii) $a_i \in \langle b_1, \ldots, b_k \rangle, i = 1, \ldots, k.$

Setze nun

$$a_{k+1} := b_{k+1} - \sum_{i=1}^{k} \lambda_i a_i$$

mit noch zu bestimmenden $\lambda_i \in \mathbb{K}$. Aus den Bedingungen $(a_{k+1}, a_j) = 0$ erhält man

$$\lambda_j = \frac{(b_{k+1}, a_j)}{\| a_j \|^2}.$$

Nach Konstruktion sind (i) und (ii) automatisch erfüllt. Wäre $a_{k+1} = 0$, so würde folgen: $b_{k+1} \in \langle a_1, \ldots, a_k \rangle \subseteq \langle b_1, \ldots, b_k \rangle$ im Widerspruch dazu, dass $b_1, \ldots, b_n$ eine Basis ist.

Auf diese Weise konstruieren wir eine orthogonale Menge $\{a_1, \ldots, a_n\}$, die nach Lemma 18.9 linear unabhängig ist. Insbesondere ist $\{a_1, \ldots, a_n\}$ eine Basis.

Abschließend normieren wir die a_i auf Länge 1, d.h. wir setzen

$$a_i' := \frac{a_i}{\| a_i \|}.$$

Dann ist $a_1', \ldots, a_n'$ eine Orthonormalbasis.

□

Definition 18.11. Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Ist $M \subseteq V$ eine Teilmenge, so ist das orthogonale Komplement von M definiert durch

$$M^\perp := \{v \in V \mid (v, m) = 0 \text{ für alle } m \in M\} \subseteq V.$$

• Für jede Teilmenge $M \subseteq V$ ist $M^\perp \subseteq V$ ist ein $\mathbb{K}$-linearer Unterraum.

Lemma 18.12. *Sei V ein $\mathbb{K}$-Vektorraum mit einem inneren Produkt. Ist $U \subseteq V$ ein linearer Unterraum mit $\dim_{\mathbb{K}} U < \infty$, so gilt*

(a) $V = U \oplus U^\perp$,

(b) $(U^\perp)^\perp = U$.

Wir schicken dem Beweis folgende Vorbemerkung voraus:

Bemerkung 18.13. Sei W ein $\mathbb{K}$-Vektorraum der Dimension $n < \infty$ mit innerem Produkt. Sei $w_1, \ldots, w_n$ eine Orthonormalbasis von W und $w \in W$. Dann gilt:

$$w = \sum_{i=1}^{n} (w, w_i) w_i.$$

Beweis. Sei $w = \sum_{i=1}^n \alpha_i w_i$. Dann folgt

$$(w, w_j) = (\sum_{i=1}^{n} \alpha_i w_i, w_j) = \sum_{i=1}^{n} \alpha_i (w_i, w_j) = \alpha_j.$$

□

Beweis. (zu Lemma 18.12)(a): Wir zeigen zunächst $V = U + U^\perp$. oE sei dazu $v \in V \setminus U$. Nach Theorem 18.10 hat der Unterraum U eine Orthonormalbasis $a_1, \ldots, a_m$. Ergänze nach dem Verfahren von Cram-Schmidt $a_1, \ldots, a_n$ zu einer Orthonormalbasis $a_1, \ldots, a_n, w$ von $\langle U, v \rangle$.

Insbesondere ist $w \in U^\perp$. Nach der Vorbemerkung ist

$$v = \sum_{i=1}^{n} (v, a_i)a_i + (v, w)w \in U + U^\perp.$$

Es ist nun noch zu zeigen: $U \cap U^\perp = \emptyset$. Sei dazu $v \in U \cap U^\perp$. Dann folgt $(v, v) = 0$, also $v = 0$.

(b): Nach Definition ist

$$(U^\perp)^\perp = \{v \in V \mid (v, w) = 0 \text{ für alle } w \in U^\perp\}.$$

Die Inklusion $U \subseteq (U^\perp)^\perp$ ist daher offensichtlich. Sei umgekehrt $v \in (U^\perp)^\perp$. Schreibe wie im Beweis zu Teil (a)

$$v = \sum_{i=1}^{n} (v, a_i)a_i + (v, w)w \in U + U^\perp.$$

Dann ist $(v, w) = 0$ und daher $v \in U$. □

19. Lineare Funktionale und Adjungierte

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit einem inneren Produkt $(\ ,\ )$. Wir charakterisieren die linearen Funktionale auf V, d.h. die Elemente von $V^* := Hom_{\mathbb{K}}(V, \mathbb{K})$, und definieren zu jedem Endomorphismus $f : V \to V$ eine Abbildung $f^* : V \to V$, so dass gilt

$$(f(v_1), v_2) = (v_1, f^*(v_2)).$$

Gilt $f = f^*$, so ist f selbstadjungiert; die selbstadjungierten Endomorphism bilden eine wichtige Klasse von linearen Abbildungen mit speziellen Eigenschaften.

Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Da das innere Produkt linear in der ersten Variable ist, definiert für ein festes $w \in V$

$$f_w : V \to \mathbb{K},\ v \mapsto f_w(v) := (v, w)$$

ein Element von $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. Das nächste Lemma besagt, dass alle linearen Funktionale auf V von dieser Form sind.

Lemma 19.1. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$, und sei $f \in V^* = Hom_{\mathbb{K}}(V, \mathbb{K})$. Dann gibt es einen eindeutig bestimmten Vektor $w \in V$, so dass gilt*

$$f(v) = f_w(v) = (v, w), \quad v \in V.$$

Beweis. Sei $\{a_1, \ldots, a_n\}$ eine Orthonormalbasis von V. Setze

$$w = \sum_{i=1}^{n} \overline{f(a_i)} \cdot a_i.$$

Für das durch den Vektor w definierte lineare Funktional f_w gilt

$$f_w(a_j) = (a_j, \sum_{i=1}^{n} \overline{f(a_i)} a_i) = (a_j, \overline{f(a_j)} a_j) = f(a_j);$$

da f und f_w auf einer Basis übereinstimmen, folgt $f = f_w$. Ist $w' \in V$ ein weiterer Vektor mit $f(v) = f_w(v) = f_{w'}(v)$, so gilt $(v, w) = (v, w')$ und damit $(v, w - w') = 0$ für $v \in V$; der Spezialfall $v = w - w'$ liefert $(w - w', w - w') = 0$, also ist $w = w'$. □

Beispiel 19.2. Das folgende Beispiel zeigt, dass die die Aussage von Lemma 19.1 für unendlich-dimensionale $\mathbb{K}$-Vektorräume im Allgemeinen nicht gilt. Sei $V = \mathbb{C}[x]$ der $\mathbb{C}$-Vektorraum der Polynome, zusammen mit dem durch die Formel

$$(p, q) = \int_0^1 p(t)\overline{q(t)}dt$$

definierten inneren Produkt. Sei $z \in \mathbb{C}$ fest gewählt und $L \in V^*$ die Abbildung $p \mapsto p(z)$. Angenommen es gibt ein $q \in V$, so dass gilt

$$L(p) = (p, q), \quad p \in V.$$

Betrachte $h(x) := x - z \in V$. Für jedes $p \in V$ ist $h(z)p(z) = 0$, so dass

$$0 = L(hp) = (hp, q) = \int_0^1 h(t)p(t)\overline{q(t)}dt.$$

Dies gilt insbesondere für $p = \overline{h}q$; es folgt

$$0 = \int_0^1 \overline{h(t)}h(t)q(t)\overline{q(t)}dt = \int_0^1 |h(t)|^2|q(t)|^2dt$$

und weiter $hq = 0$. Da $h \neq 0$ ist, ist dann $q = 0$, und die Annahme $L(p) = (p, q)$ impliziert $L : p \mapsto p(z) = (p, 0) = 0$ ist das triviale lineare Funktional, Widerspruch.

Proposition 19.3. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit einem inneren Produkt $(\ ,\ )$, und sei $f \in End_{\mathbb{K}}(V)$. Dann gibt es einen eindeutig bestimmten Endomorphismus $f^* \in End_{\mathbb{K}}(V)$, so dass*

$$(f(v_1), v_2) = (v_1, f^*(v_2)), \quad v_1, v_2 \in V.$$

Beweis. Sei $v \in V$ ein Vektor. Dann definiert $g(w) = (f(w), v)$ ein lineares Funktional auf V, und nach Lemma 19.1 gibt es einen eindeutig bestimmten Vektor $v' \in V$ mit $g(w) = f_{v'}(w) = (w, v')$ für alle $w \in V$. Setze $f^*(v) = v'$. Diese Konstruktion bestimmt f^* eindeutig und liefert

$$(fw, v) = g(w) = f_{v'}(w) = (w, v') = (w, f^*v).$$

Es bleibt zu zeigen, dass f^* linear is. Für $v_1, v_2, v_3 \in V$ ist

$$(v_1, f^*(v_2 + v_3)) = (f(v_1), v_2 + v_3) = (f(v_1), v_2) + (f(v_1), v_3) = \\ = (v_1, f^*(v_2)) + (v_1, f^*(v_3)) = (v_1, f^*(v_2) + f^*(v_3)),$$

Damit ist $0 = (v_1, f^*(v_2 + v_3) - (f^*(v_2) + f^*(v_3)))$ für $v_1 \in V$; dies gilt insbesondere für $v_1 = f^*(v_2+v_3)-(f^*(v_2)+f^*(v_3))$, so dass $f^*(v_2+v_3) = f^*(v_2) + f^*(v_3)$. Ist $\alpha \in \mathbb{K}$ ein Skalar, so folgt aufgrund von

$$(v_1, f^*(\alpha v_2)) = (f(v_1), \alpha v_2) = \overline{\alpha}(f(v_1), v_2)) = \\ \overline{\alpha}(v_1, f^*(v_2)) = (v_1, \alpha f^*(v_2))$$

genauso $f^*(\alpha v_2) = \alpha f^*(v_2)$. □

Lemma 19.4. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Dann gilt:*

(a) *Ist $f \in End_{\mathbb{K}}(V)$, und ist $A = (\alpha_{ij})$ die Matrix von f bzgl. einer Orthonormalbasis $\{a_1, \ldots, a_n\}$, so ist $\alpha_{ij} = (f(a_j), a_i)$.*

(b) *Hat f bzgl. einer Orthonormalbasis die Matrix $A = (\alpha_{ij})$, so hat f^* bzgl. dieser Basis die Matrix $A^* = (\overline{\alpha}_{ji})$ (komplex konjugiert Transponierte).*

Beweis. (a): Da $a_1, \ldots, a_n$ einen Orthonormalbasis ist, gilt für $v \in V$

$$v = \sum_{i=1}^{n} (v, a_i) a_i,$$

siehe Bemerkung 18.13. Die Matrix A ist definiert durch

$$f(a_j) = \sum_{i=1}^{n} \alpha_{ij} a_i.$$

Aus der ersten Formel ergibt sich $f(a_j) = \sum_{i=1}^{n} (f(a_j), a_i) a_i$, und Koeffizientenvergleich mit der zweiten Formel zeigt $\alpha_{ij} = (f(a_j), a_i)$.

(b): Sei $a_1, \ldots, a_n$ eine Orthonormalbasis. Sind $A = (\alpha_{ij})$ und $A^* = (\beta_{ij})$ die Matrizen von f und f^* bzgl. dieser Basis, so gilt nach (a)

$$\alpha_{ij} = (f(a_j), a_i) \text{ und } \beta_{ij} = (f^*(a_j), a_i).$$

Nach Definition von f^* gilt für diese Einträge

$$\beta_{ij} = (f^*(a_j), a_i) = \overline{(a_i, f^*(a_j))} = \overline{(f(a_i), a_j)} = \overline{\alpha_{ji}}.$$

□

Definition 19.5. Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$ (nicht unbedingt endlich-dimensional), und sei $f \in End_{\mathbb{K}}(V)$. Dann hat f eine Adjungierte f^*, falls es ein $f^* \in End_{\mathbb{K}}(V)$ gibt, so dass

$$(f(v_1), v_2) = (v_1, f^*(v_2)) \text{ für alle } v_1, v_2 \in V.$$

- Ist $\dim_{\mathbb{K}} V < \infty$, so hat jedes $f \in End_{\mathbb{K}}(V)$ eine Adjungierte.
- Existiert eine Adjungierte f^*, so ist diese eindeutig bestimmt.

Beispiele 19.6. (a) Sei $V = \mathbb{K}^{n\times n}$. Für $A = (\alpha_{ij}) \in V$ bezeichne $A^* = (\overline{\alpha}_{ji}) \in V$ die komplex konjugiere transponierte Matrix. Betrachte V mit dem innerem Produkt $(A, B) = tr(B^*A)$ von Beispiel 18.2(c). Ist $M \in V$ eine fest gewählte Matrix, so definiert Linksmultiplikation mit M, $L_M : V \to V,\ A \mapsto MA$, eine lineare Abbildung. Es gilt

$$\begin{array}{rclcl} (L_M(A), B) & = & tr(B^*(MA)) & = & tr(MAB^*) \\ & = & tr(AB^*M) & = & tr(A(M^*B)^*) = (A, L_{M^*}(B)) \end{array}$$

(dies verwendet $tr(CD) = tr(DC)$ und $(CD)^* = D^*C^*$), also ist $(L_M)^* = L_{M^*}$.
(b) Sei $V = \mathbb{C}[x]$ mit dem innerem Produkt von Beispiel 19.2. Für $f = \sum_{i=0}^n \alpha_i x^i$, schreibe $\overline{f} = \sum_{i=0}^n \overline{\alpha}_i x^i$, d.h. $\overline{f}(t) = \overline{f(t)}$ für t reell. Für $f \in V$ fest, ist die Abbildung $M_f : V \to V,\ p \mapsto fp$ linear. Es gilt

$$(M_f(p), q) = \int_0^1 f(t)p(t)\overline{q(t)}dt = \int_0^1 p(t)\overline{[\overline{f(t)}q(t)]}dt = (p, M_{\overline{f}}(q)),$$

d.h. $M_{\overline{f}}$ ist die Adjungierte zu M_f.
(c) Sei $V = \mathbb{R}^n$ versehen mit dem Standardskalarprodukt $(x, y) = \sum_{i=1}^n x_i y_i = x^t y$. Sei $A \in M_n(\mathbb{R})$ und $f_A(x) := Ax$. Dann gilt:

$$(f_A(x), y) = (Ax, y) = x^t A^t y = (x, A^t y).$$

Also ist $f_A^* = f_{A^t}$.
(d) Sei $V = \mathbb{C}^n$ versehen mit dem Standardskalarprodukt $(x, y) = \sum_{i=1}^n x_i \overline{y_i} = x^t \overline{y}$. Dann erhält man auf analoge Weise $f_A^* = f_{A^*}$.
(e) Dieses Beispiel zeigt, dass es für unendlich dimensionale Vektorräume im Allgemeinen keine adjungierte Abbildung gibt. Sei dazu $V = \mathbb{K}[x]$ wie in Beispiel(b). Betrachte die lineare Abbildung

$$D : V \to V,\ p = \sum_{i=0}^n \alpha_i x^i \mapsto p' = \sum_{i=0}^n i\alpha_i x^{i-1}.$$

Dann gibt es zu D keine adjungierte Abbildung. Beweis: Übung.

Die Beispiele zeigen, dass der Übergang $f \mapsto f^*$ ähnliche Eigenschaften hat, wie die Konjugation von komplexen Zahlen. Genauer gilt:

Lemma 19.7. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Sind $f, g \in End_{\mathbb{K}}(V)$, und $\alpha \in \mathbb{K}$, so gilt*

(a) $(f + g)^* = f^* + g^*$,
(b) $(\alpha f)^* = \overline{\alpha} f^*$,

(c) $(fg)^* = g^* f^*$,
(d) $(f^*)^* = f$.

