Lineare Algebra

gehalten 2023/24 von Prof. Dr. Weitze-Schmithüsen

UNIVERSITÄT
DES
SAARLANDES

Hinweise

Das vorliegende Skript ist nicht wertvoller als eine handschriftliche Mitschrift und ersetzt keinesfalls das eigenständige Besuchen der Vorlesung oder das selbstständige Nachbereiten. Computersatz ist kein Garant für Fehlerfreiheit!

Diese Mitschrift wird von einem Studenten erstellt, Tippfehler können natürlich nicht ausgeschlossen werden. Hinweise auf Fehler sind daher ausdrücklich erwünscht:

guenther@math.uni-sb.de

Inhaltsverzeichnis

Teil 1.

Lineare Algebra I

Kapitel I.

Grundlagen

Im Folgenden verwenden wir etliche Symbole: „$\Leftrightarrow$“ steht für „ist äquivalent“, „$\Rightarrow$“ steht für „daraus folgt“, „$:=$“ steht für „wird definiert als“, „$:\Leftrightarrow$“ steht für „wird definiert durch die Eigenschaft“.

1. Voraussetzungen aus Mengentheorie und Aussagenlogik

Für die Vorlesung setzen wir voraus:

1.1. Naive Mengenlehre

Eine Menge besteht aus Objekten. Diese werden als *Elemente* bezeichnet.

Beispiel:
- Die Menge der natürlichen Zahlen $\mathbb{N}$ (ohne 0),
- Die Menge der ganzen Zahlen $\mathbb{Z}$,
- Die Menge der rationalen Zahlen $\mathbb{Q}$,
- Die Menge der reellen Zahlen $\mathbb{R}$,
- Die Menge $M_1 = \{1, 2, 7, 11\}$,
- Die Menge $M_2 = \{\text{Saarbrücken}, \text{Neunkirchen}, \text{Bexbach}, \text{Köln}\}$,
- Die Menge $M_3 = \{\{1, 2\}, \{1, 7\}, \{1, 2, 7, 11\}\}$.

Die Schreibweise „$x \in M$“ bedeutet, dass x ein Element der Menge M ist.

Zwei wichtige Prinzipien für die Mengenlehre sind

(i) *Extensionalität*: Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente haben. Mit anderen Worten: $M_1 = M_2$ gilt genau dann, wenn gilt: $(x \in M_1 \Leftrightarrow x \in M_2)$.

(ii) *Aussonderungsaxiom*: Zu jeder Menge M_1 und jeder Aussage P über Elemente von M_1 gibt es eine Menge M_2, sodass gilt: M_2 besteht genau aus den Elementen von M_1, für die die Aussage P wahr ist. Wir schreiben

$$M_2 = \{x \in M_1 \mid P(x) \text{ ist wahr}\}.$$

Beispiel: $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$.

Es gibt genau eine Menge, die keine Elemente enthält. Diese heißt die *leere Menge* und wird mit $\varnothing$ bezeichnet.

1.2. Grundlagen der Aussagenlogik

Eine Aussage ist entweder wahr oder falsch.

Zu jeder Aussage A gibt es eine Verneinung $\neg A$ mit der folgenden Eigenschaft: Ist A wahr, dann ist $\neg A$ falsch. Ist A falsch, dann ist $\neg A$ wahr.

Aus zwei Aussagen A und B können wir neue Aussagen wie folgt bilden:

(i) Die *Konjunktion* $A \wedge B$ („A und B“): A und B ist wahr, wenn A und B wahr sind, sonst falsch. Wie erhalten die folgende Wahrheitstafel:

$B\backslash A$	w	f
w	w	f
f	f	f

(ii) Die *Disjunktion* $A \vee B$ („A oder B“): A oder B ist falsch, falls A und B falsch sind, sonst wahr. Wir erhalten die folgende Wahrheitstafel:

$B\backslash A$	w	f
w	w	w
f	w	f

(iii) Die *Implikation* $A \Rightarrow B$ („aus A folgt B“): $A \Rightarrow B$ ist falsch, falls A wahr und B falsch ist, sonst wahr. Wir erhalten die folgende Wahrheitstafel:

$B\backslash A$	w	f
w	w	w
f	f	w

A heißt dann auch *Prämisse* oder *Voraussetzung* und B die *Konklusion* oder *Folgerung*.

Beispiel: Die Aussage „Aus $2 = 5$ folgt: 6 ist ungerade“ ist eine wahre Aussage.

(iv) Die *Äquivalenz* $A \Leftrightarrow B$ („A genau dann wenn B“): $A \Leftrightarrow B$ ist genau dann wahr, falls A und B beide wahr sind oder beide falsch sind. Wir erhalten die folgende Wahrheitstabelle:

$B \backslash A$	w	f
w	w	f
f	f	w

1.3. Einige Regeln

Für Aussagen A, B und C gilt:

(i) $\neg(\neg A) \Leftrightarrow A$, d. h. $\neg(\neg A)$ ist genau dann wahr, wenn A wahr ist.

(ii) $A \wedge B \Leftrightarrow B \wedge A$ und $A \vee B \Leftrightarrow B \vee A$.

(iii) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$ und $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$.

(iv) $(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$ und $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$.

(v) $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$ und $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$. („Regeln von de Morgan“)

Diese Regeln können mit Wahrheitstafeln überprüft werden.

1.4. Aussagen über Mengen mittels Quantoren

Es sei M eine Menge. $P(x)$ sei eine Eigenschaft, die von x abhänge. Mithilfe von Quantoren lassen sich neue Aussagen gewinnen:

(i) Der *Allquantor* $\forall$: „$\forall\, x \in M : P(x)$“ („Für alle x aus M gilt $P(x)$“): Die Aussage ist wahr, falls für jedes Element $x \in M$ die Eigenschaft $P(x)$ wahr ist.

(ii) Der *Existenzquantor* $\exists$: „$\exists\, x \in M : P(x)$“ („Es existiert ein $x \in M$, für das $P(x)$ wahr ist“): Die Aussage ist wahr, falls es *mindestens* ein $x \in M$ gibt, für das die Eigenschaft $P(x)$ wahr ist.

(iii) Die eindeutige Existenz $\exists!$: „$\exists!\, x \in M : P(x)$“ („Es gibt genau ein $x \in M$, sodass $P(x)$ wahr ist“): Die Aussage ist wahr, falls es genau ein $x \in M$ gibt, sodass $P(x)$ wahr ist.

1.5. Negation vertauscht Quantoren

Es gelten die folgenden Äquivalenzen:

$$\neg(\exists\, x \in M : P(x)) \Leftrightarrow \forall\, x \in M : \neg P(x),$$
$$\neg(\forall\, x \in M : P(x)) \Leftrightarrow \exists\, x \in M : \neg P(x).$$

2. Konstruktion in der Mengentheorie

Definition I.2.1 (Teilmenge): Eine Menge M_1 heißt *Teilmenge* einer Menge M_2, falls alle Elemente aus M_1 in M_2 liegen, wir schreiben dafür $M_1 \subseteq M_2$. Das heißt: $M_1 \subseteq M_2 :\Leftrightarrow \forall\, x \in M_1 : x \in M_2$.

Definition I.2.2 (Konstruktion neuer Mengen): Seien M_1, M_2 Mengen.

(i) $M_1 \cap M_2 := \{x \mid x \in M_1 \wedge x \in M_2\}$ heißt *Durchschnitt* von M_1 und M_2.

(ii) $M_1 \cup M_2 := \{x \mid x \in M_1 \vee x \in M_2\}$ heißt *Vereinigung* von M_1 und M_2.

(iii) $M_1 \setminus M_2 := \{x \mid x \in M_1 \wedge x \notin M_2\}$ heißt *Differenzmenge*. Wir schreiben auch $M_1 - M_2$.

(iv) $M_1 \times M_2 := \{(x, y) \mid x \in M_1 \text{ und } y \in M_2\}$ heißt *kartesisches Produkt* von M_1 und M_2.

(v) Für $k \in \mathbb{N}$ heißt

$$M_1^k := \{(x_1, x_2, \ldots, x_k) \mid x_1 \in M_1 \wedge \cdots \wedge x_k \in M_1\}$$

die *k-te kartesische Potenz* von M_1.

(vi) Die *Potenzmenge* $\mathfrak{P}(M_1)$ von M_1 ist die Menge aller Teilmengen von M_1, d. h. $\mathfrak{P}(M_1) := \{M \mid M \subseteq M_1\}$.

Beispiel I.2.3: Gegeben seien die vier Mengen $M_1 := \{1, 2\}$, $M_2 := \{1, 2, 3\}$, $M_3 := \varnothing$ und $M_4 := \{1, 7, a, b\}$. Dann können wir beobachten:

(i) $M_3 \subseteq M_1 \subseteq M_2$.

(ii) $M_2 \cap M_4 = \{1\}$, $M_1 \cap M_2 = \{1, 2\}$, $M_2 \cap M_3 = \varnothing$.

(iii) $M_2 \cup M_4 = \{1, 2, 3, 7, a, b\}$, $M_1 \cup M_2 = \{1, 2\}$, $M_2 \cup M_3 = M_2$.

(iv) $M_1 \times M_4 = \{(1, 1), (1, 7), (1, a), (1, b), (2, 1), (2, 7), (2, a), (2, b)\}$.

(v) $M_2 \setminus M_4 = \{2, 3\}$,

(vi) $\mathfrak{P}(M_2) = \{\{1, 2, 3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1\}, \{2\}, \{3\}, \varnothing\}$.

Notation I.2.4: Seien M_1, M_2, M_3 Mengen.

(i) $M_1 \supseteq M_2 :\Leftrightarrow M_2 \subseteq M_1$.

(ii) $x \notin M :\Leftrightarrow \neg(x \in M)$.

(iii) $M_1 \subsetneq M_2 :\Leftrightarrow M_1 \subseteq M_2 \wedge M_1 \neq M_2$. Wir sagen, M_1 sei *echt enthalten* in M_2.

(iv) Falls $M_1 \subseteq M_2$, dann heißt die Differenzmenge $M_2 \setminus M_1$ auch das *Komplement* von M_1 in M_2 und wird auch notiert als M_1^{c} oder $C_{M_2}(M_1)$.

Bemerkung I.2.5: Seien M, M_1, M_2 Mengen. Dann gelten:

(i) $M \subseteq M$.

(ii) Es gilt $M_1 = M_2$ genau dann, wenn $M_1 \subseteq M_2$ und $M_2 \subseteq M_1$.

Beweis: (i) Wir müssen zeigen, dass gilt: Für alle $x \in M$ gilt: $x \in M$. Das ist offensichtlich wahr.

(ii) Die Richtigkeit dieser Aussage folgt aus dem Extensionalitätsprinzip. □

Proposition I.2.6: *Es seien M_1, M_2, M_3 Mengen. Dann gelten:*

(i) $(M_1 \cup M_2) \cup M_3 = M_1 \cup (M_2 \cup M_3)$, $(M_1 \cap M_2) \cap M_3 = M_1 \cap (M_2 \cap M_3)$.

(ii) $M_1 \cap M_2 = M_2 \cap M_1$, $M_1 \cup M_2 = M_2 \cup M_1$.

(iii) $M_1 \cap (M_2 \cup M_3) = (M_1 \cap M_2) \cup (M_1 \cap M_3)$ *und* $M_1 \cup (M_2 \cap M_3) = (M_1 \cup M_2) \cap (M_1 \cup M_3)$.

Beweis: Wir wollen das Extensionalitätsprinzip verwenden, um die Aussagen zu zeigen. Exemplarisch für (i): Sei $x \in (M_1 \cap M_2) \cap M_3$. Per Definition gelten die folgenden Äquivalenzen:

$$\begin{aligned} x \in (M_1 \cap M_2) \cap M_3 &\Leftrightarrow x \in (M_1 \cap M_2) \wedge x \in M_3 \\ &\Leftrightarrow (x \in M_1 \wedge x \in M_2) \wedge x \in M_3 \\ &\Leftrightarrow x \in M_1 \wedge x \in M_2 \wedge x \in M_3 \qquad \text{(Abschnitt 1 ③)} \\ &\Leftrightarrow x \in M_1 \wedge (x \in M_2 \cap M_3) \\ &\Leftrightarrow x \in M_1 \cap (M_2 \cap M_3). \end{aligned}$$

Die anderen Aussagen lassen sich auf die gleiche Weise zeigen, man sagt „analog“. □

Proposition I.2.7: *Seien M, M_1, M_2 Mengen mit $M_1 \subseteq M$ und $M_2 \subseteq M$. Dann gilt:*

(i) $M \setminus (M \setminus M_1) = C_M(C_M(M_1)) = (M_1^{\mathsf{c}})^{\mathsf{c}} = M_1$,

(ii) $M \setminus M = \varnothing$,

(iii) $M \setminus \varnothing = M$,

(iv) $(M_1 \cup M_2)^{\mathsf{c}} = M_1^{\mathsf{c}} \cap M_2^{\mathsf{c}}$,

(v) $(M_1 \cap M_2)^{\mathsf{c}} = M_1^{\mathsf{c}} \cup M_2^{\mathsf{c}}$.

Beweis: (i) Es gelten die Äquivalenzen

$$\begin{aligned} x \in M \setminus (M \setminus M_1) &\Leftrightarrow x \in M \wedge x \notin M \setminus M_1 \\ &\Leftrightarrow x \in M \wedge \neg(x \in M \wedge x \notin M_1) \\ &\Leftrightarrow x \in M \wedge (x \notin M \vee x \in M_1) && \text{(Abschnitt 1 ③)} \\ &\Leftrightarrow (x \in M \wedge x \notin M) \vee (x \in M \wedge x \in M_1) \\ &\Leftrightarrow x \in M \wedge x \in M_1 \\ &\Leftrightarrow x \in M_1 && \text{(da } M_1 \subseteq M\text{).} \end{aligned}$$

Wir haben also gezeigt: $M \setminus (M \setminus M_1) = M_1$.

(ii) Wir haben

$$x \in M \setminus M \Leftrightarrow x \in M \wedge x \notin M.$$

Dies ist immer falsch, somit ist $M \setminus M = \varnothing$.

(iii) Unter Verwendung von (ii) können wir schreiben $M \setminus \varnothing = M \setminus (M \setminus M)$, mit (i) können wir jetzt ablesen, dass $M \setminus (M \setminus M) = M = M \setminus \varnothing$.

Aussagen (iv) und (v) sind Übungsaufgaben auf dem ersten Übungsblatt.□

Proposition I.2.8: *Es seien M_1, M_2, M_3 drei Mengen. Dann gilt:*

(i) $M_1 \subseteq M_1 \cup M_2$,

(ii) $M_1 \cap M_2 \subseteq M_1$,

(iii) *Ist $M_1 \subseteq M_2$ und $M_2 \subseteq M_3$, dann ist $M_1 \subseteq M_3$,*

(iv) $M_1 \cup \varnothing = M_1$,

(v) $M_1 \cap \varnothing = \varnothing$,

(vi) *Es gilt $M_1 \subseteq M_2$ genau dann, wenn $M_1 \cap M_2 = M_1$,*

(vii) *Es gilt $M_1 \subseteq M_2$ genau dann, wenn $M_1 \cup M_2 = M_2$.*

Beweis: Siehe Beispiel in Abschnitt 3 für die Beweisstrategie. Auf dem aktuellen Übungsblatt finden sich einige der Aussagen, vom Rest sollten Sie sich als eigene Übung überzeugen. □

3. Nützliche Beweisverfahren

In diesem Abschnitt seien A, B, M_1 und M_2 Mengen.

3.1. Nachweis von Teilmengenbeziehungen

Um Teilmengenbeziehungen nachzuweisen, verwenden wir folgende Strategie: Es gilt $A \subseteq B$ genau dann, wenn für alle $a \in A$ gilt: $a \in B$.

Beispiel: Zeige, dass $M_1 \cap M_2 \subseteq M_1$.

Beweis: Sei $x \in M_1 \cap M_2$. Es gilt $x \in M_1 \cap M_2$ genau dann, wenn $x \in M_1$ und $x \in M_2$, insbesondere gilt dann $x \in M_1$. Also folgt $M_1 \cap M_2 \subseteq M_1$. □

3.2. Gleichheit von Mengen

Um Gleichheit von Mengen nachzuweisen, verwenden wir, dass gilt: $A = B$ genau dann, wenn $A \subseteq B$ und $B \subseteq A$.

Beispiel: Zeige, dass gilt: Wenn $M_1 \subseteq M_2$, dann gilt $M_1 \cap M_2 = M_1$.

Beweis: Es gelte $M_1 \subseteq M_2$.
„$\subseteq$“: $M_1 \cap M_2 \subseteq M_1$ gilt nach Abschnitt 3.1.
„$\supseteq$“: Wenn $x \in M_1$, dann gilt $x \in M_2$, da $M_1 \subseteq M_2$, also $x \in M_1$ und $x \in M_2$, per Definition also $x \in M_1 \cap M_2$. Es gilt also $M_1 \subseteq M_1 \cap M_2$.
Aus „$\subseteq$“ und „$\supseteq$“ folgt jetzt $M_1 \cap M_2 = M_1$. □

3.3. Äquivalenz von Aussagen

Seien A und B Aussagen. Um die Äquivalenz von A und B zu zeigen, wollen wir verwenden, dass gilt: $(A \Leftrightarrow B)$ gilt genau dann, wenn $(A \Rightarrow B)$ und $(B \Rightarrow A)$.

Beispiel: Zeige, dass gilt: $M_1 \subseteq M_2$ genau dann, wenn $M_1 \cap M_2 = M_1$.

Beweis: „$\Rightarrow$“: Wenn $M_1 \subseteq M_2$, dann folgt $M_1 \cap M_2 = M_1$ aus Abschnitt 3.2 .
„$\Leftarrow$“: Für M_1, M_2 gelte $M_1 \cap M_2 = M_1$ und es sei $x \in M_1$. Nach Voraussetzung gilt $x \in M_1 \cap M_2$, also $x \in M_1$ und $x \in M_2$. Insbesondere $x \in M_2$, also $M_1 \subseteq M_2$. □

Insgesamt haben wir Proposition I.2.8 (ii) und Proposition I.2.8 (vi) bewiesen.

3.4. Widerspruchsbeweis

Wir wollen verwenden, dass gilt: $(A \Rightarrow B)$ genau dann, wenn $(\neg B \Rightarrow \neg A)$. Beispiele dazu werden Sie in den Präsenzübungen kennen lernen.

3.5. Beweis durch vollständige Induktion

Seien $A(n)$ eine Aussage, die von $n \in \mathbb{N}$ abhängt, $S := \{n \in \mathbb{N} \mid A(n) \text{ ist wahr}\}$ und $s_0 \in S$. Falls weiterhin gilt „Wenn $n \in S$, dann ist $n+1 \in S$“, dann ist $A(n)$ wahr für alle $n \geq s_0$.

Beispiel: Für $n \in \mathbb{N}$ sei $T(n) := 1 + 2 + \cdots + n$. Zeige: $T(n) = \frac{1}{2}n(n+1)$ für alle $n \in \mathbb{N}$.

Beweis: Sei $S = \{n \in \mathbb{N} \mid T(n) = \frac{1}{2}n(n+1)\}$. Da $T(1) = 1 = \frac{1}{2}1 \cdot 2$ gilt $1 \in S$. Sei nun weiterhin n ein Element von S, d. h. $T(n) = \frac{1}{2}n(n+1)$. Wir wollen zeigen, dass dann auch $n+1 \in S$, d. h. $T(n+1) = \frac{1}{2}(n+1)(n+2)$. Es ist

$$\begin{aligned} T(n+1) = 1 + 2 + \cdots + n + (n+1) &= T(n) + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) \qquad (n \in S) \\ &= \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Die Aussage gilt also für alle $n \geq 1$ in $\mathbb{N}$. □

4. Abbildungen

In diesem Abschnitt seien W, X, Y, Z Mengen.

Notation: Ist $M = \{a_1, \ldots, a_n\}$ eine Menge, die aus n Elementen besteht, so heißt n die *Anzahl der Elemente von* M und wird mit $\#M$ oder $|M|$ bezeichnet.

Definition I.4.1: (i) Eine *Abbildung* oder *Funktion* $f\colon X \to Y$ ordnet jedem $x \in X$ genau ein $y \in Y$ zu. Wir schreiben dafür

$$f\colon X \longrightarrow Y, \qquad x \longmapsto y = f(x).$$

X heißt *Definitionsbereich*, Y heißt *Wertebereich*.

(ii) $\mathrm{Abb}(X, Y)$ bezeichnet die Menge aller Abbildungen von X nach Y, d. h.

$$\mathrm{Abb}(X, Y) := \{f \mid f\colon X \to Y \text{ Abbildung}\}.$$

Bemerkung I.4.2: (i) Genauer definiert man Abbildungen mengentheoretisch wie folgt: Eine Abbildung $f\colon X \to Y$ ist gegeben durch eine Teilmenge Γ_f von $X \times Y$ mit folgender Eigenschaft: $\forall\, x \in X\, \exists!\, y \in Y : (x, y) \in \Gamma_f$. Wir schreiben $x \mapsto f(x) = y$ genau dann, wenn $(x, y) \in \Gamma_f$. Γ_f heißt auch (mengentheoretischer) Graph von f.

(ii) Zwei Abbildungen $f\colon X \to Y$ und $g\colon X \to Y$ sind gleich genau dann, wenn für alle $x \in X$ gilt, dass $f(x) = g(x)$.

(iii) $\mathrm{Abb}(\varnothing, Y)$ enthält genau ein Element, dessen Graph $\Gamma_f = \varnothing$ ist. $\varnothing$ wird hier als Teilmenge von $\varnothing \times Y = \varnothing$ aufgefasst.

Beispiel I.4.3: (i) Für $X_1 := \{-1, 0, 1\}$ und $Y_1 := \{0, 1\}$ betrachte die beiden Abbildungen

$$f_1\colon X_1 \longrightarrow Y_1, \quad x \longmapsto x^3 - x, \qquad g_1\colon X_1 \longrightarrow Y_1, \quad x \longmapsto 0.$$

Dann gilt $f_1 = g_1$.

(ii) Für $X_2 := \mathbb{R}$ und $Y_2 := \mathbb{R}_{\geq 0}$[1] ist $f_2\colon X_2 \to Y_2$, $x \mapsto x^2$ ein Beispiel für eine Abbildung.

(iii) Für $X_3 = \mathbb{R}_{\geq 0}$ und $Y_3 := \mathbb{R}$ ist $f_3\colon X_3 \to Y_3$, $x \mapsto \sqrt{x}$ eine Abbildung.

(iv) Betrachte die beiden Mengen $X_4 := \{s \mid s \text{ ist Student in diesem Hörsaal}\}$, $Y_4 := \{t \mid t \text{ ist Datum eines Tages im Jahr}\}$.

$$f_4\colon X_4 \longrightarrow Y_4, \qquad s \longmapsto (\text{Geburtsdatum von } s)$$

ist eine Abbildung zwischen X_4 und Y_4.

Definition I.4.4 (Identität): Für eine beliebige Menge M heißt die Abbildung

$$\mathrm{id}_M\colon M \longrightarrow M, \qquad x \longmapsto x$$

die *Identität auf* M.

Definition I.4.5 (Urbild und Bild): Sei $f\colon X \to Y$ eine Abbildung.

(i) Für $B \subseteq Y$ heißt $f^{-1}(B) := \{x \in X \mid f(x) \in B\}$ das *Urbild* von B unter f.

(ii) Für $A \subseteq X$ heißt $f(A) := \{f(x) \mid x \in A\}$ das *Bild* von A unter f.

Beispiel I.4.6: Für die Abbildungen aus Proposition I.4.3 gilt

[1] Es ist $R_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$.

(i) Es sind $f_1^{-1}(\{0\}) = \{-1, 0, 1\}$, $f_1^{-1}(\{1\}) = \varnothing$, $f_1^{-1}(\{0, 1\}) = \{-1, 0, 1\}$, $f_1^{-1}(\varnothing) = \varnothing$ und $f_1(\{1\}) = \{0\} = f_1(\{0\}) = f_1(\{-1\})$.

(ii) $f_2^{-1}(\{y \in \mathbb{R} \mid y \geq 1\}) = \{x \in \mathbb{R} \mid x \leq -1 \text{ oder } x \geq 1\}$.

(iii) $f_3(\mathbb{R}_{\geq 0}) = \mathbb{R}_{\geq 0}$.

Definition I.4.7 (Verkettung und Einschränkung): Es sei $A \subseteq X$ eine Teilmenge.

(i) Für Abbildungen $f\colon X \to Y$ und $g\colon Y \to Z$ definieren wir die Abbildung

$$(g \circ f)\colon X \longrightarrow Z, \qquad x \longmapsto (g \circ f)(x) := g(f(x)).$$

$g \circ f$ heißt *Verkettung*, oder auch *Komposition* von f und g.

(ii) Ist $f\colon X \to Y$ eine Abbildung, dann definieren wir die Abbildung $f|_A\colon A \to Y$ gegeben durch $x \mapsto f(x)$. Diese heißt *Einschränkung* von f auf A.

Bemerkung I.4.8 (Kategorielle Eigenschaft): (i) Verkettung ist *assoziativ*, das heißt für Abbildungen $f\colon W \to X$, $g\colon X \to Y$ und $h\colon Y \to Z$ gilt:

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Wir notieren deshalb ab jetzt auch $h \circ g \circ f$ für $h \circ (g \circ f) = (h \circ g) \circ f$.

(ii) Die Identität „tut“ beim Verketten nichts, d. h. für eine Abbildung $f\colon X \to Y$ gilt: $\mathrm{id}_Y \circ f = f = f \circ \mathrm{id}_X$.

Beweis: Wir wollen die Gleichheit der beiden Abbildungen zeigen, indem wir nachweisen, dass beide Abbildungen dieselbe Wirkung auf allen Elementen des Definitionsbereichs haben.

(i) Für jedes $w \in W$ gilt:

$$\begin{aligned}(h \circ (g \circ f))(w) = h((g \circ f)(w)) &= h(g(f(w))) \\ &= (h \circ g)(f(w)) = ((h \circ g) \circ f)(w),\end{aligned}$$

also $h \circ (g \circ f) = (h \circ g) \circ f$.

(ii) Für alle $x \in X$ gilt $(\mathrm{id}_Y \circ f)(x) = \mathrm{id}_Y(f(x)) = f(x)$, also $\mathrm{id}_Y \circ f = f$. Analog erhält man, dass $f = f \circ \mathrm{id}_X$. □

Beispiel I.4.9: Die Verknüpfung der beiden Abbildungen f_2, f_3 aus Proposition I.4.3 ist $(f_3 \circ f_2) = \sqrt{x^2} = |x|$.

Definition I.4.10 (Injektiv, Surjektiv, Bijektiv): Eine Abbildung $f\colon X \to Y$ heißt

(i) *injektiv*, wenn gilt: Für alle $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$ ist schon $x_1 = x_2$. Das ist genau dann der Fall, wenn gilt: Für alle $y \in Y$ ist $\#f^{-1}(\{y\}) \leq 1$.

(ii) *surjektiv*, wenn gilt: Für alle $y \in Y$ gibt es $x \in X$, sodass $f(x) = y$. Das ist genau dann der Fall, wenn gilt: Für alle $y \in Y$ ist $\#f^{-1}(\{y\}) \geq 1$.

(iii) *bijektiv*, wenn gilt: Für alle $y \in Y$ gibt es genau ein $x \in X$, sodass $f(x) = y$. Das ist genau dann der Fall, wenn gilt: Für alle $y \in Y$ ist $\#f^{-1}(\{y\}) = 1$.

f ist bijektiv genau dann, wenn f injektiv und surjektiv ist.

Definition I.4.11 (Umkehrabbildung): Ist $f\colon X \to Y$ eine Abbildung, so heißt $g\colon Y \to X$ *Umkehrabbildung von* f, falls gelten:

(i) Für alle $x \in X$ gilt $g(f(x)) = x$, d. h. $g \circ f = \mathrm{id}_X$,

(ii) Für alle $y \in Y$ gilt $f(g(y)) = y$, d. h. $f \circ g = \mathrm{id}_Y$.

Proposition I.4.12 (Eigenschaften bijektiver Abbildungen): *Sei $f\colon X \to Y$ eine Abbildung.*

(i) *f hat eine Umkehrabbildung genau dann, wenn f bijektiv ist.*

(ii) *Falls f eine Umkehrabbildung hat, so ist diese eindeutig durch f bestimmt und ebenfalls bijektiv. Wir notieren die Umkehrabbildung dann mit f^{-1}.*

(iii) *Sind $f\colon X \to Y$, $g\colon Y \to Z$ bijektive Abbildungen, dann ist auch $g \circ f$ bijektiv und es gilt $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Beweis: (i) Siehe Übungsblatt 2.

(ii) Wir wollen die Aussage in mehreren Schritten zeigen.

Eindeutigkeit der Umkehrabbildung: Seien $g_1\colon Y \to X$ und $g_2\colon Y \to X$ zwei Umkehrabbildungen von f, d. h. $g_1 \circ f = \mathrm{id}_X = g_2 \circ f$ und $f \circ g_1 = \mathrm{id}_Y = f \circ g_2$. Dann gilt für alle $y \in Y$:

$$g_1(y) = (g_1 \circ \mathrm{id}_Y)(y) = (g_1 \circ (f \circ g_2))(y) = ((g_1 \circ f) \circ g_2)(y) = (\mathrm{id}_X \circ g_2)(y) = g_2(y).$$

Hierbei haben wir zunächst Proposition I.4.8(ii), dann die Definition der Umkehrabbildung, dann Proposition I.4.8(i), erneut die Definition der Umkehrabbildung und schließlich Proposition I.4.8(ii) verwendet. Insgesamt folgt $g_1 = g_2$.

Bijektivität der Umkehrabbildung: Sei g die Umkehrabbildung von f, d. h. es gelten $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$. f ist also die Umkehrabbildung von g und nach (i) ist g bijektiv.

(iii) Sind $f\colon X \to Y$ und $g\colon Y \to Z$ bijektive Abbildungen und $f^{-1}\colon Y \to X$ und $g^{-1}\colon Z \to Y$ die nach (i) existierenden, bijektiven Umkehrabbildungen, die nach (ii) sogar eindeutig sind. Dann gilt:

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ f \circ f^{-1} \circ g^{-1} = g \circ \mathrm{id}_Y \circ g^{-1} = g \circ g^{-1} = \mathrm{id}_Z,$$

genauso $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \mathrm{id}_X$, d. h. $f^{-1} \circ g^{-1}$ ist eine Umkehrabbildung von $g \circ f$. Aussage (i) garantiert die Bijektivität von $g \circ f$ und (ii) gibt, dass $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

Notation I.4.13: Hat $f\colon X \to Y$ die Umkehrabbildung $f^{-1}\colon Y \to X$, dann sagt man auch, f und f^{-1} sind *invers* zueinander und f^{-1} heißt auch *inverse Abbildung zu* f.

Bemerkung I.4.14: Sei $f\colon X \to Y$ eine Abbildung und $b \in Y$. Ist f bijektiv, dann hat f^{-1} zwei Bedeutungen:

(i) $f^{-1}(\{b\})$ ist das Urbild von $\{b\}$.

(ii) $f^{-1}(b)$ ist das Bild von b unter der Umkehrabbildung f^{-1}. In diesem Fall gilt dann aber $f^{-1}(\{b\}) = \{f^{-1}(b)\}$.

Beispiel I.4.15 (X^k vs. $\mathrm{Abb}(\{1,\dots,k\}, X)$): Es sei $k \in \mathbb{N}$. Die k-te kartesische Potenz $X^k = X \times \cdots \times X$ kann auf folgende Weise mit $\mathrm{Abb}(\{1,\dots,k\}, X)$ identifiziert werden:
① Definiere die Abbildung

$$F\colon X^k \longrightarrow \mathrm{Abb}(\{1,\dots,k\}, X), \quad (a_1,\dots,a_k) \longmapsto (f\colon \{1,\dots,k\} \to X, i \mapsto a_i).$$

② Definiere die Abbildung

$$G\colon \mathrm{Abb}(\{1,\dots,k\}, X) \longrightarrow X^k = X \times \cdots \times X, \quad f \longmapsto (f(1),\dots,f(k)).$$

G ist die Umkehrabbildung zu F, nach Proposition I.4.12 sind damit F und G bijektiv.

Definition I.4.16: Definiere wegen Proposition I.4.15: $X^0 := \mathrm{Abb}(\varnothing, X)$. Also besteht nach Proposition I.4.2 X^0 aus einem Punkt.

Definition I.4.17: $\mathrm{Perm}(X) := \{f\colon X \to X \mid f \text{ ist bijektiv}\}$.

Falls $X = \{a_1,\dots,a_n\}$ eine Menge mit n Elementen ist, dann ist $n!$ die Anzahl der Elemente von $\mathrm{Perm}(X)$.

5. Relationen

Sei M in diesem Abschnitt stets eine Menge.

Definition I.5.1 (Relation): Eine *(zweistellige) Relation* auf M ist eine Teilmenge $R \subseteq M \times M = M^2$. Statt $(x, y) \in R$ schreibt man auch xRy oder $x \sim_R y$.

Beispiel I.5.2: (i) $R_1 := \{(x, y) \in M \times M \mid x = y\}$ heißt die *Gleichheitsrelation* auf M. Es gilt also xR_1y genau dann, wenn $x = y$.

(ii) Beispiele für Relationen auf $M = \mathbb{R}$:

$$R_2 := \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \leq y\}, \quad R_3 := \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x < y\},$$
$$R_4 := \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \geq y\}, \quad R_5 := \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x > y\},$$
$$R_6 := \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \neq y\}.$$

Definition I.5.3 (Eigenschaften von Relationen): Es sei $R \subseteq M \times M$ eine Relation. Dann heißt R

(i) *reflexiv*, falls für alle $x \in M$ gilt: xRx,

(ii) *symmetrisch*, falls für alle $x, y \in M$ gilt: Wenn xRy, dann yRx,

(iii) *antisymmetrisch*, falls für alle $x, y \in M$ gilt: Wenn xRy und yRx, dann $x = y$,

(iv) *transitiv*, falls für alle $x, y, z \in M$ gilt: Wenn xRy und yRz, dann xRz.

Beispiel I.5.4: Für die Relationen in Proposition I.5.2 gilt:

	R_1 („=“)	R_2 („$\leq$“)	R_3 („$<$“)	R_6 („$\neq$“)
reflexiv	✓	✓	–	–
symmetrisch	✓	–	–	✓
antisymmetrisch	✓	✓	✓	–
transitiv	✓	✓	✓	–

Definition I.5.5 (Äquivalenz– und Ordnungsrelation): Eine Relation R auf M heißt

(i) *Äquivalenzrelation*, falls R reflexiv, symmetrisch und transitiv ist,

(ii) *(partielle) Ordnungsrelation*, falls R reflexiv, antisymmetrisch und transitiv ist.

Beispiel I.5.6: In Proposition I.5.2 ist R_1 eine Äquivalenzrelation und R_1, R_2 sowie R_4 sind Ordnungsrelationen.

Beispiel I.5.7 (Kongruenzrelation): Seien $M := \mathbb{Z}$ und $n \in \mathbb{N}$ gegeben. Definiere die Relation

$$R := \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a - b \text{ ist durch } n \text{ teilbar}\}.$$

Wir schreiben für aRb auch $a \equiv b \pmod{n}$ oder auch $a \equiv_n b$. Zum Beispiel sind $2 \equiv_5 12$, $2 \equiv_3 5$ oder $2 \equiv_4 -2$.

Bemerkung I.5.8: „$\equiv_n$“ ist eine Äquivalenzrelation.

Beweis: ① *Reflexivität*: Für alle $a \in \mathbb{Z}$ gilt $n \mid a - a = 0$, also $a \equiv_n a$.

② *Symmetrie*: Seien $a, b \in \mathbb{Z}$, dann gilt $a \equiv_n b$ genau dann, wenn $n \mid a - b$. Wegen $a - b = -(b - a)$ gilt in diesem Fall auch $b \equiv_n a$.

③ *Transitivität*: Seien $a, b, c \in \mathbb{Z}$, dann gilt: Wenn $a \equiv_n b$ und $b \equiv_n c$, dann gibt es $k, l \in \mathbb{Z}$, sodass $a - b = k \cdot n$ und $b - c = l \cdot n$, also

$$a - c = (a - b) + (b - c) = k \cdot n + l \cdot n = (k + l)n,$$

und damit $a \equiv_n c$. □

Bemerkung I.5.9: Für $n \in \mathbb{N}$ erhält man eine Zerlegung von $\mathbb{Z}$ in n Mengen

$$M_1 := \{a \in \mathbb{Z} \mid a \equiv_n 1\},\ M_2 := \{a \in \mathbb{Z} \mid a \equiv_n 2\}, \ldots,\ M_n = \{a \in \mathbb{Z} \mid a \equiv_n 0\}.$$

Dies sind genau die Restklassen modulo n.

Beispiel I.5.10: Für $n = 2$ sind die beiden Mengen M_1 und M_2 aus Proposition I.5.9 die Menge der ungeraden bzw. der geraden Zahlen.

Definition I.5.11: Sei „$\sim$“ eine Äquivalenzrelation auf M. Dann heißt für ein Element x von M die Teilmenge

$$[x]_\sim := \{y \in M \mid x \sim y\} \subseteq M$$

die *Äquivalenzklasse von* x bezüglich „$\sim$“.

Beispiel I.5.12: Seien $M = \mathbb{Z}$, „$\sim$“ = „$\equiv_n$“, und $x = 1$. Dann ist

$$[1]_\sim := \{y \in \mathbb{Z} \mid 1 \equiv_n y\} = \{1 + kn \mid k \in \mathbb{Z}\}.$$

Satz 1 (Zerlegung in Äquivalenzklassen): *Sei „$\sim$“ eine Äquivalenzrelation auf M. Dann gilt:*

(i) *Alle Äquivalenzklassen sind nicht leer, d. h. für alle $x \in M$ gilt $[x]_\sim \neq \varnothing$.*

(ii) *Stehen x und y in Relation, dann gilt $[x]_\sim = [y]_\sim$.*

(iii) *M ist die Vereinigung aller Äquivalenzklassen, d. h. jedes Element $x \in M$ ist in einer Äquivalenzklasse enthalten.*

(iv) *Je zwei verschiedene Äquivalenzklassen sind disjunkt, d. h. für alle x, y in M gilt: $[x]_\sim = [y]_\sim$ oder $[x]_\sim \cap [y]_\sim = \varnothing$.*

Beweis: (i) Für alle $x \in M$ gilt $x \sim x$, da „$\sim$“ reflexiv ist. Damit ist $x \in [x]_\sim$, also $[x]_\sim \neq \varnothing$.

(ii) Für jedes x' in $[x]_\sim$ gilt $x \sim x'$. Per Symmetrie von „$\sim$“ ist dann auch $x' \sim x$. Aus $x' \sim x$ und $x \sim y$ (per Voraussetzung) folgt wegen der Transitivität von „$\sim$“. dass $x' \sim y$. Wiederrum wegen Symmetrie gilt dann auch $y \sim x'$. Das heißt es gilt $x' \in [y]_\sim$ und damit $[x]_\sim \subseteq [y]_\sim$. Analog erhält man $[y]_\sim \subseteq [x]_\sim$, weshalb insgesamt $[x]_\sim = [y]_\sim$.

(iii) Dies folgt wiederum aus $x \in [x]_\sim$.

(iv) Seien $x, y \in M$ mit $[x]_\sim \cap [y]_\sim \neq \varnothing$ und sei $z \in [x]_\sim \cap [y]_\sim$. Aus (ii) folgt $[x]_\sim = [z]_\sim$ und $[z]_\sim = [y]_\sim$, weshalb $[x]_\sim = [y]_\sim$. □

Definition I.5.13: (i) Für eine Äquivalenzrelation „$\sim$“ auf M bezeichnen wir mit $M_\sim := \{[x]_\sim \mid x \in M\}$ die *Menge der Äquivalenzklassen* von M bezüglich „$\sim$“.

(ii) Die Abbildung $\pi\colon M \to M_\sim, x \mapsto [x]_\sim$ heißt *kanonische Projektion*

Definition I.5.14: Sei I eine Menge. Für jedes $i \in I$ sei eine Menge M_i gegeben. Dann definieren wir:

$$\bigcap_{i \in I} M_i := \{x \mid \forall\, i \in I : x \in M_i\}, \quad \bigcup_{i \in I} M_i := \{x \mid \exists\, i \in I : x \in M_i\}.$$

Definition I.5.15: Eine Teilmenge $P \subseteq \mathfrak{P}(M)$ heißt *Partition*, falls gelten:

(i) $\varnothing \notin P$,

(ii) $\bigcup_{A \in P} A = M$,

(iii) Für alle $A, B \in P$ gilt: Ist $A \neq B$, dann ist $A \cap B = \varnothing$.

Bei (ii) ist $I = P$ und $M_i = i$.

Korollar I.5.16 (aus Satz 1): *Ist M eine Menge und „$\sim$“ eine Äquivalenzrelation auf M, $M_\sim$ die Menge der Äquivalenzklassen, dann ist $M_\sim$ eine Partition.*

Bemerkung: Eine Umkehrung dieser Aussage zeigen Sie auf Übungsblatt 3.

6. Nachtrag und Ausblick

Für eine formale Einführung zur Mengenlehre siehe zum Beispiel die Lehrbücher von Deiser und Ebbinghaus.

Die Mengentheorie baut auf Axiomen auf, d. h. es werden Regeln definiert, die für Mengen gelten sollen. Es gibt unterschiedliche Axiomensysteme, wir verwenden das ZFC-Axiomensystem[2]. Zu den ZFC-Axiomen gehören zum Beispiel das *Extensionalitäts-Axiom*, das *Aussonderungsaxiom*, das *Leermengenaxiom*, das die Existenz der leeren Menge sichert, und das *Auswahlaxiom*. Überraschender Weise folgt das Auswahlaxiom nicht aus den restlichen Axiomen. Es ist spannend, sich klar zu machen, für welche Aussagen man das Auswahlaxiom braucht.

Nette Literatur in diesem Kontext ist „Logicomix: eine epische Suche nach Wahrheit“.

[2]Das „Z“ steht für den deutschen Mathematiker *Ernst Zermelo* (1871-1953), „F“ steht für den deutsch-israelischen Mathematiker *Adolf Abraham Haleri Fraenkel* (1891-1965) und „C“ steht für *choice* – das Auswahlaxiom.

Kapitel II.

Vektorräume und lineare Gleichungssysteme

Aus typographischen Gründen verwenden wir im Mitschrieb die Schreibweise

$$(x_1, \ldots, x_n)^t := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

für Spaltenvektoren (meistens im Fließtext). Diese Schreibweise ist sinnvoll, was sich später herausstellen wird. Ab jetzt wollen wir schreiben

$$\mathbb{R}^n := \{(x_1, \ldots, x_n)^t \mid x_1, \ldots, x_n \in \mathbb{R}^n\}.$$

1. Motivation

Beispiel II.1.1: Betrachte das folgende lineare Gleichungssystem

$$\begin{aligned} 2x_1 + 6x_2 + 4x_3 &= 8, \\ x_2 - 2x_3 &= 6, \\ x_1 + 4x_2 &= 10. \end{aligned} \tag{II.1}$$

Gesucht ist die Lösungsmenge $\mathbb{L} := \{x = (x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x \text{ erfüllt Gl. (II.1)}\}$. Wichtige Fragen sind:

(1) Hat Gl. (II.1) eine Lösung?

(2) Wenn Gl. (II.1) eine Lösung hat, wie viele gibt es? 1,2,3? Unendlich viele?

(3) Wenn es unendlich viele Lösungen gibt, wie können diese angegeben werden? Welche Struktur hat $\mathbb{L}$? Wie „groß“ ist $\mathbb{L}$?

(4) Gibt es ein allgemeines Lösungsverfahren für lineare Gleichungssysteme zur Bestimmung von $\mathbb{L}$?

Idee II.1.2: (1) *Matrix-Schreibweise*: Das lineare Gleichungssystem Gl. (II.1) ist durch folgende Daten bestimmt:

$$A = \begin{pmatrix} 2 & 6 & 4 \\ 0 & 1 & -2 \\ 1 & 4 & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 8 \\ 6 \\ 10 \end{pmatrix},$$

wir schreiben Gl. (II.1) auch als erweiterte (Koeffizienten-)Matrix $(A \mid b)$.

(2) *Reduktion auf homogenes lineares Gleichungssystem*: Ist $x' = (x'_1, x'_2, x'_3)^t$ eine Lösung von Gl. (II.1), dann gilt für jede weitere Lösung $x = (x_1, x_2, x_3)$ von Gl. (II.1):

$$2(x'_1-x_1)+6(x'_2-x_2)+4(x'_3-x_3) = 2x'_1+6x'_2+4x'_3-(2x_1+6x_2+4x_3) = 8-8 = 0,$$

analog für die anderen beiden Gleichungen in Gl. (II.1), also folgt für den Vektor $y := (x_1 - x'_1, x_2 - x'_2, x_3 - x'_3)^t$, dass y eine Lösung von

$$\begin{aligned} 2x_1 + 6x_2 + 4x_3 &= 0, \\ x_2 - 2x_3 &= 0, \\ x_1 + 4x_2 \qquad &= 0 \end{aligned} \tag{II.2}$$

ist. Es gilt also für jede Lösung x von Gl. (II.1): $x = x' + y$, wobei y eine Lösung von Gl. (II.2) ist. Wir bezeichnen Gl. (II.2) als das homogene Gleichungssystem zu Gl. (II.1). Es ist also sinnvoll, zunächst die Lösungsmenge von homogenen linearen Gleichungssystemen zu studieren.

(3) *Struktur der Lösungsmenge von homogenen linearen Gleichungssystemen*: Betrachte zwei Lösungen $x = (x_1, x_2, x_3)^t$ und $y = (y_1, y_2, y_3)^t$ von Gl. (II.2). Dann gilt für ihre Summe $v = (v_1, v_2, v_3)^t = (x_1 + y_1, x_2 + y_2, x_3 + y_3)^t$ und für jedes Vielfache $w = (w_1, w_2, w_3)^t = (\lambda x_1, \lambda x_2, \lambda x_3)^t$ mit $\lambda \in \mathbb{R}$:

$$\begin{aligned} 2v_1 + 6v_2 + 4v_3 &= 2(x_1 + y_1) + 6(x_2 + y_2) + 4(x_3 + y_3) \\ &= 2x_1 + 6x_2 + 4x_3 + 2y_1 + 6y_2 + 4y_3 = 0 + 0 = 0 \end{aligned}$$

und

$$2w_1 + 6w_2 + 4w_3 = 2(\lambda x_1) + 6(\lambda x_2) + 4(\lambda x_3) = \lambda(2x_1 + 6x_2 + 4x_3) = \lambda 0 = 0;$$

analog für die anderen beiden Gleichungen aus Gl. (II.2), also sind v und w ebenfalls Lösungen von Gl. (II.2). Es gilt also für die Lösungemenge $\mathbb{L}_h$ von Gl. (II.2):

$$x, y \in \mathbb{L}_h \Rightarrow x + y \in \mathbb{L}_h, \qquad \lambda \in \mathbb{R}, x \in \mathbb{L}_h \Rightarrow \lambda x \in \mathbb{L}_h.$$

2. Vektorräume

In diesem Abschnitt wollen wir den $\mathbb{R}^n$ verallgemeinern zu beliebigen Vektorräumen. Die Struktur des $\mathbb{R}^n$ wird bestimmt durch die *Vektoraddition*, die *Skalarmultiplikation* und den Rechenregeln, die für diese gelten.

Erinnerung II.2.1: Seien $x = (x_1, \ldots, x_n)^t, y = (y_1, \ldots, y_n)^t \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Dann sind

$$x + y := (x_1 + y_1, \ldots, x_n + y_n)^t, \qquad \lambda x := (\lambda x_1, \ldots, \lambda x_n)^t.$$

Rechenregeln im $\mathbb{R}^n$ sind zum Beispiel: Für alle $x, y \in \mathbb{R}^n$ gelten:

(i) $x + y = y + x$, (Kommutativität von „+“)

(ii) $(x + y) + z = x + (y + z)$. (Assoziativität von „+“)

Definition II.2.2 ($\mathbb{R}$-Vektorraum): Ein $\mathbb{R}$-Vektorraum ist eine Menge V zusammen mit einem ausgezeichneten Element $\mathbf{0} = \mathbf{0}_V$ und zwei Abbildungen (genannt *Verknüpfungen*)

$$\begin{aligned} +\colon V \times V &\longrightarrow V, \qquad (v, w) \longmapsto +(v, w) =: v + w, \\ \cdot\colon \mathbb{R} \times V &\longrightarrow V, \qquad (\lambda, v) \longmapsto \cdot(\lambda, w) =: \lambda \cdot v. \end{aligned}$$

die folgende Regeln („Vektorraumaxiome“) erfüllen: Für die Vektoraddition gilt

(A1) Für alle $x, y, z \in V$ ist $(x + y) + z = x + (y + z)$, („Assoziativgesetz“)

(A2) Für alle $x, y \in V$ ist $x + y = y + x$, („Kommutativgesetz“)

(A3) $\mathbf{0}$ ist ein *Nullvektor*, d. h. für alle $x \in V$ ist $x + \mathbf{0} = x = \mathbf{0} + x$,

(A4) Für alle $x \in V$ gibt es genau ein $y \in V$, sodass $x + y = \mathbf{0} = y + x$.
Für das eindeutige Element y schreibt man auch $-x$ und nennt es *Inverses von* x.

Zusätzlich gilt für die Skalarmultiplikation:

(M1) Für alle $x \in V$ ist $1x = x$, („Einsgesetz“)

(M2) Für alle $\lambda_1, \lambda_2 \in \mathbb{R}$ und $x \in V$ ist $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$,

(M3) Für alle $\lambda \in \mathbb{R}$, $x, y \in V$ ist $\lambda(x + y) = \lambda x + \lambda y$,

(M4) Für alle $\lambda_1, \lambda_2 \in \mathbb{R}$, $x \in V$ ist $\lambda_1(\lambda_2 x) = (\lambda_1 \lambda_2)x$.

Die Gesetze $(\mathrm{VA}_2) - (\mathrm{VA}_4)$ werden *Distributivgesetze* genannt. Wir schreiben den $\mathbb{R}$-Vektorraum V auch als 4-Tupel $(V, +, \cdot, \mathbf{0}_V)$.

Beispiel II.2.3: (i) $\mathbb{R}^n$ mit $\mathbf{0} = (0, \ldots, 0)^t$ und skalarer Multiplikation wie in Proposition II.2.1 ist ein $\mathbb{R}$-Vektorraum.

(ii) Ist M eine Menge, so ist $V := \mathbb{R}^M := \mathrm{Abb}(M, \mathbb{R})$ ein $\mathbb{R}$-Vektorraum mit den Verknüpfungen

$$+ \colon V \times V \longrightarrow V, \qquad (f_1, f_2) \longmapsto f_1 + f_2 := g$$

wobei gilt: Für alle $m \in M$ ist $g(m) = f_1(m) + f_2(m)$; und

$$\cdot \colon \mathbb{R} \times V \longrightarrow V, \qquad (\lambda, f) \longmapsto h$$

wobei gilt: Für alle $m \in M$ ist $h(m) = \lambda f(m)$; und $\mathbf{0}_V = f$, wobei gilt: Für alle $m \in M$ ist $f(m) = 0$.

Bemerkung: Für $M = \{1, \ldots, n\}$ liefert (ii) in Proposition II.2.3 genau den $\mathbb{R}^n$ aus (i).

(iii) Ist M eine Menge und W ein $\mathbb{R}$-Vektorraum, so ist $V := \mathrm{Abb}(M, W)$ mit $\mathbf{0} := f \colon M \to W$, $m \mapsto \mathbf{0}_W$ und den Verknüpfungen wie in (ii) ein $\mathbb{R}$-Vektorraum.

Beweis: Die Vektorraumaxiome aus Proposition II.2.2 gelten jeweils, weil die entsprechende Regel in $\mathbb{R}$ bzw. für (iii) in W gilt.

Wir zeigen exemplarisch das Nullgesetz für (iii), d. h. für alle $f \in \mathrm{Abb}(M, W)$ soll $f + \mathbf{0} = f = \mathbf{0} + f$. Diese Gleichheit von Abbildungen zeigt man wie in Proposition I.4.2. Es gilt

$$(f + \mathbf{0})(m) = f(m) + \mathbf{0}(m) = f(m) + \mathbf{0}_V = f(m),$$

d. h. $f + \mathbf{0}_V = f$. Analog erhält man $\mathbf{0}_V + f = f$, also ist die Regel (VA_3) erfüllt.

Die restlichen Beweis gehen auf die gleiche Weise und bleiben als Übung überlassen. □

Proposition II.2.4: *Ist $(V, +, \cdot, \mathbf{0}_V)$ ein $\mathbb{R}$-Vektorraum, dann gilt:*

(i) *Für alle $v, w \in V$ gibt es genau ein $x \in V$, sodass $v + x = w$. Wir definieren $w - v := x$. Das heißt $\mathbf{0}$ ist das einzige Element, das* (VA_3) *erfüllt.*

(ii) *Für alle $v, w \in V$ gilt: $w - v = w + (-v)$.*

(iii) *Für alle $\lambda \in \mathbb{R}$ und $v \in V$ gilt $\lambda\mathbf{0} = \mathbf{0} = 0v$. Ferner ist $-\mathbf{0} = \mathbf{0}$ und für alle $v \in V$ ist $0 - v = -v$.*

(iv) *Für alle* $\lambda \in \mathbb{R}$ *und* $v \in V$ *gilt* $\lambda(-v) = (-\lambda)v$. *Insbesondere gilt:* $-v = (-1)v$.

(v) *Für alle* $\lambda \in \mathbb{R}$ *und* $v, w \in V$ *gilt* $\lambda(v - w) = \lambda v - \lambda w$,

(vi) *Für alle* $\lambda_1, \lambda_2 \in \mathbb{R}$, $v \in V$ *gilt* $(\lambda_1 - \lambda_2)v = \lambda_1 v - \lambda_2 v$.

Beweis: Als Übungsaufgabe. □

Sei ab jetzt in diesem Abschnitt $(V, +, \cdot, \mathbf{0}_V)$ stets ein $\mathbb{R}$-Vektorraum.

Definition II.2.5 (Untervektorraum): Eine Teilmenge U von V heißt *Untervektorraum* von V, wenn gelten:

(i) $\mathbf{0}_V \in U$,

(ii) Für alle $x, y \in U$ gilt $x + y \in U$,

(iii) Für alle $\lambda \in \mathbb{R}$ und $x \in U$ gilt $\lambda x \in U$.

Proposition II.2.6: *Ist* U *ein Untervektorraum von* $(V, +, \cdot, \mathbf{0}_V)$, *dann gilt insbesondere:* U *ist mit den gleichen Verknüpfungen wie* V *und mit* $\mathbf{0}_V$ *ein* $\mathbb{R}$*-Vektorraum.*

Beweis: Die Axiome VA_1, VA_2, VA_3, SM_1, SM_2, SA_3 und SA_4 gelten in U, da sie für alle Vektoren in V gelten. Zu VA_4: Für alle $x \in U$ gibt es genau ein $y \in U$ mit $x + y = \mathbf{0} = y + x$. Da V ein $\mathbb{R}$-Vektorraum ist, gibt es in V ein eindeutiges Element $-x$, das das Gewünschte leistet. Nach Proposition II.2.4 ist $-x = (-1)x$, d. h. $-x \in U$. Somit gilt auch VA_4. □

Proposition II.2.7: *Sei* I *eine Menge,* $I \neq \varnothing$. *Für jedes* $i \in I$ *sei ein Untervektorraum* U_i *von* V *gegeben. Dann ist*

$$W := \bigcap_{i \in I} U_i$$

ebenfalls ein Untervektorraum von V.

Beweis: (i) Für alle $i \in I$ ist U_i ein Untervektorraum von V, d. h. für alle $i \in I$ gilt $\mathbf{0}_V \in U_i$; also $\mathbf{0}_V \in \bigcap_{i \in I} U_i$.

(ii) Seien $x, y \in W = \bigcap_{i \in I} U_i$, d. h. für alle $i \in I$ gilt $x, y \in U_i$. Da die U_i Untervektorräume sind, gilt für alle $i \in I$, dass $x + y \in U_i$, also $x + y \in W$.

(iii) Zeige analog zu (ii), dass für alle $\lambda \in \mathbb{R}$ und alle $x \in W = \bigcap_{i \in I} U_i$ gilt, dass $\lambda x \in W$. □

Proposition II.2.8 (Summe von Vektorräumen): *Seien U_1, U_2 Untervektorräume von V und*

$$U_1 + U_2 := \{x + y \mid x \in U_1, y \in U_2\}.$$

Dann ist $U_1 + U_2 \subseteq V$ ein Untervektorraum.

Beweis: (i) Es ist $\mathbf{0}_V = \mathbf{0}_V + \mathbf{0}_V \in U_1 + U_2$,

(ii) Seien $u, w \in U_1 + U_2$. Wegen der Definition von $U_1 + U_2$ gibt es $x_1, x_2 \in U_1$ und $y_1, y_2 \in U_2$, sodass $u = x_1 + y_1$ und $w = x_2 + y_2$. Jetzt ist

$$u + w = (x_1 + y_1) + (x_2 + y_2) = (x_1 + x_2) + (y_1 + y_2),$$

und da U_1, U_2 Untervektorräume von V sind, gelten $x_1 + x_2 \in U_1$, $y_1 + y_2 \in U_2$, also $u + w \in U_1 + U_2$.

(iii) Analog zu (ii). □

Definition II.2.9 (Summe von Untervektorräumen): Seien $n \in \mathbb{N}$ eine natürliche Zahl und $U_1, \ldots, U_n$ Untervektorräume von V. Dann heißt

$$U_1 + \cdots + U_n := \{x_1 + \cdots + x_n \mid x_1 \in U_1, \ldots, x_n \in U_n\}$$

die *Summe* von $U_1, \ldots, U_n$.

Aus Proposition II.2.8 folgt, dass $U_1 + \cdots + U_n$ wiederum ein Untervektorraum von V ist.

3. Matrizen

Notation: Seien $a, b \in \mathbb{Z}$ mit $a \leq b$. Dann schreiben wir

$$\sum_{i=a}^{b} f(i) := f(a) + f(a+1) + \cdots + f(b).$$

Für $a > b$ setzen wir $\sum_{i=a}^{b} f(i) := 0$. Ferner schreiben wir

$$\prod_{i=a}^{b} f(i) := f(a) \cdot f(a+1) \cdots f(b)$$

und für $a > b$ setzen wir $\prod_{i=a}^{b} f(i) := 1$.

In diesem Abschnitt seien $p, q, m, n \in \mathbb{N}$.

Beispiel: Die folgenden „Dinge“ sind Matrizen:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{R}^{2\times 3}, \qquad \begin{pmatrix} 0 & -3,75 \\ 1,01 & 2,79 \end{pmatrix} \in \mathbb{R}^{2\times 2}, \qquad \begin{pmatrix} 1 & 2 \\ 7 & 42 \\ \pi & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

Definition II.3.1 (Matrix):

(i) Eine *reelle $p \times q$-Matrix* ist eine Abbildung

$$A\colon \{1, 2, \ldots, p\} \times \{1, 2, \ldots, q\} \longrightarrow \mathbb{R}.$$

Dabei heißt p die *Anzahl der Zeilen* und q die *Anzahl der Spalten* von A und man schreibt $a_{i,j} := A((i,j))$ und notiert die Matrix suggestiv als

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,q} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,q} \\ \vdots & \vdots & \cdots & \vdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,q} \end{pmatrix}.$$

A heißt *quadratisch*, wenn $p = q$ ist.

(ii) Die Menge

$$\mathbb{R}^{p\times q} := \{A \mid A \text{ ist reelle } p \times q\text{-Matrix}\} = \mathrm{Abb}(\{1, \ldots, p\}\times\{1, \ldots, q\}, \mathbb{R})$$

heißt die Menge der reellen $p \times q$-Matrizen. Alternative Notationen sind $\mathrm{Mat}(p \times q, \mathbb{R})$, $\mathrm{Mat}(p, q)$

(iii) Wir schreiben $\mathbb{R}^p := \mathbb{R}^{p\times 1}$.

Bemerkung II.3.2: In Proposition II.3.1 kann auch $p = 0$ oder $q = 0$ zugelassen werden, wir bleiben aber meist bei $p, q \in \mathbb{N}$.

Beispiel II.3.3: (i) Sei A die Matrix in $\mathbb{R}^{p\times q}$ mit $a_{i,j} = A((i,j)) = 0$. Diese Matrix heißt *Nullmatrix* und wird mit $\mathbf{0} = \mathbf{0}_{p\times q}$ notiert.

(ii) Sei A die quadratische Matrix im $\mathbb{R}^{p\times p}$ mit

$$a_{i,j} = A((i,j)) := \delta_{i,j} = \begin{cases} 1, & i = j, \\ 0, & \text{sonst.} \end{cases}$$

$\delta_{i,j}$ heißt das *Kronecker-Symbol*. A heißt *Eins-Matrix* oder auch *Einheitsmatrix* und wird auch mit I_p oder E_p oder $\mathbb{1}_p$ notiert. („I“ kommt von *Identity*).

Definition II.3.4: Für $A, B \in \mathbb{R}^{p\times q}$ und $\lambda \in \mathbb{R}$ definieren wir

(i) $A + B := C$ mit $C((i,j)) = A((i,j)) + B((i,j))$,

(ii) $\lambda A := D$ mit $D((i,j)) = \lambda A((i,j))$.

Um die Lesbarkeit zu verbessern schreiben wir im Folgenden $A(i,j) := A((i,j))$.

Bemerkung II.3.5: Die Verknüpfungen auf $\mathbb{R}^{p\times q}$ aus Proposition II.3.4 sind die gleichen wie in Proposition II.2.3. Insbesondere wird damit $\mathbb{R}^{p\times q}$ zum $\mathbb{R}$-Vektorraum. Der 0-Vektor in diesem Vektorraum ist die Nullmatrix.

Definition II.3.6 (Matrizenmultiplikation): Seien $A \in \mathbb{R}^{p\times q}$ und $B \in \mathbb{R}^{q\times m}$. Wir definieren die Matrix $A \cdot B \in \mathbb{R}^{p\times m}$ als die Matrix C mit den Einträgen

$$C(i,k) = \sum_{j=1}^{q} A(i,j)B(j,k)$$

Beispiel II.3.7: Es ist

$$\begin{pmatrix} 1 & 2 & 3 \\ -3 & -2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1-2+6 & 0+0+3 \\ -3+2-2 & 0+0-1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ -3 & -1 \end{pmatrix}.$$

Definition II.3.8: Für $A \in \mathbb{R}^{m\times n}$ heißt die Matrix $B \in \mathbb{R}^{n\times m}$, definiert durch $B(i,j) = A(j,i)$ für alle $(i,j) \in \{1,\dots,n\}\times\{1,\dots,m\}$ die *transponierte Matrix* oder die *Transponierte* zu A. Wir schreiben für B auch A^t. Gebräuchlich ist auch $A^\top$.

Beispiel II.3.9:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Proposition II.3.10: *Es gelten (ergänzend zu denen aus Proposition II.3.5) folgende Rechenregeln für Matrizen (deren Zeilen– und Spaltenzahlen so beschaffen sind, dass die nachfolgenden Ausdrücke wohldefiniert sind)*

(i) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$, *(Assoziativität von $\cdot$)*

(ii) $A \cdot (B + C) = A \cdot B + A \cdot C$, *(Distributivitätsgesetz I)*

(iii) $(A + B) \cdot C = A \cdot C + B \cdot C$, *(Distributivitätsgesetz II)*

(iv) *Für alle* $r \in \mathbb{R}$ *ist* $A \cdot (r \cdot B) = (r \cdot A) \cdot B = r \cdot (AB)$,

(v) $(A + B)^t = A^t + B^t$, $(A \cdot B)^t = B^t \cdot A^t$ *und* $(A^t)^t = A$.

(vi) *Für die Einheitsmatrix* $I_p \in \mathbb{R}^{p\times p}$ *gilt* $I_p \cdot A = A$ *und* $B \cdot I_p = B$

Achtung: Beim Produkt wird durch Transponieren die Reihenfolge vertauscht.

Beispiel II.3.11: Matrizenmultiplikation ist *nicht* kommutativ; auch nicht für quadratische Matrizen! Für $A = \left(\begin{smallmatrix} 1 & 2 \\ -1 & 1 \end{smallmatrix}\right), B = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ sind

$$AB = \begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 3 \\ 1 & 2 \end{pmatrix}.$$

Beweis (von Proposition II.3.10): Wir zeigen exemplarisch (i). Dazu seien $A \in \mathbb{R}^{p\times q}$, $B \in \mathbb{R}^{q\times m}$, $C \in \mathbb{R}^{m\times n}$ und $(a,b) \in \{1,\dots,p\} \times \{1,\dots,n\}$. Dann ist

$$\begin{aligned}
(A\cdot(B\cdot C))(a,b) &= \sum_{x=1}^{q} A(a,x)\cdot(B\cdot C)(x,b) \\
&= \sum_{x=1}^{q} A(a,x)\cdot\Big(\sum_{y=1}^{m} B(x,y)C(y,b)\Big) \\
&= \sum_{x=1}^{q}\sum_{y=1}^{m} A(a,x)B(x,y)C(y,b) \\
&= \sum_{y=1}^{m}\sum_{x=1}^{q} \Big(A(a,x)B(x,y)\Big)C(y,b) \\
&= \sum_{y=1}^{m}(A\cdot B)(a,y)C(y,b) = ((A\cdot B)\cdot C)(a,b),
\end{aligned}$$

d. h. $A\cdot(B\cdot C) = (A\cdot B)\cdot C$. □

4. Invertierbare Matrizen

In diesem Abschnitt seien p, q, n, m stets natürliche Zahlen.

Definition II.4.1 (Invertierbare Matrizen):

(i) Eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ heißt *invertierbar* oder *regulär*, falls es $B \in \mathbb{R}^{n\times n}$ gibt mit $A\cdot B = I_n = B\cdot A$.

(ii) Wir schreiben für die Menge der invertierbaren $n\times n$-Matrizen

$$\mathrm{Gl}_n(\mathbb{R}) := \{A \in \mathbb{R}^{n\times n} \mid A \text{ ist invertierbar}\}.$$

„Gl“ steht für *General linear group*.

Definition II.4.2 (Inverse Matrix): Sei $A \in \mathbb{R}^{n\times n}$ eine reguläre Matrix. Die Matrix B aus Proposition II.4.1 ist eindeutig, d. h. es gibt genau eine Matrix $B \in \mathbb{R}^{n\times n}$ mit $A\cdot B = I_n = B\cdot A$. Wir nennen B die *inverse Matrix* oder die *Inverse* von A und schreiben dafür auch A^{-1}.

Beweis: Seien $B, B' \in \mathbb{R}^{n\times n}$ mit $A \cdot B = B \cdot A = I_n = A \cdot B' = B' \cdot A$. Dann folgt

$$B = B \cdot I_n = B \cdot A \cdot B' = I_n \cdot B' = B'. \qquad \square$$

Beispiel II.4.3 (Spezielle invertierbare Matrizen): Seien $\alpha, \beta, a_1, \ldots, a_n$ und $b_1, \ldots, b_n \in \mathbb{R}$.

(i) Ist $A = \left(\begin{smallmatrix} 1 & \alpha \\ 0 & 1 \end{smallmatrix}\right)$, dann ist $B = \left(\begin{smallmatrix} 1 & -\alpha \\ 0 & 1 \end{smallmatrix}\right)$ die Inverse von A:

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

A ist also invertierbar mit Inverser $B = A^{-1}$. Wie wirkt die Multiplikation mit A auf Matrizen?

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 & a_2 & \cdots & a_1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} a_1 + \alpha b_1 & a_2 + \alpha b_2 & \cdots & a_n + \alpha b_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix},$$

d. h. Multiplikation mit A von links bewirkt Addition des α-fachen der zweiten Zeile zur ersten Zeile.

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \cdot \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 & \alpha a_1 + b_1 \\ a_2 & \alpha a_2 + b_2 \\ \vdots & \vdots \\ a_n & \alpha a_n + b_n \end{pmatrix},$$

d. h. Multiplikation mit A von rechts bewirkt Addition des α-fachen der ersten Spalte zur zweiten Spalte.

(ii) Ist $V = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, dann ist $V \cdot V = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, d. h. V ist invertierbar mit $V^{-1} = V$. Wir wirkt die Multiplikation mit V auf Matrizen?

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_1 & a_2 & \cdots & a_1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ a_1 & a_2 & \cdots & a_1 \end{pmatrix},$$

d. h. Multiplikation von links mit V bewirkt Vertauschung von erster und zweiter Zeile.

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b_1 & a_1 \\ b_2 & a_2 \\ \vdots & \vdots \\ b_n & a_n \end{pmatrix},$$

d. h. Multiplikation von rechts mit V bewirkt Vertauschung der ersten und der zweiten Spalte.

(iii) Falls $\beta \in \mathbb{R}$ und $\alpha \neq 0 \neq \beta$, setze $D := \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$. Dann ist $E = \left(\begin{smallmatrix} \alpha^{-1} & 0 \\ 0 & \beta^{-1} \end{smallmatrix}\right)$ die Inverse von D, d. h. D ist invertierbar mit Inverser E. Wie wirkt die Multiplikation mit D auf Matrizen?

$$\begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} \begin{pmatrix} a_1 & \cdots & a_n \\ b_1 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 & \cdots & \alpha a_n \\ \beta b_1 & \cdots & \beta b_n \end{pmatrix},$$

d. h. Multiplikation mit D von links multipliziert die erste Zeile mit α und die zweite Zeile mit β, und

$$\begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} = \begin{pmatrix} \alpha a_1 & \beta b_1 \\ \vdots & \vdots \\ \alpha a_n & \beta b_n \end{pmatrix},$$

d. h. Multiplikation mit D von rechts multipliziert die erste Spalte mit α und die zweite Zeile mit β.

Definition II.4.4 (Standardbasismatrizen): Seien $i, j \in \mathbb{N}$ mit $1 \leq i \leq m$ und $1 \leq j \leq n$. Wir definieren die Matrix $E_{i,j}$ wie folgt:

$$E_{i,j} \colon \{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow \mathbb{R}, \qquad (k, l) \longmapsto \begin{cases} 1, & \text{falls } k = i \text{ und } l = j, \\ 0, & \text{sonst.} \end{cases}$$

Die Matrix $E_{i,j}$ enthält also genau eine 1 und sonst nur Nullen. Die 1 steht an der Stelle (i, j). Die Matrizen $E_{i,j}$ heißen *Standardbasismatrizen*.

Achtung: Die Standardbasismatrizen sind nicht invertierbar.

Beispiel II.4.5 (Ein paar Standardbasismatrizen): Die Matrizen

$$E_{1,2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 4},$$

$$E_{3,1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}, \qquad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

sind Standardbasismatrizen.

Bemerkung II.4.6: Für eine beliebige Matrix $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$ gilt

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} E_{i,j}.$$

Proposition II.4.7 (Rechenregeln für Standardbasismatrizen): *Sei* $E_{i,j} \in \mathbb{R}^{q\times m}$.

(i) *Für* $E_{k,l} \in \mathbb{R}^{m\times n}$ *gilt*

$$E_{i,j} \cdot E_{k,l} = \begin{cases} E_{i,l}, & \text{falls } k = j, \\ \mathbf{0}, & \text{sonst.} \end{cases}$$

Hierbei ist $\mathbf{0}$ *die Nullmatrix in* $\mathbb{R}^{q\times n}$.

(ii) *Für* $M \in \mathbb{R}^{m\times n}$ *gilt*

$$E_{i,j} \cdot M = \sum_{b=1}^{n} M(j,b) E_{i,b} \in \mathbb{R}^{q\times n}.$$

$E_{i,j} \cdot M$ *ist die Matrix, die in der* i*-ten Zeile die* j*-te Zeile von* M *enthält, und sonst nur Nullen.*

(iii) *Für* $M \in \mathbb{R}^{p\times q}$ *gilt*

$$M \cdot E_{i,j} = \sum_{a=1}^{p} M(a,i) \cdot E_{a,j} \in \mathbb{R}^{p\times m}.$$

$M \cdot E_{i,j}$ *ist die Matrix, die in der* j*-ten Spalte die* i*-te Spalte von* M *enthält, und sonst nur Nullen.*

Beweis: (i) Sei $A = E_{i,j} \cdot E_{k,l} \in \mathbb{R}^{q\times n}$. Für $(a,b) \in \{1,\dots,q\} \times \{1,\dots,n\}$ gilt dann:

$$A(a,b) = \sum_{x=1}^{m} E_{i,j}(a,x) \cdot E_{k,l}(x,b).$$

Dieser Eintrag ist 0, außer wenn $i = a$, $j = x = k$ und $l = b$. In diesem Fall ist der obige Eintrag 1. Also ist $A(a,b) = 1$ genau dann, wenn $(a,b) = (i,l)$ und $k = j$ und sonst 0.

(ii) Für eine Matrix M aus $\mathbb{R}^{m\times n}$ gilt

$$E_{i,j} \cdot M = E_{i,j} \cdot \Big(\sum_{a=1}^{m} \sum_{b=1}^{n} M(a,b) E_{a,b} \Big) = \sum_{a=1}^{m} \sum_{b=1}^{n} M(a,b) E_{i,j} E_{a,b},$$

wobei wir Proposition II.3.10 für die letzte Gleichheit verwendet haben. Weil $E_{i,j}E_{a,b} = 0$ genau dann gilt, wenn $j \neq a$, erhalten wir

$$E_{i,j} \cdot M = \sum_{b=1}^{n} M(j,b) E_{i,b}.$$

(iii) Funktioniert völlig analog zu (iii). □

Definition II.4.8 (Elementarmatrizen): Seien $a, \alpha_1, \ldots, \alpha_n$ reelle Zahlen und $i, j \in \{1, \ldots, n\}$ mit $i \neq j$. Wir definieren 3 Typen quadratischer Matrizen in $\mathbb{R}^{n\times n}$ wie folgt:

(i) *Additionsmatrizen*:

$$A_{i,j}^{\alpha} := I_n + \alpha E_{i,j}$$

Alle Einträge auf der Diagonalen der Matrix A sind 1, der Eintrag an der Stelle (i, j) ist α, alle anderen Einträge sind Null.

