Lineare Algebra

gehalten von Prof. Dr. Weitze-Schmithüsen

UNIVERSITÄT
DES
SAARLANDES

Hinweise

Hinweise zur Mitschrift 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:

s9fhguen@stud.uni-saarland.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 die selben 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 ①, ② und ④ 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 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 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 (iv) 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) Man kann äquivalent dazu Abbildungen mengentheoretisch definieren: 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: 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 Beispiel 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: (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:

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

(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 Beispiel 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: *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 $y \in Y$:

$$g_1(y) = g_1((f \circ g_2)(y)) = g_1(f(g_2(y)) = g_2(y),$$

d. h. $g_1 = g_2$.

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 $b \in Y$.

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

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

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

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

② Definiere die Abbildung

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

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

Definition I.4.16: Definiere wegen Beispiel I.4.15: $X^0 := \mathrm{Abb}(\varnothing, X)$. Also besteht nach Bemerkung 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 $X = \{a_1, \ldots, a_n\}$ eine Menge mit n Elementen ist, dann hat $\mathrm{Perm}(X)$ die Mächtigkeit $n!$.

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 Beispiel 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 Beispiel I.5.2 ist R_1 Äquivalenzrelation und R_2, 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 Bemerkung 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 „$\sim$“ = „$\equiv_n$“, $M = \mathbb{Z}$, $x = 1$. 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) *M ist die Vereinigung aller Äquivalenzklassen, d. h. jedes Element $x \in M$ ist in einer Äquivalenzklasse enthalten,*

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

(iii) Seien $x, y \in M$ mit $[x]_\sim \cap [y]_\sim \neq \varnothing$ und sei $z \in [x]_\sim = [y]_\sim$, d. h. $x \sim z$ und $y \sim z$. Wegen der Symmetrie von „$\sim$“ ist $x \sim z$ und $z \sim y$. Wegen der Transitivität von „$\sim$“ ist deshalb $x \sim y$, also $y \in [x]_\sim$; analog zeigt man $x \in [y]_\sim$, also folgt $[x]_\sim = [y]_\sim$. □

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

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

Bei (ii) ist $I = \{A \mid A \in P\}$ und $\bigcup_{A\in P} A := \bigcup_{i\in I} i$.

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“) erfüllt: Für die Vektoraddition gilt

(VA_1) Für alle $x, y, z \in V$ ist $(x + y) + z = x + (y + z)$, („Assoziativgesetz“)

(VA_2) Für alle $x, y \in V$ ist $x + y = y + x$, („Kommutativgesetz“)

(VA_3) $\mathbf{0}$ ist ein *Nullvektor*, d. h. für alle $x \in V$ ist $x + \mathbf{0} = x = \mathbf{0} + x$,

(VA_4) 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:

(SM_1) Für alle $x \in V$ ist $1x = x$, („Einsgesetz“)

(SM_2) Für alle $\lambda_1, \lambda_2 \in \mathbb{R}$ und $x \in V$ ist $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$,

(SM_3) Für alle $\lambda \in \mathbb{R}$, $x, y \in V$ ist $\lambda(x + y) = \lambda x + \lambda y$,

(SM_4) Für alle $\lambda_1, \lambda_2 \in \mathbb{R}$, $x \in V$ ist $\lambda_1(\lambda_2 x) = (\lambda_1 \lambda_2)x$.

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 skalare Multiplikation wie in Erinnerung 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 Beispiel 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 Definition 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 Bemerkung 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 $\mathbf{0} + f = f$, 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 $v + x = w$. Insbesondere gilt: $\mathbf{0}$ ist der einzige Nullvektor in V.*

(ii) *Für alle $\lambda \in \mathbb{R}$ und $v \in V$ gilt $\lambda \mathbf{0} = \mathbf{0} = 0v$.*

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

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

(v) *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 Definition 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 Definition II.3.4 sind die gleichen wie in Beispiel 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 Bemerkung 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: Sei $A \in \mathbb{R}^{n\times n}$ eine reguläre Matrix. Die Matrix B aus Definition 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: 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.

Definition II.4.4 (Elementarmatrizen): 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 *Elementarmatrizen.*

Achtung: Die Elementarmatrizen sind nicht invertierbar.

Beispiel II.4.5: 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 Elementarmatrizen.

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: *Sei $E_{i,j} \in \mathbb{R}^{q\times m}$.*

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

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

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) Es ist

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

Verwende jetzt, dass $E_{i,j}E_{a,b} = 0$ genau dann, wenn $j \neq a$ und erhalte

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

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

Definition II.4.8: Seien $a, \alpha_1, \dots, \alpha_n \in \mathbb{R}$ und $i, j \in \{1,\dots,n\}$ 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: *Die Matrizen aus Definition 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: *Die Matrizen aus Definition 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

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

(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

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

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: *Die Matrizen aus Definition II.4.8 wirken auf $M \in \mathbb{R}^{m\times n}$ bei Multiplikation von rechts wie folgt:*

(i) *$MA_{i,j}^{\alpha}$ 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) *$M\operatorname{diag}(\alpha_1,\dots,\alpha_n)$ entsteht aus M durch Multiplikation der k-ten Spalte mit α_k für alle $k \in \{1,\dots,n\}$.*

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

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

(ii) und (iii) funktionieren analog. □

Beispiel II.4.12: 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: *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: *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 Lemma 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 Lemma II.4.13) zeigen, dass die angegebenen Matrix für A^{-1} tatsächlich invers zu A ist. □

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

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 $1 \leq i \leq n$ und $1 \leq j \leq m$. 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

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

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: (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 $x, y \in \mathbb{L}^h$, dann sind also $Ax = 0$ und $Ay = 0$, d. h.

$$A(x + y) = Ax + Ay = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

und damit $x + y \in \mathbb{L}^h$. Sind $\lambda \in \mathbb{R}$ und $x \in \mathbb{L}^h$, dann ist $Ax = \mathbf{0}$, also

$$A(\lambda x) = \lambda Ax = \lambda \mathbf{0} = \mathbf{0},$$

d. h. $\lambda x \in \mathbb{L}^h$.

(ii) „$\subseteq$“: Sei $x \in \mathbb{L}(A, b)$, also $Ax = b$. Dann ist

$$A(x - x^{(s)}) = Ax - Ax^{(s)} = b - b = \mathbf{0},$$

damit gilt $v := x - x^{(s)} \in \mathbb{L}^h$ und $x = x^{(s)} + v$. „$\supseteq$“: 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 „$\Leftarrow$“: Ist $CAx = Cb$, dann ist $Ax = C^{-1}CAx = C^{-1}Cb = b$. „$\Rightarrow$“ ist klar. □

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

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

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

(iii) $C = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ *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: Wir definieren in $\mathbb{R}^n = \mathbb{R}^{n\times 1}$ für $i \in \{1, \dots, n\}$ den Vektor e_i durch $e_i((j,1)) = \delta_{i,j}$, d. h. $e_i = (\delta_{i,j})^t_{1\le j\le n}$. Der Vektor e_i heißt *Einheitsvektor*. Beachte:

(i) Der Vektor $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ 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 $s_1, \dots, s_r \in \mathbb{N}$ mit $1 \le s_1 < \cdots < s_r \le m$ mit:

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

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

Die Zahl r heißt der *Rang* von T und $s_1, \dots, s_r$ heißen die *Spaltenindizes*. T ist also in Treppenform, wenn für $1 \le i \le r$ 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 $s_1, \dots, s_r \in \mathbb{N}_0$.

Lemma II.5.8: *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 Lemma 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 Beispiel 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 $F^{(4)} = e_4 - \alpha e_2 - \gamma e_3 + 0 e_5$ 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 $F^{(6)} = e_6 - \beta_k - \delta e_3 - \varepsilon e_5$ ist eine Lösung des homogenen linearen Gleichungssystems.

Lemma II.5.11: (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. $\square$

Bemerkung II.5.12 („Der (-1)-Trick“): Die Fundamentallösungen aus Lemma II.5.11 erhalten wir wie folgt:

(i) Schreibe für $1 \leq i \leq r$ die i-te Zeile von T als s_i-te Zeile in eine neue Matrix $S \in \mathbb{R}^{m \times m}$, deren übrige Zeilen $\mathbf{0}$ sind.

(ii) Die von den Nullspalten verschiedenen Spalten der Matrix $I_n - S$ sind die Fundamentallösungen.

Beispiel II.5.13: Für das lineare Gleichungssystem aus Beispiel II.5.9 sind

$$T = \begin{pmatrix} 0 & 1 & 0 & \alpha & 0 & \beta \\ 0 & 0 & 1 & \gamma & 0 & \delta \\ 0 & 0 & 0 & 0 & 1 & \varepsilon \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \alpha & 0 & \beta \\ 0 & 0 & 1 & \gamma & 0 & \delta \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \varepsilon \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

und

$$I_n - S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\alpha & 0 & -\beta \\ 0 & 0 & 0 & -\gamma & 0 & -\delta \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\varepsilon \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

was sich mit den Ergebnissen aus Beispiel II.5.10 deckt.

Satz 2: *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 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, Lemma II.5.8 und Lemma II.5.11. □

Lemma II.5.14 (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 $A^{\alpha}_{i,j}$.

Induktionsanfang $n = 1$: A besteht nur aus einer Zeile. Ist A die Nullzeile, so sind wir fertig, denn die Nullzeile hat Treppenform und $C = (1)$ leistet das gewünscht. Ist A nicht die Nullzeile, setze $s_1 := \min\{j \in \{1, \ldots, m\} \mid a_{i,j} \neq 0\}$ und wähle $C = (\frac{1}{a_{i,j}})$, dann ist CA in Treppenform.

Induktionsschluss $n-1 \to n$: Ist $A = \mathbf{0}$, dann sind wir fertig. Ist $A \neq \mathbf{0}$, setze $s_1 := \min\{j \in \{1, \ldots, m\} \mid \exists\, i : a_{i,j} \neq 0\}$ (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:

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

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

$$\begin{pmatrix} 1 & 0 \\ 0 & \tilde{C} \end{pmatrix} \in \mathbb{R}^{n \times n} A_1 = \begin{pmatrix} 0 & \cdots & 0 & 1 & z \\ 0 & \cdots & \cdots & 0 & \\ \vdots & & & \vdots & \tilde{T} \\ 0 & \cdots & \cdots & 0 & \end{pmatrix} =: A_2$$

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

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

und leistet das Gewünschte. □

Bemerkung II.5.15: Lemma II.5.14 gibt einen Algorithmus an, wie man ein lineares Gleichungssystem in Treppenform bringen kann.

Beispiel II.5.16: Wir suchen die Lösungsmenge des folgenden linearen Glei-

chungsystems:

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

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

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

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

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.5.17: *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.5.18: *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.

Induktionsanfang $n = 1$: 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.

Induktionsschritt $n \to n+1$: 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.5.17 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

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

mit $D' \in \mathrm{Gl}_{n-1}(\mathbb{R})$ nach (Proposition 4.14) und $z \in \mathbb{R}^{1\times(n-1)}$. Es folgt jetzt, dass für alle $k \in \{1, \ldots, r\}$ gilt:

$$De_k = Dt^{(s_k)} = t'^{(s_k)} = t'^{(s'_k)} = e_k,$$

D muss also von der Form

$$D = \left(\begin{array}{c|c} I_r & B \\ \hline \mathbf{0} & C \end{array}\right)$$

sein, mit $B \in \mathbb{R}^{r\times(m-r)}$, $C \in \mathbb{R}^{(m-r)\times(m-r)}$. Damit ist $T' = DT = T$, da in T alle Zeilen ab der $(r+1)$-ten Zeile Nullzeilen sind. □

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

Beweis: Die Existenz der Gauß-Normalform zu einer Matrix $A \in \mathrm{R}^{n\times m}$ folgt aus Lemma II.5.14.

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 $T_1 = T_2$. □

Definition II.5.19: 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.5.20: 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.5.21: 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) Konstruiere $C \in \mathrm{Gl}_n(\mathbb{R})$, sodass $CA = T$ Treppenform hat (das können wir nach dem Gauß-Algorithmus Lemma II.5.14),

(ii) Brechne die Lösungsmenge $\mathbb{L}$ von $Tx = Cb$ (nach (Satz 1) ist $\mathbb{L}$ auch Lösungsmenge von $Ax = b$ nach (Proposition 5.4)).

Beispiel II.5.22: 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.5.23 (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.5.24: *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.5.25: Der Gauß-Algorithmus liefert ein Verfahren zur Bestimmung, ob eine Matrix $A \in \mathrm{R}^{n\times n}$ invertierbar ist.

Beispiel II.5.26: 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.5.27: 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.5.28: Aus Korollar II.5.23 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 Bemerkung 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: 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: *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 gelten $\varphi(g)\bullet\varphi(g^{-1}) = \varphi(g*g^{-1}) = \varphi(e_G) = e_H$ und $\varphi(g^{-1})\bullet\varphi(g) = e_H$, $\varphi(g^{-1})$ ist also invers zu $\varphi(g)$. □

Proposition III.2.4: *Es gelten:*

(i) $\mathrm{Bild}(\varphi) = \varphi(G) \subseteq H$ *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 $g_1 * g_2^{-1} \in \varphi^{-1}(H)$. Damit ist $\varphi^{-1}(H_1)$ eine Untergruppe von $(H, \bullet)$.

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

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

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

Definition III.2.5: Wir nennen

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

den *Kern* von φ.

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

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

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. $g_1 * g_2^{-1} \in \mathrm{Kern}(\varphi) = \{e_G\}$, 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:

(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: *φ ist ein Isomorphismus genau dann, wenn φ ein bijektiver Gruppenhomomorphismus ist.*

Beweis: „$\Rightarrow$“: Folgt aus der Definition.

„$\Leftarrow$“: 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: *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: (i) Es seien $(G_1, *)$, $(G_2, \bullet)$, $(G_3, \times)$ drei Gruppen und 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 $\mathrm{id} \in \mathrm{Aut}(G)$ und (i). □

3. Die symmetrische Gruppe

M sei stets eine Menge.

Erinnerung III.3.1: Es bezeichnen $\mathrm{Perm}(M) := \{f \colon M \to M \mid f \text{ ist bijektiv}\}$ und $S_n := \mathrm{Perm}(\{1, \ldots, n\})$. Beides sind Gruppen mit der Verkettung von Abbildungen als Verknüpfungen.

Beispiel III.3.2 (S_3): 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
ξ_1	2	3	1
ξ_2	3	1	2
τ_3	3	2	1

Es ist also $S_3 = \{\mathrm{id}, \tau_1, \tau_2, \tau_3, \xi_1, \xi_2\}$.

Bemerkung III.3.3: $\#(S_n) = n!$

Notation III.3.4: 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:

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \xi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Zum Beispiel sind

$$\xi_1 \circ \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \tau_3, \qquad \xi_1^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Definition III.3.6: Für $\sigma \in \mathrm{Perm}(M)$ heißt $\mathrm{Tr}(\sigma) := \{x \in M \mid \sigma(x) \neq x\}$ der *Träger* von σ.

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

Definition III.3.8: (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. Es 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.

(iv) Zyklen $\xi_1, \ldots, \xi_n$ heißen *disjunkt*, falls $\mathrm{Tr}(\xi_i) \cap \mathrm{Tr}(\xi_j) = \varnothing$ für $i \neq j$.

Beispiel III.3.9: In Beispiel III.3.2 sind τ_1, τ_2, τ_3 Transpositionen und ξ_1, ξ_2 3-Zyklen.

Beispiel III.3.10: Wir betrachten die Gruppe S_7.

(i) Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 5 & 2 & 4 & 3 & 6 & 7 \end{pmatrix} = (2, 5, 3)$$

ist ein Zyklus.

(ii) Wir können schreiben $\sigma = (2, 5, 3) = (2, 5) \circ (5, 3)$.

Bemerkung III.3.11: (i) Ist ξ ein k-Zyklus, dann ist $\mathrm{ord}(\xi) = k$ (vergleiche dazu Blatt 7, Aufgabe 2).

(ii) Für einen 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

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

Es ist $\sigma = (1) \circ (2, 12, 6, 4) \circ (3, 8, 10) \circ (5, 7, 11) \circ (9)$, also ist σ kein Zyklus.

Satz 5: *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 $\mathrm{Tr}(\sigma_1) \subseteq \mathrm{Tr}(\sigma) \setminus Z_1$, 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. □

Bemerkung III.3.13: Sind $\sigma_1, \sigma_2 \in \mathrm{Perm}(M)$ mit $\mathrm{Tr}(\sigma_1) \cap \mathrm{Tr}(\sigma_2) = \varnothing$, dann gilt $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$.

Korollar III.3.14 (aus Satz 5): *Es sei $\sigma \in S_n$. Dann gibt es $m \in \mathbb{N}_0$ mit*

$$\sigma = \tau_1 \circ \cdots \circ \tau_m$$

wobei die τ_i Transpositionen sind.

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

Definition III.3.15 (Signum): Sei $\sigma \in S_n$. Wir definieren das *Signum* von σ durch

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

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

Beispiel III.3.16: (i) Sei $\sigma = (1,2,3,4) \in S_4$. Dann ist

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

(ii) Sei $\sigma = (1,2) \in S_n$. Beachte: Es gilt $i < j$ mit $\sigma(i) > \sigma(j)$ genau dann, wenn $i = 1$ und $j = 2$, d. h. es ist $\mathrm{sign}(\sigma) = (-1)$.

Bemerkung III.3.17: (i) Da in der Definition von $\mathrm{sign}(\sigma)$ in Zähler und Nenner des Bruchs bis auf Reihenfolge und Vorzeichen die gleichen Faktoren stehen, ist $\mathrm{sign}(\sigma) \in \{\pm 1\}$.

(ii) Für $\pi \in S_n$ gilt:

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

denn π vertauscht nur die Reihenfolge der Faktoren bei der Berechnung des Signums.

Proposition III.3.18: *Die Abbildung*

$$\operatorname{sign}\colon (S_n, \circ) \to (\{-1, 1\}, \cdot), \qquad \sigma \longmapsto \operatorname{sign}(\sigma)$$

ist ein Gruppenhomomorphismus.

Beweis: Für σ und $\tau \in S_n$ gilt

$$\begin{aligned}\operatorname{sign}(\sigma \circ \tau) &= \prod_{1 \le i < j \le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\ &= \prod_{1 \le i < j \le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \prod_{1 \le i < j \le n} \frac{\tau(j) - \tau(i)}{j - i} = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau),\end{aligned}$$

damit ist alles gezeigt. □

Proposition III.3.19 (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

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

Beispiel III.3.20: In S_5 betrachte $\tau_1 = (1, 2)$ und $\tau_2 = (3, 5)$. Wähle

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix} = (1, 3) \circ (2, 4, 5).$$

Dann ist

$$\pi^{-1} \circ (1, 2) \circ \pi = (3, 1) \circ (5, 4, 2) \circ (1, 2) = (1, 3) \circ (2, 4, 5) = (1)(2)(3, 5)(4).$$

Korollar III.3.21: (i) *Ist τ eine Transposition in S_n, dann ist* $\operatorname{sign}(\tau) = -1$.

(ii) *Ist ζ ein Zyklus der Länge l in S_n, dann ist*

$$\operatorname{sign}(\zeta) = \begin{cases} -1, & \textit{falls } l \textit{ gerade ist}, \\ 1, & \textit{falls } l \textit{ ungerade ist}. \end{cases}$$

(iii) *Ist $\sigma \in S_n$ mit der Zerlegung $\sigma = \zeta_1 \circ \cdots \circ \zeta_N$ in disjunkte Zyklen $\zeta_1, \ldots, \zeta_N$. Dann ist*

$$\operatorname{sign}(\sigma) = \prod_{i=1}^{N} \operatorname{sign}(\zeta_i).$$

Beweis: (i) Proposition III.3.19 garantiert die Existenz einer Permutation $\pi \in S_n$ mit $\tau = \pi^{-1} \circ (1,2) \circ \pi$ und in dieser Situation können wir ausrechnen

$$\operatorname{sign}(\tau) = \operatorname{sign}(\pi)^{-1} \cdot \operatorname{sign}(1,2) \cdot \operatorname{sign}(\pi) = -1.$$

(ii) Sei $\zeta = (x_1, \ldots, x_l)$ mit paarweise verschiedenen $x_1, \ldots, x_l \in \{1, \ldots, n\}$. Nach Bemerkung III.3.11 können wir schreiben $\zeta = (x_1, x_2) \circ (x_2, x_3) \circ \ldots (x_{l-1}, x_l)$, (i) liefert jetzt $\operatorname{sign}(\zeta) = (-1)^{l-1}$, also die Aussage.

(iii) Klar, da sign Gruppenhomomorphismus ist. □

Bemerkung III.3.22: Korollar III.3.21 gibt ein einfaches Verfahren, um das Signum von einer Permutation $\sigma \in S_n$ zu berechnen.

Beispiel III.3.23: In S_{10}:

$$\operatorname{sign}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 4 & 1 & 2 & 9 & 8 & 3 & 5 & 10 & 6 \end{pmatrix}$$
$$= \operatorname{sign}\Big((1,7,3) \circ (2,4) \circ (5,9,10,6,8)\Big) = 1 \cdot (-1) \cdot 1 = (-1).$$

4. Ringe

Definition III.4.1 (Ring):

(i) Eine Menge R mit Verknüpfungen „$+$“, „$\cdot$“ heißt *Ring*, falls gelten:

(1) $(R, +)$ ist eine kommutative Gruppe,

(2) „$\cdot$“ ist assoziativ,

(3) Es gibt ein neutrales Element 1_R bezüglich „$\cdot$“, 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*, „·“ 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“).

(ii) Falls die Verknüpfung „·“ zusätzlich kommutativ ist, dann heißt R *kommutativer Ring*.

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 $\varnothing \neq M$ eine 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: *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 *Ringhomomorphismus* 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 $\Phi\colon (R, +_R, \cdot_R) \to (S, +_S, \cdot_S)$.

Φ ist insbesondere ein Gruppenhomomorphismus zwischen $(R, +_R)$, $(S, +_S)$. Somit gilt zum Beispiel für alle $x \in R$, dass $\Phi(-x) = -\Phi(x)$.

(ii) $\mathrm{Kern}(\Phi) = \{x \in R \mid \Phi(x) = 0_S\}$ heißt *Kern* von Φ.

(iii) Φ heißt *Ringendomorphismus*, falls $(R, +_R, \cdot_R) = (S, +_S, \cdot_S)$. Φ heißt *Ringisomorphismus*, falls es einen Ringhomomorphismus $\Psi\colon (S, +_S, \cdot_S) \to (R, +_R, \cdot_R)$ gibt, mit $\Phi \circ \Psi = \mathrm{id}_S$ und $\Psi \circ \Phi = \mathrm{id}_R$. Φ heißt *Ringautomorphismus*, falls Φ ein Ringendomorphismus und gleichzeitig ein Ringisomorphismus ist.

(iv) 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 $(S, +_S, \cdot_S)$.

Beispiel III.4.5: (i) Sei R die Menge der reellwertigen Cauchyfolgen. R ist ein Ring mit komponentenweiser Addition und Multiplikation von Folgen. Die Abbildung

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

ist ein Ringhomomorphismus.

(ii) Seien $(R, +, \cdot)$ ein Ring und $\varnothing \neq M$ eine Menge. Ferner seien $x_0 \in M$ und $R' := \mathrm{Abb}(M, R)$. Dann ist die Abbildung

$$\Phi_{x_0}\colon R' \longrightarrow R, \qquad f \longmapsto f(x_0)$$

ein Ringhomomorphismus, der sogenannte *Auswertungshomomorphismus zu* x_0.

Beweis: (1) Es gilt

$$\Phi_{x_0}(f + g) = (f + g)(x_0) = f(x_0) + g(x_0) = \Phi_{x_0}(f) + \Phi_{x_0}(g).$$

(2) Wie oben rechnet man nach, dass

$$\Phi_{x_0}(f \cdot g) = (f \cdot g)(x_0) = f(x_0) \cdot g(x_0) = \Phi_{x_0}(f) \cdot \Phi_{x_0}(g).$$

(3) Da $1_{R'}$ die Abbildung ist, die jedem Element von m das Element 1_R zuordnet, ist $1_{R'}(x_0) = 1_R$, also $\Phi_{x_0}(1_{R'}) = 1_R$. □

Bemerkung III.4.6: Sei $(R, +, \cdot)$ ein beliebiger Ring. Wir definieren $\Phi\colon \mathbb{Z} \to R$ wie folgt:

$$\Phi(z) := \begin{cases} \sum_{i=1}^{z} 1_R, & \text{falls } z > 0, \\ 0_R, & \text{falls } z = 0, \\ \sum_{i=1}^{-z} -1_R, & \text{falls } z < 0. \end{cases}$$

Dann ist Φ ein Ringhomomorphismus von $(\mathbb{Z}, +, \cdot)$ nach $(R, +, \cdot)$.

Beweis: (1) Man sieht mithilfe von Einsetzen der Definition und Fallunterscheidung ($a > 0, a = 0, a < 0$ und $b > 0, b = 0, b < 0$), dass

$$\Phi(a + b) = \Phi(a) + \Phi(b).$$

(2) Mit Fallunterscheidungen und unter Ausnutzung der Distributivitätsgesetze erhält man

$$\Phi(a \cdot b) = \Phi(a) \cdot \Phi(b).$$

(3) Es gilt nach Konstruktion, dass $\Phi(1) = 1_R$. □

Definition III.4.7: Für einen Ring $(R, +, \cdot)$ definieren wir

$$\operatorname{char}(R) := \begin{cases} 0, & \text{falls } \sum_{i=1}^{k} 1_R \neq 0_R \text{ für alle } k \in \mathbb{N}, \\ \min\{k \mid \sum_{i=1}^{k} 1_R = 0_R\}, & \text{sonst.} \end{cases}$$

$\operatorname{char}(R)$ heißt die *Charakteristik von R*.

Bemerkung III.4.8: Die Komposition von Ringhomomorphismen gibt wieder einen Ringhomomorphismus, d. h. sind R_1, R_2, R_3 Ringe und $\Phi \in \operatorname{Hom}_{\text{Ring}}(R_1, R_2)$, $\Psi \in \operatorname{Hom}_{\text{Ring}}(R_2, R_3)$, dann ist $\Psi \circ \Phi \in \operatorname{Hom}_{\text{Ring}}(R_1, R_3)$.

Dass die Komposition von Ringhomomorphismen wieder einen Ringhomomorphismus ergibt sieht man genau so ein, wie für Gruppenhomomorphismen.

Ab jetzt sei $(R, +, \cdot)$ stets ein Ring.

Definition III.4.9: (i) Ein $x \in R$ heißt *invertierbar*, falls es ein $y \in R$ gibt, sodass $x \cdot y = 1_R = y \cdot x$. In diesem Fall ist y eindeutig bestimmt, heißt *Inverses* zu x und wird mit x^{-1} notiert.

(ii) $R^\times := \{x \in R \mid x \text{ ist invertierbar}\}$ heißt *Einheitengruppe* oder auch *multiplikative Gruppe von R.*

Bemerkung III.4.10: $(R^\times, \cdot)$ ist eine Gruppe.

Beweis: (i) Wir müssen zeigen, dass „$\cdot$“ tatsächlich eine Verknüpfung auf $R^\times$ definiert, d. h. dass $R^\times$ abgeschlossen unter Multiplikation ist: Sind $x_1, x_2 \in R^\times$, dann muss gelten $x_1 x_2 \in R^\times$. 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 $x_1 \cdot x_2 \in R^\times$.

(ii) „$\cdot$“ ist per Definition 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. □

Notation III.4.11: Seien R ein kommutativer Ring, $x, y \in R$ und $y \in R^\times$. Dann schreiben wir $\frac{x}{y} := xy^{-1}$ und speziell $\frac{1}{y} := y^{-1}$.

Beispiel III.4.12: (i) $\mathbb{Z}^\times = \{\pm 1\}$.

(ii) $\mathbb{R}^\times = \mathbb{R} - \{0\}$, $\mathbb{Q}^\times = \mathbb{Q} - \{0\}$.

Definition III.4.13: Ein Ring $(K, +, \cdot)$ heißt *Körper*, falls K ein kommutativer Ring ist mit $0_K \neq 1_K$ und $K^\times = K - \{0\}$.

Satz 6: *Für den Ring* $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ *gilt*

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[a] \mid \mathrm{ggT}(a, n) = 1\}.$$

Ab jetzt schreiben wir $\overline{a} := [a]$ für die Restklasse $[a]$ von $a \in \mathbb{Z}$ modulo n.

Lemma III.4.14: *In der Gruppe* $(\mathbb{Z}/n\mathbb{Z}, +)$ *gilt:*

(i) $\mathrm{ord}(\overline{a}) = n/\mathrm{ggT}(a, n)$ *für* $a \in \mathbb{Z}$,

(ii) $\mathbb{Z}/n\mathbb{Z} = \langle \overline{a} \rangle$ *genau dann, wenn* $\mathrm{ggT}(a, n) = 1$.

Beweis: (i) Mit einer Aussage auf dem siebten Übungsblatt sehen wir, dass

$$\mathrm{ord}(\overline{a}) = \min\left\{k \in \mathbb{N} : \sum_{i=1}^{k} \overline{a} = 0\right\} = \min\{k \in \mathbb{N} \mid ka \text{ ist durch } n \text{ teilbar}\}$$
$$= \frac{n}{\mathrm{ggT}(a, n)}.$$

(ii) Es gilt $G = \langle \overline{a} \rangle$ genau dann, wenn $\mathrm{ord}(\overline{a}) = n$. Nach (i) ist das genau dann der Fall, wenn $\mathrm{ggT}(a, n) = 1$. □

Lemma III.4.15 (von Bézout): *Seien $a, b \in \mathbb{Z}$. Es gilt $\mathrm{ggT}(a, b) = 1$ genau dann, wenn es $k, l \in \mathbb{Z}$ gibt mit $1 = ak + bl$.*

Beweis: Es gilt $\mathrm{ggT}(a, b) = 1$ nach Lemma III.4.14 genau dann, wenn schon $\mathbb{Z}/b\mathbb{Z} = \langle \overline{a} \rangle$ und es gilt $\mathbb{Z}/b\mathbb{Z} = \langle \overline{a} \rangle$ genau dann, wenn $\overline{1} \in \langle \overline{a} \rangle$. Weiterhin gilt $\overline{1} \in \langle \overline{a} \rangle$ genau dann, wenn $\overline{1} = k\overline{a}$ für ein $k \in \mathbb{Z}$. Es gilt $\overline{1} = k\overline{a}$ genau dann, wenn es $k, l \in \mathbb{Z}$ gibt mit $1 = ka + lb$. □

Beweis (von Satz 6): $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ ist invertierbar genau dann, wenn es ein Element $\overline{k}$ von $\mathbb{Z}/n\mathbb{Z}$ gibt mit $\overline{a}\overline{k} = \overline{1}$. Aber $\overline{a}\overline{k} = \overline{1}$ gilt genau dann, wenn es $k, l \in \mathbb{Z}$ gibt, mit $ak + ln = 1$. Nach Lemma III.4.15 ist das äquivalent zu $\mathrm{ggT}(a, n) = 1$. □

Korollar III.4.16 (aus Lemma III.4.15): *Für $a, b \in \mathbb{Z}$ gilt: Es gibt $k, l \in \mathbb{Z}$ mit $\mathrm{ggT}(a, b) = ka + lb$.*

Beweis: Sei $g = \mathrm{ggT}(a, b)$, dann ist $\mathrm{ggT}(g^{-1}a, g^{-1}b) = 1$, und nach Lemma III.4.15 gibt es jetzt $k', l' \in \mathbb{Z}$ mit $1 = k' \cdot g^{-1}a + l'g^{-1}b$. Multiplizieren mit g liefert die Behauptung. □

Korollar III.4.17: *Ist U eine Untergruppe von $(\mathbb{Z}, +)$, dann gibt es $a_0 \in \mathbb{Z}$ mit $U = a_0\mathbb{Z}$.*

Beweis: Falls $U = \{0\}$, dann können wir U schreiben als $U = 0\mathbb{Z}$. Ist $U \neq \{0\}$, dann gibt es $a \in U$ mit $a > 0$. Das kleinste Element in $U \cap \mathbb{N}$ nennen wir a_0. Sei nun $b \in U$. Ist $b > 0$, dann liefert Korollar III.4.16, dass es $k, l \in \mathbb{Z}$ gibt mit $\mathrm{ggT}(a_0, b) = ka_0 + lb \in U$ (hierbei sind $a_0, b \in U$). Wegen der Minimalität von a_0 muss $\mathrm{ggT}(a_0, b)$ gelten, also $b \in a_0\mathbb{Z}$. Für $b < 0$ ist $-b \in a_0\mathbb{Z}$, und da $a_0\mathbb{Z}$ eine Untergruppe von $\mathbb{Z}$ ist, ist auch $b \in a_0\mathbb{Z}$. Insgesamt ist also $U = a_0\mathbb{Z}$. □

Korollar III.4.18: *Sei $\Phi\colon \mathbb{Z} \to (R, +, \cdot)$ der Ringhomomorphismus aus Bemerkung III.4.6. Dann ist $\mathrm{Kern}(\varphi) = \mathrm{char}(R)\mathbb{Z}$.*

Beweis: Der Kern von Φ ist eine Untergruppe von $(\mathbb{Z}, +)$. Ist $\mathrm{Kern}(\Phi) \neq \{0\}$, dann ist $\mathrm{char}(R)$ die kleinste positive Zahl in $\mathrm{Kern}(\Phi)$ und nach Korollar III.4.17 ist dann $\mathrm{Kern}(\Phi) = \mathrm{char}(R)\mathbb{Z}$. □

Beispiel III.4.19: Für den Ring $(\mathbb{Z}/n\mathbb{Z}, +)$ gilt insbesondere: $\mathrm{char}(\mathbb{Z}/n\mathbb{Z}) = n$.

Beispiel III.4.20: Ein Polynom p über R in der Variable X ist ein Ausdruck der Form

$$p(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n = \sum_{i=0}^{n} a_iX^i$$

mit $n \in \mathbb{N}_0$, $a_0, \ldots, a_n \in R$ und $a_n \neq 0$. n heißt der *Grad* von $p(X)$. Wir nennen

$$R[X] := \{p(X) \mid p(X) \text{ ist Polynom über } R\}$$

den *Polynomring* (in einer Variablen) über R. Seien $p(X) = \sum_{i=0}^{n} a_iX^i$ und $q(X) = \sum_{j=0}^{m} b_jX^j \in R[X]$; der Polynomring wird zu einem Ring mit den folgenden Verknüpfungen:

$$p(X) + q(X) := \sum_{i=0}^{\max\{n,m\}} (a_i + b_i)X^i, \qquad p(X) \cdot q(X) = \sum_{i=0}^{n+m} c_iX^i,$$

hierbei sind $a_i = 0$ für $i > n$, $b_i = 0$ für $i > m$ und $c_i = \sum_{k=0}^{i} a_kb_{i-k}$.

Proposition III.4.21: *Sei $\Phi\colon (R_1, +_1, \cdot_1) \to (R_2, +_2, \cdot_2)$ ein Ringhomomorphismus. Dann gilt: Φ ist ein Isomorphismus genau dann, wenn Φ bijektiv ist.*

Beweis: Die Argumente aus dem Beweis von Proposition III.2.10 gehen hier genau so durch. □

Definition III.4.22: Für einen Ring $(R, +, \cdot)$ und eine Teilmenge $T \subseteq R$ heißt T *Teilring*, falls gelten:

(i) $1_R \in T$,

(ii) Für alle $t_1, t_2 \in T$ sind $t_1 + t_2 \in T$, $t_1 \cdot t_2 \in T$.

(iii) Für alle $t \in T$ ist $-t \in T$.

Bemerkung III.4.23: In Definition III.4.22 ist insbesondere $(T, +, \cdot)$ selbst wieder ein Ring.

5. Körper

Erinnerung: Ein Ring $(K, +, \cdot)$ heißt *Körper*[1], falls K kommutativ ist und $0_K \neq 1_K$ sowie $K^\times = K - \{0\}$ gelten.

[1] Der englische Begriff für Körper ist *field*. Üblicherweise unterscheiden sich die englischen Begriffe nicht besonders von den deutschen Bezeichnungen, bei Körper ist das schon der Fall.

Beispiel III.5.1: (i) $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper.

(ii) $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ist ein Körper genau dann, wenn n eine Prinzahl ist. Das sieht man mithilfe von Satz 6 ein. Ist p eine Primzahl, so schreiben wir $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$. Dabei steht $\mathbb{F}$ für field.

Beispiel III.5.2: Es sei $R = \mathbb{R} \times \mathbb{R}$, ferner seien $(a_1, b_1), (a_2, b_2) \in R$. Auf R erklären wir die Verknüpfungen $+$ und $\cdot$ durch

$$(a_1, b_1)+(a_2, b_2) := (a_1+a_2, b_1+b_2), \qquad (a_1, b_1)\cdot(a_2, b_2) := (a_1a_2-b_1b_2, a_1b_2+b_1a_2).$$

Diese Verknüpfungen machen R zu einem Körper. Um das einzusehen ist zu prüfen, dass ...

(i) ... $(R, +)$ eine abelsche Gruppe (mit neutralem Element $(0, 0)$) ist,

(ii) ... „$\cdot$“ assoziativ ist,

(iii) ... „$\cdot$“ als neutrales Element $(1, 0)$ hat,

(iv) ... „$\cdot$“ kommutativ ist,

(v) ... $(a, b) \in R$ mit $(a, b) \neq 0$ invertierbar ist,

(vi) ... $(1, 0) \neq (0, 0)$ ist,

(vii) ... die Distributivgesetze gelten.

R ist ein besonderer Körper: Wir nennen $\mathbb{C} := R$ den Körper der *komplexen Zahlen*. Es gibt einen injektiven Ringhomomorphismus $\Phi\colon \mathbb{R} \to \mathbb{C}, a \mapsto (a, 0)$. Wir identifizieren $\mathbb{R}$ via Φ mit $\Phi(\mathbb{R})$ und fassen auf diese Weise $\mathbb{R}$ als Teilring von $\mathbb{C}$ auf. Schließlich nennen wir $\mathrm{i} := (0, 1)$ und erhalten

$$\mathrm{i}^2 = (-1, 0) = -(1, 0) = -\Phi(1).$$

Proposition III.5.3: *Es sei $(K, +, \cdot)$ ein Körper. Dann ist* $\mathrm{char}(K) = 0$ *oder* $\mathrm{char}(K) = p$ *für eine Primzahl p.*

Diese Aussage wir auf dem achten Übungsblatt bewiesen.