Beweis. Die erste Aussage (a) folgt aus

$$((f+g)(v_1), v_2) = (f(v_1) + g(v_1), v_2) = (f(v_1), v_2) + (g(v_1), v_2) =$$
$$= (v_1, f^*(v_2)) + (v_1, g^*(v_2)) = (v_1, f^*(v_2) + g^*(v_2)) = (v_1, (f^* + g^*)(v_2))$$

und der Eindeutigkeit der Adjungierten. Die zweite Behauptung (b) ist eine leichte Übung, und (c) bzw. (d) ergeben sich aus

$$((fg)(v_1), v_2) = (g(v_1), f^*(v_2)) = (v_1, g^* f^*(v_2)),$$

und

$$(v_1, f^{**}(v_2)) = (f^*(v_1), v_2) = \overline{(v_2, f^*(v_1))} = \overline{(f(v_2), v_1)} = (v_1, f(v_2))$$

□

Sei V ein $\mathbb{C}$-Vektorraum mit $\dim_{\mathbb{C}} V < \infty$ und einem innerem Produkt. Für einen Endomorphismus $f \in End_{\mathbb{C}}(V)$ setze

$$f_1 = \frac{1}{2}(f + f^*) \text{ und } f_2 = \frac{1}{2i}(f - f^*).$$

Dann sind f_1, f_2 eindeutig bestimmt, es gilt $f_1 = f_1^*,\ \ f_2 = f_2^*$, und

$$f = f_1 + i f_2,$$

d.h. f hat einen 'Realteil' und einen 'Imaginärteil'. Ist $f = f^*$, so ist $f = f_1$; die Analogie zwischen $f \mapsto f^*$ und $z \mapsto \overline{z}$ suggeriert, dass sich ein $f \in End_{\mathbb{K}}(V)$ mit $f = f^*$ "ähnlich wie eine reelle Zahl" verhält.

Definition 19.8. (a) Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt. Dann ist $f \in End_{\mathbb{K}}(V)$ selbst-adjungiert, falls eine Adjungierte existiert und $f = f^*$ ist.

(b) Eine Matrix $(\alpha_{ij}) \in \mathbb{K}^{n \times n}$ mit $\alpha_{ij} = \overline{\alpha}_{ji}$ nennt man symmetrisch (falls $\mathbb{K} = \mathbb{R}$) und hermitesch (falls $\mathbb{K} = \mathbb{C}$).

Sei V von endlicher Dimension.
• $f \in End_{\mathbb{K}}(V)$ ist genau dann selbst-adjungiert, wenn die Matrix von f bzgl. einer Orthonormalbasis symmetrisch ($\mathbb{K} = \mathbb{R}$) bzw. hermitesch ($\mathbb{K} = \mathbb{C}$) ist.

• Ist $f = f^*$, so gilt $(f(v), v) = (v, f(v)) = \overline{(f(v), v)}$, d.h. für einen selbst-adjungierten Endomorphismus f ist $(f(v), v) \in \mathbb{R}$ für alle $v \in V$.

Im Fall $\mathbb{K} = \mathbb{C}$ charakterisiert die Bedingung $(f(v), v) \in \mathbb{R}$ für $v \in V$ selbst-adjungierte Endomorphismen:

Lemma 19.9. *Sei V ein $\mathbb{C}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Sei $f \in \mathrm{End}_{\mathbb{C}}(V)$. Dann gilt:*

$$(f(v), v) \in \mathbb{R} \text{ für alle } v \in V \iff f = f^*.$$

Beweis. Übung. □

20. Positive Endomorphismen

Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt, und sei $f \in End_{\mathbb{K}}(V)$. Betrachte die Abbildung $p : V \times V \to \mathbb{K}$, die einem Paar von Vektoren v_1, v_2 den Skalar $(f(v_1), v_2)$ zuordnet. Wir möchten notwendige und hinreichende Bedingungen dafür angeben, dass p ein inneres Produkt definiert. Da $(\ ,\ )$ linear in der ersten Variable ist, ist $p(v_1, v_2)$ linear in der Variable v_1, und die Bedingungen (1) und (2) für ein inneres Produkt sind erfüllt. Die verbleibenden Bedingungen (3) und (4) sind

$$\begin{array}{lcl} p(v_1, v_2) & = & \overline{p(v_2, v_1)}, \\ p(v_1, v_1) & > & 0 \text{ für } v_1 \neq 0. \end{array}$$

Wegen $p(v_1, v_2) = (f(v_1), v_2)$, $\overline{p(v_2, v_1)} = \overline{(f(v_2), v_1)} = (v_1, f(v_2))$, und $p(v_1, v_1) = (f(v_1), v_1)$, definiert p genau dann ein inneres Produkt, wenn

$$(\#) \quad \begin{array}{lcl} (f(v_1), v_2) & = & (v_1, f(v_2)), \\ (f(v_1), v_1) & > & 0 \text{ für alle } v_1 \neq 0. \end{array}$$

Definition 20.1. Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Dann ist $f \in End_{\mathbb{K}}(V)$ positiv, wenn die Bedingungen (#) erfüllt sind.

• Nach Definiton ist $f \in End_{\mathbb{K}}(V)$ genau dann positiv, wenn $p(v_1, v_2) = (f(v_1), v_2)$ ein inneres Produkt definiert; insbesondere ist dann $(f(v), v) \in \mathbb{R}$ für alle $v \in V$.

• Die erste Bedingung in (#) besagt, dass $f = f^*$ ist. Äquivalent kann man also sagen, dass f selbst-adjungiert ist.

• Im Fall $\mathbb{K} = \mathbb{C}$ ist ein Endomorphismus f selbst-adjungiert genau dann, wenn $(f(v), v) \in \mathbb{R}$ für $v \in V$ ist, siehe Lemma 19.9. Also ist in diesem Fall f genau dann positiv, wenn $(f(v), v) > 0$ für alle $0 \neq v \in V$.

In Fall endlicher Dimension lassen sich alle inneren Produkte mittels positiver Endomorphismen beschreiben:

Lemma 20.2. *Sei V ein endlich-dimensonaler $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$. Ist ϕ ein beliebiges inneres Produkt auf V, so gibt es einen eindeutigen positiven Endomorphismus $f \in End_{\mathbb{K}}(V)$ mit*

$$\phi(v_1, v_2) = (f(v_1), v_2), \quad v_1, v_2 \in V.$$

Beweis. Sei $v_2 \in V$ fix. Dann definiert die Funktion $v_1 \mapsto \phi(v_1, v_2)$ ein lineares Funktional auf V, und nach Lemma 19.1 gibt es ein eindeutiges $v_2' \in V$ mit $\phi(v_1, v_2) = (v_1, v_2')$ für alle $v_1 \in V$. Definiere $f : V \to V$ durch $v_2 \mapsto v_2'$. Für $v_1, v_2 \in V$ gilt $\phi(v_1, v_2) = (v_1, f(v_2))$, und weiter

$$\phi(v_1, v_2) = \overline{\phi(v_2, v_1)} = \overline{(v_2, f(v_1))} = (f(v_1), v_2).$$

Man rechnet leicht nach, dass f linear ist; da ϕ ein inneres Produkt ist, ist f positiv. Ist $f' \in End_{\mathbb{K}}(V)$ ein weiterer positiver Endomorphismus mit $\phi(v_1, v_2) = (f'(v_1), v_2)$, so folgt $(f(v_1), v_2) = (f'(v_1), v_2)$ und $(f(v_1) - f'(v_1), v_2) = 0$. Für festes v_1 ist $f(v_1) - f'(v_1)$ orthogonal zu V, also der Nullvektor, dies zeigt $f = f'$. □

Bemerkung 20.3. (a) Sei $V = \mathbb{R}^n$ und $(\ ,\ )$ das Standardskalarprodukt. Dann besagt das vorherige Lemma, dass es zu jedem Skalarprodukt ϕ von V genau eine Matrix $A \in M_n(\mathbb{R})$ gibt mit $A^t = A$ und $\phi(v, w) = (Av, w) = v^t Aw$. Umgekehrt definiert jede symmetrische Matrix $A \in M_n(\mathbb{R})$ vermöge $\phi(v, w) := v^t Aw$ ein Skalarprodukt, sofern $(Av, v) = v^t Av > 0$ für alle $0 \neq v \in V$ gilt.

(b) Sei $V = \mathbb{C}^n$ und $(\ ,\ )$ das Standardskalarprodukt. Dann ist jedes innere Produkt auf V von der Form $\phi(v, w) = v^t A^t \overline{w}$ mit einer Matrix $A \in M_n(\mathbb{C})$, $A^* = A$ und $v^t A^t \overline{v} > 0$ für alle $v \neq 0$. Setzt man $B := A^t$, so sieht man das jedes innere Produkt auf V gegeben ist durch $\phi(v, w) := v^t B \overline{w}$ mit einer Matrix B, so dass $B = B^*$ und $v^t B \overline{v} > 0$ für alle $v \neq 0$.

Definition 20.4. Sei $A \in M_n(\mathbb{R})$ eine symmetrische Matrix. Dann nennt man A positiv definit, wenn für alle $0 \neq x \in V$ die Positivitätsbedingung $x^t Ax > 0$ gilt.

Proposition 20.5. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt. Dann ist $f \in End_{\mathbb{K}}(V)$ genau dann positiv, wenn es einen invertierbaren Endomorphismus $u \in End_{\mathbb{K}}(V)$ gibt, so dass gilt*

$$f = u^* u.$$

NB. Seien A und U die Matrizen von f und u bezüglich einer beliebigen Basis von V. Die Determinante der Matrix A eines positiven Endomorphismus ist positiv, denn:

$$det(A) = det(U^*)det(U) = det(\overline{U})det(U) = |det(U)|^2 > 0.$$

Beweis. $\Rightarrow$: Ist f positiv, so definiert $\phi(v_1, v_2) = (f(v_1), v_2)$ ein inneres Produkt auf V. Sei $a_1, \ldots, a_n$ eine orthonormale Basis bzgl. dem auf

V gegebenen inneren Produkt $(\ ,\ )$, und $b_1, \ldots, b_n$ eine orthonormale Basis bzgl. dem durch f definierten inneren Produkt ϕ. Es gilt also

$$\phi(b_j, b_k) = \delta_{jk} = (a_j, a_k).$$

Betrachte die lineare Abbildung $u : V \to V,\ b_j \mapsto a_j,\ j = 1, \ldots, n$. Da u eine Basis auf eine Basis abbildet, ist u invertierbar. Es gilt

$$\phi(b_j, b_k) = (a_j, a_k) = (u(b_j), u(b_k)).$$

Sind $v_1 = \sum_{j=1}^n x_j b_j$ und $v_2 = \sum_{j=1}^n y_j b_j$ zwei Vektoren in V, so folgt

$$\begin{aligned} \phi(v_1, v_2) &= \phi(\textstyle\sum_j x_j b_j, \sum_j y_j b_j) \\ &= \textstyle\sum_j \sum_k x_j \overline{y}_k \phi(b_j, b_k) \\ &= \textstyle\sum_j \sum_k x_j \overline{y}_k (u(b_j), u(b_k)) \\ &= (\textstyle\sum_j x_j u(b_j), \sum_k y_k u(b_k)) = (u(v_1), u(v_2)), \end{aligned}$$

d.h. $(f(v_1), v_2) = \phi(v_1, v_2) = (u(v_1), u(v_2)) = (u^*(u(v_1)), v_2)$ für alle $v_1, v_2 \in V$, und somit $f = u^* u$.
$\Leftarrow$: Sei $f = u^* u$ mit $u \in End_{\mathbb{K}}(V)$ invertierbar. Dann ist

$$f^* = (u^* u)^* = u^* u = f,$$

d.h. f ist selbst-adjungiert. Ist $v \in V$, so gilt $(f(v), v) = (u^*(u(v)), v) = (u(v), u(v)) \geq 0$. Ist $v \neq 0$ so folgt, da u invertierbar ist, $u(v) \neq 0$ und damit $(f(v), v) > 0$. Also ist f positiv. □

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$, und sei $f \in End_{\mathbb{K}}(V)$ ein positiver Endomorphismus. Betrachte die Matrix $A = (\alpha_{ij})$ von f bzgl. einer Orthonormalbasis $a_1, \ldots, a_n$. Da f positiv ist, ist $f = f^*$ und $(f(v), v) > 0$ für $0 \neq v$. Aus der ersten Bedingung folgt mit Lemma 19.4(b) $A = A^*$. Die zweite Bedingung besagt, dass für jeden Vektor $0 \neq v = \sum_{j=1}^n x_j a_j \in V$ gilt

$$\begin{aligned} (f(v), v) &= (\textstyle\sum_j x_j f(a_j), \sum_k x_k a_k) \\ &= \textstyle\sum_j \sum_k x_j \overline{x}_k (f(a_j), a_k) \\ &= \textstyle\sum_j \sum_k \alpha_{kj} x_j \overline{x}_k > 0. \end{aligned}$$

Also ist f genau dann positiv, wenn für die Matrix A gilt:

$$\begin{aligned} A &= A^*, \quad \text{und} \\ \textstyle\sum_j \sum_k \alpha_{kj} x_j \overline{x}_k &> 0 \quad \text{für } (x_1, \ldots, x_n) \neq 0. \end{aligned}$$

Dies ist ebenfalls äquivalent zu

$$A = A^* \text{ und } x^t A \overline{x} > 0 \text{ für alle } x \neq 0.$$

Definition 20.6. Eine Matrix $A = (\alpha_{ij}) \in \mathbb{C}^{n\times n}$ ist positiv, falls gilt: Ist $0 \neq x = (x_1, \ldots, x_n) \in \mathbb{C}$, so ist $x^t A\overline{x} \in \mathbb{R}$ und

$$x^t A\overline{x} > 0.$$

Nach Lemma 19.9 ist die erste Bedingung in der vorigen Definition äquivalent dazu, dass f selbstadjungiert ist.

Definition 20.7. Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum der Dimension n mit innerem Produkt ϕ. Sei $a_1, \ldots, a_n$ eine beliebige Basis von V. Dann heißt die Matrix

$$B = B_\phi = B_{\phi,a_1,\ldots,a_n} := (\phi(a_i, a_j))_{1\leq i,j\leq n}$$

Strukturmatrix oder Begleitmatrix von ϕ bezüglich der Basis $a_1, \ldots, a_n$.

Die Basis $a_1, \ldots, a_n$ ist genau dann orthonormal bezüglich ϕ, wenn B die Einheitsmatrix ist. Wegen $\overline{\phi(a_i, a_j)} = \phi(a_j, a_i)$ sind Strukturmatrizen stets symmetrsich bzw. hermitesch. Das folgende Lemma zeigt, dass weiter für alle $x \in \mathbb{K}^n, x \neq 0$, gilt: $x^t B\bar{x} > 0$.

Lemma 20.8. *Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum der Dimension n mit innerem Produkt ϕ. Sei $a_1, \ldots, a_n$ eine beliebige Basis von V. Seien*

$$v = \sum_{i=1}^{n} x_i a_i, \quad w = \sum_{i=1}^{n} y_i a_i$$

mit $x_i, y_i \in \mathbb{K}$. Sei B die Strukturmatrix von ϕ bezüglich der Basis $a_1, \ldots, a_n$. Dann gilt:

$$\phi(v, w) = x^t B\bar{y}.$$

Beweis. Dies ist eine einfache Rechnung. Es gilt:

$$\begin{aligned}
\phi(v, w) &= \phi(\sum_{i=1}^{n} x_i a_i, \sum_{j=1}^{n} y_j a_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \bar{y}_j \phi(a_i, a_j) \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_i \phi(a_i, a_j) \right) \bar{y}_j = x^t B\bar{y}.
\end{aligned}$$

□

Wir untersuchen das Verhalten der Strukturmatrix bei Basiswechel.