(ii) *Vertauschungsmatrizen*:

$$V_{i,j} := I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$$

$V_{i,j}$ entsteht aus der Einheitsmatrix, indem man die Einsen an den Stellen (i, i) und (j, j) ersetzt durch Einsen an der Stelle (i, j) und (j, i).

(iii) *Diagonalmatrizen*:

$$\operatorname{diag}(\alpha_1, \ldots, \alpha_n) := \sum_{i=1}^{n} \alpha_i E_{i,i}.$$

Die Einträge auf der Diagonalen sind $\alpha_1, \ldots, \alpha_n$, alle anderen Einträge sind Null.

Proposition II.4.9 (Invertierbarkeit der Elementarmatrizen): *Die Matrizen aus Proposition II.4.8 sind invertierbar, d. h. liegen in* $\mathrm{Gl}_n(\mathbb{R})$.

Beweis: Wir geben hier nur den Beweis dafür, dass Additionsmatrizen invertierbar sind; die restlichen Aussagen sind Konsequenzen der nachfolgenden Proposition. Nach Definition ist

$$A_{i,j}^{\alpha} \cdot A_{i,j}^{-\alpha} = (I_n + \alpha E_{i,j})(I_n - \alpha E_{i,j}) = I_n - \alpha E_{i,j} + \alpha E_{i,j} - \alpha^2 E_{i,j} E_{i,j}$$

und nach Voraussetzung gilt $i \neq j$, aus Proposition II.4.7 wissen wir also, dass $E_{i,j}E_{i,j} = \mathbf{0}$ und damit $A_{i,j}^{\alpha} \cdot A_{i,j}^{-\alpha} = I_n$, $A_{i,j}^{-\alpha}$ ist also die Inverse zu $A_{i,j}^{\alpha}$. Ferner hat $V_{i,j}$ die Inverse $V_{i,j}$ und $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ hat die Inverse $\operatorname{diag}(\alpha_1^{-1}, \ldots, \alpha_n^{-1})$. □

Proposition II.4.10 (Elementarmatrizen und Zeilenumformungen): *Die Matrizen aus Proposition II.4.8 wirken bei Multiplikation von links auf eine Matrix* $M \in \mathbb{R}^{n\times m}$ *wie folgt:*

(i) *$A_{i,j}^{\alpha}M$ entsteht aus M durch Addition des α-fachen der j-ten Zeile zur i-ten Zeile.*

(ii) *$V_{i,j}M$ entsteht aus M durch Vertauschen der i-ten und der j-ten Zeile.*

(iii) *$\operatorname{diag}(\alpha_1, \ldots, \alpha_n)M$ entsteht aus M durch Multiplikation der k-ten Zeile mit α_k für alle $k \in \{1, \ldots, n\}$.*

Beweis: Verwenden wir die Ergebnisse aus Proposition II.4.7, so können wir die Aussagen einfach nachrechnen:

(i) Es ist

$$A_{i,j}^{\alpha}M = (I_n + \alpha E_{i,j})M = M + \alpha E_{i,j}M \\ = M + \alpha\Big(\sum_{b=1}^{m} M(j,b)E_{i,b}\Big) = M + \sum_{b=1}^{m} \alpha M(j,b)E_{i,b}.$$

(ii) Es ist

$$\begin{aligned} V_{i,j}M &= (I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i})M \\ &= M - \Big(\sum_{b=1}^{n} M(i,b)E_{i,b}\Big) - \Big(\sum_{b=1}^{n} M(j,b)E_{j,b}\Big) + \Big(\sum_{b=1}^{n} M(j,b)E_{i,b}\Big) \\ &\qquad + \Big(\sum_{b=1}^{n} M(i,b)E_{j,b}\Big) \\ &= M - \Big(\sum_{b=1}^{n}(M(j,b) - M(i,b))E_{i,b}\Big) + \Big(\sum_{b=1}^{n}(M(i,b) - M(j,b))E_{j,b}\Big), \end{aligned}$$

in der i-ten Zeile stehen also die Einträge $M(j,b)$ und in der j-ten Zeile die Einträge $M(i,b)$.

(iii) Es ist

$$\operatorname{diag}(\alpha_1, \ldots, \alpha_n)M = \Big(\sum_{i=1}^{n} \alpha_i E_{i,i}\Big)\Big(\sum_{a=1}^{n}\sum_{b=1}^{m} M(a,b)E_{a,b}\Big) = \sum_{i=1}^{n}\sum_{b=1}^{m} \alpha_i M(i,b)E_{i,b},$$

wobei wir verwendet haben, dass $E_{i,i}E_{a,b} = \mathbf{0}$, falls $i \neq a$. In der i-ten Zeile stehen also die Einträge $\alpha_i M(i,b)$. □

Proposition II.4.11 (Elementarmatrizen und Spaltenumformungen): *Die Matrizen aus Proposition II.4.8 wirken auf $M \in \mathbb{R}^{m\times n}$ bei Multiplikation von rechts wie folgt:*

(i) *$MA^{\alpha}_{i,j}$ entsteht aus M durch Addition des α-fachen der i-ten Spalte auf die j-te Spalte.*

(ii) *$MV_{i,j}$ entsteht aus M durch Vertauschung der i-ten und j-ten Spalte.*

(iii) *$M\mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ entsteht aus M durch Multiplikation der k-ten Spalte mit α_k für alle $k \in \{1, \ldots, n\}$.*

Beweis: Verwenden wir $(MA)^t = A^t M^t$, können wir uns zum Beweisen der Aussagen auf die Aussagen in Proposition II.4.10 zurückziehen.

Zu (i): Es ist $(M \cdot A^{\alpha}_{i,j})^t = (A^{\alpha}_{i,j})^t M^t = A^{\alpha}_{j,i} M^t$. Jetzt ist $A^{\alpha}_{j,i} M^t$ aus M^t entstanden durch Addition des α-fachen der i-ten Zeile zur j-ten Zeile, also entsteht $MA^{\alpha}_{i,j}$ aus M durch Addition des α-fachen der i-ten Spalte zur j-en Spalte.

(ii) und (iii) funktionieren analog. □

Beispiel II.4.12 (Inverse in $\mathbf{SL_2}(\mathbb{R})$): Sei $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$. Dann ist A invertierbar genau dann, wenn $ad - bc \neq 0$. In diesem Fall ist

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Lemma II.4.13 (Multiplikation von Blockmatrizen): *Seien $M_1 \in \mathbb{R}^{p\times q}$ und $M_2 \in \mathbb{R}^{q\times m}$ Matrizen der folgenden Gestalt:*

$$M_1 = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad M_2 = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$$

mit Blöcken $A \in \mathbb{R}^{p_1\times q_1}$, $B \in \mathbb{R}^{p_1\times q_2}$, $C \in \mathbb{R}^{p_2\times q_1}$, $D \in \mathbb{R}^{p_2\times q_2}$, $E \in \mathbb{R}^{q_1\times m_1}$, $F \in \mathbb{R}^{q_1\times m_2}$, $G \in \mathbb{R}^{q_2\times m_1}$ und $H \in \mathbb{R}^{q_2\times m_2}$, wobei $p_1, p_2, q_1, q_2, m_1, m_2 \in \mathbb{N}$ mit $p_1 + p_2 = p$, $q_1 + q_2 = q$, $m_1 + m_2 = m$. Dann gilt

$$M_1 M_2 = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}.$$

Beweis: Folgt aus Definition der Matrizenmultiplikation. Nachrechnen! □

Proposition II.4.14 (Inverse für spezielle Blockmatrizen): *Sei $p < n$ und A in $\mathbb{R}^{n\times n}$ mit*

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

mit $I_p \in \mathbb{R}^{p\times p}$, $\mathbf{0} \in \mathbb{R}^{(m-p)\times p}$, $B \in \mathbb{R}^{p\times(n-p)}$ und $D \in \mathbb{R}^{(n-p)\times(n-p)}$. A ist invertierbar genau dann, wenn D invertierbar ist. In diesem Fall ist

$$A^{-1} = \begin{pmatrix} I_p & -BD^{-1} \\ \mathbf{0} & D^{-1} \end{pmatrix}.$$

Beweis: „⇒“: Sei A invertierbar, d. h. es gibt $M \in \mathbb{R}^{n\times n}$ mit $AM = MA = I_n$. Schreibe

$$M = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$$

mit $E \in \mathbb{R}^{p\times p}$, $F \in \mathbb{R}^{p\times(n-p)}$, $G \in \mathbb{R}^{(n-p)\times p}$ und $H \in \mathbb{R}^{(n-p)\times(n-p)}$. Mithilfe von Proposition II.4.13 lesen wir ab:

$$I_n = MA = \begin{pmatrix} E & EB + FD \\ G & GB + HD \end{pmatrix},$$

d. h. $E = I_p$, $G = \mathbf{0}$, $I_{n-p} = GB + HD = HD$ und

$$I_n = AM = \begin{pmatrix} I_p & F + BH \\ \mathbf{0} & DH \end{pmatrix},$$

d. h. $DH = I_{n-p}$. Somit ist D invertierbar mit Inverser H.

„⇐“: Man kann jetzt durch Nachrechnen (mithilfe von Proposition II.4.13) zeigen, dass die angegebenen Matrix für A^{-1} tatsächlich invers zu A ist. □

Proposition II.4.15 (Inverse von Produkt): *Sind $A, B \in \mathrm{Gl}_n(\mathbb{R})$, dann ist auch $AB \in \mathrm{Gl}_n(\mathbb{R})$ und für die Inverse gilt $(AB)^{-1} = B^{-1}A^{-1}$.*

Beweis: Es gilt $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A = AIA^{-1} = AA^{-1} = I$, und genau so ist $(B^{-1}A^{-1})(AB) = I$. □

5. Lineare Gleichungssysteme

In diesem Abschnitt seien stets $n, m \in \mathbb{N}$.

Definition II.5.1 (Lineares Gleichungssystem):

(i) Ein *reelles lineare Gleichungssystem* mit n Gleichungen und m Variablen oder Unbekannten ist ein System

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m &= b_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m &= b_2, \\ \vdots \qquad\quad \vdots \qquad\quad \cdots \qquad\quad \vdots \qquad\quad &\vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m &= b_n. \end{aligned}$$

Hierbei sind $a_{i,j}, b_i \in \mathbb{R}$ für $1 \le i \le n$ und $1 \le j \le m$. Die $a_{i,j}$ heißen *Koeffizienten* und die x_j *Variablen*.

(ii) Das lineare Gleichungssystem hat folgende *schematische Beschreibung*

$$\left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,m} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,m} & b_n \end{array}\right)$$

Die Matrix $A = (a_{i,j})$ heißt *Koeffizientenmatrix* und $(A|b)$ heißt auch *erweiterte Matrix* zum linearen Gleichungssystem.

(iii) Die Menge

$$\mathbb{L} := \mathbb{L}(A, b) := \{x = (x_1, \ldots, x_m)^t \in \mathbb{R}^m \mid x_1, \ldots, x_m \text{ erfüllen das LGS}\}$$

heißt *Lösungsmenge* und $\mathbb{L}^h := \mathbb{L}(A, \mathbf{0})$ heißt *Lösungsmenge des homogenen linearen Gleichungssystems.*

Ab jetzt seien $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$ und wir betrachten das lineare Gleichungssystem mit der schematischen Beschreibung $(A|b)$.

Bemerkung II.5.2: Es gilt $Ax = b$ genau dann, wenn $x \in \mathbb{L}(A|b)$.

Proposition II.5.3 (Struktur der allgemeinen Lösungsmenge):

(i) $\mathbb{L}^h = \mathbb{L}(A, \mathbf{0})$ *ist ein Untervektorraum von* $\mathbb{R}^m$.

(ii) *Ist* $x^{(s)}$ *eine spezielle Lösung von* $(A|b)$*, dann gilt*

$$\mathbb{L} := \mathbb{L}(A, b) = x^{(s)} + \mathbb{L}^h := \{x^s + v \mid v \in \mathbb{L}^{(h)}\}.$$

Beweis: (i) Seien $x, y \in \mathbb{L}^h$, d. h. $Ax = 0$ und $Ay = 0$. Dann haben wir $A(x + y) = Ax + Ay = \mathbf{0} + \mathbf{0} = \mathbf{0}$ und damit $x + y \in \mathbb{L}^h$. Sind $\lambda \in \mathbb{R}$ und $x \in \mathbb{L}^h$, dann ist wieder $Ax = \mathbf{0}$, sodass $A(\lambda x) = \lambda Ax = \lambda\mathbf{0} = \mathbf{0}$, d. h. $\lambda x \in \mathbb{L}^h$.

(ii) „$\subseteq$“: Für x in $\mathbb{L}(A, b)$ gilt $A(x - x^{(s)}) = Ax - Ax^{(s)} = b - b = \mathbf{0}$, sodass $v := x - x^{(s)}$ zu $\mathbb{L}^h$ gehört, und $x = x^{(s)} + v$ ist.
„$\supseteq$“: Für $x = x^{(s)}+v$ mit $v \in \mathbb{L}^h$ gilt $Ax = A(x^{(s)}+v) = Ax^{(s)}+Av = b+\mathbf{0} = b$, also $x \in \mathbb{L}$. □

Proposition II.5.4 (Lösungsstrategie für lineare Gleichungssysteme): *Sei* C *eine reguläre* $n \times n$*-Matrix mit reellen Einträgen. Dann gilt für* $x \in \mathbb{R}^n$*:*

$$Ax = b \Longleftrightarrow CAx = Cb.$$

Beweis: Für „⇐“: Ist $CAx = Cb$, dann ist $Ax = C^{-1}CAx = C^{-1}Cb = b$. „⇒“ ist klar. □

Korollar II.5.5 (Elementare Zeilenumformungen): *Wählt man in Proposition II.5.4 (mit der Notation aus Proposition II.4.8)*

(i) $C = A_{i,j}^{\alpha}$,

(ii) $C = V_{i,j}$ *oder*

(iii) $C = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ *mit* $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ *und* $\alpha_1 \cdots \alpha_n \neq 0$,

dann erhält man die elementaren Zeilenumformungen

(i) *Addition des α-Fachen der j-ten Gleichung zur i-ten Gleichung,*

(ii) *Vertauschen der i-ten und der j-ten Gleichung bzw.*

(iii) *Multiplizieren der i-ten Gleichung mit α_i.*

Die elementaren Zeilenumformungen verändern also die Lösungsmenge des linearen Gleichungssystems nicht.

Definition II.5.6 (Einheitsvektoren): Für $1 \leq i \leq n$ definieren wir in $\mathbb{R}^n = \mathbb{R}^{n\times 1}$ den Vektor e_i durch $e_i((j,1)) = \delta_{i,j}$, d. h. $e_i = (\delta_{i,j})_{1\leq j\leq n}^t$. Der Vektor e_i heißt *Einheitsvektor*. Beachte:

(i) Der Vektor $x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$ lässt sich schreiben als $x = \sum_{i=1}^n x_i e_i$.

(ii) Ist $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$, dann ist $Ae_j = \sum_{t=1}^n a_{t,j} e_t$, die j-te Spalte von A.

Definition II.5.7 (Treppenform, Gauß-Normalform, Rang): Eine Matrix $T = (t_{i,j}) \in \mathbb{R}^{n\times m}$ hat *Treppenform* bzw. *Gauß-Normalform* genau dann, wenn gilt: Es gibt $r \in \mathbb{N}_0$ und $s_1, \ldots, s_r \in \mathbb{N}$ mit $1 \leq s_1 < \cdots < s_r \leq m$ mit:

(i) Für alle $i \in \{1, \ldots, r\}$ gilt: $t_{i,s_i} = 1$, für $k \neq i$ ist $t_{k,s_i} = 0$ und für $k < s_i$ ist $t_{i,k} = 0$,

(ii) Für alle $i \geq r+1, j \in \{1, \ldots, m\}$ gilt $t_{i,j} = 0$.

Die Zahl r heißt der *Rang* von T und $s_1, \ldots, s_r$ heißen die *Spaltenindizes*. T ist also in Treppenform, wenn für $1 \leq i \leq r$ die s_i-te Spalte von T der Einheitsvektor e_i ist, links von der 1 an der Stelle (i,j) nur Nullen stehen und ab der $(r+1)$-ten Zeile alle Zeilen Nullzeilen sind.

Ab jetzt sei $T = (t_{i,j}) \in \mathbb{R}^{n\times m}$ eine Matrix in Treppenform vom Rang r und mit Spaltenindizes $s_1, \ldots, s_r \in \mathbb{N}_0$.

Lemma II.5.8 (Spezielle Lösung für Treppenform): *Für $b \in \mathbb{R}^n$ gilt: Das lineare Gleichungssystem mit der schematischen Beschreibung $(T|b)$ ist lösbar genau dann, wenn $b_{r+1} = \cdots = b_n = 0$. In diesem Fall ist der folgende Vektor $x^{(s)} \in \mathbb{R}^n$ eine Lösung:*

$$x^{(s)} := \sum_{i=1}^{r} b_i e_{s_i}.$$

Beweis: „⇒“: Eine Gleichung der Form $0 = b_i$ mit $b_i \neq 0$ hat keine Lösung.
„⇐“: Es ist

$$Tx^{(s)} = T\Big(\sum_{i=1}^{r} b_i e_{s_i}\Big) = \sum_{i=1}^{r} b_i T e_{s_i} = \sum_{i=1}^{r} b_i e_i = \sum_{i=1}^{n} b_i e_i = b$$

wobei wir verwendet haben, dass $b_{r+1} = \cdots = b_n = 0$, d. h. $b_i e_i = 0$ für $i > r$.□

Beispiel II.5.9: Betrachte das lineare Gleichungssystem

$$\begin{aligned} x_2 + 0x_3 + \alpha x_4 + 0x_5 + \beta x_6 &= b_1 \\ x_3 + \gamma x_4 + 0x_5 + \delta x_6 &= b_2 \\ x_5 + \varepsilon x_6 &= b_3 \\ 0 &= b_4 \end{aligned}$$

Dieses lineare Gleichungssystem hat Rang $r = 3$ und die Spaltenindizes $s_1 = 2$, $s_2 = 3$ und $s_3 = 5$. Nach Proposition II.5.8 ist das lineare Gleichungssystem genau dann lösbar, wenn $b_4 = 0$ gilt. In diesem Fall ist eine Lösung

$$x^{(s)} = \sum_{i=1}^{r} b_i e_{s_i} = b_1 e_2 + b_2 e_3 + b_3 e_5 = (0, b_1, b_2, 0, b_3)^t.$$

Beispiel II.5.10: In Proposition II.5.9 erhalten wir folgende homogene Lösungen:

$$\begin{aligned} x_2 &= -\alpha x_4 - \beta x_6 \\ x_3 &= -\gamma x_4 - \delta x_6 \\ x_5 &= \qquad - \varepsilon x_6 \end{aligned}$$

Wähle

(i) $x_1 = 1$, $x_4 = 0$, $x_6 = 0$: Dann sind $x_2 = x_3 = x_5 = 0$ und $F^{(1)} = e_1$ ist eine Lösung des homogenen linearen Gleichungssystems.

(ii) $x_1 = 0$, $x_4 = 1$, $x_6 = 0$: Dann sind $x_2 = -\alpha$, $x_3 = -\gamma$, $x_5 = 0$ und es ist $F^{(4)} = e_4 - \alpha e_2 - \gamma e_3 + 0e_5$ eine Lösung des homogenen linearen Gleichungssystems.

(iii) $x_1 = 0$, $x_4 = 0$, $x_6 = 1$: Dann sind $x_2 = -\beta$, $x_3 = -\delta$, $x_5 = -\varepsilon$ und $F^{(6)} = e_6 - \beta e_2 - \delta e_3 - \varepsilon e_5$ ist eine Lösung des homogenen linearen Gleichungssystems.

Lemma II.5.11 (Lösungsmenge des homogenen linearen Gleichungssystems):

(i) *Seien für* $j \in J := \{1, \ldots, m\} - \{s_1, \ldots, s_r\}$

$$F^{(j)} := e_j - \sum_{i=1}^{r} t_{i,j} e_{s_i}.$$

Dann ist jedes $F^{(j)}$ *eine Lösung des homogenen linearen Gleichungssystems, d. h.* $Tx = \mathbf{0}$. *Die* $F^{(j)}$ *heißen* Fundamentallösungen.

(ii) *Für die Lösungsmenge* $\mathbb{L}^h = \mathbb{L}(T, \mathbf{0})$ *gilt*

$$\mathbb{L}^h = \left\{ \sum_{j \in J} \lambda_j F^{(j)} : \lambda_j \in \mathbb{R} \right\}$$

Weiterhin gilt: Für jedes $v \in \mathbb{L}^h$ *ist die Darstellung* $v = \sum_{j \in J} \lambda_j F^{(j)}$ *eindeutig.*

Beweis: (i) Es gilt

$$TF^{(j)} = Te_j - \sum_{i=1}^{r} t_{i,j} Te_{s_i} = Te_j - \sum_{i=1}^{r} t_{i,j} e_i = Te_j - \sum_{i=1}^{n} t_{i,j} e_i = \mathbf{0},$$

da $t_{i,j} = 0$ für $i \geq r + 1$.

(ii) „$\supseteq$“: Folgt aus (i) zusammen mit der Tatsache, dass $\mathbb{L}^h$ ein Untervektorraum von $\mathbb{R}^m$ ist.

„$\subseteq$“: Sei $x = \sum_{i=1}^{m} x_i e_i \in \mathbb{L}^h$, d. h. $Tx = \mathbf{0}$. Setze $v := x - \sum_{j \in J} x_j F^{(j)}$. Für dieses v gilt dann $Tv = \mathbf{0}$, d. h. $v \in \mathbb{L}^h$. Wenn wir jetzt zeigen können, dass $v = \mathbf{0}$, dann haben wir gezeigt, dass wir x auf die behauptete Art und Weise schreiben können. Wir haben bereits $v_j = 0$ für $j \in J$ und wollen verwenden, dass $Tv = \mathbf{0}$. Die i-te Zeile von Tv ist $\sum_{k=1}^{m} t_{i,k} v_k$ und wir wissen, dass $v_k = 0$ für $k \in J$. Für $i \in \{1, \ldots, r\}$ gilt $t_{i,s_i} = 1$ und $t_{i,k} = 0$, falls $k \notin J$ und $k \neq s_i$, d. h.

$$0 = \sum_{k=1}^{m} t_{i,k} v_k = 1 v_{s_i}$$

für alle $i \in \{1, \ldots, r\}$, v muss also der Nullvektor sein und wir sind fertig. □

Satz 2 (Lösungsmenge eines linearen Gleichungssystems): *Ist $T \in \mathbb{R}^{n\times m}$ eine Treppenmatrix vom Rang r mit Spaltenindizes $s_1, \ldots, s_r$ und $b \in \mathbb{R}^n$, dann gilt für die Lösungsmenge $\mathbb{L}$ des zugehörigen linearen Gleichungssystems*

$$\mathbb{L} = x^{(j)} + \left\{ \sum_{j\in J} \lambda_j F^{(j)} : \lambda_j \in \mathbb{R} \right\}$$

wobei $x^{(j)} = \sum_{i=1}^{r} b_i e_{s_i}$ eine spezielle Lösung ist und $F^{(j)} = e_j - \sum_{i=1}^{r} t_{i,j} e_{s_i}$ für $j \in J = \{1, \ldots, m\} - \{s_1, \ldots, s_n\}$ die Fundamentallösungen des linearen Gleichungssystems sind.

Beweis: Folgt aus Proposition II.5.3, Proposition II.5.8 und Proposition II.5.11.□

Lemma II.5.12 (Gauß-Algorithmus): *Sei $A \in \mathbb{R}^{n\times m}$. Dann gibt es $C \in \mathrm{Gl}_n(\mathbb{R})$, sodass CA Treppenform hat.*

Beweis: Wir beweisen die Aussage per Induktion nach n (der Anzahl der Zeilen der Matrix A). Für diesen Beweis notieren wir $A_{i,j}(\alpha)$ für die Additionsmatrix $A^{\alpha}_{i,j}$.

Zum Induktionsanfang sei $n = 1$, d. h. A besteht nur aus einer Zeile. Ist A die Nullzeile, so sind wir fertig, denn die Nullzeile hat Treppenform und $C = (1)$ leistet das Gewünschte.

Ist A nicht die Nullzeile, setze $s_1 := \min\{j \in \{1, \ldots, m\} \mid a_{i,j} \neq 0\}$ und wähle $C = (\frac{1}{a_{1,s_1}})$, dann ist CA in Treppenform.

Die Aussage gelte für die natürliche Zahl $n-1$. Wir wollen zeigen, dass die Aussage dann auch für n gilt. Ist $A = \mathbf{0}$, dann sind wir fertig. Ist $A \neq \mathbf{0}$, setze $s_1 := \min\{j \in \{1, \ldots, m\} \mid \exists\, i : a_{1,j} \neq 0\}$ (dies ist die Nummer der „linkesten“ Spalte, die keine Nullspalte ist) und $i_0 := \min\{i \in \{1, \ldots, n\} \mid a_{i,s_1} \neq 0\}$, d. h. A ist von der folgenden Form:

$$A = \begin{pmatrix} 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & a_{i_0,s_1} & \vdots & & \vdots \\ \vdots & & \vdots & * & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & * & \cdots & * \end{pmatrix}$$

Setze $A_1 := V_{1,i_0} \cdot (\prod_{i \neq i_0} A_{i,i_0}(-a_{i,s_1})) \cdot \operatorname{diag}(1, \ldots, 1, a_{i_0,s_1}^{-1}, 1, \ldots, 1) \cdot A$. Dann hat A_1 folgende Form:

$$A_1 = \left(\begin{array}{cccc|c} 0 & \cdots & 0 & 1 & z \\ \hline 0 & \cdots & 0 & 0 & \\ \vdots & & \vdots & \vdots & \tilde{A} \\ 0 & \cdots & 0 & 0 & \end{array}\right)$$

Nach Induktionsvoraussetzung gibt es eine Matrix $\tilde{C} \in \mathrm{Gl}_{n-1}(\mathbb{R})$ mit $\tilde{C}\tilde{A} = \tilde{T}$ mit einer Matrix in Treppenform $\tilde{T}$ und Rang $\tilde{r}$ und Spaltenindizes $\tilde{s}_1, \ldots, \tilde{s}_{\tilde{r}}$. Nach (Proposition 4.14) ist die Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & \tilde{C} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

invertierbar, also in $\mathrm{Gl}_n(\mathbb{R})$. Nach (Lemma 4.13) ist schließlich

$$\begin{pmatrix} 1 & 0 \\ 0 & \tilde{C} \end{pmatrix} A_1 = \begin{pmatrix} 0 & \cdots & 0 & 1 & z \\ 0 & \cdots & \cdots & 0 & \\ \vdots & & & \vdots & \tilde{T} \\ 0 & \cdots & \cdots & 0 & \end{pmatrix} =: A_2$$

Was wir jetzt noch erreichen müssen, ist, dass $z_{\tilde{s}_j} = 0$ für $i = 1, \ldots, \tilde{r}$. Die Matrix

$$T := \prod_{i=2}^{\tilde{r}+1} A_{1,i}(-z_{\tilde{s}_{i-1}+s_1}) A_2$$

hat jetzt Treppenform. Unsere Matrix C ist

$$C := \prod_{i=2}^{\tilde{r}+1} A_{1,i}(-z_{\tilde{s}_{i-1}+s_1}) \cdot \begin{pmatrix} 1 & 0 \\ 0 & \tilde{C} \end{pmatrix} \cdot V_{1,i_0} \cdot \prod_{i \neq i_0} A_{i,i_0}(-a_{i,s_1}) \cdot \operatorname{diag}(1, \ldots, 1, a_{i_0,s_1}^{-1}, 1, \ldots, 1)$$

und leistet das Gewünschte. □

Bemerkung II.5.13: Proposition II.5.12 gibt einen Algorithmus an, wie man ein lineares Gleichungssystem in Treppenform bringen kann.

Beispiel II.5.14: Wir suchen die Lösungsmenge des folgenden linearen Gleichungsystems:

$$\left(\begin{array}{ccccc|c} 0&0&0&0&0&0\\ 0&2&4&2&0&6\\ 0&3&6&-3&-6&-3\\ 0&0&0&1&1&2 \end{array}\right) \xrightarrow{\frac{1}{2}\cdot \mathrm{II}} \left(\begin{array}{ccccc|c} 0&0&0&0&0&0\\ 0&1&2&1&0&3\\ 0&3&6&-3&-6&-3\\ 0&0&0&1&1&2 \end{array}\right)$$

$$\xrightarrow{\mathrm{I}\leftrightarrow\mathrm{II}} \left(\begin{array}{ccccc|c} 0&1&2&1&0&3\\ 0&0&0&0&0&0\\ 0&3&6&-3&-6&-3\\ 0&0&0&1&1&2 \end{array}\right) \xrightarrow{\mathrm{III}-3\cdot\mathrm{I}} \left(\begin{array}{ccccc|c} 0&1&2&1&0&3\\ 0&0&0&0&0&0\\ 0&0&0&-6&-6&-12\\ 0&0&0&1&1&2 \end{array}\right)$$

$$\xrightarrow{-\frac{1}{6}\cdot\mathrm{III}} \left(\begin{array}{ccccc|c} 0&1&2&1&0&3\\ 0&0&0&0&0&0\\ 0&0&0&1&1&2\\ 0&0&0&1&1&2 \end{array}\right) \xrightarrow{\mathrm{II}\leftrightarrow\mathrm{III}} \left(\begin{array}{ccccc|c} 0&1&2&1&0&3\\ 0&0&0&1&1&2\\ 0&0&0&0&0&0\\ 0&0&0&1&1&2 \end{array}\right)$$

$$\xrightarrow{\mathrm{IV}-\mathrm{II}} \left(\begin{array}{ccccc|c} 0&1&2&1&0&3\\ 0&0&0&1&1&2\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{array}\right) \xrightarrow{\mathrm{I}-\mathrm{II}} \left(\begin{array}{ccccc|c} 0&1&2&0&-1&1\\ 0&0&0&1&1&2\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{array}\right)$$

die Lösungsmenge ist also

$$\mathbb{L} = \begin{pmatrix} 0\\1\\0\\2\\0 \end{pmatrix} + \left\{ \lambda_1 \begin{pmatrix} 1\\0\\0\\0\\0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0\\-2\\1\\0\\0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0\\1\\0\\-1\\1 \end{pmatrix} : \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}.$$

Lemma II.5.15 (Multiplikation mit invertierbaren Matrizen): *Sei $A \in \mathrm{Gl}_n(\mathbb{R})$. Dann gilt für $v \in \mathbb{R}^n$: Ist $v \neq \mathbf{0}$, dann ist $Av \neq \mathbf{0}$.*

Beweis: Ist $Av = \mathbf{0}$, so ist $v = A^{-1}Av = A^{-1}\mathbf{0} = \mathbf{0}$. □

Lemma II.5.16 (Eindeutigkeit der Treppenform): *Seien $T, T' \in \mathbb{R}^{n\times m}$ Matrizen in Treppenform. Ist $D \in \mathrm{Gl}_n(\mathbb{R})$ mit $T' = DT$, dann gilt $T' = T$.*

Beweis: Wir zeigen die Aussage via vollständiger Induktion.

Für den Induktionsanfang sei $n = 1$. T und T' bestehen also aus einer Zeile in $\mathbb{R}^{1\times m}$ und es gibt $d \in \mathbb{R} - \{0\}$ mit $T' = dT$. Somit gilt: T ist genau dann die Nullzeile, wenn T' die Nullzeile ist. Andernfalls muss der erste Eintrag von T und T' an der gleichen Stelle stehen. Da T und T' in Treppenform sind, ist

dieser Eintrag jeweils 1, d. h. d muss dann schon 1 sein und T und T' sind gleich.

Die Aussage gelte für die natürliche Zahl n. Wir wollen zeigen, dass die Aussage dann auch für $n+1$ gilt. Die Matrix T habe Rang r und Spaltenindizes $s_1, \ldots, s_r$ und T' habe den Rang r' und Spaltenindizes $s'_1, \ldots, s'_r$, sie haben also die Form

$$T' = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}, \qquad T = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}$$

Wir schreiben $T' = (t'^{(1)}, t'^{(2)}, \ldots, t'^{(m)})$, $T = (t^{(1)}, \ldots, t^{(m)})$, wobei $t'^{(1)}, \ldots, t'^{(m)}$, $t^{(1)}, \ldots, t^{(m)} \in \mathbb{R}^n$ die Spaltenvektoren von T' bzw. T sind. Es gelten

(a) $t^{(j)} = \mathbf{0}$ für $1 \leq j \leq s_1 - 1$ und $t'^{(j)} = \mathbf{0}$ für $1 \leq j \leq s'_1 - 1$,

(b) $t^{(s_k)} = e_k$ für $1 \leq k \leq r$ und $t'^{(s'_k)} = e_k$ für $1 \leq k \leq r'$,

(c) $t_i^{(j)} = T(i,j) = 0$ für $i > r$ und $t_i'^{(j)} = T'(i,j) = 0$ für $i > r'$,

(d) $t'^{(j)} = Dt^{(j)}$ für $1 \leq j \leq m$.

Damit erhalten wir die folgenden Eigenschaften:

(i) Für $1 \leq j \leq s_1 - 1$ gilt $t'^{(j)} = Dt^{(j)} = D\mathbf{0} = \mathbf{0}$,

(ii) Nach Lemma II.5.15 ist $t'^{(s_1)} = Dt^{(s_1)} = De_1 \neq \mathbf{0}$, d. h. es ist $s'_1 = s_1$ und damit $e_1 = t'^{(s_1)} = Dt^{(s_1)} = De_1$.

Es gilt also

$$D = \left(\begin{array}{c|c} 1 & z \\ \hline 0 & \\ \vdots & \hat{D} \\ 0 & \end{array}\right)$$

mit $\hat{D} \in \mathrm{Gl}_{n-1}(\mathbb{R})$ nach (Proposition 4.14) und $z \in \mathbb{R}^{1\times(n-1)}$. Weiterhin lassen sich T und T' schreiben als

$$T = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & & \hat{T} & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}, \qquad T' = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & & \hat{L}' & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix},$$

wobei $\hat{T}' = \hat{D}\hat{T}$. Nach Induktionsvoraussetzung gilt $\hat{T} = T'$. Somit folgt die Behauptung. □

Satz 3 (Satz über die Gauß-Normalform): *Für jede Matrix $A \in \mathbb{R}^{n\times m}$ gibt es genau eine Matrix $T \in \mathbb{R}^{n\times m}$ in Treppenform, sodass gilt: Es gibt $C \in \mathrm{Gl}_n(\mathbb{R})$ mit $T = CA$.*

Beweis: Die Existenz einer solchen Matrix $T \in \mathrm{R}^{n\times m}$ folgt aus Proposition II.5.12.

Zur Eindeutigkeit: Seien $T_1, T_2 \in \mathrm{R}^{n\times m}$ in Treppenform und $C_1, C_2 \in \mathrm{Gl}_n(\mathbb{R})$ mit $T_1 = C_1A$ und $T_2 = C_2A$. Dann ist $A = C^{-1}T_1$ und $T_2 = C_2C_1^{-1}T_1$. Da $C_2C_1^{-1} \in \mathrm{Gl}_n(\mathbb{R})$ nach (Proposition 4.15), können wir mit (Lemma 5.17) schließen, dass $T_1 = T_2$. □

In der Situation von Satz 3 bezeichnen wir T auch als *Gauß-Normalform von A*.

Definition II.5.17: Sei $A \in \mathrm{R}^{n\times m}$. Der *Rang* $\mathrm{Rang}(A)$ ist der Rang r der Gauß-Normalform T von A.

Bemerkung II.5.18: Für alle $D \in \mathrm{Gl}_n(\mathbb{R})$ und $A \in \mathbb{R}^{n\times m}$ gilt: Die Gauß-Normalformen von A und DA sind gleich, und damit $\mathrm{Rang}(DA) = \mathrm{Rang}(A)$.

Beweis: Sei $C \in \mathrm{Gl}_n(\mathbb{R})$, sodass $CA = T$ Treppenform hat. Aber dann gilt $CD^{-1}(DA) = T$, sodass A und DA dieselbe Gauß-Normalform T haben. □

Fazit II.5.19: Wir erhalten folgendes Lösungsverfahren für ein lineares Gleichungssystem $Ax = b$ mit $A \in \mathrm{R}^{n\times m}$ und $b \in \mathrm{R}^n$:

(i) Bestimme $C \in \mathrm{Gl}_n(\mathbb{R})$, sodass $CA = T$ Treppenform hat (das können wir nach dem Gauß-Algorithmus Proposition II.5.12),

(ii) Berechne die Lösungsmenge $\mathbb{L}$ von $Tx = Cb$ (nach Satz 2 ist dann $\mathbb{L}$ auch die Lösungsmenge von $Ax = b$ (nach Proposition 5.4).

Beispiel II.5.20: Wir betrachten das folgende lineare Gleichungssystem:

$$T = \left(\begin{array}{cccccc|c} 0 & 1 & 0 & \alpha & 0 & \beta & b_1 \\ 0 & 0 & 1 & \gamma & 0 & \delta & b_2 \\ 0 & 0 & 0 & 0 & 1 & \varepsilon & b_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_5 \end{array}\right)$$

Dann gelten:

(i) $\mathrm{Rang}(T|b) = 3$ genau dann, wenn $b_4 = b_5 = 0$ und $(T|b)$ ist genau dann lösbar, wenn $b_4 = b_5 = 0$. Ist $b_4 \neq 0$ oder $b_5 \neq 0$, so ist $\mathrm{Rang}(T|b) = 4$.

(ii) Seien $b_4 = b_5 = 0$. Wann ist $(T|b)$ eindeutig lösbar? Genau dann, wenn es keine Fundamentallösungen gibt, d. h. $m - r = 0$, wobei m die Anzahl der Spalten von T ist.

Korollar II.5.21 (aus Satz 3): *Seien $A \in \mathbb{R}^{n\times m}$ und $b \in \mathbb{R}^n$. Wir betrachten das lineare Gleichungssystem $Ax = b$.*

(i) *$Ax = b$ ist genau dann lösbar, wenn* $\mathrm{Rang}(A) = \mathrm{Rang}(A|b)$,

(ii) *Ist das lineare Gleichungssystem lösbar, dann ist die Lösung eindeutig genau dann, wenn* $\mathrm{Rang}(A) = m$,

(iii) *$Ax = c$ ist genau dann für jedes $c \in \mathbb{R}^n$ lösbar, wenn* $\mathrm{Rang}(A) = n$.

Beweis: Wähle $C \in \mathrm{Gl}_n(\mathbb{R})$ sodass $CA = T$ Treppenform hat. Es gilt also $\mathrm{Rang}(A) = \mathrm{Rang}(T) = r$.

(i) $(A|b)$ ist genau dann lösbar, wenn $(CA|Cb)$ lösbar ist und $(CA|Cb)$ ist lösbar genau dann, wenn $(T|Cb)$ lösbar ist. $(T|Cb)$ ist lösbar genau dann, wenn $\mathrm{Rang}(T|Cb) = \mathrm{Rang}(T)$. Es gilt $\mathrm{Rang}(T|Cb) = \mathrm{Rang}(T)$ genau dann, wenn $\mathrm{Rang}(A) = \mathrm{Rang}(A|b)$, denn $\mathrm{Rang}(A) = \mathrm{Rang}(T)$, sodass nach (Bemerkung 5.20) $\mathrm{Rang}(A|b) = \mathrm{Rang}(T|Cb)$.

(ii) Sei $(A|b)$ lösbar; dann ist auch $(T|Cb)$ lösbar. $(A|b)$ ist eindeutig lösbar genau dann, wenn $(T|Cb)$ eindeutig lösbar ist, d. h. genau dann, wenn $m = r$ ist.

(iii) Diese Aussage folgt aus der Tatsache, dass $Tx = c$ für alle $c \in \mathbb{R}^n$ genau dann lösbar ist, wenn T keine Nullzeilen hat, d. h. wenn $n = r$. □

Korollar II.5.22: *Die folgenden Aussagen sind äquivalent für $A \in \mathbb{R}^{n\times n}$:*

(i) *A ist invertierbar,*

(ii) $\mathrm{Rang}(A) = n$,

(iii) *Es gibt $S \in \mathbb{R}^{n\times n}$ mit $AS = I_n$.*

Beweis: „(i) ⇒ (ii)“: Ist A invertierbar, so gibt es B in $\mathrm{Gl}_n(\mathbb{R})$ mit $BA = I_n$, d. h. I_n ist die Gauß-Normalform von A und somit $\mathrm{Rang}(A) = \mathrm{Rang}(I_n) = n$.

„(ii) ⇒ (i)“: Wir wollen verwende, dass I_n die einzige Treppenform in $\mathbb{R}^{n\times n}$ von Rang n ist. Wir haben also: Ist $\mathrm{Rang}(A) = n$, dann gibt es $C \in \mathrm{Gl}_n(\mathbb{R})$ mit $CA = I_n$, d. h. $A = C^{-1}CA = C^{-1}$, also ist A invertierbar.

„(i) ⇒ (iii)“ ist klar. „(iii) ⇒ (i)“: Gilt $AS = I_n$, dann ist $ASc = c$ für alle $c \in \mathrm{R}^n$, d. h. das lineare Gleichunssystem $Ax = c$ hat für jedes $c \in \mathrm{R}^n$ eine Lösung, also muss $\mathrm{Rang}(A) = n$ gelten nach (Korollar 5.23). □

Bemerkung II.5.23: Der Gauß-Algorithmus liefert ein Verfahren zur Bestimmung, ob eine Matrix $A \in \mathrm{R}^{n\times n}$ invertierbar ist.

Beispiel II.5.24: Prüfen Sie, ob die Matrix

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

invertierbar ist.

Lösung Wir wollen simultan die linearen Gleichungssysteme $(A|e_1)$, $(A|e_2)$ und $(A|e_3)$ lösen, d. h.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$

Definition II.5.25: Sei $A \in \mathbb{R}^{n\times m}$. Die Abbildung

$$\Phi_A \colon \mathbb{R}^m \longrightarrow \mathbb{R}^n, \qquad x \longmapsto Ax$$

erfüllt folgende Eigenschaften:

(i) Für alle $x, y \in \mathbb{R}^m$ gilt

$$\Phi_A(x+y) = A(x+y) = Ax + Ay = \Phi_A(x) + \Phi_A(y),$$

(ii) Für alle $\lambda \in \mathbb{R}$ und $x \in \mathbb{R}^m$ gilt

$$\Phi_A(\lambda x) = A(\lambda x) = \lambda Ax = \lambda \Phi_A(x).$$

Eine solche Abbildung nennt man *lineare Abbildung* oder *Vektorraumhomomorphismus*, dazu mehr später.

Bemerkung II.5.26: Aus Korollar II.5.21 folgt:

(i) Φ_A ist surjektiv genau dann, wenn $\mathrm{Rang}(A) = n$,

(ii) Φ_A ist injektiv genau dann, wenn $\mathrm{Rang}(A) = m$.

Kapitel III.

Strukturmathematik: Gruppen, Ringe, Körper

1. Gruppen

Definition III.1.1 (Verknüpfung): Es sei M eine Menge.

(i) Eine *Verknüpfung auf M* ist eine Abbildung

$$*\colon M \times M \longrightarrow M.$$

Wir schreiben für gewöhnlich $m_1 * m_2 := *(m_1, m_2)$.

(ii) Eine Verknüpfung $*$ auf M heißt

- *assoziativ*, falls für alle $a, b, c \in M$ gilt: $(a * b) * c = a * (b * c)$. Wir schreiben in diesem Fall einfach $a * b * c$.
- *kommutativ*, falls für alle $a, b \in M$ gilt: $a * b = b * a$.

Definition III.1.2 (Gruppe): Sei G eine Menge und $*$ eine Verknüpfung auf G. Das Paar $(G, *)$ heißt *Gruppe*, falls die folgenden Bedingungen erfüllt sind:

(i) $*$ ist assoziativ,

(ii) Es gibt ein $e \in G$, sodass für alle $g \in G$ gilt: $e * g = g = g * e$. In diesem Fall ist e eindeutig, denn erfüllt hat e' die gleiche Eigenschaft wie e, so gilt $e' = e * e' = e$. Das eindeutige Element e heißt *neutrales Element*.

(iii) Für alle $g \in G$ gibt es $h \in G$, sodass $g * h = e = h * g$. In diesem Fall ist h eindeutig, denn gibt es h_1 und h_2, die beide die Eigenschaft von h haben, dann gilt $h_1 = h_1 * e = h_1 * (g * h_2) = (h_1 * g) * h_2 = e * h_2 = h_2$. Das eindeutige Element h heißt dann das *Inverse* zu g und wird als g^{-1} notiert.

Definition III.1.3 (Are you Abel?): Eine Gruppe $(G, *)$ heißt *abelsch*, falls $*$ kommutativ ist.

Beispiel III.1.4: Das folgende sind alles Gruppen:

(i) $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, (abelsch)

(ii) $\mathbb{R}$-Vektorräume mit Addition, (abelsch)

(iii) $(\mathbb{Q} - \{0\}, \cdot)$, $(\mathbb{R} - \{0\}, \cdot)$, (abelsch)

(iv) $\mathrm{Gl}_n(\mathbb{R})$ mit der Matrizenmultiplikation, (nicht abelsch)

(v) Sei M eine Menge. Dann bildet

$$\mathrm{Perm}(M) := \{f \colon M \to M \mid f \text{ bijektiv}\}$$

mit der Komposition von Abbildungen eine Gruppe. Ist speziell $M = \{1, \ldots, n\}$, so schreibt man oft $S_n := \mathrm{Perm}(M)$.

Bemerkung III.1.5 (Gruppe der Kongruenzklassen): Sei $n \in \mathbb{N}$, dann ist „$\equiv_n$“ eine Äquivalenzrelation (siehe Proposition I.5.9). Setze

$$\mathbb{Z}/n\mathbb{Z} := \{[a] \mid a \in \mathbb{Z}\},$$

das heißt $\mathbb{Z}/n\mathbb{Z}$ ist die Menge der Äquivalenzklassen bezüglich „$\equiv_n$“. Es gilt also $\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \ldots, [n-1]\}$. Auf der Menge der Äquivalenzklassen erklären wir jetzt eine Verknüpfung durch

$$+ \colon \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, \qquad [a] + [b] := [a + b].$$

Diese Verknüpfung macht $(\mathbb{Z}/n\mathbb{Z}, +)$ zu einer abelschen Gruppe.

Beweis: (i) Die Verknüpfung „+“ ist wohldefiniert, d. h. hängt nicht von den gewählten Repräsentanten. Sind nämlich $a', b' \in \mathbb{Z}$ mit $a' \equiv a \pmod{n}$ und $b' \equiv b \pmod{n}$, dann gilt $n \mid (a' - a)$ und $n \mid (b' - b)$, also

$$n \mid (a' - a + b' - b) \Leftrightarrow n \mid [(a' + b' - (a + b)],$$

d. h. $a' + b' \equiv a + b \pmod{n}$ und deshalb $[a' + b'] = [a + b]$.

(ii) Die Verknüpfung ist assoziativ, da + auf $\mathbb{Z}$ assoziativ ist. $[0]$ ist das neutrale Element und für $[a]$ ist $[-a]$ das inverse Element. Da + außerdem kommutativ ist, ist $(\mathbb{Z}/n\mathbb{Z}, +)$ eine abelsche Gruppe. □

Ab jetzt schreiben wir $\overline{a}$ für die Äquivalenzklasse $[a] \in \mathbb{Z}/n\mathbb{Z}$.

Definition III.1.6 (Untergruppe): Sei $(G, *)$ eine Gruppe und H eine Teilmenge von G. H heißt *Untergruppe* von G, falls gilt:

(i) Das neutrale Element von G ist in H,

(ii) Für alle $h_1, h_2 \in H$ gilt $h_1 * h_2 \in H$, (Abgeschlossenheit unter „$*$“)

(iii) Für alle $h \in H$ gilt $h^{-1} \in H$. (Abgeschlossenheit unter Inversenbildung)

Bemerkung III.1.7: In der Situation von Definition III.1.6 gilt:

$$*\colon H \times H \longrightarrow H, \qquad (h_1, h_2) \longmapsto h_1 * h_2$$

ist eine Verknüpfung auf H und $(H, *)$ ist eine Gruppe in eigenem Recht.

Proposition III.1.8 (Untergruppenkriterium): *Es sei $(G, *)$ eine Gruppe und H eine Teilmenge von G. H ist Untergruppe von G genau dann, wenn folgende Aussagen gelten:*

(a) $H \neq \varnothing$,

(b) *Für alle $h_1, h_2 \in H$ ist $h_1 * h_2^{-1} \in H$.*

Beweis: Sei e das neutrale Element von G.

„$\Rightarrow$“: Da H eine Untergruppe von G ist, gilt $e \in H$, d. h. insbesondere $H \neq \varnothing$, also gilt (a). Seien $h_1, h_2 \in H$, dann ist nach (iii) $h_2^{-1} \in H$ und nach (ii) ist $h_1 * h_2^{-1} \in H$, d. h. (b) gilt.

„$\Leftarrow$“: Es sei $H \neq \varnothing$, d. h. es gibt $h \in H$. Nach (b) ist $e = h * h^{-1} \in H$, also gilt (i). Dann folgt aus (b): Für alle $h \in H$ ist $h^{-1} = eh^{-1} \in H$, also gilt (iii), und nun folgt: Für alle $h_1, h_2 \in H$ ist $h_2^{-1} \in H$, d. h. mit (b) ist $h_1 * h_2 \in H$, also gilt (ii). □

Bemerkung III.1.9: Sei $n \in \mathbb{Z}$. Dann ist $n\mathbb{Z} := \{nk \mid k \in \mathbb{Z}\}$ eine Untergruppe von $(\mathbb{Z}, +)$.

Beweis: Da $0 \in \mathbb{Z}$ und $0n = 0$ ist $n\mathbb{Z} \neq \varnothing$. Sind $nk_1, nk_2 \in n\mathbb{Z}$ $(k_1, k_2 \in \mathbb{Z})$, dann ist

$$nk_1 - nk_2 = n(k_1 - k_2) \in n\mathbb{Z},$$

d. h. nach Proposition III.1.8 ist $n\mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$. □

Ab jetzt sei $(G, *)$ stets eine Gruppe mit neutralem Element e.

Bemerkung III.1.10 (Schnitt von Untergruppen): Sei $\varnothing \neq I$ eine Menge, und zu jedem $i \in I$ sei G_i eine Untergruppe von $(G, *)$. Dann ist auch $\bigcap_{i \in I} G_i$ eine Untergruppe von $(G, *)$.

Beweis: (a) Für alle $i \in I$ ist $e \in G_i$, d. h. $e \in \bigcap_{i\in I} G_i$; insbesondere ist $\bigcap_{i\in I} G_i$ nicht leer.

(b) Seien $g_1, g_2 \in \bigcap_{i\in I} G_i$. Dann gilt für alle $i \in I$, dass $g_1, g_2 \in G_i$, und da jedes G_i eine Untergruppe von G ist, gilt für jedes $i \in I$, dass $g_1 * g_2^{-1} \in G_i$ ist, es ist also $g_1 * g_2^{-1} \in \bigcap_{i\in I} G_i$.

Nach dem Untergruppenkriterium Proposition III.1.8 ist $\bigcap_{i\in I} G_i$ damit eine Untergruppe von $(G, *)$. □

Definition III.1.11 (Erzeugte Gruppe, zyklische Gruppe):

(i) Seien $M \subseteq G$ und

$$I := \{H \subseteq G \mid H \text{ ist Untergruppe von } (G, *) \text{ und } M \subseteq H\}.$$

Definiere

$$\langle M\rangle := \bigcap_{H\in I} H.$$

Dann ist $\langle M\rangle$ nach Bemerkung III.1.10 eine Untergruppe von $(G, *)$. $\langle M\rangle$ heißt das *Erzeugnis von M* oder die *von M erzeugte Untergruppe von* $(G, *)$. Offensichtlich ist $\langle M\rangle$ die bezüglich Inklusion kleinste Untergruppe von $(G, *)$, die M enthält.

(ii) G heißt *zyklisch*, falls es $g \in G$ gibt mit $G = \langle\{g\}\rangle$. Wir schreiben oft $\langle g\rangle$ für $\langle\{g\}\rangle$.

Beispiel III.1.12: Es sei $(G, *) = (\mathbb{Z}/10\mathbb{Z}, +)$. Dann sind z. B. $\langle[1]\rangle = \mathbb{Z}/10\mathbb{Z}$, $\langle[2]\rangle = \{[2], [4], [6], [8], [10]\}$ und $\langle[3]\rangle = \mathbb{Z}/10\mathbb{Z}$.

Proposition III.1.13: *Für $g \in G$ ist $\langle g\rangle = \{g^k \mid k \in \mathbb{Z}\}$. Hierbei ist*

$$g^k = \begin{cases} g * g * \cdots * g & \text{falls } k \in \mathbb{N}, \\ e_G, & \text{falls } k = 0, \\ (g^{-k})^{-1}, & \text{falls } k < 0. \end{cases}$$

Definition III.1.14: (i) Die *Ordnung* der Gruppe G ist die Anzahl der Elemente von G, wir schreiben dafür $\mathrm{ord}(G)$.

(ii) Für $g \in G$ definieren wir die *Ordnung von g* als die Ordnung von $\langle g\rangle$ und schreiben dafür $\mathrm{ord}(g)$.

Beispiel III.1.15: Sei $(G, *) = (\mathbb{Z}/10\mathbb{Z}, +)$. G ist zyklisch, denn $G = \langle[1]\rangle$. Es gelten $\mathrm{org}(G) = 10$, $\mathrm{ord}([1]) = 10$, $\mathrm{ord}([2]) = 5$, $\mathrm{ord}([5]) = 2$.

Lemma III.1.16: *Sei H eine Untergruppe von $(G, *)$. Wir definieren die Relation $\sim$ auf G durch $g_1 \sim g_2 :\Leftrightarrow g_1 * g_2^{-1} \in H$. Dann gelten:*

(i) *$\sim$ ist eine Äquivalenzrelation auf G mit den Äquivalenzklassen $[g] = H * g$, wobei $H * g := \{h * g \mid h \in H\}$.*

(ii) *Für $g \in G$ ist die Abbildung*

$$F_g \colon [e_G] = H \longrightarrow [g] = H * g, \qquad h \longmapsto h * g$$

eine Bijektion.

Beweis: (i) Nachrechnen, sehr ähnlich zu (Bemerkung I.5.8).

(ii) F_g ist surjektiv, denn für alle Elemente $h * g \in H * g$ (d. h. $h \in H$) gilt: $F_g(h) = h * g$. Außerdem ist F_g injektiv, denn für alle $h_1, h_2 \in H$ mit $F_g(h_1) = F_g(h_2)$ gilt $h_1 * g = h_2 * g$, d. h. durch Verknüpfung mit $*$ mit g^{-1} von rechts gilt $h_1 = h_2$ □

Definition III.1.17: In der Situation von Lemma III.1.16 heißt $H * g$ die *Rechts-Nebenklasse* von H in G bezüglich g.

Satz 4 (Lagrange): *Es sei $(G, *)$ eine endliche Gruppe, d. h.* $\mathrm{ord}(G) < \infty$. *Für jede Untergruppe H von G gilt:* $\mathrm{ord}(H)$ *teilt* $\mathrm{ord}(G)$.

Beweis: Betrachte die Äquivalenzrelation aus Lemma III.1.16. Teil (ii) von Lemma III.1.16 garantiert, dass alle Nebenklassen die gleiche Anzahl an Elementen haben, d. h. $\#(g * H) = \#(H)$ für alle $g \in G$. Nach Satz 1 ist G die disjunkte Vereinigung seiner Nebenklassen, also

$$G = \bigcup_{g \in G} H * g$$

mit $H * g_1 = H * g_2$ oder $H * g_1 \cap H * g_2 = \varnothing$. Sei k die Anzahl der Nebenklassen, d. h. wir haben k Elemente $g_1, \ldots, g_k$ mit $G = \biguplus_{i=1}^{k} H * g_i$. Dann können wir schreiben

$$\mathrm{ord}(G) = \sum_{i=1}^{k} \#(H * g_i) = \sum_{i=1}^{k} \#(H) = k\#(H),$$

die Ordnung von H teilt also die Ordnung von G. □

Korollar III.1.18: *Sei $(G, *)$ eine endliche Gruppe, deren Ordnung eine Primzahl p ist. Dann ist G zyklisch.*

Beweis: Wähle $e_G \neq g \in G$ und setze $H = \langle g \rangle$. Dann ist $\mathrm{ord}(H) \geq 2$ und nach Satz 4 teilt $\mathrm{ord}(H)$ die Ordnung von G. Da die Ordnung von G die Primzahl p ist, muss $\mathrm{ord}(H) = p = \mathrm{ord}(G)$ gelten, also ist $G = H = \langle g \rangle$ und G ist zyklisch. □

Bemerkung III.1.19: Seien $g_1, g_2 \in G$. Dann ist $(g_1 * g_2)^{-1} = g_2^{-1} * g_1^{-1}$.

Beweis: Es gelten $g_1 * g_2 * (g_2^{-1} * g_1^{-1}) = g_1 * g_2 * g_2^{-1} * g_1^{-1} = g_1 * e_G * g_1^{-1} = e_G$ und $(g_2^{-1} * g_1^{-1}) * g_1 * g_2 = e_G$, d. h. $g_2^{-1} * g_1^{-1}$ ist das Inverse zu $g_1 * g_2$. □

2. Gruppenhomomorphismen

Definition III.2.1 (Gruppenhomomorphismus): Es seien $(G, *)$ und $(H, \bullet)$ zwei Gruppen.

(i) Eine Abbildung $\varphi\colon G \to H$ heißt *(Gruppen-)Homomorphismus* von $(G, *)$ nach $(H, \bullet)$, falls für alle $g_1, g_2 \in G$ gilt:

$$\varphi(g_1 * g_2) = \varphi(g_1) \bullet \varphi(g_2).$$

(ii) Wir bezeichnen

$$\mathrm{Hom}(G, H) := \{\varphi\colon G \to H \mid \varphi \text{ ist Homomorphismus von } (G, *) \text{ nach } (H, \bullet)\}$$

Im Folgenden notieren wir mit $\varphi\colon (G, *) \to (H, \bullet)$ Gruppenhomomorphismen von $(G, *)$ nach $(H, \bullet)$.

Beispiel III.2.2 (Erste Beispiele für Gruppenhomomorphismen): Die folgenden Abbildungen sind Gruppenhomomorphismen:

(i) $\varphi_A\colon (\mathbb{R}^m, +) \to (\mathbb{R}^n, +)$, $x \mapsto Ax$, wobei $A \in \mathbb{R}^{n \times m}$,

(ii) $\varphi\colon (\mathbb{R}, +) \to (\mathbb{R} - \{0\}, \cdot)$, $x \mapsto \exp(x)$, denn

$$\varphi(x_1 + x_2) = \exp(x_1 + x_2) = \exp(x_1) \cdot \exp(x_2) = \varphi(x_1) \cdot \varphi(x_2).$$

(iii) Für eine beliebige Gruppe $(G, *)$ und $g \in G$ ist

$$\varphi_g\colon (\mathbb{Z}, +) \longrightarrow (G, *), \qquad k \longmapsto g^k$$

ein Gruppenhomomorphismus.

(iv) Sei $(G, *) = (S_n, \circ)$. Dann ist

$$(S_n, \circ) \longrightarrow (\mathrm{Gl}_n(\mathbb{R}), \cdot), \qquad \sigma \longmapsto A_{\sigma^{-1}} \quad \text{mit } A_\sigma(i,j) = \begin{cases} 1, & \text{falls } j = \sigma(i), \\ 0, & \text{sonst.} \end{cases}$$

ein Gruppenhomomorphismus.

(v) Für beliebige Gruppen $(G_1, *)$ und $(G_2, \bullet)$ ist

$$\varphi_{1,2} \colon (G_1, *) \longrightarrow (G_2, \bullet), \qquad g \longmapsto e_{G_2}$$

ein Gruppenhomomorphismus. φ heißt der *triviale Homomorphismus*.

Beweis: Dass die Abbildungen aus (i), (ii), (iii), (v) tatsächlich Homomorphismen sind, ist offensichtlich.

Zu (iv): Berechne zunächst $A_{\sigma_1} \cdot A_{\sigma_2}$ für $\sigma_1, \sigma_2 \in S_n$. Für $i, j \in \{1, \ldots, n\}$ gilt

$$(A_{\sigma_1} \cdot A_{\sigma_2})(i,j) = \sum_{k=1}^{n} A_{\sigma_1}(i,k) A_{\sigma_2}(k,j).$$

$A_{\sigma_1}(i,k) \cdot A_{\sigma_2}(k,j)$ ist 1 genau dann, wenn $k = \sigma_1(i)$ und $j = \sigma_2(k)$, sonst 0, d. h.

$$(A_{\sigma_1} A_{\sigma_2})(i,j) = \left\{ \begin{array}{ll} 1, & \text{falls } j = \sigma_2(\sigma_1(i)), \\ 0, & \text{sonst.} \end{array} \right\} = A_{\sigma_2 \circ \sigma_1}(i,j).$$

Damit ist $A_{\sigma_1} A_{\sigma_1^{-1}} = A_{\mathrm{id}} = I_n = A_{\sigma_1^{-1}} A_{\sigma_1}$, also ist $A_{\sigma_1} \in \mathrm{Gl}_n(\mathbb{R})$ für alle $\sigma_1 \in S_n$.

Außerdem ist für alle $\sigma_1, \sigma_2 \in S_n$

$$\varphi(\sigma_1 \circ \sigma_2) = A_{(\sigma_1 \circ \sigma_2)^{-1}} = A_{\sigma_2^{-1} \circ \sigma_1^{-1}} = A_{\sigma_1^{-1}} A_{\sigma_2^{-1}} = \varphi(\sigma_1)\varphi(\sigma_2),$$

φ ist also in der Tat ein Gruppenhomomorphismus. □

Ab jetzt seien $(G, *)$, $(H, \bullet)$ stets Gruppen und $\varphi \colon (G, *) \to (H, \bullet)$ ein Gruppenhomomorphismus.

Proposition III.2.3 (Erste Eigenschaften von Gruppenhomomorphismen): *Es gelten die folgenden Aussagen:*

(i) $\varphi(e_G) = e_H$,

(ii) *Für alle $g \in G$ ist* $\varphi(g^{-1}) = (\varphi(g))^{-1}$.

Beweis: (i) Es ist $\varphi(e_G) = \varphi(e_G * e_G) = \varphi(e_G) \bullet \varphi(e_G)$, d. h. wir können schreiben $e_H = (\varphi(e_G))^{-1} \bullet \varphi(e_G) = (\varphi(e_G))^{-1} \bullet \varphi(e_G) \bullet \varphi(e_G) = \varphi(e_G)$.

(ii) Es gilt $\varphi(g) \bullet \varphi(g^{-1}) = \varphi(g * g^{-1}) = \varphi(e_G) = e_H$, und völlig analog $\varphi(g^{-1}) \bullet \varphi(g) = e_H$, d. h. $\varphi(g^{-1})$ ist invers zu $\varphi(g)$. □

Proposition III.2.4 (Bild und Urbild): *Es gelten:*

(i) $\text{Bild}(\varphi) = \varphi(G) \subseteq H$ *ist eine Untergruppe von* $(H, \bullet)$,

(ii) *Ist* $H_1 \subseteq (H, \bullet)$ *eine Untergruppe, dann ist* $\varphi^{-1}(H_1) \subseteq (G, *)$ *eine Untergruppe.*

Beweis: Die Beweisstrategie für beide Teile ist die Verwendung des Untergruppenkriteriums.

(ii) Das neutrale Element von G ist enthalten in $\varphi^{-1}(H)$, denn wir wissen $\varphi(e_G) = e_H \in H_1$, also $\varphi^{-1}(H_1) \neq \varnothing$.

Seien $g_1, g_2 \in \varphi^{-1}(H_1)$. Dann ist $\varphi(g_1 * g_2^{-1}) = \varphi(g_1) \bullet \varphi(g_2^{-1}) = \varphi(g_1) \bullet \varphi(g_2)^{-1}$, und da H_1 eine Gruppe ist, gilt $\varphi(g_1 * g_2^{-1}) \in H_1$, also $g_1 * g_2^{-1} \in \varphi^{-1}(H_1)$. Damit ist $\varphi^{-1}(H_1)$ eine Untergruppe von $(H, \bullet)$.

(i) $\text{Bild}(\varphi)$ ist nicht leer, da $\text{Bild}(\varphi) \ni \varphi(e_G) = e_H$. Für zwei Elemente $\varphi(g_1), \varphi(g_2)$ mit $g_1, g_2 \in G$ gilt

$$\varphi(g_1) \bullet \varphi(g_2)^{-1} = \varphi(g_1) \bullet \varphi(g_2^{-1}) = \varphi(g_1 * g_2^{-1}) \in \text{Bild}(\varphi),$$

also ist $\text{Bild}(\varphi)$ eine Untergruppe von $(H, \bullet)$. □

Definition III.2.5 (Kern): Die Menge

$$\text{Ker}(\varphi) = \varphi^{-1}(\{e_H\}) = \{g \in G \mid \varphi(g) = e_H\}$$

heißt *Kern* von φ.

Bemerkung III.2.6 (Kern als Untergruppe): Der Kern von φ ist nach Proposition III.2.4 eine Untergruppe von $(G, *)$.

Proposition III.2.7 (Kern vs. Injektivität): φ *ist injektiv genau dann, wenn* $\text{Kern}(\varphi) = \{e_G\}$.

Beweis: „⇒“: Klar per Definition von Injektivität und Proposition III.2.3 (i).

„⇐“: Seien $g_1, g_2 \in G$ mit $\varphi(g_1) = \varphi(g_2)$. Dann ist

$$\varphi(g_1 * g_2^{-1}) = \varphi(g_1) \bullet \varphi(g_2)^{-1} = \varphi(g_1) \bullet \varphi(g_1)^{-1} = e_H,$$

d. h. $g_1 * g_2^{-1} \in \text{Kern}(\varphi) = \{e_G\}$, also ist $g_1 * g_2^{-1} = e_G$ und damit $g_1 = g_2$. □

Beispiel III.2.8 (Fortsetzung von Beispiel III.2.2): Für die Gruppenhomomorphismen aus Beispiel III.2.2 (in gleicher Reihenfolge) ergeben sich

(i) $\mathrm{Kern}(\varphi) = \mathbb{L}(A, \mathbf{0})$, $\mathrm{Bild}(\varphi) = \{b \in \mathbb{R}^n \mid \mathbb{L}(A, b) \neq \varnothing\}$,

(ii) $\mathrm{Kern}(\varphi) = \{0\}$, $\mathrm{Bild}(\varphi) = \mathbb{R}_{>0}$,

(iii) $\mathrm{Kern}(\varphi_g) = \mathrm{ord}(g)\mathbb{Z} = \{\mathrm{ord}(g)k \mid k \in \mathbb{Z}\}$, $\mathrm{Bild}(\varphi_g) = \langle g \rangle$,

(iv) $\mathrm{Kern}(\varphi) = \{\mathrm{id}\}$, $\mathrm{Bild}(\varphi)$ heißt *Gruppe der Permutationsmatrizen*,

(v) $\mathrm{Kern}(\varphi) = G_1$, $\mathrm{Bild}(\varphi) = \{e_{G_2}\}$.

Definition III.2.9 (Typen von Homomorphismen):

(i) φ heißt *Endomorphismus von Gruppen*, falls $(G, *) = (H, \bullet)$.

(ii) φ heißt *Isomorphismus*, falls es einen Gruppenhomomorphismus

$$\psi \colon (H, \bullet) \longrightarrow (G, *)$$

mit $\psi \circ \varphi = \mathrm{id}_G$ und $\varphi \circ \psi = \mathrm{id}_H$ gibt.

(iii) φ heißt *Automorphismus*, falls φ ein Endomorphismus und ein Isomorphismus ist.

Proposition III.2.10 (Charakterisierung von Gruppenisomorphismen): *φ ist ein Isomorphismus genau dann, wenn φ ein bijektiver Gruppenhomomorphismus ist.*

Beweis: „$\Rightarrow$“: Folgt aus der Definition und Proposition I.4.12.

„$\Leftarrow$“: Wir müssen zeigen, dass für einen bijektiven Gruppenhomomorphismus die Umkehrabbildung $\psi = \varphi^{-1}$ wieder ein Gruppenhomomorphismus ist. Seien $h_1, h_2 \in H$. Dann gilt

$$\psi(h_1 \bullet h_2) = \psi\Big(\varphi[\psi(h_1)] \bullet \varphi[\psi(h_2)]\Big) = \psi\Big(\varphi[\psi(h_1) * \psi(h_2)]\Big) = \psi(h_1) * \psi(h_2)$$

wegen $\psi \circ \varphi = \mathrm{id}_G$ und $\varphi \circ \psi = \mathrm{id}_H$. □

Proposition III.2.11 (Kern als „Normalteiler“): *Für $N := \mathrm{Kern}(\varphi)$ gilt: Für alle $g \in G$, $h \in N$ ist $g * h * g^{-1} \in N$.*

Beweis: Wir rechnen nach:

$$\varphi(g * h * g^{-1}) = \varphi(g) \bullet \varphi(h) \bullet \varphi(g)^{-1} = \varphi(g) \bullet e_H \bullet \varphi(g)^{-1} = e_H. \qquad \square$$

Bemerkung III.2.12 (Verkettung von Gruppenhomomorphismen): (i) Es seien $(G_1, *)$, $(G_2, \bullet)$, $(G_3, \times)$ drei Gruppen. Ferner seien Gruppenhomomorphismen $\varphi_1 \colon (G_1, *) \to (G_2, \bullet)$ und $\varphi_2 \colon (G_2, \bullet) \to (G_3, \times)$ gegeben. Dann ist auch die Verknüpfung

$$\varphi_2 \circ \varphi_1 \colon G_1 \longrightarrow G_3$$

ein Gruppenhomomorphismus von $(G_1, *)$ nach $(G_3, \times)$.

(ii) $(\mathrm{Aut}(G), \circ)$ ist eine Untergruppe von $(\mathrm{Perm}(G), \circ)$.

Beweis: (i) Für $a, b \in G_1$ gilt

$$\begin{aligned}(\varphi_2 \circ \varphi_1)(a * b) &= \varphi_2[\varphi_1(a * b)] = \varphi_2[\varphi_1(a) \bullet \varphi_1(b)] \\ &= \varphi_2[\varphi_1(a)] \times \varphi_2[\varphi_1(b)] = (\varphi_2 \circ \varphi_1)(a) \times (\varphi_2 \circ \varphi_1)(b).\end{aligned}$$

(ii) Folgt aus Proposition I.4.12 zusammen mit der Tatsache, dass id zu $\mathrm{Aut}(G)$ gehört. □

Bemerkung III.2.13 (Nachtrag): Seien $(G, \bullet)$ eine Gruppe und g ein Element von G mit endlicher Ordnung $k_0 = \mathrm{ord}(g)$. Dann gilt:

(i) Das Erzeugnis $\langle g \rangle$ von g ist die Menge $\{1_G, g, g^2, \cdots, g^{k_0 - 1}\}$.

(ii) Die natürliche Zahl k_0 ist charakterisiert über $k_0 = \min\{k \in \mathbb{N} \mid g^k = 1_G\}$.

(iii) Es gilt genau dann $g^n = 1_G$, wenn n von k_0 geteilt wird. Um das einzusehen, kann man beispielsweise $n \pmod{k_0}$ betrachten.

3. Die symmetrische Gruppe

M sei stets eine Menge. In diesem Abschnitt möchten wir Permutationen der Menge M verstehen.

Erinnerung III.3.1 (Permutationsgruppe): Wie in Definition I.4.17 eingeführt bezeichnet $\mathrm{Perm}(M) := \{f \colon M \to M \mid f \text{ ist bijektiv}\}$, und wie in Beispiel III.1.4 definiert, ist $S_n := \mathrm{Perm}(\{1, \ldots, n\})$. Beides sind Gruppen mit der Verkettung von Abbildungen als Verknüpfungen.

Beispiel III.3.2 (Die S_3): Es sei $n = 3$. Wie in Definition I.4.17 festgehalten, ist $\#(S_3) = 3! = 6$, d. h. die oben aufgelisteten Abbildungen sind alle Elemente, die es gibt und $S_3 = \{\mathrm{id}, \tau_1, \tau_2, \tau_3, \zeta_1, \zeta_2\}$. Wir listen alle Permutationen in S_3 über ihre Wertetabelle auf:

	1	2	3
id	1	2	3
τ_1	1	3	2
τ_2	2	1	3
τ_3	3	2	1
ζ_1	2	3	1
ζ_2	3	1	2

Bemerkung III.3.3 (Ordnung der S_n): Wie bereits nach Definition I.4.17 angemerkt, ist $\#(S_n) = n!$.

Notation III.3.4 (via Wertetabelle): Wir notieren eine Permutation $\sigma \in S_n$ über ihre Wertetabelle. Wir schreiben

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}.$$

Beispiel III.3.5: Die Permutationen id, ξ_1 und τ_2 sehen in der neu eingeführten Notation folgendermaßen aus:

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \zeta_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Definition III.3.6 (Träger): Für $\sigma \in \mathrm{Perm}(M)$ heißt $\mathrm{Tr}(\sigma) := \{x \in M \mid \sigma(x) \neq x\}$ der *Träger* von σ. Zwei Permutationen σ_1, σ_2 aus $\mathrm{Perm}(M)$ mit $\mathrm{Tr}(\sigma_1) \cap \mathrm{Tr}(\sigma_2) = \varnothing$ heißen *disjunkt*.