Proposition III.5.4: *Seien $(R, +, \cdot)$ ein Ring mit $1_R \neq 0_R$ und $(K, +, \cdot)$ ein Körper. Ferner sei $\Phi\colon K \to R$ ein Ringhomomorphismus. Dann ist Φ injektiv.*

Beweis: Wir haben zu zeigen, dass $\mathrm{Kern}(\Phi) = \{0\}$. „$\supseteq$“ ist klar. Für „$\subseteq$“: Angenommen es gäbe $x \in \mathrm{Kern}(\Phi)$ mit $x \neq 0_K$. Da K ein Körper ist, gäbe es dann $y \in K$ mit $xy = yx = 1_K$, und dann wäre

$$\Phi(x)\Phi(y) = \Phi(xy) = \Phi(1_K) = \Phi(1_R), \qquad \Phi(y)\Phi(x) = \Phi(yx) = \Phi(1_K) = 1_R.$$

Da aber $x \in \mathrm{Kern}(\Phi)$, wäre $\Phi(x) = 0$, also auch $\Phi(xy) = \Phi(x)\Phi(y) = 0$ im Widerspruch zu $1_R \neq 0_R$. □

Kapitel IV.

Vektorräume und Dimensionstheorie

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,

(SM_1) Für alle $x \in V$ ist $1_K x = x$,

(SM_2) Für alle $\lambda_2, \lambda_2 \in K$ und $x \in V$ ist $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$,

(SM_3) Für alle $\lambda \in K$ und $x_1 x_2 \in V$ ist $\lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2$,

(SM_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 Definition II.2.2.

(ii) Wir können wie in Definition 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 „$\cdot$“. 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 $\Phi_A\colon K^q \to K^p, v \mapsto Av$ (vergleiche (Bemerkung II.5.27)).

(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 Beispiel 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

Beispiel IV.2.1: Es sei $K = \mathbb{F}_5$. Wir betrachten das folgende lineare Gleichungssystem:

$$\begin{aligned} x + \overline{0}y + z + \ \ w &= \overline{1} \\ z + \overline{2}w &= \overline{2} \end{aligned}$$

Zu diesem linearen Gleichungssystem gehört die erweiterte Matrix

$$\left(\begin{array}{cccc|c} \overline{1} & \overline{0} & \overline{1} & \overline{1} & \overline{1} \\ \overline{0} & \overline{0} & \overline{1} & \overline{2} & \overline{2} \end{array}\right).$$

Diese wollen wir in Treppenform bringen. Dazu ziehen wir die zweite von der ersten Zeile ab und erhalten

$$\left(\begin{array}{cccc|c} \overline{1} & \overline{0} & \overline{0} & \overline{4} & \overline{4} \\ \overline{0} & \overline{0} & \overline{1} & \overline{2} & \overline{2} \end{array}\right).$$

Eine spezielle Lösung für das lineare Gleichungssystem ist $(\overline{4}, \overline{0}, \overline{2}, \overline{0})^t$. Zur Bestimmung der Fundamentallösungen benutzen wir den (-1)-Trick und erhalten

$$S = \begin{pmatrix} \overline{1} & \overline{0} & \overline{0} & \overline{4} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} & \overline{2} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix}, \qquad I - S = \begin{pmatrix} \overline{0} & \overline{0} & \overline{0} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{3} \\ \overline{0} & \overline{0} & \overline{0} & \overline{1} \end{pmatrix}.$$

Nach (Satz 2) ist die Lösungsmenge $\mathbb{L}$ des linearen Gleichungssystems

$$\mathbb{L} = \left\{ \begin{pmatrix} \overline{4} \\ \overline{0} \\ \overline{2} \\ \overline{0} \end{pmatrix} + t \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \\ \overline{0} \end{pmatrix} + s \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{3} \\ \overline{1} \end{pmatrix} : t, s \in \mathbb{F}_5 \right\}.$$

Ab jetzt seien wieder $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein K-Vektorraum.

Definition IV.2.2:

(i) Sei $M = \{v_1, \ldots, v_n\} \subseteq V$ eine endliche Teilmenge von V. Ein Element $v \in V$ heißt *Linearkombination von* M, falls es $\lambda_1, \ldots, \lambda_n \in K$ gibt mit

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i.$$

(ii) Ist $M \subseteq V$ eine beliebige (möglicherweise unendliche) Teilmenge von V, dann heißt $v \in V$ *Linearkombination von* M, falls es eine positive ganze Zahl n sowie $v_1, \ldots, v_n \in M$ und $\lambda_1, \ldots, \lambda_n \in K$ gibt mit

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i.$$

Für $n = 0$ ist $\sum_{i=1}^{n} \lambda_i v_i = \mathbf{0}_V$.

Bemerkung IV.2.3: (i) In Definition IV.2.2 stimmen für endliche Mengen M die Definitionen in (i) und (ii) überein.

(ii) $\mathbf{0}_V$ ist Linearkombination für jedes M.

(iii) Ist $M = \varnothing$, so ist $\mathbf{0}_V$ die einzige Linearkombination von M.

Definition IV.2.4 (Basis): Sei $B = \{v_1, \ldots, v_n\}$ eine endliche Teilmenge von V. B heißt *Basis*, falls sich jeder Vektor $v \in V$ auf genau eine Weise als Linearkombination von B schreiben lässt. Das heißt:

$$\forall\, v \in V\ \exists!\, (\lambda_1, \ldots, \lambda_n) \in K^n : v = \sum_{i=1}^{n} \lambda_i v_i.$$

Definition IV.2.5: Sei M eine beliebige Menge.

(i) Für $f \in \mathrm{Abb}(M, K)$ heißt $\mathrm{Tr}(f) := \{m \in M \mid f(m) \neq 0\}$ der *Träger* von f.

(ii) Wir nennen $\mathrm{Abb}_0(M,K) := \{f \in \mathrm{Abb}(M,K) \mid \mathrm{Tr}(f) \text{ ist endliche Menge}\}$ die Menge der Abbildungen von M nach K mit endlichem Träger.

(iii) Ist $M \subseteq V$ und $\lambda \in \mathrm{Abb}_0(M,K)$, dann definiere

$$\sum_{v\in M} \lambda(v)v := \sum_{v\in \mathrm{Tr}(\lambda)} \lambda(v)v.$$

Definition IV.2.6: Eine beliebige Teilmenge $B \subseteq V$ heißt *Basis*, falls gilt:

$$\forall\, v \in V\ \exists!\, \lambda \in \mathrm{Abb}_0(B,K) : v = \sum_{w\in B} \lambda(w)w.$$

Definition IV.2.7 (Lineare (Un-)Abhängigkeit):

(i) Eine endliche Teilmenge $M = \{v_1, \ldots, v_n\}$ von V heißt *linear unabhängig*, wenn gilt:

$$\forall\, (\lambda_1, \ldots, \lambda_n) \in K^n : \sum_{i=1}^{n} \lambda_i v_i = \mathbf{0}_V \Rightarrow \lambda_1 = \cdots = \lambda_n = 0.$$

Das heißt: $\mathbf{0}_V$ lässt sich auf genau eine Weise als Linearkombination von M darstellen.

(ii) Eine beliebige Teilmenge M von V heißt *linear unabhängig*, wenn gilt:

$$\forall\, \lambda \in \mathrm{Abb}_0(M,K) : \sum_{w\in M} \lambda(w)w = \mathbf{0}_V \Rightarrow \lambda = \mathbf{0}_{\mathrm{Abb}_0(M,K)}.$$

(iii) Eine beliebige Teilmenge M von V heißt *linear abhängig*, wenn M nicht linear unabhängig ist.

Beispiel IV.2.8: (i) Sei $M := \{(1,1,0)^t, (2,1,0)^t, (1,0,0)^t\} \subseteq K^3$. M ist linear abhängig, denn

$$1\begin{pmatrix}1\\1\\0\end{pmatrix} - 1\begin{pmatrix}2\\1\\0\end{pmatrix} + 1\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

(ii) Sei $M := \{(1,0)^t, (1,1)^t\} \subseteq K^2$. M ist linear unabhängig, denn für $\lambda_1, \lambda_2 \in K$ mit $\lambda_1(1,0)^t + \lambda_2(1,1)^t = (0,0)^t$ erhalten wir das lineare Gleichungssystem

$$\begin{aligned}\lambda_1 + \lambda_2 &= 0,\\ \lambda_2 &= 0,\end{aligned}$$

dessen einzige Lösung $\lambda_1 = \lambda_2 = 0$ leicht abzulesen ist.

Definition IV.2.9: Sei $M \subseteq V$. Dann heißt

$$\mathrm{Lin}(M) := \{v \in V \mid v \text{ ist Linearkombination von } M\}$$

die *lineare Hülle von* M, oder *Spann von* M oder auch *Erzeugnis von* M. Gebräuchlich ist auch die Schreibweise $\langle M\rangle := \mathrm{Lin}(M)$ und falls $M = \{v_1, \ldots, v_n\}$, schreibt man auch $\langle v_1, \ldots, v_n\rangle := \langle\{v_1, \ldots, v_n\}\rangle$.

Bemerkung IV.2.10: In (Definition III.1.11) war uns $\langle M\rangle$ schon als Schreibweise für die von M erzeugte Untergruppe begegnet. Das sind unterschiedliche Dinge, die Notation ist nicht eindeutig. Es muss explizit benannt– oder aus dem Kontext klar werden, ob die erzeugte Untergruppe, oder der erzeugte K-Vektorraum gemeint ist.

Beispiel IV.2.11: (i) Seien $v_1 = (1,0,0)^t$, $v_2 = (0,1,0)^t$ und $v_3 = (0,0,1)^t \in K^3$. Dann sind z. B. $\langle v_1, v_2, v_3\rangle = \langle\{v_1, v_2, v_3\}\rangle = K^3$, $\langle v_1\rangle = \{(a,0,0)^t \mid a \in K\}$, oder $\langle v_1, v_2\rangle = \{(a,b,0)^t \mid a, b \in K\}$.

(ii) Seien $v_1 = (1,0,1)^t$, $v_2 = (0,2,0)^t$, $v_3 = (5,2,5)^t$, $v_4 = (1,0,0)^t$. Dann ist $v_3 = 5v_1 + 1v_2 \in \langle v_1, v_2\rangle$, aber $v_4 \notin \langle v_1, v_2\rangle$.

(iii) Ist $M = \varnothing$, dann ist $\mathrm{Lin}(M) = \{\mathbf{0}\}$.

Proposition IV.2.12: *Es sei $M \subseteq V$. Dann gelten:*

(i) $M \subseteq \mathrm{Lin}(M)$,

(ii) $\mathrm{Lin}(M)$ *ist ein Untervektorraum von V; genauer:*

$$\mathrm{Lin}(M) = \bigcap_{U \in J} U$$

mit $J = \{U \subseteq V \mid U$ ist Untervektorraum von V, $M \subseteq U\}$.

(iii) *Ist M' ebenfalls eine Teilmenge von V, dann gilt: Ist $M' \subseteq M$, dann ist auch* $\mathrm{Lin}(M') \subseteq \mathrm{Lin}(M)$.

(iv) *M ist ein Untervektorraum genau dann, wenn* $M = \mathrm{Lin}(M)$.

(v) *Es gilt* $\mathrm{Lin}(\mathrm{Lin}(M)) = \mathrm{Lin}(M)$.

(vi) *Für zwei Untervektorräume U_1, U_2 von V gilt:*

$$\mathrm{Lin}(U_1 \cup U_2) = U_1 + U_2 = \{x + y \mid x \in U_1, y \in U_2\}.$$

Beweis: (i) $x \in M$ können wir zum Beispiel schreiben als $x = 1x \in \mathrm{Lin}(M)$.

(ii) Nachzurechnen, dass $\mathrm{Lin}(M)$ ein Untervektorraum ist, bleibt dem Leser als Übung überlassen.

„$\subseteq$“: Sei $x \in \mathrm{Lin}(M)$, d. h. es gibt $n \in \mathbb{N}_0$, $v_1, \ldots, v_n \in M$ und $\lambda_1, \ldots, \lambda_n \in K$ mit $x = \sum_{i=1}^{n} \lambda_1 v_1 + \cdots + \lambda_n v_n$. Für jedes $U \in J$ ist schon $v_1, \ldots, v_n \in U$, und da U ein Vektorraum ist, gilt damit auch $x = \sum_{i=1}^{n} \lambda_i x_i \in U$. Insgesamt also $x \in \bigcap_{U \in J} U$.

„$\supseteq$“: Folgt direkt, da $\mathrm{Lin}(M) \in J$.

(iii) Klar.

(iv) „$\Leftarrow$“: Direkte Konsequenz aus (ii).

„$\Rightarrow$“: Wir müssen zeigen, dass $M = \mathrm{Lin}(M)$ ist. Aus (i) erhalten wir $M \subseteq \mathrm{Lin}(M)$. Die andere Inklusion $\mathrm{Lin}(M) \subseteq M$ folgt aus (ii), da M ein Untervektorraum von V mit $M \subseteq M$ ist.

(v) Folgt sofort aus (ii) und (iv).

(vi) Die Inklusion „$\supseteq$“ folgt aus der Definition von $U_1 + U_2$. Zu „$\subseteq$“: Aus (Proposition II.2.8) wissen wir, dass $U_1 + U_2$ ein Untervektorraum ist, außerdem ist $U_1 \cup U_2 \subseteq U_1 + U_2$, also $\mathrm{Lin}(U_1 \cup U_2) \subseteq U_1 + U_2$. □

Proposition IV.2.13 (Kriterium für Basen): *Für $B \subseteq V$ gilt: B ist eine Basis von V genau dann, wenn B linear unabhängig ist und* $\mathrm{Lin}(B) = V$.

Beweis: Die Implikation „$\Rightarrow$“ ist direkte Konsequenz der Definition einer Basis (nur die triviale Nulldarstellung).

Für „$\Leftarrow$“: 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.2.14: Es seien $V = \mathbb{R}^2$ und $B = \{(1, -1)^t, (1, 1)^t\}$. B ist linear unabhängig: Seien $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $\lambda_1(1, -1)^t + \lambda_2(1, 1)^t = (0, 0)^t$, dann ist

$$\begin{aligned} \lambda_1 + \lambda_2 &= 0 \\ -\lambda_1 + \lambda_2 &= 0 \end{aligned} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

schreiben wir $A := \left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$, dann können wir formulieren: B ist linear unabhängig genau dann, wenn das homogene lineare Gleichungssystem

$A.(\lambda_1, \lambda_2)^t = \mathbf{0}$ nur die triviale Lösung $\mathbf{0}$ hat. Nach (II.5.23) ist das genau dann der Fall, wenn der Rang(A) gleich der Anzahl der Spalten ist.

Weiter ist $\mathrm{Lin}(B) = \mathbb{R}^2$: Es gilt $\mathrm{Lin}(B) = \mathbb{R}^2$ genau dann, wenn es für alle $v \in \mathbb{R}^2$ Zahlen $\lambda_1, \lambda_2 \in \mathbb{R}$ gibt mit $\lambda_1 v_1 + \lambda_2 v_2 = v$. Das ist genau dann der Fall, wenn es für alle $v \in \mathbb{R}^2$ Zahlen $\lambda_1, \lambda_2 \in \mathbb{R}$ gibt mit $A(\lambda_1, \lambda_2) = v$. Nach (II.5.23) gilt das genau dann, wenn der Rang(A) gleich der Anzahl der Zeilen von A ist.

Jetzt prüft man nach, dass Rang$(A) = 2$, und weiß, dass B eine Basis von $\mathbb{R}^2$ ist.

Proposition IV.2.15: *Seien $v_1, \ldots, v_m \in K^n =: V$ und $A := (v_1 | \ldots | v_m)$. Dann gelten:*

(i) *$\{v_1, \ldots, v_m\}$ ist linear unabhängig genau dann, wenn* Rang$(A) = m$.

(ii) $\mathrm{Lin}(\{v_1, \ldots, v_m\}) = V$ *genau dann, wenn* Rang$(A) = n$.

(iii) *Ist $B \subseteq V$ ist eine Basis von V, so ist $\#(B) = n$.*

Beweis: Wir bemerken, dass $A \in K^{n \times m}$, d. h. A hat n Zeilen und m Spalten.

(i) $\{v_1, \ldots, v_m\}$ ist linear unabhängig genau dann, wenn $A(\lambda_1, \ldots, \lambda_m)^t = \mathbf{0}$ nur die triviale Lösung $(\lambda_1, \ldots, \lambda_m)^t = \mathbf{0}$ hat. Nach (II.5.23) gilt das genau dann, wenn Rang$(A) = m$.

(ii) $\mathrm{Lin}(\{v_1, \ldots, v_m\}) = K^n$ gilt genau dann, wenn das lineare Gleichungssystem $A(\lambda_1, \ldots, \lambda_m)^t = v$ für jedes $v \in V$ eine Lösung hat. Nach (II.5.23) gilt das genau dann, wenn Rang$(A) = n$.

(iii) B sei eine Basis von V. Nach (i) ist dann $\#(B) \leq n$. Insbesondere ist B endlich, also können wir B schreiben als $B = \{v_1, \ldots, v_m\}$ für ein $m \in \mathbb{N}$. Für die Matrix $A = (v_1 | \ldots | v_m)$ gilt dann: $m =$ Rang$(A) = n$, d. h. $m = n$ und B besteht aus n Vektoren. □

Satz 7: *Es sei V ein K-Vektorraum mit Basis B. Dann ist $V \cong \mathrm{Abb}_0(B, K)$.*

Beweis: Wir definieren die Abbildung

$$\Lambda\colon\ \mathrm{Abb}_0(B, K) \longrightarrow V, \qquad \lambda \longmapsto \sum_{b \in B} \lambda(b) b.$$

Die Abbildung Λ ist wohldefiniert und linear. Des Weiteren ist Λ surjektiv, da $\mathrm{Lin}(B) = V$ und injektiv, da B linear unabhängig– und damit Kern$(\Lambda) = \{\mathbf{0}\}$ ist. □

Korollar IV.2.16:

(i) *Besitzt V eine Basis B mit n Elementen (wobei $n \in \mathbb{N}$), dann ist $V \cong K^n$.*

(ii) *Hat V eine endliche Basis, dann sind alle Basen gleichmächtig.*

Beweis: (i) Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Aus Satz 7 wissen wir, dass $V \cong \mathrm{Abb}_0(B, K)$. Weiterhin ist

$$\mathrm{Abb}_0(B, K) \longrightarrow K^n, \qquad \lambda \longmapsto (\lambda(v_1), \ldots, \lambda(v_n))$$

ein Isomorphismus, also $V \cong \mathrm{Abb}_0(B, K) \cong K^n$.

(ii) Sei B eine endliche Basis von V. Falls $B = \varnothing$, ist $V = \{0\}$ und die Behauptung gilt. Ist $\#(B) = n \in \mathbb{N}$, dann wissen wir aus (i), dass $V \cong K^n$, außerdem wissen wir aus Proposition IV.2.15, dass jede Basis von V schon n Elemente hat. □

Bemerkung IV.2.17: (i) Die Umkehrabbildung $\Lambda^{-1}\colon V \to \mathrm{Abb}_0(B, K)$ zu Λ aus Satz 7 nennen wir auch *Koordinatenabbildung D_B zur Basis B.*

(ii) Sei $B = \{b_1, \ldots, b_n\}$. Wählen wir eine Reihenfolge auf den Elementen von B, dann erhalten wir die *geordnete Basis* $(b_1, \ldots, b_n)$. Auf diese Weise bekommt man dann einen Isomorphismus $\mathrm{Abb}_0(B, K) \to K^n$.

Satz 8 (Charakterisierende Eigenschaft einer Basis): *Es sei V ein K-Vektorraum und $B \subseteq V$. Dann sind folgende Aussagen äquivalent:*

(i) *B ist eine Basis,*

(ii) *B ist maximal unter den linear unabhängigen Teilmengen von V, d. h. B ist linear unabhängig und für alle $M \subseteq V$ mit $B \subsetneq M$ gilt: M ist linear abhängig.*

(iii) *B ist ein minimales Erzeugendensystem, d. h. $\mathrm{Lin}(B) = V$ und für alle $M \subsetneq B$ gilt $\mathrm{Lin}(M) \subsetneq V$.*

Beweis: „(i) ⇒ (ii)“: Da B eine Basis ist, ist B linear unabhängig. Sei nun M eine Teilmenge von V mit $B \subsetneq M$. Wählen wir $v \in M \setminus B$, dann gibt es $\lambda \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{w \in B} \lambda(w) w$ und wir können definieren

$$\lambda'\colon M \longrightarrow K, \qquad w \longmapsto \begin{cases} \lambda(w), & \text{für } w \in B, \\ -1, & \text{für } w = v, \\ 0, & \text{für } w \in M \setminus B \cup \{v\}. \end{cases}$$

Damit ist

$$\sum_{w\in M} \lambda'(w)w = \sum_{w\in B} \lambda(w)w - 1v + \sum_{w\in M\setminus B\cup\{v\}} 0w = \mathbf{0}_V$$

und λ' ist nicht die Nullabbildung. M ist also linear abhängig.

„(ii) $\Rightarrow$ (iii)“: Wir haben zu zeigen, dass B ganz V erzeugt, d. h. dass $\mathrm{Lin}(B) = V$. Sei $v \in V$. Ist $v \in B$, dann ist schon $v \in \mathrm{Lin}(B)$. Ist $v \in V \setminus B$, dann setze $M := B \cup \{v\}$. Da B maximal linear unabhängig ist, ist M linear abhängig, d. h. es gibt $\lambda \in \mathrm{Abb}_0(M, K)$, sodass $\sum_{w\in M} \lambda(w)w = \mathbf{0}_V$. Dabei ist $\lambda(v) \neq 0$, denn sonst wäre B linear abhängig. Setze $\mu := \lambda(v) \in K^\times$, dann ist $\mu^{-1} \sum_{w\in M} \lambda(w)w = \mathbf{0}_V$; auflösen nach v gibt

$$v = -\sum_{w\in B} \frac{1}{\mu}\lambda(w)w,$$

es ist also $v \in \mathrm{Lin}(B)$.

Es bleibt zu zeigen, dass B minimal als Erzeugendensystem ist. Angenommen, es gäbe $M' \subsetneq B$ mit $\mathrm{Lin}(M') = V$. Dann erhalten wir wie in „(i) $\Rightarrow$ (ii)“, dass B linear abhängig sein müsste; ein Widerspruch.

„(iii) $\Rightarrow$ (i)“: Wir haben zu zeigen, dass B linear unabhängig ist. Das wollen wir zeigen per Widerspruch: Angenommen, B ist linear abhängig. Dann gäbe es $\lambda \in \mathrm{Abb}_0(M, K)$ mit $\sum_{w\in B} \lambda(w)w = \mathbf{0}_V$, wobei λ nicht die Nullabbildung ist, d. h. es gäbe v_0 mit $\lambda(v_0) \neq 0$. Wir könnten dann auflösen

$$v_0 = -\sum_{w\in B\setminus\{v_0\}} \frac{\lambda(w)}{\lambda(v_0)}w,$$

d. h. es wäre $\mathrm{Lin}(B \setminus \{v_0\}) = \mathrm{Lin}(B) = V$ im Widerspruch zur Minimalität von B. Also muss B linear unabhängig– und damit eine Basis sein. □

Satz 9 (Existenz von Basen für endlich erzeugte Vektorräume): *Es sei V ein K-Vektorraum, der ein endliches Erzeugendensystem besitzt. Dann gelten:*

(i) *V hat eine Basis.*

(ii) *Jedes endliche Erzeugendensystem M von V enthält eine Basis, d. h. ist $M \subseteq V$ mit $V = \mathrm{Lin}(M)$ und $\#(M) < \infty$, dann existiert eine Teilmenge B von M, sodass B eine Basis für V ist.*

(iii) *Jede linear unabhängige Teilmenge M von V lässt sich durch Hinzunehmen endlich vieler Vektoren zu einer Basis ergänzen, d. h. ist $M \subseteq V$ linear unabhängig, dann gibt es $n_0 \in \mathbb{N}_0$ und $v_1, \ldots, v_n \in V$, sodass $M \cup \{v_1, \ldots, v_n\}$ eine Basis von V ist.*

(iv) *Je zwei Basen von V haben gleichviele Elemente.*

Lemma IV.2.18:

(i) *Ist $M \subseteq V$ linear abhängig, so gibt es $v \in M$ mit $v \in \mathrm{Lin}(M \setminus \{v\})$.*

(ii) *Hat V eine Basis B mit $\#(B) = n$, dann gilt für jede linear unabhängige Teilmenge M von V, dass $\#(M) \leq n$.*

(iii) *Ist M linear unabhängig und $v \notin \mathrm{Lin}(M)$, dann gilt: $M \cup \{v\}$ ist linear unabhängig.*

Beweis: Wir haben (i) und (iii) bereits im Beweis von Satz 8 bewiesen.

Zu (ii): Betrachte die Koordinatenabbildung $D_B \colon V \to K^n$. Aus (Bemerkung IV.2.17) beziehungsweise (Korollar IV.2.16) wissen wir, dass D_B ein Isomorphismus ist; insbesondere ist D_B injektiv. Auf Blatt 9, Aufgabe 4 haben wir gezeigt, dass $\Phi(M)$ linear unabhängig ist, d. h. $\#(M) = \#(\Phi(M)) \leq n$, wobei wir uns die Ungleichung in (Proposition IV.2.15) überlegt haben. □

Beweis (von Satz 9): (i) Folgt aus (ii).

(ii) Sei M ein endliches Erzeugendensystem von V. Ist M linear unabhängig, dann ist M Basis nach (Proposition IV.2.13). Ist M linear abhängig, dann gibt es nach Lemma IV.2.18 ein $v \in M$ mit $\mathrm{Lin}(M \setminus \{v\}) = \mathrm{Lin}(M) = V$. $M' := M \setminus \{v\}$ ist also ebenfalls ein Erzeugendensystem. Das können wir so oft wiederholen, bis ein linear unabhängiges Erzeugendensystem entstanden ist.

(iii) Sei M eine linear unabhängige Teilmenge von V. Aussage (ii) stellt sicher, dass V eine endliche Basis besitzt. Außerdem garantiert Lemma IV.2.18, dass $\#(M) \leq \#(B) = n$. Ist $\mathrm{Lin}(M) = V$, dann ist M selbst bereits eine Basis. Falls $\mathrm{Lin}(M) \subsetneq V$, wählen wir $v \in V \setminus \mathrm{Lin}(M)$. Nach Lemma IV.2.18 ist $M' := M \cup \{v\}$ weiterhin linear unabhängig. Außerdem ist $\#(M') \leq n$. Diesen Schritt können wir solange fortsetzen, bis wir nach $n - \#(M)$ Schritten eine linear unabhängige Teilmenge erhalten, die V erzeugt; also eine Basis.

(iv) Diese Aussage haben wir bereits in (Korollar IV.2.16) gezeigt. □

Definition IV.2.19 (Dimension eines endlich erzeugten Vektorraums): Sei V ein K-Vektorraum mit endlichem Erzeugendensystem. Dann wird die Anzahl der Elemente einer Basis B die *Dimension von V* genannt. Wir schreiben $\dim_K(V) = \#(B)$

Satz 9 stellt sicher, dass V eine Basis hat und dass die Dimension von V nicht von der gewählten Basis abhängt.

Definition IV.2.20 (Endlichdimensionaler Vektorraum): Der K-Vektorraum V heißt *endlichdimensional*, wenn V ein endliches Erzeugendensystem hat, sonst *unendlichdimensional*.

Proposition IV.2.21 (Monotonie der Dimension): *Es seien V ein endlichdimensionaler K-Vektorraum und U ein Untervektorraum von V. Dann ist*

$$\dim_K(U) \leq \dim_K(V).$$

Beweis: Jede linear unabhängige Teilmenge von U ist auch linear unabhängig in V, d. h. hat höchstens $n = \dim_K(V)$ viele Elemente. Durch sukzessives Hinzunehmen von Elementen aus U wie im Beweis von Satz 9 erhalten wir eine maximale linear unabhängige Teilmenge von U, also eine Basis von U mit höchstens n Elementen. □

3. Summen von Untervektorräumen und Faktorräume

In diesem Abschnitt seien stets V ein K-Vektorraum und $U_1, \ldots, U_n$ Untervektorräume von V. In (Definition II.2.9) hatten wir den Untervektorraum

$$U_1 + \cdots + U_n := \{u_1 + \cdots + u_n \mid u_i \in U_i\}$$

erklärt, die Summe der $U_1, \ldots, U_n$.

Definition IV.3.1: Die Summe $U_1 + \cdots + U_n$ heißt *direkte Summe*, wenn gilt:

$$\forall\, u_i \in U_i : (u_1 + \cdots + u_n = \mathbf{0}_V \Rightarrow u_1 = \cdots = u_n = \mathbf{0}_V).$$

Ist die Summe direkt, so schreiben wir $\bigoplus_{i=1}^n U_i = \sum_{i=1}^n U_i$.

Bemerkung IV.3.2: (i) Ist $U_1 + \cdots + U_n$ eine direkte Summe, dann gilt $U_i \cap U_j = \{\mathbf{0}\}$ für $i \neq j$, denn gäbe es $\mathbf{0} \neq v \in U_i \cap U_j$, dann wäre $v - v = \mathbf{0}$ eine nichttriviale Nulldarstellung.

(ii) Die Gegenrichtung stimmt nicht, betrachte zum Beispiel den Vektorraum $V = \mathbb{R}^2$ mit Untervektorräumen $U_1 = \langle (1,0)^t \rangle$, $U_2 = \langle (0,1)^t \rangle$ und $U_3 = \langle (1,1)^t \rangle$. In diesem Fall ist schon $U_i \cap U_j = \{\mathbf{0}\}$ für $i \neq j$, die Summe $U_1 + U_2 + U_3$ ist aber nicht direkt, da $(-1)(1,0)^t + (-1)(0,1)^t + 1(1,1)^t = \mathbf{0}$.

Beispiel IV.3.3: (i) Ist $B = \{b_1, \ldots, b_d\}$ eine Basis von V, dann haben wir: $V = \bigoplus_{i=1}^d Kb_i$ ist eine direkte Summe. Hierbei bezeichnet $Kb_i := \langle b_i \rangle$.

(ii) Seien B_i Basen der U_i und $W = \bigoplus_{i=1}^n U_i$ eine direkte Summe. Dann gilt $B_i \cap B_j = \varnothing$ für $i \neq j$, da $U_i \cap U_j = \{\mathbf{0}\}$, und $B = \bigcup_{i=1}^n B_i$ eine Basis von $W = \bigoplus_{i=1}^n U_i$. Ist W endlichdimensional, dann gilt insbesondere für die Dimension von W:

$$\dim_K(W) = \dim_K\left(\bigoplus_{i=1}^n U_i\right) = \sum_{i=1}^n \dim_K(U_i).$$

Proposition IV.3.4: *Falls $U_1, \ldots, U_n$ endlichdimensional sind, dann gilt: $\sum_{i=1}^n U_i$ ist eine direkte Summe genau dann, wenn $\dim_K(\sum_{i=1}^n U_i) = \sum_{i=1}^n \dim_K(U_i)$.*

Beweis: „$\Rightarrow$“ haben wir in Beispiel IV.3.3 (ii) gesehen. Für „$\Leftarrow$“: Wähle für jedes U_i eine Basis B_i. Die Vereinigung $B := \bigcup_{i=1}^n B_i$ erzeugt $W := \sum_{i=1}^n U_i$, außerdem gilt $\#(B) \leq \sum_{i=1}^n \#(B_i) = \sum_{i=1}^n \dim_K(U_i) = \dim_K(\sum_{i=1}^n U_i) = \dim_K(W)$. Da B die Summe erzeugt, muss B aber mindestens $\dim_K(W)$-viele Elemente haben. B muss also sogar eine Basis von W sein, und es sind $\#(B) = \dim_K(W)$ und $B_i \cap B_j = \varnothing$ für $i \neq j$.

Angenommen, wir fänden eine nicht-triviale Nulldarstellung $\mathbf{0} = v_1 + \cdots + v_n$ mit $v_i \in U_i$, wobei nicht alle $v_i = \mathbf{0}$. Dann erhielten wir $\mathbf{0}$ als Linearkombination der Vektoren in B, wobei nicht alle Koeffizienten 0 sind. Dies stünde im Widerspruch dazu, dass B eine Basis ist, die Summe muss also direkt sein. □

Definition IV.3.5 (Äquivalenz modulo Untervektorräumen): Sei U ein Untervektorraum von V. Dann erklärt

$$v \sim w :\Leftrightarrow v - w \in U$$

eine Äquivalenzrelation auf V.

„$\sim$“ ist tatsächlich eine Äquivalenzrelation, wie wir aus (Lemma III.1.16) wissen. Im Folgenden notieren wir mit „$\sim$“ stets die Äquivalenzrelation aus Definition IV.3.5. Wir setzen $[v] := \{w \in V \mid v \sim w\}$ und schreiben auch $v + U := [v]$. Ferner notieren wir mit $V/U := V/\sim = \{[v] \mid v \in V\}$ die Menge der Äquivalenzklassen.

Beispiel IV.3.6: (i) Seien $V = \mathbb{R}^3$, $v_1 = (1,0,0)^t$ und $U = \langle v_1 \rangle = \mathbb{R}v_1$. Für $v \in \mathbb{R}^3$ ist $[v] = v + U = \{v + tv_1 \mid t \in \mathbb{R}\}$ die Gerade durch v parallel zur x-Achse.

(ii) Seien $V = \mathbb{R}^3$, $v_1 = (1,0,0)^t$, $v_2 = (0,1,0)^t$ und $U = \langle v_1, v_2 \rangle$. Für $v \in \mathbb{R}^3$ ist $[v] = v + U = \{v + t_1v_1 + t_2v_2 \mid t_1, t_2 \in \mathbb{R}\}$ die Ebene durch v parallel zur xy-Ebene.

Proposition IV.3.7 (Vektorraumstruktur auf V/U): *Auf V/U erklären wir Verknüpfungen*

$$+\colon V/U \times V/U \longrightarrow V/U,\ [v]+[w] := [v+w],$$
$$\cdot\colon K \times V/U \longrightarrow V/U,\ \lambda[v] := [\lambda v].$$

Diese Verknüpfungen machen V/U zu einem K-Vektorraum.

Beweis: Zunächst haben wir zu zeigen, dass die Verknüpfungen wohldefiniert sind, d. h. nicht von gewählten Repräsentanten abhängen. Seien dazu $v, v', w, w' \in V$ mit $v \sim v'$ und $w \sim w'$, d. h. $v - v' = u_1 \in U$ und $w - w' = u_2 \in U$. Dann gilt einerseits

$$v' + w' - (v + w) = v' - v + w' - w = u_1 + u_2 \in U,$$

d. h. $v' + w' \sim v + w$, also $[v' + w'] = [v + w]$, und andererseits gilt für $\lambda \in K$

$$\lambda v' - \lambda v = \lambda(v' - v) = \lambda u_1 \in U,$$

d. h. $\lambda v' \sim \lambda v$, also $[\lambda v'] = [\lambda v]$. Dass V/U mit „+“, „$\cdot$“ wie oben definiert zu einem K-Vektorraum wird, erhält man, weil die Rechenregeln in Definition IV.1.1 für V die entsprechenden Regeln für V/U implizieren. Exemplarisch zeigen wir SM_1, also $1 \cdot [v] = [v]$ für alle $v \in V$:

$$1 \cdot [v] = [1 \cdot v] = [v]$$

wobei das erste Gleichheitszeichen von der Definition von „$\cdot$“ auf V/U– und das zweite Gleichheitszeichen von der Gültigkeit von SM_1 in V herrührt. □

Im Beweis ist maßgeblich eingegangen, dass $(V, +)$ eine abelsche Gruppe ist.

Definition IV.3.8: Die Abbildung

$$\pi\colon V \longrightarrow V/U, \qquad v \longmapsto [v]$$

heißt *kanonische Projektion* und ist eine surjektive lineare Abbildung. Der Kern von π ist $\mathrm{Kern}(\pi) = U$.

Wann definiert ein Vektorraumhomomorphismus $\Phi\colon V \to W$ eine Abbildung $V/U \to W$? Mindestens brauchen wir, dass gilt: Wenn $v_1 \sim v_2$, dann ist $\Phi(v_1) = \Phi(v_2)$. Insbesondere also: Ist $v \in U$ ($v \in U$ genau dann, wenn $v \sim 0$), dann muss $\Phi(v) = \Phi(\mathbf{0}) = \mathbf{0}$ sein, d. h. $U \subseteq \mathrm{Kern}(\Phi)$. Es wird sich herausstellen, dass das bereits genügt.

Die beschriebene Situation lässt sich mit folgendem Diagramm veranschaulichen:

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & \Phi(V) \subseteq W \\ & \searrow^{\pi} & \uparrow{\scriptstyle \overline{\Phi}} \\ & & V/U \end{array}$$

Satz 10 (Homomorphie-Satz): *Seien V, W zwei K-Vektorräume, $\Phi\colon V \to W$ eine lineare Abbildung, d. h. $\Phi \in \mathrm{Hom}_K(V, W)$, und U ein Untervektorraum von V mit $U \subseteq \mathrm{Kern}(\Phi)$.*

(i) *Es gibt genau eine surjektive lineare Abbildung $\overline{\Phi}\colon V/U \to \Phi(V)$, sodass für alle $v \in V$ gilt: $\overline{\Phi}([v]) = \Phi(v)$.*

(ii) *Wenn sogar $U = \mathrm{Kern}(\Phi)$ gilt, dann folgt, dass $\overline{\Phi}\colon V/U \to \Phi(V)$ ein Isomorphismus ist.*

Insbesondere gilt also $V/\mathrm{Kern}(\Phi) \cong \Phi(V)$.

Beweis: (i) Definiere $\overline{\Phi}\colon V/U \to W$ durch $\overline{\Phi}([v]) = \Phi(v)$ und zeige, dass dies eine wohldefinierte lineare Abbildung ist.

(1) Seien $v, v' \in V$ mit $v \sim v'$, d. h. $v - v' = u \in U$. Dann gilt

$$\Phi(v') - \Phi(v) = \Phi(v' - v) = \Phi(u) = \mathbf{0};$$

also $\Phi(v) = \Phi(v')$ und $\overline{\Phi}$ ist wohldefiniert.

(2) Für $v_1, v_2 \in V$ gilt

$$\overline{\Phi}([v_1]+[v_2]) = \overline{\Phi}([v_1+v_2]) = \Phi(v_1+v_2) = \Phi(v_1)+\Phi(v_2) = \overline{\Phi}([v_1])+\Phi([v_2]).$$

(3) Analog zeigt man für $\lambda \in K$ und $v_1 \in V$, dass $\overline{\Phi}(\lambda[v_1]) = \lambda\overline{\Phi}([v_1])$.

(ii) Es bleibt zu zeigen, dass $\overline{\Phi}$ injektiv ist, wenn $U = \mathrm{Kern}(\Phi)$. Sei dazu nun $v \in V$ mit $\overline{\Phi}([v]) = \mathbf{0}$; wegen der Definition von $\overline{\Phi}$ also $\Phi(v) = 0$, d. h. $v \in \mathrm{Kern}(\Phi) = U$, also $[v] = [0]$. Damit ist die Injektivität von $\overline{\Phi}$ gezeigt. □

Bemerkung IV.3.9: Satz 10 liefert eine Möglichkeit, V/U als Vektorraum zu bestimmen: Finde einen surjektiven Vektorraumhomomorphismus $\Phi\colon V \to W$ mit $U = \mathrm{Kern}(\Phi)$, dann gilt $V/U \cong W$.