Lemma 20.9. *Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum der Dimension n mit innerem Produkt ϕ. Seien $a_1, \ldots, a_n$ und $a'_1, \ldots, a'_n$ Basen von V mit zugehörigen Strukturmatrizen B und B'. Sei $S = (s_{ij})$ die Übergangsmatrix, d.h.*

$$a'_j = \sum_{i=1}^{n} s_{ij} a_i.$$

Dann gilt:

$$B' = S^t B \bar{S}.$$

Beweis. Es gilt

$$\begin{aligned}
\phi(a'_i, a'_j) &= \phi(\sum_{k=1}^{n} s_{ki} a_k, \sum_{l=1}^{n} s_{lj} a_l) \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} s_{ki} \bar{s}_{lj} \phi(a_k, a_l) \\
&= \sum_{l=1}^{n} \left(\sum_{k=1}^{n} s_{ki} \phi(a_k, a_l) \right) \bar{s}_{lj} \\
&= \text{Eintrag an der Stelle } (i,j) \text{ von } S^t B \bar{S}.
\end{aligned}$$

□

Proposition 20.10. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt $(\ ,\ )$, und sei $a_1, \ldots, a_n$ eine orthonormale Basis. Dann ist $f \in End_{\mathbb{K}}(V)$ genau dann positiv, wenn die Matrix $A = (\alpha_{ij})$ von f bzgl. $a_1, \ldots, a_n$ positiv ist.*

Beweis. Es gilt

$$\begin{aligned}
& f \text{ ist positiv} \\
\iff\ & f = f^* \text{ und } (f(v), v) > 0 \text{ für alle } v \neq 0 \\
\iff\ & A = A^* \text{ und } (Ax)^t \bar{x} > 0 \text{ für alle } x \neq 0 \\
\iff\ & A = A^* \text{ und } x^t A^t \bar{x} > 0 \text{ für alle } x \neq 0 \\
\iff\ & A = A^* \text{ und } \bar{x^t} A x > 0 \text{ für alle } x \neq 0 \\
\iff\ & A = A^* \text{ und } x^t A \bar{x} > 0 \text{ für alle } x \neq 0.
\end{aligned}$$

□

Korollar 20.11. *Ist $A \in \mathbb{C}^{n \times n}$, so ist A positiv genau dann, wenn es eine invertierbare Matrix $U \in \mathbb{C}^{n \times n}$ mit $A = U^* U$ gibt. Hat A reelle Einträge (d.h. $A \in \mathbb{R}^{n \times n}$), so ist $U \in \mathbb{R}^{n \times n}$ und $A = U^t U$.*

Dies ist eine direkte Konsequenz aus den vorherigen Resultaten zusammen mit Lemma 20.5. Wir geben nochmals einen Matrix basierten Beweis.

Beweis. Sei A positiv. Dann ist durch $\phi(x,y) := x^t A\bar{y}$ ein inneres Produkt auf $V = \mathbb{K}^n$ gegeben. Bezüglich der Standardbasis $e_1, \ldots, e_n$ hat ϕ die Strukturmatrix A. Sei $a_1, \ldots, a_n$ eine ON-Basis bezüglich ϕ. Dann hat ϕ bezüglich dieser Basis die Einheitsmatrix E als Strukturmatrix. Sei S die Übergangsmatrix. Dann gilt $E = S^t A\bar{S}$, und hieraus folgert man

$$A = (S^t)^{-1}\bar{S}^{-1} = (S^{-1})^t\overline{S^{-1}} = U^*U$$

mit $U := \overline{S^{-1}}$. □

In der folgenden Bemerkung fassen wir den Zusammenhang zwischen positiven Matrizen und inneren Produkten nochmals zusammen.

Bemerkung 20.12. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt ϕ. Sei $a_1, \ldots, a_n$ eine beliebige Basis von V, und B die zugehörige Strukturmatrix. Dann ist B eine positive Matrix, d.h. $B = B^*$ und $x^t B\bar{x} > 0$ für alle $x \neq 0$.

Sei umgekehrt B eine positive Matrix. Dann wird durch

$$\phi(\sum_j x_j a_j, \sum_k y_k a_k) = x^t B\overline{y}$$

ein inneres Produkt auf V definiert. Inbesondere lassen sich nach Wahl einer Basis alle möglichen inneren Produkte auf V mittels positiver Matrizen beschreiben.

Beispiele 20.13. Sei $V = \mathbb{K}^n$ mit der Standardbasis $e_1, \ldots, e_n$. Sei $B = E$ die Einheitsmatrix. Dann gilt offensichtlich $B = B^*$ und

$$x^t B\bar{x} = \sum_{i=1}^{n} x_i\bar{x}_i = \sum_{i=1}^{n} |x_i|^2 > 0$$

für alle $x \neq 0$. Wir erhalten hiermit das Standard innere Produkt auf dem $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.

Das in Beispiel 18.2(b) definierte innere Produkt auf $V = \mathbb{K}^2$: $(v_1, v_2) = x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2$ kommt von

$$B = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}.$$

Genauer: Es ist $B = B^*$, und dass B eine positive Matrix definiert lässt sich aus $det(B) = 5 > 0$ und $det(B^{(1)}) = det(1) = 1 > 0$ folgern; vgl. Lemma 20.14 unten.

Um alle inneren Produkte auf einem endlich dimensionalen Vektorraum V zu beschreiben, brauchen wir also noch ein Kriterium dafür, wenn eine Matrix $A \in M_n(\mathbb{K})$ mit $A = A^*$ positiv ist.

Theorem 20.14. *Sei $A = (\alpha_{ij}) \in \mathbb{C}^{n\times n}$ eine selbst-adjungierte Matrix. Dann ist A genau dann positiv, wenn für alle $1 \leq k \leq n$ gilt*

$$det(A^{(k)}) = det \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2k} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \alpha_{k1} & \alpha_{k2} & \cdots & \alpha_{kk} \end{pmatrix} > 0$$

Beweis. Übung. □

Für symetrische Matrizen $A \in M_n(\mathbb{R})$ gilt also:

$$A \text{ ist positiv definit } \iff det(A^{(k)}) > 0 \text{ für } 1 \leq k \leq n.$$

21. Unitäre Abbildungen

Sind V und W $\mathbb{K}$-Vektorräume mit jeweils fest gewähltem inneren Produkt, so ist eine unitäre Abbildung $V \to W$ ein Isomorphismus von Vektorräumen, der zusätzlich mit den inneren Produkten verträglich ist; insbesondere erhält eine solche Abbildung den Winkel zwischen zwei Vektoren, sowie die Norm eines Vektors. Offensichtliche unitäre Abbildungen sind, zum Beispiel, Drehungen und Spiegelungen im $\mathbb{R}^2$. Wir betrachten allgemein unitäre Abbildungen und ihre fundamentalen Eigenschaften.

Definition 21.1. Seien V und W $\mathbb{K}$-Vektorräume (über demselben Körper) mit einem (jeweils fest gewählten) innerem Produkt $(\ ,\ )_V$ bzw. $(\ ,\ )_W$. Eine lineare Abbildung $f : V \to W$ erhält innere Produkte, falls

$$(f(v_1), f(v_2))_W = (v_1, v_2)_V \quad \text{für alle } v_1, v_2 \in V.$$

• Erhält f innere Produkte, so gilt $\| f(v) \| = \| v \|$ für alle $v \in V$; insbesondere erhält f die Norm, und ist injektiv (da $f(v) = 0 \Rightarrow v = 0$).

• Ist $f : V \to W$ ein Isomorphismus von Vektorräumen, der innere Produkte erhält, so erhält f^{-1} ebenfalls innere Produkte: Ist $f(v_i) = w_i$, so ist $(f^{-1}(w_1), f^{-1}(w_2))_V = (v_1, v_2)_V = (f(v_1), f(v_2))_W = (w_1, w_2)_W$. In diesem Fall sind V und W isomorph als Vektorräume mit innerem Produkt.

Lemma 21.2. *Seien V und W $\mathbb{K}$-Vektorräume mit innerem Produkt (über demselben Körper), so dass $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$. Für eine lineare Abbildung $f : V \to W$ sind gleichwertig:*

(a) *f erhält innere Produkte,*
(b) *f ist ein Isomorphismus von Vektorräumen mit innerem Produkt,*
(c) *f bildet jede Orthonormalbasis auf eine Orthonormalbasis ab,*
(d) *f bildet eine Orthonormalbasis auf eine Orthonormalbasis ab.*

• Sind V und W zwei Vektorräume mit innerem Produkt über demselben Körper, so ist $V \cong W$ (als Vektorräume mit innerem Produkt) genau dann, wenn $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$ ist: Die lineare Abbildung $f : V \to W$, die eine Orthonormalbasis von V auf eine Orthonormalbasis von W abbildet ist ein Isomorphismus von Vektorräumen, und erhält nach (d) die inneren Produkte.

Beweis. (a)⇒ (b): Nach Annahme ist $\| f(v) \| = \| v \|$ für alle $v \in V$. Also ist f injektiv, und wegen $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$ dann auch ein Isomorphismus.

(b)⇒(c): Ist f ein Isomorphismus und $a_1, \ldots, a_n$ eine Orthonormalbasis von V, so ist $f(a_1), \ldots, f(a_n)$ eine Basis von W. Da f innere Produkte erhält ist $(f(a_i), f(a_j)) = (a_i, a_j) = \delta_{ij}$, d.h. $f(a_1), \ldots, f(a_n)$ ist eine Orthonormalbasis von W.

(c)⇒(d): Trivial.

(d)⇒(a): Sei $a_1, \ldots, a_n$ eine Orthonormalbasis von V, so dass $f(a_1), \ldots, f(a_n)$ eine Orthonormalbasis von W ist. Dann ist

$$(f(a_i), f(a_j)) = (a_i, a_j) = \delta_{ij}.$$

Sind $v_1 = \sum_{j=1}^n x_j a_j$ und $v_2 = \sum_{k=1}^n y_k a_k$ in V, so gilt

$$\begin{aligned}(f(v_1), f(v_2)) &= (\textstyle\sum_{j=1}^n x_j f(a_j), \sum_{k=1}^n y_k f(a_k)) \\ &= \textstyle\sum_j \sum_k x_j \overline{y}_k (f(a_j), f(a_k)) \\ &= \textstyle\sum_j \sum_k x_i \overline{y}_j (a_j, a_k) \\ &= (\textstyle\sum_{j=1}^n x_j a_j, \sum_{k=1}^n y_k a_k) = (v_1, v_2),\end{aligned}$$

d.h. f erhält innere Produkte. □

Beispiele 21.3. (a) Jeder n-dimensionaler $\mathbb{K}$-Vektorraum V mit innerem Produkt ist isomorph zu $\mathbb{K}^n$ mit dem üblichen inneren Produkt: Ist $a_1, \ldots, a_n$ eine Orthonormalbasis von V, so ist ein Isomorphismus durch $A : V \to \mathbb{K}^n$, $\sum_{j=1}^n x_j a_j \mapsto (x_1, \ldots, x_n)$ gegeben.

(b) Sei $V = \mathbb{R}^3$ mit dem üblichen inneren Produkt, und sei W der $\mathbb{R}$-Vektorraum der schiefsymmetrischen Matrizen in $\mathbb{R}^{3\times 3}$ (d.h. $A^t = -A$) mit dem inneren Produkt $(A, B) = \frac{1}{2} tr(AB^t)$ (der Faktor $1/2$ dient der Vereinfachung). Ist $A = (\alpha_{ij}) \in W$, so gilt für die Einträge von A

$$\alpha_{ii} = 0 \text{ und } \alpha_{ij} = -\alpha_{ji} \text{ für } i \neq j.$$

Damit hat jede Matrix in W die Form

$$A = \begin{pmatrix} 0 & -\alpha_{12} & -\alpha_{13} \\ \alpha_{12} & 0 & -\alpha_{23} \\ \alpha_{13} & \alpha_{23} & 0 \end{pmatrix}$$

und die folgenden Matrizen bilden eine Basis von W

$$A_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ A_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Für $i = 1, 2, 3$ gilt $tr(A_i A_i^t) = 2$; da weiter $tr(A_i A_j^t) = 0$ für $i \neq j$ bilden A_1, A_2, A_3 eine Orthonormalbasis von W. Die lineare Abbildung

$$f : \mathbb{R}^3 \mapsto W, e_i \mapsto A_i,$$

bildet eine Orthonormalbasis auf eine Orthonormalbasis ab, und ist daher ein Isomorphismus von Vektorräumen mit innerem Produkt.

(c) Sei V der $\mathbb{R}$-Vektorraum der stetigen Funktionen $[0, 1] \to \mathbb{R}$ mit

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

und sei $V = W$ mit dem inneren Produkt (vgl. Beispiel 18.4(b))

$$(f, g)_W = \int_0^1 f(t)g(t)t^2 dt.$$

Ist $\psi : W \to V,\ f \mapsto \psi(f) := (t \mapsto tf(t)$, so gilt $(\psi(f), \psi(g))_V = (f, g)_W$, d.h. ψ erhält innere Produkte; aber ψ ist kein Isomorphismus, da $Bild(\psi) \subsetneq W$ ist (die konstanten Funktionen liegen nicht in $Bild(\psi)$).

Lemma 21.4. *Seien V und W $\mathbb{K}$-Vektorräume (über demselben Körper) mit innerem Produkt, und $f : V \to W$ eine lineare Abbildung. Dann erhält f innere Produkte genau dann, wenn $\| f(v) \| = \| v \|$ für alle $v \in V$.*

Beweis. Übung. □

Definition 21.5. Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt. Eine unitäre Abbildung auf V ist ein Isomorphismus $V \to V$, der das innere Produkt erhält.

• Die unitären Abbildungen auf V bilden eine Gruppe $U(V)$: Sind $u_1, u_2 \in U(V)$, so ist $u_2 u_1$ invertierbar und

$$\| u_2 u_1(v) \| = \| u_1(v) \| = \| v \| .$$

Ist $u \in U(V)$, so ist auch $u^{-1} \in U(V)$; das neutrale Element in $U(V)$ ist die Identitätsabbildung Id.

• Ist $\dim_{\mathbb{K}} V < \infty$, so ist $u \in U(V)$ genau dann, wenn u eine (bzw. jede) Orthonormalbasis von V auf eine Orthonormalbasis abbildet.

Proposition 21.6. *Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt, und $u \in End_{\mathbb{K}}(V)$. Dann ist $u \in U(V)$ genau dann, wenn u^* existiert und $u^*u = uu^* = Id$ ist.*

• Existiert u^* und ist $uu^* = Id$, so ist $u^* = u^{-1}$; da u^{-1} eindeutig bestimmt ist, folgt dann wegen $uu^{-1} = u^{-1}u = Id$ auch $u^*u = Id$.

Beweis. $\Rightarrow$: Ist u unitär, so ist u invertierbar und

$$(u(v_1), v_2) = (u(v_1), uu^{-1}(v_2)) = (v_1, u^{-1}v_2),\ v_1, v_2 \in V;$$

also ist die Adjungierte $u^* = u^{-1}$.
$\Leftarrow$: Ist $uu^* = u^*u = Id$, so ist u invertierbar und $u^{-1} = u^*$. Wegen

$$(u(v_1), u(v_2)) = (v_1, u^*u(v_2)) = (v_1, v_2),\ v_1, v_2 \in V,$$

erhält u innere Produkte. □

Definition 21.7. Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist unitär, falls $A^*A = E$ ist.

• Ist $\dim_{\mathbb{K}} V < \infty$, und $u \in End_{\mathbb{K}}(V)$, so ist $u \in U(V)$ genau dann, wenn die Matrix U von u bzgl. einer (bzw. jeder) Orthonormalbasis unitär ist.

• Sei $A = (\alpha_{ij}) \in \mathbb{C}^{n\times n}$. Dann gilt bzgl. dem üblichen inneren Produkt auf $\mathbb{C}^n$: A ist unitär $\Leftrightarrow$ die Zeilen (Spalten) von A sind orthonormal.

Beispiel 21.8. Sei $V = \mathbb{C}^n$ mit standard-innerem Produkt $(x, y) = x^t\overline{y}$. Sei $B \in M_n(V)$ und $f = f_B$. Dann gilt:

$$(f_B(x), f_B(y)) = x^t B^t \overline{B}\overline{y}.$$

Also folgt leicht: f_B ist unitär $\iff B^*B = E$.

Definition 21.9. Eine Matrix $A \in \mathbb{K}^{n\times n}$ ist orthogonal, falls $A^tA = E$.