Beispiel III.3.7: In Beispiel III.3.2 ist $\mathrm{Tr}(\xi_1) = \{1, 2, 3\}$, $\mathrm{Tr}(\tau_2) = \{1, 2\}$ und $\mathrm{Tr}(\mathrm{id}) = \varnothing$.

Bemerkung III.3.8 (Zum Vertauschen disjunkter Permutationen): Sind σ_1, σ_2 disjunkte Permutationen einer Menge M, dann ist $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$. Die Disjunktheit ist zwingend notwendig für dieses Vertauschen. Beispielsweise gilt für die Permutationen ζ_1, τ_2 aus Beispiel III.3.2

$$\zeta_1 \circ \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \tau_3 \neq \tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \tau_2 \circ \zeta_1.$$

Definition III.3.9 (Zyklen und Transpositionen):

(i) Seien $x_1, \ldots, x_k \in M$. Wir definieren die folgende Permutation ξ:

$$\xi(x) = \begin{cases} x_{i+1}, & \text{falls } x = x_i \text{ und } i \in \{1, \ldots, k-1\}, \\ x_1, & \text{falls } x = x_k, \\ x, & \text{falls } x \notin \{x_1, \ldots, x_k\}. \end{cases}$$

Solch eine Permutation heißt k-Zyklus. Wir schreiben $\xi = (x_1, \ldots, x_k)$. k heißt die *Länge von* ξ. Der Träger des k-Zyklus ξ ist $\mathrm{Tr}(\xi) = \{x_1, \ldots, x_k\}$.

(ii) $\sigma \in \mathrm{Perm}(M)$ heißt Zyklus, falls es $k \in \mathbb{N}$ gibt, sodass σ ein k-Zyklus ist.

(iii) $\sigma \in \mathrm{Perm}(M)$ ist eine *Transposition*, falls σ ein 2-Zyklus ist.

Beispiel III.3.10: (i) In Beispiel III.3.2 sind τ_1, τ_2, τ_3 Transpositionen und ξ_1, ξ_2 3-Zyklen.

(ii) Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 5 & 3 \end{pmatrix}$$

aus S_5 lässt sich schreiben als $\sigma = (12) \circ (345) = (345) \circ (12)$.

(iii) Der 4-Zyklus $\sigma = (2573)$ aus S_7 lässt sich als $(25) \circ (57) \circ (37)$, also als Komposition von Transpositionen, schreiben.

Bemerkung III.3.11: Für den k-Zyklus $\xi = (x_1, \ldots, x_k)$ gilt:

$$(x_1, \ldots, x_k) = (x_1, x_2) \circ (x_2, x_3) \circ \cdots \circ (x_{k-1}, x_k).$$

Beispiel III.3.12: Wir betrachten die Gruppe S_{12} und die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 1 & 12 & 8 & 2 & 7 & 4 & 11 & 10 & 9 & 3 & 5 & 6 \end{pmatrix}.$$

Zwar ist $\sigma = (1) \circ (2, 12, 6, 4) \circ (3, 8, 10) \circ (5, 7, 11) \circ (9)$ kein Zyklus, aber σ lässt sich als Verkettung disjunkter Zyklen schreiben.

Satz 5 (über die Zerlegung von Permutationen in Zyklen): *Sei M eine endliche Menge. Jede Permutation $\sigma \in \mathrm{Perm}(M)$ ist ein Produkt von disjunkten Zyklen. Das heißt es gibt disjunkte Zyklen $\xi_1, \ldots, \xi_N$ mit $\sigma = \xi_1 \circ \xi_2 \circ \cdots \circ \xi_N$. Es gilt dann $\mathrm{Tr}(\xi_i) \subseteq \mathrm{Tr}(\sigma)$. Das leere Produkt ist zugelassen, also $N = 0$ ist erlaubt (dieses gibt $\sigma = \mathrm{id}$).*

Beweis: Wir zeigen die Aussage via Induktion über die Anzahl der Elemente im Träger von σ. Induktionsanfang: Ist $\#(\mathrm{Tr}(\sigma)) = 0$, so ist σ ist die Identität und die Behauptung gilt.

Induktionsvoraussetzung: Die Behauptung gelte für alle $\sigma' \in \mathrm{Perm}(M)$ mit $\#(\mathrm{Tr}(\sigma')) < \#(\mathrm{Tr}(\sigma))$.

Induktionsschluss: Wähle $x_0 \in \mathrm{Tr}(\sigma)$ und $k_0 := \min\{k \in \mathbb{N} \mid \sigma^k(x_0) = x_0\}$. Wir bemerken, dass $k_0 \leq \mathrm{ord}(\sigma)$. Wir erhalten $Z_1 = \{x_0, \sigma(x_0), \ldots, \sigma^{k_0-1}(x_0)\}$, und setzen $\zeta_1 := (x_0, \sigma(x_0), \ldots, \sigma^{k_0-1}(x_0))$. Dann gilt also: Wenn $x \in Z_1$, dann ist $\zeta_1(x) = \sigma(x)$. Setze jetzt $\sigma_1 := \zeta_1^{-1} \circ \sigma$. Dann gilt: Wenn $x \in Z_1$, dann ist $\sigma_1(x) = \zeta_1^{-1}(\sigma(x)) = x$ und ist $x \in M - Z_1$, dann ist $\sigma_1(x) = \zeta_1^{-1}(\sigma(x)) = \sigma(x)$, also ist $\sigma_1(x) \in M - Z_1$. Damit ist $\mathrm{Tr}(\sigma_1) \subseteq \mathrm{Tr}(\sigma) - Z_1$, also haben wir die Ungleichung $\#(\mathrm{Tr}(\sigma_1)) < \#(\mathrm{Tr}(\sigma))$. Nach der Induktionsvoraussetzung gibt es disjunkte Zyklen $\zeta_2, \ldots, \zeta_N$ mit $\sigma_1 = \zeta_2 \circ \cdots \circ \zeta_N$, außerdem gilt $\mathrm{Tr}(\zeta_2), \ldots, \mathrm{Tr}(\zeta_N) \subseteq \mathrm{Tr}(\sigma_1)$ und somit $\mathrm{Tr}(\zeta_1) \cap \mathrm{Tr}(\zeta_i) = \varnothing$ für $i \in \{2, \ldots, N\}$. Insgesamt haben wir

$$\sigma = \zeta_1 \circ \sigma_1 = \zeta_1 \circ \zeta_2 \circ \cdots \circ \zeta_N$$

und $\zeta_1, \ldots, \zeta_N$ haben disjunkte Träger. □

Im Folgenden notieren wir Permutationen auch als Produkt disjunkter Zyklen, und erhalten damit eine zweite Schreibweise.

Korollar III.3.13 (aus Satz 5): *Es sei $\sigma \in S_n$. Dann gibt es eine nicht-negative ganze Zahl m und Transpositionen $\tau_1, \ldots, \tau_m$ in S_n, sodass $\sigma = \tau_1 \circ \cdots \circ \tau_m$.*

Achtung! Weder sind die oben angegebenen Transpositionen im Allgemeinen disjunkt, noch ist m eindeutig.

Beweis: Folgt aus Satz 5 in Verbindung mit Bemerkung III.3.11. □

Definition III.3.14 (Signum): Sei $\sigma \in S_n$. Wir definieren das *Signum* von σ durch

$$\mathrm{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

Hierbei heißt $\prod_{1 \leq i < j \leq n}(\cdot) := \prod_{1 \leq i \leq n} \prod_{j=i+1}^{n}(\cdot)$.

Beispiel III.3.15: (i) Sei $\sigma = (1,2,3,4) \in S_4$. Dann ist

$$\begin{aligned}\operatorname{sign}(\sigma) &= \frac{\sigma(2)-\sigma(1)}{2-1} \cdot \frac{\sigma(3)-\sigma(1)}{3-1} \cdot \frac{\sigma(4)-\sigma(1)}{4-1} \cdot \frac{\sigma(3)-\sigma(2)}{3-2} \\ &\quad \cdot \frac{\sigma(4)-\sigma(2)}{4-2} \cdot \frac{\sigma(4)-\sigma(3)}{4-3} \\ &= \frac{3-2}{2-1} \cdot \frac{4-2}{3-1} \cdot \frac{4-3}{3-2} \cdot \frac{1-3}{4-2} \cdot \frac{1-4}{4-3} \\ &= (-1) \cdot (-1) \cdot (-1) = -1.\end{aligned}$$

An diesem Beispiel können wir sehen, dass $\operatorname{sign}(\sigma) = (-1)^k$ für $k = \#\{(i,j) \mid i < j \text{ und } \sigma(i) > \sigma(j)\}$.

(ii) Sei $\sigma_2 = (1,2) \in S_n$. Alle Faktoren in der Definition von $\operatorname{sign}(\sigma_2)$ sind 1, bis auf diejenigen, wo $i = 1, 2$. Das heißt

$$\operatorname{sign}(\sigma_2) = \prod_{j=2}^{n} \frac{\sigma(j)-2}{j-1} \prod_{j=3}^{n} \frac{\sigma(j)-1}{j-2} = \frac{\sigma(2)-2}{2-1} \prod_{j=3}^{n} \frac{j-2}{j-1} \prod_{j=3}^{k} \frac{j-1}{j-2} = -1,$$

wobei wir verwendet haben, dass $\sigma(2) = 1$, und dass $\sigma(j) = 3$ für $j \geq 3$.

Bemerkung III.3.16 (Fehlstand und Wirkung von Permutationen): (i) Wie in (III.3.15) (i) kann man sich überlegen, dass für σ in S_n und k wie im Beispiel gilt, dass $\operatorname{sign}(\sigma) = (-1)^k$. Man nennt die Tupel (i,j) der Menge, über die k definiert wird, Fehlstände, und k entsprechend die Anzahl der Fehlstände von σ.

Insbesondere ist das Signum jeder Permutation ± 1, was man der expliziten Formel möglicherweise nicht sofort ansieht.

(ii) Ist π ein weiteres Element der S_n, dann ist

$$\operatorname{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j)-\sigma(i)}{j-i} = \prod_{1 \leq i < j \leq n} \frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)}.$$

Das liegt daran, dass π nur die Reihenfolge der Faktoren vertauscht. Genauer korrespondiert der Faktor zum Tupel (i,j) auf der rechten Seite zum Tupel (i',j') auf der rechten Seite, der durch

$$(i',j') = \begin{cases} (\pi(i),\pi(j)), & \text{falls } \pi(i) < \pi(j), \\ (\pi(j),\pi(i)), & \text{falls } \pi(i) > \pi(j) \end{cases}$$

festgelegt ist.

Proposition III.3.17: *Für Permutationen σ_1, σ_2 in S_n gilt* $\operatorname{sign}(\sigma_1 \circ \sigma_2) = \operatorname{sign}(\sigma_1)\operatorname{sign}(\sigma_2)$.

Beweis: Mithilfe von Proposition III.3.16 rechnen wir die Aussage einfach nach:

$$\begin{aligned}\operatorname{sign}(\sigma_1 \circ \sigma_2) &= \prod_{1\le i<j\le n} \frac{\sigma_1(\sigma_2(j)) - \sigma_1(\sigma_2(i))}{j-i} \\ &= \prod_{1\le i<j\le n} \frac{\sigma_1(\sigma_2(j)) - \sigma_1(\sigma_2(i))}{\sigma_2(j)-\sigma_2(i)} \prod_{1\le i<j\le n} \frac{\sigma_2(j)-\sigma_2(i)}{j-i} \\ &= \operatorname{sign}(\sigma_1)\operatorname{sign}(\sigma_2). \end{aligned}$$

□

Proposition III.3.18 (Konjugationstrick): *Seien $\#(M) \ge 2$ und a, b, a', b' Elemente von M mit $a \ne b$ und $a' \ne b'$. Wähle eine Permutation π mit $\pi(a') = a$ und $\pi(b') = b$. Dann gilt:*

$$(a', b') = \pi^{-1} \circ (a, b) \circ \pi.$$

Beweis: Für $x \in M$ gilt

$$(\pi^{-1} \circ (a,b) \circ \pi)(x) = \begin{cases} b', & \text{falls } x = a' \\ a', & \text{falls } x = b' \\ x, & \text{falls } x \notin \{a', b'\}. \end{cases}$$

□

Satz 6 (über das Signum): *Sei n eine natürliche Zahl.*

(i) *Die Signumgsfunktion* $\operatorname{sign}\colon S_n \to (\{\pm 1\}, \cdot)$ *ist ein Gruppenhomomorphismus.*

(ii) *Ist τ in S_n eine Transposition, dann ist* $\operatorname{sign}(\tau) = -1$.

(iii) *Ist ξ ein Zyklus der Länge ℓ, d. h. $\xi = (x_1 \dots x_\ell)$, dann ist*

$$\operatorname{sign}(\xi) = \begin{cases} -1, & \textit{falls } \ell \in 2\mathbb{N}, \\ 1, & \textit{falls } \ell \in 2\mathbb{N}+1. \end{cases}$$

Beweis: (i) Das haben wir in Proposition III.3.17 bereits festgehalten.

(ii) Nach dem Konjugationstrick Proposition III.3.18 gibt es eine Transposition π, sodass $\tau = \pi \circ (12) \circ \pi^{-1}$. In Beispiel III.3.15 haben wir ausgerechnet, dass $\operatorname{sign}(12) = -1$ und nach Proposition III.3.17 ist $\operatorname{sign}(\tau) = \operatorname{sign}(\pi)(-1)\operatorname{sign}(\pi^{-1}) = -1$.

(iii) Sei $\xi = (x_1 \dots x_\ell)$ ein ℓ-Zyklus mit $x_1, \dots, x_\ell$ aus $\{1, \dots, n\}$. Nach Korollar III.3.13 ist $\xi = (x_1 x_2) \circ (x_2 x_3) \circ \dots \circ (x_{\ell-1} x_\ell)$ und $\operatorname{sign}(\xi) = (-1)^\ell$. □

4. Ringe

Zur Erinnerung: Gruppen verallgemeinern in einem gewissen Sinne die aus der Schule bekannte Menge $(\mathbb{R}, +)$ der reellen Zahlen zusammen mit der Addition. Sie stellen ein abstraktes Konzept dessen dar, was man benötigt, um in einer Menge „vernünftig rechnen zu können".

In diesem Abschnitt möchten genauer verallgemeinern, was aus der Schule ebenfalls bekannt ist: Reelle Zahlen kann man Addieren und Multiplizieren. Und zwar auf eine Weise, sodass sich beide Operationen „vertragen" beziehungsweise „zusammenpassen". Das wird durch die Distributivgesetze sicher gestellt.

Definition III.4.1 (Ring):

(i) Eine Menge R mit Verknüpfungen „$+$", „$\cdot$" heißt *unitärer Ring* oder *Ring mit Eins*, falls gelten:

(1) $(R, +)$ ist eine kommutative Gruppe,

(2) „$\cdot$" ist assoziativ,

(3) Es gibt ein neutrales Element 1_R bezüglich „$\cdot$", das von 0_R verschieden ist, d. h. für alle $x \in R$ ist $x \cdot 1_R = x = 1_R \cdot x$.

(4) Es gelten die *Distributivgesetze*, d. h. für alle $x, y, z \in R$ gilt

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z), \qquad (y + z) \cdot x = (y \cdot x) + (z \cdot x).$$

Wir nennen die Verknüpfung „$+$" *Addition*, „$\cdot$" heißt *Multiplikation*, das neutrale Element von $(R, +)$ heißt 0_R (oder kurz manchmal 0), das neutrale Element von $(R, \cdot)$ heißt 1_R (oder kurz manchmal 1). Für $x \in R$ heißt $-x$ das Inverse zu x bezüglich „$+$". Für $x, y \in R$ ist $x - y := x + (-y)$. Für $x, y, z \in R$ ist $x \cdot y + z := (x \cdot y) + z$ („Punkt vor Strich").

Wir kürzen „unitärer Ring" beziehungsweise „Ring mit Eins" im Folgenden ab zu „Ring".

(ii) Falls die Verknüpfung „$\cdot$" zusätzlich kommutativ ist, dann heißt R *kommutativer Ring*.

Es gibt in verschiedenen Bereichen der Mathematik unterschiedliche Auffassungen dazu, was ein vernünftiger Ringbegriff ist. Entsprechend viele verschiedene Definitionen des Begriffs gibts es, und entsprechend wichtig ist es, sich an die (in dieser Vorlesung) gewählte Konvention zu halten. Diese deckt das Spektrum an Ringen ab, die wir hier behandeln wollen.

Für den Kontext: Insbesondere in der kommutativen Algebra und algebraischen Geometrie ist man der Auffassung, dass ein Ring eine Eins haben sollte.

Ringe ohne Eins werden dort manchmal als „Rng“ bezeichnet; unter anderem um anzuzeigen, dass diesem Ring etwas „fehlt“. Außerdem werden Ringe dort in aller Regel als kommutativ vorausgesetzt.

Beispielsweise in der Theorie der Operatoralgebren ist es hingegen natürlich, sich mit nicht zwangsläufig kommutativen Ringen zu beschäftigen. Ebenfalls spielen dort Ringe ohne Eins eine gewisse Rolle.

Beispiel III.4.2: (i) $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Ringe.

(ii) $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ist ein Ring. Hierbei ist die Multiplikation erklärt durch

$$[a] \cdot [b] := [a \cdot b].$$

Dass diese Verknüpfung wohldefiniert ist, zeigt man genau wie bei der Addition.

(iii) $(\mathbb{R}^{n\times n}, +, \cdot)$ ist ein Ring.

(iv) Seien R ein beliebiger Ring und M eine nichtleere Menge. Dann wird $R' := \mathrm{Abb}(M, R) = R^M$ zu einem Ring mit den nachfolgend definierten Verknüpfungen:

$$f{+}g\colon M \longrightarrow R, \quad m \longmapsto f(m){+}g(m), \qquad f{\cdot}g\colon M \longrightarrow R, \quad m \longmapsto f(m){\cdot}g(m)$$

für $f, g \in R'$. Hierbei ist $0_{R'}$ die Abbildung von M nach R, die jedes Element von M auf 0_R abbildet, und $1_{R'}$ die Abbildung, die jedes Element von M auf 1_R abbildet.

Proposition III.4.3 (Erste Eigenschaften): *Sei $(R, +, \cdot)$ ein Ring. Dann gilt:*

(i) *Für alle $x \in R$ ist $0_R \cdot x = 0_R = x \cdot 0_R$.*

(ii) *Für alle $x, y \in R$ ist $(-x) \cdot y = -(x \cdot y) = x \cdot (-y)$.*

Beweis: (i) Es gilt $0_R \cdot x = (0_R + 0_R) \cdot x = 0_R \cdot x + 0_R \cdot x$, d. h.

$$0_R = -(0_R \cdot x) + 0_R \cdot x = -(0_R \cdot x) + 0_R \cdot x + 0_R \cdot x = 0_R \cdot x + 0_R \cdot x = 0_R \cdot x.$$

(ii) Wir wollen zeigen, dass $(-x) \cdot y$ das additive Inverse zu $x \cdot y$ ist:

$$x \cdot y + (-x) \cdot y = (x + (-x)) \cdot y = 0_R \cdot y = 0_R$$

und wegen der Kommutativität der Addition ist $(-x) \cdot y = -(x \cdot y)$. Analog zeigt man, dass $x \cdot (-y) = -(x \cdot y)$. □

Definition III.4.4 (Ringhomomorphismus):

(i) Eine Abbildung $\Phi\colon R \to S$ heißt *Homomorphismus unitärer Ringe* zwischen zwei Ringen $(R, +_R, \cdot_R)$, $(S, +_S, \cdot_S)$, wenn gelten:
   (1) Für alle $x, y \in R$ ist $\Phi(x +_R y) = \Phi(x) +_S \Phi(y)$,
   (2) Für alle $x, y \in R$ ist $\Phi(x \cdot_R y) = \Phi(x) \cdot_S \Phi(y)$,
   (3) Es gilt $\Phi(1_R) = 1_S$.

   Wir schreiben in diesem Fall $\Phi\colon (R, +_R, \cdot_R) \to (S, +_S, \cdot_S)$. Im Folgenden schreiben wir meistens einfach „Ringhomomorphismus“ statt „Homomorphismus unitärer Ringe“.

(ii) Wir schreiben

$$\begin{aligned}&\mathrm{Hom}_{\mathrm{Ring}}(R, S)\\ &\quad := \{\Phi\colon (R, +_R, \cdot_R) \to (S, +_S, \cdot_S) \mid \Phi \text{ ist Ringhomomorphismus}\}\end{aligned}$$

   für die Menge der Ringhomomorphismen von $(R, +_R, \cdot_R)$ nach $(S, +_S, \cdot_S)$.

Ist $\Phi\colon (R, +_R, \cdot_R) \to (S, +_S, \cdot_S)$ ein Homomorphismus unitärer Ringe, dann ist Φ insbesondere ein Gruppenhomomorphismus der Gruppen $(R, +_R)$ und $(S, +_S)$. Insbesondere gibt es wieder

$$\mathrm{Kern}(\Phi) = \{r \in R \mid \Phi(r) = 0_S\}, \qquad \mathrm{Im}(\Phi) = \{\Phi(r) \mid r \in R\}.$$

Allerdings handelt es sich bei $\mathrm{Kern}(\Phi)$ nicht um einen Ring im Sinne unserer Definition, da 1_R nicht zu $\mathrm{Kern}(\Phi)$ gehört. Genau wie für Gruppen zeigt man, dass Φ injektiv ist genau dann, wenn $\mathrm{Kern}(\Phi) = \{0_R\}$ ist.

Bemerkung III.4.5 (Kanonischer Homomorphismus): Sei $(R, +, \cdot)$ ein Ring. Dann gibt es einen kanonischen Homomorphismus $\chi_R\colon \mathbb{Z} \to R$ mit den Eigenschaften $\chi_R(0) = 0_R$; $\chi_R(n+1) = \chi_R(n) + 1_R$ für nichtnegative ganze Zahlen n und $\chi_R(-n) = -\chi_R(n)$ für natürliche Zahlen n. Es gilt also

$$\Phi(z) := \begin{cases} \sum_{i=1}^{z} 1_R, & \text{falls } z > 0,\\ 0_R, & \text{falls } z = 0,\\ \sum_{i=1}^{-z} -1_R, & \text{falls } z < 0. \end{cases}$$

Dann ist Φ ein Ringhomomorphismus von $(\mathbb{Z}, +, \cdot)$ nach $(R, +, \cdot)$, den wir *kanonischen Homomorphismus für R* nennen.

Definition III.4.6: Für einen Ring $(R, +, \cdot)$ definieren wir

$$\mathrm{char}(R) := \begin{cases} 0, & \text{falls } \sum_{i=1}^{k} 1_R \neq 0_R \text{ für alle } k \in \mathbb{N},\\ \min\{k \mid \sum_{i=1}^{k} 1_R = 0_R\}, & \text{sonst} \end{cases}$$

und nennen $\mathrm{char}(R)$ die *Charakteristik von R*.

Genau dann ist χ_R injektiv, wenn $\mathrm{char}(R) = 0$. In diesem Fall können wir $\mathbb{Z}$ als „Teilring von R auffassen“, beziehungsweise genauer $\mathbb{Z}$ so mit einem Teilring von R von identifizieren, dass sich diese Kopie genau wie die gewohnten ganzen Zahlen verhält.

Definition III.4.7 (Einheitengruppe): Sei $(R, +, \cdot)$ ein Ring.

(i) Sei x ein Element von R. Gibt es ein Element y von R, sodass $xy = 1_R = yx$, dann heißt x *invertierbar*. In diesem Fall ist y eindeutig bestimmt und wird mit x^{-1} notiert.

(ii) Seien x_1, x_2 Elemente von R. Sind beide invertierbar, dann auch x_1x_2 mit $(x_1x_2)^{-1} = x_2^{-1}x_1^{-1}$.

(iii) Die Menge $R^\times := \{x \in R \mid x \text{ ist invertierbar}\}$ heißt *Einheitengruppe von* R und es handelt sich tatsächlich um eine Gruppe.

Beweis: (i) Das ist dasselbe Argument wie in Definition III.1.2.

(ii) Sind $x_1, x_2 \in R^\times$, dann ist $(x_1 \cdot x_2) \cdot (x_2^{-1} \cdot x_1^{-1}) = 1_R$ und $(x_2^{-1} \cdot x_1^{-1}) \cdot (x_1 \cdot x_2) = 1_R$, d. h. $x_1 \cdot x_2$ ist invertierbar mit Inversem $x_2^{-1}x_1^{-1}$, also $x_1 \cdot x_2 \in R^\times$.

(iii) Per Definition ist „$\cdot$“ assoziativ mit neutralem Element 1_R und jedes Element von $R^\times$ hat per Konstruktion ein Inverses bezüglich „$\cdot$“, d. h. $R^\times$ ist eine Gruppe. □

Ein Polynom ist von der Form $a_0 + a_1X + \cdots + aNX$ für eine Familie $a_0, \ldots, a_N$ ($N \in \mathbb{N}$) respektive eine Folge $a_0, \ldots, a_N, 0, \ldots$ von Ringelementen.

Zur Erinnerung: Eine Abbildung $a\colon \mathbb{N} \to R$, heißt *Folge mit Gliedern in* R oder *Folge in* R. Üblicherweise schreibt man a_n statt $a(n)$ und identifiziert die Abbildung a mit der Familie ihrer Bilder $(a_n)_{n\in\mathbb{N}}$.

Gibt es ein Element r in R und eine natürliche Zahl N, sodass $a_n = r$ für jedes $n \geq N$, dann heißt die Folge $(a_n)_{n\in\mathbb{N}}$ schließlich konstant. Ist $(R, +, \cdot)$ ein Ring, ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in R und ist diese Folge schließlich konstant 0_R, dann nennt man die Folge eine schließlich konstante Nullfolge.

Definition III.4.8 (Polynom): Sei $(R, +, \cdot)$ ein Ring.

(i) Ein *Polynom p über R* ist eine schließlich konstante Nullfolge $(a_n)_{n\in\mathbb{N}}$.

(ii) Ist die zugehörige Folge in (i) die konstante Nullfolge, dann heißt p das *Nullpolynom*. Wir schreiben auch $p = \mathbf{0}$.

(iii) Wir nennen

$$\mathrm{Grad}(p) = \begin{cases} \max\{i \in \mathbb{N}_0 \mid a_i \neq 0\}, & \text{falls } p \neq \mathbf{0}, \\ -\infty, & \text{falls } p = \mathbf{0} \end{cases}$$

den *Grad des Polynoms p.*

Definition III.4.9 (Polynomring): Sei $(R, +, \cdot)$ ein Ring. Wir schreiben $R[X]$ für die Menge der Polynome über R. Auf $R[X]$ definieren wir Verknüpfungen „+“ und „·“ wie folgt: Für $p = \sum_{i=0}^{n} a_i X^i$ und $q = \sum_{j=0}^{m} b_j X^j$ setzen wir

$$p + q = \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) X^i, \qquad pq = \sum_{i=0}^{n+m}$$

wobei $c_i = \sum_{k=0}^{i} a_k b_{i-k}$ für $0 \leq i \leq n + m$. Wir halten fest, dass $a_i = 0$ und $b_j = 0$, wenn $i > n$ respektive $j > m$. Dann bildet $(R[X], +, \cdot)$ einen Ring, und wird *Polynomring über R* genannt. Dabei ist $p = 1$ (mit $\mathrm{Grad}(p) = 1$) das Einselement und $p = \mathbf{0}$ (mit $\mathrm{Grad}(p) = -\infty$) das Nullelement.

Definition III.4.10 (Körper): Sei $(R, +, \cdot)$ ein Ring. Ist R ein kommutativer Ring und ist $R^\times = R - \{0\}$ (ist also jedes von Null verschiedene Element von R multipliktativ invertierbar), dann heißt R ein *Körper.*

Beispiel III.4.11: (i) Die Mengen der rationalen- bzw. reellen Zahlen mit den gewöhnlichen Verknüpfungen sind Körper.

(ii) Bei $(\mathbb{Z}, +, \cdot)$ oder $(\mathbb{R}^{n\times n}, +, \cdot)$ handelt es sich nicht um Körper.

(iii) Der Ring $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ist im Allgemeinen auch kein Körper. Ist beispielsweise $n = 6$, dann ist $\overline{2} \cdot \overline{3} = \overline{0}$. Gäbe es jetzt ein $s \in \mathbb{Z}/6\mathbb{Z}$, sodass $s \cdot \overline{2} = \overline{1}$, dann wäre $\overline{3} = \overline{1} \cdot \overline{3} = s \cdot \overline{2} \cdot \overline{3} = s \cdot \overline{0} = \overline{0}$, was natürlich Unsinn ist.

Proposition III.4.12: *Ist p eine Primzahl, dann ist* $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ *ein Körper. Wir schreiben* $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$.

Das $\mathbb{F}$ in der oben erklärten Schreibweise rührt vom englischen Wort für das Konzept eines Körpers, nämlich „field“, her.

Beweis: Da wir bereits wissen, dass $\mathbb{Z}/p\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{p-1}\}$ ein kommutativer Ring ist, bleibt nur zu zeigen, dass jedes von Null verschiedene Element multiplikativ invertierbar ist. Sei dazu $\overline{a}$ ein von Null verschiedenes Element von $\mathbb{Z}/p\mathbb{Z}$. Die von $\overline{a}$ erzeugte Untergruppe in $(\mathbb{Z}/p\mathbb{Z}/p)$ enthält $\overline{a}$ und $\overline{0}$, ist nach dem Satz von Lagrange also ganz $\mathbb{Z}/p\mathbb{Z}$. Das bedeutet, dass auch $\overline{1}$ in $\langle\overline{a}\rangle$ liegt; es gibt also eine natürliche Zahl k, sodass $\overline{1} = \sum_{i=1}^{k} \overline{a} = \overline{k}\overline{a}$. □

Beispiel III.4.13 (Körper der komplexen Zahlen): Es bezeichne $\mathbb{C} := \mathbb{R} \times \mathbb{R}$. Auf $\mathbb{C}$ erklären wir Verknüfungen „+“ und „$\cdot$“ wie folgt: Für $c_1 = (u_1, v_1)$ und $c_2 = (u_2, v_2)$ aus $\mathbb{C}$ seien

$$c_1 + c_2 := (u_1 + u_2, v_1 + v_2), \qquad c_1 c_2 = (u_1 u_2 - v_1 v_2, u_1 v_2 + u_2 v_1).$$

Diese Verknüpfungen machen $\mathbb{C}$ zu einem Körper. Wir schreiben $\mathrm{i} := (0, 1)$, $r := (r, 0)$, wenn r eine reelle Zahl ist, und $u + \mathrm{i}v := (u, v)$. Mit dieser Schreibweise ist $\mathrm{i}^2 = (0, 1)(0, 1) = (-1, 0) = -1$.

Kapitel IV.

Vektorräume und Dimensionstheorie

In diesem Kapitel wollen wir das bereits bekannte Konzept des $\mathbb{R}$-Vektorraums verallgemeinern zum Vektorraum über einem beliebigen Körper und die Eigenschaften von Vektorräumen näher studieren.

1. Vektorräume

In diesem Abschnitt sei stets $(K, +, \cdot)$ ein Körper.

Definition IV.1.1: Ein *Vektorraum über dem Körper K* (kürzer: *K-Vektorraum*) ist eine Menge V zusammen mit einer Verknüpfung „+“ und einer „äußeren Verknüpfung“ $\cdot\colon K \times V \to V$, so dass gelten:

(VA) $(V, +)$ ist eine abelsche Gruppe,

(M_1) Für alle $x \in V$ ist $1_K x = x$,

(M_2) Für alle $\lambda_2, \lambda_2 \in K$ und $x \in V$ ist $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$,

(M_3) Für alle $\lambda \in K$ und $x_1 x_2 \in V$ ist $\lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2$,

(M_4) Für alle $\lambda_1, \lambda_2 \in K$ und $x \in V$ ist $\lambda_1 \cdot (\lambda_2 x) = (\lambda_1 \cdot \lambda_2)x$.

Beispiel IV.1.2: Die Menge

$$K^n = \{(x_1, \ldots, x_n)^t \mid x_1, \ldots, x_n \in K\}$$

ist ein K-Vektorraum mit komponentenweiser Addition und komponentenweiser skalarer Multiplikation.

Bemerkung IV.1.3: (i) Für $K = \mathbb{R}$ ist Definition IV.1.1 genau die Definition von $\mathbb{R}$-Vektorräumen aus Proposition II.2.2.

(ii) Wir können wie in Proposition II.3.1 $(p \times q)$-Matrizen mit *Einträgen in* K als Abbildungen

$$\{1, \ldots, p\} \times \{1, \ldots, q\} \longrightarrow K$$

definieren und erhalten den K-Vektorraum $K^{p \times q}$ wie in (Definition II.3.4).

(iii) Alle Aussagen aus den Abschnitten (II.2), (II.3), (II.4) und (II.5) gelten genau so für Vektorräume, Matrizen und lineare Gleichungssysteme über einem beliebigen Körper K.

Im Folgenden seien V, W stets K-Vektorräume. Bei allen auftretenden Vektorräumen notieren wir die Addition mit „+“ und die skalare Multiplikation mit „·“. Für $v \in V$ und $\lambda \in K$ schreiben wir auch einfach kurz λv statt $\lambda \cdot v$.

Definition IV.1.4 (Vektorraumhomomorphismus):

(i) Ein (Vektorraum-)Homomorphismus von V nach W ist eine Abbildung $\Phi\colon V \to W$ mit den folgenden Eigenschaften:

(1) Für alle $u, v \in V$ ist $\Phi(u + v) = \Phi(u) + \Phi(v)$,

(2) Für alle $\lambda \in K$ und $v \in V$ ist $\Phi(\lambda v) = \lambda \Phi(v)$.

Φ ist insbesondere ein Gruppenhomomorphismus von $(V, +)$ nach $(W, +)$. Statt Vektorraumhomomorphismus ist auch die Bezeichnung *(K-)lineare Abbildung* gebräuchlich.

(ii) Φ heißt *Endomorphismus*, falls $V = W$. Φ heißt *Isomorphismus* falls es eine lineare Abbildung $\Psi\colon V \to W$ gibt, die $\Psi \circ \Phi = \mathrm{id}_V$ und $\Phi \circ \Psi = \mathrm{id}_W$ leistet. Φ heißt *Automorphismus*, falls Φ ein Endomorphismus und ein Isomorphismus ist. Die Vektorräume V und W heißen *isomorph*, falls es einen Vektorraum-Isomorphismus $\Phi\colon V \to W$ gibt; man schreibt in diesem Fall $V \cong W$.

Beispiel IV.1.5: (i) Eine Matrix $A \in K^{p \times q}$ definiert die lineare Abbildung $\Phi_A\colon K^q \to K^p, v \mapsto Av$.

(ii) Die Transpositionsabbildung ${}^t\colon K^{p \times q} \to K^{q \times p}, A \mapsto A^t$ ist eine lineare Abbildung (vergleiche (Proposition II.3.10)).

(iii) Die Nullabbildung $V \to W, v \mapsto 0_W$ ist eine lineare Abbildung.

Bemerkung IV.1.6: Sei $\Phi\colon V \to W$ eine lineare Abbildung.

(i) $\mathrm{Kern}(\Phi) = \{v \in V \mid \Phi(v) = 0_W\} = \Phi^{-1}(0_W)$ ist ein Untervektorraum von V.

(ii) Φ ist injektiv genau dann, wenn $\mathrm{Kern}(\Phi) = \{0_V\}$.

(iii) Φ ist ein Isomorphismus genau dann, wenn Φ ein bijektiver Homomorphismus ist.

(iv) Verkettungen von linearen Abbildungen sind lineare Abbildungen.

Beweis: (i) Aus (Bemerkung III.2.6) wissen wir, dass $\mathrm{Kern}(\Phi)$ eine Untergruppe von $(V,+)$ ist, d. h. es gilt $0_V \in \mathrm{Kern}(\Phi)$. Ebenso aus der Situation für Gruppen ist bekannt: Sind $v_1, v_2 \in \mathrm{Kern}(\Phi)$, dann ist auch $v_1 + v_2 \in \mathrm{Kern}(\Phi)$. Seien jetzt $\lambda \in K$ und $v \in \mathrm{Kern}(\Phi)$. Dann ist $\Phi(\lambda v) = \lambda\Phi(v) = \lambda \mathbf{0}_W = \mathbf{0}_W$, d. h. $\lambda v \in \mathrm{Kern}(\Phi)$.

(ii) Folgt aus (Proposition III.2.7).

(iii) Der Beweis funktioniert analog zum Beweis von (Proposition III.2.10).

(iv) Nachrechnen. □

Erinnerung: Die Menge $\mathrm{Abb}(V,W)$ ist ein K-Vektorraum (vergleiche Proposition II.2.3).

Proposition IV.1.7: *Die Teilmenge*

$$\mathrm{Hom}_{K-\mathrm{VR}}(V,W) := \{\Phi\colon V \to W \mid \Phi \textit{ ist lineare Abbildung}\} \subseteq \mathrm{Abb}(V,W)$$

ist ein Untervektorraum. Wir schreiben kurz $\mathrm{Hom}(V,W) := \mathrm{Hom}_{K-\mathrm{VR}}(V,W)$.

2. Basen und lineare Unabhängigkeit

In diesem Abschnitt seien stets K ein Körper und V ein Vektorraum über K.

Wir erinnern uns daran, dass es für ein Element x des K^n eindeutige $\lambda_1, \ldots, \lambda_n$ aus K gibt, sodass $x = \sum_{i=1}^n \lambda_i e_i$. Hierbei sind die e_i die kanonischen Basisvektoren $e_i = (\delta_{i,j})_{1\leq j\leq n}$ des K^n.

Bezeichnen $F^{(1)}, \ldots, F^{(r)}$ die Fundamentallösungen des linearen Gleichungssystems $Ax = 0$ und gehört v zu $\mathbb{L}(A|0)$, dann lässt sich v in eindeutiger Weise schreiben als $v = \sum_{i=1}^r \alpha_i F^{(i)}$, d. h. wie oben sind die Koeffizienten $\alpha_1, \ldots, \alpha_r$ eindeutig bestimmt.

Wir möchten in diesem Abschnitt die Begriffe „Basis“ und „Dimension“ einführen und studieren. Dann werden wir zeigen, dass jeder K-Vektorraum (unter Vorbehalt) eine Basis besitzt, und dass alle Basen dieselbe Mächtigkeit besitzen. Die Dimension des Vektorraums ist deshalb definiert als die Mächtigkeit einer Basis dieses Vektorraums.

Definition IV.2.1 (Linearkombination): Seien M eine Teilmenge von V und v ein Element von V. Gibt es eine nichtnegative ganze Zahl n, Elemente $v_1, \dots, v_n$ von M und Elemente $\lambda_1, \dots, \lambda_n$ von K, sodass $v = \sum_{i=1}^{n} \lambda_i v_i$, dann heißt v eine *Linearkombination von M*.

Ist in der obigen Definition $n = 0$, dann ist die Summe leer. Diese Summe ist per Definition $\mathbf{0}_V$.

Bemerkung IV.2.2: Der Nullvektor $\mathbf{0}_V$ ist Linearkombination für jede Teilmenge von V. Außerdem ist er die einzige Linearkombination der leeren Menge.

Definition IV.2.3 (Linearkombination, überarbeitet): Ist M eine Menge und $f\colon M \to K$ eine Abbildung, dann heißt $\mathrm{Tr}(f) = \{m \in K \mid f(m) \neq 0\}$ der *Träger von f*. Wir schreiben $\mathrm{Abb}_0(M, K) = \{f \in \mathrm{Abb}(M, K) \mid \mathrm{Tr}(f) \text{ ist endlich}\}$.

Ist M eine Teilmenge von V und hat $\lambda\colon M \to K$ endlichen Träger, dann definieren wir

$$\sum_{v \in M} \lambda(v)v = \sum_{v \in \mathrm{Tr}(\lambda)} \lambda(v)v.$$

Da der Träger von λ endlich ist, ist die obige Summe endlich.

Definition IV.2.4 (Basis): Sei B eine Teilmenge von V. Falls sich jeder Vektor $v \in V$ auf genau eine Weise als Linearkombination von B schreiben lässt, d. h. wenn es für jedes v in V genau ein λ in $\mathrm{Abb}_0(B, K)$ gibt, sodass $v = \sum_{w \in B} \lambda(w)w$, dann heißt B eine *Basis von V*.

Als erstes Etappenziel möchten wir zeigen, dass jeder „endlich erzeugte“ Vektorraum eine Basis besitzt.

Definition IV.2.5 (Lineare Unabhängigkeit): Sei M eine Teilmenge von V. Falls folgt, wenn immer $\sum_{v \in M} \lambda(v)v = \mathbf{0}_V$, $\lambda \in \mathrm{Abb}_0(M, K)$, ist, dass λ die Nullabbildung ist, dann heißt M *linear unabhängig*. Ist M nicht linear unabhängig, dann heißt M *linear abhängig*.

Lineare Unabhängigkeit einer Menge M ist die Forderung danach, dass der Nullvektor auf eindeutige Weise Linearkombination von M ist. Allerdings macht das bereits alle Linearkombinationen von M eindeutig.

Beispiel IV.2.6: (i) Seien $K = \mathbb{R}$ und $M_1 = \{v_1 = (1, 2, 3)^t, v_2 = (1, -1, 1)^t\}$. Dann ist $v_3 = (1, 5, 5)^t$ Linearkombination von M, denn $v_3 = 2v_1 + 3v_2$.

(ii) Die Teilmenge $M_2 = \{(1,0)^t, (1,1)^t\}$ ist linear unabhängig in K^2. Ist nämlich $\lambda_1(1,0)^t + \lambda_2(1,1)^t = (0,0)^t$, dann ist $\lambda_1 + \lambda_2 = 0$ und $\lambda_2 = 0$, sodass auch $\lambda_1 = 0$ gelten muss. Die zugehörige Abbildung $\lambda\colon M_2 \to K$ ist gegeben durch $(1,0)^t \mapsto 0$ und $(1,1)^t \mapsto 0$, ist also die Nullabbildung.

(iii) Fügen wir M_1 aus (i) den Vektor v_3 hinzu, dann erhalten wir eine linear abhängige Menge wegen $v_3 - 2v_1 - v_2 = \mathbf{0}$.

Bemerkung IV.2.7 (Lineare Unabhängigkeit für endliche Mengen): Eine endliche Teilmenge $M = \{v_1, \ldots, v_n\}$ von V ist genau dann linear unabhängig, wenn für jede Linearkombination der Null $\sum_{i=1}^n \lambda_i v_i = \mathbf{0}_V$ von M folgt, dass $\lambda_1 = \cdots = \lambda_n = 0$.

Definition IV.2.8 (Lineare Hülle): Sei M eine Teilmenge von V. Dann heißt

$$\mathrm{Lin}(M) = \{v \in V \mid v \text{ ist Linearkombination von } M\}$$

die *lineare Hülle von M* oder auch das *Vektorraumerzeugnis von M* beziehungsweise kurz *Erzeugnis von M*. Wir schreiben auch $\langle M\rangle := \mathrm{Lin}(M)$ und $\langle v_1, \ldots, v_n\rangle := \mathrm{Lin}(\{v_1, \ldots, v_n\})$.

Beispiel IV.2.9 (Erste lineare Hüllen): (i) Es sei $M = \{e_1, e_2\}$ in K^3. Dann ist

$$\begin{aligned}\mathrm{Lin}(M) &= \left\langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \right\rangle \\ &= \left\{ \lambda_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\0\end{pmatrix} : \lambda 1, \lambda_2 \in K \right\} = \left\{ \begin{pmatrix}\lambda_1\\\lambda_2\\0\end{pmatrix} : \lambda_1, \lambda_2 \in K \right\}.\end{aligned}$$

(ii) Für $M = \varnothing$ als Teilmenge von V ist $\mathrm{Lin}(M) = \{\mathbf{0}_V\}$.

Proposition IV.2.10 (Eigenschaften der linearen Hülle): *Für eine Teilmenge M von V gilt:*

(i) $M \subseteq \mathrm{Lin}(M)$.

(ii) $\mathrm{Lin}(M)$ *ist ein Untervektorraum von V.*

(iii) *Für $\langle -\rangle$ wie definiert auf Blatt 5, Aufgabe 3, gilt* $\mathrm{Lin}(M) = \langle M\rangle$, *d. h.*

$$\mathrm{Lin}(M) = \bigcap (U \subseteq V \mid U \text{ ist Untervektorraum von } V,\ M \subseteq U\}).$$

(iv) *Ist M' eine Teilmenge von M, dann ist* $\mathrm{Lin}(M')$ *eine Teilmenge von* $\mathrm{Lin}(M)$.

(v) *M ist ein Untervektorraum von V genau dann, wenn* $M = \mathrm{Lin}(M)$. *(Entsprechend ist* $\mathrm{Lin}(\mathrm{Lin}(M)) = \mathrm{Lin}(M)$.*)*

(vi) *Für zwei Untervektorräume U_1, U_2 von V gilt:*

$$\mathrm{Lin}(U_1 \cup U_2) = U_1 + U_2 = \{x + y \mid x \in U_1, y \in U_2\}.$$

Hierbei ist $U_1 + U_2 = \{v + w \mid v \in U_1, w \in U_2\}$.

Beweis: (i) Ein Element x in M können wir zum Beispiel durch $x = 1x$ als Linearkombination von M schreiben.

(ii) Per Definition gehört $\mathbf{0}_V$ zu $\mathrm{Lin}(M)$ (siehe auch Bemerkung IV.2.2). Sind v und w Elemente von $\mathrm{Lin}(M)$, dann gibt es $v_1, \ldots, v_n$ aus M und $\lambda_1, \ldots, \lambda_n$ aus K, sodass $v = \sum_{i=1}^{n} \lambda_i v_i$. Genauso gibt es passende $v_1', \ldots, v_m'$ aus M und $\lambda_1', \ldots, \lambda_m'$ aus K, sodass $w = \sum_{i=1}^{m} \lambda_i' v_i'$. Wir dürfen $m \geq n$ annehmen, und die Darstellung von v passend mit Nullen auffüllen: $v = \sum_{i=1}^{n} \lambda_i v_i + \sum_{i=n+1}^{m} 0 v_i'$. Damit gehört $v + w = \sum_{i=1}^{n} \lambda_i v_i + \sum_{i=n+1}^{m} \lambda_i' v_i'$ zu $\mathrm{Lin}(M)$. Genauso gehört λv zu $\mathrm{Lin}(M)$.

(iii) Sei U ein Unterraum von V, der M enthält. Ist w ein Element von $\mathrm{Lin}(M)$, d. h. es gibt $v_1, \ldots, v_n$ in M und $\lambda_1, \ldots, \lambda_n$ in K, sodass $v = \sum_{i=1}^{n} \lambda_i v_i$. Weil $v_1, \ldots, v_n$ per Voraussetzung auch in U liegen und weil U unter Summen und skalarer Multiplikation abgeschlossen ist, gehört dann auch v zu U. Wir haben damit gezeigt, dass auch $\mathrm{Lin}(M)$ eine Teilmenge von U ist.

Insgesamt folgt $M \subseteq \mathrm{Lin}(M) \subseteq \langle M \rangle$ mit der Definition von $\langle M \rangle$ von Blatt 5, Aufgabe 3. Weil $\langle M \rangle$ der kleinste Untervektorraum von V ist, der M enthält, gilt auch $\langle M \rangle \subseteq \mathrm{Lin}(M)$.

(iv) Folgt aus den Definitionen.

(v) „$\Longleftarrow$“ folgt aus (ii), „$\Longrightarrow$“ folgt aus (iii).

(vi) „$\subseteq$“: Da $U_1 + U_2$ ein Unterraum von V ist, der $U_1 \cup U_2$ enthält, ist $\mathrm{Lin}(U_1 \cup U_2)$ in $U_1 + U_2$ enthalten.

„$\supseteq$“: Ist w ein Element von $U_1 + U_2$, dann ist w von der Form $w = w_1 + w_2$ für w_1 aus U_1 und w_2 aus U_2. Insbesondere gehören die w_i zu $U_1 \cup U_2$ und w ist ein Element von $\mathrm{Lin}(U_1 \cup U_2)$. □

Satz 7 (Kriterium I für Basen): *Eine Teilmenge B von V ist eine Basis von V genau dann, wenn B linear unabhängig ist mit* $\mathrm{Lin}(B) = V$.

Beweis: Die Implikation „$\Longrightarrow$“ ist direkte Konsequenz der Definition einer Basis (nur die triviale Nulldarstellung).

Für „$\Longleftarrow$“: Sei $v \in V$. Da $\mathrm{Lin}(B) = V$, ist v eine Linearkombination von B. Was wir noch zu zeigen haben ist die Eindeutigkeit dieser Linearkombination. Seien $\lambda^{(1)}, \lambda^{(2)} \in \mathrm{Abb}_0(B, V)$ mit $v = \sum_{w\in B} \lambda^{(1)}(w)w = \sum_{w\in B} \lambda^{(2)}(w)w$. Dann ist

$$\mathbf{0} = v - v = \sum_{w\in B} (\lambda^{(1)}(w) - \lambda^{(2)}(w))w.$$

Da aber B linear unabhängig ist, muss $\lambda^{(1)}(w) - \lambda^{(2)}(w) = 0$ für alle $w \in B$ gelten, d. h. es ist schon $\lambda^{(1)} = \lambda^{(2)}$ und wir haben die Eindeutigkeit gezeigt.□

Beispiel IV.2.11: Seien $V = \mathbb{R}^2$ und

$$B = \left\{ b_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, b_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Ist B eine Basis des $\mathbb{R}^2$? Nach Satz 7 haben wir zwei Dinge zu prüfen: Die lineare Unabhängigkeit von B und ob B ganz $\mathbb{R}^2$ erzeugt.

Zur linearen Unabhängigkeit: Seien reelle Zahlen λ_1, λ_2 gegeben, sodass $\lambda_1 b_1 + \lambda_2 b_2 = \mathbf{0}$. Wir erhalten für die erste bzw. die zweite Koordinate die Gleichung $\lambda_1 + \lambda_2 = 0$ bzw. $\lambda_1 - \lambda_2 = 0$, welche äquivalent sind zum linearen Gleichungssystem

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Setzen wir $A := \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, dann wissen wir: B ist linear unabhängig genau dann, wenn das homogene Gleichungssystem $Ax = \mathbf{0}$ nur die Lösung $\{\mathbf{0}\}$ hat. Nach Korollar II.5.21 ist das genau dann der Fall, wenn der Rang von A gleich der Anzahl der Spalten von A (nämlich 2) ist.

Zur Erzeugenden-Eigenschaft: Die lineare Hülle von B ist $\mathbb{R}^2$ genau dann, wenn es für jedes $v \in \mathbb{R}^2$ reelle Zahlen λ_1 und λ_2 gibt, sodass $v = \lambda_1 b_1 + \lambda_2 b_2$. Das ist genau dann der Fall, wenn es für jedes $v \in \mathbb{R}^2$ reelle Zahlen λ_1, λ_2 gibt, sodass

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = v.$$

Wiederum nach Korollar II.5.21 ist das äquivalent dazu, dass der Rang von A gleich der Anzahl der Zeilen von A (nämlich 2) ist.

Es bleibt also, den Rang von A zu bestimmen:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \xrightarrow{\text{II}-\text{I}} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \xrightarrow{\frac{1}{2}\text{II}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \xrightarrow{\text{I}-\text{II}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Da der Rang von A tatsächlich 2 ist, ist B eine Basis von $\mathbb{R}^2$.

Proposition IV.2.12 (Kriterium für Basis im K^n): *Seien K ein Körper und n, m natürliche Zahlen. Ferner seien $v_1, \ldots, v_n \in K^n$ und $A = (v_1 | \ldots | v_m)$.*

(i) *Die Menge $\{v_1, \ldots, v_m\}$ ist linear unabhängig genau dann, wenn der Rang von A die Anzahl m der Spalten von A ist.*

(ii) *Die Menge $\{v_1, \ldots, v_m\}$ ist ein Erzeugendensystem von K^n genau dann, wenn der Rang von A die Anzahl n der Zeilen von A ist.*

Korollar IV.2.13 (Dimension des K^n): *Seien K ein Körper und n eine natürliche Zahl. Jedes Basis des K^n hat genau n Elemente.*

Beweis: Sei B eine Basis des K^n. Nach Proposition IV.2.12(i) ist die Mächtigkeit von B höchstens n. Insbesondere ist B endlich. Es gibt also eine natürliche Zahl m und Vektoren $v_1, \ldots, v_m$ aus V, sodass wir B schreiben können als $B = \{v_1, \ldots, v_m\}$. Setzen wir $A = (v_1 | \ldots | v_m)$, dann wissen wir nach Proposition IV.2.12 dass $m = \mathrm{Rang}(A) = n$ gelten muss, und wir sind fertig. □

Satz 8 (Koordinatenabbildung): *Seien K ein Körper und V ein K-Vektorraum.*

(i) *Ist B eine Basis von V, dann ist $V \cong \mathrm{Abb}_0(B, K)$.*

(ii) *Ist $B = \{v_1, \ldots, v_n\}$ endlich, dann ist $V \cong K^n$.*

Beweis: (i) Die Abbildung

$$\Lambda \colon \mathrm{Abb}_0(B, K) \longrightarrow V, \qquad \lambda \longmapsto \sum_{w \in B} \lambda(b) b$$

ist linear wegen der Rechenregeln für endliche Summen und da die Vektorraumstruktur auf $\mathrm{Abb}_0(B, K)$ durch die punktweisen Verknüpfungen (punktweise Addition von Funktionen und punktweise Skalarmultiplikation für Funktionen) gegeben ist. Nach Voraussetzung gilt $\mathrm{Lin}(B) = V$, d. h. Λ ist surjektiv. Da B linear unabhängig ist und damit $\mathrm{Kern}(\Lambda) = \{\mathbf{0}\}$ gilt, ist Λ außerdem injektiv. Insgesamt ist Λ also ein Isomorphismus.

(ii) Nach (i) ist $V \cong \mathrm{Abb}_0(B, K) \cong \mathrm{Abb}_0(\{e_1, \ldots, e_n\}, K) \cong K^n$, wobei $\{e_1, \ldots, e_n\}$ die Standardbasis des K^n ist. □

Definition IV.2.14: Seien K ein Körper, n eine natürliche Zahl, V ein Vektorraum über K und $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Dann hat jede Basis von V n Elemente und wir nennen $\dim(V) := n$ die *Dimension von V*. In diesem Fall heißt V *endlichdimensional*.

Beweis: (i) Durch Satz 8 können wir uns auf das Resultat aus Proposition IV.2.13 zurückziehen. □

Satz 9 (Kriterium II für Basen): *Seien K ein Körper, V ein K-Vektorraum und $B \subseteq V$ eine Teilmenge. Dann sind äquivalent:*

(i) *Die Menge B ist eine Basis.*

(ii) *Die Menge B ist eine bezüglich Inklusion maximale linear unabhängige Teilmenge von V, d. h. B ist linear unabhängig und ist M eine weitere Teilmenge von V mit $B \subsetneq M$, dann ist M nicht linear unabhängig.*

(iii) *Die Menge B ist ein bezüglich Inklusion minimales Erzeugendensysem von V, d. h. $\mathrm{Lin}(B) = V$ und ist M' eine echte Teilmenge von B, dann ist $\mathrm{Lin}(M') \subsetneq V$.*

Beweis: „(i) $\Longrightarrow$ (ii)“: Da B nach Voraussetzung eine Basis ist, ist B insbesondere linear unabhängig. Ist M eine Teilmenge von V mit $B \subsetneq M$, dann gibt es $v \in M - B$. Da B eine Basis ist, gibt es $\lambda \in \mathrm{Abb}_0(B, K)$, sodass $v = \sum_{w \in B} \lambda(w)w$. Nun erklären wir eine Abbildung $\lambda' \colon M \to K$ mit endlichem Träger durch

$$\lambda'(w) = \begin{cases} \lambda(w), & \text{falls } w \in B, \\ -1, & \text{falls } w = v, \\ 0, & \text{sonst,} \end{cases}$$

und finden $\sum_{w \in M} \lambda'(w)w = \sum_{w \in B} \lambda(w)w - 1v = \mathbf{0}$, d. h. M ist linear abhängig.

„(ii) $\Longrightarrow$ (iii)“: Sei B eine bezüglich Inklusion maximale linear unabhängige Teilmenge von V. Wir wollen zeigen, dass dann schon $\mathrm{Lin}(B) = V$ und dass B ein minimales Erzeugendensystem von V ist.

Um zu zeigen, dass $\mathrm{Lin}(B) = V$, sei $v \in V$ gegeben. Gehört v zu B, dann auch zu $\mathrm{Lin}(B)$. Gehört v nicht zu B, setze $M := B \cup \{v\}$. Da B maximal linear unabhängig ist und $B \subsetneq M$ gilt, muss M linear abhängig sein. Es gibt also $\mathbf{0} \neq \lambda \in \mathrm{Abb}_0(M, K)$, sodass $\sum_{w \in M} \lambda(w)w = \mathbf{0}$. Für dieses λ muss gelten, dass $\alpha := \lambda(v) \neq 0$, denn sonst wäre B bereits linear abhängig. Wir dürfen also durch α teilen und erhalten

$$\begin{aligned} \sum_{w \in M} \lambda(w)w = \mathbf{0} &\Longrightarrow \frac{1}{\alpha}\Big(\sum_{w \in M} \lambda(w)w \Big) = \mathbf{0} \\ &\Longrightarrow v = -\sum_{w \in B} \frac{\lambda(w)}{\alpha} w \\ &\Longrightarrow v \in \mathrm{Lin}(B). \end{aligned}$$

Gäbe es eine echte Teilemenge M' von B mit $\mathrm{Lin}(M') = B$, dann erhielten wir in „(i) $\Longrightarrow$ (ii)“, dass B linear abhängig sein müsste. Ein Widerspruch!

„(iii) $\Longrightarrow$ (i)“: Sei B ein minimales Erzeugendensystem von V. Wir wollen zeigen, dass B dann auch linear unabhängig sein muss.

Angenommen B wäre linear abhängig. Dann gäbe es $\mathbf{0} \neq \lambda \in \mathrm{Abb}_0(B, K)$, sodass $\sum_{w \in B} \lambda(w)w = \mathbf{0}$. Da $\lambda \neq \mathbf{0}$, gäbe es $v_0 \in B$ mit $\lambda(v_0) \neq 0$. Für dieses v_0 fänden wir dann wie in „(ii) $\Longrightarrow$ (iii)“, dass

$$v_0 = -\sum_{w \in B - \{v_0\}} \frac{\lambda(w)}{\lambda(v_0)} w,$$

sodass $\mathrm{Lin}(B - \{v_0\}) = V$ folgte. Ein Widerspruch! □

Korollar IV.2.15: *Seien K ein Körper, V ein K-Vektorraum und M eine nichtleere Teilmenge von V.*

(i) *Ist M linear abhängig, dann gibt es $v \in M$ mit $v \in \mathrm{Lin}(M - \{v\})$. Insbesondere gilt $\mathrm{Lin}(M) = \mathrm{Lin}(M - \{v\})$.*

(ii) *Ist M linear unabhängig und ist $v \in V - \mathrm{Lin}(M)$, dann ist auch $M \cup \{v\}$ linear unabhängig.*

Beweis: Aussage (i) folgt aus dem Beweis von „(iii) $\Longrightarrow$ (i)“ im Beweis von Satz 9, Aussage (ii) folgt aus dem Beweis von „(ii) $\Longrightarrow$ (iii)“ im Beweis von Satz 9. □

Satz 10 (Basisergänzungssatz): *Seien K ein Körper und V ein K-Vektorraum. Hat V ein endliches Erzeugendensystem. Dann gilt:*

(i) *Der Vektorraum V hat eine Basis.*

(ii) *Jedes Erzeugendensystem von V enthält eine Basis von V.*

(iii) *Jede linear unabhängige Teilmenge von V lässt sich durch Hinzunahme endlich vieler Elemente zu einer Basis ergänzen.*

Beweis: (i) Können wir (ii) zeigen, dann gibt es (i) gratis.

(ii) Wegen Proposition IV.2.15(i) und der Endlichkeit des Erzeugendensystems haben wir nach endlich vielen Rauswürfen ein bezüglich Inklusion minimales Erzeugendensystem von V. Nach Satz 9 ist das eine Basis von V.

(iii) Nach (ii) hat V eine endliche Basis. Nach Proposition IV.2.12 und Satz 9 ist jede linear unabhängige Teilmenge von V insbesondere endlich und mithilfe von Proposition IV.2.15(ii) erhalten wir durch Hinzunahme von endlich vielen Vektoren von V eine Basis von V. □

Definition IV.2.16: Sei V ein Vektorraum über K. Besitzt V ein endliches Erzeugendensystem, dann heißt V *endlichdimensional*, sonst *unendlichdimensional*. Ist V endlichdimensional und ist B irgendeine Basis von V, dann nennen wir $\dim_K(V) := \#(B)$ die *Dimension von V über K*.

3. Lineare Fortsetzung und Abbildungsmatrix

Wegen der Rechenregeln für endliche Summen sind lineare Abbildungen dadurch eindeutig festgelegt, was sie auf einer Basis des Startvektorraumes tun. Wir werden sehen, dass uns das erlaubt, die Wirkung linearer Abbildungen zwischen endlichdimensionalen Vektorräumen durch Matrizen zu beschreiben.

Satz 11 (Fortsetzungssatz): *Seien K ein Körper, V und W zwei K-Vektorräume und B eine Basis von V.*

(i) *Jede lineare Abbildung $\phi\colon V \to W$ ist eindeutig durch ihre Einschränkung $f := \phi|_B\colon B \to W$ bestimmt.*

(ii) *Jede Abbildung $f\colon B \to W$ lässt sich auf genau eine Weise zu einer linearen Abbildung $\phi\colon V \to W$ fortsetzen, d. h. es gibt genau einen Vektorraumhomomorphismus $\phi\colon V \to W$, sodass $\phi|_B = f$. Dieser heißt* lineare Fortsetzung von f.

Beweis: (i) Sei v ein Element von V. Da B eine Basis von V ist, gibt es $\lambda_v \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{b\in B} \lambda_v(b)b$ und wir erhalten

$$\phi(v) = \phi\Big(\sum_{b\in B} \lambda_v(b)b\Big) = \sum_{b\in B} \lambda_v(b)\phi(b) = \sum_{b\in B} \lambda_v(b)f(b).$$

(ii) Für jedes Element v von V gibt es $\lambda_v \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{b\in B} \lambda_v(b)b$. Deshalb können wir $\phi\colon V \to W$ definieren durch

$$\Phi\colon V \longrightarrow W, \qquad v \longmapsto \sum_{b\in B} \lambda_v(b)b.$$

Wegen der Rechenregeln für endliche Summen liefert das eine lineare Abbildung. Die Eindeutigkeit dieser Festsetzung folgt aus (i). □

Korollar IV.3.1: *Seien K ein Körper, V ein K-Vektorraum und B eine Basis von V. Dann ist die Abbildung*

$$H\colon \mathrm{Hom}_K(V, W) \longrightarrow \mathrm{Abb}_0(B, W), \qquad \phi \longmapsto \phi|_B$$

ein Isomorphismus von K-Vektorräumen.

Beweis: Nach Satz 11 ist H bijektiv und wegen der punktweisen Verknüpfungen und der Rechenregeln für endliche Summen ist H linear. □

Wir haben in Satz 8 gesehen: Ist V ein Vektorraum über de Körper K mit Basis $B = \{b_1, \ldots, b_n\}$, dann erhalten wir einen Isomorphismus

$$\Lambda\colon K^n \longrightarrow V, \qquad (x_1, \ldots, x_n)^t \longmapsto \sum_{i=1}^{n} x_i b_i.$$

Für das, was folgt, wird die Reihenfolge der Elemente der Basis eine Rolle spielen, weswegen wir *geordnete Basen* $B = (b_1, \ldots, b_n)$ betrachten wollen.

Sind n und m natürliche Zahlen und $A \in K^{n\times m}$ eine Matrix, dann schreiben wir $L_A\colon K^m \to K^n$, $x \mapsto Ax$.

Satz 12 (Darstellungsmatrix): *Seien K ein Körper, V und W Vektorräume über K mit geordneten Basen $B = (b_1, \ldots, b_m)$ respektive $C = (c_1, \ldots, c_n)$ und $\phi\colon V \to W$ eine lineare Abbildung. Dann gibt es genau eine Matrix $A \in K^{n\times m}$, sodass*

$$D_C \circ \phi = L_A \circ D_B.$$

Die Einträge der Matrix $A = (a_{ij})$ sind bestimmt durch $a_{ij} = \lambda_{ij}$, wobei $\phi(b_j) = \sum_{i=1}^{n} \lambda_{ij} c_i$ für $1 \le j \le m$. Wir schreiben $D_{C,B}(\phi) := A$ und nennen diese Matrix die Darstellungsmatrix von ϕ bezüglich B und C.

Beweis: Wir skizziere zunächst die Situation für lineare Abbildungen $\psi\colon K^m \to K^m$: Die Darstellungsmatrix von ψ bezüglich der Standardbasen ist die Matrix $A = (\psi(e_1)|\ldots|\psi(e_m))$, denn Ae_i liefert die i-te Spalte von A, d. h. für $v = \sum_{i=1}^{m} v_i e_i$ haben wir deshalb

$$\psi(v) = \psi\Big(\sum_{i=1}^{m} v_i e_i\Big) = \sum_{i=1}^{m} v_i \psi(e_i) = \sum_{i=1}^{m} v_i A e_i = A\Big(\sum_{i=1}^{m} v_i e_i\Big) = A(v)$$

und somit $L_A = \psi$. Genau das wollen wir auch für $\phi\colon V \to W$ erreichen.

Da wir aber auf der Ebene von V und W keine Matrizen zur Verfügung haben, müssen wir in die Koordinatenvektorräume K^m bzw. K^n übersetzen. Das machen wir bei fixierten geordneten Basen mittels Koordinatenabbildungen, um schließlich die Matrix $A \in K^{m\times n}$ zu finden, für die $L_A = D_C \circ \phi \circ D_B^{-1}$.

Dadurch, dass $D_B(b_i) = e_i$ für $1 \le i \le m$, dass $L_A(D_B(b_i))$ die i-te Spalte von A liefert und wir $D_C \circ \phi = L_A \circ D_B$ erreichen wollen, wissen wir, dass die i-te Spalte von A aus den Koordinaten von $\phi(b_i)$ bezüglich C bestehen muss. Aber genau das sind die Gleichungen, die wir für die Einträge der Matrix A bereits angegeben haben. □

Bemerkung IV.3.2 (Definierende Gleichung der Abbildungsmatrix): In der Situation von 12 gilt insbesondere für alle v in V

$$D_C(\Phi(v)) = D_{C,B}(\Phi)D_B(v).$$

Beispiel IV.3.3: Seien $V = W = \mathbb{R}^2$ und $\Phi\colon V \to V$ die Spiegelung an der Diagonalen $y = x$. Betrachte die geordnete Basis

$$B = \left(b_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, b_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$$

Schreiben wir $v \in \mathbb{R}^2$ als $v = \lambda_1 b_1 + \lambda_2 b_2$, dann ist $\Phi(v) = \lambda_1 b_1 - \lambda_2 b_2$. Die Linearität von Φ ist jetzt offensichtlich. Ferner ist

$$D_B(\Phi(v)) = \begin{pmatrix} \lambda_1 \\ -\lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} D_B(v).$$

Damit lesen wir die Darstellungsmatrix ab als

$$D_{B,B}(\Phi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Proposition IV.3.4 (Darstellungsmatrix und Verkettung): *Seien V_1, V_2, V_3 K-Vektorräume mit Dimensionen n_1, n_2, n_3, $\Phi\colon V_1 \to V_2$ und $\Psi\colon V_2 \to V_3$ lineare Abbildungen und B_1, B_2, B_3 Basen von V_1, V_2 und V_3. Dann gilt für die zugehörigen Abbildungsmatrizen. Dann gilt*

$$D_{B_3,B_1}(\Psi \circ \Phi) = D_{B_3,B_2}(\Psi)D_{B_2,B_1}(\Phi).$$

Beweis: Wir betrachten das folgende Diagramm von Abbildungen

$$\begin{array}{ccccc} V_1 & \xrightarrow{\Phi} & V_2 & \xrightarrow{\Psi} & V_3 \\ {\scriptstyle D_{B_1}}\downarrow & & {\scriptstyle D_{B_2}}\downarrow & & \downarrow{\scriptstyle D_{B_3}} \\ K^{n_1} & \xrightarrow[\ell_1]{} & K^{n_2} & \xrightarrow[\ell_2]{} & K^{n_3} \end{array}$$

mit den linearen Abbildungen

$$\begin{aligned} \ell_1\colon K^{n_1} &\longrightarrow K^{n_2}, & x &\longmapsto D_{B_2,B_1}(\Phi)x, \\ \ell_2\colon K^{n_2} &\longrightarrow K^{n_3}, & x &\longmapsto D_{B_3,B_2}(\Psi)x. \end{aligned}$$

Es gilt $(D_{B_3} \circ \Psi \circ \Phi)(x) = (\ell_2 \circ D_{B_2} \circ \Phi)(x) = (\ell_2 \circ l_1 \circ D_{B_1})(x)$. Nun halten wir fest, dass $(\ell_2 \circ \ell_1)(v) = D_{B_3,B_2}(\Psi)D_{B_2,B_1}(\Phi)v$. Wegen Satz 12 erhalten wir darum $D_{B_3,B_1}(\Psi \circ \Phi) = D_{B_3,B_2}(\Psi)D_{B_2,B_1}(\Phi)$. □

Definition IV.3.5 (Basiswechselmatrix): Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum und $B = (b_1, \ldots, b_n)$, $B' = (b'_1, \ldots, b'_n)$ geordnete Basen von V. Dann heißt $D_{B',B} := D_{B',B}(\mathrm{id})$ *Basiswechselmatrix von B nach B'.*

Auch die Bezeichnung *Koordinatentransformationsmatrix von B nach B'* ist gebräuchlich. Das liegt daran, dass für alle $v \in V$ gilt, dass

$$D_{B'}(v) = D_{B',B} D_B(v).$$

Da die Einträge $\lambda_{i,j}$ der Basiswechselmatrix $D_{B',B}$ bestimmt sind durch die Gleichungen $b_j = \sum_{i=1}^{n} \lambda_{i,j} b'_i$ für $1 \leq j \leq n$, geben die Einträge dieser Matrix aber gleichzeitig an, wie die Basis B aus der Basis B' hervorgeht. Deshalb wird diese Matrix gelegentlich auch *Basiswechselmatrix von B' nach B genannt.*

Proposition IV.3.6: *Seien K ein Körper, V und W Vektorräume über K mit geordneten Basen $B = (b_1, \ldots, b_m)$ und $C = (c_1, \ldots, c_n)$ und sei $\Phi\colon V \to W$ eine lineare Abbildung. Ferner seien B' und C' weitere geordnete Basen. Dann gilt*

$$D_{C',B'}(\Phi) = D_{C',C} D_{C,B}(\Phi) D_{B,B'}.$$

Beweis: Da im folgenden Diagramm jedes der Quadrate kommutiert, können wir die Situation mit dem folgenden kommutativen Diagramm abbilden:

$$\begin{array}{ccccccc}
V & \xrightarrow{\mathrm{id}} & V & \xrightarrow{\phi} & W & \xrightarrow{\mathrm{id}} & W \\
{\scriptstyle D_{B'}}\downarrow & & {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} & & \downarrow{\scriptstyle D_{C'}} \\
K^m & \xrightarrow[L_{D_{B,B'}}]{} & K^m & \xrightarrow[L_{D_{C,B}(\phi)}]{} & K^n & \xrightarrow[L_{D_{C',C}}]{} & K^n
\end{array}$$

Wegen $\mathrm{id} \circ \phi \circ \mathrm{id} = \phi$ beschreiben die äußeren Pfeile das Diagramm der Darstellungsmatrix $D_{C',B'}(\phi)$. Die Komposition der Pfeile in der unteren Zeile liefert die Abbildung $x \mapsto D_{C',C} D_{C,B}(\phi) D_{B,B'} x$, was wir behauptet haben. □

Satz 13 (Basiswechsel und Darstellungsmatrizen): *Sei K ein Körper.*

(i) *Seien $\phi\colon V_1 \to V_2$ und $\psi\colon V_2 \to V_3$ lineare Abbildungen zwischen endlichdimensionalen K-Vektorräumen und seien B_1, B_2 und B_3 geordnete Basen der jeweiligen Vektorräume. Dann gilt*

$$D_{B_3,B_1}(\psi \circ \phi) = D_{B_3,B_2}(\psi) D_{B_2,B_1}(\phi).$$

(ii) *Sei V ein endlichdimensionaler Vektorraum über K mit geordneten Basen B und B'. Dann ist die Basiswechselmatrix $D_{B',B}$ regulär und für die Inverse gilt $D_{B',B}^{-1} = D_{B,B'}$.*

(iii) *Für $V = K^n$ und die Standardbasis $E = (e_1, \dots, e_n)$ von K^n und eine weitere geordnete Basis $B = (b_1, \dots, b_n)$ gilt $D_{E,B} = (b_1 | \dots | b_n)$.*

(iv) *Für geordnete Basen B und B' von K^n gilt*

$$D_{B',B} = D_{B',E} D_{E,B} = D_{E,B'}^{-1} D_{E,B}.$$

Beweis: (i) Die Situation können wir im Diagramm

$$\begin{array}{ccccc} V_1 & \xrightarrow{\phi} & V_2 & \xrightarrow{\psi} & V_3 \\ {\scriptstyle D_{B_1}}\downarrow & & {\scriptstyle D_{B_2}}\downarrow & & \downarrow{\scriptstyle D_{B_3}} \\ K^{n_1} & \xrightarrow[L_A]{} & K^{n_2} & \xrightarrow[L_B]{} & K^{n_3} \end{array}$$

einfangen, wobei $A = D_{B_2,B_1}(\phi)$, $B = D_{B_3,B_2}(\phi)$, $n_1 = \dim V_1$, $n_2 = \dim V_2$ und $n_3 = \dim V_3$. Für die Matrix $C := D_{B_3,B_1}(\psi \circ \phi)$ gilt jetzt

$$L_C = D_{B_3} \circ \psi \circ \phi \circ D_{B_1}^{-1} = L_B \circ L_A,$$

also folgt die Behauptung.