Beispiel IV.3.10: Nachzutragen.

Proposition IV.3.11 (Basis des Faktorraums): *Sei U ein Untervektorraum von V und B eine Basis von V, die eine Basis B_U von U enthält. Dann ist*

$$C := \{b + U = [b] \mid b \in B - B_U\}$$

eine Basis des Faktorraums V/U. Hierbei ist $b + U \neq b' + U$ für $b, b' \in B - B_U$.

Beweis: (i) Wir zeigen zunächst, dass C ein Erzeugendensystem ist. Für alle $v \in V$ gibt es $\lambda \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{b\in B} \lambda(b)b$, also $[v] = \sum_{b\in B} \lambda(b)[b]$. Da für $b \in B_U$ schon $[b] = [0]$ gilt, ist also

$$[v] = \sum_{b\in B-B_0} \lambda(b)[b].$$

(ii) Angenommen es wäre $b + U = b' + U$ für $b \neq b'$ und $b, b' \in B - B_U$. Dann wäre $u := b - b' \in U$, es gäbe also zwei verschiedene Arten, um u als Linearkombination von Vektoren aus B zu schreiben, was im Widerspruch zur Basiseigenschaft stünde.

(iii) C ist linear unabhängig: Ist $\lambda \in \mathrm{Abb}_0(C, K)$ mit $\sum_{c\in C} \lambda(c)c = [\mathbf{0}]$, dann ist $\sum_{b\in B-B_U} \lambda([b])[b] = [\mathbf{0}]$, d. h. $\sum_{b\in B-B_U} \lambda([b])b =: u \in U$. Wie in (ii) folgt: $\lambda = \mathbf{0}_{\mathrm{Abb}_0(C,K)}$. □

Beispiel IV.3.12: (i) In der Spagetthi-Konstruktion ist $B_U = \{(1, 0, 0)^t\}$ eine Basis von U und $B = \{(1, 0, 0)^t, (0, 1, 0)^t, (0, 0, 1)^t\}$ eine Basis von V, d. h.

$$C = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + U, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + U \right\}$$

ist Basis des Faktorraums.

(ii) In der Lasagne-Konstruktion ist $B_U = \{(1, 0, 0)^t, (0, 1, 0)^t\}$ eine Basis von U und $B = \{(1, 0, 0)^t, (0, 1, 0)^t, (0, 0, 1)^t\}$ eine Basis von V, d. h.

$$C = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + U \right\}$$

ist Basis des Faktorraums.

Satz 11 (Dimensionsformel): *Es sei V ein K-Vektorraum.*

(i) *Ist V endlichdimensional, dann gilt* $\dim_K(V/U) = \dim_K(V) - \dim_K(U)$.

(ii) *Sind U_1, U_2 endlichdimensionale Untervektorräume von V, dann gilt* $\dim_K(U_1 + U_2) = \dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2)$.

(iii) *Ist V endlichdimensional und $\Phi\colon V \to W$ eine lineare Abbildung, dann ist* $\dim_K(\mathrm{Kern}(\Phi)) + \dim_K(\mathrm{Bild}(\Phi)) = \dim_K(V)$.

Im Folgenden schreiben wir einfach $\dim(V) := \dim_K(V)$.

Beweis: (i) Folgt aus Proposition IV.3.11

(ii) Das kartesische Produkt $U_1 \times U_2$ ist ein K-Vektorraum den üblichen (d. h. komponentenweisen) Verknüfungen. Es gilt $\dim(U_1 \times U_2) = \dim(U_1) + \dim(U_2)$, sind nämlich $B = \{b_1, \ldots, b_n\}$ eine Basis von U_1 und $C = \{c_1, \ldots, c_m\}$ eine Basis von U_2, dann ist $\{(b_1, 0), \ldots, (b_n, 0), (0, c_1), \ldots, (0, c_m)\}$ eine Basis von $U_1 \times U_2$. Die Abbildung

$$\alpha\colon U_1 \times U_2 \longrightarrow V, \qquad (u_1, u_2) \longmapsto u_1 - u_2$$

ist linear mit $\mathrm{Bild}(\alpha) = U_1 + U_2$ und $\mathrm{Kern}(\alpha) = \{(u, v) \in U_1 \times U_2 \mid u = v\}$. Die Abbildung $\mathrm{Kern}(\alpha) \to U_1 \cap U_2$, $(u, v) \mapsto u$ ist ein Isomorphismus mit Umkehrabbildung $u \mapsto (u, u)$. Jetzt liefert Satz 10 zusammen mit (i), dass

$$\begin{aligned}\dim(U_1 + U_2) = \dim(\mathrm{Bild}(\alpha)) &= \dim(U_1 \times U_2)/\mathrm{Kern}(\alpha) \\ &= \dim(U_1 \times U_2) - \dim(\mathrm{Kern}(\alpha)) \\ &= \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).\end{aligned}$$

(iii) Erhalten wir aus Satz 10 wie folgt: Nach Satz 10 ist $\mathrm{Bild}(\Phi) \cong V/\mathrm{Kern}(\Phi)$, d. h.

$$\dim(\mathrm{Bild}(\Phi)) = \dim(V/\mathrm{Kern}(\Phi)) = \dim(V) - \dim(\mathrm{Kern}(\Phi)) \qquad \square$$

Definition IV.3.13: Sei $\Phi\colon V_1 \to V_2$ eine lineare Abbildung und V_1 endlichdimensional. Dann heißt $\mathrm{Rang}(\Phi) = \dim(\mathrm{Bild}(\Phi))$ der *Rang von* Φ. Insbesondere gilt nach Satz 11 die sogenannte *Rangformel*:

$$\mathrm{Rang}(\Phi) + \dim(\mathrm{Kern}(\Phi)) = \dim(V_1).$$

4. Lineare Fortsetzung

In diesem Abschnitt seien V und W stets K-Vektorräume und $\Phi\colon V \to W$ eine lineare Abbildung.

Satz 12 (Fortsetzungssatz):

(i) *Sei B eine Basis von V. Dann ist Φ eindeutig durch die Einschränkung $f := \Phi|_B \colon B \to W$ festgelegt.*

(ii) *Sei eine Abbildung $f \colon B \to W$ gegeben. Dann gibt es genau eine lineare Abbildung $\Phi \colon V \to W$ mit $\Phi|_B = f$. Diese heißt* lineare Fortsetzung von *f.*

Beweis: (i) Sei $v \in V$. Wir können v schreiben als $v = \sum_{b\in B} \lambda(b)b$, wobei $\lambda \in \mathrm{Abb}_0(B, K)$, und es ist

$$\Phi(v) = \Phi\Big(\sum_{b\in B} \lambda(b)b\Big) = \sum_{b\in B} \lambda(b)\Phi(b) = \sum_{b\in B} \lambda(b)f(b),$$

Φ ist also bereits vollständig durch f bestimmt.

(ii) Sei nun $f \colon B \to W$ gegeben. Dann wird durch

$$\Phi \colon V \longrightarrow W, \qquad \Phi(v) := \sum_{b\in B} \lambda(b)f(b)$$

eine Abbildung festgelegt, wobei $\lambda \in \mathrm{Abb}_0(B, K)$ die eindeutige Abbildung ist mit $v = \sum_{b\in B} \lambda(b)b$. Es bleibt zu zeigen, dass Φ in der Tat linear ist. Seien dazu $v_1, v_2 \in V$. Schreibe $v_1 = \sum_{b\in B} \lambda_1(b)b$ sowie $v_2 = \sum_{b\in B} \lambda_2(b)b$ mit Abbildungen $\lambda_1, \lambda_2 \in \mathrm{Abb}_0(B, K)$. Dann gilt $v_1 + v_2 = \sum_{b\in B} \lambda(b)b$ mit $\lambda = \lambda_1 + \lambda_2 \in \mathrm{Abb}_0(B, K)$. Jetzt ist

$$\begin{aligned}\Phi(v_1 + v_2) = \sum_{b\in B} \lambda(b)f(b) &= \sum_{b\in B} (\lambda_1(b) + \lambda_2(b))f(b) \\ &= \sum_{b\in B} \lambda_1(b)f(b) + \sum_{b\in B} \lambda_2(b)f(b) = \Phi(v_1) + \Phi(v_2).\end{aligned}$$

Völlig analog zeigt man, dass Φ die skalare Multiplikation erhält. □

Korollar IV.4.1: *Die Abbildung*

$$H \colon \mathrm{Hom}_K(V, W) \longrightarrow \mathrm{Abb}(B, W), \qquad \Phi \longmapsto \Phi|_B$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis: Das ist eine direkte Konsequenz von Satz 12. □

5. Die Abbildungsmatrix

In diesem Abschnitt seien V und W zwei endlichdimensionale K-Vektorräume, $\Phi\colon V \to W$ eine lineare Abbildung und $B = (b_1, \ldots, b_m)$, $C = (c_1, \ldots, c_n)$ *geordnete* Basen von V respektive W, d. h. die Mengen $\{b_1, \ldots, b_m\}$ und $\{c_1, \ldots, c_n\}$ sind Basen von V respektive W und die Reihenfolge der Vektoren in B und C sind vorgegeben.

Es sei daran erinnert, dass die Wahl der Basen B und C die Koordinatenabbildungen $D_B\colon V \to K^m$ und $D_C\colon W \to K^n$ liefert. Das Ziel des Abschnittes ist es, die Kommutativität des Diagramms

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} \\ K^m & \xrightarrow[v\mapsto Av]{} & K^n \end{array}$$

zu zeigen.

Proposition IV.5.1: *Für $j \in \{1, \ldots, n\}$ gilt $\Phi(b_j) = \sum_{i=1}^{n} a_{i,j}c_i$ mit $a_{i,j} \in K$. Definiere $A = (a_{i,j})_{i,j} \in K^{n\times m}$. Dann gilt $D_C(\Phi(v)) = AD_B(v)$ und die Matrix A heißt* Abbildungsmatrix *von Φ bezüglich der Basen B und C und wird mit $D_{C,B}(\Phi)$ notiert.*

Beweis: $D_B(v) = (\lambda_1, \ldots, \lambda_m)$ genau dann, wenn $v = \lambda_1 b_1 + \cdots + \lambda_m b_m$, d. h.

$$\begin{aligned}\Phi(v) = \lambda_1\Phi(b_1) + \cdots + \lambda_m\Phi(b_m) &= \lambda_1\Big(\sum_{i=1}^{n} a_{i,1}c_i\Big) + \cdots + \lambda_m\Big(\sum_{i=1}^{n} a_{i,m}c_i\Big) \\ &= \sum_{j=1}^{m}\sum_{i=1}^{n} a_{i,j}\lambda_j c_i = \sum_{i=1}^{n}\Big(\sum_{j=1}^{m} a_{i,j}\lambda_j\Big)c_i,\end{aligned}$$

d. h. $D_C(\Phi(v))$ hat als i-te Koordinate $\sum_{j=1}^{m} a_{i,j}\lambda_j$ und deshalb ist der ganze Vektor $D_C(\Phi(v)) = A(\lambda_1, \ldots, \lambda_m)^t = AD_B(v)$. □

Die Matrix $A = D_{C,B}(\Phi)$ ist also bestimmt dadurch, dass $l \circ D_B = D_C \circ \Phi$, wobei l die lineare Abbildung $v \mapsto D_{C,B}(\Phi)v$ ist.

Beispiel IV.5.2: (i) Seien $V = W = \mathbb{R}^{2\times 2}$ und $\Phi\colon V \to V, A \mapsto A^t$. Wir wählen als geordnete Basis für V und W

$$\left(b_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, b_2 := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, b_3 := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, b_4 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

Dann ist die Darstellungsmatrix von Φ bezüglich B und B

$$D_{B,B}(\Phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(ii) 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

$$D_{B,B}(\Phi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Bemerkung IV.5.3: Sei $A := D_{B,C}(\Phi) = (s_1 | \dots | s_m) = (z_1, \dots, z_n)^t \in K^{n \times m}$ mit Spaltenvektoren $s_1, \dots, s_n$ und Zeilenvektoren $z_1, \dots, z_m$. Dann sind

(i) $D_B(\mathrm{Kern}(\Phi)) = \mathbb{L}(A|\mathbf{0})$,

(ii) $D_C(\mathrm{Bild}(\Phi)) = \langle s_1, \dots, s_n \rangle$.

Proposition IV.5.4 (Der Rang ist der Rang ist der Rang): *In der Situation von Bemerkung IV.5.3 gilt:*

$$\mathrm{Rang}(\Phi) = \dim(\langle s_1, \dots, s_n \rangle) = \dim(\langle z_1, \dots, z_m \rangle) = \mathrm{Rang}(A).$$

Beweis: Da $\mathrm{Rang}(\Phi) = \dim(\mathrm{Bild}(\Phi))$ und $\mathrm{Bild}(\Phi) = \langle s_1, \dots, s_n \rangle$ gilt die erste Gleichheit. Außerdem ist $\mathrm{Rang}(A) = \mathrm{Rang}(T) =: r$, wobei T die Treppenform von A ist. Wir erinnern uns, dass $\dim(\mathbb{L}(A|\mathbf{0})) = \dim(\mathbb{L}(T|\mathbf{0})) = m - r$, wir haben also

$$r = m - \dim(\mathrm{Kern}(\Phi)) = \dim(\mathrm{Bild}(\Phi)) = \mathrm{Rang}(\Phi),$$

wobei das zweite Gleichheitszeichen wegen der Rangformel gilt. Schließlich ist $\mathrm{Rang}(A) = \mathrm{Rang}(T) = \dim(\langle z_1, \dots, z_n \rangle)$ und es gilt auch die letzte Gleichheit. □

Anwendung von Proposition IV.5.4:

- Berechne Rang(Φ) zum Beispiel dadurch, dass $A = D_{B,C}(\Phi)$ in Treppenform gebracht wird,
- Die Vektoren $v_1, \ldots, v_k \in K^k$ sind linear unabhängig genau dann, wenn $\text{Rang}[(v_1|\ldots|v_k)] = k$.

Proposition IV.5.5 (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

$$\begin{aligned} \ell_1\colon K^{n_1} &\longrightarrow K^{n_2}, \qquad x \longmapsto D_{B_2,B_1}(\Phi)x, \\ \ell_2\colon K^{n_2} &\longrightarrow K^{n_3}, \qquad x \longmapsto D_{B_3,B_2}(\Psi)x. \end{aligned}$$

Es gilt

$$(D_{B_3} \circ \Psi \circ \Phi)(x) = (\ell_2 \circ D_{B_2} \circ \Phi)(x) = (\ell_2 \circ l_1 \circ D_{B_1})(x).$$

Beachte, dass $(\ell_2 \circ \ell_1)(v) = D_{B_3,B_2}(\Psi) D_{B_2,B_1}(\Phi) v$. Wegen Proposition IV.5.1 ist $D_{B_3,B_1}(\Psi \circ \Phi) = D_{B_3,B_2}(\Psi) D_{B_2,B_1}(\Phi)$. □

Definition IV.5.6 (Basiswechselmatrix): Sei $B = (b'_1, \ldots, b'_m)$ eine weitere geordnete Basis von V. Dann heißt $M_{B',B} := D_{B',B}(\text{id})$ die *Basiswechselmatrix* von B nach B'.

Proposition IV.5.7: *In der Situation von Definition IV.5.6 gilt:*

(i) *Für alle $v \in V$ ist $D_{B'}(v) = M_{B',B} D_B(v)$.*

(ii) *Ist $C' = (c'_1, \ldots, c'_n)$ eine weitere geordnete Basis von W, dann gilt:*

$$D_{C',B'}(\Phi) = M_{C',C} D_{C,B}(\Phi) M_{B,B'}.$$

(iii) *Die Matrix $M_{B',B}$ ist invertierbar und $M_{B,B'} = M_{B',B}^{-1}$.*

Beweis: (i) Wegen Proposition IV.5.1 wissen wir: Im Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}} & V \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_{B'}} \\ K^m & \xrightarrow[x\mapsto M_{B',B}x]{} & K^m \end{array}$$

gilt $D_{B'}(v) = M_{B',B} D_B(v)$.

(ii) Die Behauptung ist klar nach Anwendung von Proposition IV.5.1 auf das Diagramm

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}} & V & \xrightarrow{\Phi} & W & \xrightarrow{\mathrm{id}} & W \\ {\scriptstyle D_{B'}}\downarrow & & {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} & & \downarrow{\scriptstyle D_{C'}} \\ K^m & \xrightarrow[x\mapsto M_{B,B'}x]{} & K^m & \xrightarrow[x\mapsto D_{C,B}(\Phi)x]{} & K^n & \xrightarrow[x\mapsto M_{C',C}x]{} & K^n \end{array}$$

(iii) Wegen Proposition IV.5.5 gilt

$$M_{B,B'} M_{B',B} = D_{B,B'}(\mathrm{id}) D_{B',B}(\mathrm{id}) = D_{B,B}(\mathrm{id}) = I. \qquad \square$$

Definition IV.5.8 (Äquivalenz von Matrizen): Zwei Matrizen $A_1, A_2 \in K^{n\times m}$ heißen *äquivalent*, falls es $S \in \mathrm{Gl}_n(K)$, $T \in \mathrm{Gl}_m(K)$ gibt mit $A_2 = T A_1 S$. Das ist genau dann der Fall, wenn es zwei lineare Abbildungen $\Phi\colon K^m \to K^n$ und Basen B, B' von K^n, sowie Basen C, C' von K^n gibt mit

$$D_{C,B}(\Phi) = A_1, \qquad D_{C',B'}(\Phi) = A_2.$$

Bemerkung IV.5.9: Aus dem Satz über die Gauß-Normalform bzw. Treppenform folgt: Zwei Matrizen $A_1, A_2 \in K^{n\times m}$ sind äquivalent genau dann, wenn sie den gleichen Rang haben, d. h. wenn $\mathrm{Rang}(A_1) = \mathrm{Rang}(A_2)$.

Kapitel V.

Endomorphismen von Vektorräumen

1. Basiswechsel für Endomorphismen

In diesem Abschnitt seien stets V ein K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\Phi\colon V \to V$ eine lineare Abbildung.

Bemerkung V.1.1: Sei $A = D_{B,B}(\Phi)$. Ist $B' = (b'_1, \ldots, b'_n)$ eine weitere geordnete Basis, so ergibt sich aus Proposition IV.5.5:

$$D_{B',B'}(\Phi) = M_{B',B} D_{B,B}(\Phi) M_{B',B}^{-1}.$$

Definition V.1.2 (Ähnlichkeit von Matrizen): Zwei Matrizen $A_1, A_2 \in K^{n\times n}$ heißen *ähnlich*, wenn es $S \in \mathrm{Gl}_n(K)$ gibt mit $A_2 = SA_1S^{-1}$. In diesem Fall schreiben wir auch $A_1 \sim A_2$. A_1 und A_2 sind genau dann ähnlich, wenn sie Abbildungsmatrizen derselben linearen Abbildung Φ bezüglich möglicherweise unterschiedlichen Basen B, B' sind.

Bemerkung V.1.3: (i) Ähnlichkeit von Matrizen ist eine Äquivalenzrelation. Die zugehörigen Äquivalenzklassen werden als „Ähnlichkeitsklassen“ bezeichnet.

(ii) Wenn $A_1 = \lambda I_n$ und $A_2 \sim A_1$, dann ist schon $A_1 = A_2$, denn es ist $S(\lambda I_n)S^{-1} = \lambda S I_n S^{-1} = \lambda I_n$.

(iii) Ist K unendlich und $n \geq 1$, dann gibt es unendlich viele Ähnlichkeitsklassen.

(iv) Ist A ähnlich zu B, so ist A äquivalent zu B (vergleiche Definition IV.3.5).

(v) Aus (iv) folgt insbesondere: $\mathrm{Rang}(A)$ ist eine Ähnlichkeitsinvariente, d. h. ist $A_1 \sim A_2$, dann ist $\mathrm{Rang}(A_1) = \mathrm{Rang}(A_2)$.

2. Φ-invariante Unterräume

In diesem Abschnitt sei V ein K-Vektorraum und $\Phi\colon V \to V$ eine lineare Abbildung.

Definition V.2.1: Ein Untervektorraum U von V heißt Φ-*invariant*, falls $\Phi(U) \subseteq U$.

In diesem Fall ist $\Phi|_U\colon U \to U$ ein Endomorphismus von U.

Bemerkung V.2.2: Sei V endlichdimensional und U ein Φ-invarianter Untervektorraum. Wähle eine geordnete Basis $B_U = (b_1, \ldots, b_e)$ von U und ergänze sie zu einer Basis $B = (b_1, \ldots, b_e, c_1, \ldots, c_f)$ von V. Dann gelten:

(i)
$$D_{B,B}(\Phi) = \begin{pmatrix} A_1 & M \\ \mathbf{0}_{f\times e} & A_2 \end{pmatrix}$$
mit $A_1 \in K^{e\times e}$, $A_2 \in K^{f\times f}$ und $M \in K^{e\times f}$. $\mathbf{0}_{f\times e}$ bezeichne die Nullmatrix in $K^{f\times e}$.

(ii) In (i) ist $A_1 = D_{B_U,B_U}(\Phi|_U)$.

Proposition V.2.3: *In der Situation von Bemerkung V.2.2 gilt:*

(i) *Wir erhalten eine wohldefinierte lineare Abbildung*
$$\overline{\Phi}\colon V/U \longrightarrow V/U, \qquad [v] \longmapsto [\Phi(v)].$$

(ii) *Betrachte die Basis* $C = \{c_1+U, \ldots, c_f+U\}$ *von* V/U. *Dann gilt* $D_C(\overline{\Phi}) = A_2$.

Beweis: (i) Es sei $\pi\colon V \to V/U$ die kanonische Projektion. Per Definition ist $[\Phi(v)] = (\pi \circ \Phi)(v)$. Da $\mathrm{Kern}(\pi \circ \Phi) \supseteq U$ ist $\overline{\Phi}$ nach Satz 10 eine wohldefinierte lineare Abbildung.

(ii) Folgt aus der Definition von $\overline{\Phi}$. □

Bemerkung V.2.4: Sei V endlichdimensional und U, W seien Φ-invariante Untervektorräume von V mit $V = U \oplus W$. Wähle geordnete Basen $B_U = (b_1, \ldots, b_e)$ von U und $B_W = (c_1, \ldots, c_f)$ von W. Wie in Aufgabe 4 auf Blatt 11 gezeigt, ist $(b_1, \ldots, b_e, c_1, \ldots, c_f)$ eine Basis von V. Dann ist
$$D_{B,B}(\Phi) = \begin{pmatrix} A_1 & \mathbf{0}_{e\times f} \\ \mathbf{0}_{f\times e} & A_2 \end{pmatrix}$$

mit $A_1 = D_{B_U,B_U}(\Phi|_U)$ und $A_2 = D_{B_W,B_W}(\Phi|_W)$.

Wenn wir V also als direkte Summe von möglichst kleinen Φ-invarianten Unterräumen schreiben können, dann erhalten wir eine Abbildungsmatrix von möglichst einfacher Form.

3. Eigenvektoren und Eigenwerte

In diesem Abschnitt seien V ein K-Vektorraum und $\Phi\colon V \to V$ eine lineare Abbildung.

Definition V.3.1 (Eigenvektor/Eigenwert eines Endomorphismus):

(i) Ein Vektor $\mathbf{0} \neq v \in V$ heißt *Eigenvektor von* Φ, falls es $\lambda \in K$ gibt mit $\Phi(v) \in \lambda v$. Das heißt genau: $\langle v \rangle$ ist ein eindimensionaler Φ-invarianter Untervektorraum.

(ii) Eine Zahl $\lambda \in K$ heißt *Eigenwert von* Φ, falls es $\mathbf{0} \neq v$ gibt mit $\Phi(v) = \lambda v$.

(iii) Die Menge $\mathrm{Spec}(\Phi) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } \Phi\}$ heißt *Spektrum von* Φ.

(iv) Für $\lambda \in K$ heißt $\mathrm{Eig}(\Phi, \lambda) := \{v \in V \mid \Phi(v) = \lambda v\}$ der *Eigenraum von* λ.

Bemerkung V.3.2: (i) Es gilt genau dann $\lambda \in \mathrm{Spec}(\Phi)$, wenn $\mathrm{Eig}(\Phi, \lambda) \supsetneq \{\mathbf{0}\}$.

(ii) Es gilt $\Phi(v) = \lambda v$ genau dann, wenn $\Phi(v) - \lambda v = \mathbf{0}$, d. h. es gilt $\Phi(v) = \lambda v$ genau dann, wenn

$$(\Phi - \lambda \,\mathrm{id})(v) = \mathbf{0}$$

Wir definieren $\Psi_\lambda := \Phi - \lambda \,\mathrm{id}\colon V \to V, v \mapsto \Phi(v) - \lambda v$. Ψ_λ ist eine lineare Abbildung und $\mathrm{Eig}(\Phi, \lambda) = \mathrm{Kern}(\Psi_\lambda)$; insbesondere ist $\mathrm{Eig}(\Phi, \lambda)$ ein Untervektorraum von V.

Definition V.3.3 (Eigenvektor/Eigenwert einer Matrix): Es seien $A \in K^{n\times n}$ und $V = K^n$.

(i) Ein Vektor $\mathbf{0} \neq v \in V$ heißt *Eigenvektor von* A, falls es $\lambda \in K$ mit $Av = \lambda v$ gibt.

(ii) Eine Zahl $\lambda \in K$ heißt *Eigenwert von* A, falls es $\mathbf{0} \neq v \in V$ mit $Av = \lambda v$ gibt.

(iii) Die Menge $\mathrm{Spec}(A) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } A\}$ heißt *Spektrum von* A.

(iv) Für $\lambda \in K$ heißt $\mathrm{Eig}(A, \lambda) := \{v \in V \mid Av = \lambda v\}$ heißt *Eigenraum zu* λ.

Bemerkung V.3.4: Seien V n-dimensional, B eine geordnete Basis V, $\Phi\colon V \to V$ wie gehabt eine lineare Abbildung und $A = D_{B,B}(\Phi) \in K^{n\times n}$. Dann sind

(i) $\mathrm{Spec}(\Phi) = \mathrm{Spec}(A)$,

(ii) Es gilt $v \in \mathrm{Eig}(\Phi, \lambda)$ genau dann, wenn $D_B(v) \in \mathrm{Eig}(A, \lambda)$.

(iii) Für Eigenraum $\mathrm{Eig}(A, \lambda)$ gilt: $\mathrm{Eig}(A, \lambda) = \mathbb{L}(A - \lambda I_n|\mathbf{0}) =: \mathrm{Kern}(A - \lambda I_n)$.

Beispiel V.3.5: (i) Sei A die Diagonalmatrix $A = (\lambda_1, \ldots, \lambda_n)$. Für die Standardbasis $E = \{e_1, \ldots, e_n\}$ gilt $Ae_i = \alpha_i e_i$, d. h. $\lambda_1, \ldots, \lambda_n$ sind die Eigenwerte. Ist $\lambda \in K \setminus \{\lambda_1, \ldots, \lambda_n\}$, dann ist $A - \lambda I_n = \mathrm{diag}(\lambda_1 - \lambda, \ldots, \lambda_n - \lambda)$ invertierbar, d. h. $\mathrm{Kern}(A - \lambda I_n) = \{\mathbf{0}\}$ und λ ist kein Eigenwert von A. Es ist also $\mathrm{Spec}(A) = \{\lambda_1, \ldots, \lambda_n\}$.

(ii) Sei $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ wie in (Beispiel IV.5.2) die Spiegelung an der Diagonalen $y = x$. Dann ist

$$D_{B,B}(\Phi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

mit $B = \{(1, 1)^t, (1, -1)^t\}$, d. h. nach (i) ist $\mathrm{Spec}(\Phi) = \{1, -1\}$.

Proposition V.3.6: *Sei $A \in K^{n\times n}$. Dann sind* $\mathrm{Spec}(A)$ *und die Dimensionen der Eigenräume Ähnlichkeitsinvarianten.*

Beweis: Sei $B = SAS^{-1}$ mit $S \in \mathrm{Gl}_n(K)$. Die Zahl λ ist Eigenwert von B zum Eigenvektor $\mathbf{0} \neq v$ genau dann, wenn $Bv = \lambda v$, per Definition von B also

$$Bv = \lambda v \Leftrightarrow SAS^{-1}v = \lambda v \Leftrightarrow AS^{-1}v = \lambda S^{-1}v,$$

d. h. λ ist Eigenwert von B zum Eigenvektor v genau dann, wenn λ Eigenwert von A zum Eigenvektor $S^{-1}v \neq \mathbf{0}$ ist. Somit gilt bereits $\mathrm{Spec}(A) = \mathrm{Spec}(B)$ und $\mathrm{Eig}(B, \lambda) = \{Sv \mid v \in \mathrm{Eig}(A, \lambda)\} = S\,\mathrm{Eig}(A, \lambda)$.

Zu den Dimensionen der Eigenräumen: Da $S \in \mathrm{Gl}_n(K)$ wissen wir von Blatt 10, Aufgabe 2, dass $\dim_K(\mathrm{Eig}(B, \lambda)) = \dim_K(\mathrm{Eig}(A, \lambda))$. □

Satz 13 (Eigenräume bilden direkte Summe): *Es seien $\lambda_1, \ldots, \lambda_k \in K$ verschiedene Eigenwerte von Φ und* $\mathrm{Eig}(\Phi, \lambda_1), \ldots, \mathrm{Eig}(\Phi, \lambda_k)$ *die zugehörigen Eigenräume. Dann ist die Summe der Eigenräume eine direkte Summe, d. h.*

$$\sum_{i=1}^{k} \mathrm{Eig}(\Phi, \lambda_i) = \bigoplus_{i=1}^{k} \mathrm{Eig}(\Phi, \lambda_i).$$

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 $u_i \in \mathrm{Eig}(\Phi, \lambda_i)$. 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 $u_1, \ldots, u_{k-1} = \mathbf{0}$, also insgesamt $u_1, \ldots, u_k = \mathbf{0}$ und die Summe ist direkt. □

Korollar V.3.7: *Ist* $\dim(V) = n$*, dann gilt* $\#(\mathrm{Spec}(\Phi)) \leq n$.

Beweis: Angenommen, $\lambda_1, \ldots, \lambda_{n+1}$ wären $n+1$ verschiedene Eigenwerte von Φ. Dann wäre

$$\dim(V) \geq \dim\left(\bigoplus_{i=1}^{n+1} \mathrm{Eig}(\Phi, \lambda_i)\right) \geq n+1,$$

ein offensichtlicher Widerspruch. □

Definition V.3.8 (Diagonalisierbarkeit):

(i) Eine Matrix $A \in K^{n\times n}$ heißt *diagonalisierbar*, falls es $\lambda_1, \ldots, \lambda_n \in K$ und $S \in \mathrm{Gl}_n(K)$ gibt mit $SAS^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

(ii) Eine lineare Abbildung Φ heißt *diagonalisierbar*, falls es eine Basis B von V gibt, sodass für $b \in B$ gilt, dass b Eigenvektor von Φ ist.

Bemerkung V.3.9: Ist $\dim(V) = n$, dann gilt: Φ ist diagonalisierbar genau dann, wenn es eine geordnete Basis B von V und Skalare $\lambda_1, \ldots, \lambda_n \in K$ gibt, sodass $D_{B,B}(\Phi) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Es gibt genau dann eine solche Basis B von V, wenn $V = \bigoplus_{\lambda\in\mathrm{Spec}(\Phi)} \mathrm{Eig}(\Phi, \lambda)$. Da die Summe der Eigenräume immer direkt ist, heißt das, dass dies genau dann eintritt, wenn $\sum_{\lambda\in\mathrm{Spec}(\Phi)} \dim(\mathrm{Eig}(\Phi, \lambda)) = n$ ist.

Beweis: Die erste Äquivalenz gilt nach Definition.

Für die zweite Äquivalenz: „⇒“: Sei $B = (b_1, \ldots, b_n)$ eine Basis aus Eigenvektoren, d. h. ist $v \in V$, dann ist v Linearkombination von Eigenvektoren von Φ, also $v \in \bigoplus_{\lambda\in\mathrm{Spec}(\Phi)} \mathrm{Eig}(\Phi, \lambda)$.

„⇐“: Sei $\mathrm{Spec}(\Phi) = \{\lambda_1, \ldots, \lambda_k\}$ (es ist $k \leq n$) und wähle Basen B_i mit B_i ist Basis von $\mathrm{Eig}(\Phi, \lambda_i)$. Dann ist $B = \bigcup_{i=1}^{k} B_i$ eine Basis von V aus Eigenvektoren.

Für die dritte Äquivalenz: Betrachte die Dimensionen. □

4. Die Determinante

In diesem Abschnitt sei stets $A \in K^{n\times n}$.

Beispiel V.4.1: In der vierten Aufgabe von Blatt 4 haben wir gesehen: Die Matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in K^{2\times 2}$ ist invertierbar genau dann, wenn $ad - bc \neq 0$.

Definition V.4.2 (Determinante für 2×2-Matrizen): Für $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in K^{2\times 2}$ heißt $\det(A) := ad - bc$ die *Determinante von A*.

Bemerkung V.4.3: Die Abbildung

$$\det\colon K^2 \times K^2 \longrightarrow K, \qquad \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right] \longmapsto \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

hat folgende Eigenschaften:

(i) Ist $(a', c')^t \in K^2$, dann ist

$$\begin{aligned} \det\left[\begin{pmatrix} a \\ c \end{pmatrix} + \begin{pmatrix} a' \\ c' \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right] &= (a + a')d - b(c + c') \\ &= ad - bc + a'd - bc' \\ &= \det\left[\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right] + \det\left[\begin{pmatrix} a' \\ c' \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right], \end{aligned}$$

analog folgt $\det[\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right), \left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right) + \left(\begin{smallmatrix} b' \\ d' \end{smallmatrix}\right)] = \det[\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right), \left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)] + \det[\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right), \left(\begin{smallmatrix} b' \\ d' \end{smallmatrix}\right)]$.

(ii) Für $\lambda \in K$ gilt

$$\begin{aligned} \det\left[\lambda\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right] &= \lambda ad - b\lambda c \\ &= \lambda(ad - bc) \\ &= \lambda \det\left[\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right] = \det\left[\begin{pmatrix} a \\ c \end{pmatrix}, \lambda\begin{pmatrix} b \\ d \end{pmatrix}\right]. \end{aligned}$$

(iii) Es gilt $\det[\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right), \left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)] = ac - ac = 0$.

(iv) Es ist $\det[\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)] = 1$.

Definition V.4.4: Eine Abbildung $D\colon (K^n)^n \to K$ heißt *Determinantenform* auf K^n, falls für alle $v_1, \ldots, v_n \in K^n$ gilt:

(i) Für $v_i' \in K^n$ mit $1 \leq i \leq n$ ist

$$\begin{aligned} &D(v_1, \ldots, v_{i-1}, v_i + v_i', v_{i+1}, \ldots, v_n) \\ &\quad = D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + D(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n), \end{aligned}$$

(ii) Für $\lambda \in K$ und $1 \leq i \leq n$ 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) 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) Für die geordnete Standardbasis $(e_1, \ldots, e_n)$ ist $D(e_1, \ldots, e_n) = 1$.

Sei $A = (v_1 | \ldots | v_n) \in K^{n \times n}$. Wir schreiben im Folgenden $D(A) := D(v_1 | \ldots | v_n)$.

Bemerkung V.4.5: Die Abbildung det aus Bemerkung V.4.3 ist eine Determinantenform auf $K^{2\times 2}$.

Ab jetzt sei D eine Determinantenform auf K^n.

Bemerkung V.4.6: (i) Seien $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in K^n$ fest. Dann ist die Abbildung

$$K^n \longrightarrow K, \qquad v \longmapsto D(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

eine lineare Abbildung.

(ii) „Addition des λ-fachen der i-ten Spalte zur j-ten Spalte“: Für $\lambda \in K$ und $i \neq j$ gilt

$$\begin{aligned} &D(v_1, \ldots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \ldots, v_n) \\ &\quad = D(v_1, \ldots, v_n) + \lambda D(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_n) = D(v_1, \ldots, v_n). \end{aligned}$$

(iii) „Vertauschen von Spalten“: Für $i < j$ gilt

$$\begin{aligned} &D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\ &\qquad = -D(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n), \end{aligned}$$

denn

$$\begin{aligned} 0 &= D(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\ &= D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\ &\quad + D(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\ &= D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\ &\quad + D(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n), \end{aligned}$$

also die Behauptung.

(iv) „Multiplizieren von Spalten mit Skalaren“: Für $\alpha_1, \ldots, \alpha_n \in K$ gilt

$$D(\alpha_1 v_1, \alpha_2 v_2, \ldots, \alpha_n v_n) = \alpha_1 \cdots \alpha_n D(v_1, \ldots, v_n).$$

Proposition V.4.7: *Seien $1 \leq i, j \leq n$ mit $i \neq j$ und $\alpha_1, \alpha_2, \ldots, \alpha_n \in K$. Dann gelten:*

(i) $D(A \cdot A_{i,j}^{\alpha}) = D(A)$,

(ii) $D(A \cdot V_{i,j}) = -D(A)$,

(iii) $D(A \cdot \mathrm{diag}(\alpha_1, \ldots, \alpha_n)) = \alpha_1 \cdots \alpha_n D(A)$.

Beweis: Folgt direkt aus Bemerkung V.4.6. □

Korollar V.4.8: (i) $D(A_{i,j}^{\alpha}) = 1$,

(ii) $D(V_{i,j}) = -1$,

(iii) $D(\mathrm{diag}(\alpha_1, \ldots, \alpha_n)) = \prod_{i=1}^{n} \alpha_i$.

Beweis: Verwende Proposition V.4.7 für $A = I_n$. □

Notation V.4.9: Im Folgenden bezeichnen wir $X \in K^{n \times n}$ als spezielle Matrix, falls X Additions- oder Vertauschungs- oder Diagonalmatrix ist.

Korollar V.4.10: *Ist X eine spezielle Matrix, dann ist $D(A \cdot X) = D(A) \cdot D(X)$.*

Erinnerung: (i) Nach Satz 3 gibt es für jede Matrix $A \in K^{n \times n}$ spezielle Matrixen $X_1, \ldots, X_k \in \mathrm{Gl}_n(K)$, sodass das Produkt

$$X_1 \cdots X_k A =: T$$

Treppenform hat.

(ii) Es ist $\mathrm{Rang}(A) = n$ genau dann, wenn $\mathrm{Rang}(T) = n$. T hat $\mathrm{Rang}(T) = n$ genau dann, wenn $T = I_n$. $T = I_n$ gilt genau dann, wenn $A \in \mathrm{Gl}_n(K)$.