Beispiele 21.10. (a) Eine 1×1 Matrix $A = (\alpha)$ ist orthogonal genau dann, wenn $\alpha = \pm 1$ ist; und unitär genau dann, wenn $\overline{\alpha}\alpha = 1$ ist; im zweiten Fall ist $|\alpha| = 1$, d.h. $\alpha = exp(i\theta)$, wobei $\theta \in \mathbb{R}$ ist.
(b) Sei $A = (\alpha_{ij}) \in \mathbb{K}^{2\times 2}$. Dann ist A orthogonal genau dann, wenn

$$\begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix} = A^t = A^{-1} = \frac{1}{\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}} \begin{pmatrix} \alpha_{22} & \alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix}$$

Wegen $A^tA = E$ gilt $\det(A) = \pm 1$; ist $det(A) = 1$, so folgt $\alpha_{11} = \alpha_{22}$ und $\alpha_{12} = -\alpha_{21}$, im Fall $det(A) = -1$ ist $\alpha_{11} = -\alpha_{22}$ und $\alpha_{12} = \alpha_{21}$. Somit ist jede orthogonale Matrix in $\mathbb{K}^{2\times 2}$ von einer der beiden Formen

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \text{ oder } \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, \quad \text{wobei } \alpha^2 + \beta^2 = 1.$$

Ist $\mathbb{K} = \mathbb{R}$ und $\det(A) = +1$, so gibt es genau einen Winkel θ mit $0 \leq \theta < 2\pi$ und

$$(a, b) = (\cos(\theta), \sin(\theta).$$

Dann beschreibt die Matrix

$$A_\theta = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$$

die Drehung im $\mathbb{R}^2$ um den Winkel θ (entgegen dem Uhrzeigersinn).

Im Fall $\det(A) = -1$ gilt $\chi_A(x) = x^2 - (\alpha^2 + \beta^2) = x^2 - 1 = (x-1)(x+1)$. Sei v_1 Eigenvektor zum Eigenwert $\lambda_1 = 1$ und v_2 Eigenvektor zum Eigenwert $\lambda_2 = -1$. Dann beschreibt A die Spiegelung an der Geraden $\mathbb{R}v_1$.

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt, und seinen $B = \{a_1, \ldots, a_n\}$ bzw. $B' = \{a'_1, \ldots, a'_n\}$ zwei Orthonormalbasen von V. Sei $U = (u_{ij}) \in \mathbb{K}^{n\times n}$ die Matrix der linearen Abbildung $V \to V,\ a_i \mapsto a'_i,\ i = 1, \ldots, n$ bzgl. der Basis B, so dass gilt

$$a'_j = \sum_{i=1}^{n} u_{ij} a_i;$$

offensichtlich ist U unitär. Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ eine lineare Abbildung, so gilt für die Matrix $A = A_{f,B}$ und $A' = A_{f,B'}$

$$A' = U^{-1}AU = U^*AU.$$

Definition 21.11. Seien $A, B \in \mathbb{C}^{n\times n}$. Dann ist A unitär ähnlich zu B, falls es eine unitäre Matrix $U \in \mathbb{C}^{n\times n}$ gibt, so dass $B = U^*AU$ ist; B ist orthogonal ähnlich zu A, falls es eine orthogonale Matrix U mit $B = U^tAU$ gibt.

• Sind B, B' orthonormale Basen von V, so ist für jede lineare Abbildung $f : V \to V$ die Matrix $A' = A_{f,B'}$ unitär ähnlich zu $A = A_{f,B}$. Ist V ein $\mathbb{R}$-Vektorraum, so sind diese Matrizen orthogonal ähnlich, mittels einer reellen orthogonalen Matrix.

22. Normale lineare Abbildungen

Wir betrachten die folgende Frage: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt, und $f \in End_{\mathbb{K}}(V)$. Wann gibt es eine orthonormale Basis von V, die aus Eigenvektoren von f besteht, d.h. wann gibt es eine orthonormale Basis, so dass die Matrix von f bzgl. dieser ON-Basis Diagonalform hat?

Gilt dies, d.h. ist $B = \{a_1, \ldots, a_n\}$ eine Orthonormalbasis, so dass

$$f(a_j) = \alpha_j a_j, \ j = 1, \ldots, n,$$

so ist die Matrix von f bzgl. dieser Basis die Diagonalmatrix mit den Diagonaleinträgen $\alpha_1, \ldots, \alpha_n$. Die Adjungierte f^* von f ist bzgl. derselben Basis durch die Diagonalmatrix mit den Einträgen $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ dargestellt. Ist V ein $\mathbb{R}$-Vektorraum, so sind die α_i reell, $\overline{\alpha}_i = \alpha_i$, und es folgt $f = f^*$; d.h. für einen reellen Vektorraum muss notwendigerweise $f = f^*$ gelten. Im komplexen Fall sind die α_i nicht unbedingt reell, jedoch muss wegen $\alpha_i\overline{\alpha}_i = \overline{\alpha}_i\alpha_i$ dann $ff^* = f^*f$ gelten. Wir werden zeigen, dass die Bedingung $ff^* = f^*f$ auch hinreichend für die Existenz einer Orthonormalbasis aus Eigenvektoren ist.

Definition 22.1. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt. Eine lineare Abbildung $f : V \to V$ ist normal, falls

$$ff^* = f^*f.$$

• Selbst-adjungierte und unitäre Abbildungen sind normal. Weiter ist ein Skalarvielfaches einer normalen Abbildung normal; jedoch sind Summen und Produkte von normalen Abbildungen in der Regel nicht normal (d.h. die normalen Abbildungen bilden keine Untergruppe von $End_{\mathbb{K}}(V)$).

Wir betrachten zunächst nur selbst-adjungierte Abbildungen:

Lemma 22.2. *Sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt. Ist ein Endomorphismus $f \in End_{\mathbb{K}}(V)$ selbst-adjungiert (d.h. $f = f^*$), so ist jeder Eigenwert von f reell. Die Eigenvektoren zu verschiedenen Eigenwerten von f sind orthogonal.*

Beweis. Sei $\alpha \in \mathbb{K}$ ein Eigenwert von f, d.h. $f(v) = \alpha v$ für ein $0 \neq v$. Die Annahme $f = f^*$ liefert $\alpha(v, v) = (\alpha v, v) = (f(v), v) = (v, f(v)) = (v, \alpha v) = \overline{\alpha}(v, v)$. Wegen $0 \neq v$ ist $(v, v) \neq 0$ und es folgt $\alpha = \overline{\alpha}$. Ist $\alpha' \in \mathbb{K}$ ein Eigenwert mit Eigenvektor v', so folgt aus $\alpha(v, v') = (f(v), v') = (v, f(v')) = (v, \alpha' v') = \overline{\alpha}'(v, v') = \alpha'(v, v')$, dass $(\alpha - \alpha')(v, v') = 0$ und somit $(v, v') = 0$ für $\alpha \neq \alpha'$. □

Lemma 22.3. *Sei $V \neq 0$ ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt. Dann hat jeder selbst-adjungierte Endomorphismus $f \in End_{\mathbb{K}}(V)$ einen Eigenvektor.*

Beweis. Sei A die Matrix zu f bezüglich einer beliebigen Basis von V. Dann sind die Nullstellen des charakteristischen Polynoms $\chi_A(x)$ genau die Eigenwerte von f. Betrachtet man $A \in M_n(\mathbb{C})$ so hat A komplexe Eigenwerte, da $\mathbb{C}$ algebraisch abgeschlossen ist. Nach Lemma 22.2 ist jeder dieser Eigenwerte reell. □

Beispiel 22.4. Sei V der $\mathbb{C}$-Vektorraum der stetigen Funktionen $[0,1] \to \mathbb{C}$ mit dem innerem Produkt

$$(f,g) = \int_0^1 f(t)\overline{g(t)}dt.$$

Dann ist $\varphi : V \to V,\ f(t) \mapsto tf(t)$ selbst-adjungiert, vgl. Beispiel 19.6(b). Ist $\varphi(f) = \alpha f$, so ist $(t-\alpha)f(t) = 0$ für $0 \leq t \leq 1$ und $f(t) = 0$ für $t \neq \alpha$. Da f stetig ist, folgt $f = 0$; also hat φ keine Eigenwerte (bzw. Eigenvektoren); insbesondere gilt Lemma 22.3 im Fall unendlicher Dimension in der Regel nicht.

Lemma 22.5. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt, und sei $f \in End_{\mathbb{K}}(V)$ ein beliebiger Endomorphismus. Ist $W \subseteq V$ ein f-invarianter linearer Unterraum (d.h. $f(W) \subseteq W$), so ist das orthogonale Komplement $W^\perp$ f^*-invariant (d.h. $f^*(W^\perp) \subseteq W^\perp$).*

Beweis. Sei $w' \in W^\perp$. Ist $w \in W$, so ist nach Annahme $f(w) \in W$. Es folgt $0 = (f(w), w') = (w, f^*(w'))$ für alle $w \in W$, also ist $f^*(w') \in W^\perp$. □

Theorem 22.6. *Sei $V \neq 0$ ein n-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt, und sei $f \in End_{\mathbb{K}}(V)$ selbst-adjungiert. Dann gibt es eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht.*

Beweis. Nach Lemma 22.3 hat f einen Eigenvektor v. Sei $a_1 = \| v \|^{-1} v$, so dass a_1 ein Eigenvektor von f der Länge 1 ist. Ist $\dim_{\mathbb{K}} V = 1$, so sind wir fertig. Nach Induktion gelte die Behauptung für V mit $0 < \dim_{\mathbb{K}} V < n$. Sei $W = \langle a_1 \rangle \subseteq V$ der von a_1 erzeugte 1-dimensionale lineare Unterraum. Da v_1 ein Eigenvektor ist, ist $f(W) \subseteq W$. Nach Lemma 22.5 ist $W^\perp$ invariant unter $f^* = f$. Der lineare Unterraum $W^\perp \subseteq V$ ist ein innerer Produktraum (mit dem inneren Produkt von V), $f|_{W^\perp}$ ist selbst-adjungiert, und $\dim_{\mathbb{K}} W^\perp = n-1$. Nach Induktion gibt es eine Orthonormalbasis $a_2, \ldots, a_n$, die aus Eigenvektoren von $f|_{W^\perp}$ besteht; jeder dieser Vektoren ist auch ein Eigenvektor von f. Da

nach Lemma 18.12 $V = W \oplus W^\perp$ ist, ist $a_1, \ldots, a_n$ eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht. □

Korollar 22.7. *Sei $A \in M_n(\mathbb{C})$ eine hermitesche Matrix (d.h. $A = A^*$). Dann gibt es eine unitäre Matrix U, so dass U^*AU eine Diagonalmatrix ist. Hat A reelle Einträge, so gibt es eine orthogonale reelle Matrix U, so dass U^tAU ist eine Diagonalmatrix ist.*

Beweis. Sei $V = \mathbb{C}^n$ mit dem üblichen inneren Produkt, und sei $f = f_A$ der Endomorphismus $x \mapsto Ax$. Bezüglich der Standardbasis $e_1, \ldots, e_n$ des $\mathbb{C}^n$ hat dann f die Matrix A. Wegen $A = A^*$ ist $f = f^*$. Sei $B = \{a_1, \ldots, a_n\}$ eine Orthonormalbasis von V, so dass $f(a_i) = \alpha_i a_i$, $i = 1, \ldots, n$, und sei $D = D_B$ die entsprechene Diagonalmatrix mit Einträgen $\alpha_1, \ldots, \alpha_n$. Ist U die Matrix mit den Spaltenvektoren $a_1, \ldots, a_n$, so ist U unitär, da die ON-Basis $e_1, \ldots, e_n$ auf die ON-Basis $a_1, \ldots, a_n$ abgebildet wird. Ferner gilt dann $D = U^{-1}AU = U^*AU$. Im Fall $\mathbb{K} = \mathbb{R}$ liefert das analoge Argument eine unitäre Matrix mit reellen Einträgen, also eine reelle orthogonale Matrix. □

Theorem 22.6, zusammen mit den Bemerkungen am Anfang dieses Kapitels besagt, dass im Fall eines endlich-dimensionalen *reellen* Vektorraums mit innerem Produkt gilt: Ist $f \in End_{\mathbb{R}}(V)$, so hat V genau dann eine Orthonormalbasis bestehend aus Eigenvektoren zu f, wenn $f = f^*$ ist. Gleichwertig: Ist $A \in M_n(\mathbb{R})$, so gibt es genau dann eine reelle orthogonale Matrix U, so dass U^tAU eine Diagonalmatrix ist, wenn $A^t = A$ ist. Diese Resultate gelten im Fall $\mathbb{K} = \mathbb{C}$ nicht, in diesem Fall besteht ein Unterschied zwischen den Bedingungen $A = A^t$ und $A = A^*$.

Beispiel 22.8. Sei $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, x \mapsto Ax$ mit $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Dann ist das charakteristische Polynom gegeben durch

$$\chi_f(x) = (x-3)(x+1).$$

Die Berechnung eines normierten Eigenvektors zum Eigenwert $\alpha_1 = -1$ liefert $a_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Der Lotraum zu $W = \langle a_1 \rangle$ wird erzeugt von $(1,1)^t$ und wir erhalten einen normierten Eigenvektor zu $\alpha_2 = 3$ durch $a_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Für die Matrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

gilt also $U^{-1} = U^t$ und

$$U^t A U = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}.$$

Wir beweisen das Analogon zu Theorem 22.6 im Fall $\mathbb{K} = \mathbb{C}$:

Lemma 22.9. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt, und sei $f \in End_{\mathbb{K}}(V)$ normal. Dann gilt: v ist genau dann Eigenvektor von f zum Eigenwert α, wenn v Eigenvektor von f^* zum Eigenwert $\overline{\alpha}$ ist.*

• Der Beweis zeigt: Ist f normal, so ist $\| f(v) \| = \| f^*(v) \|$ für alle $v \in V$.

Beweis. Sei $v \in V$. Dann folgt aus der Annahme $ff^* = f^*f$, dass

$$\begin{aligned} \| f(v) \|^2 &= (f(v), f(v)) = (v, f^*f(v)) \\ &= (v, ff^*(v)) = (f^*(v), f^*(v)) = \| f^*(v) \|^2, \end{aligned}$$

d.h. $\| f(v) \| = \| f^*(v) \|$. Ist $f \in End_{\mathbb{K}}(V)$ beliebig und $\alpha \in \mathbb{K}$, so ist

$$\begin{aligned} ((f - \alpha id)(v_1), v_2) &= (f(v_1), v_2) - (\alpha v_1, v_2) \\ &= (v_1, f^*(v_2)) - (v_1, \overline{\alpha} v_2) = (v_1, (f^* - \overline{\alpha} id)(v_2)), \end{aligned}$$

d.h. $(f - \alpha id)^* = f^* - \overline{\alpha} id$. Eine einfache direkte Rechnung zeigt nun, dass mit f normal auch $f - \alpha id$ normal ist. Somit gilt für jedes $v \in V$

$$\| (f - \alpha id)(v) \| = \| (f^* - \overline{\alpha} id) v \|;$$

insbesondere gilt $f(v) = \alpha v$ mit $v \neq 0$ genau dann, wenn $f^*(v) = \overline{\alpha} v$. □

Theorem 22.10. *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit innerem Produkt, und sei $f \in End_{\mathbb{K}}(V)$ normal. Dann hat V eine Orthonormalbasis aus Eigenvektoren von f.*

Beweis. Da V ein komplexer Vektorraum ist, hat f einen Eigenvektor a_1; wir können annehmen, dass $\| a_1 \| = 1$ ist. Sei $W = \langle a_1 \rangle \subseteq V$. Da a_1 ein Eigenvektor ist, ist W invariant unter f. Nach Lemma 22.9 ist W auch invariant unter f^*, so dass nach Lemma 22.5 $W^\perp$ invariant unter $f^{**} = f$ ist. Da $(f|_{W^\perp})^* = f^*|_{W^\perp}$ ist, ist die Einschränkung von f auf $W^\perp$ normal, und die Behauptung folgt mit einem Induktionsargument analog zum Beweis von Theorem 22.6. □

Die offensichtliche Matrixinterpretation von Theorem 22.10 ist:

Definition 22.11. Eine komplexe Matrix $A \in M_n(\mathbb{C})$ ist normal, wenn $AA^* = A^*A$ ist.