(ii) Das folgt aus (i) für $\psi = \phi = \mathrm{id}$ und $B_1 = B$, $B_2 = B'$, $B_3 = B$.

(iii) Mit $B' = E = (e_1, \dots, e_n)$ erhalten wir die bestimmenden Gleichungen $b_j = \sum_{i=1}^n \lambda_{i,j} e_i$ für die Basiswechselmatrix, d. h. die $\lambda_{i,j}$ sind genau die Koordinaten von b_j bezüglich der Standardbasis.

(iv) Folgt aus (i) und (ii). □

4. Summen von Unterräumen und Faktorräume

Seien K ein Körper und V ein Vektorraum über K. Sind $U_1, \dots, U_n$ Untervektorräume von V, dann haben wir bereits gesehen, dass

$$\sum_{i=1}^n U_i := U_1 + \dots + U_n := \{x_1 + \dots + x_n \mid x_1 \in U_1, \dots, x_n \in U_n\} \subseteq V$$

auch ein Untervektorraum von V ist. Wir nennen $U_1 + \dots + U_n$ die *Summe von* $U_1, \dots, U_n$.

Definition IV.4.1 (Direkte Summe): Seien K ein Körper, V ein Vektorraum über K und $U_1, \dots, U_n$ Untervektorräume von V. Falls gilt: „Wenn immer $u_1 \in U_1, \dots, u_n \in U_n$ mit $u_1 + \cdots + u_n = \mathbf{0}$, dann sind $u_1 = \cdots = u_n = \mathbf{0}$“, dann heißt die Summe $U_1 + \cdots + U_n$ *direkt*. In diesem Fall schreiben wir $\bigoplus_{i=1}^n U_i := \sum_{i=1}^n U_i$.

Bemerkung IV.4.2: Seien K ein Körper, V ein K-Vektorraum und $U_1, \dots U_n$ Untervektorräume von V. Ist $\sum_{i=1}^n U_i$ direkt, dann haben wir für $i, j \in \{1, \dots, n\}$ mit $i \neq j$, dass $U_i \cap U_j = \{\mathbf{0}\}$. Das sieht man so: Für $v \in U_1 \cap U_2$ haben wir $v - v + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$, d. h. $v = \mathbf{0}$ per Definition der direkten Summe. Die Umkehrung dieser Aussage gilt nicht! Sind beispielsweise $V = \mathbb{R}^2$, $U_1 = \langle (1,0)^t \rangle$, $U_2 = \langle (0,1)^t \rangle$ und $U_3 = \langle (1,1)^t \rangle$, dann haben wir zwar $U_i \cap U_j = \{\mathbf{0}\}$ für $1 \leq i, j \leq 3$, $i \neq j$, aber

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Satz 14 (über die direkte Summe): *Seien K ein Körper, V ein Vektorraum über K und $U_1, \dots, U_n$ Untervektorräume von V.*

(i) *Seien $B_1, \dots, B_n$ Basen von $U_1, \dots, U_n$. Ist die Summe der U_i direkt, dann ist $B := B_1 \cup \cdots \cup B_n$ eine Basis von $\bigoplus_{i=1}^n U_i$.*

(ii) *Seien $U_1, \dots, U_n$ endlichdimensional. Genau dann ist die Summe der U_i direkt, wenn $\dim(\sum_{i=1}^n U_i) = \sum_{i=1}^n \dim U_i$.*

Beweis: (i) Per Definition von $\bigoplus_{i=1}^n U_i$ ist B ein Erzeugendensystem. Bleibt also die lineare Unabhängigkeit zu zeigen. Sei dazu $\lambda \in \mathrm{Abb}_0(B, K)$ mit $\mathbf{0} = \sum_{b \in B} \lambda(b) b$. Setze $u_i := \sum_{b \in B_i} \lambda(b) b$. Dann ist $u_1 + \cdots + u_n = \mathbf{0}$ und da die Summe direkt ist, erzwingt das $u_1 = \cdots = u_n = \mathbf{0}$. Damit muss λ die Nullabbildung sein und B ist linear unabhängig.

(ii) „$\Longrightarrow$“ folgt aus (i). Zu „$\Longleftarrow$“: Da die U_i endlichdimensional sind, hat jeder der Vektorräume eine Basis, sagen wir $B_i \subseteq U_i$. Setze $B := \bigcup_{i=1}^n B_i$. Dann haben wir

$$\#B \leq \sum_{i=1}^n \#B_i = \sum_{i=1}^n \dim U_i = \dim\Big(\sum_{i=1}^n U_i\Big)$$

Da B ein Erzeugendensystem von $\sum_{i=1}^n U_i$ ist, gilt $\dim(\sum_{i=1}^n U_i) \leq \#B$. Nun liefert Satz 9, dass B eine Basis von $\sum_{i=1}^n U_i$ ist und die lineare Unabhängigkeit von B liefert die Direktheit der Summe. □

Definition IV.4.3 (Äquivalenz modulo Unterraum): Seien K ein Körper, V ein Vektorraum über K und $U \subseteq V$ ein Untervektorraum. Auf V wird durch

$$v_1 \sim v_2 :\Longleftrightarrow v_1 - v_2 \in U$$

eine Äquivalenzrelation, genannt *Äquivalenz modulo U*, erklärt. Für $v \in V$ bezeichnet

$$[v] := \{w \in V \mid v \sim w\} = \{w \in V \mid v - w \in U\} =: v + U$$

die Äquivalenzklasse von v und $V/U := V/\sim = \{[v] \mid v \in V\}$ bezeichnet die Menge der Äquivalenzklassen bezüglich Äquivalenz modulo U.

Proposition IV.4.4 (Quotient nach Unterraum): *Seien K ein Körper, V ein Vektorraum über K, $U \subseteq V$ ein Unterraum und V/U die Menge der Äquivalenzklassen bezüglich Äquivalenz modulo U. Auf V/U wird duch*

$$[v] + [w] := [v + w], \qquad \lambda[v] := [\lambda v]$$

eine K-Vektorraumstruktur erklärt. Zusammen mit dieser heißt V/U der Quotient von V nach U *oder* Faktorraum V/U. *Die Abbildung $\pi\colon V \to V/U$, $v \mapsto [v]$ heißt* kanonische Projektion. *Die kanonische Projektion ist surjektiv, linear, und erfüllt* $\ker \pi = U$.

Beweis: Sobald wir uns davon überzeugt haben, dass die oben angegebenen Verknüpfungen wohldefiniert sind, sehen wir sofort dass V/U ein K-Vektorraum ist, da wir repräsentantenweise rechnen und wir wissen, dass V ein Vektorraum über K ist. Für die Wohldefiniertheit ist die Unabhängigkeit von der Wahl der Repräsentanten zu prüfen.

Per Konstruktion ist die kanonische Projektion π surjektiv. Sind v und w Elemente von V, sowie α in K, dann haben wir

$$\pi(v + \alpha w) = [v + \alpha w] = [v] + [\alpha w] = [v] + \alpha[w] = \pi(v) + \alpha\pi(w),$$

also ist π linear. Wegen $[0] = 0 + U = U$ ist gerade $\ker \pi = U$. □

Satz 15: *Es seien K ein Körper, V und W Vektorräume über K, $\varphi\colon V \to W$ eine lineare Abbildung und $U \subseteq V$ ein Untervektorraum mit $U \subseteq \mathrm{Kern}(\varphi)$. Dann gibt es genau eine lineare Abbildung $\bar{\varphi}\colon V/U \to W$, die das Diagramm*

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U \\ & \searrow^{\varphi} & \downarrow{\scriptstyle \bar{\varphi}} \\ & & W \end{array}$$

kommutativ macht, d. h. $\bar{\varphi} \circ \pi = \varphi$*. Sind sogar* $U = \operatorname{Kern}(\varphi)$ *und* $W = \operatorname{Bild}(\varphi)$*, dann ist* $\bar{\varphi}$ *injektiv (und per Konstruktion surjektiv), d. h.* $V/\operatorname{Kern}(\varphi) \cong \operatorname{Bild}(\varphi)$ *vermöge* $\bar{\varphi}$*.*

Beweis: Wegen $\bar{\varphi} \circ \pi = \varphi$ haben wir keine andere Wahl, als zu definieren: $\bar{\varphi}([v]) := \varphi(v)$. Jetzt haben wir zu überprüfen, dass $\bar{\varphi}$ wohldefiniert ist, d. h. dass alle $w \in [v]$ unter φ dasselbe Bild haben.

Ist $U = \operatorname{Kern}(\varphi)$, dann ist $\bar{\varphi}$ injektiv. Wegen $\operatorname{Bild}(\bar{\varphi}) = \operatorname{Bild}(\varphi)$ ist $\bar{\varphi}$ auch surjektiv, d. h. $\bar{\varphi}\colon V/U \to \operatorname{Bild}(\varphi)$ ist ein Isomorphismus. □

Proposition IV.4.5 (Basis des Faktorraums): *Seien* K *ein Körper,* V *ein Vektorraum über* K *und* $U \subseteq V$ *ein Untervektorraum. Sind* $B' \subseteq U$ *eine Basis und* $B \subseteq V$ *eine Basis von* V*, die* B' *enthält, dann ist*

$$C := \{[b] = b + U \mid b \in B - B'\}$$

eine Basis von V/U*.*

Beweis: Zunächst zeigen wir, dass $\operatorname{Lin}(C) = V/U$. Sei dazu $v \in V$ gegeben. Da B eine Basis von V ist, gibt es $\lambda \in \operatorname{Abb}_0(B, K)$ mit $v = \sum_{b\in B} \lambda(b)b$, d. h.

$$[v] = \Big[\sum_{b\in B} \lambda(b)b\Big] = \sum_{b\in B-B'} \lambda(b)[b] + \sum_{b\in B'} \lambda(b)[b] = \sum_{b\in B-B'} \lambda(b)[b].$$

Nun zur linearen Unabhängigkeit: Sei $\lambda \in \operatorname{Abb}_0(B - B', K)$ gegeben, sodass $\sum_{b\in B-B'} \lambda([b])[b] = [0]$. Setze $u := \sum_{b\in B-B'} \lambda(b)b$. Dann ist $[u] = [0]$, d. h. u gehört zu U. Weil B' eine Basis von U ist, gibt es also $\lambda_u \in \operatorname{Abb}(B', K)$, sodass $u = \sum_{b\in B'} \lambda_u(b)b$ und damit ist

$$\sum_{b\in B-B'} \lambda(b)b - \sum_{b\in B'} \lambda_u(b)b = \mathbf{0}.$$

Da B eine Basis ist, muss nun λ die Nullabbildung sein. □

Satz 16 (Dimensionsformel): *Seien* K *ein Körper und* V *ein endlichdimensionaler Vektorraum über* K *mit* $\dim V = n$*.*

(i) *Ist* $U \subseteq V$ *ein Untervektorraum, dann ist* $\dim V/U = \dim V - \dim U$*.*

(ii) *Sind* U_1 *und* $U_2 \subseteq V$ *Untervektorräume, dann ist*

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

(iii) *Ist W ein weiterer K-Vektorraum und $\phi\colon V \to W$ linear, dann ist*

$$\dim V = \dim \operatorname{Kern} \phi + \dim \operatorname{Bild} \phi.$$

Beweis: (i) In Proposition IV.4.5 haben wir gezeigt, wie man eine Basis von V/U erhalten kann. Insbesondere haben wir die Dimension von V/U bestimmt.

(ii) Wir möchten den Homomorphiesatz anwenden, um die Behauptung zu zeigen. Dazu suchen wir uns eine geeignete surjektive lineare Abbildung mit dem richtigen Kern, nämlich

$$\alpha\colon U_1 \times U_2 \longrightarrow U_1 + U_2, \qquad (u_1, u_2) \longmapsto u_1 - u_2.$$

Ist $u_1 + u_2 \in U_1 + U_2$ vorgegeben, dann ist $\alpha(u_1, -u_2) = u_1 + u_2$, d. h. α ist surjektiv. Auch den Kern von α können wir leicht erkennen, der ist nämlich

$$\operatorname{Kern} \alpha = \{(u_1, u_2) \in U_1 \times U_2 \mid u_1 = u_2\}.$$

Wir haben somit einen Isomorphismus $U_1 \cap U_2 \to \operatorname{Kern} \alpha$, $u \mapsto (u, u)$. Schließlich können wir Dimensionen von $U_1 \times U_2$ und $U_1 + U_2$ miteinander in Verbindung bringen: Ist B_1 eine Basis von U_1 und ist B_2 eine Basis von U_2, dann erhalten wir durch $B := \{(b, 0) \mid b \in B_1\} \cup \{(0, b) \mid b \in B_2\}$ eine Basis von $U_1 \times U_2$ der Mächtigkeit $\dim U_1 + \dim U_2$. Mit (iii) erhalten wir jetzt

$$\begin{aligned} \dim U_1 + \dim U_2 &= \dim(U_1 \times U_2) \\ &= \dim \operatorname{Kern} \alpha + \dim \operatorname{Bild} \alpha = \dim(U_1 \cap U_2) + \dim(U_1 + U_2). \end{aligned}$$

(iii) Aus Satz 15 kennen wir die Isomorphie $\operatorname{Bild} \phi \cong V/\operatorname{Kern} \phi$. In den Übungen werden Sie zeigen, dass das bedeutet, dass beide Vektorräume die selbe Dimension haben müssen. Wir können mit (i) deshalb folgern:

$$\dim \operatorname{Bild} \phi = \dim(V/\operatorname{Kern} \phi) = \dim V - \dim \operatorname{Kern} \phi. \qquad \square$$

Definition IV.4.6 (Rang und Kern): Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, W ein weiterer K-Vektorraum und $\phi\colon V \to W$ linear. Dann heißt $\operatorname{Rang} \phi := \dim \operatorname{Bild} \phi$ der *Rang von ϕ*.

Seien n und m natürliche Zahlen, K ein Körper und $A \in K^{n \times m}$ gegeben. Dann heißt $\operatorname{Kern} A := \{x \in K^m \mid Ax = \mathbf{0}\}$ der *Kern der Matrix A*.

Satz 17 (Rang): *Seien K ein Körper, V und W endlichdimensionale Vektorräume über K und $\phi\colon V \to W$ linear.*

(i) *Es gilt* $\dim V = \dim \ker \phi + \operatorname{Rang} \phi$.

Haben wir geordnete Basen $B = (b_1, \ldots, b_m)$ *von* V *und* $C = (c_1, \ldots, c_n)$ *von* W*, und ist* $A := D_{C,B}(\phi)$*, dann gilt:*

(ii) *Der Kern von* ϕ *ist isomorph zum Kern der Matrix* A*, d. h. beide Definitionen von „Kern“ passen zueinander.*

(iii) *Bezeichnet* $\{e_1, \ldots, e_m\} \subseteq K^m$ *die Standardbasis, dann ist das Bild von* ϕ *isomorph zu* $\operatorname{Lin}(Ae_1, \ldots, Ae_m)$*, d. h. der linearen Hülle der Spalten der Darstellungsmatrix von* A*.*

(iv) *Die Ränge von* A *und* φ *stimmen überein, also* $\operatorname{Rang} A = \operatorname{Rang} \phi$.

(v) *Bezeichnen* $s_1, \ldots, s_m$ *die Spalten und* $z_1, \ldots, z_n$ *die Zeilen von* A*, dann ist*

$$\operatorname{Rang} A = \dim \operatorname{Lin}(s_1, \ldots, s_m) = \dim \operatorname{Lin}(z_1, \ldots, z_n).$$

Beweis: (i) Wegen der Definition des Ranges von ϕ ist das genau Satz 16(iii).

(ii) Per Definition der Darstellungsmatrix kommutiert das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} \\ K^m & \xrightarrow[\varphi_A]{} & K^n \end{array}$$

Der Untervektorraum $\ker \phi$ von V wird vom Isomorphismus D_B auf einen Untervektorraum U von K^n abgebildet. Für ein Element v des Kerns von ϕ gilt wegen der Kommutativität des Diagramms, dass

$$\mathbf{0}_{K^n} = D_C(\phi(v)) = \varphi_A(D_B(v)) = AD_B(v),$$

sodass $U \subseteq \ker \varphi_A$. Weil die Koordinatenabbildungen Isomorphismen sind, haben wir außerdem $D_C \circ \phi \circ D_B^{-1} = \varphi_A$. Für ein w aus $\ker \phi_A$ ist darum $\mathbf{0}_{K^n} = \varphi_A(w) = D_C(\phi(D_B^{-1}(w)))$, was wegen der Injektivität der Koordinatenabbildungen zeigt, dass $D_B^{-1}(v)$ zu $\ker \phi$ gehört. Es folgt $D_B^{-1}(\ker \varphi_A) \subseteq \ker \phi$ und so insgesamt $D_B(\ker \phi) = \ker \varphi_A$. Insbesondere ist $\ker \phi \cong \ker \varphi_A$.

(iii) Genau wie (ii).

(iv) Wie wir wissen, ist $\operatorname{Rang} A = m - \dim \operatorname{Kern} A$. Nach Aussage (iii) ist $\dim \operatorname{Kern} A = \dim \operatorname{Kern} \phi$, sodass $\operatorname{Rang} A = m - \dim \operatorname{Kern} \phi$. Wegen (i) ist das aber genau $\operatorname{Rang} \phi$.

(v) Das Bild der linearen Abbildung $\phi_A\colon K^m \to K^n$, $x \mapsto Ax$ ist das Erzeugnis der Spalten von A, sodass (iv) liefert: $\operatorname{Rang} A = \operatorname{Rang} \phi_A = \dim\langle s_1, \ldots, s_m\rangle$.

Sei nun T die Treppenform von A. Die Treppenform von A entsteht aus A durch Zeilenoperationen, genauer: Es gibt elementare Zeilenumformungen $Z_1, \ldots, Z_N$ (d. h. $Z_k = A^\alpha_{i,j}$, oder $Z_k = V_{i,j}$ oder $Z_k = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$, wobei α, $\alpha_1, \ldots, \alpha_n \in K^\times$), sodass $T = Z_1 \cdots Z_N \cdot A$.

Setze $A_k := Z_{k+1} \cdots Z_N \cdot A$. Für $A_{k+1} = Z_k \cdot A_k$ ist der Spann der Zeilenvektoren von A_k der Spann der Zeilenvektoren von A_{k+1}, d. h. das Erzeugnis der Zeilen von A ist gleich dem Erzeugnis der Zeilenvektoren von T. Aber dann ist auch $\dim \operatorname{Lin}(z_1, \ldots, z_n) = \operatorname{Rang} T = \operatorname{Rang} A$. □

Kapitel V.

Endomorphismen von Vektorräumen

1. Endomorphismen und Basiswechsel

Bemerkung V.1.1 (Basiswechsel): Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K mit geordneter Basis $B = (b_1, \ldots, b_n)$, $\phi\colon V \to V$ linear und $B' = (b'_1, \ldots, b'_n)$ eine weitere geordnete Basis von V. Setzen wir $A := D_{B,B}(\phi)$, $A' := D_{B',B'}(\phi)$ und $S := D_{B',B}$, dann liefert Proposition IV.3.6, dass

$$A' = D_{B',B} D_{B,B}(\phi) D_{B,B'} = SAS^{-1}.$$

Definition V.1.2 (Ähnlichkeit): Seien K ein Körper und V ein n-dimensionaler K-Vektorraum. Sind $A_1, A_2 \in K^{n\times n}$ gegeben und gibt es $S \in \mathrm{Gl}_n(K)$ mit $A_2 = SA_1S^{-1}$, dann heißen A_1 und A_2 *ähnlich*.

Bemerkung V.1.3: Zwei Matrizen $A_1, A_2 \in K^{n\times n}$ sind ähnlich genau dann, wenn sie Darstellungsmatrizen derselben linearen Abbildung ϕ, möglicherweise bezüglich unterschiedlicher Basen, sind.

Proposition V.1.4 (Rang als Ähnlichkeitsinvariante): *Seien K ein Körper, V ein n-dimensionaler K-Vektorraum und $A \in K^{n\times n}$. Der Rang von A ist eine* Ähnlichkeitsinvariante. *Das heißt: Ist $B \in K^{n\times n}$ ähnlich zu A, dann gilt* $\operatorname{Rang} A = \operatorname{Rang} B$.

2. Eigenwerte und Eigenvektoren

Definition V.2.1 (Eigenvektoren, Eigenwerte): Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, $\phi\colon V \to V$ linear und $A \in K^{n\times n}$.

(i) Sei λ ein Element von K. Gibt es $v \in V - \{\mathbf{0}\}$ mit $\phi(v) = \lambda v$, dann heißt λ ein *Eigenwert von ϕ zum Eigenvektor v*.
Gibt es $x \in K^n - \{\mathbf{0}\}$ mit $Ax = \lambda x$, dann heißt λ ein *Eigenwert von A zum Eigenvektor x*.

(ii) Für $\lambda \in K$ heißt

$$\operatorname{Eig}(\phi, \lambda) := \{v \in V \mid \phi(v) = \lambda v\}$$
$$\operatorname{Eig}(A, \lambda) := \{x \in K^n \mid Ax = \lambda x\}$$

Eigenraum zu ϕ respektive *Eigenraum zu A*.

(iii) Die Menge der Eigenwerte

$$\operatorname{Spec}\phi := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } \phi\}$$
$$\operatorname{Spec} A := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } A\}$$

heißt *Spektrum von ϕ* respektive *Spektrum von A*.

Bemerkung V.2.2: Seien K ein Körper, V ein n-dimensionaler Vektorraum über K mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear.

(i) Ist $A = D_{B,B}(\phi)$, dann gilt $\operatorname{Spec}\phi = \operatorname{Spec} A$. Ferner ist $v \in V - \{\mathbf{0}\}$ ein Eigenvektor von ϕ zum Eigenwert λ genau dann, wenn $x = D_B(v)$ ein Eigenvektor von A zum Eigenwert λ ist.

(ii) Ein $\lambda \in K$ gehört zu $\operatorname{Spec}\phi$ respektive $\operatorname{Spec} A$ genau dann, wenn $\operatorname{Eig}(\phi, \lambda) \neq \{\mathbf{0}\}$ respektive $\operatorname{Eig}(A, \lambda) \neq \{\mathbf{0}\}$.

(iii) Wir haben die Äquivalenzen

$$Av = \lambda v \iff (A - \lambda I_n)v = \mathbf{0} \iff v \in \operatorname{Kern}(A - \lambda I_n),$$

d. h. $\operatorname{Eig}(A, \lambda) = \operatorname{Kern}(A - \lambda I_n)$. Analog ist $\operatorname{Eig}(\phi, \lambda) = \operatorname{Kern}(\phi - \lambda \operatorname{id}_V)$. Insbesondere sind Eigenräume von A beziehungsweise Eigenräume von ϕ Untervektorräume von K^n beziehungsweise V.

Beispiel V.2.3: (i) Seien $\lambda_1, \ldots, \lambda_n \in K$ und $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Dann ist $\operatorname{Spec}(A) = \{\lambda_1, \ldots, \lambda_n\}$, außerdem sind die Eigenräume leicht anzugeben: $\operatorname{Eig}(A, \lambda_i) = \operatorname{Lin}(e_i)$.

(ii) Sei $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Diagonalen (siehe Proposition IV.3.3). Bezüglich der Basis $B = \{(1,1)^t, (1,-1)^t\}$ von $\mathbb{R}^2$ hat ϕ die Darstellungsmatrix

$$D_{B,B}(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

d. h. $\operatorname{Spec}\phi = \{\pm 1\}$

Definition V.2.4 (Diagonalisierbarkeit): Seien K ein Körper, V ein K-Vektorraum mit geordneter Basis $B = (b_1, \dots, b_n)$, $\phi\colon V \to V$ linear und $A \in K^{n\times n}$.

(i) Gibt es eine Basis B' von V und Elemente $\lambda_1, \dots, \lambda_n$ von K, sodass $D_{B',B'}(\phi) = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, dann heißt ϕ *diagonalisierbar*.

(ii) Gibt es $S \in \mathrm{Gl}_n(K)$ und $\lambda_1, \dots, \lambda_n \in K$ mit $SAS^{-1} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, dann heißt A *diagonalisierbar*.

Ein Endomorphismus ϕ ist diagnonalisierbar genau dann, wenn seine Darstellungsmatrix $D_{B,B}(\phi)$ diagonalisierbar ist. Das liegt daran, wie Ähnlichkeit und Basiswechsel zusammenpassen.

Proposition V.2.5 (Eigenräume bilden direkte Summe): *Es seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von ϕ und $\operatorname{Eig}(\phi, \lambda_1), \dots, \operatorname{Eig}(\phi, \lambda_k)$ die zugehörigen Eigenräume. Dann ist die Summe der Eigenräume eine direkte Summe, d. h.*

$$\sum_{i=1}^{k} \operatorname{Eig}(\phi, \lambda_i) = \bigoplus_{i=1}^{k} \operatorname{Eig}(\phi, \lambda_i).$$

Beweis: Wir zeigen die Aussage via vollständiger Induktion über k. Für $k = 0$ und $k = 1$ ist die Aussage richtig.

Für den Induktionsschritt von $k-1$ nach k sei $\mathbf{0} = u_1 + \cdots + u_k$, wobei $u_i \in \operatorname{Eig}(\phi, \lambda_i)$. Anwendung von ϕ gibt $\mathbf{0} = \phi(\mathbf{0}) = \lambda_1 u_1 + \cdots + \lambda_k u_k$, außerdem ist $\mathbf{0} = \lambda_k u_1 + \cdots + \lambda_k u_k$, also

$$\mathbf{0} = (\lambda_1 - \lambda_k)u_1 + \cdots + (\lambda_{k-1} - \lambda_k)u_{k-1} + (\lambda_k - \lambda_k)u_k.$$

Aus der Induktionsvoraussetzung erhalten wir, dass $u_1, \dots, u_{k-1} = \mathbf{0}$, also insgesamt $u_1, \dots, u_k = \mathbf{0}$ und die Summe ist direkt. □

Korollar V.2.6 (Diagonalisierbarkeitskriterium): *Seien V ein endlichdimensionaler K-Vektorraum und ϕ ein Endomorphismus von V mit Spektrum $\operatorname{Spec}(\phi) = \{\lambda_1, \dots, \lambda_k\}$.*

(i) *Es gilt $\#\operatorname{Spec}(\phi) \leq \dim V$.*

(ii) *Genau dann ist ϕ diagonalisierbar, wenn $V = \bigoplus_{i=1}^{k} \operatorname{Eig}(\phi, \lambda_i)$ und es gilt $V = \bigoplus_{i=1}^{k} \operatorname{Eig}(\phi, \lambda_i)$ genau dann, wenn $\dim V = \sum_{i=1}^{k} \dim \operatorname{Eig}(\phi, \lambda_i)$ ist.*

Beweis: (i) Wir haben uns bereits davon überzeugt, dass λ zu $\operatorname{Spec}(\phi)$ hört genau dann, wenn $\operatorname{Eig}(\phi, \lambda)$ ein echter Unterraum von V ist, das heißt, wenn $\dim \operatorname{Eig}(\phi, \lambda) \geq 1$. Weil die Summe der Eigenräume direkt ist, und da Dimension monoton ist, kann $k > \dim V$ nicht eintreten.

(ii) Ist ϕ diagonalisierbar, dann gibt es eine Basis $B = (b_1, \ldots, b_n)$ von V, bezüglich der die Darstellungsmatrix $A = D_{B,B}(\phi)$ von ϕ Diagonalgestalt besitzt. Per Definition der Darstellungsmatrix, und da die Darstellungsmatrix nach Voraussetzung Diagonalgestalt hat, ist

$$\phi(b_i) = D_B^{-1} \circ \varphi_A \circ D_B(b_i) = D_B^{-1}\varphi_A(e_i) = D_B^{-1}(A_{i,i}e_i) = A_{i,i}b_i,$$

d. h. die Elemente der Basis B sind Eigenvektoren von ϕ. Diejenigen Elemente von B, die zu den gleichen Diagonaleinträgen von A gehören, bilden linear unabhängige Teilmengen der zugehörigen Eigenräume.

Ist λ ein Eigenwert von ϕ, dann ist $\phi(v) = \lambda v$. Mit anderen Worten: Es gilt $\phi(\langle v\rangle) \subseteq \langle v\rangle$. Denn: Für w aus $\langle v\rangle$ gibt es α aus K, sodass $w = \alpha v$ gilt, und es ist $\phi(w) = \phi(\alpha v) = (\alpha\lambda)v$. Deshalb ist $\phi|_{\mathrm{Eig}(\phi,\lambda)}$ ein Endomorphismus von $\mathrm{Eig}(\phi, \lambda)$. Wählt man eine Basis B' von $\mathrm{Eig}(\phi, \lambda)$, dann hat $D_{B',B'}(\phi|_{\mathrm{Eig}(\phi,\lambda)})$ Diagonalgestalt mit Diagonaleinträgen λ. Das zeigt zusammen mit der Diagonalgestalt von A, dass diejenigen Elemente von B, die in $\mathrm{Eig}(\phi, \lambda)$ liegen, sogar eine Basis von $\mathrm{Eig}(\phi, \lambda)$ bilden.

Die zweite äquivalente Bedingung ist eine Formulierung der ersten mithilfe von Dimensionen. □

3. Determinante

Wir erinnern an die Signumsfunktion $\mathrm{sgn}\colon S_n \to \{\pm 1\}$. Wir haben uns bereits davon überzeugt, dass folgende Rechenregeln gelten: Für einen k-Zyklus σ ist $\mathrm{sgn}(\sigma) = (-1)^{k+1}$; für $\sigma_1, \sigma_2 \in S_n$ ist $\mathrm{sgn}(\sigma_1 \circ \sigma_n) = \mathrm{sgn}(\sigma_1)\,\mathrm{sgn}(\sigma_2)$.

Definition V.3.1 (Determinante): Seien K ein Körper, n eine natürliche Zahl und $A \in K^{n\times n}$. Dann heißt

$$\det A := \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

die *Determinante von A*.

Beispiel V.3.2: Seien $n = 2$ und $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in K^{2\times 2}$ gegeben. Wir haben $S_2 = \{\mathrm{id}, (12)\}$ mit $\mathrm{sgn}\,\mathrm{id} = 1$ und $\mathrm{sgn}(12) = -1$, sodass

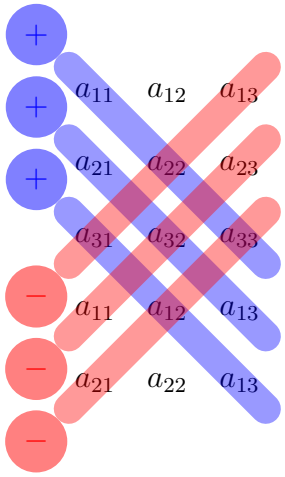
$$\det A = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = ad - bc.$$

Aus den Übungen ist bekannt, dass $\det A$ ein wichtiges Charakteristikum von A ist, das über die Invertierbarkeit von A Aufschluss gibt.

Beispiel V.3.3: Seien $n = 3$ und $A = (a_{i,j}) \in K^{3\times 3}$. Die symmetrische Gruppe vom Grad 3 ist $\{\mathrm{id}, (123), (132), (12), (13), (23)\}$ wobei die Transpositionen negatives Signum und die restlichen Permutationen positives Signum haben. Entsprechend ergibt sich

$$\begin{aligned}\det A = {} & a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ & - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2}.\end{aligned}$$

Für die obige Formel, die auch „Regel von Sarrus“ oder „Jägerzaunregel“ genannt wird, gibt es ein anschauliches Schema:



Eine Regel für $n \geq 4$ ist nicht praktikabel, da die Anzahl der Summanden explodiert. Stattdessen wird auf andere Sätze zur Berechnung von Determinanten zurückgegriffen.

Proposition V.3.4 (Eigenschaften der Determinante): *Seien K ein Körper, n eine natürliche Zahl und $D\colon \prod_{i=1}^{n} K^n \to K$, $(x_1, \ldots, x_n) \mapsto \det(x_1|\ldots|x_n)$.*

(i) *Sind $v_1, \ldots, v_n$ und v_i', $i \in \{1, \ldots, n\}$ in K^n, dann gilt*

$$\begin{aligned}& D(v_1, \ldots, v_{i-1}, v_i + v_i', v_{i+1}, \ldots, v_n) \\ & \quad = D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + D(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n).\end{aligned}$$

(ii) *Sind $v_1, \ldots, v_n$ in K^n und $\lambda \in K$, dann ist*

$$D(v_1, \ldots, v_{i-1}, \lambda v_i, v_{i+1}, \ldots, v_n = \lambda D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n).$$

(iii) *Sind $v_1, \ldots, v_n$ in K^n und gibt es $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ und $v_i = v_j$, dann ist $D(v_1, \ldots, v_n) = 0$.*

(iv) *Bezeichnet* $\{e_1, \dots, e_n\}$ *die Standardbasis, dann ist* $D(e_1, \dots, e_n) = 1$.

Beweis: Seien $v_1, \dots, v_n$ Elemente von K^n und $A = (v_1 | \dots | v_n)$.

(i) Seien i ein Element von $\{1, \dots, n\}$, $v_i' = (t_1, \dots, t_n)$ ein Vektor in K^n, $A' = (v_1 | \dots | v_{i-1} | v_i + v_i' | v_{i+1} | \dots | v_n)$ und $A'' = (v_1 | \dots | v_{i-1} | v_i' | v_{i+1} | \dots | v_n)$. Dann ist

$$\begin{aligned}\det A' &= \sum_{\sigma \in S_n} \prod_{k=1}^{n} a'_{k,\sigma(k)} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\substack{k=1 \\ \sigma(k) \neq i}}^{n} a_{k,\sigma(k)} (a_{\sigma^{-1}(i),i} + t_{\sigma^{-1}(i),i}) = \det A + \det A''.\end{aligned}$$

(ii) Seien $\lambda \in K$ und $A' = (v_1 | \dots | v_{i-1} | \lambda v_i | v_{i+1} | \dots | v_n)$. Der Faktor λ tritt in $\det A'$ in jedem Summanden genau einmal auf, d. h. $\det A' = \lambda \det A$.

(iii) Sei $v_k = v_\ell$ mit $k \neq \ell$ und $\sigma_0 := (k\ell) \in S_n$. Für $\sigma \in S_n$ sei $\sigma' := \sigma \circ \sigma_0$ (wir bemerken, dass $\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\sigma)$), ferner setze $A_n := \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$. Wir erhalten eine Bijektion $A_n \to S_n - A_n, \sigma \mapsto \sigma'$. Weiterhin gilt

$$\sigma'(i) = \begin{cases} \sigma(i), & \text{falls } i \notin \{k, \ell\}, \\ \sigma(l), & \text{falls } i = k, \\ \sigma(k), & \text{falls } i = l. \end{cases}$$

Da $v_k = v_\ell$ erhalten wir $\prod_{i=1}^n a_{i,\sigma'^{-1}(i)} = \prod_{i=1}^n a_{i,\sigma^{-1}(i)}$, eingesetzt in die Leibniz-Formel gibt das

$$\begin{aligned}\det(A) &= \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} + \sum_{\sigma \in S_n - A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \\ &= \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} + \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma') \prod_{i=1}^{n} a_{i,\sigma'(i)} \\ &= \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} - \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = 0.\end{aligned}$$

(iv) Sei $A = (e_1 | \dots | e_n)$. Wegen $a_{i,j} = \delta_{i,j}$ leistet nur id einen Beitrag in $\det A$, d. h. $\det A = \prod_{i=1}^n a_{i,i} = 1$. □

Beispiel V.3.5: Seien K ein Körper und $A = (v_1 | v_2 | v_3) \in K^{3\times 3}$ gegeben.

(i) Für $A' = (v_1 + \lambda v_2 | v_2 | v_3)$, $\lambda \in K$, gilt $\det A' = \det A$, denn

$$\det A' = \det(v_1|v_2|v_3) + \lambda \det(v_2|v_2|v_3) = \det A,$$

wobei wir für das erste Gleichheitszeichen die Eigenschaften (i) und (ii) aus Proposition V.3.4 und für das zweite Gleichheitszeichen die Eigenschaft (iv) aus der gleichen Proposition verwendet haben.

(ii) Für $A' = (v_3|v_2|v_1)$ gilt $-\det(v_1|v_2|v_3)$, denn

$$\begin{aligned} 0 &= \det(v_1 + v_3|v_2|v_1 + v_3) \\ &= \det(v_1|v_2|v_1 + v_3) + \det(v_3|v_2|v_1 + v_3) \\ &= \det(v_1|v_2|v_1) + \det(v_1|v_2|v_3) + \det(v_3|v_2|v_1) + \det(v_3|v_2|v_3) \\ &= \det(v_1|v_2|v_3) + \det(v_3|v_2|v_1), \end{aligned}$$

sodass $\det A = \det(v_1|v_2|v_3) = -\det(v_3|v_2|v_1) = -\det A'$.

(iii) Für $A' := (v_1|\lambda v_2|v_3)$ mit $\lambda \in K$ gilt nach Eigenschaft (ii) der Determinante, dass $\det A' = \lambda \det A$.

Zu den elementaren Zeilenumformungen gehörten die Matrizen

(i) $A^\alpha_{k,\ell} = I_n + \alpha E_{k,\ell}$ (Additionsmatrizen),

(ii) $V_{k,\ell} = I_n - E_{k,k} - E_{\ell,\ell} + E_{k,\ell} + E_{\ell,k}$ (Vertauschungsmatrizen),

(iii) $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ mit $\alpha_1 \cdots \alpha_n \neq 0$,

die wir im folgenden *spezielle Matrizen* nennen wollen.

Proposition V.3.6 (Determinante und spezielle Matrizen): *Seien K ein Körper, n eine natürliche Zahl und $A \in K^{n\times n}$.*

(i) *Für $A' := AA^\alpha_{k,\ell}$ ist $\det A' = \det A$.*

(ii) *Für $A' := AV_{k,\ell}$ ist $\det A' = -\det A$.*

(iii) *Für $A' := A \operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \alpha_1 \cdots \alpha_n \det A$.*

Beweis: Die gleichen Rechnungen wie in Proposition V.3.5 zeigen die Aussagen. □

Bemerkung V.3.7 (Nochmals spezielle Matrizen): Seien K ein Körper und n eine natürliche Zahl.

(i) Es gilt $\det A^\alpha_{k,\ell} = 1$, $\det V_{k,\ell} = -1$ sowie $\det \operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \alpha_1 \cdots \alpha_n$.

(ii) Ist A eine $n \times n$-Matrix mit Einträgen aus K und ist X eine spezielle Matrix, dann ist $\det(AX) = \det(A)\det(X)$.

Bemerkung V.3.8: Für die Determinante gelten folgende Rechenregeln:

(i) Entsteht A' aus A durch Addition der k-ten Spalte zur ℓ-ten Spalte ($k \neq \ell$), dann ist $\det A' = \det A$.

(ii) Entsteht A' aus A durch Vertauschung der k-ten und ℓ-ten Spalte, dann gilt $\det A' = -\det A$.

(iii) Entsteht A' aus A durch Multiplikation einer Spalte mit λ, dann ist $\det A' = \lambda \det A$.

(iv) Enthält A eine Nullspalte, dann ist $\det A = 0$.

Satz 18 (Eigenschaften der Determinante): *Seien K ein Körper, n eine natürliche Zahl und A, A_1, A_2 in $K^{n\times n}$.*

(i) *Genau dann ist* $\det A \neq 0$, *wenn* $A \in \mathrm{Gl}_n(K)$.

(ii) *Es gilt* $\det A = \det A^t$.

(iii) *Es gilt* $\det(A_1A_2) = \det A_1 \det A_2$.

(iv) *Ist* $A \in \mathrm{Gl}_n(K)$, *dann ist* $\det A^{-1} = 1/\det A$.

(v) *Die Determinante ist eine Ähnlichkeitsinvariante, d. h. ist* $S \in \mathrm{Gl}_n(K)$, *dann gilt* $\det(SAS^{-1}) = \det A$.

(vi) *Ein* λ *aus* K *ist Eigenwert von* A *genau dann, wenn* $\det(A - \lambda I_n) = 0$.

Beweis: (i) Angenommen, A gehörte nicht zu $\mathrm{Gl}_n(K)$. Es gäbe eine Treppenform T' und spezielle Matrizen $X_1, \ldots, X_k$, sodass $A^t = X_1 \cdots X_k T'$. Es wäre dann $A = T'^t X_k^t \cdots X_1^t$, wobei $\det T'^t = 0$. Außerdem wären $X_1^t, \ldots, X_k^t$ ebenfalls Vertauschungsmatrizen. Wegen Proposition V.3.6 wäre dann $\det A = 0$.

Angenommen, A gehörte zu $\mathrm{Gl}_n(K)$. Wir haben uns bereits überlegt, dass dann die Einheitsmatrix die Treppenform von A wäre, es gäbe also spezielle Matrizen $X_1, \ldots, X_k$, sodass $A = X_1 \cdots X_k I_n = I_n X_1 \cdots X_k$. Nach Proposition V.3.6 ist also $\det A = \det X_1 \cdots \det X_k \neq 0$.

(ii) Angenommen, A wäre nicht invertierbar. Dann wäre auch A^t nicht invertierbar, und nach (i) hätten wir $\det A = 0 = \det A^t$.

Angenommen, A wäre invertierbar. Dann gäbe es spezielle Matrizen $X_1, \ldots, X_\ell$, sodass $A = X_1 \cdots X_\ell$ und es wäre $A^t = X_\ell^t \cdots X_1^\ell$, d. h. nach Proposition V.3.6 wäre

$$\det A = \det X_1 \cdots \det X_\ell = \det(X_1^t) \cdots \det(X_\ell^t) = \det A^t.$$

(iii) Sind A_1 und A_2 invertierbar, dann sind sowohl A_1 als auch A_2 Produkte spezieller Matrizen und die Behauptung folgt aus Proposition V.3.6.

Ist A_1 invertierbar, aber A_2 nicht, dann ist $\operatorname{Rang}(A_2) \leq n-1$, d. h. nach der Dimensionsformel ist $\dim \operatorname{Kern} A_2 \geq 1$. Also gibt es $v \in K^n - \{\mathbf{0}\}$ mit $A_2 v = \mathbf{0}$ und dann ist erst recht $A_1 A_2 v = \mathbf{0}$, was $\dim \operatorname{Kern} A_1 A_2 \geq 1$ erzwingt. Damit ist $A_1 A_2$ nicht invertierbar, nach (i) also

$$0 \det(A_1 A_2) = \det A_1 \cdot 0 = \det A_1 \det A_2.$$

Ist A_1 nicht invertierbar, aber A_2 schon, dann ist

$$\det(A_1 A_2) = \det(A_2^t A_1^t) = \det A_2^t \det A_1^t = \det A_2 \det A_1.$$

(iv) Ist A invertierbar, dann haben wir

$$\det A^{-1} \det A = \det(A^{-1} A) = \det I_n = 1,$$

sodass $\det A^{-1} = (\det A)^{-1}$.

(v) Wegen (iv) ist $\det(SAS^{-1}) = \det S \det A (\det S)^{-1} = \det A$.

(vi) „$\Longrightarrow$“: Ist λ ein Eigenwert von A, dann gibt es einen von Null verschiedenen Vektor v in K^n, sodass $Av = \lambda v$. Das bedeutet $\mathbf{0} = Av - \lambda v = (A - \lambda I_n)v$, sodass $\ker A - \lambda I_n$ ein echter Unterraum des K^n und $A - \lambda I_n$ nicht injektiv ist. Die Zeilenstufenform von A enthält deshalb eine Nullzeile und weil die Determinante sich unter Transposition nicht ändert, ist $\det A - \lambda I_n = 0$ nach Proposition V.3.8(iv).

„$\Longleftarrow$“: Genau so. □

Bemerkung V.3.9 (Rechenregeln für Determinante): Für die Determinante gelten also die folgenden Rechenregeln:

(i) Entsteht A' aus A durch Addition der k-ten Zeile zur ℓ-ten Zeile ($k \neq \ell$), dann ist $\det A' = \det A$.

(ii) Entsteht A' aus A durch Vertauschen der k-ten und der ℓ-ten Zeile ($k \neq \ell$), dann ist $\det A' = -\det A$.

(iii) Entsteht A' aus A durch Multipilikation einer Zeile mit $\lambda \in K$, dann gilt $\det A' = \lambda \det A$.

(iv) Enthält A eine Nullzeile, dann gilt $\det A = 0$.

Definition V.3.10 (Determinante eines Endomorphismus): Seien K ein Körper, V ein n-dimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear. Dann heißt $\det \phi := \det D_{B,B}(\phi)$ die *Determinante von* ϕ.

Die Determinante eines Endomorphismus ist wohldefiniert, da wir bereits gezeigt haben, dass die Determinante eine Ähnlichkeitsinvariante ist, d. h., eine Darstellungsmatrix von ϕ bezüglich einer anderen Basis hat dieselbe Determinante. Dieselbe Überlegung zeigt, dass folgende Definition vernünftig ist:

Definition V.3.11 (Charakteristisches Polynom): Seien K ein Körper, A eine Matrix in $K^{n\times n}$, V ein n-dimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear.

(i) $\chi_A := \det(A - \lambda I_n)$ heißt *charakteristisches Polynom von A*.

(ii) $\chi_\phi := \chi_{D_{B,B}(\phi)}$ heißt *charakteristisches Polynom von ϕ*.

Beispiel V.3.12: Für $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ist

$$\chi_A = \det \begin{pmatrix} 1-X & 1 \\ 0 & 1-X \end{pmatrix} = (1-X)^2,$$

d. h. $\operatorname{Spec} A = \{1\}$.

4. Die Regel von Laplace

Definition V.4.1 (Streichmatrix): Seien K ein Körper, n eine natürliche Zahl und A in $K^{n\times n}$. Für $1 \le i, j \le n$ bezeichnet $A_{i,j} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

Beispiel V.4.2: Seien $n = 3$ und

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

Für $(i, j) = (1, 2)$ und $(i, j) = (2, 2)$ und $(i, j) = (3, 2)$ haben wir die Streichmatrizen

$$A_{1,2} = \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix}, \quad A_{2,2} = \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix}, \quad A_{3,2} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Satz 19 (Entwicklungssatz von Laplace): *Es seien K ein Körper, n eine natürliche Zahl und A in $K^{n\times n}$. Ferner sei $k \in \{1, \ldots, n\}$.*

(i) *Die Laplace-Entwicklung nach der k-ten Zeile ist*

$$\det A = \sum_{j=1}^{n} (-1)^{j+k} a_{k,j} \det(A_{k,j}).$$

(ii) *Die Laplace-Entwickung nach der k-ten Spalte ist*

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} a_{i,k} \det(A_{i,k})$$

Beispiel V.4.3: Für die Matrix A aus Proposition V.4.2 und $k = 2$ haben wir Folgendes für die Entwicklung nach der zweiten Spalte:

$$\det A = -2 \cdot \det \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} - \det \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} = -12.$$

Proposition V.4.4 (Blockmatrizen): *Seien k in $\{1, \ldots, n\}$, X in $K^{n \times n}$, Y in $K^{(n-k) \times k}$ und Z in $K^{(n-k) \times (n-k)}$. Dann ist*

$$\det \begin{pmatrix} X & \mathbf{0} \\ Y & Z \end{pmatrix} = \det X \det Z = \det \begin{pmatrix} X & Y \\ \mathbf{0} & Z \end{pmatrix}.$$

Beweis: Zunächst seien $X = I_k$ und $Y = \mathbf{0}$. Dann ist

$$\det \begin{pmatrix} I_k & \mathbf{0} \\ \mathbf{0} & Z \end{pmatrix} = \det(Z),$$

denn: Verwenden wir Spaltenumformungen, die Z in das Transponierte einer Treppenform bringen, dann bringen die selben Spaltenumformumgen $\left(\begin{smallmatrix} I_k & 0 \\ 0 & Z \end{smallmatrix}\right)$ ins Transponierte einer Treppenform, d. h. wir erhalten die Behauptung.

Nun zeigen wir die Behauptung für beliebiges X, d. h. $\det\left(\begin{smallmatrix} X & \mathbf{0} \\ \mathbf{0} & Z \end{smallmatrix}\right) = \det(X) \det(Z)$. Wir verwenden dazu

$$\begin{pmatrix} X & \mathbf{0} \\ \mathbf{0} & Z \end{pmatrix} = \begin{pmatrix} I_k & \mathbf{0} \\ \mathbf{0} & Z \end{pmatrix} \cdot \begin{pmatrix} X & \mathbf{0} \\ \mathbf{0} & I_{n-k} \end{pmatrix},$$

d. h. das oben gezeigte liefert uns die Behauptung.

Schließlich können wir die Aussage der Proposition zeigen. Ist $\det(X) = 0$, dann können wir über spezielle Matrizen erreichen, ohne die Determinante von X zu verändern, dass der Block X eine Nullzeile hat. Dann hat auch die Blockmatrix eine Nullzeile, und es gilt $\det\left(\begin{smallmatrix} X & \mathbf{0} \\ Y & Z \end{smallmatrix}\right) = 0$ nach Proposition V.3.8. Ist $\det(X) \neq 0$, dann ist $X \in \mathrm{Gl}_{n-k}(K)$. Weil

$$H := \begin{pmatrix} I_k & \mathbf{0} \\ -YX^{-1} & I_{n-k} \end{pmatrix}$$

eine Dreiecksmatrix ist, können wir $\det(H) = 1$ direkt ablesen. Außerdem gilt für $A := \left(\begin{smallmatrix} X & \mathbf{0} \\ Y & Z \end{smallmatrix}\right)$, dass $HA = \left(\begin{smallmatrix} X & \mathbf{0} \\ \mathbf{0} & Z \end{smallmatrix}\right)$. Wegen der Multiplikativität der Determinante also

$$\det \begin{pmatrix} X & \mathbf{0} \\ Y & Z \end{pmatrix} = \det(HA) = \det(H) \det(A) = \det(X) \det(Z)$$

nach dem vorher gezeigten. □

Bemerkung V.4.5 (Laplace-Entwicklung nach erster Zeile): Schreibe $A \in K^{n\times n}$ als

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ s_1 & s_2 & \cdots & s_n \end{pmatrix}$$

mit $a_{1,1}, \ldots, a_{1,n} \in K$ und $s_1, \ldots, s_n \in K^{n-1}$. Dann gilt

$$\begin{aligned} \det A &= \det \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ s_1 & s_2 & \cdots & s_n \end{pmatrix} + \cdots + \det \begin{pmatrix} 0 & \cdots & 0 & a_{1,n} \\ s_1 & \cdots & s_{n-1} & s_n \end{pmatrix} \\ &= \det \begin{pmatrix} a_{1,1} & \mathbf{0} \\ s_1 & A_{1,1} \end{pmatrix} - \det \begin{pmatrix} a_{1,2} & \mathbf{0} \\ s_2 & A_{1,2} \end{pmatrix} \\ &\qquad + \det \begin{pmatrix} a_{1,3} & \mathbf{0} \\ s_3 & A_{1,3} \end{pmatrix} + \cdots + (-1)^{n+1} \begin{pmatrix} a_{1,n} & \mathbf{0} \\ s_n & A_{1,n} \end{pmatrix} \\ &= \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det A_{1,j}. \end{aligned}$$

Beweis (von Satz 21): (i) Schreibe $A = (z_1|\ldots|z_n)^t$ als Vektor der Zeilenvektoren von A. Durch $(i-1)$ Zeilenvertauschungen können wir erreichen, dass die i-te Zeile an die Stelle der ersten Zeile rückt, d. h.

$$\det A = (-1)^{i-1} \det(z_i|z_1|\ldots|z_{i-1}|z_{i+1}|\ldots|z_n)^t,$$

sodass die vorangegangene Bemerkung die Behauptung liefert.

(ii) Durch Transposition können wir uns auf den ersten Fall zurückziehen. □

Teil 2.

Lineare Algebra II

Kapitel VI.

Die Jordan-Normalform

1. Motivation

Für diesen Abschnitt seien K ein Körper und V ein endlichdimensionaler K-Vektorraum der Dimension n. Wir möchten in diesem Abschnitt Endomorphismen $\phi\colon V \to V$ untersuchen und „zugehörige“ ϕ-invariante Zerlegungen von V studieren.

Erinnerung (Diagonalisierbarkeit): Ist $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V, dann können wir die Wirkung von ϕ mithilfe der zugehörigen Darstellungsmatrix $D_{B,B}(\phi)$ beschreiben. Diese ist festgelegt dadurch, dass sie das Diagramm

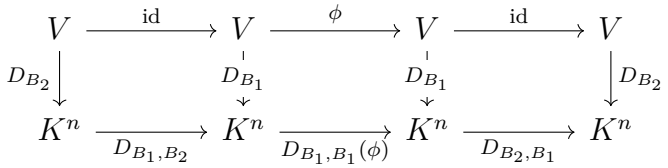
$$\begin{array}{ccc} V & \xrightarrow{\phi} & V \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_B} \\ K^n & \xrightarrow[x \mapsto D_{B,B}(\phi)x]{} & K^n \end{array}$$

kommutativ macht. Die Matrix $D_{B,B}(\phi)$ ist eine $n \times n$-Matrix mit Einträgen aus K. Dass das obige Diagramm kommutiert bedeutet genau, dass für jedes v aus V gilt: $D_B(\phi(v)) = D_{B,B}(\phi)D_B(v)$. Wegen $D_B(b_i) = e_i$ heißt das insbesondere, dass in der i-ten Spalte von $D_{B,B}(\phi)$ die Koordinaten von $\phi(b_i)$ bezüglich der Basis B stehen müssen.

Genau dann sind zwei $n \times n$-Matrizen A_1 und A_2 mit Einträgen aus K Darstellungsmatrizen desselben Endomorphismus bezüglich unterschiedlicher Basen, wenn es eine invertierbare $n \times n$-Matrix S mit Einträgen aus K gibt, sodass $A_2 = SA_1S^{-1}$. Diese Situation hatten wir in der Linearen Algebra I beschrieben, indem wir gesagt haben, A_1 und A_2 seien ähnlich.

Sind genauer B_1 und B_2 geordnete Basen von V, dann wird die oben be-

schrieben Situation im Diagramm



eingefangen, da ja $D_{B_1,B_2}^{-1} = D_{B_2,B_1}$. Hierbei haben wir schlampig nicht zwischen einer Matrix A und der zugehörigen linearen Abbildung $x \mapsto Ax$ unterschieden.

Schließlich haben wir in der Linearen Algebra I über Diagonalisierbarkeit gesprochen. Der Endormorphismus ϕ ist diagonalisierbar, wenn es eine Basis B von V bestehend aus Eigenvektorn von ϕ gibt. Das heißt für die Darstellungsmatrix $D_{B,B}(\phi)$, dass es sich um eine Diagonalmatrix, also eine Matrix der Form $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ für Elemente $\lambda_1, \ldots, \lambda_n$ von K, handelt. Wir hatten uns davon überzeugt, dass ein Endomorphismus genau dann diagonalisierbar ist, wenn $V = \bigoplus_{\lambda \in \operatorname{Spec}(\phi)} \operatorname{Eig}(\phi, \lambda)$.

Die Summe der Eigenräume ist zwar immer direkt, aber im Allgemeinen handelt es sich dabei um einen Untervektorraum von V; im Allgemeinen ist ein Endomorphismus von V nicht diagonalisierbar.

Beispiel VI.1.1 (Was kann schief gehen?): In diesem Beispiel möchten wir uns der zwei wesentlichen Probleme vergewissern, die beim Diagonalisierungsversuch auftreten können.

(i) Das charakteristische Polynom zerfällt nicht in Linearfaktoren. Sei ϕ der Endomorphismus

$$\phi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Um Schreibarbeit zu sparen, nennen wir die Matrix in der obigen Definition A. Das charakteristische Polynom von ϕ ist

$$\chi_\phi = \det(A - \lambda I_2) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1.$$

Aber weil Quadrate reeller Zahlen nicht-negativ sind, ist $\operatorname{Spec}(\phi) = \varnothing$. Über $\mathbb{R}$ haben wir entsprechend keine Chance, den Endomorphismus zu diagonalisieren. Betrachten wir den Endomorphismus $\hat{\phi}$, der mithilfe derselben Matrix auf $\mathbb{C}^2$ erklärt wird, dann ist $\operatorname{Spec}(\hat{\phi}) = \{\mathrm{i}, -\mathrm{i}\}$ und wir sind im Rennen, da wir zwei

verschiedene Eigenwerte und damit eine zweidimensionale direkte Summe von Eigenräumen haben. Genauer sind

$$\mathrm{Eig}(\hat{\phi}, \mathrm{i}) = \mathrm{Kern}(A - \mathrm{i}I_2) = \mathrm{Kern}\begin{pmatrix} -\mathrm{i} & -1 \\ 1 & -\mathrm{i} \end{pmatrix} = \left\langle \begin{pmatrix} -\mathrm{i} \\ -1 \end{pmatrix} \right\rangle = \langle b_1 \rangle$$
$$\mathrm{Eig}(\hat{\phi}, -\mathrm{i}) = \mathrm{Kern}(A + \mathrm{i}I_2) = \mathrm{Kern}\begin{pmatrix} \mathrm{i} & -1 \\ 1 & \mathrm{i} \end{pmatrix} = \left\langle \begin{pmatrix} \mathrm{i} \\ -1 \end{pmatrix} \right\rangle = \langle b_2 \rangle$$

und für $B = (b_1, b_2)$ ist $D_{B,B}(\hat{\phi}) = \mathrm{diag}(\mathrm{i}, -\mathrm{i})$.

(ii) Die direkte Summe der Eigenräume ist ein echter Unterraum von V, obwohl das charakteristische Polynom in Linearfaktoren zerfällt. Sei dazu ϕ der Endomorpismus

$$\phi\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Wieder nennen wir die Matrix aus der obigen Definition A. Das charakteristische Polynom von ϕ ist $\chi_\phi = \det(A - \lambda I_3) = (2 - \lambda)^3$, d. h. der einzige Eigenwert von ϕ ist 2. Bestimmen wir jedoch den zugehörigen Eigenraum, dann stellen wir fest, dass

$$\mathrm{Eig}(\phi, 2) = \mathrm{Kern}(A - 2I_3) = \mathrm{Kern}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle,$$

was plausibel ist, da der Rang von $A - 2I_3$ zwei ist. Wir haben entsprechend $\bigoplus_{\lambda \in \mathrm{Spec}(\phi)} \mathrm{Eig}(\phi, \lambda) = \mathrm{Eig}(\phi, 2) \subsetneq \mathbb{R}^3$, und ϕ ist nicht diagonalisierbar.

Im Folgenden möchten wir gerne zweierlei versuchen: Zum einen möchten wir zu jedem Endomorphismus ϕ von V eine „möglichst schöne“ Darstellungsmatrix J_ϕ finden. Äquivalent dazu ist das Finden eine „möglichst schönen“ Matrix J_A, die zu einer gegebenen Matrix A ähnlich ist.

Zum anderen hätten wir gerne, dass diese Matrix J_A eindeutig in einer gegebenen Ähnlichkeitsklasse ist, d. h. wir hätten gerne, dass zwei Matrizen A und A' genau dann ähnlich sind, wenn die zugehörigen J_A und $J_{A'}$ übereinstimmen.

Um nicht in Problem (i) aus Proposition VI.1.1 zu laufen, werden wir uns später auf Körper einschränken, in denen Polynome immer in Linearfaktoren zerfallen.

Definition VI.1.2 (Algebraisch abgeschlossener Körper): Sei K ein Körper. Falls es für jedes Polynom f aus $K[X]$ ein Element α aus K gibt, sodass $f(\alpha) = 0$, dann heißt K *algebraisch abgeschlossen.*

Da das Abspalten von Nullstellen den Grad reduziert, kann man leicht per Induktion zeigen, dass ein Körper genau dann algebraisch abgeschlossen ist, wenn jedes Polynom mit Koeffizienten in K in Linearfaktoren zerfällt.

Unser Ziel für dieses Kapitel wird das folgende Resultat sein:

Satz 20 (über die Jordan-Normalform): *Seien K ein algebraisch abgeschlossener Körper und V ein K-Vektorraum von Dimension n. Dann gibt es eine geordnete Basis B von V, sodass*

$$J_\phi := D_{B,B}(\phi) = \begin{pmatrix} \lambda_1 & & & & & & & & & \\ 1 & \ddots & & & & & & & & \\ & \ddots & \ddots & & & & & & & \\ & & 1 & \lambda_1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \lambda_k & & & & \\ & & & & & 1 & \ddots & & & \\ & & & & & & \ddots & \ddots & & \\ & & & & & & & 1 & \lambda_k \end{pmatrix}$$

Hierbei sind alle nicht eingetragenen Werte Null, $\lambda_1, \ldots, \lambda_r$ sind die Eigenwerte von ϕ (nicht notwendigerweise verschieden) und die Matrix J_ϕ ist eindeutig bis auf Vertauschen der Kästchen und heißt Jordan-Normalform.

2. Der Satz von Cayley-Hamilton

Das Ziel dieses Abschnitts wird zu zeigen, dass für $A \in K^{n\times n}$ mit charakteristischen Polynom $\chi_A \in K[X]$ gilt: $\chi_A(A) = 0$. Dazu sind einige Vorbemerkungen angebracht.

Bemerkung: Seien K ein Körper und V ein n-dimensionaler K-Vektorraum.

(i) Man kann quadratische Matrizen in Polynome einsetzen. Genauer: Sind $p \in K[X]$ mit $p = \sum_{i=0}^{d} a_i X^i$ und $A \in K^{n\times n}$, dann ergibt der Ausdruck

$$p(A) = a_d A^d + a_{d-1} A^{d-1} + \cdots + a_1 A + a_0 I_n$$

Sinn, denn wir können Matrizen addieren, potenzieren und mit Skalaren multiplizieren.

(ii) Wir haben $Ap(A) = p(A)A$, denn

$$\begin{aligned} Ap(A) &= A(a_dA^d + \cdots + a_1A + a_0I_n) \\ &= a_dA^{d+1} + \cdots + a_1A^2 + a_0A = (a_dA^d + \cdots + a_1A + a_0I_n)A = p(A)A. \end{aligned}$$

(iii) Sind $p_1, p_2 \in K[X]$, dann sind $(p_1 + p_2)(A) = p_1(A) + p_2(A)$ und $(p_1p_2)(A) = p_1(A)p_2(A)$. Das kann man einfach nachrechnen.

(iv) Man kann genauso Endomorphismen in Polynome einsetzen. Ist genauer $p \in K[X]$ mit $p = \sum_{i=0}^{d} a_iX^i$ und $\phi \in \mathrm{End}(V)$, dann ergibt

$$p(\phi) = a_d\phi^d + \cdots + a_1\phi + a_0 \,\mathrm{id}_V$$

Sinn. Hier steht ϕ^k für die k-fache Komposition von ϕ mit sich selbst und id_V bezeichnet die Identität von V.

(v) Sind ϕ ein Endomorphismus von V, $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V, $A := D_{B,B}(\phi)$ und $p \in K[X]$, dann gilt $D_{B,B}(p(\phi)) = p(A)$.

Beispiel VI.2.1 (Endomorphismus im Polynom): Seien $V = \mathbb{R}^2$ und $\phi\colon V \to V$, $(x, y) \mapsto (-x, y)$ die „Spiegelung an der y-Achse“. Weiter sei $p_1 = X^2 + 1$. Dann ist $p_1(\phi) = \phi^2 + \mathrm{id}_V = 2\,\mathrm{id}_V$.

Für das Polynom $p_2 = X^2 - 1$ ist $p_2(\phi) = \phi^2 - \mathrm{id}_V = \mathrm{id}_V - \mathrm{id}_V = \mathbf{0}$ die Nullabbildung.

Beispiel VI.2.2 (Zaubertrick): Sei

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

Das charakteristische Polynom von A ist $\chi_A = X^4 + 2X^3 - X^2 - 2X + 1$. Um $\chi_A(A)$ auszuwerten, müssen wir die Potenzen A^2, A^3 und A^4 ausrechnen. Diese sind

$$A^2 = \begin{pmatrix} 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & -5 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 5 \end{pmatrix},$$

$$A^3 = \begin{pmatrix} 0 & -1 & 2 & -5 \\ 0 & 2 & -5 & 12 \\ 0 & 1 & 0 & 0 \\ 1 & -2 & 5 & -10 \end{pmatrix}, \quad A^4 = \begin{pmatrix} -1 & 2 & -5 & 10 \\ 2 & -5 & 12 & -25 \\ 1 & 0 & 0 & 2 \\ -2 & 5 & 10 & 20 \end{pmatrix}.$$

Damit berechnen wir

$$\begin{aligned}\chi_A(A) &= \begin{pmatrix} -1 & 2 & -5 & 10 \\ 2 & -5 & 12 & -25 \\ 1 & 0 & 0 & 2 \\ -2 & 5 & -10 & 20 \end{pmatrix} + \begin{pmatrix} 0 & -2 & 4 & -10 \\ 0 & 4 & -10 & 24 \\ 0 & 2 & 0 & 0 \\ 2 & -4 & 10 & -20 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 5 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -5 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 2 \\ -2 & 0 & 0 & -4 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

Beispiel VI.2.3 (Charakteristisches Polynom für spezielle Matrizen): Es sei

$$A := \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 1 & 0 & 0 & \beta \\ 0 & 1 & 0 & \gamma \\ 0 & 0 & 1 & \delta \end{pmatrix}.$$

Dann ist

$$\begin{aligned}\chi_A &= \det(A - XI_4) \\ &= \det \begin{pmatrix} -X & 0 & 0 & \alpha \\ 1 & -X & 0 & \beta \\ 0 & 1 & -X & \gamma \\ 0 & 0 & 1 & \delta - X \end{pmatrix} \\ &= (-\alpha) \det \begin{pmatrix} 1 & -X & 0 \\ 0 & 1 & -X \\ 0 & 0 & 1 \end{pmatrix} + \beta \det \begin{pmatrix} -X & 0 & 0 \\ 0 & 1 & -X \\ 0 & 0 & 1 \end{pmatrix} \\ &\quad - \gamma \det \begin{pmatrix} -X & 0 & 0 \\ 1 & -X & 0 \\ 0 & 0 & 1 \end{pmatrix} + (\delta - X) \det \begin{pmatrix} -X & 0 & 0 \\ 1 & -X & 0 \\ 0 & 1 & -X \end{pmatrix} \\ &= -\alpha - \beta X - \gamma X^2 - \delta X^3 + X^4.\end{aligned}$$

Bemerkung VI.2.4: Mit (e_1, e_2, e_3, e_4) bezeichnen wir wie gewöhnlich die Standardbasis in K^4. Ferner sei A die Matrix aus Proposition VI.2.3, d. h.

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

Wir halten fest:

(i) $Ae_1 = e_2$, $Ae_2 = e_3$, $Ae_3 = e_4$,

(ii) $Ae_4 = \alpha e_1 + \beta e_2 + \gamma e_3 + \delta e_4$,

(iii) $\chi_A = -\alpha - \beta X - \gamma X^2 - \delta X^3 + X^4$.

Damit erhalten wir:

(iv) $\chi_A(A)e_1 = (-\alpha - \beta A - \gamma A^2 - \delta A^3 + A^4)e_1 = -\alpha e_1 - \beta e_2 - \gamma e_3 - \delta e_4 + A^4 e_4$ nach (i), sodass nach (ii) gilt: $Ae_1 = \mathbf{0}_{K^n}$.

(v) $\chi_A(A)e_2 = \chi_A(A)Ae_1 = A\chi_A e_1 = A\mathbf{0}_{K^n} = \mathbf{0}_{K^n}$.

Analog zeigt man, dass $\chi_A(A)e_3 = \mathbf{0}_{K^n}$ und $\chi_A(A)e_4 = \mathbf{0}_{K^n}$, sodass tatsächlich $\chi_A(A) = \mathbf{0}_{K^{n\times n}}$.

Proposition VI.2.5 (Cayley-Hamilton für spezielle Matrizen): *Sei*

$$A = \begin{pmatrix} & & & \alpha_0 \\ 1 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & 1 & \alpha_{n-1} \end{pmatrix} \in K^{n\times n}.$$

Dann gilt

(i) $\chi_A(X) = (-1)^{n+1}(\alpha_0 + \alpha_1 X + \cdots + \alpha_{n-1}X^{n-1} - X^n)$,

(ii) $\chi_A(A) = \mathbf{0}_{K^{n\times n}}$.

Beweis: Mit den Argumenten aus Proposition VI.2.4 lässt sich die Aussage zeigen. □

Um Proposition VI.2.5 weiter zu verwenden, suchen wir uns zu einem gegebenen Endomorphismus ϕ einen ϕ-invarianten Unterraum (d. h. $\phi(U) \subseteq U$) und eine Basis in U, sodass $D_{B,B}(\phi|_U)$ die Form der Matrix in Proposition VI.2.5 besitzt.

Definition VI.2.6 (ϕ-invariante Unterräume): Seien K ein Körper, V ein K-Vektorraum, $\phi \in \mathrm{End}(V)$ und U ein Untervektorraum von V. Gilt für alle $u \in U$, dass $\phi(u) \in U$, dann heißt U *ϕ-invariant*.

Proposition VI.2.7 (Minimaler ϕ-invarianter Unterraum): *Seien K ein Körper, V ein K-Vektorraum mit $\dim V = m$, $\phi \in \mathrm{End}(V)$ und $v \in V$. Sei weiter n minimal mit der Eigenschaft, dass $\{v, \phi(v), \ldots, \phi^n(v)\}$ linear abhängig ist, d. h. $\phi^n(v) = \sum_{i=0}^{n-1} \alpha_i \phi^i(v)$. Dann gilt:*

(i) *$U := \langle v, \phi(v), \ldots, \phi^{n-1}(v)\rangle$ ist ein ϕ-invarianter Untervektorraum von V, der v enthält.*

(ii) *U ist minimal bezüglich Inklusion mit der Eigenschaft aus* (i).

(iii) *$B = (v, \phi(v), \ldots, \phi^{n-1}(v))$ ist eine geordnete Basis von U und*

$$D_{B,B}(\phi|_U) = \begin{pmatrix} & & & \alpha_0 \\ 1 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & 1 & \alpha_{k-1} \end{pmatrix}$$

Beweis: (i) Wir bemerken, dass $\phi(\phi^{n-1}(v)) = \phi^n(v) = \sum_{i=0}^{n-1} \alpha_i \phi^i(v)$ zu U gehört. Ist jetzt $u \in U$, dann gibt es $c_0, \ldots, c_{n-1} \in K$ mit $u = \sum_{i=0}^{n-1} c_i \phi^i(v)$ und

$$\phi(u) = c_0\phi(v) + c_1\phi^2(v) + \cdots + c_{n-2}\phi^{n-1}(v) + c_{n-1}\phi^n(v)$$

gehört wieder zu U, d. h. U ist ϕ-invariant.

(ii) Sind W ein ϕ-invarianter Unterraum von V und $v \in W$, dann müssen auch die $\phi^i(v)$ für natürliche Exponenten i zu W gehören, d. h. $U \subseteq W$.

(iii) Per Wahl von n ist B eine Basis von U. Setzen wir $b_i := \phi^{i-1}(v)$ für $1 \leq i \leq n$, dann erhalten wir für $1 \leq i \leq n-1$, dass $\phi(b_i) = b_{i+1}$, und für $i = n$ ist $\phi(b_n)\phi^n(v) = \sum_{i=0}^{n-1} \alpha_i b_i$, was die Behauptung über die Darstellungsmatrix zeigt. □

Satz 21 (Cayley-Hamilton): *Seien n eine natürliche Zahl, K ein Körper und V ein n-dimensionaler K-Vektorraum.*

(i) *Für $A \in K^{n\times n}$ gilt $\chi_A(A) = \mathbf{0}_{K^{n\times n}}$.*

(ii) *Für $\phi \in \mathrm{End}(V)$ ist $\chi_\phi(\phi) = \mathbf{0}_{\mathrm{End}(V)}$.*

Beweis: Wir zeigen zuerst (ii). Seien χ_ϕ das charakteristische Polynom von ϕ und $\psi := p(\phi) \in \mathrm{End}(V)$. Wir wollen zeigen, dass $\psi = 0_{\mathrm{End}(V)}$. Dazu genügt es zu zeigen, dass für alle $v \in V$ gilt: $\psi(v) = \mathbf{0}_V$.

Sei also $v \in V$ gegeben. Zunächst wählen wir eine „schöne Basis“. Sei dazu k wie in Proposition VI.2.7 minimal, sodass die Menge $B' := \{v, \phi(v), \ldots, \phi^{k-1}(v)\}$ linear unabhängig ist, d. h. $U := \langle v, \phi(v), \ldots, \phi^{k-1}(v)\rangle$ ist der minimale ϕ-invariante Unterraum von V, der v enthält. Ergänze B' zu einer Basis B von V. Nach Proposition VI.2.7 ist

$$A := D_{B,B}(\phi) = \begin{pmatrix} A' & * \\ \mathbf{0} & C \end{pmatrix}, \qquad A' = \begin{pmatrix} & & & \alpha_0 \\ 1 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & 1 & \alpha_{n-1} \end{pmatrix}$$

Nun berechnen wir das charakteristische Polynom χ_A. Aus der Vorlesung „Lineare Algebra I“ wissen wir, dass die Determinante einer Blockmatrix der Gestalt von A das Produkt der Determinanten der quadratischen Blöcke A' und C ist – auch $\lambda I_n - A$ ist eine Blockmatrix von der gleichen Gestalt wie A, d. h. auch für das charakteristische Polynom gilt das. Entsprechend ist $\chi_A = \chi_{A'} \cdot \chi_C$. Setzen wir A in χ_A ein, dann erhalten wir die Blockmatrix

$$\chi_A(A) = \begin{pmatrix} \chi_{A'}(A') & * \\ \mathbf{0} & \chi_C(C) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & * \\ \mathbf{0} & \chi_C(C) \end{pmatrix},$$

denn $A^k = \left(\begin{smallmatrix} (A')^k & * \\ \mathbf{0} & C^k \end{smallmatrix}\right)$. Wir haben also gezeigt: $\psi|_U = \mathbf{0}$, insbesondere ist $\psi(v) = \mathbf{0}$.