Proposition V.4.11: *Seien $X_1, \ldots, X_k$ spezielle Matrizen, sodass $X_1 \cdots X_k \cdot A^t =: T$ in Treppenform ist.*

(i) *Es gilt $\mathrm{Rang}(A) < n$ genau dann, wenn $D(A) = 0$,*

(ii) *Ist $\mathrm{Rang}(A) = n$, dann ist $D(A) = D(X_1)^{-1} \cdots D(X_k)^{-1}$.*

Beweis: Wir verwenden, dass $A X_k^t \cdots X_1^t = T^t$.

(ii) Es gilt $\mathrm{Rang}(A) = n$ genau dann, wenn $\mathrm{Rang}(A^t) = n$ (vergleiche Proposition IV.5.4), d. h. in dieser Situation ist $T = T^t = I_n$, wir sind also in der Situation $AX_k^t \cdots X_1^t = I_n$. Jetzt liefert Proposition V.4.7, dass

$$1 = D(I_n) = D(AX_k^t \cdots X_1^t) = D(A) \cdot D(X_k^t) \cdots D(X_1^t),$$

Korollar V.4.8 liefert jetzt $D(X_i^t) = D(X_i)$, also die Behauptung.

(i) Die Implikation „$\mathrm{Rang}(A) = n \Rightarrow D(A) \neq 0$“ folgt aus Teil (ii). Sei nun $\mathrm{Rang}(A) < n$ und somit $\mathrm{Rang}(A^t) < n$. d. h. die letzte Zeile in T ist eine Nullzeile und die letzte Spalte in T^t ist eine Nullspalte. Wir können diesen Umstand beschreiben durch $T^t = T^t \cdot \mathrm{diag}(1, \ldots, 1, 0)$. Jetzt ist

$$\begin{aligned} D(A)D(X_k) \cdots D(X_1) &= D(A)D(X_k^t) \cdots D(X_1^t) \\ &= D(A \cdot X_k^t \cdots X_1^t) \\ &= D(T^t) \\ &= D(T^t \cdot \mathrm{diag}(1, \ldots, 1, 0)) \\ &= D(T^t) \cdot D(\mathrm{diag}(1, \ldots, 1, 0)) = 0. \end{aligned}$$

Da alle $D(X_i) \neq 0$ sind, folgt die Behauptung. □

Korollar V.4.12 (wichtige Eigenschaften der Determinantenform):

(i) *Es gilt $D(A) \neq 0$ genau dann, wenn $A \in \mathrm{Gl}_n(K)$. Es gilt $A \in \mathrm{Gl}_n(K)$ genau dann, wenn $\mathrm{Rang}(A) = n$. Es gilt $\mathrm{Rang}(A) = n$ genau dann, wenn die Spaltenvektoren von A linear unabhängig sind. Die Spaltenvektoren von A sind linear unabhängig genau dann, wenn die Zeilenvektoren von A linear unabhängig sind.*

(ii) *Es gibt höchstens eine Determinantenform auf K^n.*

(iii) *Die Determinantenform ist* multiplikativ*, d. h. für Matrizen $A_1, A_2 \in K^{n \times n}$ ist $D(A_1 \cdot A_2) = D(A_1) \cdot D(A_2)$.*

(iv) *Ist $A \in \mathrm{Gl}_n(K)$, dann ist $D(A^{-1}) = (D(A))^{-1}$.*

(v) *Es ist $D(A^t) = D(A)$.*

Beweis: Alle Behauptungen folgen aus Proposition V.4.11. Für (v) verwende, dass $A \in \mathrm{Gl}_n(K)$ gilt genau dann, wenn $A^t \in \mathrm{Gl}_n(K)$. Ist A invertierbar, so ist $A^t = X_1^{-1} \cdots X_k^{-1}$, d. h. $A = (X_1^{-1})^t \cdots (X_1^k)^{-1}$. Die Behauptung folgt dann aus $D(X_i^t) = D(X_i)$ nach Korollar V.4.8. □

Korollar V.4.13: *Die Rechenregeln* (i) - (iii) *aus Definition V.4.4 gelten analog für Zeilenvektoren, das heißt für* $A = (z_1, \ldots, z_n)^t$ *mit* $z_i \in K^{1\times n}$ *erhalten wir:*

(i) *Ist* $z_i' \in K^{1\times n}$*, dann ist*

$$D\left[\begin{pmatrix} z_1 \\ \vdots \\ z_i + z_i' \\ \vdots \\ z_n \end{pmatrix}\right] = D\left[\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix}\right] + D\left[\begin{pmatrix} z_1 \\ \vdots \\ z_i' \\ \vdots \\ z_n \end{pmatrix}\right],$$

(ii) *Ist* $\lambda \in K$*, so ist*

$$D\left[\begin{pmatrix} z_1 \\ \vdots \\ \lambda z_i \\ \vdots \\ z_n \end{pmatrix}\right] = \lambda D\left[\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix}\right],$$

(iii) *Ist* $z_i = z_j$ *für* $1 \leq i, j \leq n$ *und* $i \neq j$*, so ist* $D(A) = 0$.

5. Die Leibnizformel

In diesem Abschnitt sei $A \in K^{n\times n}$ eine Matrix.

Definition V.5.1 (Leibniz-Formel): Die Zahl

$$\det(A) := \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

heißt die *Determinante von A*.

Beispiel V.5.2: Für eine 2×2-Matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ ergibt die Leibniz-Formel

$$\det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = ad - bc$$

wegen $S_2 = \{\mathrm{id}, (1,2)\}$ und $\operatorname{sgn}(\mathrm{id}) = 1$, $\operatorname{sgn}(1,2) = -1$.

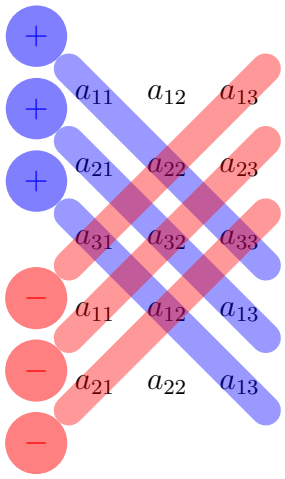
Beispiel V.5.3 (Regel von Sarrus): Für eine 3×3-Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

ist

$$\det(A) = a_{1,1}a_{2,2}, a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{3,2}a_{2,1} \\ - a_{1,3}a_{2,2}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{2,3}a_{3,2},$$

da $\operatorname{sgn}(\mathrm{id}) = \operatorname{sgn}(1,2,3) = \operatorname{sgn}(1,3,2) = 1$, $\operatorname{sgn}(1,3) = \operatorname{sgn}(1,2) = \operatorname{sgn}(2,3) = -1$. Man kann sich die Regel von Sarrus auch merken über folgendes Schema[1]:



Satz 14 (Regel von Leibniz): *Die Abbildung*

$$(K^n)^n \longrightarrow K, \qquad (v_1, \ldots, v_n) \longmapsto \det[(v_1|\ldots|v_n)]$$

ist eine Determinantenform.

Beweis: Wir überprüfen die Eigenschaften (i) - (iv) aus Definition V.4.4. Schreibe dazu $A := (v_1|\ldots|v_n)$.

(i) Seien $1 \leq l \leq n$ und $A' := (v_1|\ldots|v_{l-1}|v_l + v_l'|v_{l+1}|\ldots|v_n)$ mit einem Vektor $v_l' = (b_1, \ldots, b_n)^t \in K^n$. Dann ist

$$\begin{aligned}
\det(A') &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a'_{i,\sigma(i)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\substack{i=1 \\ \sigma(i) \neq l}}^{n} a_{i,\sigma(i)} \cdot (a_{\sigma^{-1}(l),l} + b_{\sigma^{-1}(l)}) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\substack{i=1 \\ \sigma(i) \neq l}}^{n} a_{i,\sigma(i)} b_{\sigma^{-1}(l)} \\
&= \det(A) + \det(A'')
\end{aligned}$$

mit $A'' = (v_1|\ldots|v_{l-1}|v_l'|v_{l+1}|\ldots|v_n)$.

[1] Die Skizze zum Schema stammt aus dieser Quelle.

(ii) Sei $A' = (v_1|\dots|v_{l-1}|\lambda v_l|v_{l+1}|\dots|v_n)$ mit $\lambda \in K$. Der Faktor λ tritt in $\det(A')$ in jedem Summanden genau einmal auf, d. h. $\det(A') = \lambda\det(A)$.

(iii) Sei $v_k = v_l$ mit $k \neq l$ und setze $\sigma_0 := (k,l) \in S_n$. 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

$$\sigma'(i) = \begin{cases} \sigma(i), & \text{falls } i \notin \{k,l\}, \\ \sigma(l), & \text{falls } i = k, \\ \sigma(k), & \text{falls } i = l. \end{cases}$$

Da $v_k = v_l$ 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) Es ist

$$\det(I_n) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)}.$$

Da $a_{i,j} = \delta_{i,j}$ ist $\prod_{i=1}^n a_{i,\sigma(i)} = 0$ außer für $\sigma = \mathrm{id}$, $\det(I_n) = 1 \cdot a_{1,1} \cdots a_{n,n} = 1$ gilt also. □

Die Determinantenform $(v_1, \dots, v_n) \mapsto \det[(v_1|\dots|v_n)]$ ist also die eindeutig bestimmte Determinantenform auf K^n und es gelten alle Rechenregeln aus (Abschnitt V.4). Insbesondere haben wir die wichtige Eigenschaft, dass die Determinante einer Matrix ungleich 0 ist genau dann, wenn die Matrix invertierbar ist.

Bemerkung V.5.4: det ist eine Ähnlichkeitsinvariante, denn für $A, B \in K^{n\times n}$ und $S \in \mathrm{Gl}_n(K)$ mit $B = SAS^{-1}$ ist

$$\begin{aligned} \det(B) = \det(SAS^{-1}) &= \det(S)\det(A)\det(S^{-1}) \\ &= \det(S)\det(A)\det(S)^{-1} = \det(A). \end{aligned}$$

Definition V.5.5 (Determinante eines Endomorphismus): Seien V ein K-Vektorraum der Dimension n, B eine geordnete Basis und $\Phi\colon V \to V$ eine lineare Abbildung. Definiere dann

$$\det(\Phi) = \det(D_{B,B}(\Phi)).$$

Die Definition der Determinante eines Endomorphismus hängt nach Bemerkung V.5.4 nicht von der gewählten Basis B ab.

Proposition V.5.6 (Determinante und Eigenwerte): *Sei $\lambda \in K$.*

(i) *λ ist Eigenwert von A genau dann, wenn* $\det(A - \lambda I_n) = 0$.

(ii) *In der Situation von Definition V.5.5 gilt: λ ist Eigenwert von Φ genau dann, wenn* $\det(\Phi - \lambda \operatorname{id}_V) = 0$.

Beweis: (i) λ ist genau dann ein Eigenwert von A, wenn $\operatorname{Kern}(A - \lambda I_n) \neq \{0\}$. Wegen der Dimensionsformel ist aber $\operatorname{Kern}(A - \lambda I_n) \neq \{0\}$ genau dann, wenn $A - \lambda I_n \notin \mathrm{Gl}_n(K)$, was genau dann gilt, wenn $\det(A - \lambda I_n) = 0$.

(ii) Folgt aus (i). □

Definition V.5.7: In der Situation von Definition V.5.5

(i) $\mathrm{CP}_A(X) := \det(A - XI_n) \in K[X]$ heißt *charakteristisches Polynom von A*.

(ii) $\mathrm{CP}_\Phi(X) := \det(\Phi - X \operatorname{id}_V) \in K[X]$ heißt *charakteristisches Polynom von Φ*.

Beispiel V.5.8: Es sei $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Dann ist

$$\begin{aligned}\mathrm{CP}_A(X) &= \det(A - XI_2) \\ &= \det\left[\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\right] = \det\begin{pmatrix} 1-X & 1 \\ 0 & 1-X \end{pmatrix} = (1-X)^2,\end{aligned}$$

d. h. $\operatorname{Spec}(A) = \{1\}$.

Bemerkung V.5.9: In Definition V.5.7 verwenden wir Matrizen über dem Ring $K[X]$. Es gilt: $A - XI_n \in (K[X])^{n\times n}$, Determinanten lassen sich in analoger Weise über Ringen definieren. Ist R ein Ring und $A \in R^{n\times n}$, dann gilt: Es ist $\det(A) \in R^\times$ genau dann, wenn $A \in \mathrm{Gl}_n(R)$.

Die Eigenwerte von A sind genau die Nullstellen des charakteristischen Polynoms von A.

6. Regel von Laplace

In diesem Abschnitt sei $A \in K^{n\times n}$.

Definition V.6.1: Für $1 \leq i, j \leq n$ sei $A_{i,j} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte hervorgeht.

Beispiel V.6.2: Es sei

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

Dann sind

$$A_{1,2} = \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix}, \qquad A_{2,2} = \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix}, \qquad A_{3,2} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

Satz 15 (Laplace): *Sei $k \in \{1, \ldots, n\}$. Es gelten folgende beide Rechenregeln für die Determinante:*

(i) *„Laplace-Entwicklung nach der k-ten Zeile“: Es ist*

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+k} a_{k,j} \det(A_{k,j}).$$

(ii) *„Laplace-Entwicklung nach der k-ten Spalte“: Es ist*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{i,k} \det(A_{i,k}).$$

Beispiel V.6.3: Sei A die Matrix aus Beispiel V.6.2. Für die Vorzeichen aus den Formeln in Satz 15 ist folgendes Schema nützlich:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

Entwickeln nach der zweiten Spalte ergibt

$$\begin{aligned} \det(A) &= (-1) \cdot 2 \cdot \det \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} + (-1) \cdot 1 \cdot \det \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \\ &= -12. \end{aligned}$$

Der Beweis von Satz 15 erfordert etwas Vorarbeit.

Proposition V.6.4 (Determinante von Dreiecksmatrizen): *Sei A eine* obere Dreiecksmatrix*, d. h. A sei von der Form*

$$A = \begin{pmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & d_n \end{pmatrix}.$$

Dann gilt $\det(A) = d_1 \cdots d_n$.

Die selbe Aussage gilt für untere Dreiecksmatrizen, die in naheliegender Weise definiert sind. Der Beweis der Aussage Proposition V.6.4 ist Übungsaufgabe auf Blatt 13.

Proposition V.6.5 (Determinante für Blockmatrizen): *Seien $1 \leq k \leq n$, $X \in K^{k\times k}$, $Y \in K^{(n-k)\times k}$, $Z \in K^{(n-k)\times(n-k)}$. Dann ist*

$$\det \begin{pmatrix} X & \mathbf{0} \\ Y & Z \end{pmatrix} = \det(X)\det(Z).$$

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 ist nach (Korollar IV.4.?) auch $\det\left(\begin{smallmatrix} X & \mathbf{0} \\ Y & Z \end{smallmatrix}\right) = 0$. Ist $\det(X) \neq 0$, dann ist $X \in \mathrm{Gl}_{n-k}(K)$. Nach Proposition V.6.4 ist für die Matrix

$$H := \begin{pmatrix} I_k & \mathbf{0} \\ -YX^{-1} & I_{n-k} \end{pmatrix}$$

schon $\det(H) = 1$. 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.6.6 (Laplace für eine Zeile): Sei $A \in K^{n\times n}$. Schreibe A also

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ s_1 & s_2 & \cdots & s_n \end{pmatrix}$$

mit $a_{1,i} \in K$ und $s_i \in K^{n-1}$. Dann gilt

$$\begin{aligned}
\det(A) &= \det \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ s_1 & s_2 & \cdots & s_n \end{pmatrix} + \det \begin{pmatrix} 0 & a_{1,2} & 0 & \cdots & 0 \\ s_1 & s_2 & s_3 & \cdots & s_n \end{pmatrix} \\
&\quad + \cdots + \det \begin{pmatrix} 0 & \cdots & 0 & a_{1,n} \\ s_1 & \cdots & s_{n-1} & s_n \end{pmatrix} \\
&= \det \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ s_1 & & A_{1,1} & \end{pmatrix} - \det \begin{pmatrix} a_{1,2} & 0 & \cdots & 0 \\ s_2 & & A_{1,2} & \end{pmatrix} \\
&\quad + \cdots + (-1)^{n+1} \det \begin{pmatrix} a_{1,n} & 0 & \cdots & 0 \\ s_n & & A_{1,n} & \end{pmatrix} \\
&= \sum_{j=1}^{n} (-1)^{1+j} a_{1,j} \det(A_{i,j})
\end{aligned}$$

wobei wir bei der letzten Gleichheit die Aussage aus Proposition V.6.5 verwendet haben.

Beweis (von Satz 15): (i) Schreibe $A = (z_1, \ldots, z_n)^t$. Dann gilt durch $(i-1)$-maliges Vertauschen von Zeilen, dass

$$\det(A) = (-1)^{i-1} \det(z_i, z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n)^t$$

und die Behauptung folgt mit der vorangegangenen Bemerkung.

(ii) Transponieren in (i). □

Definition V.6.7 (Adjunkte): Definiere die Matrix $A^{\#} = (\beta_{i,j}) \in K^{n\times n}$ durch $\beta_{i,j} := (-1)^{i+1} \det(A_{j,i})$. Dann heißt $A^{\#}$ die *Adjunkte von A*.

Proposition V.6.8:

(i) *Es git* $AA^{\#} = \det(A)I_n$. *Insbesondere also: Ist* $A \in \mathrm{Gl}_n(K)$, *so ist*

$$A^{-1} = \frac{1}{\det(A)} A^{\#}.$$

(ii) *Es gilt die* Cramer'sche Regel*: Sei* $A = (s_1|\cdots|s_n) \in \mathrm{Gl}_n(K)$ *und* $b \in K^n$. *Dann gilt* $Ax = b$ *genau dann, wenn* $x = (x_j)$ *mit*

$$x_j = \frac{\det(A_j)}{\det(A)}$$

wobei $A_j = (s_1|\ldots|s_{j-1}|b|s_{j+1}|\ldots|s_n)$.

Der Beweis dieser Aussage ist Übungsaufgabe auf dem nächsten Übungsblatt.

Kapitel VI.

Skalarprodukte

1. Bilinearformen

In diesem Abschnitt sei V stets ein K-Vektorraum.

Definition VI.1.1 (Multilinearform): Eine Abbildung

$$\beta\colon \prod_{i=1}^{m} V \longrightarrow K$$

heißt *m-fache Multilinearform*, falls für alle $1 \leq i \leq m$ gilt: Für alle $v_j \in V$ mit $1 \leq j \leq m$ und $i \neq j$ ist die Abbildung

$$\beta_j\colon V \longrightarrow K, \qquad v \longmapsto \beta(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

eine lineare Abbildung.

Beispiel VI.1.2: Die Determinantenabbildung ist eine n-fache Multilinearform auf dem Vektorraum K^n.

Definition VI.1.3 (Bilinearform): In der Situation von Definition VI.1.1:

(i) Ist $m = 2$, so heißt β *Bilinearform*, d. h. $\beta\colon V \times V \to K$ ist eine Bilinearform, falls für alle $v_1, v_2, w_1, w_2, v, w \in V$ und $\lambda \in K$ gelten:
   (1) $\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w)$,
   (2) $\beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2)$,
   (3) $\beta(\lambda v, w) = \lambda\beta(v, w) = \beta(v, \lambda w)$.

(ii) Eine Bilinearform β heißt *symmetrisch*, falls für alle $v, w \in V$ gilt:

$$\beta(v, w) = \beta(w, v).$$

Beispiel VI.1.4 (Einheitsform): Die Abbildung

$$\beta\colon K^n \times K^n \longrightarrow K, \qquad (x,y) \longmapsto \sum_{i=1}^{n} x_i y_i = x^t \cdot y = y^t \cdot x$$

ist eine symmetrische Bilinearform und heißt *Einheitsform.*

Bemerkung VI.1.5: (i) Für die Matrix $A \in K^{n\times n}$ ist

$$\beta_A\colon K^n \times K^n \longrightarrow K, \qquad (x,y) \longmapsto x^t A y = y^t A^t x$$

eine Bilinearform.

(ii) β_A ist symmetrisch genau dann, wenn $y^t A x = y^t A^t x$ für alle $x, y \in K^n$ gilt. Wegen $e_i^t A e_j = a_{i,j}$ ist das genau dann der Fall, wenn $A = A^t$.

(iii) Für $A = I_n$ ist β die Einheitsform.

Proposition VI.1.6: *Sei $\beta\colon K^n \times K^n \to K$ eine Bilinearform. Dann ist $\beta = \beta_A$ für $A = (a_{i,j})$ mit $a_{i,j} = \beta(e_i, e_j)$, wobei $E = (e_1, \ldots, e_n)$ die Standardbasis ist.*

Beweis: Seien $x, y \in K^n$. Dann ist

$$\begin{aligned}\beta(x,y) &= \beta\Big(\sum_{i=1}^{n} x_i e_i, \sum_{i=1}^{n} y_i e_i\Big) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \beta(e_i, e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_i a_{i,j} = x^t A y.\end{aligned}$$ □

Korollar VI.1.7 (Gram-Matrix): *Seien* $\dim(V) = n$ *und $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V. Dann heißt die Matrix $G := (g_{i,j}) \in K^{n\times n}$ mit den Einträgen $g_{i,j} := \beta(b_i, b_j)$ die* Gram-Matrix *zur Bilinearform β bezüglich der Basis B. Es gilt für alle $v, w \in V$:*

$$\beta(v,w) = D_B(v)^t \cdot G \cdot D_B(w)$$

Bemerkung VI.1.8 (Gram-Matrix unter Basiswechsel): In der Situation von Korollar VI.1.7 sei $B' = (b'_1, \ldots, b'_n)$ eine weitere geordnete Basis und G' die zu B' gehörige Gram-Matrix von β. Dann gilt

$$G' = D_{B,B'}^t \cdot G \cdot D_{B,B'}.$$

Beweis: Für $v, w \in V$ gilt

$$\begin{aligned}\beta(v,w) &= D_B(v)^t \cdot G \cdot D_B(w) \\ &= (D_{B,B'} D_{B'}(v))^t \cdot G \cdot D_{B,B'} D_{B'}(w) = D_{B'}(v)^t \cdot G' \cdot D_{B'}(w)\end{aligned}$$ □

Beispiel VI.1.9: Es sei $V := C([0,1],\mathbb{R}) = \{f\colon [0,1] \to \mathbb{R} \mid f \text{ ist stetig}\}$. Dann ist

$$\beta\colon V \times V \longrightarrow \mathbb{R}, \qquad (f,g) \longmapsto \int_0^1 f(x)g(x)\,dx$$

eine Bilinearform auf V.

Definition VI.1.10: Sei $\beta\colon V \times V \to K$ eine symmetrische Bilinearform.

(i) $v, w \in V$ heißen *orthogonal*, falls $\beta(v,w) = 0$,

(ii) Seien weiter $\dim(V) = n$ und $B = \{b_1, \ldots, b_n\}$ eine Basis von V. Dann heißt B eine *Orthogonalbasis*, falls für $i \neq j$ gilt: $\beta(b_i, b_j) = 0$

(iii) Eine Basis B wie in (ii) heißt *Orthonormalbasis*, falls B Orthogonalbasis ist und für $1 \leq i \leq n$ gilt: $\beta(b_i, b_i) = 1$.

(iv) Zwei Untervektorräume U_1, U_2 von V heißen *orthogonal*, falls für alle $u_1 \in U_1$, $u_2 \in U_2$ gilt: $\beta(u_1, u_2) = 0$.

Proposition VI.1.11 (Orthogonalsystem): *Sei β wiederum eine symmetrische Bilinearform auf V. Gilt für $v_1, \ldots, v_k \in V$, dass $\beta(v_i, v_j) = 0$ für $i \neq j$ und $\beta(v_i, v_i) \neq 0$ für $1 \leq i \leq k$, dann ist $\{v_1, \ldots, v_k\}$ linear unabhängig.*

Beweis: Seien $\lambda_1, \ldots, \lambda_k \in K$ mit $\sum_{j=1}^k \lambda_j v_j = \mathbf{0}$. Dann gilt für $1 \leq i \leq k$, dass

$$\beta(\mathbf{0}, v_i) = \beta\Big(\sum_{j=1}^k \lambda_j v_j, v_i\Big) = \sum_{j=1}^n \lambda_j \beta(v_j, v_i) = \lambda_j \beta(v_i, v_i),$$

d. h. $\lambda_i = 0$ für $1 \leq i \leq k$ und $\{v_1, \ldots, v_k\}$ ist linear unabhängig. □

2. Skalarprodukte

Definition VI.2.1: Es sei $K = \mathbb{R}$.

(i) Eine symmetrische Bilinearform β heißt *positiv definit*, falls für alle $v \in V - \{0\}$ gilt, dass $\beta(v,v) > 0$.

(ii) Ein *Skalarprodukt* auf V ist eine symmetrische, positiv definite Bilinearform. Wir schreiben $\langle v, w\rangle := \beta(v,w)$.

(iii) Ist V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf V, dann heißt $(V, \langle\cdot,\cdot\rangle)$ ein *euklidischer Vektorraum*.

Beispiel VI.2.2: (i) Die Einheitsform $(x, y) \mapsto \sum_{i=1}^{n} x_i y_i$ ist ein Skalarprodukt auf $\mathbb{R}^n$ und heißt *(euklidisches) Skalarprodukt.*

(ii) Die Bilinearform β auf $C([0,1], \mathbb{R})$ aus Beispiel VI.1.9 ist ein Skalarprodukt.

Die Abbildung $\overline{(\cdot)}\colon \mathbb{C} \to \mathbb{C}, a + \mathrm{i}b \mapsto \overline{a + \mathrm{i}b} := a - \mathrm{i}b$ heißt *komplexe Konjugation.* Sie ist ein Körperautomorphismus von $\mathbb{C}$ und erlaubt die Definition eines Betrags auf $\mathbb{C}$: Für $z \in \mathbb{C}$ heißt

$$|z| := \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2} = \sqrt{\overline{z} \cdot z}$$

der *Betrag* von z.

Definition VI.2.3: Es sei $K = \mathbb{C}$. Eine Abbildung $\beta\colon V \times V \to \mathbb{C}$ heißt eine *hermitesche Form*, falls für alle $u_1, u_2, u, v_1, v_2, v \in V$ und $\lambda \in \mathbb{C}$ gelten:

(i) $\beta(u_1 + u_2, v) = \beta(u_1, v) + \beta(u_2, v)$,

(ii) $\beta(\lambda u, v) = \lambda \beta(u, v)$,

(iii) $\beta(u, \lambda v) = \overline{\lambda} \beta(u, v)$,

(iv) $\beta(u, v) = \overline{\beta(v, u)}$.

Definition VI.2.4: Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *hermitesch*, falls $A = \overline{A^t} = A^*$.

Proposition VI.2.5:

(i) *Sei $A \in \mathbb{C}^{n \times n}$ hermitesch. Dann erklärt $\beta_A(x, y) \mapsto x^t \overline{A} \overline{y}$ eine hermitesche Form.*

(ii) *Ist β eine hermitesche Form auf $\mathbb{C}^n$, dann ist $\beta = \beta_A$ mit $A = (a_{i,j})$, wobei $a_{i,j} = \beta(e_i, e_j)$. Es sei wieder $\mathfrak{E} = \{e_1, \ldots, e_n\}$ die Standardbasis von $\mathbb{C}^n$.*

Bemerkung VI.2.6: Ist β eine hermitesche Form, dann gilt schon für alle $v \in V$, dass $\beta(v, v) \in \mathbb{R}$, da $\beta(v, v) = \overline{\beta(v, v)}$.

Definition VI.2.7: Es sei $\mathbb{K} = \mathbb{C}$ und $\beta\colon V \times V \to \mathbb{C}$ eine hermitesche Form.

(i) β heißt *positiv definit*, falls für alle $v \in V - \{0\}$ gilt, dass $\beta(v, v) > 0$.

(ii) Ein *Skalarprodukt auf V* ist eine positiv definite hermitesche Form. Wir notieren $\langle v, w \rangle := \beta(v, w)$.

(iii) Ist V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$, so heißt V ein *unitärer Raum*.

Beispiel VI.2.8: Die Abbildung

$$\mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \qquad (x,y) \longmapsto \sum_{i=1}^{n} x_i \overline{y_i}$$

heißt *Standardskalarprodukt* auf $\mathbb{C}^n$.

Im Folgenden bezeichne $\mathbb{K}$ entweder den Körper der reellen Zahlen $\mathbb{R}$, oder den Körper der komplexen Zahlen $\mathbb{C}$.

Definition VI.2.9 (Norm, Länge, Abstand):

(i) Für $v \in V$ heißt $\|v\| := \sqrt{\langle v,v\rangle}$ die *Norm* oder auch die *Länge* von v.

(ii) Für $v,w \in V$ heißt $d(v,w) := \|v-w\|$ der Abstand von v und w. Die Abbildung

$$d\colon V \times V \longrightarrow \mathbb{R}_{\geq 0}, \qquad (x,y) \mapsto d(x,y)$$

heißt *Metrik* zu $\langle\cdot,\cdot\rangle$.

Satz 16 (Cauchy-Schwarze'sche-Ungleichung):

(i) *Für alle $v,w \in V$ 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) *In* (i) *gilt Gleichheit genau dann, wenn v und w linear abhängig sind.*

Beweis: Für $v = 0$ ist (i) offensichtlich wahr.

Wir zeigen die Behauptung zunächst für den Fall, dass $\langle v,w\rangle \in \mathbb{R}$ und somit $\langle w,v\rangle = \langle v,w\rangle$. Definiere dazu die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $\lambda \mapsto \|\lambda v + w\|^2$. Jetzt ist

$$\|\lambda v + w\|^2 = \langle \lambda v + w, \lambda v + w\rangle = |\lambda|^2\langle v,v\rangle + 2\lambda\langle v,w\rangle + \langle w,w\rangle.$$

Insgesamt sehen wir: Erstens ist f eine quadratische Funktion in λ und zweitens ist $f(\lambda) \geq 0$ für alle $\lambda \in \mathbb{R}$; d. h. f hat höchstens eine reelle Nullstelle, für die Diskriminante $D = 4\langle v,w\rangle - 4\langle v,v\rangle\langle w,w\rangle$ muss also gelten $D \leq 0$; das liefert die Behauptung.

Sei jetzt allgemein $\langle v, w\rangle \in \mathbb{C}$ und setze $\alpha := \frac{\langle v,w\rangle}{|\langle v,w\rangle|}$ sowie $v' := \alpha^{-1}v$ (nach Konstruktion ist $|\alpha| = 1$). Dann ist

$$\begin{aligned}|\langle v, w\rangle|^2 = |\langle \alpha v', w\rangle|^2 = |\alpha|^2|\langle v', w\rangle|^2 &= |\langle v', w\rangle|^2 \\ &\leq \langle v', v'\rangle\langle w, w\rangle \\ &= \langle \alpha^{-1}v, \alpha^{-1}v\rangle\langle w, w\rangle \\ &= |\alpha|^{-2}\langle v, v\rangle\langle w, w\rangle = \langle v, v\rangle\langle w, w\rangle,\end{aligned}$$

wobei wir verwendet haben, dass

$$\langle v', w\rangle = \left\langle \frac{|\langle v, w\rangle|}{\langle v, w\rangle}v, w \right\rangle = \frac{|\langle v, w\rangle|}{\langle v, w\rangle}\langle v, w\rangle = |\langle v, w\rangle|,$$

also dass wir die vorher gezeigte Aussage auf $\langle v', w\rangle$ anwenden dürfen.

„$\Rightarrow$“: Ohne Einschränkung seien $v \neq 0$ und $\langle v, w\rangle \in \mathbb{R}$. Falls Gleichheit gilt, dann ist $D = 0$ und die Funktion f aus (i) hat genau eine Nullstelle λ_0. Dann gilt

$$\|\lambda_0 v + w\|^2 = \langle \lambda_0 v + w, \lambda_0 v + w\rangle = 0,$$

und wegen der positiven Definitheit von $\langle \cdot, \cdot\rangle$ ist $\lambda_0 v + w = 0$, d. h. v und w sind linear abhängig. „$\Leftarrow$“ ist klar. □

Korollar VI.2.10 (Winkeldefinition über Skalarprodukt):

(i) *Aus Satz 16 wissen wir: Für alle $v, w \in V$ ist*

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

d. h. es gibt genau ein $\alpha \in [0, \pi]$, sodass $\cos(\alpha) = \frac{\langle v,w\rangle}{\|v\|\|w\|}$. Wir nennen α den Winkel zwischen v und w *und schreiben $\alpha := \measuredangle(v, w)$.*

(ii) *v und w sind orthogonal genau dann, wenn $\langle v, w\rangle = 0$. Dies ist genau dann der Fall, wenn $\measuredangle(v, w) = \frac{\pi}{2}$.*

Proposition VI.2.11 (Eigenschaften der Norm): *Die Abbildung*

$$\|\cdot\| \colon V \longrightarrow \mathbb{R}, \qquad v \longmapsto \|v\| := \sqrt{\langle v, v\rangle}$$

hat die Eigenschaften

(i) *Für alle $v \in V$ ist $\|v\| \geq 0$ und $\|v\| = 0$ gilt genau dann, wenn $v = 0$,*

(ii) *Für alle* $\lambda \in \mathbb{K}$ *und* $v \in V$ *ist* $\|\lambda v\| = |\lambda|\|v\|$,

(iii) *Für alle* $v, w \in V$ *gilt* $\|v + w\| \leq \|v\| + \|w\|$.

Beweis: (i) und (ii) sind klar. Zu (iii): Es ist

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

wobei wir bei der vorletzten Ungleichung Satz 16 verwendet haben. □

Korollar VI.2.12 (Eigenschaften der Metrik): *Für die Abbildung*

$$d\colon V \times V \longrightarrow \mathbb{R} \qquad (v, w) \longmapsto d(v, w) = \|v - w\|$$

gelten:

(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* $d(u, v) + d(v, w) \geq d(u, w)$.

Beweis: Die Eigenschaften (i) und (ii) sind wieder klar. Zu (iii): Setze $u' := u - v$, $v' := v - w$. Jetzt ist

$$\begin{aligned}d(u, v) + d(v, w) \geq d(u, w) &\Leftrightarrow \|u - v\| + \|v - w\| \geq \|u - w\| \\ &\Leftrightarrow \|u'\| + \|v'\| \geq \|u' + v'\|,\end{aligned}$$

das heißt wir erhalten die Aussage, weil die Dreiecksungleichung für $\|\cdot\|$ gilt.□

3. Orthogonalität

In diesem Abschnitt sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Vektorraum.

Proposition VI.3.1 (Fourier-Formel): *Ist* $B = \{b_1, \ldots, b_n\}$ *eine Orthonormalbasis von* V*, dann gilt für alle* $v \in V$*:*

$$v = \sum_{i=1}^{n} \langle v, b_i\rangle b_i,$$

d. h. bezüglich dieser Basis ist $D_B(v) = (\langle v, b_i\rangle)_{1 \leq i \leq n}$.

Beweis: Schreibe $v = \sum_{j=1}^{n} \lambda_i b_i$. Dann ist

$$\langle v, b_i \rangle = \left\langle \sum_{j=1}^{n} \lambda_j b_j, b_i \right\rangle = \sum_{j=1}^{n} \lambda_j \langle v_j, v_i \rangle = \lambda_i \langle b_i, b_i \rangle = \lambda_i.$$

□

Satz 17 (Gram-Schmidt-Orthogonalisierung): *Wir erhalten eine Orthogonalbasis von V wie folgt: Ist $B = \{v_1, \ldots, v_n\}$ eine Basis von V, dann setzen wir rekursiv*

$$w_1 := v_1, \qquad w_l := v_l - \sum_{i=1}^{l-1} \frac{\langle w_i, v_l \rangle}{\langle w_i, w_i \rangle} w_i.$$

Es gelten:

(i) *Die Menge $\{w_1, \ldots, w_n\}$ ist eine Orthogonalbasis von V.*

(ii) $\mathrm{Lin}(v_1, \ldots, v_l) = \mathrm{Lin}(w_1, \ldots, w_l)$.

Korollar VI.3.2 (Existenz von Orthonormalbasen): *Jeder euklidische– oder unitäre Vektorraum V besitzt eine Orthonormalbasis.*

Beweis: Satz 17 liefert eine Orthogonalbasis $B = \{b_1, \ldots, b_n\}$ von V und

$$C = \left\{ \frac{1}{\|b_1\|} b_1, \ldots, \frac{1}{\|b_n\|} b_n \right\}$$

ist eine Orthonormalbasis von V. □

Beweis (von Satz 17): Punkt (ii) folgt direkt aus der Definition der w_i.

Aus (ii) folgt, dass $w_i \neq 0$ für $1 \leq i \leq n$. Wir wollen zeigen, dass $\langle w_i, w_j \rangle = 0$ für $1 \leq i \neq j \leq l$. Wir zeigen die Aussage mit vollständiger Induktion, der Induktionsanfang ist klar.

Für den Induktionsschritt von $l-1$ nach l: Ist $1 \leq j \leq l-1$, so ist

$$\begin{aligned} \langle w_j, w_l \rangle &= \left\langle w_j, v_l - \sum_{i=1}^{l-1} \frac{\langle w_i, v_l \rangle}{\langle w_i, w_i \rangle} w_i \right\rangle \\ &= \langle w_j, v_l \rangle - \sum_{j=1}^{l-1} \frac{\langle w_i, v_l \rangle}{\langle w_i, w_i \rangle} \langle w_j, w_i \rangle = \langle w_j, v_l \rangle - \frac{\langle w_j, v_l \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_j \rangle = 0, \end{aligned}$$

was wir zeigen wollten. □

Bemerkung VI.3.3: (i) Ist V ein $\mathbb{C}$-Vektorraum und $\beta \colon V \times V \to \mathbb{C}$ eine Abbildung, die in (Definition VI.2.3) (i), (ii) und (iii) erfüllt, so heißt β eine *Sesquilinearform*.

(ii) Für eine Bilinearform oder Sesquilinearform β und eine Basis B von V notieren wir $G_B(\beta)$ für die Gram-Matrix von β bezüglich B.

(iii) B ist eine Orthonormalbasis bezüglich β genau dann, wenn $G_B(\beta) = I$.

(iv) B ist eine Orthogonalbasis bezüglich β genau dann, wenn $G_B(\beta)$ eine Diagonalmatrix ist.

(v) Ist B' eine weitere Basis, dann gilt:

$$G_{B'}(\beta) = \begin{cases} D_{B,B'}^t \cdot G \cdot D_{B,B'}, & \text{für } \mathbb{K} = \mathbb{R}, \\ D_{B',B}^t \cdot G \cdot \overline{D}_{B,B'}, & \text{für } \mathbb{K} = \mathbb{C}. \end{cases}$$

(vi) Sei $B = \{b_1, \ldots, b_n\}$ eine Orthonormalbasis bezüglich des Standard-Skalarprodukts von $\mathbb{R}^n$ beziehungsweise $\mathbb{C}^n$. Für $A = (b_1 | \ldots | b_n) \in \mathbb{R}^{n \times n}$ beziehungsweise $A = (b_1 | \ldots | b_n) \in \mathbb{C}^n$ gilt: $A^t A = I$ beziehungsweise $A^* A = I$.