Theorem 22.12. *Sei $A \in M_n(\mathbb{C})$. Dann gibt es genau dann eine unitäre Matrix $U \in M_n(\mathbb{C})$, so dass U^*AU eine Diagonalmatrix ist, wenn A normal ist, d.h. A ist genau dann unitär ähnlich zu einer Diagonalmatrix, wenn A normal ist.*

23. Das Spektraltheorem

Sei V ein endlich-dimensionaler komplexer Vektorraum mit innerem Produkt. Wir zeigen, dass jeder normale Endomorphismus f auf V eine Linearkombination von orthogonalen Projektionen ist, d.h. es gilt

$$f = \alpha_1 p_1 + \cdots + \alpha_k p_k,$$

wobei $p_1, \ldots, p_k$ orthogonale Projektionen sind, die paarweise orthogonal sind, $p_i p_j = 0$ für $i \neq j$. Dieses Resultat lässt sich aus Theorem 22.10 herleiten; wir geben jedoch einen anderen, algebraischen Beweis.

Definition 23.1. Sei K ein beliebiger Körper und V ein K-Vektorrraum. Sei U ein Unterraum von V und $p\colon V \longrightarrow V$ eine lineare Abbildung. Dann nennt man p eine Projektion, falls $p(V) = U$ und $p|_U = \mathrm{id}_U$ gilt.

Jede Projektion erfüllt $p^2 = p$. Umgekehrt ist jede lineare Abbildung mit $p^2 = p$ eine Projektion auf $U = p(V)$.

Definition 23.2. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt. Sei $U \subseteq V$ ein Unterrraum. Dann gilt $V = U \oplus U^\perp$ und wir definieren $p = p_U\colon V \longrightarrow V$ durch $p(v) = u$, falls $v = u + w$ mit $u \in U$ und $w \in U^\perp$. Die lineare Abbildung $p = p_U$ ist die orthogonale Projektion auf U.

Lemma 23.3. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt. Sei $p\colon V \longrightarrow V$ eine Projektion auf U. Dann sind folgende Aussagen äquivalent:*

(a) *p ist die orthogonale Projektion auf U.*
(b) *$v - p(v) \in U^\perp$ für alle $v \in V$.*
(c) $\ker(p) = U^\perp$.

Beweis. "(a) $\implies$ (b)": Sei $v = u + w$ mit $u \in U$ und $w \in U^\perp$. Dann gilt: $v - p(v) = u + w - u = w \in U^\perp$.

"(b) $\implies$ (c)": Sei $v \in \ker(p)$. Dann gilt $v = v - p(v) \in U^\perp$. Sei umgekehrt $v \in U^\perp$. Wegen $v - p(v) \in U^\perp$ ist dann auch $p(v) \in U^\perp$. Da p eine Projektion auf U ist folgt $p(v) \in U \cap U^\perp = \{0\}$.

"(c) $\implies$ (a)": Sei $v = u + w$ mit $u \in U$ und $w \in U^\perp$. Dann gilt $p(v) = p(u) + p(w) = p(u) = u$, da $p|_U = \mathrm{id}_U$. □

Die folgenden zwei Lemmata beschreiben orthogonale Unterräume mittels Projektionen und deren Adjungierten, und sind der 'geometrische' Anteil des Beweises des Spektraltheorems.

Lemma 23.4. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit innerem Produkt, und sei $p : V \to V$ eine Projektion (d.h. $p^2 = p$). Dann sind folgende Aussagen gleichwertig:*

(a) *p ist normal, d.h. $pp^* = p^*p$,*
(b) *p ist selbst-adjungiert, d.h. $p = p^*$,*
(c) *p ist die Orthogonalprojektion von V auf $Bild(p)$.*

Beweis. (a)⇒(b): Wegen $pp^* = p^*p$ gilt $\| p(v') \| = \| p^*(v') \|$ für alle $v' \in V$, siehe Beweis von Lemma 22.9. Somit gilt $p(v') = 0$ genau dann, wenn $p^*(v') = 0$ ist. Sei $v \in V$, und $v' = v - p(v)$. Es folgt

$$p(v') = p(v - p(v)) = p(v) - p^2(v) = p(v) - p(v) = 0,$$

und weiter $0 = p(v') = p^*(v') = p^*(v - p(v)) = p^*(v) - p^*p(v)$. Also ist $p^* = p^*p$, und $p = p^{**} = (p^*)^* = (p^*p)^* = p^*p = p^*$.
(b)⇒(c): Zu zeigen ist: $p = p^*$ impliziert $Kern(p) = p(V)^\perp$. Sei dazu zunächst $v \in \ker(p)$ und $w = p(w_1) \in p(V)$. Dann gilt $(v, w) = (v, p(w_1)) = (p^*(v), w_1) = (p(v), w_1) = (0, w_1) = 0$. Also folgt $\ker(p) \subseteq p(V)^\perp$. Sei umgekehrt $v \in p(V)^\perp$. Dann gilt für alle $w \in V$

$$0 = (v, p(w)) = (p^*(v), w) = (p(v), w).$$

Insbesondere gilt für $w = p(v)$ also $(p(v), p(v)) = 0$, d.h. $p(v) = 0$.
(c)⇒(a): Seien $v_1, v_2 \in V$. Da p eine Orthogonalprojektion ist, ist $v_1 - p(v_1) \in Bild(p)^\perp$; klar ist $p(v_2) \in Bild(p)$. Es folgt

$$\begin{aligned} 0 = (v_1 - p(v_1), p(v_2)) &= (v_1, p(v_2)) - (p(v_1), p(v_2)) = \\ &= (v_1, p(v_2)) - (v_1, p^*p(v_2)) = (v_1, (p - p^*p)(v_2)). \end{aligned}$$

Da dies für alle $v_1, v_2 \in V$ gilt, folgt $p = p^*p$, und $p^* = (p^*p)^* = p^*p^{**} = p^*p = p$, d.h. p ist selbst-adjungiert und damit auch normal. □

Ist $V = W_1 \oplus \cdots \oplus W_k \subseteq V$ die direkte Summe von Unterräumen $W_j \subseteq V$, so hat jeder Vektor $v \in V$ eine eindeutige Darstellung als

$$v = \sum_{j=1}^{k} w_j, \ w_j \in W_j.$$

Hat V ein inneres Produkt, so ist eine orthogonale direkte Summe $V = W_1 \oplus \cdots \oplus W_k$ eine solche Zerlegung, für die zusätzlich gilt, dass für $i \neq j$ jeder Vektor in W_i orthogonal zu jedem Vektor in W_j ist.

Lemma 23.5. *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit innerem Produkt, $W_1, \ldots, W_k \subseteq V$ Unterräume, und p_j die Orthogonalprojektion von V auf W_j, $j = 1, \ldots, k$. Dann sind gleichwertig:*

(a) *$V = W_1 \oplus \cdots \oplus W_k$ ist eine orthogonale direkte Summe,*
(b) *$id = \sum_{j=1}^{k} p_j$ und $p_i p_j = 0$ für $i \neq j$,*
(c) *Ist B_j eine Orthonormalbasis von W_j, $j = 1, \ldots, k$, so ist die Vereinigung $B := \cup_j B_j$ eine Orthonormalbasis von V.*

Beweis. "(a) $\implies$ (b)" Sei $v = w_1 + \ldots + w_k$ ein beliebiger Vektor in V. Wegen $W_i^{\perp} = \oplus_{j \neq i} W_j$ und $\ker(p_i) = W_i^{\perp}$ gilt:

$$\left(\sum_{i=1}^{k} p_i \right)(v) = \sum_{i=1}^{k} \sum_{j=1}^{k} p_i(w_j) = \sum_{i=1}^{k} p_i(w_i) = \sum_{i=1}^{k} w_i = v,$$

d.h. $\sum_{i=1}^{k} p_i = \mathrm{id}_V$. Ferner gilt

$$(p_i p_j)(v) = p_i(p_j(w_1 + \ldots + w_k)) = p_i(p_j(w_j)) = p_i(w_j) = 0,$$

d.h. $p_i p_j = 0$.
"(b) $\implies$ (c)" Wegen $v = \mathrm{id}_V(v) = \sum_{i=1}^{k} p_i(v) \in W_1 + \ldots + W_k$ ist B ein Erzeugendensystem von V. Sei $B_j = \{b_{j1}, \ldots, b_{jn_j}\}$. Sei

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} \alpha_{ji} b_{ji} = 0.$$

Dann gilt für alle $l = 1, \ldots, k$

$$0 = p_l \left(\sum_{j=1}^{k} \sum_{i=1}^{n_j} \alpha_{ji} b_{ji} = 0 \right) = \sum_{i=1}^{n_l} \alpha_{li} b_{li},$$

und da B_j eine Basis ist, folgt $\alpha_{li} = 0$ für alle $1 \leq i \leq n_l$ und alle $l = 1, \ldots, k$. Damit haben wir gezeigt, dass B eine Basis ist.

Da p_j eine Orthogonalprojektion ist, gilt $p_j^* = p_j$. Weiter gilt für $v \in W_i$ und $w \in W_j$ mit $i \neq j$

$$(v, w) = (p_i(v), p_j(w)) = (v, p_i^* p_j(v)) = (v, p_i p_j(v)) = (v, 0) = 0.$$

(c) $\implies$ (a): Klar. □

Lemma 23.6. *Sei $f \in End_{\mathbb{C}}(V)$ normal. Dann gilt:*

(a) *Ist $f^2(v) = 0$, so ist $f(v) = 0$,*
(b) *Ist $q \in \mathbb{C}[x]$, so ist $q(f)$ normal,*
(c) *Das Minimalpolynom $\mu_f(x)$ hat keine mehrfachen Nullstellen.*

• Teil (a) besagt $Kern(f) \cap Bild(f) = \{0\}$: Ist $v' = f(v) \in Bild(f)$ und $f(v') = 0$, so ist $v' = 0$.

Beweis. (a): Sei $f^2(v) = 0$ und $v' = f(v)$, so dass $f(v') = 0$. Wegen $ff^* = f^*f$ ist $\| f(v') \| = \| f^*(v') \|$, also ist $f^*(v') = 0$. Es folgt $0 = (f^*(v'), v) = (v', f(v)) = (v', v')$, d.h. $v' = 0$.
(b): Sei $q = \sum_{i=0}^n \alpha_i x^i$. Dann ist $q(f) = \sum_{i=0}^n \alpha_i f^i$ und $q(f)^* = \sum_{i=0}^n \overline{\alpha_i}(f^*)^i$. Wie im Beweis von Lemma 22.9 folgt durch direkte Rechnung, dass $q(f)$ mit $q(f)^*$ kommutiert.
(c): Sei $\mu_f(x) = \prod_{i=1}^k (x - \alpha_i)$; ist $\alpha = \alpha_i$ eine mehrfache Nullstelle, so ist $\mu_f(x) = (x - \alpha)^2 g(x)$ für ein $g \in \mathbb{C}[x]$. Wegen $\mu_f(f) = 0$ folgt dann

$$(f - \alpha Id)^2 g(f)(v) = 0, \ v \in V.$$

Nach (b) ist $f - \alpha Id$ normal. Für $v \in V$ und $v' = g(f)(v)$ ist

$$(f - \alpha Id)^2(v') = (f - \alpha Id)^2 g(f)(v) = 0,$$

und (a) impliziert $(f - \alpha Id)(v') = 0$. Also ist

$$(f - \alpha Id) g(f)(v) = 0$$

für alle $v \in V$, d.h. $(f - \alpha Id)g(f) = 0$; da $\mu_f \neq 0$ das Polynom von minimalem Grad mit $\mu_f(f) = 0$ ist, ist dies ein Widerspruch und μ_f hat keine mehrfachen Nullstellen. □

Wir zeigen nun:

Theorem 23.7. *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit innerem Produkt, und sei $f \in End_{\mathbb{C}}(V)$ eine normale Abbildung. Seien $\alpha_1, \ldots, \alpha_k$ die paarweise verschiedenen Eigenwerte von f, und p_j die Orthogonalprojektion von V auf den Eigenraum $V(\alpha_j)$. Dann gilt*

(a) $f = \alpha_1 p_1 + \cdots \alpha_k p_k$,
(b) $Id = p_1 + \cdots + p_k$,
(c) $p_i p_j = 0$ *für* $i \neq j$.

Insbesondere ist nach Lemma 23.5 $V = V(\alpha_1) \oplus + \cdots + \oplus V(\alpha_k)$ eine orthogonale direkte Summe. Die Darstellung in (a) ist die Spektralauflösung von f.

NB. Hat ein beliebiger Endomorphismus f eine Darstellung als

$$f = \sum_{i=1} \alpha_i p_i,$$

mit Orthogonalprojektionen p_i, so dass $p_i p_j = 0$ für $i \neq j$ ist, so ist f normal: Wegen $p_i = p_i^*$ folgt $f^* = \sum_i \overline{\alpha}_i p_i$, und da $p_i p_j = p_j p_i = 0$ für $i \neq j$ gilt, ist $ff^* = f^*f$. Somit sind die normalen Abbildungen genau

diejenigen Endomorphismen, die eine Darstellung als Linearkombination von kommutierenden Orthogonalprojektionen haben.

Beweis. Wir beginnen mit einer Vorbemerkung. Da f normal ist, ist $\mu_f(x) = \prod_{i=1}^k (x - \alpha_i)$, wobei die $\alpha_i \in \mathbb{C}$ paarweise verschieden sind, siehe Lemma 23.6(c). Angenommen wir haben eine Darstellung wie in (a), so dass auch (c) erfüllt ist. Dann gilt

$$f^2 = (\sum_{i=1}^k \alpha_i p_i)(\sum_{j=1}^k \alpha_j p_j) = \sum_j \alpha_j^2 p_j^2 = \sum_j \alpha_j^2 p_j,$$

und allgemein $f^r = \sum_{j=1}^k \alpha^r p_j$. Ist $g \in \mathbb{C}[x]$ beliebig, so folgt

$$g(f) = \sum_{j=1}^k g(\alpha_j) p_j,$$

und für $q_j \in \mathbb{C}[x]$ mit $q_j(\alpha_i) = \delta_{ij}$ ist $q_j(f) = \sum_i q_j(\alpha_i) p_i = p_j$; dies suggeriert nach Polynomen q_j mit $q_j(\alpha_i) = \delta_{ij}$ zu suchen.

Nach dieser Vorbemerkung kommen wir zum eigentlichen Beweis. Ist $k = 1$, so ist $f = \alpha Id$ und f hat trivialerweise eine Darstellung der gewünschten Form. Sei also $k \geq 2$. Definiere $q_1, \ldots, q_k \in \mathbb{C}[x]$ durch

$$q_j(x) = \frac{(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_{j-1})(x-\alpha_{j+1})\cdots(x-\alpha_k)}{(\alpha_j-\alpha_1)(\alpha_j-\alpha_2)\cdots(\alpha_j-\alpha_{j-1})(\alpha_j-\alpha_{j+1})\cdots(\alpha_j-\alpha_k)}.$$

Dann gilt $q_j(\alpha_i) = \delta_{ij}$. Ist $g \in \mathbb{C}[x]$ beliebig vom Grad $\leq k-1$, so ist

$$g - g(\alpha_1)q_1 - g(\alpha_2)q_2 - \cdots - g(\alpha_k)q_k$$

ein Polynom in $\mathbb{C}[x]$ vom Grad $\leq k-1$ mit den k verschiedenen Nullstellen $\alpha_1, \ldots, \alpha_k$, also das Nullpolynom, d.h. es gilt

$$g = g(\alpha_1)q_1 + \cdots + g(\alpha_k)q_k.$$

Die Spezialfälle $g = 1$ und $g = x$ liefern die Identitäten

$$\begin{aligned} 1 &= q_1 + \cdots + q_k, \\ x &= \alpha_1 q_1 + \cdots + \alpha_k q_k. \end{aligned}$$

Setze $\widetilde{p}_j = q_j(f)$. Dann folgt aus den den obigen Relationen

$$\begin{aligned} Id &= \widetilde{p}_1 + \cdots + \widetilde{p}_k, \\ f &= \alpha_1 \widetilde{p}_1 + \cdots \alpha_k \widetilde{p}_k. \end{aligned}$$

Es gilt:

• $\widetilde{p}_j \neq 0$ für jedes j: Es ist $\widetilde{p}_j = q_j(f)$, wobei q_j ein Polynom vom Grad $\leq k-1$ ist. Da $\mathrm{Grad}(\mu_f) = k$ ist, kann $0 = \widetilde{p}_j = q_j(f)$ nicht gelten (μ_f hat minimalen Grad mit dieser Eigenschaft).