Aussage (i) folgt jetzt aus (ii) wie folgt: Zur gegebenen Matrix A definieren wir die lineare Abbildung $\phi\colon x \mapsto Ax$, sodass $A = D_{E,E}(\phi)$. Insbesondere gilt $\chi_A = \chi_\phi$, d. h. $\chi_A(A) = D_{E,E}(\chi_A(\phi)) = D_{E,E}(\chi_\phi(\phi)) = \mathbf{0}_{K^{n\times n}}$. □

Bemerkung VI.2.8: Sei A eine Matrix in $K^{n\times n}$. Bei der Definition des charakteristischen Polynoms $\chi_A = \det(A - XI_n)$ berechnen wir eigentlich die Determinante einer Matrix über dem Ring $R := K[X]$. Hierfür benötigen wir allgemeiner Determinanten von Matrizen über Ringen R, also $\det\colon R^{n\times n} \to R$. Diese kann genau so wie über Körpern definiert werden und es gelten auch die Regel von Laplace und die Rechenregeln für elementare Zeilen- und Spaltenoperationen.

Bemerkung VI.2.9 (Alternativer Beweis für Cayley-Hamilton): Es seien K ein Körper und $A \in K^{n\times n}$ gegeben. Wir wollen zeigen, dass $\chi_A(A) = \mathbf{0}$ in $K^{n\times n}$.

Wegen $\chi_A(X) = \det(A - XI_n)$ ist $\chi_A(A) = \det(A - AI_n) = \det(A - A) = 0$. Das ist Quatsch!

$$A - XI_n = \begin{pmatrix} a_{1,1} - X & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - X & \dots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - X \end{pmatrix} \in (K[X])^{n\times n}.$$

Ist $B \in K^{n\times n}$, dann gilt *nicht*: $\chi_A(B) = A - BI_n = A - B \in K^{n\times n}$.

3. Der Polynomring über einem Körper

Aus dem Satz von Cayley-Hamilton wissen wir für eine Matrix $A \in K^{n\times n}$, dass $\chi_A(A) = \mathbf{0}$. Gibt es außer Vielfachen von χ_A noch weitere Polynome, die das auch erfüllen? Wir werden das im folgenden Abschnitt systematisch untersuchen.

Es wird sich herausstellen, dass sich der Polynomring $K[X]$ in seinen Eigenschaften im Wesentlichen wie der Ring der ganzen Zahlen verhält, und dass es ein nicht-triviales Polynom kleinsten Grades gibt, sodass jedes Polynom, das die Matrix A annulliert, ein Vielfaches dieses kleinsten Polynoms ist.

Erinnerung VI.3.1: Sei K ein Körper. Dann heißt

$$K[X] := \left\{ \sum_{i=1}^{n} a_i X^i : n \in \mathbb{N}_0, a_i \in K \right\}$$

mit der Addition und Multiplikation von Polynomen der *Polynomring über* K.

Definition VI.3.2 (K-Algebra): Seien K ein Körper und A eine Menge mit drei Abbildungen

$$+\colon A \times A \longrightarrow A, \qquad \bullet\colon A \times A \longrightarrow A, \qquad \cdot\colon K \times A \longrightarrow A.$$

Falls gilt

(i) $(A, +, \bullet)$ ist ein Ring,

(ii) $(A, +, \cdot)$ ist ein K-Vektorraum,

(iii) „$\bullet$“ ist bilinear, d. h. für alle $x, y, z \in A$ und $\lambda \in K$ gilt

$$\begin{aligned} (x + y) \bullet z &= x \bullet z + y \bullet z, \\ x \bullet (y + z) &= x \bullet y + x \bullet z, \qquad (\lambda x) \bullet y = \lambda(x \bullet y) = x \bullet (\lambda y), \end{aligned}$$

dann heißt A eine *K-Algebra*.

Ist $(A, +, \bullet)$ ein kommutativer bzw. unitärer Ring, dann heißt A eine *kommutative K-Algebra* bzw. *unitäre K-Algebra* oder *K-Algebra mit Eins* (d. h. es gibt $1 \in A$, sodass für alle $a \in A$ gilt: $1 \bullet a = a = a \bullet 1$).

Seien A_1 und A_2 zwei K-Algebren und $\phi\colon A_1 \to A_2$ eine Abbildung. Ist ϕ ein Homomorphismus von K-Vektorräumen und ein Homomorphismus von (unitären) Ringen, dann heißt ϕ ein *Homomorphismus (unitärer) K-Algebren*.

Ist $\phi\colon A_1 \to A_2$ ein Homomorphismus (unitärer) K-Algebren und gibt es einen Homomorphismus (unitärer) K-Algebren $\psi\colon K_2 \to K_1$ mit $\phi \circ \psi = \mathrm{id}_{A_2}$ und $\psi \circ \phi = \mathrm{id}_{A_1}$, dann heißt ϕ ein *Isomorphismus (unitärer) K-Algebren*.

Wie bei Gruppen, Vektorräumen und Ringen zeigt man, dass bijektive Homomorphismen genau die Isomorphismen sind.

Beispiel VI.3.3 (Erste K-Algebren): Sei K ein Körper.

- Der Polynomring $K[X]$ mit Addition, Multiplikation und skalarer Multiplikation ist eine K-Algebra.
- Der Ring der $n \times n$- Matrizen $K^{n\times n}$ ist eine K-Algebra mit Addition, skalarer Multiplikation und Multiplikation von Matrizen.
- Für einen K-Vektorraum V ist $\mathrm{End}_K(V)$ zusammen mit Addition, skalarer Multiplikation und Verknüpfung von Endomorphismen eine K-Algebra.

Satz 22: *Sei jetzt V ein endlichdimensionaler K-Vektorraum mit geordneter Basis B. Dann ist*

$$D_{B,B}\colon \mathrm{End}_K(V) \longrightarrow K^{n\times n}, \qquad \phi \longmapsto D_{B,B}(\phi)$$

ein Isomorphismus unitärer K-Algebren.

Definition VI.3.4 (Einsetzen in Polynome): Seien K ein Körper und $(A, +, \bullet, \cdot)$ eine unitäre K-Algebra. Dann kann man die Elemente aus A in Polynome aus $K[X]$ einsetzen. Genauer: Für $p = \sum_{i=0}^{n} c_i X^i \in K[X]$ und $a \in A$ setzen wir

$$p(a) := \sum_{i=0}^{n} c_i a^i$$

wobei a^i die i-fache Multiplikation von a mit sich selbst bezeichnet und wir die Konvention $a^0 := 1_A$ verwenden.

Definition VI.3.5 (Einsetzungsmorphismus): Seien K ein Körper, $(A, +, \bullet, \cdot)$ eine unitäre K-Algebra und $a \in A$. Dann heißt

$$\varphi_a \colon K[X] \longrightarrow A, \qquad p \longmapsto p(a)$$

der *Einsetzungshomomorphismus von a*. Tatsächlich ist der Einsetzungshomomorphismus ein Homomorphismus unitärer K-Algebren, d. h. es gilt für alle $p_1, p_2 \in K[X]$ und $\lambda \in K$:

$$(p_1+p_2)(a) = p_1(a)+p_2(a), \qquad (p_1p_2)(a) = p_1(a)\bullet p_2(a), \qquad (\lambda p)(a) = \lambda p(a).$$

Definition VI.3.6: Seien K ein Körper und $p = \sum_{i=0}^{n} c_i X^i$ in $K[X]$. Die Zahl

$$\deg p := \begin{cases} \max\{n \in \mathbb{N} \mid c_n \neq 0\}, & \text{falls } p \neq 0, \\ -\infty, & \text{falls } p = 0 \end{cases}$$

heißt der *Grad des Polynoms p*. Hierbei ist „$-\infty$“ ein Symbol, das nicht zu $\mathbb{N}_0$ gehört. Für das Symbol „$-\infty$“ definieren wir folgende Rechenregeln:

(i) Für alle $k \in \mathbb{N}_0$ ist $\max\{k, -\infty\} = k$,

(ii) $\max\{-\infty, -\infty\} = -\infty$,

(iii) Für alle $k \in \mathbb{N}_0$ ist $k + (-\infty) := -\infty =: (-\infty) + k$,

(iv) $(-\infty) + (-\infty) := -\infty$,

(v) Für alle $k \in \mathbb{N}_0$ ist $-\infty < k$.

Ist p nicht das Nullpolynom, und ist $c_n \neq 0$, dann heißt c_n der *Leitkoeffizient von p*. Ist 1 der Leitkoeffizient von p, dann heißt p *normiert*.

Bemerkung VI.3.7: Seien K ein Körper und $p_1, p_2 \in K[X]$. Dann gilt:

(i) $\deg(p_1 + p_2) \leq \max\{\deg(p_1), \deg(p_2)\}$,

(ii) $\deg(p_1p_2) = \deg(p_1) + \deg(p_2)$,

(iii) Die Einheitengruppe $K[X]^\times$ (d. h. die multiplikativ invertierbaren Polynome) lässt sich folgendermaßen beschreiben:

$$\begin{aligned} K[X]^\times &= \{f \in K[X] \mid \text{Es gibt } g \in K[X] \text{ mit } gf = fg = 1\} \\ &= \{f \in K[X] \mid \deg(f) = 0\} \cong K^\times. \end{aligned}$$

Die beiden Gruppen sind isomorph vermöge $\lambda \mapsto \lambda X^0$.

(iv) Es gibt keine *Nullteiler* in $K[X]$, d. h. es gibt kein $f \in K[X] - \{0\}$, sodass es $g \in K[X] - \{0\}$ gibt mit $gf = 0$.

(v) Für unitäre kommutative Ringe R kann man genauso den Polynomring $R[X]$ definieren. In (ii) gilt dann allerdings nur noch „$\leq$“.

Proposition VI.3.8 (Polynomdivision mit Rest): *Seien K ein Körper und p_1, p_2 in $K[X]$ mit $p_2 \neq 0$. Dann gibt es Polynome $h, r \in K[X]$ mit $\deg(r) < \deg(p_2)$ und $p_1 = hp_2 + r$.*

Definition VI.3.9 (Verschwindungsideal): Seien K ein Körper, n eine natürliche Zahl und V ein K-Vektorraum.

(i) Für $A \in K^{n\times n}$ heißt die Menge $I(A) := \{f \in K[X] \mid f(A) = \mathbf{0}\}$ *Verschwindungsideal von A.*

(ii) Für $\phi \in \mathrm{End}(V)$ heißt die Menge $I(\phi) := \{f \in K[X] \mid f(\phi) = \mathbf{0}\}$ *Verschwindungsideal von ϕ.*

Bemerkung VI.3.10 (Eigenschaften des Verschwindungsideals): (i) Die Verschwindungsideale sind nicht leer, denn das Nullpolynom gehört jedenfalls immer dazu.

(ii) Nach dem Satz von Cayley-Hamilton gehört zu $I(A)$ auch das charakteristische Polynom χ_A.

(iii) Sind p_1 und p_2 in $I(A)$, dann gehört wegen Definition VI.3.5 auch die Summe $p_1 + p_2$ zu $I(A)$, denn $(p_1 + p_2(A) = p_1(A) + p_2(A) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.

(iv) Sind h in $K[X]$ und p in $I(A)$, dann gehört wegen Definition VI.3.5 auch hp zu $I(A)$, denn $(hp)(A) = h(A)p(A) = h(A)\mathbf{0} = \mathbf{0}$.

Geleitet von den Beobachtungen der vorangegangenen Bemerkung wollen wir Teilmengen von Ringen mit den Eigenschaften (i), (iii) und (iv) Ideale nennen.

Definition VI.3.11 (Ideal): Seien R ein unitärer kommutativer Ring und $\varnothing \neq I \subseteq R$ eine Teilmenge. Falls gilt:

(i) Für alle $a, b \in I$ ist $a + b \in I$,

(ii) Für alle $a \in I$ und $r \in R$ ist $ra \in I$,

dann heißt I ein *Ideal*.

Proposition VI.3.12 (Konstruktion von Idealen): *Seien R ein kommutativer unitärer Ring und $a_1, \ldots, a_k$ Elemente von R. Dann ist*

$$I := Ra_1 + \cdots + Ra_k := (a_1, \ldots, a_k) := \{r_1a_1 + \cdots + r_ka_k \mid r_1, \ldots, r_k \in R\}$$

ein Ideal. Insbesondere erhalten wir für $a \in R$, dass $Ra := \{ra \mid r \in R\} = (a)$ ein Ideal ist. Dieses heißt das von a erzeugte Hauptideal.

Beweis: (i) Seien $a = r_1a_1 + \cdots + r_ka_k$ und $\bar{a} = \bar{r}_1a_1 + \cdots + \bar{r}_ka_k$ Elemente von I. Dann ist

$$a + \bar{a} = r_1a_1 + \cdots + r_ka_k + \bar{r}_1a_1 + \cdots + \bar{r}_ka_k = (r_1 + \bar{r}_1)a_1 + \cdots + (r_k + \bar{r}_k)a_k,$$

d. h. auch $a + \bar{a}$ gehört zu I.

(ii) Für $a = r_1a_1 + \cdots + r_ka_k$ in I und $r \in R$ ist $ra = (rr_1)a_1 + \cdots + (rr_k)a_k$ in I. □

Beispiel VI.3.13: Sei R ein kommutativer unitärer Ring. Dann gibt es zwei sogenannte *triviale Ideale*; zum Einen ist $N := (0) := \{0\}$ ein Ideal von R, das sogenannte *Nullideal*, und zum Anderen ist R selbst ein Ideal in R.

Definition VI.3.14: Sei R ein kommutativer unitärer Ring.

(i) Falls für alle r_1, r_2 in R gilt: „Ist $r_1r_2 = 0$, dann ist $r_1 = 0$ oder $r_2 = 0$“, dann heißt R *nullteilerfrei*.

(ii) Sei $I \subseteq R$ ein Ideal. Gibt es $m \in I$, sodass $I = Rm = (m)$, dann heißt I ein *Hauptideal*.

(iii) Sei R zusätzlich nullteilerfrei. Ist jedes Ideal von R ein Hauptideal, dann heißt R ein *Hauptidealring*.

Satz 23 (Polynomring als Hauptidealring): *Seien K ein Körper und $K[X]$ der Polynomring über K. Dann ist $K[X]$ ein Hauptidealring, d. h. für jedes Ideal $I \subseteq K[X]$ gibt es ein Polynom $p_0 \in I$, sodass $I = (p_0) = \{hp_0 \mid h \in K[X]\}$.*

Beweis: Wir wissen bereits, dass der Polynomring nullteilerfrei und kommutativ ist. Sei nun $I \subseteq K[X]$ ein Ideal. Ist I das Nullideal, dann ist I auch ein Hauptideal. Sonst gibt es $p_0 \in I - \{0\}$, sodass $\deg(p_0) \leq \deg(p)$ für alle $p \in I - \{0\}$. Wir zeigen jetzt, dass $I = \{hp_0 \mid h \in K[X]\}$.

Die Inklusion „$\supseteq$“ ist dabei klar, da p_0 zu I gehört und damit auch alle Vielfachen von p_0.

Für „$\subseteq$“ sei nun $p \in I$ gegeben. Da wir in $K[X]$ Polynomdivision durchführen können, gibt es Polynome h und r, sodass $p = hp_0 + r$ mit $\deg(r) < \deg(p_0)$. Wegen $r = p - hp_0$ folgt aus der Minimalität des Grades von p_0 schon, dass r das Nullpolynom sein muss. Also ist $p = hp_0$ wie gewünscht. □

Für den Nachweis, dass $K[X]$ ein Hauptidealring ist, haben wir nur die Polynomdivision aus Proposition VI.3.8 benötigt. Das bedeutet, Satz 23 kann auf nullteilerfreie kommutative Ringe mit Eins verallgemeinert werden, wenn es ein Analogon zu Proposition VI.3.8 gibt, d. h. wenn es einen Begriff von Teilbarkeit mit Rest gibt. Insbesondere lässt sich der Beweis auf sogenannte *euklidische Ringe* verallgemeinern, die wir später im Kapitel „Etwas mehr Strukturmathematik“ näher kennen lernen.

Korollar VI.3.15: *Der Ring der ganzen Zahlen $\mathbb{Z}$ ist ein Hauptidealring.*

Proposition VI.3.16 (Eindeutigkeit des Idealerzeugers): *Seien K ein Körper, $K[X]$ der Polynomring über K und $I \subseteq K[X]$ ein Ideal. Dann ist der Erzeuger p_0 aus dem Beweis von Satz 23 eindeutig bis auf einen skalaren, von Null verschiedenen Faktor. Genauer: Ist $I = K[X]f = K[X]g$ mit $f, g \in K[X]$, dann gibt es $\lambda \in K^\times$ mit $g = \lambda f$.*

Beweis: Für das Nullideal $I = (0)$ gilt die Aussage. Sei also jetzt $I \neq (0)$ mit $I = K[X]f = K[X]g$. Dann gibt es $h, h' \in K[X]$ mit $g = h'f$ und $f = hg$, d. h. $f = hg = hh'f$, sodass $(1 - hh')f = 0$. Weil f nach Voraussetzung nicht das Nullpolynom ist und $K[X]$ nullteilerfrei ist, muss $1 - hh' = 0$ sein, sodass $1 = hh'$. Damit gehören h und h' zu $K[X]^\times$ und $h' = \lambda X^0$ mit $\lambda \in K^\times$ wie gewünscht. □

Für den Beweis von Proposition VI.3.16 haben wir nur benötigt, dass $R = K[X]$ nullteilerfrei ist und haben daraus erhalten: Zwei Erzeuger desselben Hauptideals sind gleich bis auf Multiplikation mit einer Einheit. Das bedeutet: Ist $(a) = (b)$, dann gibt es ein s in $R^\times$ mit $b = sa$.

Definition VI.3.17 (Teiler in $K[X]$): Seien K ein Körper und $f, g \in K[X]$.

(i) Gibt es $h \in K[X]$ mit $g = hf$, dann heißt f ein *Teiler von g*. In diesem Fall schreiben wir „$f \mid g$“.

(ii) Gilt für alle $h \in K[X]$ mit $h \mid f$ und $h \mid g$, dass $h \in K[X]^\times$, dann heißen f und g *teilerfremd*.

(iii) Sei $f \in K[X] - (K^\times \cup \{0\})$. Gilt für alle h_1 und h_2 in $K[X]$ mit $f = h_1 h_2$, dass $h_1 \in K[X]^\times$ oder $h_2 \in K[X]^\times$, dann heißt f *irreduzibel*.

Ein Polynom heißt irreduzibel, falls es keine *echten* Teiler hat – durch Einheiten kann man immer teilen. Beispielsweise sind die Teiler von $1 = 1 \cdot X^0$ alle $a \cdot X^0$ für $a \in K^\times$. Genau so wird $(X - 2)$ von allen $a(X - 2)$ (a aus $K^\times$) geteilt.

Satz 24 (Lemma von Bézout): *Seien K ein Körper und $f, g \in K[X]$. Die Polynome f und g sind teilerfremd genau dann, wenn es h_1 und $h_2 \in K[X]$ mit $1 = h_1 f + h_2 g$ gibt.*

Beweis: „$\Longleftarrow$“: Sei $h \in K[X]$ mit $h \mid f$ und $h \mid g$, d. h. es gibt f' und g' in $K[X]$, sodass $f = hf'$ und $g = hg'$. Dann ist

$$1 = h_1 h f' + h_2 h g' = h(h_1 f' + h_2 g'),$$

sodass h eine Einheit sein muss.

„$\Longrightarrow$“: Sei $I = (f, g) = K[X]f + K[X]g$. Nach Proposition VI.3.12 ist I ein Ideal in $K[X]$. Wegen Satz 23 ist I außerdem ein Hauptideal, d. h. $I = (p_0)$ für ein $p_0 \in K[X]$. Weil f und g zu I gehören, ist p_0 ein Teiler von f und ein Teiler von g und da f und g teilerfremd sind, ist p_0 eine Einheit. Es gibt also $p_0' \in K[X]^\times$ mit $p_0' p_0 = 1$. Insbesondere gehört 1 zu I, d. h. es gibt h_1 und $h_2 \in K[X]$, sodass $1 = h_1 f + h_2 g$ wie gewünscht. □

Die Implikation „$\Longleftarrow$“ gilt in allen kommutativen unitären Ringen. Die Implikation „$\Longrightarrow$“ hat nur verwendet, dass $K[X]$ ein Hauptidealring ist. Somit gilt das Lemma von Bézout in allen Hauptidealringen, insbesondere auch im Ring der ganzen Zahlen $\mathbb{Z}$.

Korollar VI.3.18: *Seien a und b ganze Zahlen. Genau dann sind a und b teilerfremd, wenn es ganze Zahlen k und ℓ gibt, sodass $1 = ka + \ell b$.*

Definition VI.3.19 (Nullstellenmenge): Sei f ein Polynom mit Koeffizienten in K. Die Menge $\mathrm{Nst}(f) = \{a \in K \mid f(a) = 0\}$ heißt *Nullstellenmenge von f*.

Ist g ein weiteres Polynom mit Koeffizienten in K und wird f von g geteilt, dann ist $\mathrm{Nst}(g) \subseteq \mathrm{Nst}(f)$.

Proposition VI.3.20: *Seien K ein Körper und $f \in K[X]$.*

(i) *Genau dann ist $a \in K$ eine Nullstelle von f, wenn $(X - a)$ ein Teiler von f ist.*

(ii) *Gibt es $a_1, \ldots, a_n \in K$, sodass $f = \prod_{i=1}^{n}(X - a_i)$ und ist $h \in K[X]$ ein normierter Teiler von f, dann ist $h = \prod_{j=1}^{k}(X - a_{i_j})$, wobei $i_1, \ldots, i_k$ in $\{1, \ldots, n\}$ mit $1 \le i_1 < \cdots < i_k \le n$.*

Beweis: (i) „$\Longleftarrow$“: Angenommen, $f = g \cdot (X - a)$ für ein $g \in K[X]$. Dann ist $f(a) = 0 \cdot g(a) = 0$, da Einsetzen ein Algebrenhomomorphismus ist.

„$\Longrightarrow$“: Polynomdivision ergibt, dass $f = g \cdot (X - a) + r$ mit $\deg(r) \le 0$. Aber das heißt $r = cX^0$ mit $c \in K$ und wegen $0 = f(a) = g(a) \cdot 0 + r(a) = c$ muss sogar $c = 0$ gelten. Damit ist $(X - a)$ ein Teiler von f, wie gewünscht.

(ii) Wir schreiben $f = gh$ für ein Polynom g mit Koeffizienten in K, d. h. $(X - a_1) \cdots (X - a_n) = gh$. Weil K ein Körper ist, folgt aus $0 = f(a_1) = g(a_1)h(a_1)$, dass $g(a_1) = 0$ oder $h(a_1) = 0$. Diese Tatsache benutzen wir, um iterativ wie folgt die Funktion $\theta\colon \{a_1, \ldots, a_n\} \to \{0, 1\}$ zu definieren:

$$\theta(a_1) = \begin{cases} 0, & \text{falls } g(a_1) = 0, \\ 1, & \text{falls } h(a_1) = 0. \end{cases}$$

Im ersten Fall ist $g = (X - a_1)\tilde{g}$ für ein Polynom $\tilde{g}$ mit Koeffizienten in K, im zweiten Fall genau so für h. Weiter setzen wir im ersten Fall $\tilde{h} = h$ und im zweiten Fall $\tilde{g} = g$. Für $\tilde{g}$ und $\tilde{h}$ haben wir dann $(X - a_2) \cdots (X - a_n) = \tilde{g}\tilde{h}$.

Iterativ erhalten wir $1 = \tilde{g}\tilde{h}$ für Polynome $\tilde{g}$ und $\tilde{h}$, für die

$$g = \prod_{i:\theta(a_i)=0} (X - a_i)\tilde{g} \quad \text{respektive} \quad h = \prod_{i:\theta(a_i)=1} (X - a_i)\tilde{h} \tag{VI.1}$$

gilt. Insbesondere ist $\deg(\tilde{g}) = \deg(\tilde{h}) = 0$, und aus Gl. (VI.1) zusammen mit der Normiertheit von h folgt $\tilde{h} = 1$. Dann ist Gl. (VI.1) genau das Behauptete. □

Korollar VI.3.21 (Kriterium Teilerfremdheit): *Seien K ein Körper, $a \in K$ und $n \in \mathbb{N}_0$. Ferner seien f und g Polynome in $K[X]$ mit $f = (X - a)^n$, und $g(a) \neq 0$. Dann sind f und g teilerfremd.*

Beweis: Sei t ein normierter Teiler von f und g. Nach Proposition VI.3.20(ii) gibt es eine nichtnegative ganze Zahl k, sodass $t = (X - a)^k$. Wegen Proposition VI.3.20(i) ist $k = 0$, und so t eine Einheit in $K[X]$. □

4. Das Minimalpolynom

In diesem Abschnitt wollen wir das kleinste Polynom kennenlernen, das einen Endomorphismus bzw. eine Matrix annulliert. Dieses Polynom heißt *Minimalpolynom*. Wir werden feststellen, dass die Nullstellenmenge des Minimalpolynoms mit der des charakteristischen Polynoms übereinstimmt, also gleich dem Spektrum des Endomorphismus bzw. der Matrix ist.

Bemerkung VI.4.1: Seien n eine natürliche Zahl, K ein Körper, V ein K-Vektorraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus bzw. $A \in K^{n\times n}$. Da $K[X]$ ein Hauptidealring ist, haben die Verschwindungsideale $I(\phi)$ respektive $I(A)$ eindeutige normierte Erzeuger kleinstens Grades m_ϕ respektive m_A.

Definition VI.4.2 (Minimalpolynom): Seien n eine natürliche Zahl, K ein Körper, V ein K-Vektorraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus bzw. $A \in K^{n\times n}$. Der Erzeuger m_ϕ respektive m_A des Verschwindungsideals $I(\phi)$ respektive $I(A)$ heißt *Minimalpolynom von* ϕ respektive *Minimalpolynom von* A.

Insbesondere ist das Minimalpolynom dasjenige Polynom f kleinsten nichtnegativen Grades, das ϕ respektive A annulliert, d. h. das $f(\phi) = \mathbf{0}$ respektive $f(A) = \mathbf{0}$ leistet.

Bemerkung VI.4.3: Aus dem Satz von Cayley-Hamilton folgt, dass m_ϕ ein Teiler des charakteristischen Polynoms χ_ϕ ist und genauso, dass m_A ein Teiler des charakteristischen Polynoms χ_A ist.

Proposition VI.4.4: *Seien n eine natürliche Zahl, K ein Körper, V ein K-Vektorraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus bzw. $A \in K^{n\times n}$. Dann gilt:*

(i) $\mathrm{Nst}(m_\phi) = \mathrm{Nst}(\chi_\phi) = \mathrm{Spec}(\phi)$,

(ii) $\mathrm{Nst}(m_A) = \mathrm{Nst}(\chi_A) = \mathrm{Spec}(A)$.

Beweis: Wir zeigen (ii), (i) geht völlig analog. Wie in Proposition VI.3.19 bemerkt, haben wir $\mathrm{Nst}(m_A) \subseteq \mathrm{Nst}(\chi_A)$. Bleibt also „$\supseteq$“ zu zeigen.

Sei dazu $\lambda \in \mathrm{Nst}(\chi_A)$. Dann ist λ ein Eigenwert von A, d. h. $m_A(\lambda)$ ist ein Eigenwert von $m_A(A) = \mathbf{0}$. Wir haben also $m_A(\lambda) = 0$, d. h. $\lambda \in \mathrm{Nst}(m_A)$. □

Beispiel VI.4.5: Es sei K ein Körper.

(i) Für $A = \mathrm{diag}(3, 3, 1)$ist $\chi_A = (X-3)^2(X-1)$ das zugehörige charakteristische Polynom. Wir setzen $p = (X-3)(X-1)$. Nach Proposition I.4.4 und Proposition I.3.20 wird χ_A von p geteilt und setzen wir A in p ein, so erhalten wir

$$p(A) = (A - 3I_3)(A - I_3) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{0},$$

d. h. p ist das Minimalpolynom.

(ii) Es seien $\lambda_1, \ldots, \lambda_n \in K$ und $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Das charakteristische Polynom von A ist $\chi_A = \prod_{i=1}^n (X - \lambda_i)$, aber das Minimalpolynom von A ist $m_A = \prod_{\lambda\in\mathrm{Spec}(A)}(X - \lambda)$. Insbesondere haben wir $\deg m_A = \#\,\mathrm{Spec}(A)$ und $\deg \chi_A = n$.

(iii) Seien $a \in K$ und

$$A = \begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix} \in K^{4\times 4}.$$

Da A eine untere Dreiecksmatrix ist, ist $\chi_A = \det(A - XI_4) = (X-a)^4$. Wir haben

$$A - aI_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

und da wir die Potenzen von $(A - aI_4)$ schon ausgerechnet haben, wissen wir bereits, dass erst $(A - aI_4)^4 = \mathbf{0}$. In diesem Fall ist also das Minimalpolynom gleich dem charakteristischen Polynom.

(iv) Betrachte $\Phi\colon K^{n\times n} \to K^{n\times n}$, $A \mapsto A^t$. Wir wissen, dass $(A^t)^t = A$, sodass $\Phi^2 = \mathrm{id}_{K^{n\times n}}$ gilt. Das Polynom $X^2 - 1$ annulliert also Φ. Ein Polynom kleineren nicht-negativen Grades kann nicht annullieren, sodass $m_\Phi = X^2 - 1$. Da $\dim K^{n\times n} = n^2$ ist auch $\deg \chi_\Phi = n^2$, die Eigenwerte von Φ sind ± 1. Überlegen Sie sich, wie das charakteristische Polynom von Φ aussieht!

Um uns das Leben leichter zu machen, möchten wir im Folgenden das charakteristische Polynom in teilerfremde Faktoren zerlegen und diese zuerst bearbeiten.

Notation VI.4.6: Seien K ein Körper, n eine natürliche Zahl und $A \in K^{n\times n}$. Mit $\phi_A\colon K^n \to K^n$ bezeichnen wir die zugehörige lineare Abbildung und wir setzen $\mathrm{Bild}(A) := \mathrm{Bild}(\phi_A)$ sowie $\mathrm{Kern}(A) := \mathrm{Kern}(\phi_A)$.

Lemma VI.4.7 (Zerlegungslemma für Matrizen): *Seien K ein Körper, n eine natürliche Zahl, $A \in K^{n\times n}$ und g_1, g_2 sowie g Polynome in $K[X]$.*

(i) *Die Räume* $\mathrm{Kern}(g(A))$ *und* $\mathrm{Bild}(g(A))$ *sind A-invariante Untervektorräume des K^n.*

(ii) *Sind g_1 und g_2 teilerfremd, dann gilt*

$$\mathrm{Kern}(g_1(A)) \cap \mathrm{Kern}(g_2(A)) = \{\mathbf{0}\}, \qquad \mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A)) = K^n.$$

(iii) *Ist $(g_1 g_2)(A) = \mathbf{0}$, also $g_1 g_2 \in I(A)$, dann gilt*

$$\mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_1(A)), \qquad \mathrm{Bild}(g_1(A)) \subseteq \mathrm{Kern}(g_2(A)).$$

(iv) *Sind g_1 und g_2 teilerfremd und $g_1g_2 \in I(A)$, dann haben wir*

$$K^n = \mathrm{Kern}(g_1(A)) \oplus \mathrm{Kern}(g_2(A))$$

sowie $\mathrm{Kern}(g_1(A)) = \mathrm{Bild}(g_2(A))$ *und* $\mathrm{Kern}(g_2(A)) = \mathrm{Bild}(g_1(A))$.

Beweis: (i) Gehört v zu $\mathrm{Kern}(g(A))$, dann ist $g(A)v = \mathbf{0}$. Die Matrizen A und $g(A)$ kommutieren, weshalb wir $g(A)Av = Ag(A)v = A\mathbf{0} = \mathbf{0}$ haben, d. h. Av gehört zu $\mathrm{Kern}(g(A))$.

Gehört v zu $\mathrm{Bild}(g(A))$, dann gibt es irgendein $w \in K^n$ mit $v = g(A)w$. Wieder ist $Av = Ag(A)w = g(A)Aw$, und das ist ein Element von $\mathrm{Bild}(g(A))$.

(ii) Aus dem Lemma von Bézout wissen wir, dass es Polynome h_1 und h_2 mit $1 = h_1g_1 + h_2g_2$ gibt. Setzen wir A in diese Darstellung ein, so erhalten wir

$$I_n = h_1(A)g_1(A) + h_2(A)g_2(A) = g_1(A)h_1(A) + g_2(A)h_2(A).$$

Für $v \in \mathrm{Kern}(g_1(A)) \cap \mathrm{Kern}(g_2(A))$ erhalten wir mit der ersten Gleichheit aus der obigen Gleichung, dass

$$v = h_1(A)g_1(A)v + h_2(A)g_2(A)v = h_1(A)\mathbf{0} + h_2(A)\mathbf{0} = \mathbf{0}$$

und für $v \in K^n$ ist wegen der zweiten Gleichheit in der ersten Gleichung, dass $v = g_1(A)h_1(A)v + g_2(A)h_2(A)v$, d. h. v gehört zu $\mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A))$.

(iii) Es sei $v = g_2(A)w$ ein Element von $\mathrm{Bild}(g_2(A))$. Dann ist

$$g_1(A)v = g_1(A)g_2(A)w = (g_1g_2)(A)w = \mathbf{0},$$

sodass v zu $\mathrm{Kern}(g_1(A))$ gehört und deshalb $\mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_1(A))$. Die zweite Aussage zeigt man genau mit Vertauschung der Indizes.

(iv) Aus (ii) wissen wir, dass $K^n = \mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A))$. Aus (iii) wissen wir außerdem, dass $\mathrm{Bild}(g_i(A)) \subseteq \mathrm{Kern}(g_j(A))$ für $i \in \{1, 2\}$ und $j \in \{1, 2\} - \{i\}$. Wieder wegen (ii) ist $K^n = \mathrm{Kern}(g_2(A)) \oplus \mathrm{Kern}(g_1(A))$, d. h. wir haben $\mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_2(A)) \oplus \mathrm{Kern}(g_1(A))$. Aus Dimensionsgründen folgern wir jetzt

$$\mathrm{Bild}(g_1(A)) = \mathrm{Kern}(g_2(A)), \qquad \mathrm{Bild}(g_2(A)) = \mathrm{Kern}(g_1(A)). \qquad \square$$

Korollar VI.4.8 (Zerlegungslemma für Endomorphismen): *Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $\phi\colon V \to V$ eine lineare Abbildung und $f = g_1g_2$ ein Polynom über K mit $f \in I(\phi)$ und teilerfremden g_1, g_2. Dann gilt*

$$V = \mathrm{Kern}(g_1(\phi)) \oplus \mathrm{Kern}(g_2(\phi))$$

sowie $\mathrm{Kern}(g_1(\phi)) = \mathrm{Bild}(g_2(\phi))$ *und* $\mathrm{Kern}(g_2(\phi)) = \mathrm{Bild}(g_1(\phi))$. *Hierbei sind* $\mathrm{Kern}(g_1(\phi))$ *und* $\mathrm{Kern}(g_2(\phi))$ *zwei ϕ-invariante Unterräume von V.*

Korollar VI.4.9: *Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $A \in K^{n\times n}$ eine Matrix und $\phi\colon V \to V$ ein Endomorphismus. Zerfällt m_A beziehungsweise m_ϕ in Linearfaktoren, d. h. $m_A = \prod_{i=1}^{r}(X-\lambda_i)^{e_i}$ beziehungsweise $m_\phi = \prod_{i=1}^{r}(X-\lambda_i)^{e_i}$, dann gilt*

$$K^n = \bigoplus_{i=1}^{r} \operatorname{Kern}\left((A-\lambda_i I_n)^{e_i}\right) \quad \textit{bzw.} \quad V = \bigoplus_{i=1}^{r} \operatorname{Kern}\left((\phi-\lambda_i \operatorname{id}_V)^{e_i}\right).$$

Wir schreiben $H_i = \operatorname{Kern}((\phi-\lambda_i \operatorname{id}_V)^{e_i})$.

$H_1, \ldots, H_r$ werden wir später als Haupträume wiedersehen. Wir halten fest, dass die obige Aussage genau so für das charakteristische Polynom gilt.

5. Nilpotente Endormorphismen

In diesem Abschnitt leiten wir ⟨ Satz 1 ⟩ für nilpotente Endomorphismen her. Als entscheidende Bausteine dafür zeigen wir in Proposition I.5.8 und Korollar I.5.9, dass sich ein endlichdimensionaler Vektorraum V, gegeben einen nilpotenten Endomorphismus ϕ, in eine direkte Summe ϕ-zyklischer Unterräume zerlegen lässt.

In diesem Abschnitt seien stets V ein n-dimensionaler Vektorraum über dem Körper K, ϕ ein Endomorphismus von V und A eine $n \times n$-Matrix mit Einträgen aus K.

Definition VI.5.1 (Nilpotent): Gibt es eine natürliche Zahl k, sodass $\phi^k = \mathbf{0}$, dann heißt ϕ *nilpotent*. Gibt es ein solches k, sodass $A^k = \mathbf{0}$, dann heißt A *nilpotent*.

Beispiel VI.5.2 (Lieblings-nilpotente-Matrix): Sei

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

Nach Proposition VI.4.5(iii) ist $A^4 = \mathbf{0}$, d. h. A ist nilpotent.

Bemerkung VI.5.3: Ist A eine nilpotente Matrix beziehungsweise ϕ ein nilpotenter Endomorphismus, dann ist $\chi_A = (-1)^n X^n$ beziehungsweise $\chi_\phi = (-1)^n X^n$. Das entsprechende Minimalpolynom ist X^k für eine natürliche Zahl $k \leq n$.

Proposition VI.5.4 (Spektrum nilpotenter Endomorphismen): *Ist A beziehungsweise ϕ nilpotent, dann ist* $\mathrm{Spec}(A) = \{0\}$ *beziehungsweise* $\mathrm{Spec}(\phi) = \{0\}$.

Beweis: Wir zeigen die Behauptung ausschließlich für Matrizen, für Endomorphismen greifen dieselben Argumente. Ist A nilpotent, dann gibt es eine natürliche Zahl k, sodass $A^k = \mathbf{0}$. Ist λ ein Eigenwert von A, dann gibt es einen von Null verschiedenen Vektor v in V, sodass $Av = \lambda v$. Dann ist auch $A^k v = \mathbf{0} = \lambda^k v$, d. h. $\lambda^k = 0$ muss gelten, weil v von Null verschieden ist. Und weil K nullteilerfrei ist, war λ selbst bereits Null. □

Definition VI.5.5 (Zyklischer Unterraum): Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler Vektorraum über K, ϕ ein Endomorphismus von V und U ein Untervektorraum von V. Ist U ein ϕ-invarianter Unterraum und gibt es einen Vektor u_0 in U sowie eine natürliche Zahl k, sodass $U = \mathrm{Lin}(u_0, \phi(u_0), \ldots, \phi^{k-1}(u_0))$, dann heißt U ein *ϕ-zyklischer Unterraum*.

Das heißt: U ist der kleinste ϕ-zyklische Unterraum, der u_0 enthält (vergleiche Proposition VI.2.7).

Bemerkung VI.5.6 (Entdeckung des größten Jordan-Kästchens): Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler K-Vektorraum und $\phi\colon V \to V$ ein nilpotenter Endomorphismus mit Minimalpolynom von Grad $d = \deg(m_\phi)$. Dann ist U ein ϕ-zyklischer Unterraum $U = \mathrm{Lin}(u_0, \phi(u_0), \ldots, \phi^{d-1}(u_0))$, $B = (u_0, \phi(u_0), \ldots, \phi^{d-1}(u_0))$ ist eine Basis von U und die Darstellungsmatrix von $\phi|_U$ bezüglich B ist

$$J := D_{B,B}(\phi|_U) = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} =: J_d.$$

Im Folgenden verfahren wir nach dem Prinzip „Teile und Herrsche“. Wir zerlegen $V = U \oplus W$, wobei U ein maximaler ϕ-zyklischer Unterraum von V ist, und zerlegen dann W weiter.

Definition VI.5.7 (ϕ-invariantes Komplement): Seien U ein ϕ-invarianter Unterraum von V und W ein weiterer Unterraum von V. Falls $V = U \oplus W$ und falls W ebenfalls ϕ-invariant ist, dann heißt W ein *ϕ-invariantes Komplement*.

Proposition VI.5.8: *Seien K ein Körper, n eine natürliche Zahl, V ein Vektorraum über K der Dimension n, $\phi\colon V \to V$ ein nilpotenter Endomorphismus, d der Grad des Minimalpolynoms von ϕ und $u_0 \in V$ mit $\phi^{d-1}(u_0) \neq \mathbf{0}$. Dann hat der ϕ-zyklische Unterraum $U = \mathrm{Lin}(u_0, \phi(u_0), \ldots, \phi^{d-1}(u_0))$ ein ϕ-invariantes Komplement.*

Beweis: Wir schreiben

$$\mathfrak{M} := \{W' \mid W' \subseteq V \text{ ist } \phi\text{-invarianter Unterraum mit } W' \cap U = \{\mathbf{0}\}\}$$

und halten fest, dass $\mathfrak{M}$ wegen $\{\mathbf{0}\} \in \mathfrak{M}$ nichtleer ist. Wir wählen ein Element W von $\mathfrak{M}$ mit maximaler Dimension.

Wir wollen zeigen, dass W das gesuchte ϕ-invariante Komplement ist, d. h. dass $V = U \oplus W$. Das erreichen wir, indem wir „$U \oplus W \subsetneq V$“ zu einem Widerspruch führen.

Angenommen, $U \oplus W$ wäre ein echter Unterraum von V. Dann fänden wir ein $v' \in V - (U \oplus W)$. Wegen $\phi^d = \mathbf{0}$ wäre insbesondere $\phi^d(v') = \mathbf{0}$, und $\mathbf{0}$ gehörte zu $U \oplus W$. Wegen $v \notin U \oplus W$ hätten wir außerdem $\phi^0(v) = v \notin U \oplus W$. Sei also nun $\ell \in \{1, \ldots, d\}$ minimal mit der Eigenschaft, dass $\phi^\ell(v') \in U \oplus W$. Wir schreiben $v := \phi^{\ell-1}(v')$.

Wegen $\phi(v) \in U \oplus W$ fänden wir eindeutige $u' \in U$ und $w \in W$, sodass $\phi(v) = u' + w$ und da wir eine Basis von U kennen, könnten wir dieses u' mit geeigneten $\alpha_0, \ldots, \alpha_{d-1} \in K$ schreiben als $u' = \sum_{i=0}^{d-1} \alpha_i \phi^i(u_0)$. Anwendung von ϕ^{d-1} auf $\phi(v)$ ergäbe

$$\mathbf{0} = \alpha_0 \phi^{d-1}(u_0) + \phi^{d-1}(w),$$

was wegen $U \cap W = \{\mathbf{0}\}$ zunächst erzwänge, dass $\alpha_0 \phi^{d-1}(u_0) = \mathbf{0}$ sowie $\phi^{d-1}(w) = \mathbf{0}$, und schließlich wegen $\phi^{d-1}(u_0) \neq \mathbf{0}$, dass $\alpha_0 = 0$ gelten müsste. Mit $u := \sum_{i=1}^{d-1} \alpha_i \phi^{i-1}(u_0)$ erhielten wir so, dass $\phi(v) = \phi(u) + w$.

Das aber erlaubte es uns, w zu schreiben als $w = \phi(v) - \phi(u) = \phi(v - u)$, d. h. $\phi(v - u)$ gehörte zu W. Außerdem gehörte $v - u$ nicht zu W, denn andernfalls müsste schon v zu $U \oplus W$ gehört haben. Wir könnten also W echt vergrößern zu $W + \mathrm{Lin}(v - u)$ (d. h. $W \subsetneq W + \mathrm{Lin}(v - u)$).

Wäre jetzt w' ein Element von $(W + \mathrm{Lin}(v - u)) \cap U$, dann gäbe es $w \in W$ und ein $c \in K$, sodass $w' = w + c(v - u)$ ein Element von U wäre. Wäre $c \neq 0$, dann erhielten wir, dass $v = c^{-1}(w' + cu - w)$ zu $U \oplus W$ gehörte, was wir ausgeschlossen haben. Es müsste also $c = 0$ gelten, sodass $w' = w$ in Wahrheit zu W gehörte, was wegen $W \cap U = \{\mathbf{0}\}$ zur Folge hätte, dass $w' = \mathbf{0}$.

Insgesamt hätten wir ein Element W' von $\mathfrak{M}$ erhalten, das W echt enthielte, was der Wahl von W widerspräche. Es muss also $V = U \oplus W$ gelten und W ist das gewünschte ϕ-invariante Komplement von U. □

Korollar VI.5.9 (Zerlegung in zyklische Unterräume): *Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n und $\phi\colon V \to V$ ein nilpotenter Endomorphismus. Dann ist V die direkte Summe ϕ-zyklischer Unterräume.*

Die Aussage zeigt man per vollständiger Induktion mit der Aussage aus Proposition VI.5.8.

Korollar VI.5.10 (Jordan-Normalform für nilpotente Matrizen): *Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler Vektorraum über K und $\phi\colon V \to V$ ein nilpotenter Endomorphismus. Dann gibt es eine Basis B von V, sodass*

$$D_{B,B}(\phi) = \begin{pmatrix} J_{d_1} & & \\ & \ddots & \\ & & J_{d_s} \end{pmatrix}$$

wobei $d_1 \geq d_2 \geq \cdots \geq d_s \geq 1$. Hierbei sind die Matrizen J_{d_i} definiert wie in Proposition VI.5.6 und heißen Jordan-Kästchen *für ϕ.*

Bemerkung VI.5.11: Sei $J_d = \sum_{i=1}^{d} E_{i,i-1}$ das Jordan-Kästchen aus Proposition VI.5.6. Dann gilt $J_d^\ell = \sum_{i=\ell+1}^{d} E_{i,i-\ell}$, d. h. J_d^ℓ hat Einsen auf der ℓ-ten unteren Nebendiagonale. Insbesondere ist

$$\operatorname{Rang}(J_d^\ell) = \begin{cases} d-\ell, & \text{falls } d-\ell \geq 0, \\ 0, & \text{sonst.} \end{cases}$$

Das zeigt man per Induktion mithilfe der Formel $E_{i,j}E_{k,l} = \delta_{k,j}E_{i,\ell}$.

Proposition VI.5.12 (Kenngrößen für die Jordan-Normalform): *Seien K ein Körper, n eine natürliche Zahl, V ein Vektorraum der Dimension n über K, und ϕ ein nilpotenter Endomorphismus von V. Für die Darstellungsmatrix $D_{B,B}(\phi) = J$ aus Proposition VI.5.10 gilt:*

(i) *Die Summe der $d_1, \ldots, d_s$ ist n.*

(ii) *Sei m_k die Anzahl der Jordan-Kästchen der Länge k und $r_k = \operatorname{Rang}(\phi^k)$. Dann gilt für jedes k, dass $m_k = r_{k-1} - 2r_k + r_{k+1}$.*

(iii) *Das größte Jordan-Kästchen hat Größe $\deg(m_\phi)$.*

(iv) *Die Anzahl s der Jordan-Kästchen ist $\dim \operatorname{Eig}(\phi, 0)$.*

Beweis: (i) Klar.

(ii) Für das Jordan-Kästchen J_{d_i} gilt $\operatorname{Rang}(J_{d_i}^k) = \max\{0, \ldots, d_i - k\}$. Damit erhalten wir

$$r_k = \operatorname{Rang}(J_{d_i}^k) = \operatorname{Rang}(J^k) = \sum_{i=1}^{s} \operatorname{Rang}(J_{d_i}^k) = \sum_{d=k+1}^{n} m_d(d-k),$$

weil nur die Kästchen der Länge $k+1$ und größer überleben. Es ist also

$$r_{k-1} - r_k = \sum_{d=k}^{n} m_d(d-(k-1)) - \sum_{d=k+1}^{n} m_d(d-k) = m_k + \sum_{d=k+1}^{n} m_d = \sum_{d=k}^{n} m_d,$$

sodass $m_k = \sum_{d=k}^{n} m_d - \sum_{d=k+1}^{n} m_d = r_{k-1} - r_k - (r_k - r_{k+1})$ wie behauptet.

(iii) Sei d der Grad von m_ϕ, d. h. $\phi^d = \mathbf{0}$. Aus (ii) folgt für $k \geq d+1$, dass $m_k = 0$ und $m_d = r_{d-1} = \operatorname{Rang}(\phi^{d-1}) > 0$.

(iv) Wir haben

$$\begin{aligned}\sum_{d=1}^{n} m_d &= r_0 - r_1 \\ &= \operatorname{Rang}(\phi^0) - \operatorname{Rang}(\phi) \\ &= n - (n - \dim \operatorname{Kern}(\phi)) = \dim \operatorname{Kern} \phi = \dim \operatorname{Eig}(\phi, 0).\end{aligned}$$ □

6. Jordan-Normalform

Für diesen Abschnitt seien stets V ein n-dimensionaler Vektorraum über dem Körper K, ϕ ein Endomorphismus von V mit zugehörigem Minimalpolynom m_ϕ und Spektrum $\operatorname{Spec}(\phi) = \{\lambda_1, \ldots, \lambda_r\}$, d. h. $m_\phi = f \prod_{i=1}^{r} (X - \lambda_i)^{e_i}$ für irgendein irreduzibles Polynom f in $K[X]$ (insbesondere hat f keine Nullstellen in K).

Ist der zugrundeliegende Körper K algebraisch abgeschlossen, dann ist f aus der obigen Produktdarstellung eine Einheit.

Definition VI.6.1 (Hauptraum): Sei λ_i ein Eigenwert von ϕ. Der Vektorraum $H_{\lambda_i} = \operatorname{Kern}(\phi - \lambda_i \operatorname{id}_V)^{e_i}$ der *Hauptraum zu* λ_i. Ist λ ein Element von $K - \operatorname{Spec}(\phi)$, dann setzen wir $H_\lambda = \{\mathbf{0}\}$.

Bemerkung VI.6.2 (Summe der Haupträume): Aus Proposition VI.4.8 folgt direkt, dass die Summe der Haupträume $H_{\lambda_1} \oplus \cdots \oplus H_{\lambda_r}$ direkt ist. Ist K darüber hinaus algebraisch abgeschlossen, dann folgt außerdem, dass diese Summe der ganze Raum V ist. Ferner sind die Haupträume ϕ-invariant.

Proposition VI.6.3 (Minimaler invarianter Unterraum, der Eigenraum enthält): *Seien λ ein Eigenwert von ϕ und $V = U \oplus W$ eine Zerlegung von V in ϕ-invariante Unterräume, sodass* $\mathrm{Eig}(\phi, \lambda) = \mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)$ *in U enthalten ist. Dann sind bereits alle Potenzen* $\mathrm{Eig}(\phi, \lambda)^k$ *für natürliche Zahlen k in U enthalten.*

Beweis: Wir gehen in zwei Schritten vor. Zunächst zeigen wir, dass ein Unterraum U genau dann ϕ-invariant ist, wenn er $\phi - \lambda \,\mathrm{id}_V$-invariant ist. Dann zeigen wir per Induktion die Aussage zu den Potenzen.

Zum ersten Schritt: „$\Longrightarrow$“: Sei u ein Element von U. Dann haben wir $(\phi - \lambda \,\mathrm{id}_V)(u) = \phi(u) - \lambda u$, was zu U gehört.

„$\Longleftarrow$“ folgt wegen $\phi = (\phi - \lambda \,\mathrm{id}_V) - (-\lambda) \,\mathrm{id}_V$ aus „$\Longrightarrow$“.

Zum zweiten Schritt: Für $k = 1$ gilt die Aussage per Voraussetzung. Die Aussage gelte für die natürliche Zahl k. Sei v ein Element von $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^{k+1}$. Wegen $V = U \oplus W$ gibt es eindeutig bestimmte Vektoren u aus U und w aus W, sodass $v = u + w$. Das heißt

$$\mathbf{0} = (\phi - \lambda \,\mathrm{id}_V)^{k+1}(v) = (\phi - \lambda \,\mathrm{id}_V)^{k+1}(u) + (\phi - \lambda \,\mathrm{id}_V)^{k+1}(v),$$

wobei rechts nach (i) je ein Element von U und eins von W stehen. Weil aber $U \cap W = \{\mathbf{0}\}$ per Voraussetzung ist, ist $(\phi - \lambda \,\mathrm{id}_V)^{k+1}(w) = \mathbf{0}$, was bedeutet, dass $(\phi - \lambda \,\mathrm{id}_V)^k(w)$ in $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)$ liegt – und dieser Kern ist nach Voraussetzung in U enthalten.

Andererseits ist der Vektor $(\phi - \lambda \,\mathrm{id}_V)^k(w)$ nach (i) ein Element von W, also gilt $(\phi - \lambda \,\mathrm{id}_V)^k(w) = \mathbf{0}$ wegen $U \cap W = \{\mathbf{0}\}$. Nach Induktionsvoraussetzung ist w in $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^k$, also in U, enthalten, d. h. w liegt in $U \cap W = \{\mathbf{0}\}$, $v = u$ und liegt in U. □

Wir halten fest, dass die Räume $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^k$ eine aufsteigende Kette von Unterräumen bilden, d. h. $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V) \subseteq \mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^2 \subseteq \ldots$

Korollar VI.6.4 (Hauptraum enthält alle Potenzen des Eigenraums): *Seien λ_i ein Eigenwert von ϕ und e_i sein Exponent im Minimalpolynom m_ϕ. Dann ist*

$$H_{\lambda_i} = \mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^{e_i} = \bigcup_{k \in \mathbb{N}} \mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^k,$$

d. h., für jedes $k \geq e_i$ gilt $H_{\lambda_i} = \mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^k$. Man sagt auch, die Kette $(\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^k)_{k \in \mathbb{N}}$ *werde* stationär.

Beweis: Der Hauptraum H_{λ_i} ist ϕ-invariant nach Proposition VI.6.2 und der Rest der Behauptung ist Proposition VI.6.3. □

Definition VI.6.5 (Algebraische Vielfachheit): Sei λ ein Eigenwert von ϕ. Der Exponent des größten Teilers $(X-\lambda)^k$ des charakteristischen Polynom χ_ϕ heißt *algebraische Vielfachheit von* λ und wird mit $\mu_a(\phi,\lambda)$ bezeichnet. Den Exponent des größten Teilers $(X-\lambda)^k$ des Minimalpolynoms m_ϕ bezeichnen wir mit $e(\phi,\lambda)$.

Proposition VI.6.6 (Dimension der Haupträume): *Sei λ ein Eigenwert von ϕ. Dann ist* $\dim H_\lambda = \mu_a(\phi,\lambda)$ *und $\phi|_{H_\lambda}$ hat das charakteristische Polynom* $(X-\lambda)^{\mu_a(\phi,\lambda)}$.

Beweis: Weil λ ein Eigenwert von ϕ ist, können wir $\chi_\phi = (X-\lambda)^e g$ schreiben, wobei λ keine Nullstelle von g ist. Das bedeutet nach Proposition VI.3.20, dass $(X-\lambda)^e$ und g teilerfremd sind.

Per Zerlegungslemma (Proposition VI.4.8) zerfällt V als direkte Summe ϕ-invarianter Unterräume $V = H_\lambda \oplus W$, und $H_\lambda = \operatorname{Kern}(\phi - \lambda \operatorname{id}_V)^e$.

Nun betrachten wir die Einschränkungen $\phi_1 = \phi|_{H_\lambda}$ und $\phi_2 = \phi|_W$, um deren charakteristische Polynome zu bestimmen. Wegen $V = H_\lambda \oplus W$ gilt $\chi_\phi = \chi_{\phi_1}\chi_{\phi_2}$.

Per Definition von H_λ ist $(\phi - \lambda \operatorname{id}_V)^e v = \mathbf{0}$ für jedes v aus H_λ, was bedeutet, dass $(\phi - \lambda \operatorname{id}_{H_\lambda})^e = \mathbf{0}$. Der Endomorphismus $\psi_1 = \phi_1 - \lambda \operatorname{id}_{H_\lambda}$ ist also nilpotent, und $\chi_{\psi_1} = X^{\dim H_\lambda}$. Darum gilt $\chi_{\phi_1} = (X-\lambda)^{\dim H_\lambda}$.

Jetzt haben wir die zwei Zerlegungen $(X-\lambda)^e g = \chi_{\phi_1}\chi_{\phi_2} = (X-\lambda)^{\dim H_\lambda}\chi_{\phi_2}$, und einerseits ist $\dim H_\lambda \le e$, da $(X-\lambda)$ sonst ein Teiler von g wäre, was wir ausgeschlossen haben, und andererseits ist $\dim H_\lambda \ge e$, weil $(X-\lambda)$ sonst ein Teiler von χ_{ϕ_2} wäre, was ebenfalls nicht sein kann. Es folgt $e = \dim H_\lambda$, wie behauptet. □

Bemerkung VI.6.7 (Eigenwerte und Haupträume): Seien λ ein Eigenwert von ϕ, H_λ der zugehörige Hauptraum und $V = H_\lambda \oplus W$ eine Zerlegung von V in ϕ-invariante Unterräume. Dann ist λ kein Eigenwert von $\phi|_W$, das Spektrum von ϕ ist die disjunkte Vereinigung von $\{\lambda\}$ und $\operatorname{Spec}(\phi|_W)$, und ist λ' ein von λ verschiedener Eigenwert von ϕ, dann ist der Hauptraum $H_{\lambda'}$ in W enthalten.

Beweis: Weil der Eigenraum $\operatorname{Eig}(\phi,\lambda)$ im Hauptraum H_λ enthalten ist, ist die erste Behauptung klar.

Die Inklusion „$\{\lambda\} \cup \operatorname{Spec}(\phi|_W) \subseteq \operatorname{Spec}(\phi)$“ ist klar. Für die Inklusion „$\supseteq$“ seien λ' ein Eigenwert von ϕ' und v ein zugehöriger Eigenvektor. Wir können v in eindeutiger Weise schreiben als $v = u + w$ für u aus H_λ und w aus V, und erhalten

$$\phi(u) + \phi(w) = \phi(v) = \lambda' v = \lambda' u + \lambda' w.$$

Da per Voraussetzung $\phi(u)$ in H_λ und $\phi(w)$ in W liegt, und das genauso für $\lambda' u$ respektive $\lambda' w$ gilt, ist per Direktheit der Summe $\phi(u) = \lambda' u$ sowie $\phi(w) = \lambda' w$. Wegen $H_\lambda \cap H_{\lambda'} = \mathbf{0}$ muss $u = \mathbf{0}$ gelten, d. h. $v = w$, sodass $\mathrm{Eig}(\phi, \lambda')$ in W enthalten ist.

Schließlich folgt aus „$\mathrm{Eig}(\phi, \lambda') \subseteq W$“ zusammen mit Proposition VI.6.3 die letzte Behauptung. □

Satz 25 (Wann gibt es eine Zerlegung in Haupträume?): *Die folgenden Aussagen sind äquivalent:*

(i) *Der Vektorraum V zerfällt in die direkte Summe $V = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} H_\lambda$ der Haupträume.*

(ii) *Das charakteristische Polynom χ_ϕ zerfällt in Linearfaktoren.*

(iii) *Das Minimalpolynom m_ϕ zerfällt in Linearfaktoren.*

Beweis: „(i) $\Longrightarrow$ (ii)“: Proposition VI.6.6 sagt $\chi|_{H_\lambda} = (X - \lambda)^{\dim H_\lambda} = (X - \lambda)^{\mu_a(\phi,\lambda)}$, außerdem sind die Haupträume alle ϕ-invariant. Es muss damit

$$\chi_\phi = \prod_{\lambda \in \mathrm{Spec}(\lambda)} \chi_{\phi|_{H_\lambda}} = \prod_{\lambda \in \mathrm{Spec}(\lambda)} (X - \lambda)^{\mu_a(\phi,\lambda)}$$

gelten.

„(ii) $\Longrightarrow$ (iii)“: Weil das Minimalpolynom ein Teiler des charakteristischen Polynoms ist, greift Proposition VI.3.20.

„(iii) $\Longrightarrow$ (i)“: Folgt aus Proposition VI.4.8 und Proposition VI.6.1. □

Korollar VI.6.8 (Zerlegung über algebraisch abgeschlossenem Körper): *Ist K ein algebraisch abgeschlossener Körper (wie beispielsweise der Körper der komplexen Zahlen $\mathbb{C}$), dann ist $V = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} H_\lambda$.*

Ist eine (und damit alle) Bedingung aus Satz 25 erfüllt, und bezeichnet $\mu_a(\phi, \lambda)$ die algebraische Vielfachheit des Eigenwerts λ von ϕ, dann gilt

$$\chi_\phi = \prod_{\lambda \in \mathrm{Spec}(\phi)} (X - \lambda)^{\mu_a(\phi,\lambda)}.$$

Ferner ist $m_\phi = \prod_{\lambda \in \mathrm{Spec}(\phi)} (X - \lambda)^{\delta_\lambda}$ für Exponenten $\delta_\lambda \leq \mu_a(\phi, \lambda)$.

Bemerkung VI.6.9: (i) Seien K der Körper der reellen Zahlen $\mathbb{R}$ und ϕ der Endormorphismus von $\mathbb{R}^2$, der durch

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

erklärt wird. Dann ist $A^2 = -I_2$ und $A^2 + I_2 = \mathbf{0}$, sodass $m_\phi = X^2 + 1$, was keine reelle Nullstelle besitzt. Weil das Spektrum von ϕ leer ist, gilt für die Summe der Haupträume, dass $\bigoplus_{\lambda \in \mathrm{Spec}(\phi)} H_\lambda = \{\mathbf{0}\}$, was ein echter Unterraum des $\mathbb{R}^2$ ist.

(ii) Ist K der Körper der komplexen Zahlen $\mathbb{C}$ und betrachten wir die lineare Abbildung mit derselben Abbildungsvorschrift wie in (i), dann zerfällt das Minimalpolynom in Linearfaktoren; es ist $m_\phi = X^2 + 1 = (X - \mathrm{i})(X + \mathrm{i})$ und damit $\mathrm{Spec}(\phi) = \{\pm \mathrm{i}\}$. Bereits in einem vorherigen Beispiel ausgerechnet haben wir

$$\mathbb{C}^2 = H_{\mathrm{i}} \oplus H_{-\mathrm{i}} = \mathrm{Eig}(\phi, \mathrm{i}) \oplus \mathrm{Eig}(\phi, -\mathrm{i}) = \left\langle \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} \mathrm{i} \\ -1 \end{pmatrix} \right\rangle .$$

Bemerkung VI.6.10: Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K und ϕ ein Endomorphismus von V.

(i) Ist λ ein Eigenwert von ϕ, dann ist $\phi - \lambda \,\mathrm{id}_V$ nilpotent auf $H(\phi, \lambda)$.

(ii) Für jede Basis B von V gilt $D_{B,B}(\phi) = D_{B,B}(\phi - \lambda \,\mathrm{id}_V) + \lambda I_n$.

Beweis: (i) Der Hauptraum H_λ ist $\mathrm{Kern}(\phi - \lambda \,\mathrm{id})^e$, wobei e die algebraische Vielfachheit von $(X - \lambda)$ im Minimalpolynom m_ϕ ist. Das bedeutet, dass $(\phi - \lambda \,\mathrm{id})^e|_{H_\lambda}$ die Nullabbildung ist, also dass $(\phi - \lambda \,\mathrm{id})|_{H_\lambda}$ nilpotent ist.

(ii) Zu fixierter Basis B ist $D_{B,B} \colon \mathrm{End}(V) \to K^{n \times n}$ ein Homomorphismus von Vektorräumen, sodass

$$D_{B,B}(\phi - \lambda \,\mathrm{id}_V) = D_{B,B}(\phi) - \lambda D_{B,B}(\mathrm{id}_V) = D_{B,B}(\phi) - \lambda I_n. \qquad \square$$

Definition VI.6.11 (Jordankästchen): Seien K ein Körper, $\lambda \in K$ und d eine natürliche Zahl. Dann heißt

$$J_d(\lambda) = \lambda I_d + J_d(0) = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 1 & \lambda & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}$$

Jordankästchen der Größe d zum Eigenwert λ.

Satz 26 (über die Jordan-Normalform): *Seien K ein Körper, V ein K-Vektorraum der Dimension n und ϕ ein Endomorphismus von V mit* $\operatorname{Spec}(\phi) = \{\lambda_1, \ldots, \lambda_\ell\}$. *Zerfällt das charakteristische Polynom in Linearfaktoren, dann gibt es eine Basis B von V mit*

$$D_{B,B}(\phi) = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_\ell \end{pmatrix},$$

wobei für $1 \leq i \leq \ell$ gilt: Es gibt natürliche Zahlen $d_{1,i}, \ldots, d_{k_i,i}$ mit $d_{1,i} \geq \cdots \geq d_{k_i,i}$, sodass D_i auf der Diagonalen die Jordankästchen $J_{d_{1,i}}(\lambda_i), \ldots, J_{d_{k_i,i}}(\lambda_i)$ stehen hat.

Die Matrix D_i heißt Jordan-Block *zum Eigenwert λ_i und die Matrizen $J_{d_{j,i}}(\lambda_i)$ heißen* Jordan-Kästchen zum Eigenwert λ_i.

Das heißt k_i ist die Anzahl der Jordan-Kästchen zum Eigenwert λ und $d_{j,i}$ ist die Länge des j-ten Jordan-Kästchen zum Eigenwert λ_i. Wir nennen D_i den *Jordan-Block* zum Eigenwert λ_i.

Beweis: Nach Satz 25 haben wir die Zerlegung $V = H_{\lambda_1} \oplus \cdots \oplus H_{\lambda_r}$ von V in die ϕ-invarianten Haupträume. Ausgehend davon betrachten wir die Einschränkungen $\phi_i = \phi|_{H_{\lambda_i}}$ auf die Haupträume.

In Proposition VI.6.10 haben wir uns überlegt, dass die $\phi_i - \lambda_i \operatorname{id}_{H_{\lambda_i}}$ nilpotent sind, und dass $D_{B,B}(\phi_i) = D_{B,B}(\phi_i - \lambda_i \operatorname{id}_{H_{\lambda_i}}) + \lambda I$ für jede Basis B des Hauptraums H_{λ_i} gilt.

Mit Proposition VI.5.10 erhalten wir je eine Basis B von H_{λ_i}, sodass $D_{B,B}(\phi_i)$ die behauptete Form hat. □

Definition VI.6.12 (Geometrische Vielfachheit): Seien V ein n-dimensionaler Vektorraum über K und λ ein Eigenwert von ϕ. Dann heißt $\mu_g(\phi, \lambda) = \dim \operatorname{Eig}(\phi, \lambda)$ die *geometrische Vielfachheit des Eigenwerts λ*.

Proposition VI.6.13 (Kenngrößen): *Mit der Notation aus Satz 26 gilt für einen Eigenwert λ von ϕ:*

(i) *Die Mächtigkeit von* $\operatorname{Spec}(\phi)$ *entspricht der Anzahl der Jordan-Blöcke. Die Basisvektoren in B zum Jordan-Block D_i bilden eine Basis B_i des Hauptraums $H(\phi, \lambda_i)$ mit $D_{B_i,B_i}(\phi|_{H_{\lambda_i}}) = D_i$:*

(ii) *Die Größe des Jordan-Blocks D_i ist die Dimension des Hauptraums $H(\phi, \lambda_i)$, welche der algebraischen Vielfachheit $\mu_a(\phi, \lambda_i) = \alpha_\lambda$ entspricht.*

(iii) *Die Anzahl der Jordan-Kästchen im Jordan-Block D_i entspricht der Dimension des Eigenraums* $\operatorname{Eig}(\phi, \lambda_i)$. *Diese nennt man auch* geometrische Vielfachheit $\mu_G(\phi, \lambda_i)$ von λ_i.

(iv) *Die Größe des größten Jordan-Kästchens im Jordan-Block D_i ist der Exponent δ_{λ_i} von $(X-\lambda_i)$ im Minimalpolynom m_ϕ.*

(v) *Seien $m_d(\lambda)$ die Anzahl der Jordan-Kästchen der Länge d im Jordan-Block D_λ und für $k \in \mathbb{N}$ bezeichne $r_k = \operatorname{Rang}(\phi|_{H_\lambda} - \lambda \operatorname{id}|_{H_\lambda})^k$. Dann gilt*

$$m_d(\lambda) = r_{d-1} - 2r_d + r_{d+1}.$$

Weiter gilt für alle $k \in \mathbb{N}_0$, dass $r_{k-1} - r_k = \sum_{d=k}^{n} m_d$.

Beweis: Die Behauptungen folgen aus dem Beweis von Satz 25, Proposition VI.6.6 und Proposition VI.5.12. □

Proposition VI.6.13 enthält ein „Kochrezept“ zur Berechnung der Jordan-Normalform eines Endomorphismus.

Beispiel VI.6.14: Wir untersuchen die Jordan-Normalform

$$A = \begin{pmatrix} \boxed{\begin{matrix} 3 & & \\ 1 & 3 & \\ & 1 & 3 \end{matrix}} & & \\ & \boxed{\begin{matrix} 3 & \\ 1 & 3 \end{matrix}} & \\ & & \boxed{\begin{matrix} 2 & \\ 1 & 2 \end{matrix}} \end{pmatrix} \in \mathbb{C}^{7\times 7}.$$

Zunächst ist $\chi_A = (X-3)^5(X-2)^2$ das charakteristische Polynom von A. Das heißt $\operatorname{Spec}(A) = \{2,3\}$, und $\dim H_3 = 5 = \mu_a(A,3)$, $\dim H_2 = 2 = \mu_a(A,2)$ lesen wir für die Haupträume ab.

Nun zu den Kenngrößen zum Eigenwert $\lambda = 3$. Der zugehörige Eigenraum ist

$$\operatorname{Kern}(A-3I) = \left\langle \begin{pmatrix} 0\\0\\1\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\0\\1\\0\\0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0\\0\\1\\1\\1\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\-1\\0\\1\\0\\0 \end{pmatrix} \right\rangle = \operatorname{Kern} \begin{pmatrix} \boxed{\begin{matrix} 0 & & \\ 1 & 0 & \\ & 1 & 0 \end{matrix}} & & \\ & \boxed{\begin{matrix} 0 & \\ 1 & 0 \end{matrix}} & \\ & & \boxed{\begin{matrix} 2 & \\ 1 & 2 \end{matrix}} \end{pmatrix}$$

sodass es zwei Jordan-Kästchen zum Eigenwert 3 geben muss. Zum Hauptraum:

$$\operatorname{Kern}(A-3I)^2 = \operatorname{Kern}\left(\begin{array}{ccccccc} \boxed{0} & & & & & & \\ \boxed{0} & \boxed{0} & & & & & \\ \boxed{1} & \boxed{0} & \boxed{0} & & & & \\ & & & \boxed{0} & & & \\ & & & \boxed{0} & \boxed{0} & & \\ & & & & & \boxed{4} & \\ & & & & & \boxed{4} & \boxed{4} \end{array}\right) = \left\langle \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\\0\\0\end{pmatrix} \right\rangle$$

hat Dimension 4, d. h. wir müssen den Kern für $(A-3I)^3$ auch noch bestimmen. Dieser ist

$$\operatorname{Kern}(A-3I)^3 = \operatorname{Kern}\left(\begin{array}{ccccccc} \boxed{0} & & & & & & \\ \boxed{0} & \boxed{0} & & & & & \\ \boxed{0} & \boxed{0} & \boxed{0} & & & & \\ & & & \boxed{0} & & & \\ & & & \boxed{0} & \boxed{0} & & \\ & & & & & \boxed{8} & \\ & & & & & \boxed{12} & \boxed{8} \end{array}\right) = \left\langle \begin{pmatrix}1\\0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\\0\\0\end{pmatrix} \right\rangle$$

und hat Dimension 3, sodass $\delta_3 = 3$ der Exponent zu λ_3 im Minimalpolynom m_ϕ ist. Weil es der Jordan-Block zum Eigenwert 3 ein 5×5-Block ist, es 2 Kästchen geben muss und das größte ein 3×3-Kästchen ist, ist jetzt schon klar, dass das verbleibende Kästchen ein 2×2-Kästchen ist. Für eine größere Jordan-Normalform müsste man potentiell weitere Untersuchungen anstellen.

Benennen wir $\operatorname{Kern}(A-3I)^i$ mit K_i, dann haben wir $K_1 \subsetneq K_2 \subsetneq K_3 = H_3$ mit $\dim K_1 = 2$, $\dim K_2 = 4$ und $\dim K_3 = 5$ bestimmt. Jetzt brauchen wir noch eine Basis für den Hauptraum H_3. Dazu wählen wir ein b_1 in $K_3 - K_2$, zum Beispiel $b_1 = e_1$. Dann ist $b_2 = (A-3I)b_1 = e_2$ in $K_2 - K_1$ und $b_3 = (A-3I)b_2 = e_3$ in K_1. Weil $(A-3I)b_3 = \mathbf{0}$ gilt, haben wir so eine Basis fürs erste Jordan-Kästchen gefunden. Wegen $\dim K_3 - \dim K_1 = 1$ sind wir „fertig“ mit dem Kästchen der Länge 3.