Bemerkung VI.3.4: (i) Seien U_1, U_2 Untervektorräume von V mit $U_1 \perp U_2$. Dann gilt $U_1 + U_2 = U_1 \oplus U_2$.

(ii) Für den Untervektorraum U von V heißt

$$U^\perp := \{v \in V \mid \langle u, v \rangle = 0 \text{ für alle } u \in U\}$$

das *orthogonale Komplement von* U. Es gilt $V = U \oplus U^\perp$.

Dieses Aussagen betrachten Sie genauer auf dem vierzehnten Übungsblatt.

Definition VI.3.5 (Orthogonale Projektion): Sei U ein Untervektorraum von V. Die Abbildung

$$\pi \colon V = U \oplus U^\perp \longrightarrow U, \qquad v = u + u' \longmapsto u$$

mit $u \in U, u' \in U^\perp$ heißt *orthogonale Projektion*. Es gilt $v - \pi(v) \perp \pi(v)$.

Proposition VI.3.6 (Satz des Pythagoras): *Seien $v, w \in V$. Gilt $v \perp w$, dann ist*

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

Diese Aussage zeigt man durch einfaches Nachrechnen.

4. Orthogonale und unitäre Matrizen

Ziel dieses Abschnittes soll sein, die Abbildungen zu studieren, die Skalarprodukte– und damit Längen, Winkel und Orthogonalität erhalten. Im Folgenden bezeichne $(V, \langle\cdot,\cdot\rangle)$ einen euklidischen- beziehungsweise unitären Vektorraum der Dimension n, $\Phi\colon V \to V$ einen Endomorphismus und A eine reelle- beziehungsweise komplexe $n \times n$-Matrix.

Definition VI.4.1: Φ heißt *orthogonal* beziehungsweise *unitär*, falls für alle $v, w \in V$ gilt:

$$\langle \Phi(v), \Phi(w)\rangle = \langle v, w\rangle.$$

Bemerkung VI.4.2: Sei Φ orthogonal beziehungsweise unitär. Dann folgt:

(i) Für alle $v \in V$ ist $\|\Phi(v)\| = \|v\|$,

(ii) Für $v, w \in V$ gilt $v \perp w$ genau dann, wenn $\Phi(v) \perp \Phi(w)$.

(iii) Φ ist injektiv (und da $\dim(V) < \infty$ auch bijektiv).

(iv) Φ^{-1} ist ebenfalls orthogonal beziehungsweise unitär.

(v) Ist $\Psi\colon V \to V$ ebenfalls ein euklidischer– beziehungsweise unitärer Endomorphismus, dann ist $\Psi \circ \Phi$ wiederum euklidisch beziehungsweise unitär.

Bemerkung VI.4.3: Die Abbildung $\mathbb{K}^n \to \mathbb{K}^n, x \mapsto Ax$ ist euklidisch beziehungsweise unitär bezüglich des Standard-Skalarprodukts genau dann, wenn für alle $x, y \in \mathbb{K}^n$ gilt:

$$x^t\overline{y} = (Ax)^t\overline{(Ay)} = x^tA^t\overline{A}\overline{y}.$$

Das gilt genau dann, wenn $A^t\overline{A} = I$, was genau dann gilt, wenn $A^*A = I$.

Definition VI.4.4 (Orthogonale– bzw. unitäre Matrizen):

(i) Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, falls $A^tA = I$. Das ist genau dann der Fall, wenn $A \in \mathrm{Gl}_n(\mathbb{R})$ ist und $A^{-1} = A^t$ gilt.

(ii) $A \in \mathbb{C}^{n\times n}$ heißt *unitär*, falls $A^*A = I$. Das ist genau dann der Fall, wenn $A \in \mathrm{Gl}_n(\mathbb{C})$ und $A^{-1} = A^*$ gilt.

Proposition VI.4.5 (Charakterisierung orthogonaler– bzw. unitärer Matrizen):

(i) *Die folgenden Aussagen sind äquivalent:*

(1) *A ist orthogonal beziehungsweise unitär,*

(2) *Die Spaltenvektoren von A bilden eine Orthonormalbasis des $\mathbb{K}^n$ bezüglich des Standard-Skalarprodukts.*

(3) *Die Zeilenvektoren von A bilden eine Orthonormalbasis des $\mathbb{K}^n$ bezüglich des Standard-Skalaprodukts.*

(ii) *Sei B eine Orthonormalbasis von V. Φ ist orthongonal bzw. unitär genau dann, wenn $D_{B,B}(\Phi)$ orthogonal bzw. unitär ist.*

(iii) *Ist λ ein Eigenwert von A und ist A orthogonal oder unitär, dann ist $|\lambda| = 1$.*

Beweis: (i) Es gelten die Äquivalenzen

$$AA^* = I \Leftrightarrow A^*A = I \Leftrightarrow A^t\overline{A} = I.$$

Die letzte Gleichung bedeutet, dass die Spalten von A eine Orthonormalbasis des $\mathbb{K}^n$ sind, wie wir uns vorher schon überlegt haben; die mittlere Gleichung bedeutet, dass A orthogonal beziehungsweise unitär ist und die linke Gleichung bedeutet, dass die Zeilenvektoren von A eine Orthonormalbasis des $\mathbb{K}^n$ bilden.

(ii) Folgt aus (Bemerkung VI.4.3), da $G_B(\langle\cdot,\cdot\rangle) = I$ für eine Orthonormalbasis B.

(iii) Ist λ ein Eigenwert von A, dann gibt es $v \in \mathbb{K}^n - \{\mathbf{0}\}$, mit $Av = \lambda v$, wir können also schreiben

$$0 \neq \|v\| = \|Av\| = \|\lambda v\| = |\lambda|\|v\|,$$

was die Behauptung zeigt. □

Bemerkung VI.4.6: Sind B_1, B_2 Orthonormalbasen von V, dann gilt: D_{B_1,B_2} ist orthogonal beziehungsweise unitär. Das folgt direkt aus (Proposition 4.5 (i)).

Definition VI.4.7: (i) Wir nennen die Menge

$$O(n) := \{A \in \mathrm{Gl}_n(\mathbb{R}) \mid A \text{ ist orthogonal}\}$$

die *orthogonale Gruppe*, in der Tat ist $O(n)$ eine Untergruppe von $\mathrm{Gl}_n(\mathbb{R})$.

(ii) Wir nennen die Menge

$$U(n) := \{A \in \mathrm{Gl}_n(\mathbb{C}) \mid A \text{ ist unitär}\}$$

die *unitäre Gruppe*, in der Tat ist $U(n)$ eine Untergruppe von $\mathrm{Gl}_n(\mathbb{R})$.

Satz 18: *Es seien* $\mathbb{K} = \mathbb{C}$ *und* $\Phi\colon V \to V$ *ein unitärer Endormorphismus. Dann hat* Φ *eine Orthonormalbasis aus Eigenvektoren von* Φ*. Insbesondere ist* Φ *diagonalisierbar.*

Korollar VI.4.8: *Ist* $A \in U(n)$*, dann gibt es* $S \in U(n)$ *mit*

$$S^t A \overline{S} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Wir bemerken, dass für die Matrix S aus Korollar VI.4.8 $S^t = \overline{S}^{-1}$ gilt, da S unitär ist.

Für den Beweis von Satz 18 brauchen wir den sogenannten *Fundamentalsatz der Algebra*, der hier nur zitiert, nicht aber bewiesen werden soll.

Satz (Fundamentalsatz der Algebra): *Jedes Polynom* $p \in \mathbb{C}[X]$*, dessen Grad größergleich Eins ist, hat mindestens eine Nullstelle in* $\mathbb{C}$*. Es gilt*

$$p(X) = c \prod_{i=1}^{d} (X - \lambda_i)$$

mit $c, \lambda_1, \ldots, \lambda_d \in \mathbb{C}$*.*

Lemma VI.4.9 („Haupt-Trick für Satz 18“): *Seien* Φ *unitär und* v *ein Eigenvektor von* Φ *zum Eigenwert* λ *und* $U = \langle v \rangle$*. Dann ist* $U^\perp$ *schon* Φ*-invariant.*

Beweis: Sei $w \in U^\perp$. Dann ist

$$\langle v, \Phi(w) \rangle = \langle \lambda^{-1} \Phi(v), \Phi(w) \rangle = \lambda^{-1} \langle v, w \rangle = 0,$$

d. h. $\Phi(w) \in U^\perp$. □

Beweis (von Satz 18): Wir zeigen die Aussage via Induktion nach $\dim V = n$. Für $n = 1$ ist die Aussage klar. Für den Induktionsschluss von $n - 1$ nach n sei $p(X) = \mathrm{CP}_\Phi(X) \in \mathbb{C}[X]$ das charakteristische Polynom von Φ. Mit dem Fundamentalsatz der Algebra sehen wir, dass p eine Nullstelle $\lambda \in \mathbb{C}$ hat. Sei jetzt v ein Eigenvektor zum Eigenwert λ. Aus (Proposition IV.4.5) wissen wir, dass $|\lambda| = 1$, also $\lambda \neq 0$. (Lemma 4.8) liefert uns jetzt, dass wir V zerlegen können als $V = \langle v \rangle \oplus \langle v \rangle^\perp$ und $\langle v \rangle$ sowie $\langle v \rangle^\perp$ sind Φ-invariante Unterräume. Die Induktionsvoraussetzung angwendet auf $\Phi|_{\langle v \rangle^\perp}$ liefert eine Orthonormalbasis $\{v_2, \ldots, v_n\}$ von $\langle v \rangle^\perp$ aus Eigenvektoren von Φ. Setze $v_1 := \|v\|^{-1} v$, dann ist $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V aus Eigenvektoren von Φ. □

Satz 19 (Variante des Spektralsatz für orthogonale Endomorphismen): *Seien* $\mathbb{K} = \mathbb{R}$ *und* $\Phi\colon V \to V$ *ein orthogonaler Endomorphismus. Dann gibt es eine Orthonormalbasis* B *von* V *mit*

$$D_{B,B}(\Phi) = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & A_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & A_m \end{pmatrix},$$

wobei

$$A_i = \begin{pmatrix} \cos(\alpha_i) & -\sin(\alpha_i) \\ \sin(\alpha_i) & \cos(\alpha_i) \end{pmatrix} \in O(2)$$

mit $\alpha_i \in [0, 2\pi] - \{0, \pi, 2\pi\}$. *Erinnerung: Es gilt* $\sin^2(\alpha) + \cos^2(\alpha) = 1$ *für alle* $\alpha \in \mathbb{R}$.

Korollar VI.4.10: Φ *lässt sich zerlegen in (1)* id*, (2) Spiegelungen und (3) Drehungen in einer Ebene.*

Beispiel VI.4.11: Es seien $n = 2$, $A \in \mathbb{R}^{2\times 2}$ orthogonal und

$$\phi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad (x, y)^t \longmapsto A.(x, y)^t.$$

Fassen wir ϕ als Abbildung als Abbildung $\phi\colon \mathbb{C}^2 \to \mathbb{C}^2$ auf, dann wissen wir aus Satz 18, dass es eine Orthonormalbasis $B = \{b_1, b_2\}$ von $\mathbb{C}^2$ aus Eigenvektoren zu den Eigenwerten $\lambda_1, \lambda_2 \in \mathbb{C}$. Im Folgenden ist komplexe Konjugation komponentenweise zu verstehen. Ist $\lambda_1 \in \mathbb{C} - \mathbb{R}$, dann gilt

$$A\bar{b_1} = \overline{Ab_1} = \overline{\lambda_1 b_1} = \overline{\lambda_1} \cdot \bar{b}_1,$$

d. h. $\bar{b}_1$ ist Eigenvektor zum Eigenwert $\overline{\lambda_1} = \lambda_2$; ohne Einschränkung ist also $b_2 = \bar{b}_1$. Setzen wir $c_1 := \mathrm{Re}(b_1) = \frac{1}{2}(b_1 + \bar{b}_1)$ und $c_2 = \mathrm{Im}(b_1) = \frac{1}{2\mathrm{i}}(b_1 - \bar{b}_1)$, dann ist auch $C = \{c_1, c_2\}$ eine Basis von $\mathbb{C}^2$. Nun kann man nachrechnen, dass

$$D_{C,C}(\Phi) = \begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix}.$$

Da aber $|\lambda|^2 = 1 = \mathrm{Re}(\lambda)^2 + \mathrm{Im}(\lambda)^2$ gibt es $\alpha \in [0, 2\pi]$ mit

$$D_{C,C}(\Phi) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Beweis (von Satz 19): Wir wollen ϕ „komplexifizieren“ wie in Beispiel VI.4.11. Der Spektralsatz (Satz 18) garantiert uns dann die Existenz einer Orthonormalbasis $B = \{b_1, \ldots, b_n\} \subseteq \mathbb{C}^n$ aus Eigenvektoren von ϕ zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$.

Wir können die Eigenwerte so sortieren, dass

$$\lambda_1, \ldots, \lambda_r = 1, \qquad \lambda_{r+1}, \ldots, \lambda_s = -1, \quad \text{und} \quad \lambda_{s+1}, \ldots, \lambda_n \in \mathbb{C} - \mathbb{R}.$$

Wie in Beispiel VI.4.11 sehen wir: Ist v Eigenvektor von Φ zum Eigenwert λ, dann ist $\overline{v}$ ein Eigenvektor von Φ zum Eigenwert $\overline{\lambda}$. Ohne Einschränkung seien $\lambda_{s+2} = \overline{\lambda}_{s+1}, \lambda_{s+4} = \overline{\lambda}_{s+3}, \ldots, \lambda_n = \overline{\lambda}_{n-1}$ sowie $b_{s+2} = \overline{b}_{s+1}, \ldots, b_n = \overline{b}_{n-1}$. Wir verwenden für $U := \langle b_{s+2i-1}, b_{s+2i} \rangle$ den Trick aus Beispiel VI.4.11, ferner setzen wir

$$c_{s+2i-1} := \mathrm{Re}(b_{s+2i-1}), \qquad c_{s+2i} := \mathrm{Im}(b_{s+2i-1}).$$

Genau wie in Beispiel VI.4.11 sehen wir: $C := (b_1, \ldots, b_s, c_{s+1}, c_{s+2}, \ldots, c_{n-1}, c_n)$ ist eine Basis von $\mathbb{C}^n$ und $D_{C,C}(\Phi)$ ist von der gewünschten Form. □

5. Hauptachsentransformation und Anwendungen

Im Folgenden bezeichne $\langle \cdot, \cdot \rangle$ das Standard-Skalarprodukt, $\|\cdot\|$ die zugehörige Norm sowie $\mathfrak{E} := \{e_1, \ldots, e_n\}$ die Standardbasis von $\mathbb{K}^n$.[1]

Satz 20 (Hauptachsentransformationssatz):

(i) *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann gibt es $S \in O(n)$ mit*

$$S^{-1}AS = S^tAS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

wobei $\lambda_i \in \mathbb{R}$.

(ii) *Sei $A \in \mathbb{C}^{n \times n}$ hermitesch. Dann gibt es $S \in U(n)$ mit*

$$S^{-1}AS = S^*AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

mit $\lambda_i \in \mathbb{R}$!

Proposition VI.5.1: *Eigenwerte hermitescher Matrizen sind reell.*

Beweis: Sei $A \in \mathbb{C}^{n \times n}$ hermitesch und $\lambda \in \mathbb{C}$ ein Eigenwert von A zum Eigenvektor v. Dann ist

$$\lambda \|v\|^2 = \lambda \langle v, v \rangle = \lambda v^t \overline{v} = (Av)^t \overline{v} = v^t A^t \overline{v} = v^t \overline{A} \overline{v} = v^t \overline{Av} = v^t \overline{\lambda v} = \overline{\lambda} \|v\|^2.$$

Da $0 \neq v$ ist $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. □

[1] Es ist wieder $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Lemma VI.5.2 (Haupttrick für Satz 20): *Sei $A \in \mathbb{C}^{n\times n}$ hermitesch, λ ein Eigenwert von A zum Eigenvektor $v \in \mathbb{C}^n$. Dann gilt: $\langle v\rangle^\perp$ ist invariant unter $v \mapsto Av$.*

Beweis: Sei $w \in \langle v\rangle^\perp$. Dann gilt

$$\langle v, Aw\rangle = v^t\overline{Aw} = v^tA^t\overline{w} = \langle Av, w\rangle = \langle \lambda v, w\rangle = \lambda\langle v, w\rangle = 0,$$

d. h. $Aw \in \langle v\rangle^\perp$. □

Beweis (von Satz 20): Teil (ii) können wir dank Lemma VI.5.2 ganz genau wie Satz 18 beweisen.

Für Teil (i) fassen wir A als hermitesche Matrix in $\mathbb{C}^{n\times n}$ auf, nach Proposition VI.5.1 sind dann aber die Eigenwerte von A reell und der Beweis für (ii) funktioniert auch über $\mathbb{R}$. □

Fazit VI.5.3 (Kriterien für Diagonalisierbarkeit):

(i) Für einen beliebigen Körper K haben wir gesehen: $A \in K^{n\times n}$ ist diagonalisierbar genau dann, wenn es eine Basis aus Eigenvektoren gibt. Es gibt genau dann eine Basis aus Eigenvektoren, wenn

$$\sum_{\lambda\in\mathrm{Spec}(A)} \dim \mathrm{Eig}(A, \lambda) = n.$$

(ii) Ist $B = \{w_1, \ldots, w_n\}$ eine Basis aus Eigenvektoren und $B = (w_1|\ldots|w_n)$, dann hat $B^{-1}AB$ Diagonalgestalt.

(iii) Symmetrische und hermitesche Matrizen sind diagonalisierbar nach Satz 20 mit einer Orthonormalbasis aus Eigenvektoren.

(iv) Ist $A \in U(n)$, dann ist A diagonalisierbar mit einer Orthonormalbasis aus Eigenvektoren (Satz 18).[2]

Notation VI.5.4: Es seien K ein beliebiger Körper, W ein endlich-dimensionaler K-Vektorraum, $\Phi\colon W \to W$ ein Endomorphismus und $\lambda \in \mathrm{Spec}(\Phi)$. Dann heißen

(i) $\mu_g(\Phi, \lambda) := \dim \mathrm{Eig}(\Phi, \lambda)$ heißt *geometrische Vielfachheit.*

[2]Es gibt unterschiedliche Konventionen für hermitesche Formen, was die Komponente betrifft, aus der Skalare komplex konjugiert „gezogen“ werden, entsprechend für das Standard-Skalarprodukt auf $\mathbb{C}^n$. In der Physik ist die Konvention gebräuchlich, die Skalare aus der ersten Komponente komplex konjugiert zu ziehen.

(ii) $\mu_a(\Phi, \lambda) :=$ (Exponent von $(X - \lambda)$ im charakteristischen Polynom $\mathrm{CP}_\Phi(X)$) heißt *algebraische Vielfachheit.*

Der Endomorphismus Φ ist diagonalisierbar genau dann, wenn für alle $\lambda \in \mathrm{Spec}(A)$ gilt: $\mu_g(\Phi, \lambda) = \mu_a(\Phi, \lambda)$.

Definition VI.5.5: (i) Eine symmetrische Matrix $G \in \mathbb{R}^{n\times n}$ heißt *positiv definit*, falls für alle $x \in \mathbb{R}^n - \{\mathbf{0}\}$ gilt: $x^t G x > 0$. Das heißt G ist positiv definit genau dann, wenn β_G ein Skalarprodukt ist.

(ii) Eine hermitesche Matrix $G \in \mathbb{C}^{n\times n}$ heißt *positiv definit*, falls für alle $x \in \mathbb{C}^n - \{\mathbf{0}\}$ gilt: $x^t G \overline{x} > 0$. Das heißt G ist positiv definit genau dann, wenn β_G ein Skalarprodukt ist.

Korollar VI.5.6 (aus Satz 20): *Eine symmetrische beziehungsweise hermitesche Matrix G ist positiv definit genau dann, wenn* $\mathrm{Spec}(G) \subseteq [0, \infty)$.

Beweis: Ist eine Konsequenz aus Satz 20, (Bemerkung VI.3.3) und (Bemerkung VI.4.6). □

Satz 21 (Minoren-Kriterium): *Seien $G = (g_{i,j})$ eine symmetrische beziehungsweise hermitesche $n \times n$-Matrix und für $1 \le k \le n$ setze*

$$G_k := \begin{pmatrix} g_{1,1} & \cdots & g_{1,k} \\ \vdots & \ddots & \vdots \\ g_{k,1} & \cdots & g_{k,k} \end{pmatrix}.$$

Dann gilt: G ist positiv definit genau dann, wenn für $1 \le k \le n$ gilt $\det G_k > 0$.

Beweis: Wir zeigen die Behauptung durch vollständige Induktion nach der Größe von G. Der Induktionsanfang für $n = 1$ ist klar. Für den Induktionsschluss von $n - 1$ nach n sei $\beta = \beta_G$, d. h. $\beta(x, y) = x^t G \overline{y}$.

Beachte: $\beta_k := \beta|_{\langle e_1,\ldots,e_k\rangle}$ und $G_k = G_{(e_1,\ldots,e_k)}(\beta_k)$.

„⇒“: G ist positiv definit genau dann, wenn β positiv definit ist. Ist β positiv definit, so auch β_k und nach der Induktionsvoraussetzung wissen wir, dass für $1 \le k \le n - 1$ gilt $\det(G_k) > 0$. Das Orthogonalisierungsverfahren (Satz 17) sichert uns die Existenz einer Orthonormalbasis B für β und damit ist $G_B(\beta) = I_n$. Dann gilt:

$$G = G_{\mathfrak{E}}(\beta) = D^t_{B,\mathfrak{E}} G_B(\beta) \overline{D}_{B,\mathfrak{E}}.$$

Setze $S := D_{B,\mathfrak{E}}$. Dann ist

$$\det(G) = \det(S^t) 1 \det(\overline{S}) = \det(S)\overline{\det(S)} = |\det(S)|^2 > 0.$$

„⇐“: Nach Induktionsvoraussetzung ist G_{n-1} positiv definit und damit auch β_{k-1}. Nach Satz 17 gibt es eine Orthonormalbasis $\langle e_1, \ldots, e_{n-1}\rangle$, diese heiße $\{v_1, \ldots, v_{n-1}\}$. Setze

$$v_n := e_n - \sum_{i=1}^{n-1} \beta(v_i, e_n) v_i,$$

nach Satz 17 ist dann $\beta(v_n, v_i) = 0$; setze $B := \{v_1, \ldots, v_n\}$. Bezüglich dieser Basis ist $G' := G_B(\beta) = \mathrm{diag}(1, \ldots, 1, c)$ mit $c \in \mathbb{C}$. Jetzt finden wir

$$0 < \det(G) = \det(S^t \cdot G' \cdot S) = |\det(S)|^2 \cdot \det(G') = |\det(S)|^2 c,$$

also muss $c > 0$ gelten und β ist positiv definit. □

Satz 22 (Trägheitssatz von Sylvester): *Sei β eine symmetrische beziehungsweise unitäre Form auf einem eukldischen beziehungsweise unitären Vektorraum $(V, \langle\cdot,\cdot\rangle)$. Dann gelten:*

(i) *Es gibt eine Basis B mit*

$$G_B(\beta) = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0),$$

wobei k-mal 1, l-mal -1 und m-mal 0 auf der Diagonalen vorkommen.

(ii) *Die Zahlen k, l, m sind Invarianten von β, d. h. sie sind gleich für alle solchen Basen B. Das Tripel (k, l, m) heißt* Signatur *von β.*

Beweis: (i) Satz 20 liefert die Existenz einer Orthonormalbasis $B = (v_1, \ldots, v_n)$ mit $G_B(\beta) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ mit $\lambda_i \in \mathbb{R}$. Definiere

$$\overline{v}_i := \begin{cases} |\lambda|^{-\frac{1}{2}} v_i, & \text{falls } \lambda_i \neq 0, \\ v_i, & \text{sonst.} \end{cases}$$

Die Basis $\{\overline{v}_1, \ldots, \overline{v}_n\}$ (nach Umsortieren) leistet das Gewünschte.

(ii) Es ist $\mathrm{Rang}(G_B(\beta)) = k + l$. Setze

$$W_1 := \{v \in K^n \mid \beta(v, v) > 0\} \quad \text{und} \quad W_2 := \{v \in K^n \mid \beta(v, v) < 0\}.$$

Wir wissen, dass $k \leq \dim W_1$, $l \leq \dim W_2$ und dass $W_1 \cap W_2 = \{\mathbf{0}\}$. Damit erhalten wir

$$k + l \leq \dim(W_1) + \dim(W_2) = \dim(W_1 + W_2) \leq k + l,$$

insgesamt also $\dim(W_1) = k$, $\dim(W_2) = l$. □

Definition VI.5.7 (Quadratische Form):

(i) Eine *quadratische Form in n Variablen* ist eine Funktion

$$q\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n, \qquad x \longmapsto x^t \cdot A \cdot x + b^t \cdot x + c$$

mit $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ und $c \in \mathbb{R}$, d. h.

$$q(x) = \sum_{i,j=1}^{n} a_{i,j} x_i x_j + \sum_{i=1}^{n} b_i x_i + c.$$

(ii) Die Menge $Q := Q(A,b,c) := \{x \in \mathbb{R}^n \mid q(x) = 0\}$ heißt *Quadrik* zu (A,b,c).

Korollar VI.5.8 (aus Satz 20): *Sei $Q \in \mathbb{R}^n$ eine Quadrik zu $(A, \mathbf{0}, c)$. Dann existiert $S \in O(n)$, so dass für $\Phi\colon \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Sx$ gilt:*

$$\Phi^{-1}(Q) = \left\{ a_1 x_1^2 + \cdots + a_r x_r^2 = \left\{ \begin{array}{c} 1 \\ 0 \end{array} \right\} \right.$$

mit $0 \le r \le n$ und $a_1, \ldots, a_r \in \mathbb{R}$.

Teil 2.

Lineare Algebra II

Kapitel VII.

Jordansche Normalform

1. Motivation

Die Leitfrage für die Lineare Algebra I war „Welche Struktur hat die Lösungsmenge $\mathbb{L}$ eines linearen Gleichungssystems $Ax = b$?“ Diese Frage haben wir in der Linearen Algebra I auch beantwortet und festgestellt, dass $\mathbb{L} = x_s + \mathbb{L}_h$, wobei $\mathbb{L}_h$ der Kern der linearen Abbildung $x \mapsto Ax$ ist.

Strukturelle Werkzeuge, die wir unterwegs verwendet haben, sind

(1) Ist $\Phi\colon K^m \to K^n$ eine lineare Abbildung, so gibt es $A \in K^{n\times m}$ mit $\Phi(x) = Ax$ für alle $x \in K^m$,

(2) Sind V, W endlichdimensionale Vektorräume der Dimension m bzw. n, so ist $V \cong K^n$ und $W \cong K^m$. Für eine lineare Abbildung $\Phi\colon V \to W$ haben wir gesehen, dass es eine eindeutig bestimmte Matrix $A \in K^{n\times m}$ gibt, die das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} \\ K^m & \xrightarrow[\varphi\colon x\mapsto Ax]{} & K^n \end{array}$$

kommutativ macht. Wir erhalten die Isomorphismen D_B, D_C durch Wahl von geordneten Basen in V bzw. W. Die Matrix A hatten wir mit $D_{C,B}(\Phi)$ bezeichnet. Insgesamt haben wir damit außerdem gezeigt: $\mathrm{Hom}_K(V, W) \cong K^{n\times m}$.

Ab jetzt betrachten wir den Spezialfall $V = W$ mit $\dim V = n$.

- Ist $\Phi\colon V \to V$ linear und B eine geordnete Basis, so bezeichnen wir $D_B(\Phi) := A$.

- Haben wir unterschiedliche Basen geordnete B, C, dann ist
$$D_C(\Phi) = D_{C,B}(\mathrm{id}) D_B(\Phi) D_{C,B}(\mathrm{id})^{-1}.$$
- Wann gibt es eine „sehr schöne“ Normalform für Φ? Wir haben bezeichnet:
  - $\mathrm{Spec}(\Phi) = \{\lambda_1, \ldots, \lambda_k\}$, die Menge der Eigenwerte von Φ,
  - $\mathrm{Eig}(\Phi, \lambda) = \mathrm{Kern}(\Phi - \lambda\,\mathrm{id})$, der Eigenraum von Φ zum Eigenwert λ.

Satz: *Es gilt* $\sum_{i=1}^{k} \dim \mathrm{Eig}(\Phi, \lambda_i) = \dim V$ *genau dann, wenn* Φ *diagonalisierbar ist.*

In diesem Fall gibt es eine Basis B von V mit
$$D_B(\Phi) = \mathrm{diag}(a_1, \ldots, a_n) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

Wir wollen jetzt eine „schöne“ Normalform für jede lineare Abbildung $\Phi\colon V \to V$ finden. Das geht aber nur für spezielle Körper, zum Beispiel $K = \mathbb{C}$.

Satz 23: *Sei* V *ein* $\mathbb{C}$*-Vektorraum von Dimension* n *und* $\Phi \in \mathrm{End}_{\mathbb{C}}(V)$*. Dann gibt es eine geordnete Basis* B *von* V *mit*
$$D_B(\Phi) = \begin{pmatrix} \lambda_1 & & & & & & & \\ 1 & \ddots & & & & & & \\ & 1 & \lambda_1 & & & & & \\ & & & \ddots & & & & \\ & & & & \lambda_k & & & \\ & & & & 1 & \lambda_k & & \\ & & & & & \ddots & \ddots & \\ & & & & & & 1 & \lambda_k \end{pmatrix} =: J_\Phi$$
Hierbei sind alle nicht angegebenen Einträge 0 und $\lambda_1, \ldots, \lambda_k$ *die Eigenwerte von* Φ *(nicht zwangsläufig paarweise verschieden). Die Matrix* J_Φ *ist eindeutig bis auf Vertauschen der Kästchen und heißt* Jordan-Normalform.

2. Der Satz von Cayley-Hamilton

In diesem Abschnitt sind stets K ein Körper, V ein endlichdimensionaler K-Vektorraum, und $(e_1, \ldots, e_n)$ die Standardbasis von K^n.

Bemerkung: Man kann Matrizen in Polynome einsetzen: Ist $f = a_d X^d + a_{d-1}X^{d-1} + \cdots + a_1 X^1 + a_0 X^0 \in K[X]$, dann ist

$$f(A) := a_d A^d + a_{d-1}A^{d-1} + \cdots + a_1 A^1 + a_0 I_n$$

wobei $A \in K^{n\times n}$ und I_n die $n\times n$-Einheitsmatrix ist. Dabei gilt $Af(A) = f(A)A$.

Auf die gleiche Weise können wir Endomorphismen von V in Polynome einsetzen:

$$f(\Phi) = a_d\Phi^d + a_{d-1}\Phi^{d-1} + \cdots + a_1\Phi + a_0\,\mathrm{id}_V$$

für $\Phi \in \mathrm{End}(V)$. Hierbei bezeichnet Φ^k, $k \in \mathbb{N}_0$, die k-fache Komposition von Φ mit sich selbst; also $\Phi^0 = \mathrm{id}_V$, $\Phi^1 = \Phi$ und $\Phi^k = \Phi \circ \Phi^{k-1}$ für $k \geq 2$.

Beispiel VII.2.1 („kleiner Zaubertrick“): Wir betrachten die Matrix

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

(1) Bestimmung des charakteristischen Polynoms $\chi_A(X) = \mathrm{CP}_A(X)$:

$$\begin{aligned}
&\det(A - XI_4) \\
&= \det \begin{pmatrix} -X & 0 & 0 & -1 \\ 1 & -X & 0 & 2 \\ 0 & 1 & -X & 1 \\ 0 & 0 & 1 & -2-X \end{pmatrix} \\
&= (-1)(-1)\det \begin{pmatrix} 1 & -X & 0 \\ 0 & 1 & -X \\ 0 & 0 & 1 \end{pmatrix} + 2\det \begin{pmatrix} -X & 0 & 0 \\ 0 & 1 & -X \\ 0 & 0 & 1 \end{pmatrix} \\
&\qquad +(-1)(1)\det \begin{pmatrix} -X & 0 & 0 \\ 1 & -X & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-2-X)\det \begin{pmatrix} -X & 0 & 0 \\ 1 & -X & 0 \\ 0 & 1 & -X \end{pmatrix} \\
&= 1 - 2X - 1X^2 + 2X^3 + X^4
\end{aligned}$$

(2) Einsetzen von A in CP_A: $\mathrm{CP}_A(A) = I_4 - 2A - A^2 + 2A^3 + A^4$

Es sind

$$A^2 = \begin{pmatrix} 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & -3 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 5 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & -1 & 2 & 5 \\ 0 & 2 & -5 & 12 \\ 0 & 1 & 0 & 0 \\ 1 & -2 & 5 & -10 \end{pmatrix}$$

$$A^4 = \begin{pmatrix} -1 & 2 & -5 & 10 \\ 2 & -5 & 12 & -25 \\ 1 & 0 & 0 & 2 \\ -2 & 5 & -10 & 20 \end{pmatrix}.$$

Damit berechnen wir

$$\begin{aligned} \mathrm{CP}_A(A) = &\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 02 & \\ -2 & 0 & 0 & -4 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & 4 \end{pmatrix} \\ &+ \begin{pmatrix} 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 5 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & 5 \end{pmatrix} + \begin{pmatrix} 0 & -2 & 4 & -10 \\ 0 & 4 & -10 & 24 \\ 0 & 2 & 0 & 0 \\ 2 & -4 & 10 & -20 \end{pmatrix} \\ &+ \begin{pmatrix} -1 & 2 & -5 & 10 \\ 2 & -5 & 12 & -25 \\ 1 & 0 & 0 & 2 \\ -2 & 5 & -10 & 20 \end{pmatrix} \\ = &\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

(3) Eine verallgemeinerbare Begründung für (2): Wir beobachten, dass

(A) $Ae_1 = e_2$, $Ae_2 = e_3 = A^2e_1$, $Ae_3 = e_4 = A^3e_1$,

(B) $Ae_4 = -a_0e_1 - a_1e_2 - a_2e_3 - a_3e_4$,

(C) $\mathrm{CP}_A(X) = a_0 + a_1X + a_2X^2 + a_3X^3 + X^4$, also

$$\mathrm{CP}_A(A) = a_0I_4 + a_1A + a_2A^2 + a_3A^3 + A^4,$$

(D)

$$\begin{aligned} \mathrm{CP}_A(A)e_1 &= (a_0I_4 + a_1A + a_2A^2 + a_3A^3 + A^4)e_1 \\ &= a_0e_1 + a_1e_2 + a_2e_3 + a_3e_4 + Ae_4 = 0. \end{aligned}$$

(E) $\mathrm{CP}_A(A)e_2 = \mathrm{CP}_A(A)Ae_1 = A\,\mathrm{CP}_A(A)e_1 = 0$, und analog sieht man $\mathrm{CP}_A(A)e_3 = 0$ sowie $\mathrm{CP}_A(A)e_4 = 0$.

Proposition VII.2.2: *Sei*

$$A = \begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \\ & & \ddots & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix} \in K^{n\times n},$$

wobei ausgelassene Einträge gleich Null sind. Dann gelten:

(i) $\mathrm{CP}_A(X) = (-1)^n(a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + X^n)$,

(ii) $\mathrm{CP}_A(A) = \mathbf{0}$.

Beweis: Verfahre wie in Beispiel VII.2.1:

- Berechne das charakteristische Polynom von A per Laplace-Entwicklung nach der letzten Spalte,
- Es gelten:
  (A) $Ae_1 = e_2$, $A^2e_1 = e_3$, …, $A^{n-1}e_1 = e_n$,
  (B) $A^ne_1 = Ae_n = -a_0e_1 - a_1e_2 - \cdots - a_{n-1}e_n$,
  (C) $\mathrm{CP}_A(A) = (-1)^n(a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1} + A^n)$,
  (D) Daraus folgt

$$\mathrm{CP}_A(A)e_1 = (-1)^n(a_0e_1 + a_1e_2 + \cdots + a_{n-1}e_n + A^ne_1) = 0,$$

  (E) $\mathrm{CP}_A(A)e_k = \mathrm{CP}_A(A)A^{k-1}e_1 = A^{k-1}\,\mathrm{CP}_A(A)e_1 = 0$ für $2 \le k \le n$,

  Aus (D) und (E) folgt: $\mathrm{CP}_A(A) = \mathbf{0}$. □

Proposition VII.2.3: *Seien $\Phi \in \mathrm{End}(V)$ und $v \in V$. Wähle n minimal, sodass $\{v, \Phi(v), \ldots, \Phi^n(v)\}$ linear abhängig ist, also*

$$\Phi^n(v) = -a_0v - a_1\Phi(v) - \cdots - a_{n-1}\Phi^{n-1}(v).$$

Dann gelten:

(i) *$U := \langle v, \Phi(v), \ldots, \Phi^{n-1}(v)\rangle$ ist (bezüglich Inklusion) minimaler Φ-invarianter Untervektorraum, der v enthält,*

(ii) $B = (v, \Phi(v), \Phi^2(v), \ldots, \Phi^{n-1}(v))$ *bildet eine geordnete Basis von* U,

(iii) *Die Einschränkung von* Φ *auf* U *hat bezüglich* U *die folgende Darstellungsmatrix:*

$$D_{B,B}(\Phi|_U) = \begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \\ & & \ddots & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}$$

Beweis: (ii) Folgt aus der Definition von n und U.

(i) U ist Φ-invariant, denn für $u \in U$ gilt $u = c_0 v + c_1\Phi(v) + \cdots + c_{n-1}\Phi^{n-1}(v)$. Jetzt ist

$$\Phi(u) = c_0\Phi(v) + c_1\Phi^2(v) + \cdots + c_{n-1}\Phi^n(v);$$

und da $\Phi(v), \ldots, \Phi^{n-1}(v) \in U$ sowie $\Phi^n(v) = -a_0 v - a_1\Phi(v) - \cdots - a_{n-1}\Phi^{n-1}(v)$, ist $\Phi(u) \in U$.

U ist minimal, denn ist W ein weiterer Φ-invarianter Untervektorraum mit $v \in W$, dann gilt $\Phi(v) \in W$, also $\Phi^2(v) \in W$, ..., also $\Phi^{n-1}(v) \in W$ und damit $U \subseteq W$.

(iii) Wir setzen $b_i := \Phi^{i-1}(v)$ für $1 \leq i \leq n$. Dann ist $\Phi(b_i) = b_{i+1}$ für $1 \leq i \leq n$, also

$$\begin{aligned} \Phi(b_n) &= \Phi^n(v) \\ &= -a_0 v - a_1\Phi(v) - \cdots - a_{n-1}\Phi^{n-1}(v) = -a_0 v_0 - \cdots - a_{n-1} b_{n-1}. \end{aligned}$$ □

Satz 24 (Cayley-Hamilton):

(i) *Seien* V *ein endlichdimensionaler* K*-Vektorraum und* $\Phi \in \mathrm{End}(V)$*. Dann gilt* $\mathrm{CP}_\Phi(\Phi) = \mathbf{0}_{\mathrm{End}(V)}$.