• $\widetilde{p}_i \widetilde{p}_j = 0$ für $i \neq j$: Für $i \neq j$ gilt $\mu_f | q_i q_j$, also ist $g\mu_f = q_i q_j$ für ein $g \in \mathbb{C}[x]$, und $0 = g(f)\mu_f(f) = q_i(f)q_j(f) = \widetilde{p}_i \widetilde{p}_j$.

• $\widetilde{p}_j$ ist eine Projektion: Wegen $Id = \widetilde{p}_1 + \cdots + \widetilde{p}_k$ und $\widetilde{p}_i\widetilde{p}_j = 0$ für $i \neq j$ folgt $\widetilde{p}_j = \widetilde{p}_j\widetilde{p}_1 + \cdots \widetilde{p}_j\widetilde{p}_k = \widetilde{p}_j^2$ für alle j.

• Die $\widetilde{p}_j$ sind Orthogonalprojektionen: Nach Definiton ist $\widetilde{p}_j = q_j(f)$, so dass $\widetilde{p}_j$ nach Lemma 23.6(b) normal, und nach Lemma 23.4 einen Orthogonalprojektion ist.

• $\widetilde{p}_j(V) = V(\alpha_j)$: Ist $v \in V(\alpha_j)$, so ist $f(v) = \alpha_j v$ und wegen

$$f(v) = (\sum_{i=1}^{k} \alpha_i\widetilde{p}_i)(v) = \sum_{i=1}^{k} \alpha_i\widetilde{p}_i(v)$$

und

$$\alpha_j v = \alpha_j(\sum_{i=1}^{k} \widetilde{p}_i)(v) = \sum_{i=1}^{k} \alpha_j\widetilde{p}_i(v)$$

folgt $\sum_{i=1}^{k}(\alpha_i - \alpha_j)\widetilde{p}_i(v) = 0$. Wegen Lemma 23.5 wissen wir bereits, dass

$$V = \widetilde{p}_1(V) \oplus \ldots \oplus \widetilde{p}_k(V)$$

eine direkte orthogonale Summe ist. Es folgt daher $\widetilde{p}_i(v) = 0$ für $i \neq j$ und $\widetilde{p}_j(v) = v$. Ist umgekehrt $0 \neq v \in Bild(\widetilde{p}_j)$, so ist $v = \widetilde{p}_j(v)$ und $f(v) = f\widetilde{p}_j(v) = (\sum_{i=1}^{k} \alpha_i\widetilde{p}_i)\widetilde{p}_j(v) = \alpha_j\widetilde{p}_j(v) = \alpha_j v$.

Insgesammt haben wir also gezeigt, dass $\widetilde{p}_j = p_j$ gilt. □

Beispiel 23.8. Wir betrachten das Beispiel 22.8 und schreiben

$$U^tAU = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Es folgt:

$$\begin{aligned} A &= -1 \cdot U \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} U^t + 3 \cdot U \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} U^t \\ &= -1 \cdot \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + 3 \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= -1 \cdot P_1 + 3 \cdot P_2 \end{aligned}$$

mit den Orthogonalprojektionen

$$P_1 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad P_2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Verwenden wir die zur Bestimmung der Orthogonalprojektionen die Methode des letzten Beweises, so ergibt sich:

$$\begin{aligned} q_1(x) &= \frac{x-3}{-1-3} = -\frac{1}{4}(x-3), \\ q_2(x) &= \frac{x+1}{3+1} = \frac{1}{4}(x+1) \end{aligned}$$

und wir erhalten

$$\begin{aligned} \widetilde{P}_1 &= = -\frac{1}{4}(A-3E) = P_1, \\ \widetilde{P}_2 &= = \frac{1}{4}(A+E) = P_2. \end{aligned}$$

24. Dualräume

Dieser Abschnitt stellt in weiten Teilen eine Wiederholung des früheren gleichnamigen Abschnitts dar. Da wir in der Linearen Algebra II mittels Zornschem Lemma Basen in beliebigen Vektorrräumen studiert haben, sind manche der Resultate allgemeiner formuliert.

Sind V und W K-Vektorräume, so ist die Menge der linearen Abbildungen $Hom_K(V,W)$ nach Lemma 6.1(a) mittels der punkteweisen Addition und Skalarmultiplikation ebenfalls wieder ein K-Vektorraum. Ist $\{a_i \mid i \in I\}$ eine Basis von V und sind $\{c_i \mid i \in I\}$ beliebige Elemente von W, so bestimmt nach Lemma 5.4(a) die Zuordnung $f(a_i) = c_i$ ein Element $f \in \mathrm{Hom}_K(V,W)$.

Wir betrachten im folgenden den Spezialfall, wenn $W = K$ ist.

Definition 24.1. Sei V ein K-Vektorraum. Dann ist der K-Vektorraum

$$V^* = Hom_K(V,K)$$

der Dualraum zu V. Wir bezeichnen die Elemente von V^* als Linearformen auf V.

• Nach Proposition 5.4(a) ist V^* ein K-Vektorraum. Ist $\{a_1,\dots,a_n\}$ eine Basis von V, so bilden nach Proposition 5.4(b) die $f_i \in V^*$ mit

$$f_i(a_j) = \delta_{ij},\ i = 1,\dots,n$$

eine Basis von V^*; die Basis $\{f_1,\dots,f_n\}$ ist die zu $\{a_1,\dots,a_n\}$ duale Basis (und umgekehrt); insbesondere ist $\dim_K V = \dim_K V^* = n$.

Beispiele 24.2. (a) Sei $V = \mathbb{R}^2$ mit Basis $\{a_1,a_2\}$, $a_1 = (1,0)$ und $a_2 = (1,1)$. Dann ist die zu $\{a_1,a_2\}$ duale Basis $\{f_1,f_2\}$ gegeben durch

$$f_1(a_1) = 1,\ f_1(a_2) = 0,\ f_2(a_1) = 0 \text{ und } f_2(a_2) = 1.$$

Die zu der Standardbasis $\{e_1,e_2\}$ von V duale Basis hat die Form $\{f_1',f_2'\}$ mit $f_1'(e_1) = 1$, $f_1'(e_2) = 0$, $f_2'(e_1) = 0$ und $f_2'(e_2) = 1$. Dabei ist f_1 dual zu e_1, aber es ist $f_1 \neq f_1'$, da $f_1'(e_2) = 0$ und $f_1(e_2) = f_1(v_1) - f_1(v_2) = -1$ gilt; man benötigt somit zur Erstellung der dualen Basis die 'volle' Basis.

(b) Sei $V = \mathbb{R}[x]$ der $\mathbb{R}$-Vektorraum der Polynome mit reellen Koeffizienten und Basis $\{\, x^i \mid i \in \mathbb{N}_0\}$. Seien $f_i \in V^*$ mit $f_i(x^j) = \delta_{ij}$. Dann

lässt sich die Linearform $f \in V^*$ mit $f(x^i) = 1$ für alle i nicht als Linearkombination der f_i darstellen, d.h. $f \notin \langle f_i |\ i \in \mathbb{N}_0 \rangle$; insbesondere bilden die zu der Basis $\{x^i \mid i \in \mathbb{N}_0\}$ von V dualen Elemente keine Basis von V^*.

Lemma 24.3. *Sei V ein K-Vektorraum und $U \subseteq V$ ein linearer Unterraum. Ist $v \in V \setminus U$, so gibt es ein $f \in V^*$ mit $f(u) = 0$ für $u \in U$ und $f(v) = 1$.*

Beweis. Sei $\{u_j : j \in J\}$ eine Basis von U. Dann ist $\{u_j : j \in J\} \cup \{v\}$ linear unabhängig und lässt sich zu einer Basis

$$\{u_j : j \in J\} \cup \{v\} \cup \{v_i : i \in I\}$$

von V ergänzen. Definiere nun f vermöge $f(u_j) := 0, f(v) := c, f(v_i) := c_i$ mit beliebigen $c, c_i \in K$, $c \neq 0$. □

Proposition 24.4. *Sei V ein K-Vektorraum und $V^{**} = (V^*)^*$. Setze*

$$T : V \to V^{**},\ (Tv)(f) = f(v),\ v \in V,\ f \in V^*.$$

(a) *T ist ein Monomorphismus,*

(b) *Ist $\dim_K V = n < \infty$, so ist T ein Isomorphismus.*

Beweis. (a): Nachrechnen zeigt, dass $Tv \in V^{**}$ und T linear ist; zum Beispiel: Die Abbildung Tv ist additiv, da für $f_1, f_2 \in V^*$ gilt

$$(Tv)(f_1 + f_2) = (f_1 + f_2)(v) = f_1(v) + f_2(v) = (Tv)(f_1) + (Tv)(f_2).$$

Wir zeigen: $0 \neq v \in V \Rightarrow Tv \neq 0$ (d.h. T ist ein Monomorphismus). Ist $0 \neq v$, so gibt es nach Lemma 24.3 gibt es ein $f \in V^*$ mit $f(v) = 1$. Es folgt $(Tv)(f) = f(v) = 1 \neq 0$ und somit $Tv \neq 0$.

(b): Ist $\dim_K V = n$, so gilt $\dim_K V^* = \dim_K V^{**} = n$. Nach Lemma 5.11 ist der Monomorphismus T ein Isomorphismus. □

Definition 24.5. Sei V ein K-Vektorraum und sei $W \subseteq V$ ein linearer Unterraum

(1) Für $f \in V^*$ definiere die Restriktion $(Rf) \in U^*$ durch

$$(Rf)(w) = f(w),\ w \in W.$$

(2) Für $g \in (V/W)^*$ definiere die Inflation $Ig \in V^*$ durch

$$(Ig)(v) = g(v + W),\ v \in V.$$

(3) Ist $M \subseteq V$ eine Teilmenge, so setze

$$M^\perp = \{f \in V^* \mid f(m) = 0 \text{ für alle } m \in M\},$$

d.h. $M^\perp \subseteq V^*$ sind die Linearformen, die M annullieren.

(4) Ist $S \subseteq V^*$ eine Teilmenge, so setze

$$S^\top = \{v \in V \mid s(v) = 0 \text{ für alle } s \in S\},$$

d.h. $S^\top \subseteq V$ sind die Vektoren, die von Linearformen in S annulliert werden.

Bemerkung 24.6. $M^\perp$ und $S^\top$ sind lineare Unterräume. Es gilt:

$$M^\perp = \langle M \rangle^\perp, \quad S^\top = \langle S \rangle^\top.$$

Beispiel 24.7. Sei $V = \mathbb{R}^2$ und sei $M = \{(0,1)\} \subseteq V$. Dann ist

$$M^\perp = \{f \in V^* \mid f(0,1) = 0\}.$$

Ist $\{e_1, e_2\}$ die Standardbasis von V, so ist die duale Basis $\{f_1, f_2\}$ mit

$$f_1(e_1) = 1,\ f_1(e_2) = 0,\ f_2(e_1) = 0,\ f_2(e_2) = 1$$

eine Basis von V^*. Sei $v = (\alpha_1, \alpha_2) = \alpha_1 e_1 + \alpha_2 e_2$ ein Element von V und $f = \beta_1 f_1 + \beta_2 f_2$ ein Element von V^*. Dann ist

$$f(v) = (\beta_1 f_1 + \beta_2 f_2)(\alpha_1 e_1 + \alpha_2 e_2) = \beta_1 \alpha_1 + \beta_2 \alpha_2.$$

Für $v = e_2$ ist $\alpha_1 = 0$ und $\alpha_2 = 1$. Ist $f(e_2) = 0$, so folgt wegen $0 = f(e_2) = \beta_2$ dann $\beta_2 = 0$, d.h. f hat die Form $f = \beta_1 f_1$ und

$$M^\perp = \langle f_1 \rangle.$$

Betrachte $S = \{f_1\} \subseteq V^*$. Wegen $f_1(v) = \alpha_1$ folgt

$$S^\top = \{v \in V \mid f_1(v) = 0\} = \{v \in V \mid v = \alpha_2 e_2\} = \langle e_2 \rangle.$$

Lemma 24.8. *Sei V ein K-Vektorraum und $W \subseteq V$ ein linearer Unterraum. Dann gilt*

(a) *$R : V^* \to W^*$ ist linear mit $\ker(R) = W^\perp$.*

(b) *Ist $\dim_K V = n < \infty$, so ist R ein Epimorphismus. Es gilt*

$$\dim_K W^\perp = n - \dim_K W.$$

(c) *$I : (V/W)^* \to W^\perp$ ist ein Isomorphismus*

Beweis. (a): Nach Definition ist $(Rf)(w) = f(w)$ für $f \in V^*$ und $w \in W$. Wegen $f \in V^*$ ist $f|_W \in W^*$ und $R : V^* \to W^*,\ f \mapsto Rf$ ist eine Abbildung. Man rechnet nach, dass R linear ist. Weiter ist

$$\ker(R) = \{f \in V^* \mid Rf = f|_W = 0\ \} = W^\perp.$$

(b): Sei $\{w_1, \ldots, w_r\}$ eine Basis von W und $\{w_1, \ldots, w_r, v_1, \ldots, v_s\}$ eine Basis von V. Erweitere $f \in W^*$ auf $g \in V^*$ durch $g(w_i) = f(w_i)$ und $g(v_i) = 0$. Damit ist $g(w) = f(w)$ für $w \in W$, d.h. $f = Rg$ und R ist surjektiv. Wegen $Bild(R) = W^*$, $\ker(R) = W^\perp$, und $\dim_K W = \dim_K W^*$ liefert der Homomorphiesatz die Identität

$$\dim_K W^\perp = n - \dim_K W.$$

(c): Für $g \in (V/W)^*$ und $w \in W$ ist $(Ig)(w) = g(w+W) = g(W) = 0$, also ist $Ig \in W^\perp$ und I definiert eine Abbildung $(V/W)^* \to W^\perp$; Nachrechnen zeigt, dass I linear ist. Sei $f \in W^\perp$. Dann definiert $\bar{f}(v+W) = f(v)$ ein Element von $(V/W)^*$ mit $I\bar{f} = f$, also ist I surjektiv. Sei $g \in \ker(I)$. Nach Definition von I gilt für $v \in V$ dann

$$0 = (Ig)(v) = g(v+W),$$

d.h. $g = 0$. Somit ist I injektiv. □

Theorem 24.9. *(Dualitätssatz) Sei V ein K-Vektorraum. Dann gilt*

(a) *Ist $W \subseteq V$ ein linearer Unterraum, so ist $W^{\perp\top} = W$.*

(b) *Sei $\dim(V) < \infty$. Ist $S \subseteq V^*$ ein linearer Unterraum, so ist $S^{\top\perp} = S$.*

(c) *Sind W_1 und W_2 lineare Unterräume von V (V beliebig), so ist*

$$(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp.$$

Falls $\dim(V) < \infty$, so gilt auch

$$(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp.$$

Beweis. (a): Es ist $W^{\perp\top} = \{v \in V \mid f(v) = 0 \text{ für } f \in W^\perp\} \supseteq W$. Ist $v \in W^{\perp\top} \setminus W$, so gibt es nach Lemma 24.3 ein $f \in V^*$ mit $f(v) = 1$ und $f(w) = 0$ für $w \in W$. Also ist $f \in W^\perp$. Wegen $0 = f(v) = 1$ ist dies ein Widerspruch, d.h. es gilt Gleichheit $W = W^{\perp\top}$.
(b): Nach Definition von $S^{\top\perp}$ gilt $S^{\top\perp} \supseteq S$. Da S und $S^{\top\perp}$ lineare Unterräume sind genügt zu zeigen, dass $\dim_K S^{\top\perp} = \dim_K S$ ist. Nach Proposition 24.4 ist $T : V \to V^{**}$, $(Tv)(f) = f(v)$, $v \in V, f \in V^*$ ein Isomorphismus. Insbesondere ist $(Tv)(s) = s(v)$ für $s \in S$. Dies liefert

$$S^\top = \{v \in V \mid s(v) = 0 \text{ für } s \in S\} = \{v \in V \mid (Tv)(s) = 0 \text{ für } s \in S\}.$$

Da T ein Isomorphismus ist folgt

$$\begin{aligned} \dim_K S^\top &= \dim_K T(\{v \in V \mid (Tv)(s) = 0 \text{ für } s \in S\}) \\ &= \dim_K \{f \in V^{**} \mid f(s) = 0 \text{ für } s \in S\} \\ &= \dim_K S^\perp \\ &= \dim_K V^* - \dim_K S, \end{aligned}$$

wobei die letzten Gleichung aus Lemma 24.8(b) (angewandt auf $W = S$ und $V = V^*$) folgt. Nach nochmaliger Anwendung von Lemma 24.8(b) (mit $W = S^\top$ und $V = V$) folgt wegen $\dim_K V = \dim_K V^*$ dann

$$\dim_K S^{\top\perp} = \dim_K V - \dim_K S^\top = \dim_K S,$$

und damit $S = S^{\top\perp}$.