Fürs zweite Kästchen wählen wir ein b_4 in K_2, sodass b_4 weder in $\langle b_1, b_2, b_3\rangle = \langle e_1, e_2, e_3\rangle$ noch in $K_1 = \langle e_3, e_5\rangle$ liegt. Beispielsweise $b_4 = e_4$ wäre eine mögliche Wahl. Dann ist $b_5 = (A-3I)b_4 = e_5$ in K_1, und $(A-3I)e_5 = \mathbf{0}$, d. h. $\langle e_1, \ldots, e_5\rangle$ spannt den gesamten Hauptraum auf.

Korollar VI.6.15 (Eindeutigkeit und Anwendbarkeit): *Die Matrix aus Satz 26 ist eindeutig bis auf Vertauschung der Jordan-Blöcke. Wir nennen sie* die Jordan-Normalform *von ϕ.*

Ist K algebraisch abgeschlossen, dann ist die Voraussetzung für Satz 26 erfüllt, d. h. dann hat ϕ eine Jordan-Normalform.

Beweis: Die erste Behauptung folgt aus Proposition VI.6.13, die zweite Aussage haben wir in Proposition VI.6.8 schon gezeigt. □

Korollar VI.6.16 (Jordan-Normalform für Matrizen): *Seien A eine Matrix in $K^{n\times n}$, $\mathrm{Spec}(A) = \{\lambda_1, \ldots, \lambda_r\}$ und χ_A zerfalle in Linearfaktoren. Dann gibt es eine reguläre $n \times n$-Matrix B mit Einträgen aus K, sodass $J(A) = BAB^{-1}$ eine Jordan-Normalform wie in Satz 26 ist. Diese Jordan-Normalform ist eindeutig bis auf Vertauschung der Jordan-Blöcke und es gelten die analogen Aussagen zu Proposition VI.6.13.*

Korollar VI.6.17 (Diagonalisierbarkeit):

(i) *Genau dann ist ϕ beziehungsweise A diagonalisierbar, wenn m_ϕ beziehungsweise m_A in paarweise verschiedene Linearfaktoren zerfällt.*

(ii) *Ist ϕ diagonalisierbar, und ist U ein ϕ-invarianter Unterraum, dann ist auch $\phi|_U$ diagonalisierbar.*

Beweis: (i) Genau dann ist ϕ beziehungsweise A diagonalisierbar, wenn m_ϕ beziehungsweise m_A in Linearfaktoren zerfällt (Satz 25) und die Größe des größten Jordan-Kästchens eines Eigenwerts, also der Exponent des zugehörigen Eigenwerts im Minimalpolynom, 1 ist (Proposition VI.6.13).

(ii) Es bezeichne ψ die Einschränkung von ϕ auf U. Wegen $m_\phi(\psi) = \mathbf{0}$ wird m_ϕ von m_ψ geteilt, d. h. es gibt ein Polynom h in $K[X]$, sodass $m_\phi = m_\psi h$. Weil m_ϕ nach (i) in paarweise verschiedene Linearfaktoren zerfällt, zerfällt auch m_ψ in paarweise verschiedene Linearfaktoren nach Proposition VI.3.20. Wiederum nach (i) ist ψ damit diagonalisierbar. □

Für die zweite Aussage braucht man die ϕ-invarianz von U wirklich. Am besten versuchen Sie, sich das an einem Beispiel klar zu machen.

Erinnerung VI.6.18 (Konjugiertheit): Seien A und B in $K^{n\times n}$. Gibt es eine Matrix S in $\mathrm{Gl}_n(K)$, sodass $B = SAS^{-1}$, dann heißen A und B *konjugiert* oder auch *ähnlich*. In diesem Fall schreiben wir $A \sim_{\mathrm{konj}} B$.

Ist A eine $n \times n$-Matrix mit komplexen Einträgen, dann ist A konjugiert zu einer Matrix $J(A)$ in Jordan-Normalform. Diese ist eine untere Dreiecksmatrix, die die Eigenwerte von A auf der Diagonalen und Nullen beziehungsweise Einsen auf der ersten Nebendiagonale trägt.

Die Differenz N von $J(A)$ und der Matrix D, deren Einträge die Diagonaleinträge von $J(A)$ und sonst Nullen sind, ist eine nilpotente Matrix N, d. h. $J(A) = N + D$ für eine nilpotente und eine diagonalisierbare Matrix. Die beiden „Teile“ von $J(A)$ kommutieren mit A, d. h. $DA = AD$ und $NA = AN$. Überraschenderweise ist diese Zerlegung eindeutig.

Proposition VI.6.19 (Jordan-Zerlegung): *Sei A eine $n \times n$-Matrix mit Einträgen aus einem Körper K und das charakteristische Polynom χ_A zerfalle in Linearfaktoren. Dann gibt es eindeutige $n \times n$-Matrizen D und N, sodass $A = D + N$, sodass D diagonalisierbar und N nilpotent ist, und sodass $DA = AD$, $NA = AN$. Diese Zerlegung von A heißt* Jordan-Zerlegung, Jordan-Chevalley-Zerlegung *oder auch* Dunford-Zerlegung.

Beweis: Die Existenz der Jordan-Zerlegung folgt aus Satz 26. Der Vollständigkeit halber halten wir fest, dass eine $n \times n$-Matrix N mit Einträgen aus K genau dann nilpotent ist, wenn SNS^{-1} für jede reguläre Matrix S in $K^{n\times n}$ nilpotent ist. Ist nämlich $N \in K^{n\times n}$ nilpotent, dann gibt es eine natürliche Zahl k, sodass $N^k = \mathbf{0}$ und für irgendeine reguläre $n \times n$-Matrix S mit Einträgen aus K gilt

$$(SNS^{-1})^k = (SNS^{-1})(SNS^{-1}) \cdots (SNS^{-1}) = SN^kS^{-1} = \mathbf{0}.$$

Zum Nachweis der Eindeutigkeit haben wir folgende Strategie: Die Abbildung $\phi_D \colon x \mapsto Dx$ leistet $\phi_{D|H_\lambda} = (x \mapsto \lambda x)$, und wir zeigen, dass ϕ_D dadurch vollständig bestimmt ist; also auch D und N.

Zunächst haben wir für jedes λ in K, weil D und A vertauschen, dass $(A - \lambda I_n)D = D(A - \lambda I_n)$, und so auch $(A - \lambda I_n)^k D = D(A - \lambda I_n)^k$.

Als nächstes zeigen wir die D-invarianz jedes Hauptraums H_λ, wobei λ in $\mathrm{Spec}(A)$ liegt. Seien λ aus $\mathrm{Spec}(A)$ und $H_\lambda = H(\lambda, A)$ der Hauptraum von A zum Eigenwert λ. Für jedes v aus H_λ gibt es einen Exponenten k, sodass $(A - \lambda I_n)^k v = \mathbf{0}$, d. h. $(A - \lambda I_n)^k Dv = D(A - \lambda I_n)^k v = D\mathbf{0} = \mathbf{0}$. Es ist deshalb Dv ein Element von H_λ.

Schließlich zeigen wir, dass $\phi_{D|H_\lambda} = (x \mapsto \lambda x)$. Nach Proposition VI.6.17 ist die Einschränkung $\phi_{D|H_\lambda}$ diagonalisierbar. Es genügt uns deshalb zu zeigen, dass $\mathrm{Spec}(\phi_{D|H_\lambda}) = \{\lambda\}$ ist.

Sei α ein Eigenwert von D zum von Null verschiedenen Eigenvektor v von D. Dann ist $Av = (D + N)v = Dv + Nv = \alpha v + Nv$, d. h. $(A - \alpha I_n)v = Nv$. Per

Induktion zeigt man dann, dass für jedes natürliche k auch $(A-\alpha I_n)^k v = N^k v$. Weil aber N nilpotent ist, gibt es einen Exponenten k, sodass $N^k v = \mathbf{0}$, also sodass $(A-\alpha I_n)^k v = \mathbf{0}$. Damit gehört v zum Hauptraum $H_\alpha = H(\alpha, A)$ von A zum Eigenwert α. Weil v dann im Schnitt $H_\alpha \cap H_\lambda$ liegt, und v per Voraussetzung von Null verschieden war, folgt $\alpha = \lambda$. □

Kapitel VII.

Multilineare Algebra – Teil 1

1. Multilineare Abbildungen

Definition VII.1.1 (Multilineare Abbildung): Seien K ein Körper und $V_1, \ldots, V_n$ und W Vektorräume über K. Ferner sei $M\colon V_1 \times \cdots \times V_n \to W$ eine Abbildung. Ist für jedes $i \in \{1, \ldots, n\}$ und jedes $(n-1)$-Tupel $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ aus $V_1 \times \cdots \times V_{i-1} \times V_{i+1} \times \cdots \times V_n$ die Abbildung

$$V_i \longrightarrow W, \qquad v \longmapsto M(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

linear, dann heißt M eine *n-fach multilineare Abbildung* oder kurz *multilineare Abbildung*.

Eine 1-multilineare Abbildung ist eine lineare Abbildung, eine 2-mulitilineare Abbildung heißt *bilinear* und ist $W = K$, dann spricht man von *Multilinearformen*.

Bemerkung VII.1.2: Seien K ein Körper und V_1, V_2 und W Vektorräume über K. Genau dann ist eine Abbildung $\beta\colon V_1 \times V_2 \to W$ bilinear, wenn für alle $v_1, v_1' \in V$, $v_2, v_2' \in V_2$ und $\lambda \in K$ gilt:

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

Ist nur ein Argument einer multilinearen Funktion der Nullvektor, dann ist das Bild des entsprechenden Tupels unter der multilinearen Abbildung die Null im Bild.

Beispiel VII.1.3: (i) Für einen Körper K ist

$$\det\colon K^n \times \cdots \times K^n \longrightarrow K, \qquad (v_1, \ldots, v_n) \longmapsto \det(v_1|\ldots|v_n)$$

eine n-fache Multilinearform.

(ii) Die skalare Multiplikation $K \times V \to V$, (λ, v) ist eine bilineare Abbildung.

(iii) Auf K^3 ist das *Kreuzprodukt*

$$K^3 \times K^3 \longrightarrow K^3, \qquad \left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \right) \longmapsto \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

eine bilineare Abbildung.

(iv) Für zwei K-Vektorräume V und W ist die Einsetzungabbildung

$$\mathrm{Hom}(V, W) \times V \longrightarrow W, \qquad (\phi, v) \longmapsto \phi(v)$$

eine bilineare Abbildung.

(v) Seien p, q, r und s natürliche Zahlen. Dann ist

$$K^{p \times q} \times K^{q \times r} \times K^{r \times s} \longrightarrow K^{p \times s}, \qquad (A, B, C) \longmapsto ABC$$

eine dreifach multilineare Abbildung.

Definition VII.1.4 (Vertauschungseigenschaften): Seien K ein Körper, n eine natürliche Zahl und V und W Vektorräume über K. Ferner sei $M \colon V^n \to W$ eine multilineare Abbildung.

(i) Gilt für alle $\sigma \in S_n$ und $v_1, \ldots, v_n \in V$, dass

$$M(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = M(v_1, \ldots, v_n),$$

dann heißt M *symmetrisch*.

(ii) Gilt für alle $\sigma \in S_n$ und $v_1, \ldots, v_n \in V$, dass

$$M(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \mathrm{sgn}(\sigma) M(v_1, \ldots, v_n),$$

dann heißt M *schiefsymmetrisch*.

(iii) Gilt für alle $v_1, \ldots, v_n \in V$, für die es $i \neq j$ mit $v_i = v_j$ gibt, dass $M(v_1, \ldots, v_n) = \mathbf{0}$ gilt, dann heißt M *alternierend*.

Proposition VII.1.5 (Schiefsymmetrisch vs. alternierend): *Seien K ein Körper, V und W Vektorräume über K, n eine natürliche Zahl und $M \colon V^n \to W$ eine multilineare Abbildung.*

(i) *Ist M alternierend, dann ist M schiefsymmetrisch.*

(ii) *Ist M schiefsymmetrisch und ist* $\operatorname{char}(K) \neq 2$*, dann ist M auch alternierend.*

Beweis: (i) Sei M alternierend. Aus der Linearen Algebra I ist bekannt, dass die symmetrische Gruppe S_n von Transpositionen erzeugt wird, d. h. für jedes $\sigma \in S_n$ gibt es Transpositionen $\tau_1, \ldots, \tau_k$ mit $\sigma = \tau_1 \circ \cdots \circ \tau_k$ und $\operatorname{sgn}(\sigma) = \prod_{i=1}^{k} \operatorname{sgn}(\tau_i) = (-1)^k$. Es genügt also, Schiefsymmetrie für Transpositionen nachzuweisen.

Seien also $1 \leq i < j \leq n$ und $\tau = (ij)$. Für alle $v_1, \ldots, v_n$ aus V gilt:

$$\begin{aligned}
0 &= M(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\
&= M(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) \\
&\quad + M(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\
&\quad + M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) \\
&\quad + M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\
&= M(v_1, \ldots, v_n) + M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n),
\end{aligned}$$

sodass $M(v_1, \ldots, v_n) = (-1)M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)$.

(ii) Seien M schiefsymmetrisch und $1 \leq i < j \leq n$. Für alle $v_1, \ldots, v_n$ in V ist

$$\begin{aligned}
&M(v_1, \ldots, v_{i+1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) \\
&\qquad = -M(v_1, \ldots, v_{i+1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n),
\end{aligned}$$

sodass $2M(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) = \mathbf{0}$, woraus wegen $\operatorname{char}(K) \neq 2$ die Behauptung folgt. □

Beispiel VII.1.6: (i) Die Determinante $\det\colon \prod_{i=1}^{n} K^n \to K$ ist alternierend und somit auch schiefsymmetrisch.

(ii) Das Kreuzprodukt $\times\colon K^3 \times K^3 \to K^3$ ist alternierend und schiefsymmetrisch.

(iii) Mit $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ wird der Körper mit zwei Elementen bezeichnet. Das Produkt

$$s\colon \mathbb{F}_2 \times \mathbb{F}_2 \longrightarrow \mathbb{F}_2, \qquad (a, b) \longmapsto ab$$

ist symmetrisch, da die Multiplikation auf $\mathbb{F}_2$ kommutativ ist. Außerdem ist das Produkt schiefsymmetrisch, weil $[-1] = [1]$, jedoch nicht alternierend, da $s(1, 1) = 1 \neq 0$.

2. Bilinearformen

Im Folgenden bezeichnet wir für einen K-Vektorraum V mit

$$\mathrm{Bl}(V,K) := \{\beta\colon V \times V \longrightarrow K \mid \beta \text{ Bilinearform}\}$$

den K-Vektorraum der K-Bilinearformen auf V. Es handelt sich um einen Untervektorraum von $\mathrm{Abb}(V \times V, K)$, d. h. $\mathrm{Bl}(V,K)$ wird zu einem Vektorraum mit den punktweisen Verknüpfungen.

Bemerkung VII.2.1: Seien K ein Körper, V ein Vektorraum über K und β eine Bilinearform auf V.

(i) Genau dann ist β symmetrisch, wenn für alle $v_1, v_2 \in V$ gilt, dass $\beta(v_1,v_2) = \beta(v_2,v_1)$.

(ii) Genau dann ist β schiefsymmetrisch, wenn für alle $v_1, v_2 \in V$ gilt, dass $\beta(v_1,v_2) = -\beta(v_2,v_1)$.

(iii) Genau dann ist β alternierend, wenn für alle $v \in V$ gilt, dass $\beta(v,v) = 0$.

Beispiel VII.2.2 (Einheitsform): Seien K ein Körper und n eine natürliche Zahl. Dann erklärt

$$\beta\colon K^n \times K^n \longrightarrow K, \qquad (x,y) \longmapsto x^t y = \sum_{i=1}^{n} x_i y_i = y^t x$$

eine symmetrische Bilinearform auf K^n.

Definition VII.2.3 (Skalarprodukte über den reellen Zahlen): Seien $K = \mathbb{R}$ und $\beta\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine Bilinearform.

(i) Gilt für alle $v \in V - \{\mathbf{0}\}$, dass $\beta(v,v) > 0$, dann heißt β *positiv definit*.

(ii) Ist β eine positiv definite, symmetrische Bilinearform, dann heißt β ein *Skalarprodukt*.

Beispiel VII.2.4 (für Skalarprodukte): (i) Für $K = \mathbb{R}$ ist die Einheitsform aus Proposition VII.2.2 ein Skalarprodukt auf $\mathbb{R}^n$ und heißt *Standardskalarprodukt* oder auch *euklidisches Skalarprodukt*.

(ii) Es bezeichne $V = C([0,1])$ den Vektorraum der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$. Auf V wird durch

$$\beta\colon V \times V \longrightarrow \mathbb{R}, \qquad (f,g) \longmapsto \int_0^1 f(x)g(x)\,dx$$

ein Skalarprodukt erklärt.

Bemerkung VII.2.5: Seien K ein Körper und $A = (a_{i,j}) \in K^{n\times n}$. Durch

$$\beta_A \colon K^n \times K^n \longrightarrow K, \qquad (x, y) \longmapsto x^t A y = y^t A^t x$$

wird eine Bilinearform erklärt. Es bezeichne $\{e_1, \ldots, e_n\}$ die Standardbasis des K^n. Für die Bilinearform β_A gilt $\beta_A(e_i, e_j) = e_i^t A e_j = a_{i,j}$. Insbesondere folgt für zwei $n \times n$-Matrizen A, A' mit $A \neq A'$, dass $\beta_A \neq \beta_{A'}$. Wir erhalten also einen injektiven Homomorphismus von K-Vektorräumen

$$K^{n\times n} \hookrightarrow \mathrm{Bl}(K^n, K), \qquad A \longmapsto \beta_A.$$

Genau dann ist die Bilinearform β_A symmetrisch, wenn $A = A^t$ gilt und genau dann ist β_A schiefsymmetrisch, wenn $A = -A^t$ ist. Für $A = I_n$ ist β_A genau die Einheitsform aus Proposition VII.2.2.

Proposition VII.2.6: *Seien K ein Körper und $\beta \colon K^n \times K^n \to K$ eine Bilinearform. Dann gibt es eine Matrix $A = (a_{i,j}) \in K^{n\times n}$, sodass $\beta = \beta_A$. Die Einträge der Matrix A sind bestimmt durch $a_{i,j} = \beta(e_i, e_j)$.*

Die Abbildung $K^{n\times n} \hookrightarrow \mathrm{Bl}(K^n, K)$, $A \mapsto \beta_A$ aus der vorangegangenen Bemerkung ist also sogar ein Isomorphismus.

Beweis: Seien $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$ in K^n. Dann ist

$$\beta(x, y) = \beta\Big(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\Big) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \beta(e_i, e_j) = x^t A y,$$

d. h. die Wirkung von β ist eindeutig durch die Werte $\beta(e_i, e_j)$, $1 \leq i, j \leq n$, festgelegt. Die Matrix A heißt *Gram-Matrix zu β bezüglich der Standardbasis.*□

In der Linearen Algebra I haben wir uns davon überzeugt, dass wir durch Wahl einer (geordneten) Basis $B = (b_1, \ldots, b_n)$ in einem K-Vektorraum V der Dimension n einen Isomorphismus $V \to K^n$ erhalten, via

$$v = \sum_{i=1}^n v_i b_i \longmapsto D_B(v) = \sum_{i=1}^n v_i e_i = (v_1, \ldots, v_n)^t.$$

Proposition VII.2.7 (Gram-Matrix): *Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$, $\beta \colon V \times V \to K$ eine Bilinearform und $G := (g_{i,j})_{1\leq i,j\leq n} \in K^{n\times n}$ die Matrix mit den Einträgen $g_{i,j} := \beta(b_i, b_j)$. Dann gilt für alle $v, w \in V$:*

$$\beta(v, w) = D_B(v)^t G D_B(w).$$

Die Matrix G heißt Gram-Matrix von β bezüglich B.

Die Aussage zeigt man genau so wie die Behauptung aus Proposition VII.2.6, die Eindeutigkeit folgt wie in Proposition VII.2.5.

Proposition VII.2.8: *Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$, $\beta\colon V \times V \to K$ eine Bilinearform und sei $B' = (b'_1, \ldots, b'_n)$ eine weitere geordnete Basis von V. Ferner bezeichnen G die Gram-Matrix von β bezüglich B und G' die Gram-Matrix von β bezüglich B'. Dann gilt*

$$G' = D_{B,B'}^t G D_{B,B'}$$

wobei $D_{B,B'}$ bestimmt ist durch $D_B(v) = D_{B,B'} D_{B'}(v)$.

Beweis: Für Elemente v, w von V ist

$$\begin{aligned}\beta(v,w) &= D_B(v)^t G D_B(w)\\ &= (D_{B,B'} D_{B'}(v))^t G (D_{B,B'} D_{B'}(w)) = D_{B'}(v)^t D_{B,B'}^t G D_{B,B'} D_{B'}(w),\end{aligned}$$

d. h. $G' = D_{B,B'}^t G D_{B,B'}$ wegen der Eindeutigkeit der Gram-Matrix. □

Definition VII.2.9 (Orthogonalität): Seien K ein Körper, V ein Vektorraum über K und $\beta\colon V \times V \to K$ eine symmetrische Bilinearform.

(i) Zwei Elemente v, w von V mit $\beta(v,w) = 0$ heißen *orthogonal*.

(ii) Für eine Teilmenge $M \subseteq V$ heißt

$$M^\perp := \{v \in V \mid \text{Für alle } w \in M \text{ ist } \beta(w,v) = 0\}$$

das *orthogonale Komplement von M in V*. Es handelt sich wegen der Bilinearität von β um einen Untervektorraum von V. Außerdem gilt $M \subseteq (M^\perp)^\perp$.

(iii) Seien U_1 und U_2 Untervektorräume von V. Gilt für alle $v \in U_1$ und alle $w \in U_2$, dass $\beta(v,w) = 0$, dann schreiben wir $U_1 \perp U_2$. Insbesondere gilt: $U \perp U^\perp$.

Beweis: Wir wollen zeigen, dass in Situation von (ii) tatsächlich $M \subseteq (M^\perp)^\perp$. Sei dazu $v \in M$ gegeben. Für alle $w \in M^\perp$ ist $0 = \beta(v,w) = \beta(w,v)$, sodass $v \in (M^\perp)^\perp$. □

Definition VII.2.10 (Orthogonalsystem und Orthogonalbasis): Seien K ein Körper, V ein Vektorraum über K und $\beta\colon V \times V \to K$ eine symmetrische Bilinearform. Sind $v_1, \ldots, v_k$ Elemente von V, sodass für alle $1 \leq i, j \leq k$ mit $i \neq j$ gilt $\beta(v_i, v_j) = 0$, dann heißt $(v_1, \ldots, v_k)$ ein *Orthogonalsystem bezüglich β*. Gilt

sogar $\beta(v_i, v_j) = \delta_{ij}$, dann heißt $(v_1, \ldots, v_k)$ ein *Orthonormalsystem bezüglich* β. Ist β aus dem Kontext klar, dann lassen wir den Zusatz „bezüglich β“ auch weg.

Ist $(v_1, \ldots, v_k)$ eine Basis und gleichzeitig ein Orthogonalsystem beziehungsweise ein Orthonormalsystem, dann heißt $(v_1, \ldots, v_k)$ eine *Orthogonalbasis* beziehungsweise eine *Orthonormalbasis*.

Auch für Bilinearformen $\beta\colon V \times V \to K$ die nicht symmetrisch sind, lässt sich das Konzept von Orthogonalität erklären – allerdings müssen wir dann unterscheiden zwischen Linksorthogonalität und Rechtsorthogonalität, d. h. wir erhalten Mengen ${}^{\perp}M$ und $M^{\perp}$. Darauf wollen wir aber im Rahmen dieser Vorlesung nicht weiter eingehen.

Definition VII.2.11 (Anisotrop): Seien K ein Körper, V ein Vektorraum über K und $\beta\colon V \times V \to K$ eine symmetrische Bilinearform auf V. Gibt es ein $v \in V - \{\mathbf{0}\}$ mit $\beta(v, v) = 0$, dann heißt β *isotrop*. Ist β nicht isotrop, dann heißt β *anisotrop*.

Satz 27 (Fourierformel): *Seien K ein Körper, V ein Vektorraum über K, $\beta\colon V \times V \to K$ eine symmetrische anisotrope Bilinearform und $(v_1, \ldots, v_k)$ ein Orthogonalsystem bezüglich β, $v_1, \ldots, v_k \neq \mathbf{0}$. Dann gilt:*

(i) *Ist $v \in \operatorname{Lin}(v_1, \ldots, v_k)$, d. h. gibt es $\lambda_1, \ldots, \lambda_k \in K$ mit $v = \sum_{i=1}^{k} \lambda_i v_i$, dann gilt für $1 \leq i \leq k$:*

$$\lambda_i = \frac{\beta(v, v_i)}{\beta(v_i, v_i)}.$$

(ii) *Die Vektoren $v_1, \ldots, v_k$ sind linear unabhängig.*

Beweis: Zu (i): Wir haben

$$\beta(v, v_i) = \beta\Big(\sum_{j=1}^{k} \lambda_j v_j, v_i\Big) = \sum_{j=1}^{k} \lambda_j \beta(v_j, v_i) = \lambda_i \beta(v_i, v_i),$$

und da β anisotrop ist, dürfen wir durch $\beta(v_i, v_i)$ teilen, was die Behauptung liefert. Aussage (ii) ist nun eine direkte Konsequenz. □

Satz 28 (Orthogonalisierungsverfahren nach Gram, Schmidt): *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K mit Basis $(v_1, \ldots, v_n)$ und $\beta\colon V \times V \to K$ eine symmetrische anisotrope Bilinearform. Rekursiv definieren wir Vektoren $w_1, \ldots, w_n$ durch*

$$w_1 := v_1, \qquad w_\ell := v_\ell - \sum_{i=1}^{\ell-1} \frac{\beta(w_i, v_\ell)}{\beta(w_i, w_i)} w_i.$$

Dann gilt:

(i) *Das Tupel* $(w_1, \ldots, w_n)$ *ist eine Orthogonalbasis von* V.

(ii) *Für jedes* $1 \leq \ell \leq n$ *gilt* $\mathrm{Lin}(w_1, \ldots, w_\ell) = \mathrm{Lin}(v_1, \ldots, v_\ell)$.

Beweis: Aussage (ii) folgt induktiv aus der Definition der w_ℓ, denn der Vektor $\sum_{i=1}^{\ell-1} \beta(w_i, v_\ell)/\beta(w_i, w_i) w_i$ gehört zu $\mathrm{Lin}(w_1, \ldots, w_{\ell-1}) = \mathrm{Lin}(v_1, \ldots, v_{\ell-1})$.

Zu Aussage (i): Wir zeigen per Induktion über $1 \leq \ell \leq n$ für $1 \leq i, j \leq \ell$ mit $i \neq j$, dass $\beta(w_i, w_j) = 0$. Für den Induktionsanfang ist nichts zu zeigen. Die Aussage gelte nun für $\ell - 1$. Per Induktionsvoraussetzung gilt also für $1 \leq i, j \leq \ell - 1$ und $i \neq j$, dass $\beta(w_i, w_j) = 0$. Für beliebiges $1 \leq j \leq \ell - 1$ ist

$$\begin{aligned}
\beta(w_j, w_\ell) &= \beta\Big(w_j, v_\ell - \sum_{i=1}^{\ell-1} \frac{\beta(w_i, v_\ell)}{\beta(w_i, w_i)} w_i\Big) \\
&= \beta(w_j, v_\ell) - \sum_{i=1}^{\ell-1} \frac{\beta(w_i, v_\ell)}{\beta(w_i, w_i)} \beta(w_j, w_i) \\
&= \beta(w_j, v_\ell) - \frac{\beta(w_j, v_\ell)}{\beta(w_j, w_j)} \beta(w_j, w_j) = \beta(w_j, v_\ell) - \beta(w_j, v_\ell) = 0.
\end{aligned}$$

Insbesondere liefert Satz 27 zusammen mit (ii), dass $(w_1, \ldots, w_n)$ eine Orthogonalbasis bildet. □

Definition VII.2.12 (Orthogonale Projektion): Seien K ein Körper, V ein Vektorraum über K, β eine symmetrische anisotrope Bilinearform.

(i) Sind U_1 und U_2 orthogonale Unterräume von V, dann ist $U_1 + U_2 = U_1 \oplus U_2$, d. h. die Summe ist direkt.

(ii) Ist V ein endlichdimensionaler Vektorraum, dann ist $V = U \oplus U^\perp$. Die Abbildung

$$\pi \colon V = U \oplus U^\perp \longrightarrow U, \qquad v = u + u' \longmapsto u$$

heißt *orthogonale Projektion*.

Proposition VII.2.13 (Satz des Pythagoras): *Seien K ein Körper, V ein Vektorraum über K und β eine symmetrische Bilinearform. Sind v und w Elemente von V mit $\beta(v, w) = 0$, dann ist*

$$\beta(v + w, v + w) = \beta(v, v) + \beta(w, w).$$

Beweis: Mit der Bilinearität von β rechnen wir nach:

$$\beta(v + w, v + w) = \beta(v, v) + \beta(v, w) + \beta(w, v) + \beta(w, w) = \beta(v, v) + \beta(w, w).$$

□

3. Linearformen und der Dualraum

In diesem Abschnitt möchten wir Linearformen, d. h. lineare Abbildungen $f\colon V \to K$ von einem K-Vektorraum V in den Grundkörper K, systematisch untersuchen und den Dualraum $V^* := \mathrm{Hom}(V, K)$ einführen. Dabei erhalten wir folgende zentrale Ergebnisse: Erstens ist V ein endlichdimensionaler Vektorraum, dann ist V isomorph zu seinem Dualraum. Jedoch gibt es nicht einen eindeutigen Isomorphismus $V \to V^*$, für jede Basis von V oder jede „geeignete“ Bilinearform erhalten wir einen solchen. Zweitens gibt es einen natürlichen Morphismus von V nach V^{**}, den sogenannten *Bidualraum von V*.

Definition VII.3.1 (Linearformen, Dualraum): Seien K ein Körper und V ein endlichdimensionaler K-Vektorraum.

(i) Eine lineare Abbildung $f\colon V \to K$ heißt *Linearform* oder *lineares Funktional*.

(ii) Die Menge $V^* := \mathrm{Hom}_K(V, K) = \mathrm{Hom}(V, K) = \{f\colon V \to K \text{ linear}\}$ heißt *Dualraum von V*.

Als Untervektorraum von $\mathrm{Abb}(V, K)$ ist $\mathrm{Hom}(V, K)$ ein K-Vektorraum (natürlich mit den punktweisen Verknüpfungen).

Beispiel VII.3.2 (für Linearformen): (i) Sei V der K-Vektorraum K^3. Dann ist

$$f\colon V \longrightarrow K, \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 + 2x_2 + 3x_3$$

eine Linearform auf V, wie aus der Linearen Algebra I bekannt. Allgemein gilt $(K^n)^* = K^{1\times n}$.

(ii) Sei V der Vektorraum der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$. Dann ist

$$I\colon C([0,1]) \longrightarrow \mathbb{R}, \qquad h \longmapsto \int_0^1 h(x)\,dx$$

eine Linearform. Für ein festes $x_0 \in [0,1]$ ist außerdem auch

$$f_{x_0} := \left.\frac{d}{dx}\right|_{x=x_0} \colon C([0,1]) \longrightarrow \mathbb{R}, \qquad h \longmapsto h'(x_0)$$

eine Linearform auf V. Genauso ist Auswertung in x_0, d. h. $A_{x_0}\colon h \mapsto h(x_0)$ eine Linearform auf V.

(iii) Ist V ein K-Vektorraum, ist $\beta\colon V \times V \to K$ eine Bilinearform und ist $w \in V$ ein Element, dann ist $L_w\colon V \to K$, $v \mapsto \beta(v, w)$ eine Linearform.

(iv) Auf $V = P_2 = \{p \in \mathbb{R}[X] \mid \deg(p) \leq 2\}$ erklärt $A_3\colon P_2 \to \mathbb{R}$, $p \mapsto p(3)$ eine Linearform. Wir können A_3 auch schreiben als $a_2X^2 + a_1X + a_0 \mapsto 9a_2 + 3a_1 + a_0$.

Bemerkung VII.3.3 (Beschreibung des Dualraums V^*): Seien K ein Körper und V ein K-Vektorraum mit Basis B. Nach dem Fortsetzungssatz aus der Linearen Algebra I ist die Abbildung

$$R_B\colon V^* = \mathrm{Hom}(V, K) \longrightarrow \mathrm{Abb}(B, K), \qquad f \longmapsto f|_B$$

ein Isomorphismus. Insbesondere gilt $V^* \cong \mathrm{Abb}(B, K)$. Außerdem haben wir in der Linearen Algebra I die Koordinatenabbildung

$$D_B\colon V \longrightarrow \mathrm{Abb}_0(B, K) = \{h\colon B \to K \mid \#\{b \in B \mid h(b) \neq 0\} < \infty\}$$

als Isomorphismus von K-Vektorräumen kennengelernt. Insgesamt erhalten wir so einen injektiven Homomorphismus

$$\Theta_b := R_B^{-1} \circ D_B\colon V \longrightarrow V^*, \qquad v = \sum_{b \in B} \lambda(b)b \mapsto f$$

wobei f die eindeutige Linearform mit $f(b) = \lambda(b)$ ist.

Bemerkung VII.3.4: Sei nun V ein endlichdimensionaler K-Vektorraum mit Basis $B = (b_1, \ldots, b_n)$. In diesem Fall gibt es keinen Unterschied zwischen $\mathrm{Abb}(B, K)$ und $\mathrm{Abb}_0(B, K)$, und zu den Abbildungen aus der vorherigen Bemerkung erhalten wir folgendes Diagramm:

$$\begin{array}{ccccc}
V & \cong & \mathrm{Abb}_0(B, K) = \mathrm{Abb}(B, K) & \cong & V^* \\
\sum_{i=1}^n f(b_i)b_i & \longmapsto & f|_B & \longmapsto & f \\
v = \sum_{i=1}^n \lambda_i b_i & \longmapsto & (f\colon B \to K, \quad b_i \mapsto \lambda_i) & \longmapsto & (F\colon V \to K)
\end{array}$$

wobei wir in der letzten Zeile mit $F\colon V \to K$ die eindeutige lineare Fortsetzung von f meinen. Das bedeutet, wir erhalten einen Isomorphismus $\Theta_B\colon V \to V^*$, der $v = \sum_{i=1}^n \lambda_i b_i$ auf die eindeutig bestimmte Linearform f_v mit der Eigenschaft $f_v(b_i) = \lambda_i$ abbildet. Insbesondere ist in diesem Fall V isomorph zu V^*.

Definition VII.3.5 (Duale Basis): Seien V ein endlichdimensionaler K-Vektorraum mit Basis $B = (b_1, \ldots, b_n)$ und $\Theta_B \colon V \to V^*$ der Isomorphismus aus Proposition VII.3.4. Für $1 \leq i \leq n$ schreiben wir $b_i^* = \Theta_B(b_i)$, d. h. b_i^* ist die durch $b_i^*(b_j) = \delta_{i,j}$ eindeutig bestimmte Linearform.

Das Tupel $B^* = (b_1^*, \ldots, b_n^*)$ ist dann eine geordnete Basis von V^*, die *zu B duale Basis B^**.

Auswertung von b_i^* auf einem Vektor v von V liefert den Koeffizenten von b_i in der Linearkombination von v bezüglich der Basis B, d. h. ist $v = \sum_{j=1}^n \lambda_j b_j$, dann ist $b_i^*(v) = \lambda_i$.

Ist $f \colon V \to K$ linear, dann gilt für die Koordinaten von f bezüglich B^*:

$$D_{B^*}(f) = \Big(f(b_1), \ldots, f(b_n)\Big)^t.$$

Beweis: Die erste Behauptung zeigt man durch direktes Nachprüfen der Definitionen. Die zweite Behauptung folgt, da Θ_B ein Isomorphismus ist. Zur dritten Behauptung: Da B eine Basis von V ist, können wir v in eindeutiger Weise schreiben als $v = \sum_{j=1}^n \lambda_j b_j$. Für ein i aus $\{1, \ldots, n\}$ erhalten wir

$$b_i^*(v) = b_i^*\Big(\sum_{j=1}^n \lambda_j b_j\Big) = \sum_{j=1}^n \lambda_j b_i^*(b_j) = \sum_{j=1}^n \lambda_j \delta_{i,j} = \lambda_i.$$

Zur vierten Behauptung: Wir müssen zeigen, dass $f = \sum_{j=1}^n f(b_i) b_i^*$ gilt. Dazu prüfen wir, dass beide Abbildungen punktweise dasselbe tun. Wie wir uns gerade überlegt haben, ist $v = \sum_{i=1}^n b_i^*(v) b_i$ und

$$f(v) = f\Big(\sum_{i=1}^n b_i^*(v) b_i\Big) = \sum_{i=1}^n b_i^*(v) f(b_i) = \sum_{i=1}^n f(b_i) b_i^*(v) = \Big(\sum_{i=1}^n f(b_i) b_i^*\Big)(v).$$

□

Beispiel VII.3.6: Es sei $V = P_2 = \{p \in \mathbb{R}[X] \mid \deg(p) \leq 2\}$ zusammen mit der Standardbasis $B = (b_0, b_1, b_2) = (1, X, X^2)$. Für ein $p = a_0 + a_1 X + a_2 X^2$ aus V ist $b_i^*(p) = a_i$, was uns die b_i^* als Abbildungen beschreibt.

Wollen wir das Funktional $A_2 \colon P_2 \to \mathbb{R}$, $p \mapsto p(2)$ in der dualen Basis ausdrücken, dann müssen wir nach der vorangegangenen Bemerkung die Werte von A_2 auf b_0, b_1, b_2 bestimmen. Es sind $A_2(b_0) = 1$, $A_2(b_1) = 2$, $A_2(b_2) = 4$, d. h. $D_{B^*}(A_2) = (1, 2, 4)^t$, also $A_2 = b_0^* + 2b_1^* + 4b_2^*$.

Definition VII.3.7 (Duale Abbildung): Seien K ein Körper, V_1 und V_2 zwei K-Vektorräume und $\phi \colon V_1 \to V_2$ eine lineare Abbildung. Dann ist

$$\phi^* \colon V_2^* \longrightarrow V_1^*, \qquad f \longmapsto f \circ \phi$$

eine lineare Abbildung und heißt die *zu ϕ duale Abbildung* oder einfach *duale Abbildung zu ϕ*.

Beweis: Seien f und g beliebige Elemente von V_2^* und λ irgendein Element von K. Wir wollen überprüfen, ob ϕ^* linear ist, d. h. ob $\phi^*(f+g) = \phi^*(f) + \phi^*(g)$ und $\phi^*(\lambda f) = \lambda\phi^*(f)$. Diese Gleichheit von Abbildungen von V_1 nach K rechnen wir auf einem beliebigen Element von V_1 nach. Für so ein v ist

$$\begin{aligned}\phi^*(f+g)(v) &= ((f+g)\circ\phi)(v)\\ &= (f+g)(\phi(v))\\ &= f(\phi(v)) + g(\phi(v)) + \phi^*(f)(v) + \phi^*(g)(v) = (\phi^*(f) + \phi^*(g))(v),\end{aligned}$$

sowie

$$\phi^*(\lambda f)(v) = ((\lambda f)\circ\phi)(v) = \lambda f(\phi(v)) = \lambda\phi^*(f)(v). \qquad \square$$

Beispiel VII.3.8: Seien $V = P_3 = \{p \in \mathbb{R}[X] \mid \deg(p) \leq 3\}$, $B = (1, X, X^2, X^3)$ und $\phi\colon V \to V$ die Ableitung, d. h.

$$\phi(a_3X^3 + a_2X^2 + a_1X + a_0) \longmapsto 3a_3X^2 + 2a_2X + a_1.$$

Die Darstellungsmatrix von ϕ bezüglich B können wir einfach ausrechnen und erhalten

$$D_{B,B}(\phi) = \begin{pmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

Wie können wir jetzt die duale Abbildung $\phi^*\colon V^* \to V^*$ beschreiben? Die duale Abbildung ϕ^* schickt ein Funktional f auf die Präkomposition mit ϕ, d. h. die Abbildung, die ein Polynom p auf $f(\phi(p)) = f(p')$ schickt. Für die duale Basis $b_0^*, \ldots, b_3^*$ erhalten wir

$$\begin{aligned}\phi^*(b_0^*) &= (p \mapsto b_0^*(p') = a_1) = b_1^*\\ \phi^*(b_1^*) &= (p \mapsto b_1^*(p')) = 2a_2) = 2b_2^*\\ \phi^*(b_2^*) &= (p \mapsto b_2^*(p') = 3a_3) = 3b_3^*\\ \phi^*(b_3^*) &= (p \mapsto b_3^*(p') = 0) = 0\end{aligned}$$

d. h. die Abbildungsmatrix von ϕ^* bezüglich B^* ist gegeben durch

$$D_{B^*,B^*}(\phi^*) = \begin{pmatrix} 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 3 & 0\end{pmatrix}.$$

Proposition VII.3.9 (Abbildungsmatrix für duale Abbildung): *Seien V_1 und V_2 endlichdimensionale Vektorräume über K mit Basen $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$, $\phi\colon V_1 \to V_2$ eine lineare Abbildung und $A := D_{C,B}(\phi)$. Dann ist die Darstellungsmatrix $D_{B^*,C^*}(\phi^*)$ der dualen Abbildung $\phi^*\colon V_2^* \to V_1^*$ gegeben durch $D_{B^*,C^*}(\phi^*) = A^t$.*

Beweis: Für die Darstellungsmatrizen von ϕ und ϕ^* haben wir die Diagramme

$$\begin{array}{ccc} V_1 & \xrightarrow{\phi} & V_2 \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} \\ K^n & \xrightarrow[D_{C,B}(\phi)]{} & K^m \end{array} \qquad \begin{array}{ccc} V_2^* & \xrightarrow{\phi^*} & V_1^* \\ {\scriptstyle D_{C^*}}\downarrow & & \downarrow{\scriptstyle D_{B^*}} \\ K^m & \xrightarrow[D_{B^*,C^*}(\phi)]{} & K^n \end{array}$$

mit Matrizen $D_{C,B}(\phi) = (x_{i,j})$ in $K^{m\times n}$ und $D_{B^*,C^*}(\phi^*) = (y_{i,j})$ in $K^{n\times m}$. Die bestimmenden Gleichungen für die Matrixeinträge sind $\phi(b_i) = \sum_{j=1}^m x_{i,j}c_j$ sowie $\phi^*(c_i^*) = \sum_{j=1}^n y_{i,j}b_j^*$. Einerseits ist $\phi^*(c_i^*)(b_j) = \sum_{k=1}^n y_{i,k}b_k^*(b_j) = y_{i,j}$, und andererseits ist $\phi^*(c_i^*)(b_j) = (c_i^* \circ \phi)(b_j) = c_i^*(\phi(b_j)) = c_i^*(\sum_{k=1}^m x_{j,k}c_k) = x_{j,i}$. Das zeigt die Behauptung über die Abbildungsmatrizen. □

Definition VII.3.10 (Einsetzungshomomorphismus): Seien K ein Körper, V ein K-Vektorraum und $v \in V$ gegeben. Die Abbildung

$$A_v\colon V^* \longrightarrow K, \qquad f \longmapsto f(v)$$

nennen wir *Einsetzungshomomorphismus*. Der Einsetzungshomomorphismus A_v ist eine Linearform, d. h. A_v gehört zu $(V^*)^*$.

Proposition VII.3.11 (Einbettung in Bidualraum): *Seien K ein Körper und V ein K-Vektorraum. Setze*

$$\psi\colon V \longrightarrow V^{**}, \qquad v \longmapsto A_v,$$

wobei A_v den Einsetzungshomomorphismus aus Proposition VII.3.10 bezeichnet. Dann gilt:

(i) *Die Abbildung ψ ist linear.*

(ii) *Die Abbildung ist injektiv.*

(iii) *Ist V endlichdimensional, dann ist ψ ein Isomorphismus.*

Vermöge ψ kann V als Untervektorraum von V^{**} aufgefasst werden, und im endlichdimensionalen Fall wird V sogar mit V^{**} identifiziert. Es gibt moralisch gesprochen eine intrinsische Verbindung von V und V^{**}. Das ist eine erheblich bessere Situation als für den Dualraum, der im Allgemeinen „nur“ über Basen mit dem ursprünglichen Raum zusammenhängt.

Beweis: (i) Seien v_1 und v_2 in V und $\lambda \in K$. Dann gilt

$$\begin{aligned}\psi(v_1 + v_2) &= A_{v_1+v_2} \\ &= (f \mapsto f(v_1 + v_2) = f(v_1) + f(v_2)) = A_{v_1} + A_{v_2} = \psi(v_1) + \psi(v_2).\end{aligned}$$

Analog rechnet man nach, dass $\psi(\lambda v) = \lambda\psi(v)$.

(ii) Sei $v \in \mathrm{Kern}(\psi)$, d. h. $A_v = \mathbf{0}$. Dann gilt für alle $f \in V^*$, dass $f(v) = \mathbf{0}$. Es gibt eine Basis B von V. (Wir werden am Ende der Vorlesung zeigen, dass das auch für unendlichdimensionale Vektorräume gilt; dazu benötigt man das sogenannte Lemma von Zorn). Für $b \in B$, definieren wir $b^* \colon V \to K$ durch lineare Fortsetzung von $b \mapsto 1$ und $b' \mapsto 0$ für $b' \in B - \{b\}$. Wir können v als Linearkombination von B schreiben, etwa $v = \sum_{b\in B} \lambda(b)b$. Für jedes $b \in B$ ist einerseits $b^*(v)\lambda(b)$, und andererseits, da v nach Voraussetzung im Kern von ψ liegt, auch $\lambda(b) = 0$. Es gilt also $\lambda(b) = 0$ für jedes b aus B und v muss der Nullvektor sein.

(iii) Ist V endlichdimensional, dann gilt $\dim V = \dim V^*$. Außerdem ist per Definition $V^{**} = (V^*)^*$, sodass $\dim V = \dim V^{**}$ folgt. Damit ist ψ auch surjektiv, also ein Isomorphismus. □

Ist V mit einer anisotropen Bilinearform ausgestattet, dann vermittelt uns das eine intrinsische Verbindung von V und V^*; wir können in dem Fall V auf kanonische Weise, d. h. ohne Bezugnahme auf Basen, in seinen Dualraum einbetten. Das skizzieren wir im Folgenden.

Proposition VII.3.12 (Kanonische Einbettung in Dualraum via Bilinearform): *Seien K ein Körper, V ein Vektorraum über K und $\beta \colon V \times V \to K$ eine anisotrope Bilinearform. Für ein Element w von V ist $L_w \colon V \to K$, $v \mapsto \beta(v, w)$ eine Linearform. Für*

$$\Theta_\beta \colon V \longrightarrow V^*, \qquad w \longmapsto L_w$$

gilt:

(i) *Die Abbildung Θ_β ist linear.*

(ii) *Die Abbildung Θ_β ist injektiv.*

(iii) *Ist V endlichdimensional, dann ist Θ_β ein Isomorphismus.*

Beweis: (i) Das ist klar wegen der Bilinearität von β.

(ii) Ist $w \in V$ mit $L_w = \mathbf{0}$, dann gilt für alle $v \in V$, dass $L_w(v) = \mathbf{0}$, d. h. für alle $v \in V$ ist $\beta(v, w) = 0$. Insbesondere ist dann auch $\beta(w, w) = 0$, sodass die Anisotropie $w = \mathbf{0}$ erzwingt.

(iii) Das folgt aus $\dim V^* = \dim V$. □

Bei Anwendungen von Proposition VII.3.12 ist β häufig ein Skalarprodukt über $\mathbb{R}$. Die Aussage Proposition VII.3.12 ist ein Spezialfall des Satzes von Riesz.

Einen Vektorraum V über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ zusammen mit einem Skalarprodukt heißt Prä-Hilbertraum. Auf V wird durch

$$d(v, w) := \langle v - w, v - w \rangle^{1/2}$$

eine Metrik erklärt. Ist (V, d) ein vollständiger metrischer Raum, dann heißt V *Hilbertraum.*

Satz (von Riesz): *Seien V ein Hilbertraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $\Theta_{\langle \cdot, \cdot \rangle}$ die Abbildung aus VII.3.12. Dann ist*

$$\Theta_{\langle \cdot, \cdot \rangle} \colon V \hookrightarrow (V^*)^{\mathrm{top}}$$

ein Isomorphismus. Hierbei bezeichnet $(V^*)^{\mathrm{top}} = \{f \in V^* \mid f \text{ ist stetig}\}$ den topologischen Dualraum von V.

Die Aussage des Satzes von Riesz gilt auch für unendlichdimensionale Hilberträume; allerdings nicht für Prä-Hilberträume. Der Beweis braucht den im endlichdimensionalen Fall natürlichen Projektionssatz $V = U \oplus U^\perp$, wobei U ein im analytischen Sinne abgeschlossener Untervektorraum von V ist, der sich auf Hilberträume verallgemeinert.

Satz 29 (Zusammenfassung zum Dualraum): *Seien K ein Körper und V ein K-Vektorraum.*

(i) *Jede Basis B von V definiert durch $\Theta_B \colon V \hookrightarrow V^*$, $v = \sum_{b \in B} \lambda(b) b \mapsto \sum_{\lambda \in B} \lambda(b) b^*$ eine Identifikation von V mit einem Untervektorraum des Dualraums V^*. Ist V endlichdimensional, dann ist Θ_B ein Isomorphismus.*

(ii) *Jede anisotrope Bilinearform $\beta \colon V \times V \to K$ definiert durch $\Theta_\beta \colon V \hookrightarrow V^*$, $w \mapsto L_w = \beta(\cdot, w)$ eine Identifikation von V mit einem Untervektorraum von V^*. Ist V endlichdimensional, dann ist Θ_β ein Isomorphismus.*

(iii) *Durch $V \hookrightarrow V^{**}$, $v \mapsto A_v = (f \mapsto f(v))$ erhalten wir eine kanonische Einbettung von V in V^{**}. Ist V endlichdimensional, dann handelt es sich um einen Isomorphismus.*

Beweis: (i) Bemerkungen II.3.3 und II.3.4.

(ii) Proposition II.3.12.

(iii) Proposition II.3.11. □

Der Dualraum verdankt seinen Namen dem lateinischen Wort „dual“, welches „zweifach“, oder „zwei betreffend“ heißt. Hier ist dual am besten zu verstehen als „gegenläufig“, oder „Pfeile drehen sich um“ – gemeint sind Abbildungspfeile.

Zum Beispiel in der Analysis oder der Geometrie haben Dualräume vielfältige Anwendungen. Ein prominentes Beispiel sind Tangentialräume und Kotangentialräume. Ist beispielsweise X eine Fläche im $\mathbb{R}^3$ und p ein Punkt der Fläche X, dann heften wir einen $\mathbb{R}$-Vektorraum T_pX, den Tangentialraum an X in p, an diese Fläche an. Diesen Vektorraum können wir uns vorstellen als den Graph der linearen Abbildung, die die Fläche X in p „am besten approximiert“.[1] Auf natürliche Weise ist dieser Tangentialraum in den umgebenden $\mathbb{R}^3$ eingebettet und erbt so das Standardskalarprodukt.

Ist $f\colon X \to \mathbb{R}$ eine stetig differenzierbare Funktion, und gehört v zu T_pX, dann heißt $D_vf(p)$ die Richtungsableitung in Richtung p. Wir erhalten eine Abbildung

$$Df(p)\colon T_pX \longrightarrow \mathbb{R}, \qquad v \longmapsto D_vf(p).$$

Diese Abbildung $Df(p)$ gehört zum Dualraum von T_pX, dem sogenannten Kotangentialraum, der üblicherweise mit T_p^*X bezeichnet wird. Nach dem Satz von Riesz gibt es genau einen Vektor $w \in T_pX$, sodass für alle $v \in T_pX$ gilt: $D_vf(p) = \langle v, w\rangle$. Dieser spezielle Vektor heißt Gradient von f in p.

4. Tensorprodukte

Seien V_1, V_2 und W Vektorräume über dem selben Körper K. In diesem Abschnitt möchten wir „die Mutter aller bilinearen Abbildungen $V_1 \times V_2 \to W$“ kennen lernen. Die führt zum Begriff des Tensorprodukts $V_1 \otimes_K V_2$.

Dieses Konzept hat viele Anwendungen in der Analysis, der Geometrie und wird insbesondere in vielen Ingenieurswissenschaften häufig gebraucht.

Für einen K-Vektorraum V und einen Untervektorraum U haben wir in der Linearen Algebra I den besonderen Vektorraum V/U kennengelernt. Dieser hieß „Quotientenvektorraum“ oder auch „Faktorraum“ und war definiert als die Menge der Äquivalenzklassen $V/U = \{[v]_\sim \mid v \in V\}$ bezüglich der auf V erklärten Äquivalenzrelation

$$v \sim w :\Longrightarrow v - w \in U,$$

[1] Ist X beispielsweise der Graph einer differenzierbaren Funktion, dann ist der Tangentialraum in einem Punkt der Graph der totalen Ableitung dieser differenzierbaren Funktion an einer geeigneten Stelle.

welchen wir per $[v]+[w] := [v+w]$ und $\lambda[v] := [\lambda v]$ eine K-Vektorraumstruktur aufgeprägt haben. Für ein $v \in V$ ist die Äquivalenzklasse $[v]$ genau die Menge $v+U = \{v+u \mid u \in U\}$. Die Restklassenabbildung $\pi\colon V \to V/U$, $v \mapsto [v]$ heißt *kanonische Projektion* und ist per Konstruktion der K-Vektorraumstruktur von V/U trivialerweise eine lineare Abbildung.

Außerdem haben wir den Homomorphiesatz in der Linearen Algebra I kennengelernt: Für K-Vektorräume V und W, eine lineare Abbildung $\phi\colon V \to W$ und einen Untervektorraum $U \subseteq V$ mit $U \subseteq \mathrm{Kern}(\phi)$ haben wir das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U \\ & \searrow^{\phi} & \downarrow{\scriptstyle \bar{\phi}} \\ & & W \end{array}$$

Das heißt es gibt genau eine lineare Abbildung $\bar{\phi}\colon V/U \to W$ mit $\bar{\phi} \circ \pi = \phi$.

Bemerkung VII.4.1: Seien K ein Körper, n und m natürliche Zahlen, $V_1 := K^n$, $V_2 := K^m$ und $T := K^{n\times m}$. Ferner sei

$$\tau\colon V_1 \times V_2 \longrightarrow T, \qquad (x, y) \longmapsto xy^t.$$

Dann gilt:

(i) Die Abbildung τ is bilinear.

(ii) Für die Standardbasen $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_m)$ die Standardbasen von K^n beziehungsweise K^m gilt $\tau(e_i, e'_j) = E_{i,j}$, wobei $E_{i,j}$ die Elementarmatrix in $K^{n\times m}$ bezeichnet, die durch $E_{i,j} = (\delta_{i,k}\delta_{j,\ell})_{1\le k,\ell\le n}$ definiert ist.

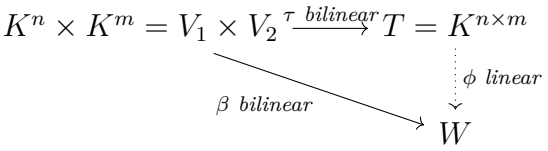
Beispiel VII.4.2: Seien $V_1 = \mathbb{R}^2$, $V_2 = \mathbb{R}^3$ und $T = \mathbb{R}^{2\times 3}$. Für die Abbildung τ aus Proposition VII.4.1 haben wir beispielsweise

$$\tau\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 0 & -2 \end{pmatrix}$$

$$\tau\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Proposition VII.4.3 (Universelle Abbildungseigenschaft): *Die Abbildung τ aus Proposition VII.4.1 hat folgende Eigenschaft: Für jeden K-Vektorraum W und*

jede bilineare Abbildung $\beta\colon V_1 \times V_2 \to W$ gibt es genau eine lineare Abbildung $\phi\colon T \to W$ mit $\phi \circ \tau = \beta$, d. h. wir haben das kommutative Diagramm



Beweis: Wieder bezeichne $(e_1, \ldots, e_n)$ beziehungsweise $(e'_1, \ldots, e'_m)$ die Standardbasis von K^n beziehungsweise K^m. Die Menge der Elementarmatrizen $\{E_{i,j} \mid 1 \le i \le n, 1 \le j \le n\}$ bildet eine Basis von $T = K^{n\times m}$.

Nach dem Fortsetzungssatz für lineare Abbildungen gibt es genau eine lineare Abbildung $\phi\colon T \to W$ mit $\phi(E_{i,j}) = \beta(e_i, e'_j)$ und für dieses ϕ gilt tatsächlich $\phi \circ \tau = \beta$. Seien dazu $v = \sum_{i=1}^n \lambda_i e_i$ in V_1 und $w = \sum_{j=1}^m \lambda'_j e'_j$ in V_2. Dann ist

$$\begin{aligned}
\phi(\tau(v,w)) &= \phi\Big(\tau\Big(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^m \lambda'_j e'_j\Big)\Big) \\
&= \phi\Big(\sum_{i=1}^n \sum_{j=1}^m \lambda_i \lambda_j \tau(e_i, e'_j)\Big) \\
&= \sum_{i=1}^n \sum_{j=1}^m \lambda_i \lambda'_j \phi(\tau(e_i, e'_j)) \\
&= \sum_{i=1}^n \sum_{j=1}^m \lambda_i \lambda'_j \beta(e_i, e_{j'}) = \beta\Big(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^m \lambda'_j e'_j\Big) = \beta(v,w).
\end{aligned}$$

Außerdem ist ϕ eindeutig, da insbesondere $\phi \circ \tau(e_i, e'_j) = \beta(e_i, e'_j)$ gelten muss.□

Beispiel VII.4.4: Seien $V_1 = V_2 = K^n$ und $\tau\colon K^n \times K^n \to K^{n\times n}$, $(x, y) \mapsto xy^t$. Ferner sei $\beta\colon K^n \times K^n \to K$ die Einheitsform, d. h. $\beta(x, y) = x^t y$. Was ist die Abbildung $\phi\colon K^{n\times n} \to K$ aus der vorangegangenen Proposition? Auf den Elementarmatrizen $E_{i,j}$ muss ϕ folgendes machen:

$$\phi(E_{i,j}) = \beta(e_i, e_j) = e_i^t e_j = \delta_{i,j}.$$

Für eine Matrix $A = (a_{i,j}) \in K^{n\times n}$ ist also

$$\phi(A) = \phi\Big(\sum_{i=1}^n \sum_{j=1}^n a_{i,j} E_{i,j}\Big) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} \delta_{i,j} = \sum_{i=1}^n a_{i,i},$$

d. h. $\phi(A)$ gibt die Summe der Diagonaleinträge von A zurück.

Definition VII.4.5: Seien K ein Körper, n eine natürliche Zahl und $A = (a_{i,j})$ in $K^{n\times n}$ eine Matrix. Dann heißt

$$\operatorname{tr}(A) := \sum_{i=1}^{n} a_{i,i} \in K$$

die *Spur von* A. Die Abbildung $\operatorname{tr}\colon K^{n\times n} \to K$, $A \mapsto \operatorname{tr}(A)$ ist eine lineare Abbildung, d. h. insbesondere $\operatorname{tr} \in (K^{n\times n})^*$.

Das Tupel (T, τ) aus Proposition VII.4.3 nennt man auch ein *Tensorprodukt von* V_1 *und* V_2. Dadurch, dass wir uns bereits für beliebige natürliche Zahlen n und m davon überzeugt haben, dass K^n und K^m ein Tensorprodukt haben, haben wir allgemeiner für endlichdimensionale Vektorräume die Existenz von Tensorprodukten etabliert. Es ist allerdings ein natürliches Bedürfnis, solche Tensorprodukte auch für unendlichdimensionale Vektorräume zu haben, wie zum Beispiel für Polynomringe. Damit wollen wir uns im Folgenden beschäftigen.

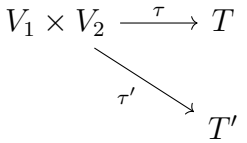
Definition VII.4.6 (Tensorprodukt): Seien K ein Körper, V_1, V_2 und T Vektorräume über K und $\tau\colon V_1 \times V_2 \to T$ eine bilineare Abbildung. Gibt es für jeden K-Vektorraum W und jede bilineare Abbildung $\beta\colon V_1 \times V_2 \to W$ genau eine lineare Abbildung $\phi\colon T \to W$ mit $\phi \circ \tau = \beta$, dann heißt das Tupel (T, τ) ein *Tensorprodukt von* V_1 *und* V_2 *über* K.

Notation VII.4.7: In der Situation von Proposition VII.4.6 schreiben wir für das Tensorprodukt $V_1 \otimes_K V_2 := V_1 \otimes V_2 := T$; für $v_1 \in V_1$, $v_2 \in V_2$ schreiben wir $v_1 \otimes v_2 := \tau(v_1, v_2)$ und nennen $v_1 \otimes v_2$ einen reinen Tensor.

Wir werden zeigen: Je zwei K-Vektorräume V und W haben ein Tensorprodukt, und Tensorprodukte sind eindeutig bis auf Isomorphie.

Proposition VII.4.8 (Es kann nur einen geben): *Seien K ein Körper, V_1 und V_2 Vektorräume über K. Ein Tensorprodukt (T, τ) von V_1 und V_2 ist eindeutig bis auf eindeutige Isomorphie, d. h. sind (T, τ) und (T', τ') Tensorprodukte von V_1 und V_2 über K, dann gibt es einen eindeutigen Isomorphismus $\phi\colon T \to T'$ mit $\tau' = \phi \circ \tau$.*

Beweis: Wir befinden uns in der Situation



und sowohl $\tau\colon V_1 \times V_2 \to T$ als auch $\tau'\colon V_1 \times V_2 \to T'$ sind bilinear. Über die universelle Abbildungseigenschaft von (T, τ) erhalten wir also eine eindeutige lineare Abbildung $\phi\colon T \to T'$ mit $\phi \circ \tau = \tau'$ und wegen der universellen Abbildungseigenschaft von (T', τ') gibt es eine eindeutige lineare Abbildung $\psi\colon T' \to T$ mit $\psi \circ \tau' = \tau$.

Wir haben also $(\psi \circ \phi) \circ \tau = \psi \circ \tau' = \tau$ und außerdem ist $\mathrm{id} \circ \tau = \tau$. Wegen der Eindeutigkeitsaussage in der universellen Abbildungseigenschaft für (T, τ) erhalten wir $\psi \circ \phi = \mathrm{id}$. Analog erhalten wir, dass $\phi \circ \psi = \mathrm{id}$. Damit ist ϕ ein eindeutiger Isomorphismus mit Inversem ψ. □

Erinnerung VII.4.9 (Abbildungen mit endlichem Träger): Es seien M eine Menge und K ein Körper. Für eine Abbildung $f\colon M \to K$ heißt die Menge $\mathrm{Tr}(f) := \{m \in M \mid f(m) \neq 0\}$ der *Träger von* f. Wir schreiben

$$\mathrm{Abb}_0(M, K) := \{f\colon M \to K \mid \#\,\mathrm{Tr}(f) < \infty\} \subseteq \mathrm{Abb}(M, K)$$

für den Untervektorraum der Abbildungen mit endlichem Träger.

Für $m \in M$ bezeichne $f_m\colon M \to K$ die Abbildung mit

$$f_m(m') = \begin{cases} 1, & \text{falls } m = m', \\ 0, & \text{sonst.} \end{cases}$$

Dann bildet die Menge $\{f_m \mid m \in M\}$ eine Basis von $\mathrm{Abb}_0(M, K)$. Mit anderen Worten: Wir haben M einen K-Vektorraum zugeordnet in der Art, dass diejenigen Elemente, die zu Elementen von M korrespondieren, eine Basis dieses Vektorraums bilden. Man nennt $\mathrm{Abb}_0(M, K)$ deshalb auch *freien K-Vektorraum über M*.

Beispielsweise für $M = \mathbb{Z}$ ist $\mathrm{Abb}_0(M, K)$ die Menge der endlichen K-wertigen Folgen indiziert über $\mathbb{Z}$. Für $M = \mathbb{R}^2$ und $(0, 0) \in \mathbb{R}^2$ ist

$$f_{(0,0)}\colon \mathbb{R} \times \mathbb{R}, \qquad (x, y) \longmapsto \begin{cases} 1, & \text{falls } x = 0 \text{ und } y = 0, \\ 0, & \text{sonst.} \end{cases}$$

Satz 30 (Existenz von Tensorprodukten): *Seien V_1 und V_2 Vektorräume über K. Dann existiert ein Tensorprodukt (T, τ) von V_1 und V_2 über K, welches eindeutig bis auf eindeutige Isomorphie ist. Wir schreiben $V_1 \otimes_K V_2 = V_1 \otimes V_2$ für T und für v_i aus V_i vereinbaren wir die Schreibweise $v_1 \otimes v_2 = \tau(v_1, v_2)$.*

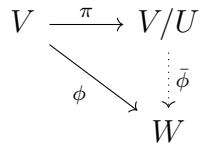
Für endlichdimensionale Vektorräume V_1 und V_2 liefert Proposition VII.4.1 das Gewünschte. In dieser Situation haben wir $\{(e_i, e'_j) \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ auf die Basis $\{E_{i,j} \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ von $K^{n \times m}$ abgebildet.

Erinnerung VII.4.10 (Homomorphiesatz): Für einen K-Vektorraum V und einen Untervektorraum U haben wir in der Linearen Algebra I den besonderen Vektorraum V/U kennengelernt. Dieser hieß „Quotientenvektorraum“ oder auch „Faktorraum“ und war definiert als die Menge der Äquivalenzklassen $V/U = \{[v]_\sim \mid v \in V\}$ bezüglich der auf V erklärten Äquivalenzrelation

$$v \sim w :\Longrightarrow v - w \in U,$$

welcher wir per $[v]+[w] := [v+w]$ und $\lambda[v] := [\lambda v]$ eine K-Vektorraumstruktur aufgeprägt haben. Für ein $v \in V$ ist die Äquivalenzklasse $[v]$ genau die Menge $v+U = \{v+u \mid u \in U\}$. Die Restklassenabbildung $\pi\colon V \to V/U$, $v \mapsto [v]$ heißt *kanonische Projektion* und ist per Konstruktion der K-Vektorraumstruktur von V/U trivialerweise eine lineare Abbildung.

Außerdem haben wir den Homomorphiesatz in der Linearen Algebra I kennengelernt: Für K-Vektorräume V und W, eine lineare Abbildung $\phi\colon V \to W$ und einen Untervektorraum $U \subseteq V$ mit $U \subseteq \operatorname{Kern}(\phi)$ haben wir das kommutative Diagramm



Das heißt es gibt genau eine lineare Abbildung $\bar{\phi}\colon V/U \to W$ mit $\bar{\phi} \circ \pi = \phi$.

Beweis (von Satz 30): Im ersten Schritt möchten wir einen (zu großen) Kandidaten für unser Tensorprodukt einführen, nämlich den freien K-Vektorraum F über der Menge $V_1 \times V_2$ zusammen mit der kanonischen Injektion $V_1 \times V_2 \hookrightarrow F$. Wir setzen also

$$F := \operatorname{Abb}_0(V_1 \times V_2, K), \qquad \tau_1\colon V_1 \times V_2 \longrightarrow F, \quad (v_1, v_2) \longmapsto f_{(v_1,v_2)}$$

mit $f_{(v_1,v_2)}$ wie in Proposition VII.4.9. Bis jetzt gut an diesen Kandidaten für T und τ ist, dass wir eine Basis $\{f_{(v_1,v_2)} \mid (v_1, v_2) \in V_1 \times V_2\}$ von F kennen, d. h. für jeden K-Vektorraum W und jede Abbildung $\beta\colon V_1 \times V_2 \to W$ erhalten wir per Fortsetzungssatz eine eindeutige lineare Abbildung $\hat{\phi}\colon F \to W$ mit $\hat{\phi} \circ \tau_1 = \beta$. (Dazu definieren wir $\hat{\phi}$ auf der Basis durch $\hat{\phi}(f_{(v_1,v_2)}) = \beta(v_1, v_2)$.) Schlecht ist, dass τ_1 weit entfernt davon ist, bilinear zu sein.

Im zweiten Schritt möchten wir unseren Kandidaten verbessern. Für alle $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$ und $\alpha_1, \alpha_2 \in K$ brauchen wir, dass

$$\tau(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') = \alpha_1\alpha_2\tau(v_1, v_2) + \alpha_1\tau(v_1, v_2') + \alpha_2\tau(v_1', v_2) + \tau(v_1', v_2').$$

Wir wollen im Folgenden durch geeignete Quotientenbildung „erzwingen“, dass genau das gilt, was wir uns wünschen.[2] Sei R der Untervektorraum von V, der von den Elementen

$$\{f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')} - \alpha_1\alpha_2 f_{(v_1,v_2)} - \alpha_1 f_{(v_1,v_2')} - \alpha_2 f_{(v_1',v_2)} - f_{(v_1',v_2')}\}$$

erzeugt wird. Wir setzen $T := F/R$, bezeichnen mit π die kanonische Projektion $\pi\colon F \to T = F/R$ und definieren $\tau\colon V_1 \times V_2 \to T = F/R$ durch $\tau := \pi \circ \tau_1$, also $(v_1, v_2) \mapsto [f_{(v_1,v_2)}] = f_{(v_1,v_2)} + R$. Dieses τ ist tatsächlich bilinear: Für $v_1, v_1' \in V$, $v_2, v_2' \in V'$ und $\alpha_1, \alpha_2 \in K$ gilt per Definition von $T = F/R$ und per Definition der Vektorraumstruktur auf T, dass

$$\begin{aligned}
&\tau(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') \\
&\quad = [f_{(\alpha v_1 + v_1', \alpha_2 v_2 + v_2')}] \\
&\quad = [\alpha_1 f_{(v_1,v_2')} + \alpha_2 f_{(v_1',v_2)} + f_{(v_1',v_2')} + \alpha_1\alpha_2 f_{(v_1,v_2)}] \\
&\quad = \alpha_1\alpha_2 [f_{(v_1,v_2)}] + \alpha_1 [f_{(v_1,v_2')}] + \alpha_2 [f_{(v_1',v_2)}] + [f_{(v_1',v_2')}] \\
&\quad = \alpha_1\alpha_2 \tau(v_1, v_2) + \alpha_1 \tau(v_1, v_2') + \alpha_2 \tau(v_1', v_2) + \tau(v_1', v_2').
\end{aligned}$$

Durch die Definition von T und τ haben wir die folgende Erweiterung des Diagramms von oben:

$$\begin{array}{ccccc}
V_1 \times V_2 & \xrightarrow{\tau_1} & F & \xrightarrow{\pi} & T = F/R \\
& \searrow^{\beta} & \downarrow \exists!\, \hat{\phi} & \swarrow_{\phi} & \\
& & W & &
\end{array}$$

Können wir zeigen, dass $R \subseteq \operatorname{Kern}(\hat{\phi})$, dann folgt aus dem Homomorphiesatz, dass es eine lineare Abbildung $\phi\colon T \to W$ gibt, die das rechte Dreieck im Diagramm kommutativ macht, d. h. die $\phi \circ \pi = \hat{\phi}$ leistet. Insgesamt ergibt das

$$\phi \circ \tau = \phi \circ \pi \circ \tau_1 = \hat{\phi} \circ \tau_1 = \beta,$$

d. h. wenn $R \subseteq \operatorname{Kern}(\hat{\phi})$, dann ist (T, τ) ein Tensorprodukt von V_1 und V_2. Um zu zeigen, dass R im Kern von $\hat{\phi}$ liegt, genügt es, das für die Erzeuger nachzurechnen. Für $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$ und $\alpha_1, \alpha_2 \in K$ gilt

$$\begin{aligned}
&\hat{\phi}(f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')}) \\
&\quad = \beta(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') \\
&\quad = \alpha_1\alpha_2 \beta(v_1, v_2) + \alpha_1 \beta(v_1, v_2') + \alpha_2 \beta(v_1', v_2) + \beta(v_1', v_2') \\
&\quad = \alpha_1\alpha_2 \hat{\phi}(f_{(v_1,v_2)}) + \alpha_1 \hat{\phi}(f_{(v_1,v_2')}) + \alpha_2 \hat{\phi}(f_{(v_1',v_2)}) + \hat{\phi}(f_{(v_1',v_2')})
\end{aligned}$$

[2] Gelegentlich nennt man so etwas auch „Pippi-Langstrumpf-Mathematik“.

was mit der Linearität von $\hat{\phi}$ die Behauptung liefert. Schließich ist ϕ eindeutig, weil $\{[f_{(v_1,v_2)}] \mid v_1 \in V_1, v_2 \in V_2\}$ ein Erzeugendensystem für T ist. □

Da wir $v_1 \otimes v_2 = \tau(v_1, v_2) = \pi(f_{(v_1,v_2)})$ vereinbart haben, und die Menge $\{f_{(v_1,v_2)} \mid v_1 \in V_1, v_2 \in V_2\}$ eine Basis von F ist, ist $\{v_1 \otimes v_2 \mid v_1 \in V_1, v_2 \in V_2\}$ ein Erzeugendensystem von $T = V_1 \otimes_K V_2$. Das bedeutet: Nicht jedes Element von $V_1 \otimes V_2$ ist von der Form $v_1 \otimes v_2$; ein beliebiges Element von $V_1 \otimes V_2$ können wir aber als endliche Summe von Elementen der Form $v_1 \otimes v_2$ schreiben.

Bemerkung VII.4.11 (Rechenregeln für reine Tensoren): Seien V_1 und V_2 Vektorräume über K und $V_1 \otimes V_2$ ihr Tensorprodukt. Sind v_1, v_1' Elemente von V_1, v_2 und v_2' Elemente von V_2 und λ in K, dann gilt:

(i) $(\lambda v_1) \otimes v_2 = \lambda(v_1 \otimes v_2) = v_1 \otimes (\lambda v_2)$,

(ii) $(v_1 + v_1') \otimes v_2 = v_1 \otimes v_2 + v_1' \otimes v_2$ und $v_1 \otimes (v_2 + v_2') = v_1 \otimes v_2 + v_1 \otimes v_2'$.