(ii) *Für* $A \in K^{N\times N}$ *ist* $\mathrm{CP}_A(A) = \mathbf{0}_{K^{N\times N}}$.

Beweis: (i) Wir schreiben $f := \mathrm{CP}_\Phi \in K[X]$ und $\psi := f(\Phi) \in \mathrm{End}(V)$. Wir wollen zeigen, dass ψ die Nullabbildung ist, indem wir zeigen $\psi(v) = \mathbf{0}_V$ für alle $v \in V$.

(1) Suchen einer „schönen“ Basis: Sei ein beliebiges $v \in V$ vorgegeben und n gewählt wie in Proposition VII.2.3, d. h. $U := \langle v, \Phi(v), \ldots, \Phi^{n-1}(v)\rangle$ ist minimaler Φ-invarianter Untervektorraum mit $v \in U$. Sei $B' := (v, \Phi(v), \ldots, \Phi^{n-1}(v))$. Ergänze B' zu einer Basis B von V; nach Proposition VII.2.3 ist dann

$$A := D_{B,B}(\Phi) = \begin{pmatrix} A' & * \\ \mathbf{0} & C \end{pmatrix}$$

mit Blockmatrizen $A' := D_{B',B'}(\Phi|_U) \in K^{n\times n}$, $* \in K^{n\times(N-n)}$, $\mathbf{0} \in K^{(N-n)\times n}$ und $C \in K^{(N-n)\times(N-n)}$.

(2) Berechnung von $f = \mathrm{CP}_\Phi = \mathrm{CP}_A$:

$$f(X) = \det(A - XI_n) = \det(A' - XI_N)\det(C - XI_{(N-n)}),$$

nach Proposition VII.2.3 ist $f(A') = \mathbf{0}$, und deshalb (bitte versuchen Sie, diesen Schritt nachzuvollziehen)

$$f(A) = \begin{pmatrix} f(A') & * \\ \mathbf{0} & f(C) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & * \\ \mathbf{0} & f(C) \end{pmatrix},$$

wir haben also gezeigt, dass $f(\Phi)|_U = 0$ und insbesondere gilt für unser v: $f(\Phi)(v) = 0 = \psi(v) = 0$.

(ii) Folgt aus (i), wenn wir $\Phi\colon v \mapsto Av \in \mathrm{End}(V)$ wählen. □

„Alternativer Beweis“: Wir wollen zeigen, dass $\mathrm{CP}_A(A) = \mathbf{0}$ für $A \in K^{N\times N}$. Es gilt $\mathrm{CP}_A(X) = \det(A - XI_N)$, also ist $\mathrm{CP}_A(A) = \det(A - A) = \det(\mathbf{0}) = 0$.

Das ist offensichtlich Unsinn. Das kann man schon daran sehen, dass in der letzten Gleichheit links eine Matrix, und rechts eine Zahl steht. Das Problem ist: $A - XI_N$ lebt in $K[X]^{N\times N}$, ausgeschrieben mit $A = (a_{i,j})$:

$$A - XI_N = \begin{pmatrix} a_{1,1} - X & a_{1,2} & \dots & a_{1,N-1} & a_{1,N} \\ a_{2,1} & a_{2,2} - X & \dots & a_{2,N-1} & a_{2,N} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{N,1} & a_{N,2} & \dots & a_{N,N-1} & a_{N,N} - X \end{pmatrix},$$

dort können wir keine Matrix einsetzen.

Definition VII.2.4: $A \in K^{n\times n}$ heißt *nilpotent*, falls es $k \in \mathbb{N}$ gibt mit $A^k = \mathbf{0}$.

Beispiel VII.2.5: Sei $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. Dann ist $A \neq \mathbf{0}$, aber $A^2 = \mathbf{0}$.

Proposition VII.2.6: *Es sei $A \in K^{n\times n}$.*

(i) *Ist A nilpotent, dann gilt* $\mathrm{Spec}(A) \subseteq \{0\}$,

(ii) *Ist $K = \mathbb{C}$, dann gilt: A ist nilpotent genau dann, wenn* $\mathrm{Spec}(A) = \{0\}$.

Beweis: (i) Sei A nilpotent mit $A^k = \mathbf{0}$. Sei weiter $\lambda \in K$ ein Eigenwert von A, d. h. es gibt $v \in K^n - \{0\}$ mit $Av = \lambda v$. Für dieses v gilt

$$\mathbf{0} = A^k v = A^{k-1}\lambda v = \lambda A^{k-1} v = \cdots = \lambda^k v,$$

d. h. $\lambda^k = 0$, also $\lambda = 0$.

(ii) „⇒“: Folgt aus (i), weil das charakteristische Polynom $\mathrm{CP}_A(X) \in \mathbb{C}[X]$ nach dem Fundamentalsatz der Algebra eine Nullstelle in $\mathbb{C}$ hat.

„⇐“: Wegen des Fundamentalsatz der Algebra (induktiv) zerfällt das charakteristische Polynom in Linearfaktoren, also

$$\mathrm{CP}_A(X) = (X - \lambda_1) \cdots (X - \lambda_n)$$

mit geeigneten $\lambda_1, \ldots, \lambda_n \in K$. Nach Voraussetzung ist $\lambda_1 = \cdots = \lambda_n = 0$, also $\mathrm{CP}_A(X) = X^n$. Nach dem Satz von Cayley-Hamilton ist $\mathbf{0} = \mathrm{CP}_A(A) = A^n$, A ist also nilpotent. □

Beispiel VII.2.7: Gegeben sei die Matrix

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

Das charakteristische Polynom von A ist $\mathrm{CP}_A(X) = \det(\begin{smallmatrix} -X & -1 \\ 1 & -X \end{smallmatrix}) = X^2 + 1$ und dieses hat keine Nullstellen über $\mathbb{R}$, d. h. $\mathrm{Spec}(A) = \varnothing$.

3. Der Polynomring $K[X]$

Das Ziel dieses Abschnitts wird es sein, ein Polynom f von minimalem Grad zu finden, sodass $f(A) = \mathbf{0}$ für eine Matrix A. Ein solches f nennt man Minimalpolynom. Minimalpolynome werden in der Theorie der Jordan-Normalform eine wichtige Rolle spielen.

In diesem Abschnitt bezeichne K stets einen Körper.

Erinnerung VII.3.1: Der Polynomring

$$K[X] := \left\{ \sum_{i=0}^{n} a_i X^i : n \in \mathbb{N}_0, a_i \in K \right\}$$

ist ein kommutativer Ring mit Addition und Multiplikation von Polynomen.

Welche Art von Objekten kann man in Polynome einsetzen? Wir brauchen Addition, Potenzen (also Multiplikation, da Potenzen alleine nicht genügen) und Multiplikation mit Skalaren (aus einem Körper K). Mengen, in denen man so „rechnen“ kann, nennt man K-Algebren.

Definition VII.3.2 (K-Algebra):

(i) Eine *K-Algebra* ist eine Menge A mit 3 Abbildungen

$$+\colon A \times A \longrightarrow A, \quad \bullet\colon A \times A \longrightarrow A, \quad \cdot\colon K \times A \longrightarrow A,$$

sodass gelten:

(A) $(A, +, \bullet)$ ist ein Ring,[1]

(B) $(A, +, \cdot)$ ist ein K-Vektorraum,

(C) $\bullet$ ist K-bilinear, d. h. für alle $x, y, z \in A$ und $\lambda \in K$ gilt

$$(x + y) \bullet z = x \bullet z + y \bullet z, \quad x \bullet (y + z) = x \bullet y + x \bullet z,$$
$$\lambda \cdot (x \bullet y) = (\lambda \cdot x) \bullet y = x \bullet (\lambda \cdot y).$$

Bemerkung: Die ersten beiden Gleichungen von links folgen aus (A).

(ii) Ist die Verknüpfung $\bullet$ aus (i) eine kommutative Verknüpfung, so nennt man $(A, +, \bullet)$ eine *kommutative Algebra* und hat der Ring $(A, +, \bullet)$ ein Einselement (d. h. es gibt $1 \in A$ mit $1 \bullet a = a \bullet 1$ für alle $a \in A$), so heißt A eine *Algebra mit 1*.

(iii) Seien A_1, A_2 zwei K-Algebren und $\Phi\colon A_1 \to A_2$ eine Abbildung. Dann heißt Φ ein *K-Algebren-Homomorphismus*, wenn gelten:

- Φ ist Ringhomomorphismus für die Ringe $(A_1, +, \bullet)$, $(A_2, +, \bullet)$,
- Φ ist lineare Abbildung für die K-Vektorräume $(A_1, +, \cdot)$, $(A_2, +, \cdot)$.

Beispiel VII.3.3 (*K*-Algebren): Das Folgende sind K-Algebren:

(i) Der Polynomring $K[X]$ mit Addition und Multiplikation von Polynomen und skalarer Multiplikation,

(ii) $K^{n\times n}$ mit Addition und Multiplikation von Matrizen und skalarer Multiplikation,

(iii) $\mathrm{End}_K(V)$ für einen K-Vektorraum V mit Addition und Verknüpfung von Abbildungen sowie skalarer Multiplikation,

(iv) K selbst (zum Beispiel als Spezialfall von (ii) wegen $K \cong K^{1\times 1}$).

Definition VII.3.4 (Einsetzungshomomorphismus): Ist $(A, +, \bullet)$ eine K-Algebra mit 1, so kann man die Elemente von A in Polynome aus $K[X]$ einsetzen.

[1] Im Gegensatz zur Definition von *Ring* aus Lineare Algebra I wollen wir explizit Ringe ohne Eins zulassen und nennen Ringe, die eine Eins haben, *Ringe mit Eins*.

Genauer: Für $f = \sum_{i=0}^{n} c_i X^i \in K[X]$ und $a \in A$ definieren wir

$$f(a) := \sum_{i=0}^{n} c_i \cdot a^i,$$

wobei $a^0 := 1$, $a^1 = a$ und $a^k = a \bullet a^{k-1}$ für $k \geq 2$, $k \in \mathbb{N}$. Wir erhalten so für jedes $a \in A$ einen K-Algebren-Homomorphismus $\varphi_a \colon K[X] \to A, f \mapsto f(a)$. Dieser heißt der *Einsetzungshomomorphsimus zu a*.

Definition VII.3.5 (Grad): Sei $f = \sum_{i=0}^{n} c_i X^i \in K[X]$.

(i) Wir nennen

$$\deg(f) := \begin{cases} n, & \text{falls } c_n \neq 0, \\ -\infty, & \text{falls } f = 0_{K[X]}. \end{cases}$$

den *Grad von f*; $-\infty$ ist hier ein Symbol das nicht zu $\mathbb{N}_0$ gehört.

(ii) Für $k \in \mathbb{N}_0$ definieren wir $\max\{k, -\infty\} := k$, $\max\{-\infty, -\infty\} := -\infty$, $k + (-\infty) := -\infty =: (-\infty) + k$, und $-\infty + -\infty := -\infty$, $-\infty < k$.

(iii) Ist f nicht das Nullpolynom ist $c_n \neq 0$, so heißt c_n der *Leitkoeffizient* von f. f heißt *normiert*, falls $c_n = 1$.

Bemerkung VII.3.6: Seien $f, g \in K[X]$. Dann gelten:

(i) $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$,

(ii) $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis: Man zeigt die Aussagen mithilfe der Definition. Für (ii) betrachte man den Leitkoeffizienten von $f \cdot g$.

Für kommutative Ringe mit 1 kann man analog den Polynomring $R[X]$ mit Koeffizienten aus R definieren. In (ii) gilt dann im Allgemeinen aber nur noch „$\leq$“. □

Proposition VII.3.7 (Polynom-Division mit Rest): *Seien $f, g \in K[X]$ mit $g \neq 0$. Dann gibt es $h, r \in K[X]$ mit $\deg(r) < \deg(g)$ und $f = g \cdot h + r$.*

Beweis: Ist f das Nullpolynom, so stimmt die Behauptung wenn h und r das Nullpolynom sind. Man zeigt die Behauptung per Induktion nach $n = \deg(f)$, da diese Tatsache aus der Schule bekannt ist, wird der Beweis hier ausgespart.□

Wir wollen uns jetzt der Frage annehmen, welche Polynome A annullieren.

Definition VII.3.8 (Verschwindungsideal): Es seien $A \in K^{n\times n}$ und $\Phi \in \mathrm{End}(V)$. Dann heißt

$$I(A) := \{f \in K[X] \mid f(A) = \mathbf{0}\}$$

das *Verschwindungsideal von A*.

Bemerkung VII.3.9 (Eigenschaften des „Verschwindungsideals“):
(i) Das Nullpolynom ist in $I(A)$.
(ii) Das charakteristische Polynom $\mathrm{CP}_A(X)$ ist in $I(A)$ (Satz von Cayley-Hamilton).
(iii) Sind $f_1, f_2 \in I(A)$, dann ist auch $f_1 + f_2 \in I(A)$.
(iv) Sind $f \in I(A)$ und $h \in K[X]$, dann ist $h \cdot f \in I(A)$.

Definition VII.3.10 (Ideal): Sei R ein kommutativer Ring mit 1. Eine nichtleere Teilmenge $I \subseteq R$ heißt *Ideal*, falls gelten:

(i) Sind $a, b \in I$, dann ist auch $a + b \in I$,
(ii) Sind $a \in I$ und $r \in R$, dann ist auch $r \cdot a \in I$.

Proposition VII.3.11 (Konstruktion von Idealen): *Seien R ein kommutativer Ring mit 1 und $a_1, \ldots, a_n \in R$. Dann gilt:*

$$I := \{r_1 \cdot a_1 + \cdots + r_n \cdot a_n \mid r_1, \ldots, r_n \in R\}$$

ist ein Ideal von R. Wir schreiben $I = Ra_1 + \cdots + Ra_n$.

Beweis: Wir haben die Eigenschaften aus Definition VII.3.10 nachzuprüfen.

(i) Für $r_1, \ldots, r_n$ und $r'_1, \ldots, r'_n \in R$ gilt:

$$\begin{aligned}(r_1 \cdot a_1 + \cdots + r_n \cdot a_n) + (r'_1 \cdot a_1 + \cdots + r'_n \cdot a'_n)& \\ = (r_1 + r'_1) \cdot a_1 + \cdots + (r_n + r'_n) \cdot a_n \in I.&\end{aligned}$$

(ii) Seien $r \in R$ und $r_1a_1 + \cdots + r_na_n \in I$. Dann gilt

$$r \cdot (r_1 \cdot a_1 + \cdots + r_n \cdot a_n) = (r \cdot r_1) \cdot a_1 + \cdots + (r \cdot r_n) \cdot a_n \in I,$$

da $r \cdot r_i \in R$ für $1 \leq i \leq n$. □

Insbesondere ist für ein $a \in R$ die Menge $Ra := \{r \cdot a \mid r \in R\}$ ein Ideal in R.

Definition VII.3.12 (Hauptideal): Sei R ein kommutativer Ring mit 1.

(i) Ein Ideal $I \subseteq R$ heißt *Hauptideal*, falls es $m \in I$ gibt mit

$$I = \{r \cdot m \mid r \in R\} = Rm.$$

(ii) R heißt *Hauptidealring*, falls jedes Ideal in R ein Hauptideal ist.

Satz 25: *Der Polynomring $K[X]$ ist ein Hauptidealring, d. h. für jedes Ideal I in $K[X]$ gibt es ein $f \in I$ mit $I = \{h \cdot f \mid h \in K[X]\}$.*

Beweis: Sei I ein Ideal von $K[X]$. Wir wollen zeigen, dass es dann $f_0 \in I$ gibt mit $I = \{h \cdot f_0 \mid h \in K[X]\}$.

Ist $I = (0)$, dann leistet das Nullpolynom das gewünschte, d. h. $f_0 = 0$ in diesem Fall.

Andernfalls wählen wir $f_0 \in I$ so, dass $\deg(f_0)$ minimal ist. Auf jeden Fall ist $K[X]f_0 = \{h \cdot f_0 \mid h \in K[X]\}$ enthalten in I, da I ein Ideal ist. Es bleibt zu zeigen, dass $K[X]f_0 \supseteq I$; sei dazu $f \in I$. Nach Proposition VII.3.7 gibt es $h, r \in K[X]$, sodass $f = h \cdot f_0 + r$ und $\deg(r) < \deg(f_0)$. Es ist $r = f - h \cdot f_0 \in I$, wegen der Minimalität des Grades von f_0 muss r also das Nullpolynom sein und $f = h \cdot f_0$. □

Proposition VII.3.13 (Eindeutigkeit des Erzeugers): *Für ein gegebenes Ideal I ist das f_0 in Satz 25 bis auf einen skalaren Faktor eindeutig bestimmt.*

Beweis: Für $I = (0)$ stimmt die Aussage. Seien nun $I \neq (0)$ und $f_0, g_0 \in I - \{0\}$ mit $I = K[X]f_0 = K[X]g_0$. Dann gibt es $h, h' \in K[X]$ mit $g_0 = h \cdot f_0$ und $f_0 = h' \cdot g_0$, d. h. $f_0 = h' \cdot h \cdot f_0$ und $\deg(f_0) = \deg(h) + \deg(h') + \deg(f_0)$. Es muss also gelten $\deg(h) + \deg(h') = 0$, also auch $\deg(h) = \deg(h') = 0$ und wir haben alles gezeigt. □

Im Folgenden werden wir feststellen, dass wir viele Tatsachen bezüglich Teilbarkeit, die wir von den ganzen Zahlen gewohnt sind, genau so für den Polynomring vorfinden.

Bemerkung VII.3.14: (i) Wir können die Einheitengruppe von $K[X]$ mithilfe von Bemerkung VII.3.6 charakterisieren:

$$\begin{aligned} K[X]^\times &= \{f \in K[X] \mid \text{Es gibt } g \in K[X] \text{ mit } f \cdot g = g \cdot f = 1\} \\ &= \{f \in K[X] \mid \deg(f) = 0\}. \end{aligned}$$

(ii) Wir können K mit $K[X]^\times \cup \{0\}$ identifizieren via $a \mapsto a \cdot X^0$.

(iii) Sind $f, g \in K[X]$ mit $f \cdot g = 0$, dann gilt $f = 0$ oder $g = 0$ (siehe Bemerkung VII.3.6).

Definition VII.3.15: Seien $f, g \in K[X]$.

(i) f heißt *Teiler* von g, falls es $h \in K[X]$ gibt mit $g = h \cdot f$. Wir schreiben in diesem Fall $f \mid g$.

(ii) f und g heißen *teilerfremd*, falls für alle $h \in K[X]$ mit $h \mid f$ und $h \mid g$ schon gilt, dass $h \in K[X]^\times$.

(iii) Sei nun $f \notin K[X]^\times$. Dann heißt f *irreduzibel*, falls für alle $h_1, h_2 \in K[X]$ mit $f = h_1 \cdot h_2$ gilt, dass $h_1 \in K[X]^\times$ oder $h_2 \in K[X]^\times$.

Eigenschaft (iii) heißt f hat keine „echten“ Teiler.

Proposition VII.3.16 (Lemma von Bezout): *Seien $f, g \in K[X]$. f und g sind teilerfremd genau dann, wenn es $h_1, h_2 \in K[X]$ gibt mit $1 = h_1 \cdot f + h_2 \cdot g$.*[2]

Beweis: „$\Leftarrow$“: Sei $h \in K[X]$ mit $h \mid f$ und $h \mid g$, d. h. es gilt $f = h \cdot \hat{f}$ und $g = h \cdot \hat{g}$ mit $\hat{f}, \hat{g} \in K[X]$. Dann ist

$$1 = h_1 \cdot h \cdot \hat{f} + h_2 \cdot h \cdot \hat{g} = h(h_1 \cdot \hat{f} + h_2 \cdot \hat{g}),$$

d. h. $0 = \deg(1) = \deg(h) + \deg(h_1 \cdot \hat{f} + h_2 \cdot \hat{g})$, also $\deg(h) = 0$ und damit $h \in K[X]^\times$.

„$\Rightarrow$“: Sei $I := \{h_1 \cdot f + h_2 \cdot g \mid h_1, h_2 \in K[X]\}$. Nach Proposition VII.3.11 ist I ein Ideal, nach Satz 25 ist I sogar ein Hauptideal, d. h. es gibt $m \in K[X]$ mit $I = K[X] \cdot m = \{h \cdot m \mid h \in K[X]\}$. Da $f, g \in I$ teilt m insbesondere f und g, also ist $m \in K[X]^\times$, da f und g teilerfremd sind. Es ist also $\deg(m) = 0$ und $m = a \cdot X^0$ mit einem $a \in K[X] - \{0\}$. Außerdem gilt $1X^0 = a^{-1} \cdot (aX^0) \in I$, da I ein Ideal ist. □

Proposition VII.3.17 (Linearfaktoren von Polynomen): *Sei $f \in K[X]$.*

(i) *Für $a \in K$ gilt $f(a) = 0$ genau dann, wenn f von $(X - a)$ geteilt wird.*

(ii) *Falls $f = (X - a_1) \cdots (X - a_n)$ mit $a_1, \ldots, a_n \in K$, dann gilt: Ist $h \in K[X]$ ein normiertes Polynom, das f teilt, dann ist h von der Gestalt*

$$h = (X - a_{i_1}) \cdots (X - a_{i_k})$$

mit $1 \leq i_1 < i_2 < \cdots < i_k \leq n$.

[2] Hierbei identifizieren wir 1 mit $1X^0$ wie in Bemerkung VII.3.14 beschrieben.

Beweis: (i) „⇐“: Ist $f = g \cdot (X - a)$ mit $g \in K[X]$, dann ist

$$f(a) = g(a)(a - a) = g(a) \cdot 0 = 0.$$

„⇒“: Ist $f(a) = 0$, dann ergibt Polynomdivision durch $(X - a)$:

$$f = h \cdot (X - a) + r$$

mit $h, r \in K[X]$ und $\deg(r) \leq 0$, d. h. $r = cX^0$. Wegen $0 = f(a) = 0 + c$ muss $c = 0$ sein, also $r = 0$ und $(X - a)$ teilt f.

(ii) Wir zeigen die Behauptung per vollständiger Induktion auf den Grad von f. Für $n = \deg(f) = 0$ gilt $f = 1$, also $\deg(h) = 0$ nach Bemerkung VII.3.6 und da h normiert ist auch $h = 1$.

Sei die Aussage für $n \in \mathbb{N}$ gezeigt. Schreibe $f = g \cdot h$ mit $g \in K[X]$. Dann gilt: Ist

$$0 = f(a_1) = g(a_1) \cdot h(a_1),$$

dann ist $g(a_1) = 0$ oder $h(a_1) = 0$. Gilt $g(a_1) = 0$, dann wissen wir aus (i), dass $g = (X - a_1) \cdot \tilde{g}$ mit einem $\tilde{g} \in K[X]$, also

$$(X - a_1) \cdot (X - a_2) \cdots (X - a_n) = (X - a_1) \cdot \tilde{g} \cdot h$$

und damit

$$(X - a_1) \cdot [(X - a_2) \cdots (X - a_n) - \tilde{g} \cdot h] = 0.$$

Nach Bemerkung VII.3.14 ist also $\tilde{g} \cdot h = (X - a_2) \cdots (X - a_n)$ und die Induktionsvoraussetzung liefert die Behauptung.

Ist $h(a_1) = 0$, dann ist $h = (X - a_1) \cdot \tilde{h}$ mit $\tilde{h} \in K[X]$ und mit denselben Schritten wie zuvor finden wir dass $(X - a_2) \cdots (X - a_n)$ von $\tilde{h}$ geteilt wird und die Induktionsvoraussetzung liefert die Behauptung. □

Korollar VII.3.18 (aus Proposition VII.3.17): *Seien $f, g \in K[X]$ mit $f = (X - a)^n$, wobei $a \in K$ und $n \in \mathbb{N}_0$ und $g(a) \neq 0$. Dann sind f und g teilerfremd.*

Beweis: Sei h ein gemeinsamer Teiler von f und g. Wegen Proposition VII.3.17 ist h von der Form $h = (X - a)^m$ mit $0 \leq m \leq n$. Da g von h geteilt wird, muss $h(a) \neq 0$ gelten, da $g(a) \neq 0$, also $m = 0$, d. h. $h = (X - a)^0 = 1$. □

Definition VII.3.19 (Nullstellenmenge): Sei $f \in K[X]$. Wir nennen

$$\mathrm{Nst}(f) := \{a \in K \mid f(a) = 0\}$$

die *Nullstellenmenge von f*.

Bemerkung VII.3.20: Sind $f, g \in K[X]$ sodass g von f geteilt wird, dann ist $\mathrm{Nst}(f) \subseteq \mathrm{Nst}(g)$.

4. Das Minimalpolynom

In diesem Abschnitt seien V ein K-Vektorraum der Dimension n, $A \in K^{n\times n}$ und $\Phi \in \mathrm{End}(V)$.

Bemerkung VII.4.1 (Existenz von Minimalpolynom):

(i) Es gibt ein eindeutiges normiertes Polynom $m_A \in K[X]$ mit

$$I(A) = K[X] \cdot m_A = \{f \cdot m_A \mid f \in K[X]\}.$$

Insbesondere gilt: Das Minimalpolynom m_A teil das charakteristische Polynom CP_A und $m_A(A) = \mathbf{0}$ (vergleiche Satz 25 und Proposition VII.3.13).

(ii) Analog gibt es ein eindeutig bestimmtes normiertes Polynom $m_\Phi \in K[X]$ $I(\Phi) = K[X] \cdot m_\Phi$. Insbesondere wird CP_Φ von m_Φ geteilt und $m_\Phi(\Phi) = \mathbf{0}$.

(iii) Ist $A = D_B(\Phi)$, dann ist $m_A = m_\Phi$.

Definition VII.4.2 (Minimalpolynom): Das Polynom m_A (bzw. m_Φ) aus Bemerkung VII.4.1 heißt *Minimalpolynom von A* (bzw. Φ).

Proposition VII.4.3 (Minimalpolynom und Spektrum): *Die Nullstellenmenge von m_A (bzw. m_Φ) ist das Spektrum von A (bzw. von Φ), also*

$$\mathrm{Spec}(A) = \{\lambda \in K \mid m_A(\lambda) = 0\} = \mathrm{Nst}(m_A),$$
$$\mathrm{Spec}(\Phi) = \{\lambda \in K \mid m_\Phi(\lambda)\} = \mathrm{Nst}(m_\Phi).$$

Beweis: „$\supseteq$“: Ist $\lambda \in \mathrm{Nst}(m_A)$, dann gilt nach Bemerkung VII.3.20 schon, dass $\lambda \in \mathrm{Nst}(\mathrm{CP}_A) = \mathrm{Spec}(A)$. Genauso für Φ.

„$\subseteq$“: Ist λ ein Eigenwert von A, dann ist $m_A(\lambda)$ ein Eigenwert von $m_A(A)$ (wie in den Übungsgruppen besprochen), also ist $m_A(\lambda)$ ein Eigenwert der Nullmatrix $\mathbf{0}$, d. h. es ist $m_A(\lambda) = 0$ und damit $\lambda \in \mathrm{Nst}(m_A)$. □

Beispiel VII.4.4: Seien $V = K^{n\times n}$ und $\Phi\colon A \mapsto A^t$. Dann gilt $\Phi^2 = \mathrm{id}$, also $\Phi^2 - \mathrm{id} = \mathbf{0}$. Für $f = X^2 - 1$ gilt damit $f(\Phi) = \mathbf{0}$ und da es kein Polynom kleineren Grades gibt, das die Nullabbildung wird, sobald man Φ darin einsetzt, ist $f = m_\Phi$. Beachte, dass $\deg(\mathrm{CP}_\Phi) = \dim V = n^2$, aber $\deg(m_\Phi) = 2$.

Erinnerung VII.4.5: Für $A \in K^{n\times n}$ bezeichnen

(i) $\Phi_A\colon K^n \to K^n, x \mapsto Ax$ die zugehörige lineare Abbildung,

(ii) $\mathrm{Bild}(A) := \mathrm{Bild}(\Phi_A) = \{Ax \mid x \in K^n\}$ das Bild von A,

(iii) $\mathrm{Kern}(A) := \mathrm{Kern}(\Phi_A) = \{x \in K^n \mid Ax = \mathbf{0}\}$ den Kern von A.

Lemma VII.4.6 (Zerlegungslemma - Matrixversion): *Seien $g_1, g_2 \in K[X]$. Dann gelten die folgenden Aussagen:*

(i) *Sind g_1 und g_2 teilerfremd, dann sind*

$$\mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A)) = K^n, \quad \mathrm{Kern}(g_1(A)) \cap \mathrm{Kern}(g_2(A)) = \{\mathbf{0}\},$$

(ii) *Ist $(g_1 \cdot g_2)(A) = \mathbf{0}$, dann ist $\mathrm{Bild}(g_1(A)) \subseteq \mathrm{Kern}(g_2(A))$,*

(iii) *Sind $f := g_1 \cdot g_2 \in I(A)$ und g_1, g_2 teilerfremd, dann gelten*

$$K^n = \mathrm{Kern}(g_1(A)) \oplus \mathrm{Kern}(g_2(A)),$$
$$\mathrm{Kern}(g_1(A)) = \mathrm{Bild}(g_2(A)), \quad \mathrm{Kern}(g_2(A)) = \mathrm{Bild}(g_1(A)),$$

(iv) *Die Räume $U_1 := \mathrm{Kern}(g_1(A))$ und $U_2 := \mathrm{Kern}(g_2(A))$ sind A-invariant (d. h. es gilt $AU_i \subseteq U_i$ für $1 \leq i \leq 2$).*

Beweis: (i) Sind g_1, g_2 teilerfremd, dann gibt es nach dem Lemma von Bézout zwei Polynome $h_1, h_2 \in K[X]$, sodass $1 = h_1 \cdot g_1 + h_2 \cdot g_2$, es ist also

$$I_n = h_1(A) \cdot g_1(A) + h_2(A) \cdot g_2(A),$$

d. h. für alle $x \in K^n$ ist $x = I_n x = h_1(A) \cdot g_1(A)x + h_2(A) \cdot g_2(A)x$, was wir zeigen wollten.

(ii) Für $x \in K^n$ gilt

$$g_2(A)(g_1(A)x) = (g_1(A) \cdot g_2(A))x = (g_1 \cdot g_2)(A)x = \mathbf{0}_{K^{n\times n}}x = \mathbf{0}.$$

(iii) Nach (i) ist $K^n = \mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A))$, nach (ii) gilt für $i \neq j$, $i, j \in \{1, 2\}$, dass $\mathrm{Bild}(g_i(A)) \subseteq \mathrm{Kern}(g_j(A))$ und schließlich wissen wir aus (i), dass $\mathrm{Kern}(g_1(A)) \cap \mathrm{Kern}(g_2(A)) = \{\mathbf{0}\}$, d. h.

$$K^n = \mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_1(A)) \oplus \mathrm{Kern}(g_2(A)) \subseteq K^n,$$

was wir zeigen mussten.

(iv) Sei $v \in U_i = \mathrm{Kern}(g_i(A))$. Wir wollen zeigen, dass Av wieder in U_i liegt. Wegen $g_i(A) \cdot Av = A \cdot g_i(A)v = A\mathbf{0} = \mathbf{0}$ ist das der Fall. □

Korollar VII.4.7 (Zerlegungslemma - Endomorphismenversion): *Sei* $f = g_1 \cdot g_2 \in K[X]$ *mit* $f \in I(\Phi)$, *d. h.* $f(\Phi) = \mathbf{0}$, *wobei* $g_1, g_2 \in K[X]$ *teilerfremd seien. Dann gilt*

$$V = \operatorname{Kern}(g_1(\Phi)) \oplus \operatorname{Kern}(g_2(\Phi)).$$

Hierbei sind $\operatorname{Kern}(g_1(\Phi))$ *und* $\operatorname{Kern}(g_2(\Phi))$ *zwei* Φ*-invariante Unterräume.*

Korollar VII.4.8 (Zerlegungslemma für charakteristische Polynome): *Zerfällt* $\mathrm{CP}_\Phi \in K[X]$ *in Linearfaktoren, d. h. ist* $\mathrm{CP}_\Phi = \prod_{i=1}^{r}(X - \lambda_i)^{e_i}$ *mit Exponenten* $e_1, \ldots, e_r \in \mathbb{N}$ *und paarweise verschiedenen* $\lambda_1, \ldots, \lambda_r \in K$, *dann gilt:*

$$V = \bigoplus_{i=1}^{r} \operatorname{Kern}\left((\Phi - \lambda_i \operatorname{id})^{e_i}\right)$$

ist eine Zerlegung von V *in eine direkte Summe von* Φ*-invarianten Unterräumen.*

Bemerkung VII.4.9 (Reduktion auf den nilpotenten Fall): Seien

$$H := \operatorname{Kern}((\Phi - \lambda \cdot \operatorname{id})^e)$$

mit $\lambda \in K$, $e \in \mathbb{N}$ und $\psi := \Phi - \lambda \cdot \operatorname{id}$. Dann ist $\psi|_H$ nilpotent und $\Phi|_H = \psi|_H + \lambda \cdot \operatorname{id}$.

5. Nilpotente Endomorphismen

In diesem Abschnitt seien stets V ein endlichdimensionaler K-Vektorraum der Dimension n, $\Phi \in \operatorname{End}(V)$ und $A \in K^{n\times n}$.

Bemerkung VII.5.1 (Minimal– und charakteristisches Polynom, Spektrum):

Sei Φ nilpotent, d. h. es gibt $k \in \mathbb{N}$ mit $\Phi^k = \mathbf{0}$. Nach (Proposition I.3.17) ist das Minimalpolynom m_Φ ein Teiler von $\mathrm{CP}_\Phi(X) = X^k$, d. h. $m_\Phi = X^d$ mit $d \leq k$. Nach (Proposition I.4.3) sind weiter $\operatorname{Spec}(\{0\})$ und $\mathrm{CP}_\Phi(X) = X^n$ (insbesondere gilt in Proposition I.2.6 (i) „=“).

Proposition VII.5.2 (Existenz eines Φ-zyklischen Unterraums): *Es bezeichne* d *den Grad des Minimalpolynoms von* Φ. *Ist* Φ *nilpotent, so hat* V *einen* d*-dimensionalen* Φ*-zyklischen Untervektorraum* U, *d. h. es gibt* $u_0 \in V$ *mit*

(i) $U = \langle u_0, \Phi(u_0), \Phi^2(u_0), \ldots, \Phi^{d-1}(u_0)\rangle$,

(ii) U *ist* Φ*-invariant,*

(iii) *U ist d-dimensional.*

Beweis: Wähle $u_0 \in V$ mit $\Phi^{d-1}(u_0) \neq \mathbf{0}$ (das geht, weil $\Phi^{d-1} \neq \mathbf{0}_{\mathrm{End}(V)}$) und setze $U := \langle u_0, \Phi(u_0), \Phi^2(u_0), \ldots, \Phi^{d-1}(u_0)\rangle$. Da $\Phi^d(u_0) = \mathbf{0}$ ist U tatsächlich Φ-invariant. Außerdem gilt offensichtlich $\dim U \leq d$. Andererseits gilt für $f := \mathrm{CP}_{\Phi|_U}$, dass $f(X) = X^{\dim U}$ und nach dem Satz von Cayley-Hamilton ist $f(\Phi|_U) = \mathbf{0}$. Wegen $\Phi^{d-1}(u_0) \neq 0$ gilt aber $(\Phi|_U^{d-1}) \neq \mathbf{0}$ und deshalb muss $\dim U \geq d$ gelten. □

Bemerkung VII.5.3: In der Situation von Proposition VII.5.2 ist

$$B = (u_0, \Phi(u_0), \ldots, \Phi^{d-1}(u_0))$$

eine geordnete Basis von U und es ist

$$D_B(\Phi|_U) = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & 0 \end{pmatrix} =: J_d.$$

Beispiel VII.5.4: Für J_4 sind

$$J_4^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad J_4^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

und $J_4^4 = \mathbf{0}$.

Proposition VII.5.5: *Sei J_d wie in Bemerkung VII.5.3, d. h. $J_d := \sum_{i=2}^{n} E_{i,i-1}$. Dann gilt $(J_d)^e = \sum_{i=e+1}^{d} E_{i,i-e}$. Insbesondere ist*

$$\mathrm{Rang}((J_d)^e) = \begin{cases} d-e, & \text{falls } e \leq d, \\ 0, & \text{falls } e \geq d. \end{cases}$$

Beweis: Wir zeigen die Behauptung durch vollständige Induktion. Für $e = 1$ stimmt die Behauptung. Die Aussage gelte jetzt für $e \in \mathbb{N}$. Dann gilt

$$(J_d)^{e+1} = \Big(\sum_{i=e+1}^{d} E_{i,i-e}\Big) \cdot \Big(\sum_{j=2}^{d} E_{j,j-1}\Big) = \sum_{i=e+1}^{d} \sum_{j=2}^{d} E_{i,i-e}E_{j,j-1} = \sum_{i=e+2}^{d} E_{i,i-(e+1)},$$

denn $E_{i,i-e} \cdot E_{j,j-1} \neq \mathbf{0}$ genau dann, wenn $i - e = j \geq 2$. □

Proposition VII.5.6 (Φ-invariantes Komplement): *Sei Φ ein nilpotenter Endomorphismus. Der Φ-zyklische Unterraum U aus Proposition VII.5.2 besitzt ein Φ-invariantes Komplement W, d. h. es gibt einen Unterraum $W \subseteq V$ mit $V = U \oplus W$ und $\Phi(W) \subseteq W$.*

Beweis: Wie zuvor bezeichne $d = \deg(m_\Phi)$, d. h. $\Phi^d = \mathbf{0}$. Sei

$$\mathfrak{M} := \{\tilde{W} \mid \tilde{W} \text{ ist Untervektorraum von } V,\ \Phi(W) \subseteq W \text{ und } U \cap \tilde{W} = \{\mathbf{0}\}\}.$$

Es ist $\mathfrak{M} \neq \varnothing$, denn $\{\mathbf{0}\} \in \mathfrak{M}$. Wähle in $\mathfrak{M}$ ein Element W mit maximaler Dimension. Wir wollen uns jetzt davon überzeugen, dass $V = U \oplus W$. Angenommen, es gälte $U \oplus W \subsetneq V$, d. h. es gäbe $\tilde{v} \in V - U \oplus W$. Hätten wir so ein $\tilde{v}$, dann könnten wir $v \in V$ konstruieren, das $v \notin U \oplus W$ und $\Phi(v) \in U \oplus W$ leistete und für dieses v gälte dann $\Phi^d(\tilde{v}) = \mathbf{0} \in U \oplus W$. Wäre e der kleinste Exponent mit $\Phi^e(\tilde{v}) \in U \oplus W$, dann leistete $v := \Phi^{e-1}(\tilde{v})$ das gewünschte.