(c): Die Gleichheit $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ folgt direkt aus der Definition. Sei nun $\dim(V) < \infty$. Wir wollen $W_1^\perp + W_2^\perp = (W_1 \cap W_2)^\perp$ beweisen. Die Inklusion $\subseteq$ folgt wieder direkt aus der Definition. Es reicht also zu zeigen, dass

$$\dim(W_1^\perp + W_2^\perp) = \dim(W_1 \cap W_2)^\perp$$

gilt. Hierzu werden wir mehrfach die Formel

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

sowie Lemma 24.8(b). Es folgt

$$\begin{aligned}
\dim(W_1^\perp + W_2^\perp) &= \dim(W_1^\perp) + \dim(W_2^\perp) - \dim(W_1^\perp \cap W_2^\perp) \\
&= \dim(V) - \dim(W_1) + \dim(V) - \dim(W_2) - \dim(W_1 + W_2)^\perp \\
&= 2\dim(V) - \dim(W_1) - \dim(W_2) - \dim(V) + \dim(W_1 + W_2) \\
&= \dim(V) - \dim(W_1) - \dim(W_2) + \dim(W_1 + W_2) \\
&= \dim(V) - \dim(W_1 \cap W_2) \\
&= \dim(W_1 \cap W_2)^\perp.
\end{aligned}$$

□

Als eine Anwendung des Dualitätssatzes betrachten wir affine Unterräume eines n-dimensionalen K Vektorraums V. Nach Definition hat ein m-dimensionaler affiner Unterraum von V die Form $H = a + W$, wobei $a \in V$ und $W \subseteq V$ ein m-dimensionaler linearer Unterraum ist. Ist $\dim_K W = n - 1$, so ist $H = a + W$ eine affine Hyperebene.

Die affinen Hyperebenen lassen sich mittels Linearformen wie folgt beschreiben: Sei $0 \neq f \in V^*$ und sei $\alpha \in K$. Wegen $0 \neq f$ gibt es ein $w \in V$ mit $f(w) \neq 0$. Setze $a = \alpha f(w)^{-1} w$, so dass $f(a) = \alpha$ ist. Wegen $\dim_K \langle f \rangle^\top = \dim_K \{v \in V \mid f(v) = 0\} = \dim_K \ker(f) = n - 1$ definiert

$$H_{f,\alpha} = \{v \in V \mid f(v) = \alpha\} = a + \ker(f)$$

eine affine Hyperebene in V. Umgekehrt hat jede affine Hyperebene $H \subseteq V$ diese Form: Sei dazu $H = a + W$. Wegen $\dim W = n - 1$ ist $\dim W^\perp = 1$ und $W^\perp$ wird von einem $0 \neq f \in V^*$ erzeugt. Für $\alpha = f(a)$ folgt

$$H = H_{f,\alpha},$$

denn für $a + w \in H$ gilt offensichtlich $f(a + w) = f(a) + f(w) = \alpha + 0 = \alpha$. Ist umgekehrt $v = a + (v - a) \in H_{f,\alpha}$, so ist zu zeigen, dass $v - a \in W$. Wegen $f(v - a) = f(v) - f(a) = \alpha - \alpha = 0$ folgt $v - a \in \langle f \rangle^\perp = W^{\perp^\top} = W$ (nach Satz 24.9).

Dieser Sachverhalt läßt sich verallgemeinern.

Theorem 24.10. *Sei V ein n-dimensionaler K-Vektorraum. Dann ist jeder m-dimensionale affine Unterraum der Durchschnitt von $n-m$ affinen Hyperebenen.*

• Die Aussage der Theorems ist für $V = \mathbb{R}^n$, $n = 2, 3$, geometrisch offensichtlich: Im Fall $V = \mathbb{R}^2$ sind die affine Hyperebenen die Geraden im $\mathbb{R}^2$. Das Theorem besagt, dass jeder Punkt ein Schnitt von 2 Geraden ist. Für $V = \mathbb{R}^3$ sind die affinen Hyperebenen die Flächen im $\mathbb{R}^3$; jeder Punkt ist der Schnitt von 3 Flächen, und jede Gerade ist der Schnitt von 2 Flächen.

Beweis. Sei $H = a + W \subseteq V$, $a \in V$, $W \subseteq V$ ein m-dimensionaler linearer Unterraum. Nach Lemma 24.8(b) ist $\dim_K W^\perp = n - m$. Sei $\{f_1, \ldots, f_{n-m}\}$ eine Basis von $W^\perp$; für $i = 1, \ldots, n-m$ setze $\alpha_i = f_i(a)$. Dies liefert $n - m$ Hyperebenen H_{f_i,α_i}. Wir zeigen $H = \cap_{i=1}^{n-m} H_{f_i,\alpha_i}$.

Ist $v = a + w \in H$, so gilt für $i = 1, \ldots, n - m$ nach Definition

$$f_i(v) = f_i(a) + f_i(w) = f_i(a) + 0 = \alpha_i.$$

Also ist $v \in \cap_{i=1}^{n-m} H_{f_i,\alpha_i}$. Ist umgekehrt $v \in \cap_{i=1}^{n-m} H_{f_i,\alpha_i}$, so folgt wegen $f_i(v - a) = 0$ dann $v - a \in \langle f_i \rangle^\top$ und somit $v - a \in \langle f_1, \ldots, f_{n-m} \rangle^\top$. Der Dualitätssatz 24.9(a) zeigt $\langle f_1, \ldots, f_{n-m} \rangle^\top = W^{\perp\top} = W$, also ist $v - a \in W$ und $v \in a + W = H$. □

Bemerkung 24.11. Für lineare Gleichungssysteme ergibt sich mittels Dualräume folgende geometrische Interpretation: Für ein System

$$(\mathrm{L}): \qquad \sum_{j=1}^{n} \alpha_{ij} x_j = \beta_i, \quad i = 1, \ldots, m,$$

betrachte die m Gleichungen separat. Die i-te Gleichung hat die Form

$$(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})(x_1, \ldots, x_n)^t = \beta_i$$

Interpretiert man $x = (x_1, \ldots, x_n)^t$ als Vektor in $V = K^n$, so definiert $a_i = (\alpha_{i1}, \ldots, \alpha_{in})$ eine lineare Abbildung

$$f_i \colon V \to K, \quad x \mapsto a_i x.$$

Damit lässt sich das System (L) schreiben als

$$(\mathrm{L}): \qquad f_i(x) = \beta_i, \quad i = 1, \ldots, m, \quad f_i \in V^*, \ x \in V.$$

Die Lösungsmenge von (L) ist genau der Schnitt der m Hyperebenen H_{f_i,β_i}. Das System (L) hat eine Lösung falls dieser Schnitt nicht leer ist, und eine eindeutige Lösung, falls dieser Schnitt aus genau einem Punkt besteht.

Bemerkung 24.12. Als Anwendung geben wir ein Verfahren zur Berechnung des Durchschnitts $U_1 \cap U_2$ von linearen Unterräumen $U_1, U_2 \subseteq V$, $\dim(V) = n < \infty$, an. Nach Wahl einer Basis von V können wir oE $V = K^n$ annehmen. Sei

$$U_1 = \langle a_1, \ldots, a_s \rangle \text{ und } U_2 = \langle b_1, \ldots, b_t \rangle$$

mit Spaltenvektoren a_i und b_j. Wir bestimmen nun Linearformen

$$f_1, \ldots, f_k, g_1, \ldots g_l \in V^*,$$

so dass gilt

$$U_1 = \langle f_1, \ldots, f_k \rangle^\top, \quad U_2 = \langle g_1, \ldots, g_l \rangle^\perp.$$

Dann gilt $U_1 \cap U_2 = \langle f_1, \ldots, f_k, g_1, \ldots, g_l \rangle^\perp$.

Dazu seien $A = (a_1, \ldots, a_k) \in K^{n \times k}$ und $B = (b_1, \ldots, b_l) \in K^{n \times l}$ die Matrizen mit den Spalten a_i bzw. b_j. Der Lösungsraum der linearen Gleichungssysteme

$$A^t x = 0 \text{ und } B^t y = 0$$

liefert dann die Linearformen $f_1, \ldots, f_k$ bzw. $g_1, \ldots, g_l$.

Sei nun $C \in K^{(k+l) \times n}$ die Matrix mit den Zeilen $f_1, \ldots, f_k, g_1, \ldots, g_l$. Dann ist der Lösungsraum des linearen Gleichungsystems $Cz = 0$ genau der Durchschnitt $U_1 \cap U_2$.

Beispiel 24.13. Seien

$$\begin{aligned} U_1 &= \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix} \right\rangle, \\ U_1 &= \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\rangle. \end{aligned}$$

Dann sind

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 2 & 4 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

und die Lösung der entsprechenden linearen Gleichungssysteme liefert

$$f_1 = (-2, 1, 0), \quad g_1 = (-1, 1, 0).$$

Es gilt also $U_1 = \langle f_1 \rangle^\top$ und $U_2 = \langle g_1 \rangle^\top$. Also ist

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

und die Lösung von $Cz = 0$ liefert

$$U_1 \cap U_2 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Definition 24.14. Seien V, W K-Vektorräume und sei $g \in \mathrm{Hom}_K(V, W)$. Wir definieren

$$g^* : W^* \to V^*,\ f \mapsto f \circ g.$$

Die Abbildung g^* heißt die zu g duale Abbildung.

- Man zeigt leicht, dass $g^* \colon W^* \longrightarrow V^*$ eine lineare Abbildung ist.
- Sind V, W K-Vektorräume, $g, h \in \mathrm{Hom}_K(V, W)$ und $\alpha \in K$, so gilt

$$(g + h)^* = g^* + h^* \text{ und } (\alpha g)^* = \alpha g^*.$$

- Seien V_i K-Vektorräume, $i = 1, 2, 3$, $g \in \mathrm{Hom}_K(V_1, V_2)$ und $h \in \mathrm{Hom}_K(V_2, V_3)$. Dann gilt für das Kompositum der dualen Abbildungen

$$(h \circ g)^* = g^* \circ h^*.$$

Wir zeigen: Hat $g \in \mathrm{Hom}_K(V, W)$ bzgl. Basen von V und W die Matrix (α_{ij}), so hat g^* bzgl. der dualen Basen die Matrix $(\alpha_{ij})^t$.

Lemma 24.15. *Seien V, W K-Vektorräume, $\{v_1, \ldots, v_n\}$ eine Basis von V, $\{w_1, \ldots, w_m\}$ eine Basis von W, und $g \in \mathrm{Hom}_K(V, W)$ mit*

$$g(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i, \quad j = 1, \ldots, n.$$

Ist $\{f_1, \ldots, f_n\}$ die duale Basis zu $\{v_1, \ldots, v_n\}$ und $\{g_1, \ldots, g_m\}$ die duale Basis zu $\{w_1, \ldots, w_m\}$, so gilt für die duale Abbildung g^*

$$g^*(g_i) = \sum_{j=1}^{n} \alpha_{ij} f_j, \quad i = 1, \ldots, m.$$

Beweis. Nach Definition der dualen Abbildung gilt

$$(g^*(g_i))(v_k) = g_i(g(v_k)) = g_i(\sum_{l=1}^{m} \alpha_{lk} w_l) = \sum_{l=1}^{m} \alpha_{lk}(g_i(w_l)).$$

Nach Definition gilt $g_i(w_l) = \delta_{il}$ und es folgt

$$(g^*(g_i))(v_k) = \alpha_{ik}.$$

Wegen $f_j(v_k) = \delta_{jk}$ folgt andererseits

$$\left(\sum_{j=1}^{n} \alpha_{ij} f_j\right)(v_k) = \sum_{j=1}^{n} \alpha_{ij} f_j(v_k) = \alpha_{ik}.$$

□

Ist V ein K-Vektorraum, so gilt nach Definition 24.5(3) und (4)

$$M^\perp = \{f \in V^* \mid f(m) = 0 \text{ für alle } m \in M\} \text{ für } M \subseteq V,$$
$$S^\top = \{v \in V \mid s(v) = 0 \text{ für alle } s \in S\} \text{ für } S \subseteq V^*.$$

Proposition 24.16. *Seien V, W K-Vektorräume und sei $g \in \mathrm{Hom}_K(V, W)$. Dann gilt*

(a) $\ker(g^*) = \mathit{Bild}(g)^\perp$.
(b) $\ker(g) = \mathit{Bild}(g^*)^\top$.
(c) *g ist Epimorphismus $\Leftrightarrow$ g^* ist Monomorphismus.*
(d) *g ist Monomorphismus $\Leftrightarrow$ g^* ist Epimorphismus.*
(e) *g ist Isomorphismus $\Leftrightarrow$ g^* ist Isomorphismus.*

Beweis. (a): Aus den Definitionen folgt

$$\begin{aligned} \ker(g^*) &= \{f \in W^* \mid g^*(f)(v) = f(g(v)) = 0 \text{ für alle } v \in V\} \\ &= \{f \in W^* \mid f(w) = 0 \text{ für alle } w \in \mathit{Bild}(g)\} \\ &= \mathit{Bild}(g)^\perp. \end{aligned}$$

(b): Eine Richtung ist offensichtlich, da

$$\begin{aligned} \mathit{Bild}(g^*)^\top &= \{v \in V \mid g(v) = 0 \text{ für alle } g \in \mathit{Bild}(g^*)\} \\ &= \{v \in V \mid g^*(f)(v) = f(g(v)) = 0 \text{ für alle } f \in W^*\} \\ &\supseteq \ker(g). \end{aligned}$$

Ist $v \in \mathit{Bild}(g^*)^\top$, so gilt $g^*(f)(v) = f(g(v)) = 0$ für alle $f \in W^*$. Aus Lemma 24.3, angewandt auf $V = \mathit{Bild}(g)$ und $U = \{0\}$, folgt $g(v) = 0$, also ist $v \in \ker(g)$ und es gilt Gleichheit.

Die Teile (c), (d) und (e) sind Übungblatt 12 der Linearen Algebra I, Aufgabe 4. □

25. Bilinearformen und Dualität

Definition 25.1. Seien V, W zwei K-Vektorräume. Eine Abbildung $\beta : V \times W \longrightarrow K$ heißt Bilinearform, wenn sie linear in jedem Argument

ist, d.h.

$$\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w), \forall v_1, v_2 \in V, w \in W$$
$$\beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2), \forall v \in V, w_1, w_2 \in W$$
$$\beta(av, w) = a\beta(v, w) = \beta(v, aw), \forall v \in V, w \in W, a \in K.$$

Beispiel 25.2. 1) Sei V ein beliebiger K-Vektorraum und V^* sein Dualraum. Dann ist

$$\beta\colon V^* \times V \longrightarrow K, \quad (f, v) \mapsto f(v),$$

eine Bilinearform.