Das liegt einfach an unserer Schreibweise. Wir haben uns auf $v_1 \otimes v_2 = \tau(v_1, v_2)$ verständigt, und $\tau\colon V_1 \times V_2 \to V_1 \otimes V_2$ ist bilinear.

Proposition VII.4.12 (Basis des Tensorprodukts): *Seien K ein Körper, V_1 und V_2 zwei Vektorräume über K und B eine Basis von V_1 sowie C eine Basis von V_2. Dann ist*

$$D := \{b \otimes c \mid b \in B, c \in C\} \subseteq V_1 \otimes V_2$$

eine Basis von $V_1 \otimes V_2$.

Beweis: Wir haben zu zeigen, dass D das Tensorprodukt $V_1 \otimes V_2$ erzeugt und, dass D linear unabhängig ist. Wir kennen bereits ein Erzeugendensystem für $V_1 \otimes V_2$, nämlich $\{v_1 \otimes v_2 \mid v_1 \in V_1, v_2 \in V_2\}$. Wir zeigen für die Erzeugendeneigenschaft von D, dass D dieses Erzeugendensystem erzeugt. Seien dazu $v_1 \in V_1$ und $v_2 \in V_2$. Dann gibt es eindeutige Linearkombinationen $v_1 = \sum_{i=1}^{n} r_i b_i$ und $v_2 = \sum_{j=1}^{m} s_j c_j$ und

$$v_1 \otimes v_2 = \Big(\sum_{i=1}^{n} r_i b_i\Big) \otimes \Big(\sum_{j=1}^{m} s_j c_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} r_i s_j b_i \otimes c_j,$$

d. h. $v_1 \otimes v_2$ gehört zu $\mathrm{Lin}(D)$ wie gewünscht.

Nun zu linearen Unabhängigkeit: Sei $\mathbf{0} = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j} b_i \otimes c_j$ eine Nulldarstellung. Wegen der universellen Abbildungseigenschaft von $V_1 \otimes V_2$ gilt für jede Bilinearform $\beta\colon V_1 \times V_2 \to K$ und die zugehörige lineare Abbildung $\phi_\beta\colon V_1 \otimes V_2 \to K$, dass

$$0 = \phi_\beta(\mathbf{0}) = \phi_\beta\Big(\sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j} b_i \otimes c_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j}\beta(b_i, c_j).$$

Seien $1 \leq k \leq n$ und $1 \leq \ell \leq m$. Insbesondere für die Bilinearform erklärt durch

$$\beta_{(k,\ell)} \colon B \times C \longrightarrow K, \qquad (b_i, c_j) \longmapsto \begin{cases} 1, & \text{falls } (i,j) = (k,\ell), \\ 0, & \text{sonst.} \end{cases}$$

gilt die obige Gleichung, d. h. $0 = \sum_{i=1}^{n} \sum_{j=1}^{m} r_{i,j} \beta_{(k,\ell)}(b_i, c_j) = r_{k,\ell}$ und somit ist D linear unabhängig. □

Korollar VII.4.13 (Dimension des Tensorprodukts): *Seien K ein Körper und V_1 und V_2 endlichdimensionale Vektorräume über K. Dann gilt*

$$\dim(V_1 \otimes V_2) = (\dim V_1)(\dim V_2).$$

Bemerkung VII.4.14 (Induzierter Morphismus): Seien K ein Körper, V_1, V_2, V_1' und V_2' Vektorräume über K, sowie $\phi \colon V_1 \to V_2$ und $\psi \colon V_1' \to V_2'$ lineare Abbildungen. Ferner bezeichne $\phi \times \psi$ die Abbildung $V_1 \times V_1' \to V_2 \times V_2'$, $(v_1, v_1) \mapsto (\phi(v_1), \psi(v_1))$. Dann haben wir das Diagramm

$$\begin{array}{ccc} V_1 \times V_1' & \xrightarrow{\tau_1} & V_1 \otimes V_1' \\ \downarrow{\scriptstyle \phi\times\psi} & \searrow^{\tau_2 \circ (\phi\times\psi)} & \downarrow{\scriptstyle \phi\otimes\psi} \\ V_2 \times V_2' & \xrightarrow[\tau_2]{} & V_2 \otimes V_2' \end{array}$$

und die Abbildung längs des roten Pfeils ist bilinear. Das heißt für den roten Pfeil erhalten wir eine eindeutige lineare Abbildung von $V_1 \otimes V_1'$ nach $V_2 \otimes V_2'$, die wir suggestiv mit $\phi \otimes \psi$ bezeichnen wollen, die dadurch festgelegt ist, dass für alle $v_1 \in V_1$ und $v_1' \in V_1'$ gilt:

$$(\phi \otimes \psi)(v_1 \otimes v_1') = \phi(v_1) \otimes \psi(v_1').$$

An der Stelle weisen wir darauf hin, dass das die Abbildung (durch lineare Fortsetzung) wirklich „nur“ eindeutig festlegt, da wir nur angeben, was die Abbildung auf reinen Tensoren tun soll. Allgemeine Elemente von $V_1 \otimes V_1'$ sind aber endliche Summen reiner Tensoren.

Bemerkung VII.4.15 (Abbildungsmatrix von $\phi \otimes \psi$): Seien alle Vektorräume in Proposition VII.4.14 endlichdimensional mit Basen $B = (b_1, \ldots, b_m)$ von V_1, $C = (c_1, \ldots, c_n)$ von V_2, $B' = (b_1', \ldots, b_{m'}')$ von V_1' und $C' = (c_1, \ldots, c_{n'}')$ von V_2'. Ferner bezeichne $D_{C,B}(\phi) = (\beta_{i,j}) \in K^{n \times m}$ und $D_{C',B'}(\psi) = (\gamma_{k,\ell}) \in K^{n' \times m'}$ die Darstellungsmatrizen. Für die Basen

$$\begin{aligned} D_1 &:= (b_1 \otimes b_1', \ldots, b_1 \otimes b_{m'}', b_2 \otimes b_1', \ldots, b_2 \otimes b_{m'}', \ldots, b_m \otimes b_1', \ldots, b_m \otimes b_{m'}'), \\ D_2 &:= (c_1 \otimes c_1', \ldots, c_1 \otimes c_{n'}', \ldots, c_n \otimes c_1', \ldots, c_n \otimes c_{n'}') \end{aligned}$$

und die Darstellungsmatrix $A := D_{D_2,D_1}(\phi \otimes \psi)$ gilt

$$A = \begin{pmatrix} \beta_{1,1}D_{C',B'}(\psi) & \beta_{1,2}D_{C',B'}(\psi) & \cdots & \beta_{1,m}D_{C',B'}(\psi) \\ \beta_{2,1}D_{C',B'}(\psi) & \beta_{2,2}D_{C',B'}(\psi) & \dots & \beta_{2,m}D_{C',B'}(\psi) \\ \vdots & \vdots & & \vdots \\ \beta_{n,1}D_{C',B'}(\psi) & \beta_{n,2}D_{C',B'}(\psi) & \cdots & \beta_{n,m}D_{C',B'}(\psi) \end{pmatrix}.$$

Man nennt A auch das Kroneckerprodukt von $D_{C,B}(\phi)$ und $D_{C',B'}(\psi)$.

Beweis: Für $1 \leq i \leq m$ und $1 \leq j \leq m'$ gilt

$$\begin{aligned}(\phi \otimes \psi)(b_i \otimes b'_j) &= \phi(b_i) \otimes \psi(b'_j) \\ &= \Big(\sum_{i=1}^{n} \beta_{k,i} c_k\Big) \otimes \Big(\sum_{\ell=1}^{n'} \gamma_{\ell,j} c'_\ell\Big) = \sum_{k=1}^{n}\sum_{\ell=1}^{n'} \beta_{k,i}\gamma_{\ell,j} c_k \otimes c'_\ell.\end{aligned}$$

Die Zeilen von A sind indiziert durch die Basiselemente $c_k \otimes c'_\ell$ und die Spalten von A sind indiziert durch die Basiselemente $b_i \otimes b'_j$. □

Proposition VII.4.16: *Seien K ein Körper, $(A, +, \circ, \cdot)$ und $(B, +, \circ, \cdot)$ zwei Algebren über K und $A \otimes_K B = A \otimes B$ das Tensorprodukt von A und B als K-Vektorräume. Dann wird das Tensorprodukt $A \otimes B$ zu einer K-Algebra durch*

$$\begin{aligned}\bullet\colon A \otimes B \times A \otimes B &\longrightarrow A \otimes B, \\ \Big(\sum_{i=1}^{m} a_i \otimes b_i, \sum_{j=1}^{n} a'_j \otimes b'_j\Big) &\longmapsto \sum_{i=1}^{m}\sum_{j=1}^{n} (a_i \circ a'_j) \otimes (b_i \circ b'_j).\end{aligned}$$

Insbesondere gilt für $a, a' \in A$, $b, b' \in B$ und die reinen Tensoren $a \otimes b$, $a' \otimes b'$, dass $(a \otimes b) \bullet (a' \otimes b') = (a \circ a') \otimes (b \circ b')$.

Beweis: Wir haben zu zeigen, dass die Multiplikation „$\circ$“ auf $A \otimes B$ wohldefiniert ist. Dass diese Verknüpfung $A \otimes B$ zu einer K-Algebra macht, ist eine Standardrechnung, wobei man auf die Algebreneigenschaften von A und B zurückführt.

Wir erinnern uns daran, dass $A \otimes B$ per Konstruktion der Vektorraum F/R ist, wobei $F = \mathrm{Abb}_0(A \times B)$, für den $\{f_{(a,b)} \mid a \in A, b \in B\}$ eine Basis bildet, und $R \subseteq F$ der Unterraum ist, der von Elementen der Form $f_{(ra_1+a_2, sb_1+b_2)} - rsf_{(a_1,b_1)} - rf_{(a_1,b_2)} - sf_{(a_2,b_1)} - f_{(a_2,b_2)}$ erzeugt wird. Seien $r, s \in K$, $a, a_1, a_2 \in A$ und $b, b_1, b_2 \in B$. Dann haben wir

$$\begin{aligned}&(ra_1 + a_2) \otimes (sb_1 + b_2) \bullet (a \otimes b) \\ &\quad = (ra_1 + a_2) \circ a \otimes (sb_1 + b_2) \circ b \\ &\quad = rsa_1 \circ a \otimes b_1 \circ b + ra_1 \circ a \otimes b_2 \circ b + a_2 \circ a \otimes sb_1 \circ b + a_2 \circ a \otimes b_2 \circ b \\ &\quad = (rsa_1 \otimes b_1 + ra_1 \otimes b_2 + a_2 \otimes sb_1 + a_2 \otimes b_2) \bullet (a \otimes b).\end{aligned}$$ □

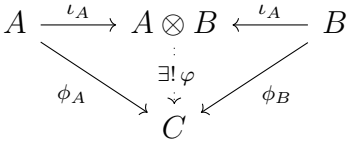
Proposition VII.4.17: *Seien K ein Körper, A und B kommutative unitäre K-Algebren.*

(i) *Die Abbildungen*

$$\iota_A\colon A \longrightarrow A \otimes B, \qquad a \longmapsto a \otimes 1$$
$$\iota_B\colon B \longrightarrow A \otimes B, \qquad b \longmapsto 1 \otimes b$$

sind K-Algebren-Homomorphismen.

(ii) *Das Tensorprodukt $A \otimes B$ hat die folgende universelle Abbildungseigenschaft: Für jede kommutative unitäre K-Algebra C und K-Algebren-Homomorphismen $\phi_A\colon A \to C$, $\phi_B\colon B \to C$ gibt es genau einen K-Algebren-Homomorphismus $\varphi\colon A \otimes B \to C$, der das folgende Diagramm kommutativ macht:*



Beweis: (i) Für a_1, a_2 aus A gilt

$$\iota_A(a_1 \circ a_2) = a_1 \circ a_2 \otimes 1 = (a_1 \otimes 1) \bullet (a_2 \otimes 1) = \iota_A(a_1) \bullet \iota_A(a_2),$$

ferner ist $\iota_A(1) = 1 \otimes 1$ die Eins von $A \otimes B$. Die weiteren Eigenschaften folgen aus der Bilinearität des Tensorprodukts.

(ii) Die Abbildung

$$\beta\colon A \times B \longrightarrow C, \qquad (a, b) \longmapsto \phi_A(a) \circ \phi_A(b)$$

ist bilinear, was uns einen eindeutigen Homomorphismus $\varphi\colon A \otimes B \to C$ von K-Vektorräumen liefert, der $\varphi(a \otimes b) = \phi_A(a) \circ \phi_B(b)$ erfüllt. Dieses φ ist sogar ein K-Algebren-Homomorphismus, denn für $a_1, a_2 \in A$ und $b_1, b_2 \in B$ gilt

$$\begin{aligned}
\varphi((a_1 \otimes b_1) \bullet (a_2 \otimes b_2)) &= \varphi(a_1 \circ a_2 \otimes b_1 \circ b_2) \\
&= \phi_A(a_1 \circ a_2) \circ \phi_A(b_1 \circ b_2) \\
&= \phi_A(a_1) \circ \phi_A(a_2) \circ \phi_B(b_1) \circ \phi_B(b_2) \\
&= \phi_A(a_1) \circ \phi_B(b_1) \circ \phi_A(a_2) \circ \phi_B(b_2) \\
&= \varphi(a_1 \otimes b_1) \circ \varphi(a_2 \otimes b_2). \qquad \square
\end{aligned}$$

Möchte man auch nicht-kommutative K-Algebren betrachten, so fordert man bei der universellen Abbildungseigenschaft, dass für $a \in A$ und $b \in B$ gilt: $\phi_A(a) \circ \phi_B(b) = \phi_B(b) \circ \phi_A(a)$.

Kapitel VIII.

Euklidische und unitäre Vektorräume

In diesem Kapitel möchten wir Vektorräume über $\mathbb{K}$ (d. h. dem Körper der reellen oder komplexen Zahlen), die mit einem Skalarprodukt ausgestattet sind, untersuchen. Das Skalarprodukt beschert uns eine Norm, und damit einen Abstandsbegriff auf dem entsprechenden Vektorraum. Außerdem können wir im reellen Fall Winkel messen, und haben im Allgemeinen das Konzept von Orthogonalität zur Verfügung.

Die Frage, welche speziellen linearen Abbildungen die zusätzliche Struktur – die Skalarprodukte bzw. die Längenbegriffe – berücksichtigen, führt uns auf die sogenannten Spektralsätze.

1. Die Spektralsätze

Definition VIII.1.1 (Adjungierte Matrix): Sei $A = (a_{i,j})$ eine komplexe $n \times m$-Matrix. Mit $\overline{A}$ bezeichnen wir die eintragsweise komplex konjugierte Matrix, d. h. $\overline{A} = (\overline{a}_{i,j})$. Die Matrix $A^* = \overline{A}^t$ heißt *zu A adjungierte Matrix*.

Erinnerung VIII.1.2: Sei $A = (a_{i,j})$ eine $n \times n$-Matrix. Sind die Einträge von A reell und ist $A^t = A$, so heißt A *symmetrisch*. Sind die Einträge komplex und ist $A^* = A$, dann heißt A *selbstadjungiert* beziehungsweise *hermitesch*. Sind die Einträge reell, ist A invertierbar und gilt $A^{-1} = A^t$, dann heißt A *orthogonal*. Sind die Einträge komplex, ist A invertierbar und ist $A^{-1} = A^*$, so heißt A *unitär*.

Beispiel VIII.1.3 (Drehkästchen): Für einen reellen Winkel α beschreibt

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

eine Drehung der Ebene um den Winkel α. Wir haben

$$AA^t = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

das heißt A ist orthogonal.

Definition VIII.1.4 (Normale Matrizen): Ist A eine $n \times n$-Matrix mit reellen Einträgen und ist $A^tA = AA^t$, dann heißt A *normal*. Sind die Einträge komplex und gilt $A^*A = AA^*$, dann heißt A *normal*.

Beispiel VIII.1.5 (für normale Matrizen): Ist A eine reelle symmetrische $n \times n$-Matrix, dann ist A tautologisch normal. Genau so sind selbstadjungierte Matrizen normal. Ist A orthgonal, d. h. $A^{-1} = A^t$, dann ist A ebenfalls normal. Ebenfalls sind unitäre Matrizen normal.

Satz 31 (Spektralsatz): *Sei A eine $n \times n$-Matrix mit komplexen Einträgen. Ist A normal, so gibt es einen unitären Basiswechsel S und komplexe Zahlen $\lambda_1, \ldots, \lambda_n$, sodass $A = S\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)S^{-1}$.*

Satz 32 (Hauptachsentransformationssatz):

(i) *Sei A in $\mathbb{C}^{n\times n}$ hermitesch. Dann gibt es* reelle *Zahlen $\lambda_1, \ldots, \lambda_n$ und eine unitäre $n \times n$-Matrix S, sodass $S^{-1}AS = S^*AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

(ii) *Sei A in $\mathbb{R}^{n\times n}$ symmetrisch. Dann gibt es reelle Zahlen $\lambda_1, \ldots, \lambda_n$ und eine orthogonale $n \times n$-Matrix S, sodass $S^{-1}AS = S^tAS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

Satz 33 (Isometrienormalform):

(i) *Sei A in $\mathbb{C}^{n\times n}$ unitär. Dann gibt es eine komplexe Zahlen $\lambda_1, \ldots, \lambda_n$ mit $|\lambda_i| = 1$ für $1 \le i \le n$ und eine unitäre Matrix S in $\mathbb{C}^{n\times n}$ mit $S^{-1}AS = S^*AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

(ii) *Sei A in $\mathbb{R}^{n\times n}$ orthogonal. Dann gibt es reelle Zahlen $\alpha_1, \ldots, \alpha_k$ und eine orthogonale $n \times n$-Matrix S, sodass*

$$S^{-1}AS = S^tAS = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & A_{\alpha_1} & & \\ & & & & & & & \ddots & \\ & & & & & & & & A_{\alpha_k} \end{pmatrix},$$

wobei A_{α_i} das 2×2-Drehkästchen zum Winkel α_i bezeichnet.

2. Euklidische und unitäre Vektorräume

In diesem Abschnitt möchten wir Skalarprodukte, die wir bis jetzt nur für reelle Vektorräume gesehen haben, auch für komplexe Vektorräume einführen. Wie im reellen Fall haben wir dann einen Längen- und Abstandsbegriff (sowie einen abgeschwächten Winkelbegriff) zur Hand.

Erinnerung VIII.2.1: Seien V ein reeller Vektorraum und $\beta\colon V \times V \to \mathbb{R}$ eine Abbildung. Ist β bilinear, symmetrisch und positiv definit, dann heißt β ein *Skalarprodukt auf V*.

Diese Definition wollen wir passend verallgemeinern. Um einen sinnvollen Definitheitsbegriff zu erhalten, können wir nicht verbatim dieselbe Definition verwenden, da $\mathbb{C}$ nicht angeordnet ist und so „$\beta(v,v) \geq 0$“ erstmal keinen Sinn ergibt.

Ist $\mathbb{R}^n$ ausgestattet mit dem Standard-Skalarprodukt, und ist $x = (x_1, \ldots, x_n)^t$, dann erklärt $\|x\| = (\sum_{i=1}^n x_i^2)^{1/2}$ die von $\langle \cdot, \cdot \rangle$ induzierte Norm.

Für $n = 1$ ist die gewöhnliche Multiplikation auf $\mathbb{R}$ das Standard-Skalarprodukt und $|x| = \sqrt{x^2}$ die zugehörige Norm. In $\mathbb{C}$ haben wir $|z| = \sqrt{z\overline{z}}$, was uns im Folgenden inspirieren soll.

Beispiel VIII.2.2: Auf $V = \mathbb{C}^n$ betrachten wir die Funktion $h\colon (z,w) \mapsto z\overline{w}$. Für dieses h gilt $h(cz_1 + z_2, w) = ch(z_1, w) + h(z_2, w)$, und

$$h(z, cw_1 + w_2) = z\overline{(cw_1 + w_2)} = \overline{c}z\overline{w}_1 + z\overline{w}_2 = \overline{c}h(z, w_1) + h(z, w_2),$$

d. h. h ist nicht bilinear. Außerdem gilt $h(z,w) = z\overline{w}$ sowie $h(w,z) = w\overline{z}$, d. h. $h(z,w) = \overline{h(w,z)}$. Was wir dadurch jedoch gewonnen haben ist, dass $h(z,z) = \overline{h(z,z)}$ für jedes z in $\mathbb{C}^n$, d. h. über positive Definitheit von h zu sprechen ergibt Sinn.

Definition VIII.2.3: Seien V ein $\mathbb{C}$-Vektorraum und $s\colon V \times V \to \mathbb{C}$ eine Abbildung. Gilt für alle $\lambda \in \mathbb{C}$, $u_1, u_2, u, v_1, v_2, v \in V$, dass

(i) $s(\lambda u_1 + u_2, v) = \lambda s(u_1, v) + s(u_2, v)$,

(ii) $s(u, \lambda v_1 + v_2) = \bar{\lambda} s(u, v_1) + s(u, v_2)$,

dann heißt s eine „Sesquilinearform“.[1]

Ist $h\colon V \times V \to \mathbb{C}$ eine Sesquilinearform und gilt für alle $v, w \in V$, dass $h(v,w) = \overline{h(w,v)}$, dann heißt h eine *hermitesche Form*.[2]

Beispiel VIII.2.4: (i) Die Abbildung

$$h\colon \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \qquad x^t\bar{y} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_n \end{pmatrix} = \sum_{i=1}^{n} x_i \bar{y}_i$$

ist eine hermitesche Form.

(ii) Ist $G \in \mathbb{C}^{n\times n}$ und bezeichnet

$$h_G\colon \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \qquad (x,y) \longmapsto x^t G \bar{y},$$

dann ist h_G eine Sesquilinearform. Genau dann ist h_G hermitesch, wenn $\bar{G}^t = G$.

Proposition VIII.2.5: *Ist $h\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ eine Sesquilinearform, dann ist $h = h_G$ für die Matrix $G = (g_{i,j})$ mit $g_{i,j} = h(e_i, e_j)$.*

Wie immer bezeichnet hier $(e_1, \ldots, e_n)$ die Standardbasis des $\mathbb{C}^n$. Die Aussage zeigt man ganz analog zur folgenden Proposition II.2.6.

Proposition VIII.2.6: *Seien V ein endlichdimensionaler $\mathbb{C}$-Vektorraum, $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V und $h\colon V\times V \to \mathbb{C}$ eine Sesquilinearform. Die durch $g_{i,j} = h(b_i, b_j)$ definierte Matrix $G = (g_{i,j})$ in $\mathbb{C}^{n\times n}$ heißt* Gram-Matrix von h bezüglich der Basis B. *Sie hat folgende Eigenschaften:*

(i) *Für alle v, w in V ist $h(v,w) = D_B(v)^t G \overline{D_B(w)}$, d. h. das folgende Diagramm kommutiert:*

$$\begin{array}{ccc} V \times V & \xrightarrow{\ h\ } & \mathbb{C} \\ {\scriptstyle D_B\times D_B}\downarrow & & \| \\ \mathbb{C}^n \times \mathbb{C}^n & \xrightarrow[h_G]{} & \mathbb{C} \end{array}$$

(ii) *Ist $B' = (b'_1, \ldots, b'_n)$ eine weitere Basis von V und ist G' die Gram-Matrix von h bezüglich B', dann gilt $G' = D^t_{B,B'} G \bar{D}_{B,B'}$.*

[1] „Sesqui“ ist das lateinische Wort für „anderthalb“, eine Sesquilinearform ist also „anderthalbfach linear“.

[2] Die hermitsche Form verdankt ihren Namen dem französischen Mathematiker Charles Hermite (1822-1901).

Bemerkung VIII.2.7: Seien V ein $\mathbb{C}$-Vektorraum und $h\colon V \times V \to \mathbb{C}$ eine hermitesche Form auf V, dann gilt für alle v in V, dass $h(v,v) = \overline{h(v,v)}$. Insbesondere ist $h(v,v)$ eine relle Zahl.

Definition VIII.2.8: Seien V ein $\mathbb{C}$-Vektorraum und $h\colon V \times V \to \mathbb{C}$ eine hermitesche Form auf V.

(i) Gilt für alle $v \in V - \{0\}$, dass $h(v,v) > 0$, dann heißt h *positiv definit*.

(ii) Eine positiv definite hermitesche Sesquilinearform $h\colon V \times V \to \mathbb{C}$ heißt *Skalarprodukt auf* V. In diesem Fall schreibt man für gewöhnlich $\langle v, w\rangle$ für $h(v,w)$.

Üblicherweise schreibt man $\mathbb{K}$ stellvertretend für den Körper der reellen Zahlen $\mathbb{R}$ oder den Körper der komplexen Zahlen $\mathbb{C}$.

Definition VIII.2.9 (Prähilbertraum, euklidischer- oder unitärer Vektorraum): Ein $\mathbb{K}$-Vektorraum mit Skalaprodukt heißt *Prähilbertraum*. Ein $\mathbb{R}$-Vektorraum mit Skalarprodukt heißt *euklidischer Vektorraum*. Ein $\mathbb{C}$-Vektorraum mit Skalaprodukt heißt *unitärer Vektorraum*.

Manchmal wird in der Literatur für euklidische- oder unitäre Vektorräume endlichdimensionalität verlangt.

Definition VIII.2.10 (Norm, Länge, Abstand): Sei $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum.

(i) Für $v \in V$ heißt $\|v\| := \langle v, v\rangle^{1/2}$ die *Norm* oder *Länge von* v.

(ii) Für $v, w \in V$ heißt $d(v,w) := \|v - w\|$ *Abstand von* v *und* w. Die Abbildung
$$d\colon V \times V \longrightarrow \mathbb{R}_{\geq 0}, \qquad (v,w) \longmapsto d(v,w)$$
heißt *Metrik zu* $\langle\cdot,\cdot\rangle$.

Satz 34 (Cauchy-Schwarz'sche Ungleichung): *Sei* $(V, \langle\cdot,\cdot\rangle)$ *ein Prähilbertraum.*

(i) *Für* $v, w \in V$ *gilt* $|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$*, und somit* $|\langle v, w\rangle| \leq \|v\|\|w\|$.

(ii) *Gleichheit in (i) gilt genau dann, wenn* v *und* w *linear abhängig sind.*

Beweis: (i) Für $w = \mathbf{0}$ stimmen beide Behauptungen. Wir dürfen also annehmen, dass $w \neq \mathbf{0}$. Wir betrachten zunächst den Fall „$\langle v, w\rangle \in \mathbb{R}$“ und definieren $f\colon \mathbb{R} \to \mathbb{R}$ durch $\lambda \mapsto \|v + \lambda w\|^2$. Dann ist
$$\begin{aligned} f(\lambda) = \langle v + \lambda w, v + \lambda w\rangle &= \langle v, v\rangle + \lambda\langle w, v\rangle + \lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle \\ &= \langle v, v\rangle + 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle \end{aligned}$$

Die Zahlen $a := \langle v, v\rangle$, $b := 2\langle v, w\rangle$ und $c := \langle w, w\rangle$ sind reelle Zahlen, außerdem ist $a \neq 0$ per Voraussetzung. Das heißt, f ist in Wahrheit eine quadratische Funktion – mit Schulwissen erkennt man, dass der zugehörige Graph eine „nach oben geöffnete Parabel“ ist. Per Definition der Funktion f gilt für alle $\lambda \in \mathbb{R}$, dass $f(\lambda) \geq 0$, d. h. f hat höchstens eine Nullstelle, was äquivalent zu „$D = b^2 - 4ac \leq 0$“ ist. Ausgeschrieben erhalten wir also

$$4|\langle v, w\rangle|^2 - 4\langle w, w\rangle\langle v, v\rangle \leq 0,$$

wobei die Betragsstriche um $\langle v, w\rangle$ keine Rolle spielen, da $\langle v, w\rangle$ nach Voraussetzung eine reelle Zahl ist. Umstellen liefert die Behauptung.

Nun betrachten wir den allgemeinen Fall „$\langle v, w\rangle \in \mathbb{C}^\times$“. Die komplexe Zahl $\alpha := \langle v, w\rangle / |\langle v, w\rangle|$ hat Betrag Eins. Wir setzen $v' := \alpha^{-1}v$ und erhalten daraus

$$\langle v', w\rangle = \left\langle \frac{|\langle v, w\rangle|}{\langle v, w\rangle} v, w \right\rangle = |\langle v, w\rangle|.$$

Nun können wir schreiben $|\langle v, w\rangle|^2 = |\langle \alpha v', w\rangle|^2 = |\alpha|^2 |\langle v', w\rangle|^2$. Weil $\langle v', w\rangle$ nach der obigen Gleichung eine reelle Zahl ist, liefert der erste Schritt, dass

$$|\langle v, w\rangle|^2 \leq \langle v', v'\rangle\langle w, w\rangle = \langle \alpha^{-1}v, \alpha^{-1}v\rangle\langle w, w\rangle = |\alpha^2|\langle v, v\rangle\langle w, w\rangle,$$

wie gewünscht.

(ii) Im ersten Schritt von (i) gilt Gleichheit in der Cauchy-Schwarz'schen Ungleichung genau dann, wenn $D = 0$ oder $w = \mathbf{0}$. Die Diskriminante ist Null genau dann, wenn f eine Nullstelle λ hat, was wegen der positiven Definitheit bedeutet, dass $v = -\lambda w$. In beiden Fällen sind v und w linear abhängig.

Im zweiten Schritt von (i) gilt Gleichheit in der Cauchy-Schwarz'schen Ungleichung genau dann, wenn v' und w linear abhängig sind. Wegen der Definition von v' als Vielfaches von v gilt Gleichheit also genau dann, wenn v und w linear abhängig sind. □

Proposition VIII.2.11: *Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und*

$$\|\cdot\| \colon V \longrightarrow \mathbb{R}, \qquad v \longmapsto \|v\| := \langle v, v\rangle^{1/2}.$$

Die Abbildung $\|\cdot\|$ hat folgende Eigenschaften:

(i) *Für alle $v \in V$ ist $\|v\| \geq 0$ und $\|v\| = 0$ gilt genau dann, wenn $v = \mathbf{0}$.*

(ii) *Für alle $v \in V$ und $\lambda \in \mathbb{K}$ ist $\|\lambda v\| = |\lambda|\|v\|$.*

(iii) *Für alle* $v, w \in V$ *gilt* $\|v + w\| \leq \|v\| + \|w\|$.

Eigenschaft (i) wird *Positive Definitheit*, (ii) wird *Homogenität* und (iii) wird *Dreiecksungleichung* genannt.

Beweis: Aussagen (i) und (ii) kann man direkt nachrechnen. Zu (iii): Es gilt

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

Wegen der Monotonie der Wurzel folgt die Behauptung. □

Definition VIII.2.12: Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $N\colon V \to \mathbb{R}$ mit den Eigenschaften (i), (ii) und (iii) aus Proposition VIII.2.11 heißt *Norm auf* V. Das Tupel $(V, \|\cdot\|)$ heißt *normierter Vektorraum über* $\mathbb{K}$ oder kurz *normierter Vektorraum*.

Proposition VIII.2.13: *Seien* $(V, \|\cdot\|)$ *ein normierter Vektorraum und*

$$d\colon V \times V \longrightarrow \mathbb{R}, \qquad (v, w) \longmapsto d(v, w) := \|v - w\|.$$

Die Abbildung d *hat folgende Eigenschaften:*

(i) *Für alle* $v, w \in V$ *ist* $d(v, w) \geq 0$ *und* $d(v, w) = 0$ *genau dann, wenn* $v = w$.

(ii) *Für alle* $v, w \in V$ *ist* $d(v, w) = d(w, v)$.

(iii) *Für alle* $u, v, w \in V$ *ist* $d(u, v) + d(v, w) \geq d(u, w)$.

Eigenschaft (i) heißt *Positive Definitheit* und Eigenschaft (iii) heißt *Dreiecksungleichung*.

Beweis: Aussage (i) folgt aus der positiven Definitheit der Norm und Aussage (ii) folgt aus der Homogenität der Norm. Für (iii) müssen wir nur einmal die Dreiecksungleichung für die Norm verwenden; für $u, v, w \in V$ ist nämlich

$$d(u, v) + d(v, w) = \|u - v\| + \|v - w\| \geq \|u - v + v - w\| = \|u - w\| = d(u, w)$$

wie gewünscht. □

Definition VIII.2.14: Sei X eine Menge. Eine Abbildung $d\colon X \times X \to \mathbb{R}$ mit den Eigenschaften (i), (ii), (iii) aus Proposition VIII.2.13 heißt *Metrik auf X*.

Achtung: Nicht jede Norm kommt von einem Skalarprodukt und nicht jede Metrik wird von einer Norm induziert!

Korollar VIII.2.15 (Winkeldefinition): *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum. Dann gilt:*

(i) *Für alle $v, w \in V - \{0\}$ ist $-1 \leq \langle v, w\rangle/(\|v\|\|w\|) \leq 1$, d. h. es gibt genau ein $\alpha \in [0, \pi]$ mit $\cos(\alpha) = \langle v, w\rangle/(\|v\|\|w\|)$. Wir schreiben $\measuredangle(v, w) := \alpha$.*

(ii) *Zwei Vektoren $v, w \in V$ sind orthogonal genau dann, wenn $\langle v, w\rangle = 0$, also wenn $\alpha = \measuredangle(v, w) = \pi/2$.*

3. Orthogonale und unitäre Endomorphismen

In diesem Abschnitt möchten wir die strukturerhaltenden Abbildungen eines Prähilberträums $(V, \langle\cdot,\cdot\rangle)$ kennenlernen, das heißt wir suchen Abbildungen $\phi\colon V \to V$, die die Vektorraumstruktur erhalten (solche nennt man linear) die gleichzeitig das Skalarprodukt erhalten, d. h. für alle $v, w \in V$ soll gelten: $\langle v, w\rangle = \langle\phi(v), \phi(w)\rangle$ – solche Abbildungen erhalten dann auch Längen, beziehungsweise sogar Winkel im euklidischen Fall.

Definition VIII.3.1: Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und $\phi\colon V \to V$ ein Endomorphismus. Ist $\mathbb{K} = \mathbb{R}$ und gilt für alle $v, w \in V$, dass $\langle\phi(v), \phi(w)\rangle = \langle v, w\rangle$, dann heißt ϕ *orthogonal bezüglich* $\langle\cdot,\cdot\rangle$ oder kurz *orthogonal*. Ist $\mathbb{K} = \mathbb{C}$ und gilt für alle $v, w \in V$, dass $\langle\phi(v), \phi(w)\rangle = \langle v, w\rangle$, dann heißt ϕ *unitär bezüglich* $\langle\cdot,\cdot\rangle$ oder kurz *unitär*.

Bemerkung VIII.3.2: Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und $\phi \in \mathrm{End}(V)$ orthogonal beziehungsweise unitär.

(i) Für alle $v \in V$ gilt $\|\phi(v)\| = \|v\|$. Abbildungen mit dieser Eigenschaft nennt man *Isometrie*.

(ii) Seien $v, w \in V$. Genau dann gilt $\langle v, w\rangle = 0$, wenn $\langle\phi(v), \phi(w)\rangle = 0$. Das heißt, orthogonale respektive unitäre Abbildungen erhalten die Orthogonalität.

(iii) Isometrien sind injektiv. Ist nämlich $v \in V$ mit $\phi(v) = \mathbf{0}$, dann ist $\|\phi(v)\| = 0$, d. h. $\|v\| = 0$ und wegen den Eigenschaften der Norm heißt das $v = \mathbf{0}$.

(iv) Ist ϕ regulär, dann ist auch ϕ^{-1} orthogonal beziehungsweise unitär.

(v) Kompositionen orthogonaler beziehungsweise unitärer Abbildungen sind orthogonal beziehungsweise unitär. Genauer: Ist auch $\psi\colon V \to V$ orthogonal beziehungsweise unitär, dann ist $\phi \circ \psi$ und $\psi \circ \phi$ orthogonal beziehungsweise unitär.

Proposition VIII.3.3 (Matrix-Version von Orthogonal/Unitär): *Seien V ein endlichdimensionaler Prähilbertraum, B eine Basis von V, ϕ ein Endomorphismus von V, $A = D_{B,B}(\phi)$ und $G = G_B(\langle\cdot,\cdot\rangle)$ die Gram-Matrix bezüglich B. Dann gilt:*

(i) *Genau dann ist ϕ orthogonal beziehungsweise unitär, wenn $A^t G\overline{A} = G$.*

(ii) *Ist B eine Orthonormalbasis, dann ist ϕ orthogonal beziehungsweise unitär genau dann, wenn $A = D_{B,B}(\phi)$ orthogonal beziehungsweise unitär im Sinne von Proposition VIII.1.2 ist.*

Beweis: (i) Per charakterisierender Eigenschaft der Abbildungsmatrix haben wir für v aus V, dass $D_B(\phi(v)) = AD_B(v)$. Orthogonalität respektive Unitarität von ϕ bedeutet genau, dass $\langle\phi(v), \phi(w)\rangle = \langle v, w\rangle$ für alle v und w aus V gilt. Als letztes benötigen wir die Definition der Gram-Matrix; für alle v und w aus V leistet die nämlich $\langle v, w\rangle = D_B(v)^t G_B D_B(w)$ für v und w aus V. Damit erhalten wir für jedes Paar v und w aus V, dass

$$\begin{aligned}\langle\phi(v), \phi(w)\rangle &= D_B(\phi(v))^t G_B \overline{D_B(\phi(w))} \\ &= \Big(D_{B,B}(\phi) D_B(v)\Big)^t G_B \overline{\Big(D_{B,B}(\phi) D_B(w)\Big)} \\ &= D_B(v)^t \Big(D_{B,B}(\phi)^t G_B \overline{D_{B,B}(\phi)}\Big)\overline{D_B(w)}.\end{aligned}$$

Wegen $\langle v, w\rangle = D_B(v) G_B \overline{D_B(w)}$ ist ϕ orthogonal beziehungsweise unitär, wenn die letzte Zeile gleich $D_B(v)^t G_B \overline{D_B(w)}$ ist. Setzt man Paare aus $B \times B$ ein, dann erhält man die einzelnen Matrixeinträge der beiden Matrizen.

(ii) Das ist ein reiner Vergleich von Aussage (i) und Proposition VIII.1.2. □

Erinnerung VIII.3.4 (Komplexe Konjugation): Mit stumpfem Nachrechnen zeigt man für komplexe Zahlen z_1, z_2, dass $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2$ sowie $\overline{z_1 z_2} = \overline{z}_1 \overline{z}_2$. Sind A und B Matrizen geeigneter Größe, dann ist auch $\overline{A + B} = \overline{A} + \overline{B}$ – klar, da komponentenweise addiert wird – und $\overline{AB} = \overline{A} \cdot \overline{B}$, denn

$$(\overline{AB})_{i,j} = \overline{\sum_k a_{i,k} b_{k,i}} = \sum_k \overline{a}_{i,k} \overline{b}_{k,j} = (\overline{A} \cdot \overline{B})_{i,j}.$$

Proposition VIII.3.5 (Charakterisierung orthogonaler und unitärer Matrizen): *Sei A in $\mathbb{K}^{n\times n}$. Die folgenden Aussagen sind äquivalent:*

(i) *Die Matrix A ist orthogonal beziehungsweise unitär.*

(ii) *Die Zeilen von A bilden eine Orthonormalbasis des $\mathbb{K}^n$.*

(iii) *Die Spalten von A bilden eine Orthonormalbasis des $\mathbb{K}^n$.*

Beweis: Wir zeigen exemplarisch die Äquivalenz „(i) $\Longleftrightarrow$ (ii)“. Angenommen, A ist unitär, d. h. $A^* = \overline{A}^t = A^{-1}$. Insbesondere gilt dann $AA^* = I_n$, und wir erhalten

$$\delta_{i,j} = (I_n)_{i,j} = (AA^*)_{i,j} = (e_i^t A)(A^* e_j) = (A^t e_i)^t \overline{(A^t e_j)} = \langle A^t e_i, A^t e_j\rangle. \qquad \square$$

Bemerkung VIII.3.6 (Basiswechsel in Prähilberträumen): Seien V ein endlichdimensionaler Prähilbertraum, $C = (c_1, \ldots, c_n)$ eine Orthonormalbasis von V und $B = (b_1, \ldots, b_n)$ eine weitere Basis von V. Nach Satz 27 gilt dann $b_i = \sum_{j=1}^n \langle b_i, c_j\rangle c_j$ für $1 \le i \le n$; dem kommutativen Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}} & V \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} \\ \mathbb{K}^n & \xrightarrow[D_{C,B}(\mathrm{id})]{} & \mathbb{K}^n \end{array}$$

entnimmt man also $D_{C,B}(\mathrm{id}) = (\alpha_{i,j})_{i,j} = (\langle b_i, c_j\rangle)_{i,j}$. Mit dieser Charakterisierung der Einträge und Proposition VIII.3.5 kann man zeigen: Genau dann ist eine solche Basiswechselmatrix unitär, wenn auch B eine Orthonormalbasis ist.

Definition VIII.3.7 (Orthogonale und unitäre Gruppen): Sei n eine natürliche Zahl. Die Mengen

$$O(n) = \{A \in \mathbb{R}^{n\times n} \mid A^t A = I_n\}, \qquad SO(n) = \{A \in O(n) \mid \det(A) = 1\},$$
$$U(n) = \{A \in \mathbb{C}^{n\times n} \mid A^* A = I_n\}, \qquad SU(n) = \{A \in U(n) \mid \det(A) = 1\},$$

heißen *(spezielle) orthogonale Gruppe* respektive *(spezielle) unitäre Gruppe*. Es handelt sich um Gruppen.

Bemerkung VIII.3.8 (Eigenwerte orthogonaler bzw. unitärer Matrizen): Ist A in $O(n)$ beziehungsweise $U(n)$ und ist λ ein Eigenwert von A, dann liegt λ in $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$. Es gibt dann nämlich einen von Null verschiedenen Eigenvektor v, und aus $\|v\| = \|Av\| = \|\lambda v\| = |\lambda|\|v\|$ können wir deshalb $|\lambda| = 1$ folgern.

4. Adjungierte Abbildung

Dem Bottom-up-Ansatz der letzten Abschnitte folgend lernen wir in diesem Abschnitt das Äquivalent zum Adjungieren von Matrizen auf der Ebene von linearen Abbildungen kennen. Dazu erinnern wir uns an die außergewöhnliche Beziehung eines Prähilbertraums zu seinem Dualraum.

Erinnerung VIII.4.1 (Riesz'sche Isomorphismus): Seien V ein endlichdimensionaler Prähilbertraum und für v in V sei $\Theta_v\colon V \to \mathbb{K}$, $w \mapsto \langle w, v\rangle$. Dann ist $\Theta\colon V \to V^*$, $v \mapsto \Theta_v$ ein antilinearer Monomorphismus und damit Isomorphismus. Stattet man V^* mit dem dualen Skalarprodukt $\langle \alpha, \beta\rangle_* = \langle \Theta^{-1}(\beta), \Theta^{-1}(\alpha)\rangle$ aus, dann ist Θ sogar eine Isometrie.

Die Injektivität von Θ folgt direkt aus der Anisotropie des Skalarprodukts, welche liefert: Sind v_1, v_2 aus V und gilt für alle w aus V, dass $\langle v_1, w\rangle = \langle v_2, w\rangle$, dann ist bereits $v_1 = v_2$.

Bemerkung VIII.4.2 (Eigenschaft der adjungierten Matrix): Seien x in $\mathbb{K}^m$, y in $\mathbb{K}^n$ und A in $\mathbb{K}^{n\times m}$. Mit den jeweiligen Standard skalarprodukten auf $\mathbb{K}^n$ und $\mathbb{K}^m$ gilt die Gleichungskette

$$\langle Ax, y\rangle = (Ax)^t\overline{y} = x^tA^t\overline{y} = x^t\overline{A}^*\overline{y} = x^t\overline{(A^*y)} = \langle x, A^*y\rangle.$$

Ausgehend von dieser Gleichung suchen wir zu einer gegebenen linearen Abbildung zwischen Prähilberträumen eine neue lineare Abbildung, die das obige leistet.

Definition VIII.4.3 (Adjungierte Abbildung): Seien V und W Prähilberträume und $\phi\colon V \to W$ linear. Gibt es eine lineare Abbildung $\psi\colon W \to V$, sodass für alle v in V und w in W gilt, dass $\langle \phi(v), w\rangle_W = \langle v, \psi(w)\rangle_V$, dann heißt ψ die *zu ϕ adjungierte Abbildung*. Falls ψ existiert, so ist ψ eindeutig.

Proposition VIII.4.4 (Existenz der adjungierten Abbildung): *Seien V und W endlichdimensionale Prähilberträume und $\phi\colon V \to W$ linear. Dann existiert die adjungierte Abbildung $\phi^*\colon W \to V$.*

Beweis: Wir möchten eine lineare Abbildung $\psi\colon W \to V$ mit den geforderten Eigenschaften konstruieren. Dazu verwenden wir Proposition VIII.4.1. Sei also w in W gegeben. Dann ist $\alpha_w\colon V \to \mathbb{K}$, $v \mapsto \langle \phi(v), w\rangle$ linear. Es gibt also genau ein Element $\psi_w = \Theta^{-1}(\alpha_w)$ in V, sodass

$$\alpha_w(v) = \langle \phi(v), w\rangle_W = \langle v, \psi_w\rangle_V = \Theta_{\psi_w}(v).$$

Die Zuordnung $w \mapsto \psi_w$ ist linear, da Θ und die beteiligten Skalarprodukte in der zweiten Komponente antilinear sind. □

Für die Existenz der Adjungierten haben wir nur Proposition VIII.4.1 gebraucht, d. h. für lineare Abbildungen zwischen Prähilberträumen ohne die Dimensionseinschränkung, für die Proposition VIII.4.1 ebenfalls gilt, gibt es auch adjungierte Abbildungen.

Proposition VIII.4.5 (Adjungierte und duale Abbildung): *Seien V und W endlichdimensionale Prähilberträume mit Identifikationen $\Theta_V\colon V \to V^*$ respektive $\Theta_W\colon W \to W^*$ und $\phi\colon V \to W$ linear. Bezeichnet $\phi^\vee\colon W^* \to V^*$ die duale Abbildung von ϕ, dann kommutiert folgendes Diagramm:*

$$\begin{array}{ccc} W & \xrightarrow{\phi^*} & V \\ {\scriptstyle\Theta_W}\downarrow & & \downarrow{\scriptstyle\Theta_V} \\ W^* & \xrightarrow[\phi^\vee]{} & V^* \end{array}$$

Beweis: Damit das obige Diagramm kommutiert, müssen die Abbildungen $\Theta_V \circ \phi^*, \phi^\vee \circ \Theta_W\colon W \to V^*$ übereinstimmen. Jedem w in W muss also jeweils dieselbe Linearform auf V zugeordnet werden, d. h. wir müssen punktweise auf V überprüfen. Sei w in W gegeben. Für jedes v aus V haben wir dann

$$\begin{aligned} \Big((\Theta_V \circ \phi^*)(w)\Big)(v) &= \Big(\Theta_V(\phi^*(w))\Big)(v) \\ &= \langle v, \phi^*(w)\rangle \\ &= \langle \phi(v), w\rangle = \Theta_W(w)(\phi(v)) = \Big(\phi^\vee \circ \Theta_W(w)\Big)(v). \qquad \square \end{aligned}$$

Die obige Proposition rechtfertigt die Doppelbelegung von „*“ für duale und adjungierte Abbildung, selbst wenn das beim Erstkontakt einen potentiellen Stolperstein darstellt.

Definition VIII.4.6 (Selbstadjungiert, Normal): Seien V ein Prähilbertraum, $\phi\colon V \to V$ linear und $\phi^*\colon V \to V$ existiere. Falls $\phi^* = \phi$, dann heißt ϕ *selbstadjungiert*. Falls $\phi^* \circ \phi = \phi \circ \phi^*$, dann heißt ϕ *normal*.

Definition VIII.4.7 (Orthogonal bzw. Unitär): Seien V und W endlichdimemsionale Prähilberträume und $\phi\colon V \to W$ linear. Gilt $\phi^* \circ \phi = \mathrm{id}_V$ und $\phi \circ \phi^* = \mathrm{id}_W$, so heißt ϕ *orthogonal* beziehungsweise *unitär*.

Bemerkung VIII.4.8: Ist $\phi\colon V \to W$ orthogonal beziehungsweise unitär, dann gilt für alle v und w in V, dass $\langle v, w\rangle_V = \langle \phi^*(\phi(v)), w\rangle_V = \langle \phi(v), \phi(w)\rangle_W$. Insbesondere ist ϕ eine Isometrie. Außerdem ist ϕ als invertierbare lineare Abbildung mit Inverser $\phi^{-1} = \phi^*$ surjektiv, und ϕ^* ist ebenfalls eine Isometrie.

Schließlich bedeutet das, dass $D_B\colon V \to \mathbb{K}^n$ orthogonal beziehungsweise unitär ist genau dann, wenn B eine Orthonormalbasis ist.

Korollar VIII.4.9: *Seien V und W endlichdimensionale Prähilberträume mit geordneten Basen $B = (b_1, \dots, b_m)$ und $C = (c_1, \dots, c_n)$, und $\phi\colon V \to W$ sei linear. Wenn B und C Orthonormalbasen sind, dann gilt $D_{C,B}(\phi)^* = D_{B,C}(\phi^*)$.*

Beweis: Wir sind in der Situation

$$\begin{array}{ccccc} V & \xrightarrow{\phi} & W & \xrightarrow{\psi} & V \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} & & \downarrow{\scriptstyle D_B} \\ \mathbb{K}^m & \xrightarrow[D_{C,B}(\phi)]{} & \mathbb{K}^n & \xrightarrow[D_{C,B}(\phi)^*]{} & \mathbb{K}^m \end{array}$$

und wollen uns davon überzeugen, dass dasjenige ψ, welches das obige Diagramm kommutativ macht, ϕ^* ist. Das bedeutet $\psi = D_B^{-1} \circ (v \mapsto D_{C,B}(\phi)^* v) \circ D_C$. Wir bezeichnen mit $\langle \cdot, \cdot \rangle_E$ beziehungsweise $\langle \cdot, \cdot \rangle_{E'}$ die Standardskalarprodukte auf $\mathbb{K}^m$ beziehungsweise $\mathbb{K}^n$. Dann haben wir

$$\begin{aligned} \langle \phi(v), w \rangle_W &= \langle D_C(\phi(v)), D_C(w) \rangle_{E'} \\ &= \langle D_{C,B}(\phi) D_B(v), D_C(w) \rangle_{E'} \\ &= \langle D_B(v), D_{C,B}(\phi)^* D_B(v) \rangle_E = D_B(v), D_B(\psi(w)) \rangle_E = \langle v, \psi(w) \rangle_V. \end{aligned}$$

sodass $D_{C,B}(\phi)^*$ wirklich $D_{B,C}(\phi^*)$. □

Korollar VIII.4.10 (Endomorphismen- und Matrixbegriffe): *Seien V ein endlichdimensionaler Prähilbertraum, $\phi\colon V \to V$ linear und B eine Orthonormalbasis von V.*

(i) *Genau dann ist ϕ selbstadjungiert, wenn $D_{B,B}(\phi) = D_{B,B}(\phi)^*$, d. h. wenn die Darstellungsmatrix symmetrisch respektive hermitesch ist.*

(ii) *Genau dann ist ϕ normal, wenn $D_{B,B}(\phi)$ normal ist.*

(iii) *Genau dann ist ϕ orthogonal beziehungsweise unitär, wenn $D_{B,B}(\phi)$ orthogonal beziehungsweise unitär ist.*

5. Beweis der Spektralsätze

In diesem Abschnitt möchten wir zeigen, dass normale Matrizen per unitärem Basiswechsel diagonalisiert werden können, d. h. dass es eine zugehörige Orthonormalbasis aus Eigenvektoren gibt.

Das tun wir, indem wir eine viel stärkere Aussage zeigen, nämlich: In einer $\mathbb{C}$-Unteralgebra von $\mathbb{C}^{n\times n}$, die gewisse Eigenschaften erfüllt, sind alle Matrizen gleichzeitig diagonalisierbar.

Um diese Aussage zu zeigen, verschaffen wir uns geeignete invariante Zerlegungen.

Proposition VIII.5.1 (Invariante Unterräume für kommutierende Matrizen): *Seien A und B zwei $n \times n$-Matrizen, sodass $AB = BA$.*

(i) *Sind die Einträge von A und B aus einem beliebigen Körper K, dann sind die Eigenräume von A Invariant unter B, d. h., ist λ ein Eigenwert von A, dann ist $B\,\mathrm{Eig}(A,\lambda) \subseteq \mathrm{Eig}(A,\lambda)$.*

(ii) *Sind die Einträge von A und B komplexe Zahlen, dann sind die orthogonalen Komplemente der Eigenräume von A invariant unter B^*, d. h. für einen Eigenwert λ von A gilt $B^*\,\mathrm{Eig}(A,\lambda)^\perp \subseteq \mathrm{Eig}(A,\lambda)$.*
Insbesondere erhalten wir für eine normale Matrix A eine invariante Zerlegung $\mathbb{C}^n = \mathrm{Eig}(A,\lambda) \oplus \mathrm{Eig}(A,\lambda)^\perp$.

In Teil (ii) der obigen Proposition ist $\mathbb{C}^n$ mit dem Standardskalarprodukt ausgestattet.

Beweis: (i) Ist x ein Eigenvektor von A zum Eigenwert λ, dann haben wir $ABx = BAx = B\lambda x = \lambda Bx$, d. h. Bx liegt im Eigenraum $\mathrm{Eig}(A,\lambda)$.

(ii) Sind x ein Element von $\mathrm{Eig}(A,\lambda)^\perp$ und y in $\mathrm{Eig}(A,\lambda)$, dann gilt

$$\langle B^*x, y\rangle = (B^*x)^t\overline{y} = x^t\overline{B}\overline{y} = \langle x, By\rangle = 0,$$

weil ja By nach (i) ein Eigenvektor von A zum Eigenwert λ ist, und x in $\mathrm{Eig}(A,\lambda)^\perp$ liegt.
Der Zusatz folgt, da A sowohl mit sich selbst als auch (nach Voraussetzung) mit A^* kommutiert. □

Wie eingangs erwähnt ist $\mathbb{C}^{n\times n}$ zusammen mit Matrizenaddition, Matrizenmultiplikation und Skalarmultiplikation eine $\mathbb{C}$-Algebra. Zu zwei kommutierenden Matrizen A und B suchen wir nun die kleinste $\mathbb{C}$-Unteralgebra U, die A und B enthält. Jedenfalls müssen Ausdrücke der Form $\sum_{i,j} x_{i,j}A^iB^j$ in U enthalten sein.

Proposition VIII.5.2 (Von kommutierenden Matrizen erzeugte Algebra): *Seien K ein Körper und A, B kommutierende $n \times n$-Matrizen mit Einträgen aus K. Dann ist*

$$\mathfrak{A} = \left\{ \sum_{i=0}^{m}\sum_{j=0}^{k} c_{i,j}A^iB^j : m,k \in \mathbb{N}, c_{i,j} \in K \right\}$$

eine kommutative K-Unteralgebra von $K^{n\times n}$. Wir schreiben $K[A,B] = \mathfrak{A}$ und nennen $K[A,B]$ die von A und B erzeugte K-Unteralgebra.

Beweis: Man sieht sofort, dass $\mathfrak{A}$ ein Untervektorraum ist. Weil A und B vertauschen, haben wir $(A^iB^j)(A^rB^s) = A^{i+r}B^{j+s} = (A^rB^s)(A^iB^j)$, d. h. Produkte solcher Elemente bleiben in $\mathfrak{A}$ und die Reihenfolge der Faktoren spielt keine Rolle. Weil die Elemente von $\mathfrak{A}$ Linearkombinationen solcher Elemente sind, ist $\mathfrak{A}$ unter Multiplikation abgeschlossen und kommutativ. □

Satz 35 (Simultane Orthonormalbasis): *Sei $\mathfrak{A} \subseteq \mathbb{C}^{n\times n}$ eine kommutative Unteralgebra, die unter Adjunktion abgeschlossen ist (d. h. für A aus $\mathfrak{A}$ ist auch A^* in $\mathfrak{A}$ enthalten). Dann gibt es eine Orthonormalbasis $B = \{b_1, \ldots, b_n\}$ des $\mathbb{C}^n$ bezüglich des Standardskalarprodukts von simultanen Eigenvektoren, d. h. für jedes $A \in \mathfrak{A}$ sind $b_1, \ldots, b_n$ Eigenvektoren von A.*

Beweis: Wir zeigen die Aussage per Induktion nach der Dimension n. Für $n = 1$ ist nichts zu zeigen, die kanonische Basis $B = \{1\}$ leistet offensichtlich das Gewünschte.

Die Aussage gelte nun für eine natürliche Zahl $n - 1$ größergleich Eins. Ist $\mathfrak{A} = \mathbb{C}I_n = \{\mathrm{diag}(\lambda, \ldots, \lambda) \mid \lambda \in \mathbb{C}\}$ oder $\mathfrak{A} = \{\mathbf{0}\}$, dann gilt die Behauptung mit der Standardbasis.

Ist $\mathfrak{A}$ keine der beiden oben genannten Algebren, dann gibt es $A_0 \in \mathfrak{A} - \mathbb{C}I_n$. Weil $\mathbb{C}$ algebraisch abgeschlossen ist, hat A_0 einen Eigenwert λ und $\mathbb{C}^n$ zerfällt in $\mathbb{C}^n = \mathrm{Eig}(A, \lambda) \oplus \mathrm{Eig}(A, \lambda)^\perp$. Weil irgendeine andere Matrix $B \in \mathfrak{A}$ mit A kommutiert, ist wegen Proposition VIII.5.1 auch $B\,\mathrm{Eig}(A, \lambda) \subseteq \mathrm{Eig}(A, \lambda)$ sowie $B^*\,\mathrm{Eig}(A, \lambda)^\perp \subseteq \mathrm{Eig}(A, \lambda)^\perp$, d. h. sowohl $\mathrm{Eig}(A, \lambda)$ als auch $\mathrm{Eig}(A, \lambda)^\perp$ sind invariant unter allen Elementen von $\mathfrak{A}$.

Da $\dim \mathrm{Eig}(A, \lambda) > 0$ können wir die Induktionsvoraussetzung auf $\mathrm{Eig}(A, \lambda)$ und $\mathrm{Eig}(A, \lambda)^\perp$ anwenden und erhalten Orthonormalbasen B_1 von $\mathrm{Eig}(A, \lambda)$ sowie B_2 von $\mathrm{Eig}(A, \lambda)^\perp$ aus simultanen Eigenvektoren für $\mathfrak{A}$. Die Vereinigung $B := B_1 \cup B_2$ liefert eine Orthonormalbasis von $\mathbb{C}^n$ aus simultanen Eigenvektoren von $\mathfrak{A}$. □

Bemerkung VIII.5.3: Ist $\mathfrak{A}$ eine kommutative $\mathbb{R}$-Unteralgebra von $\mathbb{R}^{n\times n}$ und haben alle Matrizen in $\mathfrak{A}$ ausschließlich reelle Eigenwerte, dann gibt es sogar eine Orthonormalbasis des $\mathbb{R}^n$ aus simultanen Eigenvektoren.

Das zeigt man genau wie in Satz 35. Man verwendet, dass alle Matrizen über $\mathbb{C}$ diagonalisierbar sind und der Eigenwert von A_0 nach Voraussetzung in $\mathbb{R}$ liegt.

Beweis: „$\Longrightarrow$“ Sei A normal, d. h. $AA^* = A^*A$, und sei $\mathfrak{A} = \mathbb{C}[A, A^*]$ die von A und A^* erzeugte $\mathbb{C}$-Unteralgebra von $\mathbb{C}^{n\times n}$ aus Proposition VIII.5.2. Dann erfüllt $\mathfrak{A}$ die Voraussetzungen von Satz 35, und wir erhalten eine Orthonormalbasis B bestehend aus Eigenvektoren von A. Nach Proposition VIII.3.6 ist

die Basiswechselmatrix $S = D_{E,B} = D_{E,B}(\mathrm{id})$ unitär und für die Abbildung $\phi_A \colon x \mapsto Ax$ gilt

$$A = D_{E,E}(\phi_A) = D_{E,B} D_{B,B}(\phi_A) D_{B,E} = S \operatorname{diag}(\lambda_1, \ldots, \lambda_n) S^{-1},$$

wobei $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A sind.

„$\Longleftarrow$“: Seien S eine unitäre $n \times n$-Matrix und $\lambda_1, \ldots, \lambda_n$ komplexe Zahlen, sodass $S^* A S = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Dann ist

$$\begin{aligned} S^* A A^* S &= S^* A S S^* A^* S \\ &= \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \operatorname{diag}(\bar{\lambda}_1, \ldots, \bar{\lambda}_n) = (S^* A^* S)(S^* A S) = S^* A^* A S, \end{aligned}$$

d. h. A ist in der Tat normal. □

Bemerkung VIII.5.4 (Diagonalisierbarkeit bei reellen Eigenwerten): Sei A eine reelle symmetrische $n \times n$-Matrix und alle Eigenwerte von A seien reell. Dann gibt es eine Orthonormalbasis des $\mathbb{R}^n$ bestehend aus Eigenvektoren von A.

Beweis: Sei $\mathfrak{A} = \mathbb{R}[A, A^t]$. Nach Satz 35 gibt es eine Orthonormalbasis B des $\mathbb{C}^n$ aus simultanen Eigenvektoren, d. h. es gibt eine unitäre Matrix S, sodass

$$\mathfrak{A}^S = \{SBS^{-1} \mid B \in \mathfrak{A}\} \subseteq \{\operatorname{diag}(\alpha_1, \ldots, \alpha_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{C}\}.$$

Die Matrizen $A_1 = SAS^{-1}$ und $A_2 = A_1^t = A_1^* = SA^tS^{-1}$ haben reelle Einträge, und $\mathfrak{A}^S = \mathbb{R}[A_1, A_2]$. Das bedeutet, dass $\mathfrak{A}^S$ in $\mathbb{R}^{n \times n}$ enthalten ist, also $\mathfrak{A}^S \subseteq \{\operatorname{diag}(\alpha_1, \ldots, \alpha_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$. Mit Proposition VIII.5.3 folgt die Behauptung. □

Satz 36 (Hauptachsentransformation für Endomorphismen): *Seien $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler euklidischer oder unitärer Vektorraum und ϕ ein Endomorphismus von V. Ist ϕ selbstadjungiert, dann gibt es eine Orthonormalbasis von V bestehend aus Eigenvektoren von ϕ und alle Eigenwerte sind reell.*

Proposition VIII.5.5: *Eigenwerte selbstadjungierter Endomorphismen sind reell.*

Beweis: Seien ϕ ein selbstadjungierter Endomorphismus eines euklidischen oder unitären Vektorraums, λ ein Eigenwert von ϕ und v ein von Null verschiedener Eigenvektor zum Eigenwert λ. Dann haben wir

$$\lambda \|v\|^2 = \lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle \phi(v), v \rangle = \langle v, \phi(v) \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \|v\|^2,$$

und da $\|v\|$ per Voraussetzung nicht Null ist, muss λ reell sein. □

Bemerkung VIII.5.6 (Eigenwerte orthogonaler Matrizen): Sei A eine orthogonale Matrix. Ist λ Eigenwert zum Eigenvektor v, dann ist $\overline{\lambda}$ Eigenwert zum Eigenvektor $\overline{v}$, denn $A\overline{v} = \overline{Av} = \overline{\lambda v} = \overline{\lambda} \cdot \overline{v}$.

Satz 37 (Isometrienormalform für Endomorphismen): *Seien $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler euklidischer oder unitärer Raum und ϕ ein orthogonaler oder unitärer Endomorphismus von V.*

(i) *Ist V unitär, dann gibt es eine Orthonormalbasis von V aus Eigenvektoren von ϕ. Für die Eigenwerte λ_i von ϕ gilt $|\lambda_i| = 1$.*

(ii) *Ist V euklidisch, dann gibt es eine Orthonormalbasis B von V, sodass*

$$D_{B,B}(\phi) = \operatorname{diag}(1, \dots, 1, -1, \dots, -1, A_1, \dots, A_k)$$

mit Drehmatrizen $A_i = \left(\begin{smallmatrix} \cos(\alpha_i) & -\sin(\alpha_i) \\ \sin(\alpha_i) & \cos(\alpha_i) \end{smallmatrix}\right)$ und reellen Zahlen α_i.

Beispiel VIII.5.7 (Der Spektralsatz in Dimension 2): Seien $V = \mathbb{R}^2$ ausgestattet mit dem Standardskalarprodukt und $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch $x \mapsto Ax$ für eine orthogonale Matrix $A \in \mathrm{O}(2) \subseteq \mathrm{U}(n)$. Aufgefasse als komplexe Matrix liefert **??** erhalten wir eine Orthonormalbasis $B = (b_1, b_2)$ des $\mathbb{C}^2$ bestehend aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \lambda_2 \in \mathbb{C}$, die beide komplexen Betrag 1 haben.

Bezeichne $\phi'\colon \mathbb{C}^2 \to \mathbb{C}^2$ die Abbildung $x \mapsto Ax$. Bezüglich B ist die Darstellungsmatrix von ϕ' sehr einfach, nämlich $D_{B,B}(\phi') = \operatorname{diag}(\lambda_1, \lambda_2)$, und außerdem ist $\phi'|_{\mathbb{R}^2} = \phi$, wobei wir $\mathbb{R}^2$ mit einem Unterraum des $\mathbb{C}^2$ identifizieren.

Gehört λ_1 zu $\mathbb{C}-\mathbb{R}$, dann ist $A\bar{b_1} = \overline{Ab_1} = \overline{\lambda_1 b_1} = \bar{\lambda}_1\bar{b}_1$, d. h. $\bar{b}_1$ ist Eigenvektor zum Eigenwert $\bar{\lambda}_1 \neq \lambda_1$. Weil A nur zwei Eigenwerte hat, muss das heißen, dass $\lambda_2 = \bar{\lambda}_1$; wir dürfen also annehmen dass $\bar{b}_1 = b_2$. Wir definieren

$$c_1 := \operatorname{Re}(b_1) = \frac{1}{2}(b_1 + \bar{b}_1) = \frac{1}{2}(b_1 + b_2), \qquad c_2 := \operatorname{Im}(b_1) = \frac{1}{2\mathrm{i}}(b_1 - \bar{b}_1) = \frac{1}{2\mathrm{i}}(b_1 - b_2).$$

Dann ist $C = (\sqrt{2}c_1, \sqrt{2}c_2)$ eine Orthonormalbasis des $\mathbb{C}^2$ und

$$D_{B,C} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix}, \qquad D_{C,B} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix}.$$

Außerdem ist C sogar eine Orthonormalbasis des $\mathbb{R}^2$, d. h. die Einträge von c_1

und c_2 sind reell. Wir haben

$$\begin{aligned} D_{C,C}(\phi') &= D_{C,B} D_{B,B}(\phi') D_{B,C} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \bar{\lambda}_1 \end{pmatrix} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \lambda_1 & \bar{\lambda}_1 \\ \mathrm{i}\lambda_1 & -\mathrm{i}\bar{\lambda}_1 \end{pmatrix} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix} \\ &= \begin{pmatrix} 2\operatorname{Re}(\lambda_1) & 2\operatorname{Im}(\lambda_1) \\ -2\operatorname{Im}(\lambda_1) & 2\operatorname{Re}(\lambda_1) \end{pmatrix}. \end{aligned}$$

Wegen $|\lambda_1| = 1$ ist auch $|\lambda_1|^2 = \operatorname{Re}(\lambda_1)^2 + \operatorname{Im}(\lambda_1)^2 = 1$, d. h. es gibt $\varphi \in \mathbb{R}$ mit $\cos(\varphi) = \operatorname{Re}(\lambda_1)$ und $\sin(\varphi) = -\operatorname{Im}(\varphi)$. Wir können $D_{C,C}(\phi')$ also schreiben als

$$D_{C,C}(\phi') = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Da C eine Orthonormalbasis des $\mathbb{R}^2$ ist, gilt außerdem $D_{C,C}(\phi) = D_{C,C}(\phi')$.

Gehört λ_1 zu $\mathbb{R}$, dann ist auch λ_2 eine reelle Zahl, da $\det A = \lambda_1\lambda_2$ und wegen $|\lambda_1| = |\lambda_2| = 1$ haben wir $\lambda_1, \lambda_2 \in \{\pm 1\}$. Nach **??** gibt es eine Orthonormalbasis C des $\mathbb{R}^2$, sodass

$$D_{C,C}(\phi) \in \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}.$$

Beweis (von Satz 37): (i) Folgt direkt aus dem Spektralsatz **??**.

(ii) Wir wählen eine Orthonormalbasis von V und identifizieren V mit $\mathbb{R}^n$ vermöge D_B, wir dürfen also ohne Einschränkung annehmen, dass $V = \mathbb{R}^n$ versehen mit dem Standardskalarprodukt und dass $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ die Abbildung $x \mapsto Ax$ mit $A \in \mathrm{O}(n)$ ist. Nach **??** gibt es eine Orthonormalbasis $B' = (b'_1, \dots, b'_n)$ des $\mathbb{C}^n$ bestehend aus Eigenvektoren von A, sodass b_i Eigenvektor von A zum Eigenwert λ_i ist.

Wir sortieren $b'_1, \dots, b'_n$ so um, dass $\lambda_1 = \dots = \lambda_r = 1$, $\lambda_{r+1} = \dots = \lambda_{r+s} = -1$ und $\lambda_{s+1}, \dots, \lambda_n \in \mathbb{C} - \mathbb{R}$. Wie im vorangegangenen Beispiel definieren wir $\phi'\colon \mathbb{C}^n \to \mathbb{C}^n$ durch $x \mapsto Ax$. Wie im vorangegangenen Beispiel erhalten wir: Ist v ein Eigenvektor von ϕ' zum Eigenwert λ, dann ist $\bar{v}$ ein Eigenvektor von ϕ' zum Eigenwert $\bar{\lambda}$.

Durch Umsortierung können wir erreichen, dass $\lambda_{s+1} = \bar{\lambda}_{s+2}, \dots, \lambda_{n-1} = \bar{\lambda}_n$ und $b_{s+1} = \bar{b}'_{s+2}, \dots, b'_{n-1} = \bar{b}'_n$.

Verfahren wir für $U := \operatorname{Lin}(b'_{s+2i-1}, b'_{s+2i})$ wie im vorangegangenen Beispiel, d. h. setzen wir

$$c_{s+2i-1} := \sqrt{2}\operatorname{Re}(b'_{s+2i-1}), \qquad c_{s+2i} := \sqrt{2}\operatorname{Im}(b'_{s+2i-1}),$$

dann ist (c_{s+2i-1}, c_{s+2i}) eine Orthonormalbasis von U und durch Zusammensetzung all dieser Basen zu $C := (b_1, \ldots, b_s, c_{sn}, \ldots, c_{n-1}, c_n)$ erhalten wir eine Orthonormalbasis des $\mathbb{R}^n$, sodass $D_{B,B}(\phi)$ die angegebene Form hat. □

Den Rest des Abschnittes möchten wir schönen Anwendungen der Hauptachsentransformation für Bilinearformen widmen. Im Folgenden bezeichne stets $h\colon V \times V \to \mathbb{K}$ eine Bilinearform beziehungsweise Sesquilinearform, und zu einer gewählten Basis B von V sei stets $G = G_B(h)$ die Gram-Matrix von h bezüglich B. Wir erinnern daran, dass die Gram-Matrix von h bezüglich einer anderen Basis C über

$$G_C(h) = D_{B,C}^t G_B(h) \overline{D}_{B,C}$$

zusammenhängt.

Definition VIII.5.8 (Definitheit): Gilt für alle v aus $V - \{0\}$, dass $h(v,v) > 0$ respektive $h(v,v) < 0$, so heißt h *positiv definit* respektive *negativ definit*. Sind die Ungleichungen nicht strikt, dann spricht man von *positiver Semidefinitheit* respektive *negativer Semidefinitheit*. Ist h weder positiv semidefinit noch negativ semidefinit, d. h. gibt es v und w aus $V-\{0\}$, sodass $h(v,v) > 0$ und $h(w,w) < 0$, dann heißt h *indefinit*.

Beispiel VIII.5.9: Angenommen, die Gram-Matrix G von h bezüglich B ist eine Diagonalmatrix. Genau dann ist h positiv definit respektive negativ definit, wenn alle Diagonaleinträge positiv respektive negativ sind. Gibt es einen positiven und einen negativen Diagonaleintrag, dann ist h indefinit.

Bemerkung VIII.5.10: Wegen der Transformationseigenschaft der Gram-Matrix, an die wir oben erinnert haben, gilt: G ist positiv definit bzw. negativ definit bzw. indefinit genau dann, wenn $A^t G \overline{A}$ für jedes A aus $\mathrm{Gl}_n(\mathbb{K})$ positiv definit bzw. negativ definit bzw. indefinit ist.

Korollar VIII.5.11 (Eigenwerte und Definitheit): *Sei A eine symmetrische beziehungsweise hermitesche Matrix. Dann gilt:*

(i) *Genau dann ist A positiv definit, wenn* $\mathrm{Spec}(A) \subseteq (0, \infty)$.

(ii) *Genau dann ist A negativ definit, wenn* $\mathrm{Spec}(A) \subseteq (-\infty, 0)$.

(iii) *Genau dann ist A indefinit, wenn es einen positiven und einen negativen Eigenwert gibt.*

Beweis: Nach dem Satz über die Hauptachsentransformation ist A als symmetrische respektive hermitesche Matrix diagonalisierbar, sodass wir die Behauptung aus Proposition VIII.5.10 und Proposition VIII.5.9 ablesen können. □

Definition VIII.5.12 (Definitheit für Matrizen): Sei A eine symmetrische beziehungsweise hermitesche Matrix in $\mathbb{K}^{n\times n}$. Für jedes x aus $\mathbb{K}^n$ ist dann $x^tA\overline{x}$ eine reelle Zahl.