Für dieses v könnten wir dann schreiben $\Phi(v) = \Phi(\tilde{u}) + w$ mit $\tilde{u} \in U$ und $w \in W$. Wegen $\Phi(v) \in U \oplus W$ fänden wir eine Darstellung

$$\Phi(v) = \sum_{i=0}^{d-1} a_i \Phi^i(u_0) + w$$

mit $a_0, \ldots, a_{d-1} \in K$ und $w \in W$. Anwendung von Φ^{d-1} auf $\Phi(v)$ gäbe dann

$$\mathbf{0} = a_0 \Phi^{d-1}(u_0) + \Phi^{d-1}(w)$$

mit $\Phi^{d-1}(u_0) \in U$ und $\Phi^{d-1}(w)$. Wegen $U \cap W = \{\mathbf{0}\}$ und $\Phi^{d-1}(u_0) \neq \mathbf{0}$ müsste dann gelten, dass $a_0 = 0$ und in unserer vorherigen Darstellung hätten wir $\tilde{u} := \sum_{i=1}^{d-1} a_i \Phi^i(u_0)$.

Setzten wir jetzt $\hat{W} := \langle v - \tilde{u} \rangle$, dann wäre $\hat{W}$ auch ein Φ-invarianter Unterraum (da $\Phi(W) \subseteq W$ und $\Phi(v - \tilde{u}) = \Phi(v) - \Phi(\tilde{u}) = w \in W$) und außerdem gälte $W \subsetneq \hat{W}$, denn $v \in V - U \oplus W$ und $\tilde{u} \in U$, d. h. $v - \tilde{u} \notin W$. Schließlich gälte auch $\hat{W} \cap U = \{\mathbf{0}\}$: Für $\hat{w} \in \hat{W} \cap U$ hätten wir $\hat{w} = w' + c(v - \tilde{u}) \in U$ mit $w' \in W$ und $c \in K$. Für diese Darstellung müsste aber $c \neq 0$ gelten (andernfalls könnten wir ja schreiben $v = c^{-1}(\hat{w} + c\tilde{u} - w) \in U \oplus W$, was nicht sein kann), also hätten wir $\hat{w} = w' \in U \cap W$, d. h. $\hat{w} = \mathbf{0}$. Die Existenz von $\hat{W}$ stünde aber im Widerspruch der Maximalität der Dimension von W. □

Korollar VII.5.7 (Zerlegung in Φ-invariante Unterräume): *Ist Φ nilpotent, so ist V direkte Summe von Φ-zyklischen Unterräumen.*

Beweis: Das können wir per Induktion unter Verwendung von Proposition VII.5.6 zeigen. □

Korollar VII.5.8 (Jordan-Normalform für nilpotente Endomorphismen): *Ist Φ nilpotent, so gibt es eine geordnete Basis B von V mit*

$$D_B(\Phi) = \begin{pmatrix} J_{d_1} & & & \\ & J_{d_2} & & \\ & & \ddots & \\ & & & J_{d_k} \end{pmatrix}$$

wobei $d_1 \geq d_2 \geq \cdots \geq d_k \geq 1$. Die J_{d_i} heißen Jordan-Kästchen.

Beweis: Folgt aus Korollar VII.5.7 und Bemerkung VII.5.3. □

Proposition VII.5.9: *Für die Matrix $D_B(\Phi) =: A$ in Korollar VII.5.8 gelten:*

(i) $d_1 + \cdots + d_k = n$,

(ii) *Für die Anzahl m_k von Jordan-Kästchen der Größe k gilt:*

$$m_k = \operatorname{Rang}(A^{k-1}) - 2\operatorname{Rang}(A^k) + \operatorname{Rang}(A^{k+1}).$$

(iii) *Das größte Jordan-Kästchen hat die Größe $d = \deg(m_\Phi)$.*

Beweis: (i) ist klar und (iii) folgt aus (ii).

Zu (ii): Wegen Proposition VII.5.5 wissen wir, dass für das Jordan-Kästchen J_p der Größe p gilt: $J_p^e = \mathbf{0}$ für $p \leq e$ und $\operatorname{Rang}(A^e) = \sum_{p \geq e+1}^{n} (p-e) \cdot m_p$. Insbesondere gilt für $1 \leq k \leq n$:

$$\begin{aligned} \operatorname{Rang}(A^{k-1}) &= m_k + 2m_{k+1} + 3m_{k+2} + 4m_{k+3} + \ldots + (n-k-1)m_n, \\ \operatorname{Rang}(A^k) &= \quad m_{k+1} + 2m_{k+2} + 3m_{k+3} + \ldots + (n-k)m_n, \\ \operatorname{Rang}(A^{k+1}) &= \quad m_{k+2} + 2m_{k+3} + \ldots + (n-(k+1))m_n, \end{aligned}$$

und das gibt die Behauptung, denn $i - 2(i-1) + (i-2) = 0$. □

Korollar VII.5.10 (Jordan-Normalform für nilpotente Matrizen): *Ist A eine nilpotente Matrix, dann gibt es $B \in \mathrm{Gl}_n(K)$ mit*

$$BAB^{-1} = \begin{pmatrix} J_{d_1} & & \\ & \ddots & \\ & & J_{d_k} \end{pmatrix}$$

wobei $d_1, \ldots, d_k \in \mathbb{N}$, $d_1 \geq d_2 \geq \cdots \geq d_k \geq 1$. Es gelten die Aussagen für die d_i's aus Proposition VII.5.9; die J_d sind Jordan-Kästchen wie in Bemerkung VII.5.3 definiert.

6. Jordan'sche Normalform

Für diesen Abschnitt seien V ein K-Vektorraum der Dimension n, $\Phi \in \mathrm{End}(V)$ und $A \in K^{n\times n}$.

Wir erinnern uns, dass Φ diagonalisierbar ist genau dann, wenn

$$V = \bigoplus_{\lambda\in\mathrm{Spec}(\Phi)} \mathrm{Eig}(\Phi, \lambda).$$

Die Eigenräume $\mathrm{Eig}(\Phi, \lambda)$ sind Φ-invariante Unterräume von V.

Ist Φ nicht diagonalisierbar, hat $\mathrm{Eig}(\Phi, \lambda)$ für $\lambda \in \mathrm{Spec}(\Phi)$ im Allgemeinen kein Φ-invariantes Komplement.

Beispiel VII.6.1: Wir betrachten die lineare Abbildung $\phi\colon K^3 \to K^3, x \mapsto Ax$ mit

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

Es sind $\mathrm{CP}_\Phi(X) = (-1)^3X^3$ und $\mathrm{Eig}(\Phi, 0) = \langle e_3\rangle$. Für ein $x = (x_1, x_2, x_3)$ gilt $Ax = (0, x_1, x_2)^t$, $A^2x = (0, 0, x_1)^t \in \mathrm{Eig}(\Phi, 0)$, also hat $\mathrm{Eig}(\Phi, 0)$ kein Φ-invariantes Komplement.

Wir suchen jetzt die kleinsten Φ-invarianten Unterräume H_λ mit den Eigenschaften

- $\mathrm{Eig}(\Phi, \lambda) \subseteq H_\lambda$,
- H_λ ist Φ-invariant,
- H_λ hat ein Φ-invariantes Komplement.

Bemerkung VII.6.2: Sei $U \subseteq V$ ein Untervektorraum. U ist Φ-invariant genau dann, wenn U für ein $\lambda \in K$ invariant unter $(\Phi - \lambda\,\mathrm{id})$ ist. Das gilt genau dann, wenn U für alle $\lambda \in K$ invariant unter $(\Phi - \lambda\,\mathrm{id})$ ist.

Beweis: Für „$\Rightarrow$“: Ist $v \in U$, dann ist $(\Phi - \lambda\,\mathrm{id})v = \Phi(v) - \lambda v \in U$ für alle $\lambda \in K$. Die andere Richtung zeigt man analog. □

Proposition VII.6.3: *Sei $V = U \oplus W$ eine Zerlegung von V in Φ-invariante Untervektorräume mit $\mathrm{Eig}(\Phi, \lambda) \subseteq U$. Dann ist $\mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^k) \subseteq U$ für alle $k \in \mathbb{N}$.*

Beweis: Wir zeigen die Behauptung via Induktion nach k. Für $k = 1$ gilt $\mathrm{Eig}(\Phi, \lambda) = \mathrm{Kern}(\Phi - \lambda\,\mathrm{id}) \subseteq U$ nach Voraussetzung.

Sei die Aussage für ein $k \in \mathbb{N}$ gezeigt. Wir wollen zeigen, dass dann gilt: $\mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^{k+1}) \subseteq U$. Für $v \in \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^{k+1})$ ist $(\Phi - \lambda\,\mathrm{id})^{k+1}v = \mathbf{0}$, d. h. $(\Phi - \lambda\,\mathrm{id})^k v \in \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})) = \mathrm{Eig}(\Phi, \lambda) \subseteq U$.

Schreiben wir jetzt $v = u + w$ mit $u \in U$, $w \in W$, dann ist

$$U \oplus W \ni (\Phi - \lambda\,\mathrm{id})^k(u) + (\Phi - \lambda\,\mathrm{id})^k(w) = (\Phi - \lambda\,\mathrm{id})^k(u + w) \in U,$$

es muss also, da $U \oplus W$ eine direkte Summe ist, gelten, dass $(\Phi - \lambda\,\mathrm{id})^k w = \mathbf{0}$. Es ist also $w \in \mathrm{Kern}(\Phi - \lambda\,\mathrm{id})^k \subseteq U$ nach Induktionsvoraussetzung, damit ist $w = \mathbf{0}$ und $v \in U$. □

Definition VII.6.4: Wir nennen

$$H(\Phi, \lambda) := \bigcup_{k=0}^{\infty} \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^k)$$

den *Hauptraum von* Φ *zum Eigenwert* λ.

Bemerkung VII.6.5: (i) Es ist

$$\mathrm{Eig}(\Phi, \lambda) = \mathrm{Kern}(\Phi - \lambda\,\mathrm{id}) \subseteq \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^2) \subseteq \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^3) \subseteq \ldots,$$

weshalb die Haupträume tatsächlich Vektorräume sind.

(ii) Die Unterräume $H(\Phi, \lambda)$ sind Φ-invariant, denn für eine natürliche Zahl k gilt $v \in \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^k)$ genau dann, wenn $(\Phi - \lambda\,\mathrm{id})^k v = \mathbf{0}$. Es ist

$$((\Phi - \lambda\,\mathrm{id})^k \circ \Phi)(v) = (\Phi \circ (\Phi - \lambda\,\mathrm{id})^k)(v),$$

d. h. $\mathbf{0} = \Phi(\mathbf{0}) = \Phi(v) \in \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^k)$.

Proposition VII.6.6 (Eigenschaften der Haupträume):

(i) *Die Haupträume* $H(\Phi, \lambda)$, $\lambda \in \mathrm{Spec}(A)$, *haben* Φ*-invariante Komplemente.*

(ii) $H(\Phi, \lambda) = \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^e)$, *wobei* e *die algebraische Vielfachheit von* $\lambda \in \mathrm{Spec}(A)$ *ist,*

(iii) $\dim(H(\Phi, \lambda)) = e$.

Beweis: (i) Es gibt $g \in K[X]$ mit $\mathrm{CP}_\Phi(X) = (X - \lambda)^e g$, wobei $g(\lambda) \neq 0$. Wir setzen $H := \mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^e)$ und $W := \mathrm{Kern}(g(\Phi))$. Nach Korollar VII.4.7 wissen wir, dass $V = H \oplus W$ eine Zerlegung von V in Φ-invariante Unterräume von V ist. Mit (ii) ergibt das die Behauptung aus (i).

(ii) Es ist $g(\Phi)|_W \equiv 0$ und $g(\lambda) \neq 0$, d. h. $\lambda \notin \mathrm{Spec}(\Phi|_W)$. Damit ist $(\Phi - \lambda\,\mathrm{id})^n|_W$ injektiv für alle $n \in \mathbb{N}$. Es ist also für alle $k \in \mathbb{N}$ schon

$$\mathrm{Kern}((\Phi - \lambda\,\mathrm{id})^{e+k}) = H.$$

Das macht unseren Beweis für (i) vollständig.

(iii) Es ist $V = H \oplus W$ eine direkte Summe von Φ-invarianten Unterräumen, wir können also schreiben

$$\mathrm{CP}_\Phi(X) = \mathrm{CP}_{\Phi|_H}(X) \cdot \mathrm{CP}_{\Phi|_W}(X). \tag{VII.1}$$

Nach Korollar VII.4.7 ist $\mathrm{CP}_{\Phi|_H}(X) = (X-\lambda)^n$ für ein $n \in \mathbb{N}$. Wegen $g(\lambda) \neq 0$ ist $e \leq n$, wegen Gl. (VII.1) ist aber auch $e \geq n$, also ist $e = n$. Damit ist $\dim(H) = \deg(\mathrm{CP}_{\Phi|_H}(X)) = e$. □

Korollar VII.6.7: *Der Hauptraum $H(\Phi, \lambda)$ ist der kleinste Φ-invariante Unterraum von V, der $\mathrm{Eig}(\Phi, \lambda)$ enthält, und ein Φ-invariantes Komplement hat.*

Beweis: Direkte Konsequenz aus Proposition VII.6.3 und Proposition VII.6.6.□

Proposition VII.6.8 (Direktheit der Summe der Haupträume): *Seien $\lambda_1, \ldots, \lambda_k$ Eigenwerte von Φ. Dann ist*

$$\sum_{i=1}^{k} H(\Phi, \lambda_i) = \bigoplus_{i=1}^{k} H(\Phi, \lambda_i).$$

Beweis: Wir zeigen die Behauptung per Induktion nach k. Für $k = 0$ und $k = 1$ ist die Aussage klar. Sei die Aussage jetzt gezeigt für die natürliche Zahl $k-1$ und $v_i \in H_i$ mit $\sum_{i=1}^{k} v_i = \mathbf{0}$. Wir haben zu zeigen, dass $v_i = \mathbf{0}$ für $1 \leq i \leq k-1$. Sei $e > 0$ mit $(\Phi - \lambda_k\,\mathrm{id})^e v_k = \mathbf{0}$. Dann ist

$$\mathbf{0} = \sum_{i=1}^{k} (\Phi - \lambda\,\mathrm{id})^e(v_i) = \sum_{i=1}^{k-1} (\Phi - \lambda\,\mathrm{id})^e(v_i),$$

nach Bemerkung VII.6.2 gilt $(\Phi - \lambda_k\,\mathrm{id})^e v_i \in H_i$ und nach der Induktionsvoraussetzung gilt damit $(\Phi - \lambda_k\,\mathrm{id})^e(v_i) = \mathbf{0}$ für $1 \leq i \leq k-1$. Für ein $1 \leq i \leq k-1$ gibt es für $v_i \in H_i$ ein $f_i > 0$ mit $(\Phi - \lambda_i\,\mathrm{id})^{f_i}(v_i) = \mathbf{0}$. Die Polynome $(X-\lambda_k)^e$ und $(X-\lambda_i)^{f_i}$ sind teilerfremd, d. h. wir finden $g, h \in K[X]$ mit

$$g(X-\lambda_k)^e + h(X-\lambda_i)^{f_i} = \mathrm{id}\,.$$

Für v_i ist

$$v_i = \mathrm{id}(v_i) = g(\Phi)(\Phi - \lambda_k\,\mathrm{id})^e(v_i) + h(\Phi)(\Phi - \lambda_i)^{f_i}(v_i) = \mathbf{0},$$

also $v_i = \mathbf{0}$ für $1 \leq i \leq k-1$. Dann ist aber auch $v_k = \mathbf{0}$ und die Summe $\sum_{i=1}^{k} H(\Phi, \lambda_i)$ ist direkt. □

Proposition VII.6.9: *Die folgenden Aussagen sind äquivalent:*

(i) $V = \bigoplus_{\lambda \in \mathrm{Spec}(\Phi)} H(\Phi, \lambda)$,

(ii) $\mathrm{CP}_\Phi(X) = \prod_{\lambda \in \mathrm{Spec}(\Phi)} (X - \lambda)^{\dim(H(\Phi,\lambda))}$,

(iii) $\mathrm{CP}_\Phi(X)$ *zerfällt in Linearfaktoren.*

(iv) *Das Minimalpolynom* m_Φ *zerfällt in Linearfaktoren.*

Beweis: „(i) $\Rightarrow$ (ii)“: In Beweis von Proposition VII.6.6 haben wir gesehen, dass $\mathrm{CP}_{\Phi|_{H(\Phi,\lambda)}}(X) = (X-\lambda)^{\dim H(\Phi,\lambda)}$. Da die Haupträume eine direkte Summe bilden und Φ-invariante Unterräume von V sind, wissen wir, dass

$$\mathrm{CP}_\Phi = \prod_{\lambda \in \mathrm{Spec}(\Phi)} \mathrm{CP}_{\Phi|_{H(\Phi,\lambda)}} = \prod_{\lambda \in \mathrm{Spec}(\Phi)} (X - \lambda)^{\dim H(\Phi,\lambda)}.$$

„(ii) $\Rightarrow$ (iii)“: Klar.

„(iii) $\Rightarrow$ (iv)“: Klar, da m_Φ ein Teiler des charakteristischen Polynoms ist.

„(iv) $\Rightarrow$ (i)“: Es bezeichne im Folgenden $N_m := \mathrm{Nst}(m_\Phi(X))$. Wir zeigen die Behauptung per Induktion nach der Anzahl k der Nullstellen von m_Φ.

Für $k = 0$ muss $m_\Phi \in K[X]^\times$ gelten und dann ist $V = \{\mathbf{0}\}$. Für $k = 1$ ist $m_\Phi = (X - \lambda)^e$ mit $e > 0$ und $H(\Phi, \lambda) = V$.

Sei die Aussage jetzt gezeigt für die natürliche Zahl $k - 1$. Wir schreiben

$$m_\Phi = \prod_{i=1}^{k} (X - \lambda_i)^{e_i}.$$

Wie im Beweis von Proposition VII.6.6 gibt es einen Φ-invarianten Unterraum W von V mit $V = H(\Phi, \lambda) \oplus W$ und es ist $m_\Phi|_W(\lambda) \neq 0$. Zudem gilt

$$m_\Phi = m_{\Phi|_W} \cdot m_{\Phi|_{H(\Phi,\lambda)}}$$

Da $m_\Phi|_W(\lambda) \neq 0$ hat $m_{\Phi|_W}$ weniger als k Nullstellen, d. h. wir können die Induktionsvoraussetzung auf $m_{\Phi|_W}$ anwenden und sehen, dass W eine direkte Summe von Haupträumen ist. Damit ist also auch V eine direkte Summe von Haupträumen. □

Beispiel VII.6.10: Es seien $K = \mathbb{R}$, $\Phi\colon K^2 \to K^2$, $x \mapsto Ax$ mit $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Es ist $A^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Es ist $m_\Phi(X) = X^2 + 1$, dieses Polynom hat keine reellen Nullstellen, d. h. $\mathrm{Spec}(A) = \varnothing$ und $H(\Phi, \lambda) = \{\mathbf{0}\}$ für $\lambda \in K$.

Definition VII.6.11 (Jordankästchen): Für $\lambda \in K$, $d \in \mathbb{N}$ sei

$$J_d(\lambda) = \lambda I_d + J_d(0) = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}.$$

Die Matrix $J_d(\lambda)$ heißt *Jordankästchen der Länge d zum Eigenwert λ.*

Bemerkung VII.6.12: Sei $\Phi \in \mathrm{End}(V)$ mit $\mathrm{CP}_\Phi(X) = \prod_{\lambda \in \mathrm{Spec}(\Phi)} (X - \lambda)^{e_\lambda}$. Nach Proposition VII.6.9 ist $V = \bigoplus_{\lambda \in \mathrm{Spec}(\Phi)} H(\Phi, \lambda)$; nach Proposition VII.6.6 ist $H(\Phi, \lambda) = \mathrm{Kern}((\Phi - \lambda \,\mathrm{id})^{e_\lambda})$ für $\lambda \in \mathrm{Spec}(\Phi)$.

Sei $\lambda \in \mathrm{Spec}(\Phi)$. Es ist $\Psi_\lambda := (\Phi - \lambda \,\mathrm{id})|_{H(\Phi,\lambda)}$ nilpotent, d. h. Ψ_λ wird durch eine nilpotente Abbildungsmatrix beschrieben.

Nach Korollar VII.5.8 und Korollar VII.5.10 können wir die Abbildungsmatrix in Jordan-Normalform bringen (durch Basiswechsel), d. h. es gibt eine Basis B, sodass

$$\lambda I_{e_\lambda} + \tilde{A} = D_B(\Phi|_{H(\Phi,\lambda)}) = D_B(\Psi_\lambda + \lambda \,\mathrm{id}),$$

wobei $\tilde{A}$ wie in Korollar VII.5.10.

Satz 26 (Jordan'sche Normalform): *Seien V ein K-Vektorraum der Dimension n und $\phi \in \mathrm{End}\, V$. Ferner sei $\mathrm{Spec}(\phi) = \{\lambda_1, \ldots, \lambda_\ell\}$ und es gelte*

$$\mathrm{CP}_\phi(X) = \prod_{i=1}^{\ell} (X - \lambda_i)^{\mu_a(\lambda_i)}.$$

Dann existiert eine Basis B von V und für alle $\lambda_i \in \mathrm{Spec}(\phi)$ existieren k_i und $d_{1,i} \geq d_{2,i} \geq \cdots \geq d_{k,i} \geq 1$ in $\mathbb{N}$, sodass

$$D_{B,B}(\phi) = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_\ell \end{pmatrix}.$$

Für $1 \leq i \leq \ell$ ist hierbei

$$D_i = \begin{pmatrix} J_{d_{1,i}}(\lambda_i) & & \\ & \ddots & \\ & & J_{d_{k,i}}(\lambda_i) \end{pmatrix}.$$

Wir nennen $m_d(\lambda)$ die Anzahl der Jordankästchen der Länge d zum Eigenwert λ.

Beweis: Die nötigen Ergebnisse haben wir in den Abschnitten 5 und 6 zusammengetragen (vergleiche insbesondere Bemerkung VII.6.12). □

Proposition VII.6.13 (Kenngrößen): *Wir verwenden die Notation aus Satz 26. Für* $\lambda \in \operatorname{Spec}(\phi)$ *gilt:*

(i) *Es ist* $\#\operatorname{Spec}(\phi) = \ell$ *die Anzahl der Jordanblöcke* D_i*. Die Basisvektoren in* B *zu* D_i *bilden eine Basis* B_i *von* $H(\phi, \lambda_i)$ *mit* $D_{B_i,B_i}(\phi|_{H(\phi,\lambda_i)}) = D_i$*.*

(ii) *Die Länge eines Jordanblocks* D_i *ist gerade* $\dim H(\phi, \lambda_i) = \mu_a(\lambda)$*.*

(iii) *Die Anzahl der Jordankästchen im Jordanblock* D_i *ist gerade* $\dim \operatorname{Eig}(\phi, \lambda_i)$*, die geometrische Vielfachheit* $\mu_g(\lambda)$ *des Eigenwerts* λ_i*.*

(iv) *Die Länge des längsten Jordankästchens im Jordanblock* D_i *ist der Exponent von* λ_i *im Minimalpolynom.*

(v) *Für alle* $d \in \mathbb{N}$ *gilt:*

$$m_d(\lambda) = \operatorname{Rang}((\phi - \lambda \operatorname{id})^{d-1}) - 2\operatorname{Rang}((\phi - \lambda \operatorname{id})^{d}) + \operatorname{Rang}((\phi - \lambda \operatorname{id})^{d+1}).$$

Beweis: (i) Dies haben wir gezeigt in Proposition VII.6.9.

(ii) Das haben wir in Proposition VII.6.6 eingesehen.

(iii) Wir wissen

$$\begin{aligned}\dim \operatorname{Eig}(\phi, \lambda) &= \dim \operatorname{Eig}(\phi - \lambda \operatorname{id}, \mathbf{0}) \\ &= \dim \operatorname{Kern}(\phi - \lambda \operatorname{id}) = \#\text{ Jordankästchen.}\end{aligned}$$

(iv) $\phi - \lambda \operatorname{id}$ ist nilpotent auf $H(\phi, \lambda)$. Die Länge des längsten Jordankästchens ist gerade die Dimension des größten zyklischen Unterraums von $H(\phi, \lambda)$. Das ist aber gerade der sogenannte Nilpotenzindex von $\phi - \lambda \operatorname{id}|_{H(\phi,\lambda)}$, d. h. $\min\{d \in \mathbb{N} \mid A^d = 0\}$. Das ist genau der Exponent von λ im Minimalpolynom von ϕ.

(v) $\phi - \lambda \operatorname{id}|_{H(\Phi,\lambda)}$ ist nilpotent, d. h. wir sind in der Situation, Proposition VII.5.9 verwenden zu können. □

Proposition VII.6.14: *Die Jordan'sche Normalform ist eindeutig bis auf Reihenfolge der Jordanblöcke.*

Beweis: Die Kenngrößen aus Proposition VII.6.13 sind eindeutig. □

Korollar VII.6.15 (Jordan'sche Normalform für Matrizen): *Sei $A \in K^{n\times n}$ mit $\mathrm{Spec}(A) = \{\lambda_1, \ldots, \lambda_\ell\}$ und $\mathrm{CP}_A(X) = \prod_{i=1}^{\ell}(X-\lambda_i)^{\mu_a(\lambda_i)}$. Dann gibt es $B \in \mathrm{Gl}_n(K)$, sodass $J(A) := BAB^{-1}$ Jordan'sche Normalform (wie in Satz 26 beschrieben) hat.*

Beweis: Folgt aus Satz 26 mit $\phi\colon K^n \to K^n, v \mapsto Av$. □

Beispiel VII.6.16: Sei

$$A = \begin{pmatrix} 0 & -6 & 0 & 3 & 3 \\ 1 & 0 & -2 & 0 & 1 \\ 0 & -4 & 0 & 2 & 2 \\ 2 & -1 & -4 & 1 & 2 \\ 0 & -2 & 0 & 1 & 1 \end{pmatrix}.$$

Wir wollen im Folgenden die Jordan'sche Normalform von A bestimmen.

(1) *Berechnung von* CP_A: Es ist

$$\mathrm{CP}_A(X) = \det \begin{pmatrix} X & 6 & 0 & -3 & -3 \\ -1 & X & 2 & 0 & -1 \\ 0 & 4 & X & -2 & -2 \\ -2 & 1 & 4 & X-1 & -2 \\ 0 & 2 & 0 & -1 & X-1 \end{pmatrix} = X^3(X-1)^2.$$

(2) *Bestimmung von* $\mathrm{Spec}(A)$, μ_a *und* μ_g: Offensichtlich ist $\mathrm{Spec}(A) = \{0, 1\}$, ferner sind $\mu_a(0) = 3$, sowie $\mu_a(1) = 2$. Weiter ist $\mathrm{Rang}(A) = 3$, also ist $\dim \mathrm{Kern}(A) = \dim \mathrm{Eig}(\phi, 0) = 2$, also $\mu_g(0) = 2 \neq 3 = \mu_a(0)$.

Bringt man $A - I_5$ in Treppenform, so kann man $\mathrm{Rang}(A - I_5) = 4$ ablesen, was $\dim \mathrm{Kern}(A - I_5) = \dim \mathrm{Eig}(A, 1) = 1$ bedeutet, also $\mu_g(1) = 1 \neq 2 = \mu_a(1)$.

Da die geometrischen– und algebraischen Vielfachheiten nicht übereinstimmen, ist A nicht diagonalisierbar.

(3) *Finde die Jordan'sche Normalform*: Aus Proposition VII.6.13 lesen wir ab

$$J(A) = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

und $\mu_A(X) = X^2(X-1)^2$ ist das Minimalpolynom von A.

Proposition VII.6.17 (Jordan-Zerlegung): *Sei $A \in K^{n\times n}$ mit Jordan'scher Normalform $J(A)$. Dann gibt es eine eindeutige Zerlegung $A = D + N$ mit:*

(i) *D ist diagonalisierbar,*

(ii) *N ist nilpotent,*

(iii) *$DA = AD$, $NA = AN$.*

Diese Zerlegung heißt Jordan-Zerlegung (oder auch Jordan-Chevalley-Zerlegung oder Dunford-Zerlegung).

Beweis: Setze $A' := J(A)$, dann ist $A' = D' + N'$, wobei D' eine Diagonalmatrix und N' eine Matrix mit Einsen auf der Nebendiagonale ist. Man kann nachrechnen, dass gilt $D'A' = A'D'$, $N'A' = A'N'$. Nach Korollar VII.6.15 gibt es $S \in \mathrm{Gl}_n(K)$ mit $SA'S^{-1} =: A$. Setze $D := SD'S^{-1}$, $N := SN'S^{-1}$; D ist diagonalisierbar und $N^k = (SN'S^{-1})^k = SN'^kS^{-1} = 0$ für ausreichend großes k (da N' nilpotent ist), d. h. N ist nilpotent. Ferner ist

$$A = S(D' + N')S^{-1} = SD'S^{-1} + SN'S^{-1} = D + N,$$

womit wir (i) und (ii) gezeigt haben.

Für (iii): Es ist

$$DA = SD'S^{-1}SA'S^{-1} = SD'A'S^{-1} = SA'D'S^{-1} = SA'S^{-1}SD'S^{-1} = AD,$$

genau so für $NA = AN$. Es bleibt zu zeigen, dass die Zerlegung eindeutig ist. Sei dazu $A = \hat{D} + \hat{N}$ eine weitere Jordan-Zerlegung. Wegen (iii) gilt für alle $k \in \mathbb{N}$ und $\lambda, \alpha \in K$:

$$\hat{D}(A - \lambda I)^k = (A - \lambda I)^k \hat{D}, \qquad \hat{N}(A - \alpha I)^k = (A - \alpha I)^k \hat{N}.$$

Außerdem ist der Hauptraum $H(A, \lambda)$ für alle $\lambda \in \mathrm{Spec}(A)$ invariant unter $\hat{D}$, denn $v \in H(A, \lambda)$ genau dann wenn es $k \in \mathbb{N}$ gibt mit $(A - \lambda I)^k v = \mathbf{0}$, d. h. in diesem Fall ist

$$(A - \lambda I)^k \hat{D}(v) = \hat{D}(A - \lambda I)^k v = \hat{D}\mathbf{0} = \mathbf{0},$$

und $\hat{D}v \in H(A, \lambda)$.

Schließlich gilt für alle $\lambda \in \mathrm{Spec}(A)$, dass $\hat{D}|_{H(A,\lambda)} = D|_{H(A,\lambda)}$, denn:

$$V = H(A, \lambda) \oplus \bigoplus_{\lambda \neq \lambda' \in \mathrm{Spec}(A)} H(A, \lambda')$$

ist eine $\hat{D}$-invariante Zerlegung von V als direkte Summe, also ist $\hat{D}|_{H(A,\lambda)}$ diagonalisierbar. Ist $v \in H(A,\lambda)$ ein Eigenvektor von $\hat{D}$ zum Eigenwert α, dann ist $Av = (\hat{D} + \hat{N})v = \alpha v + \hat{N}v$, d. h. $(A - \alpha I)v = \hat{N}v$. Wegen unserer vorherigen Überlegung haben wir für alle $k \in \mathbb{N}$, dass $(A - \alpha I)^k v = \hat{N}v$. Da $\hat{N}$ nilpotent ist, gibt es $k \in \mathbb{N}$, sodass $\hat{N}^k v = \mathbf{0}$, also $(A - \alpha I)^k v = \mathbf{0}$, es ist also $v \in H(A,\alpha)$. Aber jetzt muss $\alpha = \lambda$ gelten, da die Summe der Haupträume direkt ist und $v \neq \mathbf{0}$.

Insgesamt haben wir also gesehen, dass $\hat{D} = D$ und damit auch $\hat{N} = N$. $\square$

Kapitel VIII.

Euklidische und unitäre Vektorräume

Erinnerung: • Ein euklidischer Vektorraum ist ein endlichdimensionaler $\mathbb{R}$-Vektorraum zusammen mit einem Skalarprodukt,

• Ein unitärer Vektorraum ist ein endlichdimensionaler $\mathbb{C}$-Vektorraum zusammen mit einem Skalarprodukt.

1. Singulärwertzerlegung

In diesem Abschnitt seien $V = \mathbb{R}^n$ mit dem Standardskalarprodukt $\langle x, y\rangle = x^t y$ oder $V = \mathbb{C}^n$ mit dem Standardskalarprodukt $\langle x, y\rangle = x^t \overline{y}$.

Erinnerung VIII.1.1 (Orthogonale bzw. unitäre Matrizen):

(i) $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal* genau dann, wenn eine (und somit alle) der folgenden äquivalenten Bedingungen erfüllt ist:

(1) Die lineare Abbildung $x \mapsto Ax$ erhält das Standardskalarprodukt, d. h. für alle $x, y \in \mathbb{R}^n$ gilt $\langle Ax, Ay\rangle = \langle x, y\rangle$.

(2) Es ist $A^t A = I$, insbesondere ist A invertierbar mit $A^{-1} = A^t$.

(3) Die Spaltenvektoren von A bilden eine Orthonormalbasis des $\mathbb{R}^n$ bezüglich des Standardskalarprodukts.

(4) Die Zeilenvektoren von A bilden eine Orthonormalbasis des $\mathbb{R}^n$ bezüglich des Standardskalarprodukts.

(ii) $A \in \mathbb{C}^{n\times n}$ heißt *unitär* genau dann, wenn eine (und somit alle) der folgenden äquivalenten Bedingungen erfüllt ist:

(1) Die lineare Abbildung $x \mapsto Ax$ erhält das Standardskalarprodukt, d. h. für alle $x, y \in \mathbb{C}^n$ gilt $\langle Ax, Ay\rangle = \langle x, y\rangle$.

(2) Es ist $A^*A = I$, insbesondere ist A invertierbar mit $A^{-1} = A^*$.
(3) Die Spaltenvektoren von A bilden eine Orthonormalbasis des $\mathbb{C}^n$ bezüglich des Standardskalarprodukts.
(4) Die Zeilenvektoren von A bilden eine Orthonormalbasis des $\mathbb{C}^n$ bezüglich des Standardskalarprodukts.

Hierbei ist $A^* = \overline{A}^t$.

(iii) Wir nennen

$$O(n) := \{A \in \mathbb{R}^{n\times n} \mid A \text{ orthogonal}\}, \quad U(n) := \{A \in \mathbb{C}^{n\times n} \mid A \text{ unitär}\}$$

die orthogonale bzw. unitäre Gruppe. Beide sind Untergruppen bezüglich Matrizenmultiplikation von $\mathrm{Gl}_n(\mathbb{R})$ bzw. $\mathrm{Gl}_n(\mathbb{C})$.

Erinnerung VIII.1.2 (Diagonalisierbarkeit mit Basiswechsel aus $O(n)$, $U(n)$): Die hier zusammengestellten Aussagen finden sich im „Lineare Algebra I"-Mitschrieb aus dem Wintersemester 2018/2019 in den Sätzen 18 und 20.

(i) Ist $A \in U(n)$, dann gibt es $S \in U(n)$ mit

$$S^*AS = S^{-1}AS = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

(ii) Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, dann gibt es $S \in O(n)$ mit

$$S^tAS = S^{-1}AS = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

(iii) Ist $A \in \mathbb{C}^{n\times n}$ hermitesch, dann gibt es $S \in U(n)$ mit

$$S^*AS = S^{-1}AS = \mathrm{diag}(\lambda_1, \dots, \lambda_n),$$

wobei $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

Im Folgenden wollen wir Matrizen bis auf Äquivalenz untersuchen, d. h. wir wollen unterschiedliche Basiswechsel im Definitionsbereich und im Bildbereich zulassen.

Satz 27 (Singulärwertzerlegung): *Sei $A \in \mathbb{R}^{m\times n}$ eine beliebige Matrix.*

(i) *Ist $m \le n$, dann gilt: Es gibt $U_1 \in O(m)$, $U_2 \in O(n)$ und $\mu_1, \dots, \mu_n \in \mathbb{R}_{\ge 0}$ mit*

$$U_1AU_2 = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \mu_2 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \mu_n & 0 & \cdots & 0 \end{pmatrix} =: \Sigma \in \mathbb{R}^{m\times n}. \qquad \text{(VIII.1)}$$

Die Zahlen $\mu_1, \dots, \mu_n$ sind eindeutig durch die Matrix A bestimmt und heißen die Singulärwerte von A. *Die Quadrate $\mu_1^2, \dots, \mu_m^2$ sind die Eigenwerte der symmetrischen Matrix $AA^t \in \mathbb{R}^{m\times m}$.*

(ii) *Ist $n < m$, dann erhalten wir eine analoge Aussage zu (i) mit*

$$U_1 A U_2 = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mu_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} =: \Sigma \in \mathbb{R}^{m\times n}. \tag{VIII.2}$$

Die Zahlen $\mu_1^2, \dots, \mu_n^2$ sind die Eigenwerte von $A^tA \in \mathbb{R}^{n\times n}$.

Lemma VIII.1.3: *Sei $A \in \mathbb{R}^{n\times m}$.*

(i) *A^tA ist symmetrisch und ist $\lambda \in \mathrm{Spec}(A^tA)$, dann ist $\lambda \geq 0$.*

(ii) *Es ist* $\mathrm{Kern}(A^tA) = \mathrm{Kern}(A)$ *und also insbesondere* $\mathrm{Rang}(A^tA) = \mathrm{Rang}(A)$.

Beweis: (i) Die Symmetrie von A^tA ist klar. Ist λ ein Eigenwert von A^tA zum Eigenvektor v, dann gilt

$$\|Av\|^2 = \langle Av, Av\rangle = v^tA^tAv = \langle v, A^tAv\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2,$$

d. h. $\lambda \geq 0$.

(ii) Es ist $v \in \mathrm{Kern}(A^tA)$ genau dann, wenn $A^tAv = \mathbf{0}$, aber $A^tAv = \mathbf{0}$ genau dann, wenn $\langle Av, Av\rangle = \langle v, A^tAv\rangle = \mathbf{0}$, es ist also $Av = \mathbf{0}$ also ist $v \in \mathrm{Kern}(A)$. Die Argumentation kann man umdrehen, also ist $\mathrm{Kern}(A^tA) = \mathrm{Kern}(A)$. □

Beweis (von Satz 27): Wir zeigen den zweiten Teil des Satzes und erhalten dann den ersten Teil durch Transponieren. Wir wollen verwenden, dass A^tA symmetrisch ist, d. h. es gibt $S \in O(n)$ mit $S^t(A^tA)S = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_i \geq 0$ für $1 \leq i \leq n$, vergleiche Lemma VIII.1.3. Durch Umsortieren der Spalten von S^t können wir erreichen, dass $\lambda_1, \dots, \lambda_p \geq 0$ und $\lambda_i = 0$ für $p \leq i \leq n$ mit einem $p \in \{1, \dots, n\}$. Nach Lemma VIII.1.3 ist $p = \mathrm{Rang}(A^tA) = \mathrm{Rang}(A)$.