2) Sei $K = \mathbb{R}$ und V ein innerer Produktraum mit innerem Produkt $(\ ,\ )$. Dann wird durch $\beta(v_1, v_2) := (v_1, v_2)$ eine Bilinearform

$$\beta\colon V \times V \longrightarrow K$$

definiert. Jedes innere Produkt ist also insbesondere eine Bilinearform.

3) Sei K wieder beliebig. Dann wird durch

$$\beta\colon K^n \times K^n \longrightarrow K, \quad \beta(x, y) := \sum_{i=1}^{n} x_i y_i,$$

eine Bilinearform definiert.

Definition 25.3. Sei $\beta : V \times W \longrightarrow K$ eine Bilinearform. Dann heißt β ausgeartet in der ersten Variablen, falls es ein $v \in V, v \neq 0$, gibt, so daß $\beta(v, w) = 0$ für alle $w \in W$. Analog hierzu heißt β ausgeartet in der zweiten Variablen, falls es ein $w \in W, w \neq 0$, gibt, so daß $\beta(v, w) = 0$ für alle $v \in V$.

Beispiele 25.4. In 1) - 3) nehmen wir Bezug auf die Beispiele 25.2.

1) Hier ist β offensichtlich nicht ausgeartet in der ersten Variablen. Die Form ist ebenfalls nicht ausgeartet in der zweiten Variablen aufgrund von Lemma 24.3.

2) Wegen $(v, v) > 0$ für alle $v \neq 0$, ist jedes innere Produkt in beiden Variablen nicht ausgeartet.

3) Setzt man $y = e_i$ (Standardbasis), so folgt aus $0 = \beta(x, e_i) = x_i$ sofort, dass β nicht ausgeartet im ersten Argument ist. Analog sieht man ein, dass die Form auch im zweiten Argument nicht ausgeartet ist.

4) Sei $V = \{f\colon \mathbb{R} \longrightarrow \mathbb{R} \mid f \text{ stetig } \}$ der VR der stetigen Funktionen auf $\mathbb{R}$. Sei

$$\beta\colon V \times V \longrightarrow \mathbb{R}, \quad \beta(f, g) := \int_0^1 f(x)g(x)dx.$$

Dann ist β ausgeartet in beiden Argumenten.

5) Sei

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

und

$$\beta\colon K^3 \times K^2 \longrightarrow K, \quad \beta(x,y) := x^t B y.$$

Dann ist β ausgeartet in der ersten Variablen, da für $x = (0,0,1)^t$ offensichtlich $\beta(x,y) = 0$ für alle $y \in \mathbb{R}^2$ gilt. Andererseits ist β nicht ausgeartet in der zweiten Variablen, da $\beta(e_1, y) = y_1$ und $\beta(e_2, y) = y_2$ gilt.

Wie lineare Abbildungen zwischen endlich dimensionalen Vektorräumen, so lassen sich auch Bilinearformen durch eine Matrix beschreiben.

Definition 25.5. Sei $\dim(V) = n < \infty$ und $\dim(W) = m < \infty$. Seien $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ Basen von V und W und sei $\beta : V \times W \longrightarrow K$ eine Bilinearform. Dann heißt die $n \times m$-Matrix

$$B = (\beta(v_i, w_j))_{i=1,\ldots,n, j=1,\ldots,m}$$

die Strukturmatrix von β bezüglich der Basen $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$.

Falls $V = W$ so spricht man von der Strukturmatrix bezüglich der Basis $v_1, \ldots, v_n$.

Bemerkung 25.6. Die Strukturmatrix ist abhängig von der Wahl der Basen. Wir werden später diese Abhängigkeit genau analysieren.

Beispiele 25.7. 1) Sei V ein endlich-dimensionaler VR mit Basis $v_1, \ldots, v_n$. Sei $f_1, \ldots, f_n \in V^*$ die dazu duale Basis von V^*. Dann ist die Strukturmatrix der Bilinearform $\beta\colon V^* \times V \longrightarrow K, \beta(f, v) := f(v)$, durch die Einheitsmatrix gegeben.

2) Sei $K = \mathbb{R}$ und V ein endlich-dimensionaler $\mathbb{R}$-VR mit innerem Produkt $(\ ,\ )$. Sei $\beta\colon V \times V \longrightarrow \mathbb{R}$ gegeben durch $\beta(v_1, v_2) := (v_1, v_2)$. Sei $a_1, \ldots, a_n$ eine ON-Basis von V. Dann ist die Strukturmatrix von β bezüglichder Basis $a_1, \ldots, a_n$ durch die Einheitsmatrix gegeben.

3) Sei $A \in K^{n \times m}$ und

$$\beta\colon K^n \times K^m \longrightarrow K, \quad (x,y) \mapsto x^t A y.$$

Dann ist die Strukturmatrix von β bezüglich der Standardbasen durch die Matrix A gegeben.

Der folgende Satz zeigt, daß die Bilinearform β durch ihre Strukturmatrix vollständig bestimmt ist.

Theorem 25.8. *Sei* $\dim(V) = n < \infty$ *und* $\dim(W) = m < \infty$. *Seien* $v_1, \ldots, v_n$ *und* $w_1, \ldots, w_m$ *Basen von* V *und* W *und sei* $\beta : V \times W \longrightarrow K$ *eine Bilinearform. Sei* B *die zugehörige Strukturmatrix. Sei*

$$\begin{aligned} v &= x_1 v_1 + \ldots + x_n v_n, x_i \in K, \\ w &= y_1 w_1 + \ldots + y_m w_m, y_j \in K. \end{aligned}$$

Dann gilt:

$$\beta(v, w) = (x_1, \ldots, x_n) B \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Umgekehrt definiert jede Matrix $B \in K^{n,m}$ *auf diese Weise eine Bilinearform* $\beta : V \times W \longrightarrow K$.

Beweis. Die Bilinearität von β impliziert

$$\beta(v, w) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j \beta(v_i, w_j) = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} \beta(v_i, w_j) y_j.$$

Hieraus folgt die Behauptung. □

Korollar 25.9. *Sei* $\dim(V) = \dim(W) = n < \infty$ *und* $\beta : V \times W \longrightarrow K$ *eine Bilinearform mit Strukturmatrix* B *(bezüglich gewählter Basen). Dann sind die folgenden Aussagen äquivalent:*

a) β *ist ausgeartet im ersten Argument.*
b) β *ist ausgeartet im zweiten Argument.*
c) $\det(B) = \det(B^t) = 0$.
d) $\mathrm{rg}(B) = \mathrm{rg}(B^t) < n$.
e) $\ker(B) \neq \{0\}$.
f) $\ker(B^t) \neq \{0\}$.

Wie das Beispiel 25.4, 5) zeigt, ist die Voraussetzung $\dim(V) = \dim(W)$ von entscheidender Bedeutung.

Beweis. Sei β ausgeartet im ersten Argument. Dies ist genau dann der Fall, wenn es ein $v = \sum_{i=1}^{n} x_i v_i \neq 0$ gibt, so dass für alle $w \in W$ die Gleichheit $\beta(v, w) = 0$ gilt. Dies ist äquivalent dazu, dass es ein $0 \neq x \in K^n$ gibt, so dass $x^t B y = 0$ für alle $y \in K^n$ ist. Dies ist äquivalent zur Existenz eines $0 \neq x \in K^n$ mit $x^t B = 0$, was wiederum äquivalent zur Aussage $\ker(B^t) \neq 0$ ist. Dies ist äquivalent zu $\det(B^t) = 0$. Wegen $\det(B) = \det(B^t)$, ist dies wiederum äquivalent zu $\ker(B) \neq 0$ und wir können jetzt wie eben schließen, dass dies äquivalent dazu ist, dass β ausgeartet im zweiten Argument ist. □

Während die Aussagen a) und b) unabhängig von der Wahl von Basen sind, scheinen die restlichen Aussagen von B und damit von der Basiswahl abhängig zu sein. Der folgende Satz zeigt jedoch, daß dies nicht der Fall ist.

Theorem 25.10. *Sei* $\dim(V) = n < \infty$ *und* $\dim(W) = m < \infty$. *Seien* $v_1, \ldots, v_n$ *und* $w_1, \ldots, w_m$ *Basen von* V *und* W *und sei* $\beta : V \times W \longrightarrow K$ *eine Bilinearform. Sei* B *die zugehörige Strukturmatrix. Seien* $v'_1, \ldots, v'_n$ *und* $w'_1, \ldots, w'_m$ *weitere Basen mit Übergangsmatrizen* $S \in \mathrm{Gl}_n(K)$ *und* $T \in \mathrm{Gl}_m(K)$. *Sei* B' *die Srukturmatrix bezüglich der neuen Basen. Dann gilt:*

$$B' = S^t B T.$$

Beweis. Es sei

$$v'_i = \sum_{j=1}^{n} s_{ji} v_j, \quad w'_j = \sum_{l=1}^{m} t_{lj} w_l.$$

Dann folgt

$$\begin{aligned} \beta(v'_i, w'_j) &= \beta\left(\sum_{j=1}^{n} s_{ji} v_j, \sum_{l=1}^{m} t_{lj} w_l\right) \\ &= \sum_{j=1}^{n}\sum_{l=1}^{m} s_{ji} t_{lj} \beta\left(v_j, w_l\right) \\ &= \sum_{l=1}^{m}\left(\sum_{j=1}^{n} s_{ji} \beta(v_j, w_l\right) t_{lj} \end{aligned}$$

Hieraus folgt die Behauptung. □

Von besonderer Bedeutung ist der Spezialfall $V = W, \dim(V) = n < \infty$. Hier gilt $B' = S^t B S$, wobei $v_1, \ldots, v_n$ und $v'_1, \ldots, v'_n$ Basen von V sind mit Übergangsmatrix S.

Definition 25.11. a) Sei $\beta : V \times W \longrightarrow K$ eine Bilinearform. Dann heißt

$$\beta^t : W \times V \longrightarrow K, \quad \beta^t(w, v) = \beta(v, w)$$

die zu β transponierte Bilinearform.

b) Eine Bilinearform $\beta : V \times V \longrightarrow K$ heißt symmetrisch, falls $\beta^t = \beta$.

Bemerkung 25.12. a) Falls β bezüglich gewählter Basen die Strukturmatrix B hat, so hat β^t bezüglich dieser Basen die Strukturmatrix B^t.

b) $\beta : V \times V \longrightarrow K$ ist genau dann symmetrisch, wenn die Strukturmatrix B symmetrisch ist.

Definition 25.13. a) Sei $\beta : V \times V \longrightarrow K$ eine symmetrische Bilinearform. Ein Vektor $v \in V$ heißt orthogonal zu $w \in V$, falls $\beta(v, w) = 0$.

b) Sei $X \subseteq V$. Dann heißt

$$X^{\perp} = \{v \in V \mid \beta(v, x) = 0, \forall x \in X\}$$

das orthogonale Komplement von X bezüglich β.

Einfache Beobachtungen:

a) $X^{\perp} = \langle X \rangle^{\perp}$.

b) $X^{\perp}$ ist ein linearer Unterraum von V.

Die Beweise dieser Beobachtungen als auch der Beweis des folgenden Satzes (vgl. Lemma 24.8 und Satz 24.9) sind völlig analog zu den entsprechenden Aussagen für Dualräume.

Theorem 25.14. *Sei $\beta : V \times V \longrightarrow K$ eine nicht ausgeartete symmetrische Bilinearform auf dem endlich dimensionalen Vektorraum V. Sei $X \subseteq V$ ein linearer Unterraum. Dann gilt:*

$$\dim(V) = \dim(X) + \dim(X^{\perp}).$$

Außerdem gilt: $\left(X^{\perp}\right)^{\perp} = X$.

Schließlich wollen wir den Zusammenhang von nicht ausgearteten Bilinearformen zu Dualräumen untersuchen.

Lemma 25.15. *Sei $\beta \colon V \times W \longrightarrow K$ eine nicht-ausgeartete Bilinearform. Dann sind die folgenden Aussagen äquivalent:*

$$(a) \qquad \dim(V) < \infty,$$
$$(b) \qquad \dim(W) < \infty.$$

Falls diese äquivalenten Bedingungen erfüllt sind, so gilt ferner:

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

.

Beweis. Wir zeigen zunächst, dass die Abbildung

$$W \longrightarrow V^*, \quad w \mapsto f_w,$$

wobei $f_w(v) := \beta(v, w)$, injektiv ist. Man sieht leicht ein, dass die Abbildung linear ist. Es gilt:

$$f_w = 0 \iff \beta(v, w) = 0 \text{ für alle } v \in V.$$

Da β nicht ausgeartet im zweiten Argument ist, folgt $w = 0$. Also hat $w \mapsto f_w$ trivialen Kern und ist damit injektiv.

Sei nun $\dim(V) < \infty$. Dann folgt $\dim(W) \leq \dim(V^*) = \dim(V)$. Da β nach Voraussetzung auch nicht ausgeartet im ersten Argument ist, zeigt man analog $\dim(V) \leq \dim(W^*) = \dim(W)$. □

Theorem 25.16. *Sei* $\dim(V) = \dim(W) < \infty$ *und* $\beta\colon V \times W \longrightarrow K$ *eine nicht-ausgeartete Bilinearform. Dann ist die Abbildung*

$$W \longrightarrow V^*, \quad w \mapsto f_w, \text{ wobei } f_w(v) := \beta(v, w),$$

ein Isomorphismus von $K - VR$.

Beweis. Wir haben bereits im Beweis zum Lemma 25.15 eingesehen, dass die Abbildung injektiv ist. Daher folgt die Isomorphie bereits aus $\dim(V^*) = \dim(V) = \dim(W)$, wobei die zweite Gleichheit ebenfalls in 25.15 enthalten ist. □

In diesem Zusammenhang sind auch die folgenden Ergebnisse von Bedeutung.

Lemma 25.17. *Sei* $\beta\colon V \times W \longrightarrow K$ *eine Bilinearform, die nicht ausgeartet in der zweiten Variablen ist. Sei* $W_1 \subseteq W$ *ein linearer Unterraum und*

$$W_1^\perp := \{v \in V \mid \beta(v, w_1) = 0 \text{ for all } w_1 \in W_1\}.$$

Dann ist $W_1^\perp \subseteq V$ *ein linearer Unterraum und* β *induziert eine wohldefinierte und nicht-ausgeartete Bilinearform*

$$\begin{aligned} \bar{\beta}\colon V/W_1^\perp \times W_1 &\longrightarrow K, \\ (v + W_1^\perp, w_1) &\mapsto \beta(v, w_1). \end{aligned}$$

Beweis. Einfache Verifikation zeigt, dass $W_1^\perp \subseteq V$ ein linearer Unterraum ist. Sei $v_1 \in W_1^\perp$. Da dann

$$\beta(v + v_1, w_1) = \beta(v, w_1) + \beta(v_1, w_1) = \beta(v, w_1)$$

für alle $v \in V, w_1 \in W_1$ gilt, ist $\bar{\beta}$ wohldefiniert.

Sei $\bar{\beta}(v + W_1^\perp, w_1) = 0$ für alle $w_1 \in W_1$. Dann ist $\beta(v, w_1) = 0$ für alle w_1 und damit $v \in W_1^\perp$. Somit ist $v + W_1^\perp = 0$ in $V/W_1^\perp$ und $\bar{\beta}$ nicht ausgeartet im ersten Argument.

Sei $\bar{\beta}(v + W_1^\perp, w_1) = 0$ für alle $v \in V$. Dann ist $\beta(v, w_1) = 0$ für alle $v \in V$. Da β nicht ausgeartet im zweiten Argument ist, folgt $w_1 = 0$. Also ist auch $\bar{\beta}$ nicht ausgeartet im zweiten Argument. □

Falls also $\beta\colon V \times W \longrightarrow K$ eine Bilinearform ist, die in der zweiten Variablen nicht ausgeartet ist, und zudem ein linearer Unterraum $W_1 \subseteq W$ mit $\dim(W_1) < \infty$ gegeben ist, so folgt aus Satz 25.16 zusammen mit dem letzten Lemma:

$$\left(V/W_1^\perp\right)^* \simeq W_1.$$