(i) Gilt für alle x aus $\mathbb{K}^n - \{0\}$, dass $x^tA\overline{x} > 0$, so heißt A *positiv definit*.

(ii) Gilt für alle x aus $\mathbb{K}^n - \{0\}$, dass $x^tA\overline{x} < 0$, so heißt A *negativ definit*.

(iii) Gibt es x und y in $\mathbb{K}^n - \{0\}$, sodass $x^tA\overline{x} > 0$ und $y^tA\overline{y} < 0$, dann heißt A indefinit.

Kapitel IX.

Etwas mehr Strukturmathematik

1. Gruppenaktionen

Wie oft in der Mathematik kann man ein tieferes Verständnis für Objekte gewinne, indem man passende Abbildungen zwischen ihnen betrachtet. So kann man Gruppen durch Homomorphismen in spezielle Symmetriegruppen besser verstehen. Das ist das sogenannte Konzept einer *Gruppenaktion*. Wir werden diesen Abschnitt darauf verwenden, zu verstehen, wieso der einleitende Satz richtig ist.

Definition IX.1.1: Seien G eine Gruppe, X eine nichtleere Menge und $\alpha\colon G \times X \to X$ eine Abbildung. Falls für alle x aus X gilt, dass $\alpha(e_G, x) = x$, und falls für alle g, h aus G sowie x aus X gilt, dass $\alpha(\alpha(g,h),x) = \alpha(g,\alpha(h,x))$, dann heißt α eine *Linksaktion von G auf X*. Man sagt auch *Gruppenoperation* oder *Gruppenwirkung*.

Notation IX.1.2: Sei $\alpha\colon G \times X \to X$ eine Gruppenaktion. Üblicherweise schreibt man $g \bullet x$ oder sogar gx, falls die Aktion aus dem Kontext klar ist, anstelle von $\alpha(g,x)$. Dadurch erhält die zweite Bedingung an die Gruppenaktion die leichter verdauliche Gestallt $(gh)x = g(hx)$.

Bemerkung IX.1.3: Sei $\alpha\colon G\times X \to X$ eine Gruppenaktion. Sind x, y Elemente von X und gilt $gx = y$, dann ist bereits $x = g^{-1}y$.

Beispiel IX.1.4 (für Gruppenaktionen): (i) Die Abbildung $\alpha_1\colon \mathrm{Gl}(n,K) \times K^n \to K^n$, $(A,v) \mapsto Av$ erklärt eine Gruppenaktion, da $I_n v = v$ für jedes v aus K^n, und da Matrizenmultiplikation assoziativ ist.

(ii) $\alpha_2\colon S_n \times \{1,\ldots,n\} \to \{1,\ldots,n\}$, $(\sigma,k) \mapsto \sigma(k)$ ist eine Gruppenaktion.

(iii) Für die Gruppe $\mathbb{Q}$ und ihre Untergruppe $\mathbb{Z}$ definiert $(k, q) \mapsto k + q$ eine Gruppenaktion. Ist allgemeiner G eine Gruppe, und ist U eine Untergruppe von G, dann agiert U via Linksmultiplikation auf G:

$$\alpha_3 \colon U \times G \longrightarrow G, \qquad (u, x) \longmapsto ux.$$

Es ist nämlich $\alpha_3(e_G, x) = e_G x = x$ und $\alpha_3(u_1 u_2, x) = u_1(u_2 x) = \alpha_3(u_1, \alpha_3(u_2, x))$.

(iv) Die Gruppe $\mathbb{Z}$ agiert auf $\mathbb{Z}/n\mathbb{Z}$ durch

$$\alpha_4 \colon \mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, \qquad (k, [r]) \longmapsto [k + r].$$

(v) Permutation der Koordinaten erklärt eine Aktion auf dem K^n, genauer:

$$\alpha_5 \colon S_n \times K^n \longrightarrow K^n, \qquad (\sigma, (v_1, \ldots, v_n)^t) \longmapsto (v_{\sigma(1)}, \ldots, v_{\sigma(n)})^t$$

ist eine Gruppenaktion. Klarerweise ist $\mathrm{id} \bullet v = v$ für jedes v aus K^n, und sind σ, τ zwei Permutationen, dann gilt

$$(\sigma \circ \tau) \bullet \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_{\sigma\circ\tau(1)} \\ \vdots \\ v_{\sigma\circ\tau(n)} \end{pmatrix} = \sigma \bullet \begin{pmatrix} v_{\tau(1)} \\ \vdots \\ v_{\tau(n)} \end{pmatrix}.$$

(vi) Seien V ein $\mathbb{R}$-Vektorraum und w ein von Null verschiedener Vektor in V. Dann ist

$$\alpha_w : \mathbb{R} \times V \longrightarrow V, \qquad (\lambda, v) \longmapsto \lambda w + v$$

eine Gruppenaktion.

Proposition IX.1.5: *Sei $\alpha \colon G \times X \to X$ eine Gruppenaktion. Dann wird durch „$x \sim y$, falls es g in G mit $gx = y$ gibt“ eine Äquivalenzrelation auf X erklärt.*

Beweis: Da für jedes x aus X gilt, dass $e_G x = x$, steht jedes x zu sich selbst in Relation.

Zur Symmetrie: Seien x und y Elemente von X, sodass $x \sim y$. Dann gibt es irgendein Gruppenelement g, sodass $gx = y$. Damit gilt $g^{-1}y = g^{-1}(gx) = (g^{-1}g)x$, sodass auch $y \sim x$.

Zur Transitivität: Seien x, y und z Elemente von X mit $x \sim y$ und $y \sim z$. Das heißt es gibt Gruppenelemente g und h, sodass $y = gx$ und $z = hy$. Weil G auf X agiert, heißt das $z = hy = h(gx) = (hg)x$, also auch $x \sim z$. □

Definition IX.1.6: Sei $\alpha \colon G \times X \to X$ eine Gruppenaktion und „$\sim$“ die zugehörige Äquivalenzrelation.

(i) Für ein Element x von X heißt $Gx = [x]_\sim = \{gx \mid g \in G\}$ die *Bahn von* x.

(ii) Gibt es nur eine Bahn für die Aktion, dann heißt die Aktion *transitiv*.

(iii) Für ein Element x von X heißt $\mathrm{Stab}(x) = \{g \in G \mid gx = x\}$ der *Stabilisator von* x.

(iv) Sei x ein Element von X. Gilt für alle g aus G, dass $gx = x$, dann heißt x ein *Fixpunkt der Aktion*. In diesem Fall ist $Gx = \{x\}$ und $\mathrm{Stab}(x) = G$.

(v) Gibt es für jedes g in $G - \{e_G\}$ ein x aus X mit $gx \neq x$, dann heißt die Operation *treu*.

(vi) Gilt für jedes x in X, dass $\mathrm{Stab}(x) = \{e_G\}$, dann heißt die Aktion *frei*.

Beispiel IX.1.7: (i) Für die Aktion α_1 hat ein Vektor v aus $K^n - \{0\}$ die Bahn $\mathrm{Gl}_n(K)v = K^n - \{0\}$, weil wir jeden von Null verschiedenen Vektor zu einer Basis ergänzen, und dann einen Basiswechsel machen können. Die Bahn des Nullvektors ist die Menge mit dem Nullvektor, da Multiplikation mit einer Matrix eine lineare Abbildung ist.

Für die Aktion α_6 un einen Vektor v ist $\mathbb{R} \bullet_{\alpha_w} v$ eine affine Gerade mit Richtungsvektor w und Stützvektor w.

(ii) Die Gruppenaktion α_1 ist nicht transitiv, denn wir haben zwei Bahnen. Die Aktion α_2 ist transitiv, denn für $1 \leq k \leq n$ gehört die Transposition $(1k)$ zu S_n, sodass $k = (1k) \bullet 1$ in $S_n \bullet 1$ liegt.

(iii) Für die Aktion α_1 ist $\mathrm{Stab}(0) = \mathrm{Gl}_n(K)$ und

$$\mathrm{Stab}(e_1) = \left\{ \begin{pmatrix} 1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} \in \mathrm{Gl}_n(K) \right\}.$$

Für α_2 ist $\mathrm{Stab}(n) \cong S_{n-1}$. Ist für α_6 der Vektor w von Null verschieden, dann gilt $\mathrm{Stab}(v) = \{0\}$. Für die Aktion α_3 ist $\mathrm{Stab}(g) = \{u \in U \mid ug = g\} = \{e_G\}$.

(iv) Für die Aktion α_1 ist Null der einzige Fixpunkt. Die Aktion α_2 hat keinen Fixpunkt. Bei der Aktion α_5 sind alle Fixpunkte von der Form $\{(\lambda, \ldots, \lambda)^t \mid \lambda \in K\}$.

(v) Die Aktion α_1 ist treu. Ist nämlich $Aei = e_i$ für $1 \leq i \leq n$, dann ist $A = I_n$. Die Aktion α_4 hingegen ist nicht treu, denn $n \cdot [k] = [n + k] = k$ für jedes Element $[k]$ von $\mathbb{Z}/n\mathbb{Z}$.

(vi) Die Aktion α_1 ist nicht frei, denn es gibt einen Fixpunkt. Für $n \geq 3$ gilt $(12) \bullet 3 = 3$, sodass α_2 nicht frei ist. Ist bei der Aktion α_6 der Vektor w von Null verschieden, dann ist α_6 frei.

Proposition IX.1.8: *Seien G eine Gruppe, X eine nichtleere Menge und $\mathrm{Sym}(X) = \{\sigma\colon X \to X \text{ bijektiv}\}$ die Symmetriegruppe von X. Dann gibt es eine $1:1$-Korrespondenz zwischen Gruppenaktionen $G \times X \to X$ und Gruppenhomomorphismen $G \to \mathrm{Sym}(X)$:*

Ist $\alpha\colon G \times X \to X$ eine Gruppenaktion, dann ist $\sigma_g = \alpha(g, \cdot)\colon X \to X$, $x \mapsto \alpha(g, x)$ eine bijektive Selbstabbildung von X und $\sigma\colon G \to \mathrm{Sym}(X)$, $g \mapsto \sigma_g$ ist ein Gruppenhomomorphismus.

Ist umgekehrt $\sigma\colon G \to \mathrm{Sym}(X)$ ein Gruppenhomomorphismus, dann liefert $\alpha\colon G \times X \to X$, $(g, x) \mapsto \sigma(g)(x)$ eine Gruppenaktion.

Beweis: Wir zeigen die Bijektivität von σ_g, indem wir nachrechnen, dass $\sigma_{g^{-1}}$ die zugehörige Inverse ist. Für jedes x in X ist nämlich $\sigma_g \circ \sigma_{g^{-1}}(x) = g \bullet (g^{-1}(x)) = (gg^{-1})(x) = x$.

Nun zeigen wir, dass σ ein Gruppenhomomorphismus ist. Zu nächst gilt für jedes x in X, dass $\sigma(e_g)(x) = e_G \bullet x = x$, d. h. $\sigma(e_g) = \sigma_{e_g} = \mathrm{id}_X$.

Sind g und h Elemente von G und ist x ein Element von X, dann haben wir

$$\sigma(gh)(x) = (gh) \bullet x = g \bullet (h \bullet x) = \sigma_g \circ \sigma_h(x) = \sigma(g) \circ \sigma(h)(x),$$

also $\sigma(gh) = \sigma(g) \circ \sigma(h)$. Genau so rechnet man nach, dass die angegebene Art, aus einem Homomorphismus eine Gruppenaktion zu machen, invers zur ersten Konstruktion ist. □

Bemerkung IX.1.9: Ist $\alpha\colon G \times X \to X$ eine Gruppenaktion und ist $\sigma\colon G \to \mathrm{Sym}(X)$ der zugehörige Homomorphismus, dann ist α treu genau dann, wenn σ injektiv ist.

Bemerkung IX.1.10: Sei $\alpha\colon G \times X \to X$ eine Gruppenaktion. Da Bahnen Äquivalenzklassen sind, bilden sie eine Partition von X . Ist X endlich und $Gx_1, \ldots, Gx_k$ die Menge der Bahnen – $x_1, \ldots, x_k$ ist also ein Repräsentantensystem der Bahnen – so gilt

$$\#X = \sum_{i=1}^{k} \#Gx_i.$$

Satz 38 (Bahnformel): *Seien G eine endliche Gruppe und $\alpha\colon G \times X \to X$ eine Gruppenaktion. Dann gilt $\#G = \#\,\mathrm{Stab}_G(x) \#Gx$ für jedes x in X.*

Beweis: Wir betrachten die Äquivalenzrelation „$g \sim h$ genau dann, wenn $g \bullet x = h \bullet x$“ auf G. Genau dann steht g in Relation zu h, wenn $g^{-1}h$ in $\mathrm{Stab}_G(x)$ liegt, also wenn h zu $g \bullet \mathrm{Stab}_G(x) = \{gf \mid f \in \mathrm{Stab}_G(x)\}$ gehört. Nach der ersten Charakterisierung von $g \sim h$ haben wir genau so viele Äquivalenzklassen wie Elemente in der Bahn Gx, und nach der letzten Charakterisierung hat jede Äquivalenzklasse genau $\#\,\mathrm{Stab}_G(x)$ viele Elemente. Das heißt es gilt

$$\#G = \sum \#[g_i] = \sum \#\,\mathrm{Stab}_G(x) = \#\,\mathrm{Stab}_G(x)\#Gx$$

wie behauptet. □

2. Teilbarkeit in Ringen

Das Ziel dieses Abschnittes ist das systematische Studium der Konzepte „Teilbarkeit“ und „größter gemeinsamer Teiler“ um beispielsweise zu erkennen, auf welche Ringe sich das wichtige Lemma von Bézout verallgemeinert.

Definition IX.2.1: Sei $(R, +, \cdot)$ ein kommutativer unitärer Ring.

(i) Seien $a, b \in R$. Gibt es $c \in R$ mit $b = ca$, dann sagen wir *a teile b* und schreiben $a \mid b$.

(ii) Sei $a \in R - \{0\}$. Gibt es $c \in R - \{0\}$ mit $ac = 0$, dann heißt a ein *echter Nullteiler*.

(iii) Gibt es in R keine echten Nullteiler, dann heißt R *nullteilerfrei*. Das heißt: Sind $a, b \in R$ mit $ab = 0$, dann ist $a = 0$ oder $b = 0$. Ein nullteilerfreier kommutativer unitärer Ring heißt auch Integritätsring oder Integritätsbereich.

(iv) Sei $a \in R$. Gibt es $b \in R$ mit $ab = 1$, dann heißt a *invertierbar*. Die Menge $R^\times := \{u \in R \mid u \text{ ist invertierbar}\}$ heißt *Einheitengruppe* oder *multiplikative Gruppe von R* und ist eine Gruppe mit Multiplikation.

Bemerkung IX.2.2 (Kürzungsregel, Einheitenteiler): Sei R ein kommutativer unitärer Ring.

(i) Genau dann ist R nullteilerfrei, wenn für alle $a, c, c' \in R$ mit $a \neq 0$ gilt: Wenn $ac = ac'$, dann ist $c = c'$.

(ii) Teiler von Einheiten sind selbst Einheiten.

Beweis: (i) Ist $ac = ac'$, dann ist $a(c - c') = 0$, d. h. wegen der Nullteilerfreiheit von R ist $c - c' = 0$ und damit $c = c'$.

(ii) Seien $a \in R$, $u \in R^\times$ und a teile u. Dann gibt es $c \in R$ mit $u = ac$ und $u' \in R$ mit $uu' = 1$. Aber dann erhalten wir $a(cu') = uu' = 1$, d. h. a ist eine Einheit. □

Bemerkung IX.2.3 (Teilbarkeit ist fast eine Ordnungsrelation): Seien R ein kommutativer unitärer nullteilerfreier Ring und $a, b, c \in R$. Dann ist Teilbarkeit fast eine Ordnungsrelation auf R:

(i) Jedes $a \in R$ teilt sich selbst, denn $a = 1a$.

(ii) Sind $a, b, c \in R$ mit $a \mid b$ und $b \mid c$, dann gibt es $a', b' \in R$ mit $b = aa'$ und $c = bb'$, d. h. $c = bb' = aa'b'$ und damit teilt a auch c.

(iii) Sind $a, b \in R$ mit $a \mid b$ und $b \mid a$, dann gibt es a' und b' aus R mit $b = a'a$ und $a = b'b$, d. h. $b = a'a = a'b'b$, sodass $1 = a'b'$, d. h. $a'b'$ gehört zu $R^\times$.

Uns stört, dass wir in (iii) keine Gleichheit, sondern nur Gleichheit bis auf Multiplikation mit einer Einheit haben. Der passende Ausweg aus dieser misslichen Lage sind Äquivalenzrelationen: Wir sollten diejenigen Elemente zusammenfassen, die sich nur durch Multiplikation mit einer Einheit unterscheiden.

Proposition IX.2.4 (Assoziiertheit und Teilbarkeit): *Seien R ein kommutativer unitärer nullteilerfreier Ring und $a, b \in R$. Es gilt $a \mid b$ und $b \mid a$ genau dann, wenn es $u \in R^\times$ mit $b = ua$ gibt. Solche Elemente a, b nennen wir* zueinander assoziiert*. Assoziiertheit ist eine Äquivalenzrelation auf R.*

Sind $a, b \in R$ und $a', b' \in R$ so, dass a und a' sowie b und b' assoziiert sind, dann gilt $a \mid b$ genau dann, wenn $a' \mid b'$.

Wir nennen $R^\times a := \{ua \mid u \in R^\times\} = [a]_\sim$ die Assoziiertheitsklasse von a*. Teilbarkeit ist eine Ordnungrelation auf der Menge der Assoziiertheitsklassen.*

Beweis: Ist $a \mid b$ und $b \mid a$, dann liefert Proposition IX.2.3(iii) ein $u \in R^\times$ mit $b = ua$. Ist $b = ua$, dann gelten sowohl $a \mid b$ als auch $b \mid a$ wegen $u^{-1}b = a$. Dass Assoziiertheit eine Äquivalenzrelation ist, ist klar.

Sind $a, b \in R$ und $a', b' \in R$ so, dass a und a' sowie b und b' assoziiert sind, dann wissen wir, dass $a' \mid a$, $a \mid b$ und $b \mid b'$. Wegen Proposition IX.2.3 folgt daraus $a' \mid b'$.

Dass Teilbarkeit eine Ordnungsrelation auf der Menge der Assozziertheitsklassen ist, ist klar – genau das war das Ziel. □

Definition IX.2.5 (Größter gemeinsamer Teiler): Seien R ein kommutativer unitärer nullteilerfreier Ring und $a, b, g \in R$. Gilt $g \mid a$, $g \mid b$ und gilt für alle $g' \in R$ mit $g' \mid a$ und $g' \mid b$, dass $g \mid g'$, dann heißt g ein *größter gemeinsamer Teiler von a und b.*

Bemerkung IX.2.6: Seien $a, b, g, g' \in R$ und g ein größter gemeinsamer Teiler von a, b. Genau dann ist g' assoziiert zu g, wenn g' auch ein größter gemeinsamer Teiler von a und b ist, d. h. größte gemeinsame Teiler sind nur bis auf Assoziiertheit eindeutig. Wir schreiben in diesem Fall $g = \mathrm{ggT}(a, b)$ beziehungsweise $R^\times g = \mathrm{ggT}(a, b)$.

Erinnerung IX.2.7: Seien R ein Ring und $I \subseteq R$ eine nichtleere Teilmenge.

(i) Gilt für alle $a, b \in I$ und $r \in R$, dass $a + b \in I$ und $ra \in I$, so heißt I ein *Ideal in R*.

(ii) Gibt es $a_1, \ldots, a_n$ mit $I = \{r_1 a_1 + \cdots + r_n a_n \mid r_1, \ldots, r_n\}$, dann schreiben wir $I := \langle a_1, \ldots, a_n \rangle$.

(iii) Gibt es $g \in R$ mit $I = \langle g \rangle$, dann heißt I ein *Hauptideal*.

(iv) Sind alle Ideale in R Hauptideale, dann ist R ein Hauptidealring.

(v) Die Ringe $\mathbb{Z}$ und $K[X]$ (für einen Körper K) sind Hauptidealringe.

Bemerkung IX.2.8: Seien R ein Integritätsring und $a, b, d \in R$. Es gilt $d \mid a$ und $d \mid b$ genau dann, wenn $\langle a, b \rangle \subseteq \langle d \rangle$.

Beweis: Es gilt $\langle a, b \rangle \subseteq \langle d \rangle$ genau dann, wenn $a, b \in \langle d \rangle$, was per Definition gerade bedeutet, dass $d \mid a$ und $d \mid b$. □

Definition IX.2.9 (Teilerfremdheit): Seien R ein Integritätsring und a, b Elemente von R, die einen größten gemeinsamen Teiler besitzen. Ist $\mathrm{ggT}(a, b) = 1$, dann heißen a und b *teilerfremd*.

Satz 39 (Verallgemeinerter Bézout): *Seien R ein Integritätsring, $a, b, g \in R$ und $\langle a, b \rangle$ ein Hauptideal. Es gilt genau dann $g = \mathrm{ggT}(a, b)$, wenn $\langle g \rangle = \langle a, b \rangle$.*

Beweis: „$\Longrightarrow$“: Ist $g = \mathrm{ggT}(a, b)$, dann gilt insbesondere $g \mid a$ und $g \mid b$, d. h. wir haben per Proposition IX.2.8, dass $\langle a, b \rangle \subseteq \langle g \rangle$. Weil $\langle a, b \rangle$ nach Voraussetzung ein Hauptideal ist, gibt es g' in R, sodass $\langle g' \rangle = \langle a, b \rangle$. Wieder mit Proposition IX.2.8 erhalten wir $g' \mid a$ und $g' \mid b$. Weil g ein größter gemeinsamer Teiler von a und b ist, wird g von g' geteilt, d. h. g liegt in $\langle g' \rangle$ und damit ist $\langle g \rangle \subseteq \langle g' \rangle = \langle a, b \rangle$; also gilt die behauptete Gleichheit.

„$\Longleftarrow$“: Ist $\langle g \rangle = \langle a, b \rangle$, dann gibt Proposition IX.2.8, dass $g \mid a$ und $g \mid b$. Für $g' \in R$ mit $g' \mid a$ und $g' \mid b$ liefert Proposition IX.2.8 $\langle g' \rangle \supseteq \langle a, b \rangle = \langle g \rangle$, d. h. $g' \mid g$. □

Korollar IX.2.10: *Aus Satz 39 folgt insbesondere:*

(i) *Ist R ein Hauptidealring, dann gibt es für je zwei Elemente $a, b \in R$ einen größten gemeinsamen Teiler.*

(ii) *Ist R ein Hauptidealring und sind $a, b, g \in R$ mit $g = \mathrm{ggT}(a, b)$, dann gibt es $k, \ell \in R$, sodass $g = ka + \ell b$. Insbesondere gilt das Lemma von Bézout in beliebigen Hauptidealringen.*

Eine ganz besondere Sorte von Hauptidealringen sind die euklidischen Ringe, von denen wir bereits an ein paar Stellen in der Vorlesung gesprochen haben. Wir wollen kurz motivieren, was diese Ringe besonders macht.

Definition IX.2.11 (Euklidischer Ring): Es seien R ein Integritätsring und $\varphi\colon R \to \mathbb{N}$ eine Funktion. Falls $\varphi(r) = 0$ genau dann, wenn $r = 0$, und wenn für je zwei a und b aus R, $b \neq 0$, ein c aus R gibt, sodass $\varphi(a - bc) < \varphi(b)$, dann heißt φ eine *Gradfunktion* und das Paar (R, φ) ein *euklidischer Ring*.

Die Beispiele für euklidische Ringe, die uns in der Vorlesung bereits begegnet sind, sind der Ring ganzen Zahlen $\mathbb{Z}$ mit dem Betrag als Gradfunktion, oder der Polynomring $K[X]$ über einem Körper K mit der Gradfunktion

$$\varphi\colon K[X] \longrightarrow \mathbb{N}, \qquad f \longmapsto \begin{cases} \deg(f) + 1, & \text{falls } f \neq 0, \\ 0, & \text{falls } f = 0. \end{cases}$$

Der Vollständigkeit halber halten wir fest:

Proposition IX.2.12: *Sei R ein euklidischer Ring. Dann ist R ein Hauptidealring.*

Für den Beweis der Aussage für $\mathbb{Z}$ mussten wir genau so mit Rest Teilen können, wie es uns die Gradfunktion jetzt erlaubt.

Über das Lemma von Bézout weiß man um die Existenz größter gemeinsamer Teiler für je zwei Elemente eines Hauptidealrings. Für praktische Anwendungen sind euklidische Ringen aber die besseren Ringe, da das systematische Teilen mit Rest ein Verfahren ermöglicht, um konstruktiv größte gemeinsame Teiler zu bestimmen. Das ist der sogenannte euklidische Algorithmus, auf den wir aus Zeitgründen in dieser Vorlesung verzichten.

In allgemeineren Ringen muss es nicht zu je zwei verschiedenen Elementen einen größten gemeinsamen Teiler geben.

Beispiel IX.2.13: Sei R der Ring $\mathbb{Z}[\sqrt{-5}]$.[1] Dieser Missbrauch von Notation ist gerechtfertigt, da die beiden Wurzeln von -5 die komplexen Zahlen $\pm \mathrm{i}\sqrt{5}$

[1] Es handelt sich um den Ganzheitsring des algebraischen Zahlkörpers $\mathbb{Q}(\sqrt{-5})$.

sind, und $\mathbb{Z}[\mathrm{i}\sqrt{5}] = \mathbb{Z}[-\mathrm{i}\sqrt{5}]$; für den Ring spielt es keine Rolle, welche der beiden Wurzeln wir wählen.

So oder so: Der Ring R ist ein Teilring im Körper $\mathbb{C}$, weshalb wir auf R den multiplikativen komplexen Betrag zur Verfügung haben.

Ist ε eine Einheit von R, dann gibt es eine Einheit μ von R mit $\varepsilon\mu = 1$. Über den komplexen Betrag erhalten wir $|\varepsilon||\mu|^2 = 1^2$ als Gleichung in den natürlichen Zahlen, weshalb ε und μ komplexen Betrag 1 haben müssen. Ist umgekehrt ε ein Element von R mit komplexem Betrag 1, dann können wir das Inverse per Formel angeben. Das bedeutet, $R^\times = \{r \in R \mid |r|^2 = 1\} = \{\pm 1\}$.

Für $a = 6$ und $b = 4 + 2\sqrt{-5}$ rechnet man elementar nach, dass

$$a = 3\cdot 2 = (1+\sqrt{-5})(1-\sqrt{-5}), \qquad b = 2(2+\sqrt{-5}) = (1-\sqrt{-5})(-1+\sqrt{-5}),$$

d. h. a und b werden gleichzeitig von 2 und $1-\sqrt{-5}$ geteilt. Ferner sind a und b offensichtlich nicht assoziiert, wir kennen ja die Einheitengruppe, und per Koeffizientenvergleich von $a(x+y\sqrt{-5})$ und b beziehungsweise umgekehrt sieht man, dass weder a von b noch b von a geteilt wird.

Wäre jetzt d ein größter gemeinsamer Teiler von a und b, dann müsste d von 2 und $1-\sqrt{-5}$ geteilt werden, es gäbe also x und y in R mit $d = 2x = (1-\sqrt{-5})y$. Weil der komplexe Betrag multiplikativ ist, hätten wir

$$|d|^2 = |2|^2|x|^2 = 4|x|^2 = |1-\sqrt{-5}|^2|y|^2 = 6|y|^2$$

als Gleichungen in den natürlichen Zahlen, weshalb $|d|^2$ wegen eindeutiger Primfaktorzerlegung in $\mathbb{N}$ von der Form $|d|^2 = 12k$, $k \in \mathbb{N}$, sein müsste. Es sind $|a|^2 = |b|^2 = 36$, und weil d nach Voraussetzung ein Teiler von a und b ist, müsste es s und t in R geben, dass $ds = a$, $dt = b$. Wieder mithilfe des Betrags könnten wir daraus folgern, dass $12k|s|^2 = 36 = 12k|t|^2$.

Wäre s oder t eine Einheit, dann hießen die obigen Gleichungen, dass a respektive b zu d assoziiert wären, was aber nicht sein kann, da weder a von b noch b von a geteilt wird. Das Betragsquadrat könnte deshalb nur die Werte $4, 5, \ldots$ annehmen, und es kann kein entsprechendes d geben.

3. Moduln

In diesem Abschnitt verallgemeinern wir das Konzept eines Vektorraums über einem Körper zu einem Modul über einem Ring. Wir werden sehen. dass es an einigen Stellen in dieser Vorlesung bereits sehr natürlich gewesen wäre, diesen Begriff zu verwenden.

Definition IX.3.1 (Modul): Seien R ein kommutativer unitärer Ring, $(M,+)$ eine abelsche Gruppe und $\cdot\colon R\times M\to M$ eine Abbildung. Falls für alle m,n aus M und r,s aus R gelten:

(i) $1_R m = m$,

(ii) $r(sm) = (rs)m$,

(iii) $r(m+n) = rm + rn$ und $(r+s)m = rm + sm$,

dann heißt $(M,+,\cdot)$ ein *R-Modul* oder *Modul über R*.

Die Definition von „Modul über R“ ist exakt dieselbe wie die für „Vektorraum über K“. Dass „die meisten“ Elemente des „agierenden Monoids“ nicht mehr invertierbar sind, hat gravierende Konsequenzen.

Beispiel IX.3.2 (vorherige Auftritte): (i) Sei R ein kommutativer unitärer Ring. Genau auf den Idealen von R agiert R durch Linksmultiplikation, d. h. genau die Ideale von R sind die in R enthaltenen R-Moduln. Für Körper treten hier keine interessanten Phänomene auf; für Ringe schon!

(ii) Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K und $\varphi\colon V\to V$ linear. Dann wird V via $(f,v)\mapsto f(\varphi)(v)$ zu einem $K[X]$-Modul. Die Strukturtheorie dieses Moduls führt im Fall von $K=\mathbb{C}$ auf die Theorie der Jordanschen Normalform. Für andere Körper gibt es schwächere Normalformtheorien.

(iii) Sei G eine abelsche Gruppe. Dann machen die Rechenregeln in G die Gruppe zu einem $\mathbb{Z}$-Modul. Ein $\mathbb{Z}$-Modul ist umgekehrt eine einfache abelsche Gruppe. Insbesondere ist $\mathbb{Z}/n\mathbb{Z}$ ein $\mathbb{Z}$-Modul.

(iv) Genau wie für Körper werden R^n und $\mathrm{Abb}(M,R)$ mit den naheliegenden Verknüpfungen zu R-Moduln.

Bemerkung IX.3.3 (Bekannte Konzepte):

- Seien R ein kommutativer unitärer Ring und M, N Moduln über R. Eine Abbildung $\varphi\colon M\to N$ mit $\varphi(m_1+\alpha m_2) = \varphi(m_1)+\alpha\varphi(m_2)$ für alle m_1,m_2 in M und α in R heißt *linear* oder *R-Modulhomomorphismus*.
- Ist M ein Modul über R und ist N eine Untergruppe von M, die bezüglich Skalarmultiplikation abgeschlossen ist, dann heißt N ein *R-Untermodul von M*.
- Für eine lineare Abbildung $\varphi\colon M\to N$ sind $\operatorname{Kern}\varphi$ und $\operatorname{Im}\varphi$ wie gewohnt Untermoduln der entsprechenden Moduln.

- Der Homomorphiesatz und der Fortsetzungssatz für lineare Abbildungen gilt genau so für Moduln, da die Beweise zu keiner Zeit gebraucht haben, dass ein Körper zugrundeliegt. Das bedeutet außerdem, dass es für zwei R-Moduln M und N ein Tensorprodukt $M \otimes_R N$ gibt.
- Für eine Teilmenge X des R-Moduls M ist

$$\mathrm{Lin}(X) = \bigcap (N \subseteq M \text{ Untermodul} \mid X \subseteq N) = \left\{ \sum_{i=1}^{n} \alpha_i x_i : n \in \mathbb{N}, \alpha_i \in R, x_i \in X \right\}.$$

die lineare Hülle von X. Gibt es in $\mathrm{Lin}(X)$ nur die triviale Nulldarstellung, dann heißt X *linear unabhängig*. Ist X linear unabhängig mit $\mathrm{Lin}(X) = M$, dann heißt X eine *Basis von* M. Hat M eine Basis, dann heißt M *frei*.

Bemerkung IX.3.4 (Was ist anders?): (i) *Moduln müssen nicht frei sein.* Ist beispielsweise n eine natürliche Zahl, dann können wir $\mathbb{Z}/n\mathbb{Z}$ als Modul über sich selbst, oder aber als $\mathbb{Z}$-Modul auffassen. Tun wir letzteres, dann gibt es keine nichtleere linear unabhängige Teilmenge von $\mathbb{Z}/n\mathbb{Z}$, weil die Vielfachen von n jedes Element von $\mathbb{Z}/n\mathbb{Z}$ „annullieren“. Als $\mathbb{Z}/n\mathbb{Z}$-Modul ist er frei, zum Beispiel mit Basis $\{1\}$.

(ii) *Untermoduln freier Moduln müssen nicht frei sein.* Ist R kein Hauptidealring und ist I ein Ideal, das von mehr als einem Element erzeugt wird, dann ist I als R-Modul nicht frei. In den Übungen haben wir beispielsweise gesehen, dass $\mathbb{Z}[X]$ kein Hauptidealring ist, indem wir gezeigt haben, dass $\langle 2, X\rangle$ kein Hauptideal ist.

(iii) *Die Anzahl der Erzeuger muss nicht monoton sein.* Für Beispiele braucht man bösartige Ringe, aber das kann wirklich schief gehen. Der Polynomring $\mathbb{Z}[X_n \mid n \in \mathbb{N}]$ in abzählbar vielen Veränderlichen ist als Modul über sich selbst frei, da er von der linear unabhängigen Menge $\{1\}$ erzeugt wird. Das Ideal, das von den Polynomen ohne konstanten Term erzeugt wird, ist aber nicht endlich erzeugt, da jedes Polynom nur endlich viele Terme enthält, deren Koeffizienten von Null verschieden sind.

(iv) *Tensorprodukte können degenerieren.* Sind $M = \mathbb{Z}/2\mathbb{Z}$, $N = \mathbb{Z}/3\mathbb{Z}$ und $h\colon M \times N \to X$ eine bilineare Abbildung in irgendeinen anderen $\mathbb{Z}$-Modul, dann gilt

$$h([a],[b]) = h([3][a],[b]) = h([a],[3][b]) = 0,$$

das heißt $M \otimes_{\mathbb{Z}} N$ ist der Nullmodul.

Ein paar beruhigende Worte zum Abschluss: R^n und $\mathrm{Abb}_0(M,R)$ sind freie R-Moduln. Die aus der Vektorraumsituation vertrauten kanonischen Basen sind auch Basen für die jeweiligen Moduln.

Definition IX.3.5 (Algebra): Seien R ein kommutativer unitärer Ring und A ein weiterer Ring. Gibt es eine Abbildung $\circ\colon R \times A \to A$, sodass A ein R-Modul ist und die Ringmultiplikation auf A eine R-bilineare Abbildung, dann heißt A eine *R-Algebra*.

Sind A_1, A_2 zwei R-Algebren, ist $\varphi\colon A_1 \to A_2$ eine Abbildung und gilt für alle a, b aus A_1 und r aus R, dass

$$\varphi(a+b) = \varphi(a) + \varphi(b), \qquad \varphi(a \circ b) = \varphi(a) \circ \varphi(b), \qquad \varphi(ra) = r\varphi(a),$$

dann heißt φ ein *Homomorphismus von R-Algebren*.

4. Multilineare Algebra - Teil 2

Erinnerung IX.4.1: Seien R ein kommutativer unitärer Ring, M und N Moduln über R, n eine nichtnegative ganze Zahl, und $h\colon M^n \to N$ eine multilineare Abbildung.

(i) Gilt für alle $\sigma \in S_n$ und $m_1, \ldots, m_n \in M$, dass

$$h(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) = h(m_1, \ldots, m_n),$$

dann heißt h *symmetrisch*.

(ii) Gilt für alle $m_1, \ldots, m_n \in M$, für die es $i \neq j$ mit $m_i = m_j$ gibt, dass $h(m_1, \ldots, m_n) = \mathbf{0}_N$, dann heißt h *alternierend*.

Wir erinnern uns, dass schiefsymmetrische multilineare Abbildungen spezielle alternierende Abbildungen sind. Wir folgen der Konvention $M^0 = R$, $M^1 = M$ und entsprechend $M^n = M^{n-1} \times M$ für $n \geq 2$.

Definition IX.4.2: Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl. Dann schreiben wir:

(i) $\mathrm{Mult}^n_M(N) := \{h\colon M^n \to N \text{ multilinear}\}$,

(ii) $\mathrm{Sym}^n_M(N) := \{h\colon M^n \to N \text{ symmetrisch und multilinear}\}$,

(iii) $\mathrm{Alt}^n_M(N) := \{h\colon M^n \to N \text{ alternierend und multilinear}\}$.

Alle drei sind R-Moduln zusammen mit den punktweisen Verknüpfungen von Abbildungen.

Satz 40 (Tensorpotenzen): *Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl.*

(i) *Es gibt einen R-Modul $T^n(M)$ zusammen mit einer multilinearen Abbildung $t\colon M^n \to T^n(M)$ mit folgender universellen Abbildungseigenschaft: Für alle R-Moduln N und $h \in \mathrm{Mult}^n_M(N)$ gibt es genau eine lineare Abbildung $\phi\colon T^n(M) \to N$ mit $\phi \circ t = h$.*

(ii) *Es gibz einen R-Modul $S^n(M)$ zusammen mit einer symmetrischen multilineare Abbildung $s\colon M^n \to S^n(M)$ mit folgender universellen Abbildungseigenschaft: Für alle R-Moduln N und $h \in \mathrm{Sym}^n_M(N)$ gibt es genau eine lineare Abbildung $\phi\colon S^n \to N$ mit $\phi \circ s = h$.*

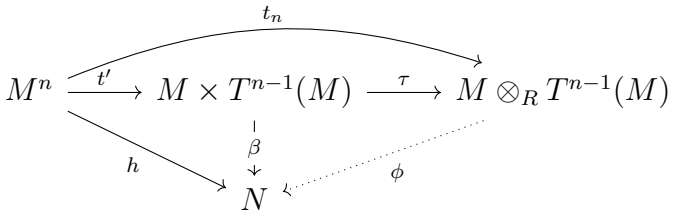
(iii) *Es gibt einen R-Modul $\bigwedge^n(M)$ zuammen mit einer alternierenden multilinearen Abbildung $a\colon M^n \to \bigwedge^n(M)$ mit folgender universellen Abbildungseigenschaft: Für alle R-Moduln N und alle $h \in \mathrm{Alt}^n_M(N)$ gibt es genau eine lineare Abbildung $\phi\colon \bigwedge^n(M) \to N$ mit $\phi \circ a = h$.*

Definition IX.4.3: In der Situation von Satz 40 heißen $T^n(M)$ die *n-te Tensorpotenz von M*, $S^n(M)$ die *n-te symmetrische Potenz von M* und $\bigwedge^n(M)$ die *n-te äußere Potenz von M*.

Beweis (von Satz 40): (i) Wir wollen per vollständiger Induktion vorgehen. Für $n = 0$ setzen wir $T^0(M) := R$, für $n = 1$ setzen wir $T^1(M) := M$ und für $n \geq 2$ setzen wir $T^n(M) := M \otimes_R T^{n-1}(M)$. Ferner setzen wir $t_0 = \mathrm{id}_R$, $t_1 := \mathrm{id}_M$ und für $n \geq 2$ definieren wir

$$t = t_n\colon M^n \longrightarrow T^n(M), \qquad (m_1, \ldots, m_n) \longmapsto m_1 \otimes t_{n-1}(m_2, \ldots, m_n).$$

Nun zeigen wir per vollständiger Induktion die Gültigkeit der universellen Abbildungseigenschaft. Für $n = 0$ und $n = 1$ ist alles klar. Die Aussage gelte nun für ein $n \geq 2$ und es sei $h\colon M^n \to N$ eine multilineare Abbildung. Dann sind wir in der Situation



und wir möchten die Abbildung ϕ konstruieren.

Zur Existenz der Abbildung ϕ: Wir definieren $t'\colon M^n \to M \times T^{n-1}(M)$ durch $(m_1, \ldots, m_n) \mapsto (m_1, t_{n-1}(m_2, \ldots, m_n))$. Für jedes $m_1 \in M$ ist dann $h_{m_1} := h(m_1, \cdot)\colon M^{n-1} \to N$, $(m_2, \ldots, m_n) \mapsto h(m_1, \ldots, m_n)$ multilinear,

d. h., per Induktionsvoraussetzung gibt es einen Vektorraumhomomorphismus $\phi_{m_1}\colon T^{n-1}(M) \to N$ mit $\phi_{m-1} \circ t_{n-1} = h_{m_1}$.

Nun definieren wir $\beta\colon M \times T^{n-1}(M) \to N$ per $(m_1, \omega) \mapsto \phi_{m_1}(\omega)$. Dann ist $\beta \circ t' = h$ und β ist bilinear.

Sei nun $\tau\colon M \times T^{n-1}(M) \to M \otimes_R T^{n-1}(M)$ die Tensorabbildung. Per Konstruktion ist dann auch $\tau \circ t' = t_n$. Wegen der universellen Abbildungseigenschaft des Tensorprodukts existiert eine lineare Abbildung $\phi\colon M \otimes_R T^{n-1}(M) \to N$ mit $\phi \circ \tau = \beta$, d. h. $\phi \circ t_n = \phi \circ \tau \circ t' = \beta \circ t' = h$.

Die Eindeutigkeit von ϕ folgt jetzt aus der Eindeutigkeit von β und ϕ.

(ii) Wir würden gerne als Kandidaten $T^n(M)$ zusammen mit t_n hernehmen. Leider ist t_n im Allgemeinen nicht symmetrisch, d. h. für $\sigma \in S_n - \{\mathrm{id}\}$ und $m_1, \ldots, m_n \in M$ ist im Allgemeinen

$$t(m_1, \ldots, m_n) = m_1 \otimes \cdots \otimes m_n \neq m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} = t(m_{\sigma(1)}, \ldots, m_{\sigma(n)}).$$

Das Problem können wir aber gut aus der Welt schaffen: Wir teilen den Untermodul

$$U_1 := \mathrm{Lin}(\{m_1 \otimes \cdots \otimes m_n - m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} \mid \sigma \in S_n, m_1, \ldots, m_n \in M\})$$

aus $T^n(M)$ heraus und erhalten $S^n(M) := T^n(M)/U_1$ zusammen mit der kanonischen Projektion $\pi_1\colon T^n(M) \to T^n(M)/U_1 = S^n(M)$. Die Komposition $s := \pi_1 \circ t\colon M^n \to S^n(M)$ ist jetzt per Konstruktion sowohl multilinear, als auch symmetrisch.

Für ein multilineares symmetrisches $h\colon M^n \to N$ haben wir genau eine lineare Abbildung $\phi'\colon T^n(M) \to N$ mit $\phi' \circ t_n = h$. Wegen

$$\begin{aligned}\phi'(m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)}) &= \phi'(t(m_{\sigma(1)}, \ldots m_{\sigma(n)})) \\ &= h(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) \\ &= h(m_1, \ldots, m_n) = \phi'(m_1 \otimes \cdots \otimes m_n)\end{aligned}$$

ist ϕ' symmetrisch, d. h. $\phi'(U_1) = \{\mathbf{0}_N\}$. Der Homomorphiesatz liefert uns jetzt eine lineare Abbildung $\phi\colon S^n(M) \to N$ mit $\phi \circ \pi_1 = \phi'$. Dieses ϕ leistet wegen der universellen Abbildungseigenschaft von $T^n(M)$ und der Surjektivität von π_1 das Gewünschte und ist eindeutig.

(iii) Analog zu (ii) definieren wir

$$U_2 := \mathrm{Lin}(\{m_1 \otimes \cdots \otimes m_n \mid m_1, \ldots, m_n \in M, \exists\, i \neq j : m_i = m_j\})$$

und setzen $\bigwedge^n(M) := T^n(M)/U_2$. Es bezeichne $\pi_2\colon T^n(M) \to \bigwedge^n(M)$ die kanoische Projektion auf den Quotientenvektorraum. Dann erfüllen $\bigwedge^n(M)$ und $a := \pi_2 \circ t$ die behauptete universelle Abbildungseigenschaft. □

Bemerkung IX.4.4 (Trivialitäten): Natürlich haben wir $T^0(M) = S^0(M) = \bigwedge^0(M) = R$ sowie $T^1(M) = S^1(M) = \bigwedge^1(M) = M$.

Bemerkung IX.4.5: Ähnlich wie beim Tensorprodukt kann jeweils aus der universellen Abbildungseigenschaft zeigen, dass die Potenzen $T^n(M)$, $S^n(M)$ und $\bigwedge^n(M)$ eindeutig bis auf eindeutige Isomorphie sind.

Notation IX.4.6: Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl. Dann schreiben wir:

(i) $m_1 \otimes \cdots \otimes m_n = t(m_1, \ldots, m_n)$,

(ii) $m_1 \odot \cdots \odot m_n = s(m_1, \ldots, m_n)$,

(iii) $m_1 \wedge \cdots \wedge m_n = a(m_1, \ldots, m_n)$.

Bemerkung IX.4.7 (Erzeuger und Rechenregeln): Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl.

(i) Der R-Modul $T^n(M)$ wird erzeugt von der Menge der reinen Tensoren $\{m_1 \otimes \cdots \otimes m_n \mid m_1, \ldots, m_n \in M\}$. Entsprechend wird $S^n(M)$ erzeugt von $\{m_1 \odot \cdots \odot m_n \mid m_1, \ldots, m_n \in M\}$ und $\bigwedge^n(M)$ von $\{m_1 \wedge \cdots \wedge m_n \mid m_1, \ldots, m_n\}$.

(ii) Für $m_1, \ldots, m_n \in M$, $m_i' \in M$ und $r \in R$ gilt

$$\begin{aligned} &m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i + r m_i' \otimes m_{i+1} \otimes \cdots \otimes m_n \\ &\quad = m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i \otimes m_{i+1} \otimes \cdots \otimes m_n \\ &\qquad + r m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i' \otimes m_{i+1} \otimes \cdots \otimes m_n \end{aligned}$$

und da auch „$\odot$“ sowie „$\wedge$“ multlinear sind, gilt die gleiche Aussage auch für diese Produkte.

(iii) Für alle $m_1, \ldots, m_n \in M$ und $\sigma \in S_n$ gilt

$$m_1 \odot \cdots \odot m_n = m_{\sigma(1)} \odot \cdots \odot m_{\sigma(n)}, \qquad m_1 \wedge \cdots \wedge m_n = \operatorname{sgn}(\sigma) m_{\sigma(1)} \wedge \cdots \wedge m_{\sigma(n)}.$$

Gibt es ferner $i \neq j$ mit $m_i = m_j$, dann ist $m_1 \wedge \cdots \wedge m_n = 0$.

Beispiel IX.4.8: Seien $M = \mathbb{R}^3$ zusammen mit der Standardbasis (e_1, e_2, e_3) und Für $v = 3e_1 + 2e_2$, $w = e_1 + e_2 + 5e_3$.

(i) In $T^2(M)$ gilt

$$\begin{aligned} v \otimes w &= (3e_1 + 2e_2) \otimes (e_1 + e_2 + 5e_3) \\ &= 3e_1 \otimes e_1 + 3e_1 \otimes e_2 + 15e_1 \otimes e_3 + 2e_2 \otimes e_1 + 2e_2 \otimes e_2 + 10e_2 \otimes e_3. \end{aligned}$$

(ii) In $S^2(M)$ haben wir

$$\begin{aligned} v \odot w &= (3e_1 + 2e_2) \odot (e_1 + e_2 + 5e_3) \\ &= 3e_1 \odot e_1 + 5e_1 \odot e_2 + 15e_1 \odot e_3 + 2e_2 \odot e_2 + 10e_2 \odot e_3. \end{aligned}$$

(iii) In $\bigwedge^2(M)$ ist

$$v \wedge w = (3e_1 + 2e_2) \wedge (e_1 + e_2 + 5e_3) = e_1 \wedge e_2 + 15e_1 \wedge e_3 + 10e_2 \wedge e_3.$$

Proposition IX.4.9: *Seien R ein kommutativer unitärer Ring und M ein freier R-Modul vom Rang r mit Basis $(b_1, \ldots, b_r)$. Dann gilt:*

(i) *$T^n(M)$ ist freier R-Modul mit Basis $\{b_{i_1} \otimes \cdots \otimes b_{i_n} \mid 1 \leq i_1, \ldots, i_n \leq r\}$.*

(ii) *$S^n(M)$ ist freier R-Modul mit Basis $\{b_1^{\nu_1} \odot \cdots \odot b_r^{\nu_r} \mid \sum_{i=1}^r \nu_i = n\}$.*[2]

(iii) *$\bigwedge^n(M)$ ist freier R-Modul mit Basis $\{b_{i_1} \wedge \cdots \wedge b_{i_n} \mid 1 \leq i_1 < \cdots < i_n \leq r\}$.*

Beweis: Dass die Mengen jeweils Erzeugendensysteme sind folgt aus Proposition IX.4.7.

(i) Folgt aus Proposition III.3.11 (die dort verwendeten Argumente gehen auch für Moduln durch).

(ii) Fehlt.

(iii) Ist $n > r$, so ist $\bigwedge^n(M) = \{\mathbf{0}\}$, d. h. die Behauptung ist wahr. Ist $n = r$, so stimmt die Behauptung nach Proposition III.6.8.

Sei nun $n < r$ und seien $r_{(i_1,\ldots,i_n)} \in R$ für $1 \leq i_1 < \cdots < i_n \leq r$ mit

$$\sum_{1 \leq i_1 < \cdots < i_n \leq r} r_{(i_1,\ldots,i_n)} b_{i_1} \wedge \cdots \wedge b_{i_n} = \mathbf{0}.$$

Wir wollen für jeden n-Tupel $\boldsymbol{j} = (j_1, \ldots, j_n)$ mit $1 \leq j_1 < \cdots < j_n \leq r$ zeigen, dass $r_{\boldsymbol{j}} = \mathbf{0}$. Wähle dazu $\sigma_{\boldsymbol{j}} \in S_r$ sodass $\sigma_{\boldsymbol{j}}(1) = j_1, \ldots, \sigma_{\boldsymbol{j}}(n) = j_n$ und $\sigma_{\boldsymbol{j}}(n+1), \ldots, \sigma_{\boldsymbol{j}}(r)$ die Werte in $\{1, \ldots, r\} - \{j_1, \ldots, j_n\}$ sind.

Jetzt ist

$$b_{j_1} \wedge \cdots \wedge b_{j_n} \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} = (-1)^\ell b_1 \wedge \cdots \wedge b_r \tag{IX.1}$$

mit einem Exponent $\ell \in \mathbb{N}$ und es ist $b_{i_1} \wedge \cdots \wedge b_{i_n} \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} = \mathbf{0}$, falls $(i_1, \ldots, i_n) \neq (j_1, \ldots, j_n)$. Einsetzen in Gl. (IX.1) gibt

$$\begin{aligned} \mathbf{0} &= \Big(\sum_{1 \leq i_1 < \cdots < i_n \leq r} r_{(i_1,\ldots,i_n)} b_{i_1} \wedge \cdots \wedge b_{i_n} \Big) \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} \\ &= (-1)^\ell \cdot r_{(i_1,\ldots,i_n)} b_1 \wedge \cdots \wedge b_r, \end{aligned}$$

d. h. $r_{(i_1,\ldots,i_n)} = 0$, also die lineare Unabhängigkeit. □

[2] Natürlich meinen wir mit $b_i^{\nu_i}$ das symmetrische Produkt $b_i \odot \cdots \odot b_i$ mit ν_i-vielen Faktoren.

Korollar IX.4.10 (Dimensionen der Potenzen): *Seien K ein Körper und V ein d-dimensionaler K-Vektorraum. Dann gilt:*

(i) *$T^n(V)$ ist ein K-Vektorraum der Dimension d^n.*

(ii) *$S^n(V)$ ist ein K-Vektorraum der Dimension $\binom{d+n-1}{n}$.*

(iii) *$\bigwedge^n(V)$ ist ein K-Vektorraum der Dimension $\binom{d}{n}$.*

Beweis: Wir können die in Proposition IX.4.9 angegebenen Basen mithilfe der bekannten Urnenmodelle einfach abzählen: Die angegebenen Dimensionen sind die Möglichkeiten, aus einer Urne mit d Kugeln n Kugeln zu ziehen...

- ...mit Zurücklegen und mit Beachtung der Reihenfolge für $T^n(V)$,
- ...mit Zurücklegen und ohne Beachtung der Reihenfolge für $S^n(V)$,
- ...ohne Zurücklegen und ohne Beachtung der Reihenfolge für $\bigwedge^n(V)$. □

Proposition IX.4.11 (Induzierte Abbildungen auf Potenzen): *Seien M_1 und M_2 Moduln über R und $\varphi\colon M_1 \to M_2$ ein Homomorphismus von R-Moduln. Dann gibt es eindeutige R-lineare Abbildungen $T^n(\varphi)\colon T^n(M_1) \to T^n(M_2)$, $S^n(\varphi)\colon S^n(M_1) \to S^n(M_2)$ und $\bigwedge^n(\varphi)\colon \bigwedge^n(M_1) \to \bigwedge^n(M_2)$, sodass für alle $m_1, \ldots, m_n$ aus M_1 gilt:*

(i) $T^n(\varphi)(m_1 \otimes \cdots \otimes m_n) = \varphi(m_1) \otimes \cdots \otimes \varphi(m_n)$,

(ii) $S^n(\varphi)(m_1 \odot \cdots \odot m_n) = \varphi(m_1) \odot \cdots \odot \varphi(m_n)$,

(iii) $\bigwedge^n(\varphi)(m_1 \wedge \cdots \wedge m_n) = \varphi(m_1) \wedge \cdots \wedge \varphi(m_n)$.

Beweis: Wir zeigen exemplarisch (iii). Wir sind in der Situation

$$\begin{array}{ccc} M_1^n & \xrightarrow{a} & \bigwedge^n(M_1) \\ & \searrow^{h} & \downarrow{\scriptstyle ?} \\ & & \bigwedge^n(M_2) \end{array}$$

mit $h\colon M_1^n \to \bigwedge^n(M_2)$, $(m_1, \ldots, m_n) \mapsto \varphi(m_1) \wedge \cdots \wedge \varphi(m_n)$. Dieses h ist multilinear und alternierend, weshalb die universelle Abbildungseigenschaft von $\bigwedge^n(M_1)$ eine eindeutige lineare Abbildung $\Phi\colon \bigwedge^n(M_1) \to \bigwedge^n(M_2)$ mit $\Phi \circ a = h$ liefert. Die Kommutativität des obigen Diagramms bedeutet gerade, dass Φ die Abbildung $\bigwedge^n(\varphi)$ aus der Behauptung ist. □

Proposition IX.4.12 (Funktorialität): *Sind $\varphi_1\colon M_1 \to M_2$ und $\varphi_2\colon M_2 \to M_3$ Homomorphismen von R-Moduln, dann sind $T^n(\varphi_2 \circ \varphi_1) = T^n(\varphi_2) \circ T^n(\varphi_1)$, $S^n(\varphi_2 \circ \varphi_1) = S^n(\varphi_2) \circ S^n(\varphi_1)$ und $\bigwedge^n(\varphi_2 \circ \varphi_1) = \bigwedge^n(\varphi_2) \circ \bigwedge^n(\varphi_1)$. Ferner gilt für jeden R-Modul M, dass $T^n(\mathrm{id}_M) = \mathrm{id}_{T^n(M)}$, $S^n(\mathrm{id}_M) = \mathrm{id}_{S^n(M)}$ und $\bigwedge^n(\mathrm{id}_M) = \mathrm{id}_{\bigwedge^n(M)}$.*

Beweis: Exemplarisch zeigen wir die Aussage für T^n. Sind $m_1, \ldots, m_n$ in M_1, dann gilt

$$\begin{aligned}\Big(T^n(\varphi_2) \circ T^n(\varphi_1)\Big)(m_1 \otimes \cdots \otimes m_n) \\ &= T^n(\varphi_2)(\varphi_1(m_1) \otimes \cdots \otimes \varphi_1(m_n)) \\ &= (\varphi_2 \circ \varphi_1)(m_1) \otimes \cdots \otimes (\varphi_2 \circ \varphi_1)(m_n) = T^n(\varphi_2 \circ \varphi_1)(m_1 \otimes \cdots \otimes m_n).\end{aligned}$$

Wegen der Eindeutigkeitsaussage in Proposition IX.4.11 ist deshalb $T^n(\varphi_2 \circ \varphi_1) = T^n(\varphi_2) \circ T^n(\varphi_1)$.

Weil $T^n(M)$ von reinen Tensoren erzeugt wird, genügt es, reine Tensoren zu überprüfen. Sind also $m_1, \ldots, m_n$ in M und ist $m_1 \otimes \cdots \otimes m_n$ der zugehörige reine Tensor in $T^n(M)$, dann gilt

$$T^n(\mathrm{id}_M)(m_1 \otimes \cdots \otimes m_n) = \mathrm{id}_M(m_1) \otimes \cdots \otimes \mathrm{id}_M(m_n) = m_1 \otimes \cdots \otimes m_n,$$

sodass $T^n(\mathrm{id}_M) = \mathrm{id}_{T^n(M)}$ wie behauptet. □

Bemerkung IX.4.13 (Es kann nur eine geben): Sei M der Modul R^n und sei $(e_1, \ldots, e_n)$ die Standardbasis von M. Dann ist $\bigwedge^n(M)$ ein freier R-Modul mit Basis $e_1 \wedge \cdots \wedge e_n$, und es gibt genau eine lineare Abbildung $\Phi\colon \bigwedge^n(M) \to R$ mit $\Phi(e_1 \wedge \cdots \wedge e_n) = 1$. Per universeller Abbildungseigenschaft gibt es genau eine multilineare Abbildung $h\colon R^n \times \cdots \times R^n \to R$, sodass $h(e_1, \ldots, e_n) = 1$.

Definition IX.4.14 (Determinanten-Abbildung): Die Abbildung

$$\det\colon R^n \times \cdots \times R^n \longrightarrow R, \qquad (m_1, \ldots, m_n) \longmapsto \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) m_{1,\sigma(1)} \cdots m_{n,\sigma(n)}$$

wobei m_i der Vektor $(m_{i,1}, \ldots, m_{i,n})^t$ in R^n ist, ist alternierend und multilinear mit $\det(e_1, \ldots, e_n) = 1$, ist also die Abbildung aus Proposition IX.4.13. Zugehörig zu det definieren wir

$$\det\colon R^{n \times n} \longrightarrow R, \qquad A \longmapsto \det(Ae_1, \ldots, Ae_n).$$

5. Tensor, symmetrische und äußere Algebra

In diesem Abschnitt wollen wir – ganz ähnlich wie für Monome – Produkte von Elementen aus $T^\ell(M)$ mit Elementen von $T^k(M)$ erklären. Das wird uns eine Algebra $T(M)$ verschaffen, in der alle Potenzen $T^\ell(M)$ leben. Diese Algebra heißt *Tensoralgebra zu M*. Dasselbe funktioniert auch für die symmetrischen und äußeren Potenzen und liefert die *symmetrische* respektive *äußere Algebra zu M*.

In diesem Abschnitt sei R stets ein kommutativer unitärer Ring.

Definition (Unendliche Summe von R-Moduln): Seien I eine Menge und $(M_i)_{i\in I}$ eine Familie von R-Moduln. Dann heißt

$$\begin{aligned}\bigoplus_{i\in I} M_i &= \{(m_i)_{i\in I} \mid \text{Für jedes } i\in I \text{ ist } m_i \in M_i,\ m_i \neq \mathbf{0} \text{ nur für endlich viele } i\}\\ &= \{f \in \mathrm{Abb}(I, \bigcup M_i) \mid f(i)\in M_i, f(i)\neq \mathbf{0} \text{ nur für endlich viele } i\}\end{aligned}$$

die *direkte Summe der M_i*. Diese wird zusammen mit den komponentenweisen Verknüpfungen selbst zu einem Modul über R. Wir schreiben $\sum_{i\in I} m_i$ für das Element $(m_i)_{i\in I}$ von $\bigoplus_{i\in I} M_i$.

Die Elemente von $\bigoplus_{i\in I} M_i$ sind „schließlich konstante Nullfamilien“, sodass $\sum_{i\in I} m_i$ in Wahrheit immer eine endliche Summe ist.

Proposition IX.5.1: *Seien k und ℓ nicht-negative ganze Zahlen. Dann gibt es eine eindeutige bilineare Abbildung $h_{k,\ell}\colon T^k(M)\times T^\ell \longrightarrow T^{k+\ell}(M)$ mit der Eigenschaft*

$$h_{k,\ell}(m_1\otimes\cdots\otimes m_k, \tilde{m}_1\otimes\cdots\tilde{m}_\ell) = m_1\otimes\cdots\otimes m_k\otimes\tilde{m}_1\otimes\cdots\otimes\tilde{m}_\ell.$$

Beweis: Für den Beweis der Aussage verwenden wir die universelle Abbildungseigenschaft zweimal.

Zunächst definieren wir für jedes Tupel $\tilde{m} = (\tilde{m}_1,\dots,\tilde{m}_\ell)$ aus M^ℓ die Abbildung $h_{\tilde{m}}\colon M^k \to T^{k+\ell}(M)$ durch

$$(m_1,\dots,m_k) \longmapsto m_1\otimes\cdots\otimes m_k\otimes\tilde{m}_1\otimes\cdots\otimes\tilde{m}_\ell.$$

Weil $h_{\tilde{m}}$ multilinear ist, liefert die universelle Abbildungseigenschaft von $T^k(M)$ eine eindeutige lineare Abbildung $\varphi_{\tilde{m}}\colon T^k(M)\to T^{k+\ell}(M)$, die durch

$$\varphi_{\tilde{m}}(m_1\otimes\cdots\otimes m_k) = m_1\otimes\cdots\otimes m_k\otimes\tilde{m}_1\otimes\cdots\otimes\tilde{m}_\ell$$

festgelegt ist. So bekommen wir eine Abbildung $h\colon M^\ell \to \mathrm{Hom}_R(T^k(M), T^{k+\ell}(M))$, $\tilde{m} \mapsto \varphi_{\tilde{m}}$. Wegen der charakterisierenden Eigenschaft von $\varphi_{\tilde{m}}$ ist dieses h multilinear, und die universelle Abbildungseigenschaft von $T^\ell(M)$ gibt uns eine eindeutige lineare Abbildung

$$\varphi\colon T^\ell(M) \longrightarrow \mathrm{Hom}_R(T^k(M), T^{k+\ell}(M)) \quad \text{mit} \quad \varphi(\tilde{m}_1 \otimes \cdots \otimes \tilde{m}_\ell) = \varphi_{\tilde{m}},$$

wobei $\tilde{m} = (\tilde{m}_1, \ldots, \tilde{m}_\ell)$. Für die Abbildung

$$h_{k,\ell}\colon T^k(M) \times T^\ell(M) \longrightarrow T^{k+\ell}(M), \qquad (m, \tilde{m}) \longmapsto \varphi(\tilde{m})(m)$$

gilt dann

$$\begin{aligned} h_{k,\ell}(m_1 \otimes \cdots \otimes m_k, \tilde{m}_1 \otimes \cdots \otimes \tilde{m}_\ell) &= \varphi(\tilde{m}_1 \otimes \cdots \otimes \tilde{m}_\ell)(m_1 \otimes \cdots \otimes m_k) \\ &= \varphi_{\tilde{m}}(m_1 \otimes \cdots \otimes m_k) \\ &= m_1 \otimes \cdots \otimes m_k \otimes \tilde{m}_1 \otimes \cdots \otimes \tilde{m}_\ell, \end{aligned}$$

und $h_{k,\ell}$ ist per Konstruktion bilinear, leistet also das Gewünschte. □

Definition IX.5.2 (Die Hauptakteure): Wir nennen $T(M) = \bigoplus_{n=0}^\infty T^n(M)$, $S(M) = \bigoplus_{n=0}^\infty S^n(M)$, $\bigwedge(M) = \bigoplus_{n=0}^\infty \bigwedge^n(M)$ die *Tensoralgebra*, *symmetrische Algebra* respektive *äußere Algebra zu M*.

Definition IX.5.3 (Multiplikation auf Tensoralgebra): Die Festsetzung

$$m \otimes \tilde{m} = \sum_{i \in \mathbb{N}_0} \sum_{j \in \mathbb{N}_0} h_{i,j}(m_i, \tilde{m}_j)$$

für $m = \sum_{i\in\mathbb{N}_0} m_i$, $\tilde{m} = \sum_{j\in\mathbb{N}_0} \tilde{m}_j$ definiert eine Multiplikation „$\otimes$“ auf $T(M)$.

Insbesondere gilt für reine Tensoren $m = m_1 \otimes \cdots \otimes m_k$ und $\tilde{m} = \tilde{m}_1 \otimes \cdots \otimes \tilde{m}_\ell$, dass $m \otimes \tilde{m} = m_1 \otimes \cdots \otimes m_k \otimes \tilde{m}_1 \otimes \cdots \otimes \tilde{m}_\ell$.

Definition IX.5.4 (Tensor, symmetrische, äußere Algebra): Mit der Multiplikation „$\otimes$“ wird $T(M)$ zu einer unitären R-Algebra. Auf analoge Weise erhalten wir Multiplikationen „$\odot$“ auf $S(M)$ und „$\wedge$“ auf $\bigwedge(M)$, die $S(M)$ respektive $\bigwedge(M)$ zu unitären R-Algebren machen.

Die R-Algebra $T(M)$ heißt *Tensoralgebra zu M*, $S(M)$ heißt *symmetrische Algebra zu M* und $\bigwedge(M)$ heißt *äußere Algebra zu M*.

Kapitel X.

Unendlichdimensionale Vektorräume und Zornsches Lemma

1. Motivation

Erinnerung: Sei V ein Vektorraum über dem Körper K. Eine Teilmenge B von V heißt Basis genau dann, wenn jedes Element von V eine eindeutige Linearkombination der Elemente von B ist. Nach Satz 7 aus der Linearen Algebra I heißt das gerade, dass B linear unabhängig und erzeugend ist, oder, mit den Worten von Satz 9, dass B eine bezüglich Inklusion maximale linear unabhängige Teilmenge von V ist.

Die Potenzmenge $\mathfrak{P}(V)$ von V ist bezüglich Inklusion geordnet, und in $\mathfrak{P}(V)$ sitzt

$$\mathfrak{S} = \{M \subseteq V \mid M \text{ ist linear unabhängig}\}.$$

Wonach wir fragen, wenn wir uns über die Existenz einer Basis von V Gedanken machen, ist ein maximales Element von $\mathfrak{S}$.

Als erster naiver Ansatz beginnen wir mit einer linear unabhängigen Teilmenge M von V. Falls M noch nicht erzeugend ist, dann wählen wir ein v in $V - \mathrm{Lin}(M)$, und erhalten eine neue, größere linear unabhängige Teilmenge $\tilde{M} = M \cup \{v\}$ von V (siehe Korollar II.2.15, Lineare Algebra I). Mit $\tilde{M}$ fahren wir nun genau so fort.

Ist V ein endlich erzeugter Vektorraum, dann terminiert dieses Verfahren, d. h. endet mit einer erzeugenden Menge M_k, die linear unabhängig ist, also einer Basis. In diesem Fall erhalten wir eine Kette $M \subseteq M \cup \{v_1\} \subseteq \cdots \subseteq M \cup \{v_1, \ldots, v_k\}$, wobei $M_i = M \cup \{v_1, \ldots, v_i\}$.

Proposition X.1.1: *Sei K eine Kette in $\mathfrak{S}$. Dann hat K eine obere Schranke in $\mathfrak{S}$.*

Beweis: Seien $v_1, \ldots, v_k$ Elemente von B'_K. Dann gibt es Mengen $S_1, \ldots, S_k$ aus K, sodass $v_i \in S_i$. Weil K total geordnet ist, können wir je zwei Mengen vergleichen; genauer: Für $i, j \in \{, \ldots, k\}$ gilt $S_i \subseteq S_j$ oder $S_j \subseteq S_i$. Wir finden deshalb einen Index i_0, sodass für $1 \leq i \leq k$ gilt: $S_i \subseteq S_{i_0}$. Insbesondere sind $v_1, \ldots, v_k$ Elemente von S_{i_0}. Da S_{i_0} per Konstruktion linear unabhängig ist, ist auch $\{v_1, \ldots, v_k\}$ linear unabhängig. □

Aus Proposition X.1.1 folgt direkt, dass jede Kette in $\mathfrak{S}$ eine obere Schranke hat. Wir werden sehen, dass das impliziert, dass $\mathfrak{S}$ ein maximales Element enthält. Dies wird aus dem Lemma von Zorn folgen.

2. Das Lemma von Zorn

Definition X.2.1: Seien X eine Menge und $\leqslant$ eine Relation auf X.

(i) Ist „$\leqslant$“ reflexiv, antisymmetrisch und transitiv, dann heißt die Relation eine *Ordungsrelation*.

(ii) Sind zusätzlich je zwei Elemente vergleichbar, d. h. gilt für $x, y \in X$ dass $x \leqslant y$ oder $y \leqslant x$, dann heißt „$\leqslant$“ *Totalordnung*.

(iii) Ist K eine nichtleere Teilmenge von X, sodass die Einschränkung von „$\leqslant$“ eine Totalordnung auf K ist, dann heißt K eine *Kette*.

Definition X.2.2 (Maximale, minimale, größte und kleinste Elemente): Seien $(X, \leqslant)$ eine geordnete Menge und x_0 ein Element von X.

(i) Gilt für alle $x \in X$ mit $x_0 \leqslant x$, dass $x = x_0$, dann heißt x *maximal*.

(ii) Gilt für alle $x \in X$ mit $x_0 \geqslant x$, dass $x_0 = x$, dann heißt x *minimal*.

(iii) Gilt für alle $x \in X$, dass $x \leqslant x_0$, dann heißt x_0 ein *größtes Element*.

(iv) Gilt für alle $x \in X$, dass $x \geqslant x_0$, dann heißt x_0 ein *kleinstes Element*.

Bemerkung X.2.3: Eine geordnete Menge X kann viele maximale oder minimale Elemente haben, aber nur ein größtes beziehungsweise kleinstes Element.

Ist x_0 ein größtest Element, dann ist x_0 das einzige maximale Element; genau so für das kleinste Element.

Definition X.2.4 (Obere Schranke, induktiv geordnet): Sei $(X, \leqslant)$ eine geordnete Menge.

(i) Seien S eine Teilmenge von X und x_0 ein Element von X. Gilt für alle $s \in S$, dass $s \leqslant x_0$, dann heißt x_0 eine *obere Schranke*. Das kleinste Element in $\{x \in X \mid x \text{ ist obere Schranke von } S\}$ heißt *kleinste obere Schranke*.

(ii) Hat jede Kette in X eine obere Schranke, dann heißt X *induktiv geordnet*. Hat jede Kette in X eine kleinste obere Schranke, dann heißt X eine *strikt induktiv geordnet*.

Beispiel X.2.5: (i) Die Menge der natürlichen Zahlen mit dem gewöhnlichen „kleiner-gleich“ ist nicht induktiv geordnet, denn $\mathbb{N}$ selbst ist eine Kette, die keine obere Schranke hat.

(ii) Seien M eine Menge und $X = \mathfrak{P}(M)$ zusammen mit der Inklusion. Für jede Teilmenge $\mathfrak{S}$ von $\mathfrak{P}(M)$ gibt es eine kleinste obere Schranke S_0, nämlich $S_0 := \bigcup_{S \in S_0} S$. Insbesondere ist $(\mathfrak{P}(M), \subseteq)$ strikt induktiv geordnet.

Definition X.2.6: Sei $(X, \leqslant)$ eine geordnete Menge. Hat jede nichtleere Teilmenge von X ein kleinstes Element, dann heißt „$\leqslant$“ eine *Wohlordnung*.

Beispiel X.2.7: (i) Die totalgeordente Menge $(\mathbb{N}, \leq)$ ist wohlgeordnet („Prinzip des kleinsten Täters“).

(ii) Die partiell georndete Menge $(\mathfrak{P}(M), \subseteq)$ ist nicht wohlgeordnet, falls M mindestens zwei Elemente hat. Sind nämlich m_1 und m_2 zwei verschiedene Elemente von M, dann gehört $\{\{m_1\}, \{m_2\}\}$ zu $\mathfrak{P}(M)$, und dieses hat kein kleinstes Element.

Wir betrachten folgende Aussagen:

(i) **Auswahlaxiom:** Seien M eine nichtleere Menge, I eine nichtleere Indexmenge und $(M_i)_{i \in I}$ eine Familie nichtleerer Teilmengen von M. Dann gibt es eine Abbildung $f\colon I \to M$, sodass $f(i) \in M_i$. Dieses f nennt man *Auswahlfunktion*.

(ii) **Zorn’sches Lemma:** Sei $(M, \leqslant)$ eine nichtleere geordnete Menge. Ist $(M, \leqslant)$ induktiv geordnet, dann gibt es ein maximales Element in M.

(iii) **Wohlordnungssatz:** Jede Menge M hat eine Totalordnung bezüglich der sie wohlgeordnet ist.

Satz 41: *Das Auswahlaxiom, das Zorn’sche Lemma und der Wohlordnungssatz sind logisch äquivalent.*

Das Auswahlaxiom ist das Axiom in den ZFC-Axiomen, das nicht zu den ZF-Axiomen gehört. Man kann zeigen, dass das Zorn’sche Lemma und damit die Existenz von Basen in beliebigen Vektorräumen nicht aus den gewöhnlichen ZF-Axiomen folgen.