Seien $e_1, \dots, e_n$ die Standardbasisvektoren von $\mathbb{R}^n$ und $v_i = Se_i \in \mathbb{R}^n$ die Spalten von S. Die v_i, $1 \leq i \leq n$ sind die Eigenvektoren von A^tA zum Eigenwert

λ_i und $\{v_1, \ldots, v_n\}$ bildet eine Orthonormalbasis von $\mathbb{R}^n$ bezüglich $\langle \cdot, \cdot \rangle$. Es ist also

$$\langle Av_i, Av_j \rangle_{\mathbb{R}^m} = e_i^t S^t (A^t A) S e_j = e_i^t \mathrm{diag}(\lambda_1, \ldots, \lambda_n) e_j = \delta_{i,j} \lambda_i.$$

Insbesondere ist $Av_i \neq \mathbf{0}$ für $1 \leq i \leq p$ und $Av_i = \mathbf{0}$ für $p \leq i \leq n$. Wir definieren $\mu_i := \lambda_i^{1/2}$ für $1 \leq i \leq n$ und $w_i = 1/\mu_i Av_i \in \mathbb{R}^m$ für $1 \leq i \leq p$. Dann gilt:

$$\langle w_i, w_j \rangle = \frac{1}{\mu_i \mu_j} \langle Av_i, Av_j \rangle = \delta_{i,j},$$

$\{w_1, \ldots, w_p\}$ ist also ein Orthonormalsystem im $\mathbb{R}^m$. Wir ergänzen dieses zu einer Orthonormalbasis $\{w_1, \ldots, w_p, w_{p+1}, \ldots, w_m\}$ des $\mathbb{R}^m$. Es gilt insbesondere: $Av_i = \mu_i w_i$ für $1 \leq i \leq n$. Setze $\tilde{U}_1 = (w_1 | \ldots | w_m) \in O(m)$ und $U_1 := \tilde{U}_1^{-1}$ und $U_2 := (v_1 | \ldots | v_n) \in O(n)$. Dann gilt für $1 \leq i \leq n$:

$$U_1 A U_2 e_i = U_1 A v_i = \mu_i U_1 w_i = \mu_i e_i,$$

d. h. $U_1 A U_2$ hat die behauptete Form.

Es bleibt zu zeigen, dass Σ eindeutig ist: Ist $U_1 A U_2$ von der behaupteten Gestalt, dann ist $\mathrm{diag}(\mu_1^2, \ldots, \mu_n^2) = U_2^t A^t U_1^t U_1 A U_2 = U_2^t (A^t A) U_2$, $\mu_1^2, \ldots, \mu_n^2$ sind also die Eigenwerte von $A^t A$ und damit eindeutig bestimmt. □

Bemerkung VIII.1.4: Die Singulärwertzerlegung geht analog über $\mathbb{C}$ mit Matrizen $U_1 \in U(m)$, $U_2 \in U(n)$ und $\mu_1, \ldots, \mu_n \in \mathbb{R}_{\geq 0}$.

Beispiel VIII.1.5: Seien

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad \Sigma^+ = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \\ 0 & 0 \end{pmatrix}.$$

Dann sind $\Sigma\Sigma^+ = I_2$ und $\Sigma^+\Sigma = \mathrm{diag}(1, 1, 0)$.

Proposition VIII.1.6 (Pseudoinverse): *Sei $A \in \mathbb{R}^{m \times n}$ mit Singulärwertzerlegung $A = U_1 \Sigma U_2$ wie in Satz 27. Sei weiter $A^+ := U_2^t \Sigma^+ U_1^t$ mit*

$$(\Sigma^+)_{i,j} := \begin{cases} \frac{1}{\mu_i}, & \text{falls } i = j \leq p, \\ 0, & \text{sonst.} \end{cases}$$

Dann gelten:

(i) $A^+ A = U_1 \left(\begin{smallmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{smallmatrix} \right) U_1^t \in \mathbb{R}^{m \times m}$,

(ii) $A A^+ = U_2^t \left(\begin{smallmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{smallmatrix} \right) U_2 \in \mathbb{R}^{n \times n}$,

(iii) $AA^+A = A$ *und* $A^+AA^+ = A^+$,

(iv) AA^+ *und* A^+A *sind symmetrisch.*

Eine Matrix A^+*, die die Eigenschaften (iii) und (iv) erfüllt, heißt Pseudoinverse.*

Beweis: (i) Es sind $AA^+ = U_1\Sigma U_2 U_2^t \Sigma^+ U_1^t = U_1 \Sigma\Sigma^+ U_1^t$ und wegen

$$\Sigma\Sigma^+ = \begin{pmatrix} I_p & \mathbf{0}_{p\times m-p} \\ \mathbf{0}_{m-p\times p} & \mathbf{0}_{m-p\times m-p} \end{pmatrix}$$

ist $U_1\Sigma\Sigma^+U_1^t = U_1(\begin{smallmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{smallmatrix})U_1^t$ wie gewünscht.

(ii) Analog.

(iii) Wir rechnen nach

$$AA^+A = U_1\Sigma U_2 U_2^t \Sigma^+ U_1^t U_1 \Sigma U_2 = U_1 \Sigma\Sigma^+\Sigma U_2 = U_1\Sigma U_2 = A,$$

analog zeigt man $A^+AA^+ = A^+$.

(iv) Es ist

$$(AA^t)^t = (U_1\Sigma\Sigma^+U_1^t)^t = U_1(\Sigma\Sigma^+)^t U_1^t = U_1(\Sigma\Sigma^+)U_1^t = AA^+,$$

analog für A^+A. □

Bemerkung VIII.1.7: A^+ wie in Proposition VIII.1.6 wird auch Moore-Pensroe-Pseudoinverse genannt. Erfüllt A^+ nur die Bedingung aus (iii) nennt man sie oft schon Pseudoinverse. Diese spielen eine große Rolle in der Numerik, da man damit spezielle Lösungen von Linearen Gleichungssystemen berechnet, siehe Übungsblatt 4, Aufgabe 3.

2. Normale Matrizen

Definition VIII.2.1 (Normale Matrizen): Sei $A \in \mathbb{C}^{n\times n}$. A heißt *normal*, falls $AA^* = A^*A$.

Satz 28 (Spektralsatz für normale Matrizen): *Für* $A \in \mathbb{C}^{n\times n}$ *gilt: A ist normal genau dann, wenn es* $S \in U(n)$ *und Zahlen* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ *gibt mit*

$$SAS^{-1} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Vor dem Beweis sammeln wir ein paar vorbereitende Aussagen.

Notation VIII.2.2: Seien K ein Körper, $A \in K^{n\times n}$ sowie $U \subseteq K^n$. Dann schreiben wir $AU := \{Ax \mid x \in U\}$ für das Bild von U unter A.

Proposition VIII.2.3 (Invariante Unterräume für kommutierende Matrizen):

(i) *Seien K ein Körper, $A, B \in K^{n\times n}$ mit $AB = BA$ und $\lambda \in \operatorname{Spec}(A)$. Dann gilt: $B \operatorname{Eig}(A,\lambda) \subseteq \operatorname{Eig}(A,\lambda)$, d. h. $\operatorname{Eig}(A,\lambda)$ ist B-invariant.*

(ii) *Seien nun $A, B \in \mathbb{C}^{n\times n}$ mit $AB = BA$ und $\lambda \in \operatorname{Spec}(A)$. Dann gilt: $B^* \operatorname{Eig}(A,\lambda)^\perp \subseteq \operatorname{Eig}(A,\lambda)^\perp$, d. h. $\operatorname{Eig}(A,\lambda)^\perp$ ist B^*-invariant.*

Beweis: (i) Sei $x \in \operatorname{Eig}(A,\lambda)$. Wir wollen sehen, dass $Bx \in \operatorname{Eig}(A,\lambda)$. Das können wir aber sofort nachrechnen: $ABx = BAx = B\lambda x = \lambda Bx$.

(ii) Sei $x \in \operatorname{Eig}(A,\lambda)^\perp$. Wir wollen sehen, dass für alle $y \in \operatorname{Eig}(A,\lambda)$ gilt: $\langle B^*x, y\rangle = 0$. Es ist

$$\langle B^*x, y\rangle = x^t(B^*)^t\overline{y} = x^t\overline{By} = \langle x, By\rangle = 0,$$

da $\operatorname{Eig}(A,\lambda)$ ja nach (i) B-invariant ist. Damit gilt $B^*x \in \operatorname{Eig}(A,\lambda)^\perp$. □

Proposition VIII.2.4 (Von A, A^* erzeugte $\mathbb{C}$-Algebra): *Sei $A \in \mathbb{C}^{n\times n}$ normal. Wir definieren*

$$\mathfrak{A} := \mathbb{C}[A, A^*] := \left\{ \sum_{i=0}^{n}\sum_{j=0}^{k} b_{i,j}A^i(A^*)^j : m, k \in \mathbb{N}_0, b_{i,j} \in \mathbb{C}\right\}.$$

(i) *$\mathfrak{A}$ ist eine $\mathbb{C}$-Unteralgebra von $\mathbb{C}^{n\times n}$,*

(ii) *$\mathfrak{A}$ ist eine kommutative Algebra,*

(iii) *$\mathfrak{A}$ ist unter Adjunktion abgeschlossen, d. h. ist $M \in \mathfrak{A}$, dann auch M^*.*

Beweis: (i) Wir sehen ein, dass $\mathfrak{A}$ ein $\mathbb{C}$-Vektorraum ist. Explizit wollen wir nachrechnen, dass $\mathfrak{A}$ unter Matrizenmultiplikation abgeschlossen ist. Seien $M_1 = \sum_{i=0}^{m}\sum_{j=0}^{k} b_{i,j}A^i(A^*)^j$ und $M_2 = \sum_{r=0}^{m'}\sum_{s=0}^{k'} b'_{r,s}A^r(A^*)^s$. Dann ist, da A normal ist,

$$M_1M_2 = \sum_{i=0}^{m}\sum_{j=0}^{k}\sum_{r=0}^{m'}\sum_{s=0}^{k'} b_{i,j}b'_{r,s}A^i(A^*)^jA^r(A^*)^s = \sum_{i=0}^{m}\sum_{j=0}^{k}\sum_{r=0}^{m'}\sum_{s=0}^{k'} b_{i,j}b'_{r,s}A^{i+r}(A^*)^{j+s},$$

also $M_1M_2 \in \mathfrak{A}$.

(ii) Der obigen Rechnung für M_1M_2 können wir $M_1M_2 = M_2M_1$ direkt ansehen, d. h. $\mathfrak{A}$ ist kommutativ.

(iii) Sei $M = \sum_{i=0}^{m}\sum_{j=0}^{k} b_{i,j}A^i(A^*)^j \in \mathfrak{A}$. Dann ist

$$M^* = \sum_{i=0}^{m}\sum_{j=0}^{k} \overline{b}_{i,j}A^j(A^*)^i \in \mathfrak{A}.$$

□

Satz 29 (Simultane Orthonormalbasen): *Ist $\mathfrak{A} \subseteq \mathbb{C}^{n\times n}$ eine $\mathbb{C}$-Unteralgebra, die kommutativ– und unter Adjunktion abgeschlossen ist, so gibt es eine Orthonormalbasis $B = \{b_1, \ldots, b_n\}$ von $\mathbb{C}^n$ von* simultanen Eigenvektoren*, d. h. für alle $A \in \mathfrak{A}$ sind $b_1, \ldots, b_n$ Eigenvektoren von A.*

Beweis: Wir zeigen die Behauptung durch vollständig Induktion nach n.

Für $n = 1$ ist die Aussage trival. Wir können $B = \{1\}$ wählen, dann hat jede Matrix in $\mathbb{C}^{1\times 1}$ Diagonalgestalt bezüglich dieser Basis.

Für den Induktionssschritt $n-1$ nach n müssen wir Fälle unterscheiden:

(i) Ist $\mathfrak{A} = \mathbb{C}I_n = \{\mathrm{diag}(c, \ldots, c) \mid c \in \mathbb{C}\}$ oder $\mathfrak{A} = \{0\}$, ist die Behauptung klar.

(ii) Ist $\mathfrak{A}$ eine andere Algebra als die Algebren in (i), dann können wir $A_0 \in \mathfrak{A}$ mit $A_0 \notin \mathbb{C}I_n$ wählen. Sei λ ein Eigenwert von A_0 (diesen gibt es, da $\mathbb{C}$ algebraisch abgeschlossen ist).

Für $V_\lambda := \mathrm{Eig}(A_0, \lambda)$ gilt: $0 < \dim V_\lambda < n$. Nach Proposition VIII.2.3 wissen wir für alle $A \in \mathfrak{A}$: $AV_\lambda \subseteq V_\lambda$ und $A^*V_\lambda^\perp \subseteq V_\lambda^\perp$. Nach Induktionsvoraussetzung gibt es eine simultane Orthonormalbasis B_1 von V_λ und eine simultane Orthonormalbasis B_2 von $V_\lambda^\perp$ unter $\mathfrak{A}_1 := \{A|_{V_\lambda} \mid A \in \mathfrak{A}\}$ und $\mathfrak{A}_2 := \{A|_{V_\lambda^\perp} \mid A \in \mathfrak{A}\}$.

$B_1 \cup B_2$ ist jetzt eine Orthonormalbasis von $\mathbb{C}^n = V_\lambda \oplus V_\lambda^\perp$, $B = B_1 \cup B_2$ leistet das Gewünschte. □

Beweis (für Satz 28): „⇒“: Die nötige Arbeit haben wir in Proposition VIII.2.4 und Satz 29 geleistet.

„⇐“: Es gelte $SAS^* = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ mit $S \in U(n)$. Dann ist

$$SAA^*S^* = SAS^*(SAS^*)^* = \mathrm{diag}(|\lambda_1|^2, \ldots, |\lambda_n|^2) = (SAS^*)^*SAS^* = SA^*AS^*,$$

also $AA^* = SS^*AA^*SS^* = SS^*A^*AS^*S = A^*A$. □

3. Die adjungierte Abbildung

Für zwei Vektoren $x, y \in \mathbb{C}^n$ und eine Matrix $A \in \mathbb{C}^{n\times n}$ haben wir bereits gesehen, dass

$$\langle Ax, y\rangle = (Ax)^t\overline{y} = x^tA^t\overline{y} = x^t\overline{A^*}\overline{y} = \langle x, A^*y\rangle.$$

Im Folgenden wollen wir uns überlegen, was „Sternen“ mit linearen Abbildungen tut.

In diesem Abschnitt seien $\mathbb{K}$ der Körper der reellen– oder komplexen Zahlen, V und W zwei $\mathbb{K}$-Vektorräume mit Skalarprodukten $\langle\cdot,\cdot\rangle_V$ beziehungsweise $\langle\cdot,\cdot\rangle_W$.

Lemma VIII.3.1: *Sind v_1, $v_2 \in V$, sodass für alle $v \in V$ gilt: $\langle v_1, v\rangle_V = \langle v_2, v\rangle_V$. Dann folgt $v_1 = v_2$.*

Beweis: Nach Voraussetzung ist $\langle v_1 - v_2, v_1\rangle_V = \langle v_1 - v_2, v_2\rangle_V$, also

$$\langle v_1 - v_2, v_1 - v_2\rangle_V = 0,$$

d. h. $v_1 = v_2$. □

Proposition VIII.3.2: *Sei $\Phi\colon V \to W$ eine lineare Abbildung. Für jedes $w \in W$ gibt es höchstens einen Vektor $\Phi^*(w) \in V$ mit der folgenden Eigenschaft: Für alle $v \in V$ gilt*

$$\langle\Phi(v), w\rangle_W = \langle v, \Phi^*(w)\rangle_W.$$

Beweis: Gäbe es ein zweites Element $\tilde{\Phi}^*(w)$ mit den beschriebenen Eigenschaften, dann gälte für alle $v \in V$:

$$\langle v, \tilde{\Phi}^*(w)\rangle_V = \langle\Phi(v), w\rangle_W = \langle v, \Phi^*(w)\rangle_V,$$

also $\Phi^*(w) = \tilde{\Phi}^*(w)$ nach Lemma VIII.3.1. □

Definition VIII.3.3: Falls in Proposition VIII.3.2 für jedes $w \in W$ ein solches $\Phi^*(w)$ existiert, dann heißt die Abbildung $\Phi^*\colon W \to V$, $w \mapsto \Phi^*(w)$ die *adjungierte Abbildung zu* Φ.

Proposition VIII.3.4 (Linearität der Adjungierten): *Falls die adjungierte Abbildung Φ^* existiert, so ist sie linear.*

Beweis: Für alle $w_1, w_2 \in W$ und $\alpha \in \mathbb{K}$ sowie für alle $v \in V$ gilt

$$\begin{aligned}\langle v_1, \Phi^*(\alpha w_1 + w_2)\rangle_V &= \langle \Phi(v), \alpha w_1 + w_2\rangle_W \\ &= \overline{\alpha}\langle \Phi(v), w_1\rangle_W + \langle \Phi(v), w_2\rangle_W \\ &= \overline{\alpha}\langle v, \Phi^*(w_1)\rangle_V + \langle v, \Phi^*(w)\rangle_V = \langle v, \alpha_1\Phi^*(w_1) + \Phi^*(w_2)\rangle,\end{aligned}$$

nach Lemma VIII.3.1 ist also $\Phi^*(\alpha w_1 + w_2) = \alpha\Phi^*(w_1) + \Phi^*(w_2)$. □

Ab jetzt schreiben wir kurz $\langle \cdot, \cdot\rangle$ statt $\langle \cdot, \cdot\rangle_V$ beziehungsweise $\langle \cdot, \cdot\rangle_W$.

Definition VIII.3.5: Sei nun $\Phi\colon V \to V$ eine lineare Abbildung.

(i) Φ heißt *selbstadjungiert*, falls $\Phi^* = \Phi$,

(ii) Φ heißt *normal*, falls $\Phi^* \circ \Phi = \Phi \circ \Phi^*$.

Φ heißt *orthogonal* beziehungsweise. *unitär*, falls für alle Vektoren $v_1, v_2 \in V$ gilt: $\langle \Phi(v), \Phi(w)\rangle = \langle v_1, v_2\rangle$. Orthgonale beziehungsweise unitäre Abbildungen sind (als Isometrien) injektiv; ist V endlichdimensional, so sind orthogonale beziehungsweise unitäre Abbildungen damit bijektiv.

Bemerkung VIII.3.6: Sei $\Phi\colon V \to V$ eine lineare Abbildung.

(i) Ist Φ selbstadjungiert, dann ist Φ normal.

(ii) Ist Φ orthogonal beziehungsweise unitär und V endlichdimensional, so ist $\Phi^* = \Phi^{-1}$, denn für alle $v, w \in V$ gilt

$$\langle \Phi(v), w\rangle = \langle \Phi(v), \Phi(\Phi^{-1}(w))\rangle = \langle v, \Phi^{-1}(w)\rangle.$$

Damit ist insbesondere Φ normal.

Im Folgenden wollen wir uns davon überzeugen, dass es Φ^* gibt, wenn V und W endlichdimensional sind und dass A^* und Φ^* zusammen passen. Für die Darstellungsmatrizen müssen wir dazu aber immer Orthonormalbasen fixieren.

Erinnerung VIII.3.7: Seien V ein n-dimensionaler– und W ein m-dimensionaler $\mathbb{K}$-Vektorraum mit geordneten Basen $B = (b_1, \ldots, b_n)$ bzw. $C = (c_1, \ldots, c_m)$. Die Koordinatenabbildung $D_B\colon V \to \mathbb{K}^n$ ist die Umkehrabbildung der Abbildung $\mathbb{K}^n \to V$, $(x_1, \ldots, x_n) \mapsto \sum_{i=1}^n x_i b_i$, entsprechend $D_C\colon W \to \mathbb{K}^m$. Für $\Phi\colon V \to W$ kommutiert das Diagramm, wenn $A = D_{C,B}(\Phi)$ und genau dadurch wird A definiert:

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} \\ \mathbb{K}^n & \xrightarrow[x\mapsto Ax]{} & \mathbb{K}^m \end{array}.$$

Es bezeichne $\langle \cdot, \cdot \rangle_S$ das Standardskalarprodukt auf $\mathbb{K}^n$.

Bemerkung VIII.3.8: Sei B eine Basis von V. Für alle $v, w \in V$ gilt genau dann $\langle v, w \rangle = \langle D_B(v), D_B(w) \rangle_S$, wenn B eine Orthonormalbasis ist.

Proposition VIII.3.9: *Es seien V und W zwei endlichdimensionale $\mathbb{K}$-Vektorräume und $\Phi\colon V \to W$ eine lineare Abbildung.*

(i) *Die Adjungierte $\Phi^*\colon W \to V$ existiert,*

(ii) *Sind $B = (b_1, \ldots, b_m)$ und $C = (c_1, \ldots, c_n)$ Orthonormalbasen von V beziehungsweise W, dann gilt:*

$$D_{B,C}(\Phi^*) = (D_{C,B}(\Phi))^*.$$

Beweis: Sei $A = D_{C,B}(\Phi)$. Wir betrachten die Situation

$$\begin{array}{ccccc} V & \xrightarrow{\Phi} & W & \xrightarrow{\Psi} & V \\ {\scriptstyle D_B}\downarrow & & {\scriptstyle D_C}\downarrow & & \downarrow{\scriptstyle D_B} \\ \mathbb{K}^m & \xrightarrow[x \mapsto Ax]{} & \mathbb{K}^n & \xrightarrow[x \mapsto A^*x]{} & \mathbb{K}^m \end{array},$$

wir setzen also $\Psi\colon W \to V$, $w \mapsto D_B^{-1}(A^* D_C(w))$. Für $v \in V$ und $w \in W$ ist dann (unter Verwendung von Bemerkung VIII.3.8)

$$\begin{aligned} \langle v, \Psi(w) \rangle &= \langle v, D_B^{-1}(A^* D_C(w)) \rangle \\ &= \langle D_B(v), A^* D_C(w) \rangle_S \\ &= \langle A D_B(v), D_C(w) \rangle_S = \langle D_C(\Phi(v)), D_C(w) \rangle = \langle \Phi(v), w \rangle, \end{aligned}$$

d. h. Ψ ist die Adjungierte von Φ. Nach Definition von Ψ kommutiert das folgende Diagramm:

$$\begin{array}{ccc} W & \xrightarrow{\Psi} & V \\ {\scriptstyle D_C}\downarrow & & \downarrow{\scriptstyle D_B} \\ \mathbb{K}^m & \xrightarrow[x \mapsto A^*x]{} & \mathbb{K}^n \end{array}$$

also genau $A^* = D_{B,C}(\Phi^*)$. □

Korollar VIII.3.10: *Sind V ein n-dimensionaler $\mathbb{K}$-Vektorraum, $\Phi \in \mathrm{End}(V)$ und B eine Orthonormalbasis von V, dann gelten:*

(i) *Φ is selbstadjungiert genau dann, wenn $D_{B,B}(\Phi)$ hermitesch ist,*

(ii) *Φ ist normal genau dann, wenn $D_{B,B}(\Phi)$ normal ist,*

(iii) *Φ ist orthogonal bzw. unitär genau dann, wenn $D_{B,B}(\Phi)$ orthogonal bzw. unitär ist.*

Kapitel IX.

Multilineare Algebra

1. Dualraum

In diesem Abschnitt seien stets K ein Körper und V ein K-Vektorraum.

Definition IX.1.1 (Linearformen und Dualraum):

(i) Eine *Linearform auf* V ist eine lineare Abbildung von V nach K.

(ii) Der Vektorraum $V^* := \mathrm{Hom}_K(V, K) = \mathrm{Hom}(V, K)$ heißt *Dualraum von* V.

Der Daulraum V^* ist also der K-Vektorraum der Linearformen auf V.

Erinnerung IX.1.2: Sind $f, g \in V^*$ und $\alpha \in K$, dann sind die Funktionen $f+g$ und αf für alle $v \in V$ definiert durch

$$(f+g)(v) := f(v) + g(v), \qquad (\alpha f)(v) := \alpha f(v).$$

Beispiel IX.1.3: Setze $V = P_2 = \{p \in \mathbb{R}[X] \mid \deg(p) \leq 2\}$. V hat zum Beispiel $B = \{1, X, X^2\}$ als Basis, insbesondere ist $\dim V = 3$, und es sind zum Beispiel $p_1 = 1 + 2X + 3X^2$ und $p_2 = X - X^2$ Elemente von V.

Der Dualraum von V ist $V^* = \{f\colon P_2 \to \mathbb{R} \mid f \text{ lineare Abbildung}\}$; zum Beispiel

$$f\colon P_2 \longrightarrow \mathbb{R}, \qquad p \mapsto \int_0^1 p(X)\, dX$$

ist ein Element von V^*. Es sind

$$f(p_1) = \int_0^1 1 + 2X + 3X^2\, dX = [X + X^2 + X^3]_0^1 = 3,$$

$$f(p_2) = \int_0^1 X - X^2\, dX = \left[\frac{1}{2}X^2 - \frac{1}{3}X^3\right]_0^1 = \frac{1}{6}.$$

Bemerkung IX.1.4: (i) Nach dem Fortsetzungssatz für lineare Abbildungen (siehe Korollar IV.4.1 aus Lineare Algebra I) gilt für eine gegebene Basis B von V, dass $V^* \cong \mathrm{Abb}(B, \mathbb{K})$.

(ii) Falls V endlichdimensional ist mit $\dim V = n$, so gilt $\dim V^* = n$ und $V \cong V^*$. Die Isomorphie $V \cong V^*$ muss für unendlichdimensionale Vektorräume nicht gelten.

Proposition IX.1.5: *Sei nun V endlichdimensional mit Basis $B = \{b_1, \ldots, b_n\}$. Für $i \in \{1, \ldots, n\}$ definieren wir*

$$b_i^* \colon V \longrightarrow K, \qquad b_j \longmapsto \delta_{i,j}.$$

Dann bildet $\{b_1^*, \ldots, b_n^*\}$ eine Basis von V^*.

Beweis: (1) Wir zeigen zunächst, dass $\{b_1^*, \ldots, b_n^*\}$ ein Erzeugendensystem für V^* ist. Sei $f \in V^*$. Definieren wir $\alpha_i := f(b_i)$, so gilt $f = \sum_{i=1}^n \alpha_i b_i^*$, denn sowohl f als auch $\sum_{i=1}^n \alpha_i b_i^*$ sind lineare Abbildungen, und sie stimmen auf der Basis B überein.

(2) Nun zeigen wir, dass $\{b_1^*, \ldots, b_n^*\}$ linear unabhängig ist. Seien dazu $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum_{i=1}^n \alpha_i b_i^* = \mathbf{0}$. Für $j \in \{1, \ldots, n\}$ gilt $0 = \sum_{i=1}^n \alpha_i b_i^*(b_j) = \alpha_j$, was wir zeigen wollten. □

Definition IX.1.6: Die Basis $B^* := \{b_1^*, \ldots, b_n^*\}$ aus Proposition IX.1.5 heißt *duale Basis von V zu B*.

Beispiel IX.1.7: In Beispiel IX.1.3 wähle $B = \{b_0 = 1, b_1 = X, b_2 = X^2\}$. Dann erhalten wir die duale Basis $B^* = \{b_0^*, b_1^*, b_2^*\}$ mit $b_0^*(a + bX + cX^2) = a$, $b_1^*(a + bX + cX^2) = b$ umd $b_2^*(a + bX + cX^2) = c$.

Für f wie in Beispiel IX.1.3 erhalten wir

$$f(1) = \int_0^1 1 \, dX = 1, \quad f(X) = \int_0^1 X \, dX = \frac{1}{2}, \quad f(X^2) = \int_0^1 X^2 \, dX = \frac{1}{3},$$

also $f = b_0^* + \frac{1}{2} b_1^* + \frac{1}{3} b_2^*$.

Korollar IX.1.8: *In der Situation von Proposition IX.1.5 gilt: Wir erhalten einen Isomorphismus $h = h_B \colon V \to V^*$ als lineare Fortsetzung von $b_i \mapsto b_i^*$.*

Proposition IX.1.9 (Duale Abbildung): *Seien V und W beliebige K-Vektorräume und $\Phi \colon V \to W$ eine lineare Abbildung. Dann erhalten wir folgende lineare Abbildung:*

$$\Phi^* \colon W^* \longrightarrow V^*, \qquad f \longmapsto f \circ \Phi.$$

Beweis: Nachrechnen der Linearität wird Übungsaufgabe auf dem fünften Übungsblatt.

Achtung: Φ^* hat eine doppelte Bedeutung als duale Abbildung und als adjungierte Abbildung. Hier ist die *duale Abbildung* gemeint. □

Definition IX.1.10: In der Situation von Proposition IX.1.9 heißt Φ^* die *duale Abbildung zu* Φ.

Bemerkung IX.1.11 (Koeffizienten von dualem Vektor): Seien V ein K-Vektorraum mit Basis $\{b_1, \ldots, b_n\}$ und $f \in V^*$. Wir suchen die Koeffizienten $\alpha_i \in K$, $1 \leq i \leq n$, in $f = \sum_{i=1}^n \alpha_i b_i^*$. Es gilt $f = \sum_{i=1}^n \alpha_i b_i^*$ genau dann, wenn $\alpha_i = f(b_i)$ für $1 \leq i \leq n$ (vergleiche Beweis von Proposition IX.1.5).

Beispiel IX.1.12 (Polynome): (i) Es sei $P_d := \{p(X) \in K[X] \mid \deg(p) \leq d\}$. Für die Basis $B_d = \{b_d = X^d, b_{d-1} = X^{d-1}, \ldots, b_0 = 1\}$ ist die duale Basis die Menge $B_d^* = \{b_d^*, \ldots, b_0^*\}$ mit

$$b_i^* \colon \mathbb{R}[X] \longrightarrow \mathbb{R}, \qquad \sum_{i=0}^{d} a_i X^i \longmapsto a_i.$$

Für $V := P_3$ und $W := P_2$ sei $\Phi \colon V \to W$, $p(X) \mapsto p'(X)$. Die duale Abbildung ist dann $\Phi^* \colon W^* \to V^*$, $f \mapsto f \circ \Phi$ mit $(f \circ \Phi)(p(X)) = f(p'(X))$. Ist zum Beispiel f die Abbildung definiert durch $q(X) \mapsto \int_0^1 q(X)\, dX$, dann ist $\Phi^*(f) \in V^*$ mit

$$p(X) = \sum_{i=0}^{3} a_i X^i \mapsto \int_0^1 p'(X)\, dX = p(1) - p(0) = a_3 + a_2 + a_1;$$

es ist also $\Phi^*(f) = b_3^* + b_2^* + b_1^*$ (hierbei ist $B = \{b_0, b_1, b_2, b_3\}$ die Basis von P_3 von oben).

(ii) Sei weiter $C = \{c_0 = 1, c_1 = X, c_2 = X^2\}$ die entsprechende Basis von P_2. Die Bilder von der b_i unter Φ sind

$$\Phi(b_0) = 0, \qquad \Phi(b_1) = 1 = c_0, \qquad \Phi(b_2) = 2x = 2c_1, \qquad \Phi(b_3) = 3X^2 = 3c_2,$$

die Darstellungsmatrix $A = D_{C,B}(\Phi) \in \mathbb{R}^{4\times 3}$ ist also

$$D_{C,B}(\Phi) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(iii) Es bezeichnen $B^* = \{b_0^*, b_1^*, b_2^*, b_3^*\}$ und $C^* = \{c_0^*, c_1^*, c_2^*\}$ die entsprechenden dualen Basen und $\Phi^*\colon W^* = P_2^* \to V^* = P_3^*$, $f \mapsto \Phi^*(f) = f \circ \Phi$ die duale Abbildung. Für ein Polynom $p = \sum_{i=0}^3 a_i X^i$ finden wir

$$p = \sum_{i=0}^{3} a_i X^i \overset{\Phi}{\longmapsto} 3a_3X^2 + 2a_2X + a_1$$
$$3a_3X^2 + 2a_2X + a_1 \overset{f}{\longmapsto} 3a_3f(X^2) + 2a_2f(X) + a_1,$$

d. h. $\Phi^*(f) = 3f(X^2)b_3^* + 2f(X)b_2^* + f(1)b_1^*$; insbesondere sind

$$\Phi^*(c_0^*) = b_1^*, \qquad \Phi^*(c_1^*) = 2b_2^*, \qquad \Phi^*(c_2^*) = 3b_3^*,$$

also ist

$$B := D_{B^*,C^*}(\Phi^*) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = A^t \in \mathbb{R}^{3\times 4}.$$

Beispiel IX.1.13 (Dualraum von K^n): Ist $V = K^n$, dann ist

$$V^* = \operatorname{Hom}(K^n, K) \cong K^{1\times n}.$$

Wir fixieren den Isomorphismus $K^{1\times n} \to \operatorname{Hom}(K^n, K)$, $A \mapsto (x \mapsto Ax)$ und identifizieren auf diese Weise $(K^n)^*$ mit $K^{1\times n}$.

Zur Standardbasis $E_n = \{e_1, \ldots, e_n\}$ von K^n ist die duale Basis von $(K^n)^*$ die Menge $E^* = \{e_1^t, \ldots, e_n^t\}$. Ist $\Phi\colon K^n \to K^m$, $x \mapsto Ax$ mit $A \in K^{m\times n}$, dann erhalten wir für die duale Abbildung:

$$\Phi^*\colon K^{1\times m} \cong (K^m)^* \to (K^n)^* \cong K^{1\times n}, \qquad (\alpha_1, \ldots, \alpha_m) \longmapsto (\alpha_1, \ldots, \alpha_m)A$$

Insbesondere gilt

$$\begin{array}{ccc} (\alpha_1, \ldots, \alpha_m) & K^{1\times m} \xrightarrow{\Phi^*} K^{1\times n} & (\alpha_1, \ldots, \alpha_n) \cdot A \\ \big\downarrow & \big\downarrow D_{E_m^*} \qquad\qquad \big\downarrow D_{E_n^*} & \big\downarrow \\ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} & K^m \xrightarrow[x\mapsto A^t x]{} K^n & A^t \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \end{array}$$

das heißt wir haben ausgerechnet, dass $D_{E_n^*,E_m^*}(\Phi^*) = A^t$.

Proposition IX.1.14 (Abbildungsmatrix für die duale Abbildung): *Seien V und W endlichdimensionale K-Vektorräume mit Basen $B = \{b_1, \ldots, b_m\}$ beziehungsweise $C = \{c_1, \ldots, c_m\}$ und $A = D_{C,B}(\Phi)$. Dann gilt für die duale Abbildung $\Phi^*\colon W^* \to V^*$, dass $D_{B^*,C_*}(\Phi^*) = A^t$.*

Beweis: Seien $f \in W^*$ und $\alpha_i = f(c_i)$, $1 \leq i \leq n$. Betrachte die beiden folgenden kommutativen Diagramme:

$$\begin{array}{ccccc} V & \xrightarrow{\Phi} & W & \xrightarrow{f} & K \\ \downarrow{\scriptstyle D_B} & & \downarrow{\scriptstyle D_C} & & \downarrow{\scriptstyle \mathrm{id}} \\ K^m & \xrightarrow[x\mapsto Ax]{} & K^n & \xrightarrow[y\mapsto(\alpha_1,\ldots,\alpha_n)y]{} & K \end{array}$$

also $(f \circ \Phi)(v) = (\alpha_1, \ldots, \alpha_n) \cdot A \cdot D_B(v)$ und

$$\begin{array}{c} f \\ \downarrow \\ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \end{array} \qquad \begin{array}{ccc} W^* & \xrightarrow{\Phi^*} & V^* \\ \downarrow{\scriptstyle D_{C^*}} & & \downarrow{\scriptstyle D_{B^*}} \\ K^n & \longrightarrow & K^m \end{array} \qquad \begin{array}{c} f \circ \Phi \\ \downarrow \\ \begin{pmatrix} (f\circ\Phi)(b_1) \\ \vdots \\ (f\circ\Phi)(b_m) \end{pmatrix} \end{array}$$

Das erste kommutative Diagramm gibt uns

$$((f \circ \Phi)(b_1)| \ldots |(f \circ \Phi)(b_m)) = (\alpha_1, \ldots, \alpha_n) \cdot A \cdot I,$$

d. h.

$$\begin{pmatrix} (f \circ \Phi)(b_1) \\ \vdots \\ (f \circ \Phi)(b_m) \end{pmatrix} = A^t \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},$$

was wir zeigen wollten. □

Definition IX.1.15 (Bidualraum): Es heißt $V^{**} := (V^*)^* = \mathrm{Hom}(V^*, K)$ der *Bidualraum von* V.

Beispiel IX.1.16 (Einsetzungshomomorphismus): Zu $v \in V$ erhalten wir einen Einsetzungshomomorphismus

$$\Lambda(v)\colon V^* \longrightarrow K, \qquad f \longmapsto f(v).$$

Die Abbildung ist linear (nachrechnen). Mit diesen Einsetzungshomomorphismen können wir eine lineare Abbildung

$$\Lambda\colon V \longrightarrow V^{**}, \qquad v \longmapsto \Lambda(v)$$

erklären.

Proposition IX.1.17 (V als Unterraum von V^{**}): *Sei V endlichdimensional. Der Homomorphismus $\Lambda\colon V \to V^{**}$, $v \mapsto \Lambda(v)$ aus Beispiel IX.1.16 ist injektiv und damit bijektiv, da* $\dim V^{**} = \dim V^* = \dim V$.

Beweis: Sei $w \in V$ mit $w \neq \mathbf{0}$. Wir wollen zeigen, dass dann auch $\Lambda(w) \neq \mathbf{0}_{V^*}$. Wir schreiben $V = \langle w \rangle \oplus U$ (das ist möglich wegen des Basisergänzungssatzes). Jedes Element $v \in V$ lässt sich eindeutig schreiben als $v = \lambda w + u$ mit $\lambda \in K$ und $u \in U$.

Die Abbildung $\Pi_w \colon V \to K$, $v = \lambda w + u \mapsto \lambda$ ist linear, d. h. $\Pi_w \in V^*$. Außerdem ist $\Lambda(w)(\Pi_w) = \Pi_w(w) = 1$, d. h. $\Lambda(w) \neq \mathbf{0}_{V^*}$. □

Bemerkung IX.1.18: Die Abbildung Λ ist auch für unendlichdimensionale Vektorräume injektiv. Dies folgt aus der Existenz von Basen und der Gültigkeit des Basisergänzungssatzes für unendlichdimensionale Vektorräume, womit wir uns am Ende der Vorlesung genauer beschäftigen werden. Im Allgemeinen ist Λ jedoch *nicht* bijektiv.

2. Multilineare Abbildungen

In diesem Abschnitt sei stets K ein Körper.

Definition IX.2.1 (Multilineare Abbildung): Seien $V_1, \ldots, V_n$ und W Vektorräume über K. Eine Abbildung

$$M \colon V_1 \times V_2 \times \cdots \times V_n \longrightarrow W$$

heißt *n-fach multilineare Abbildung*, wenn für jedes $i \in \{1, \ldots, n\}$ und jede Wahl von Vektoren $v_j \in V_j$ (mit $1 \leq j \leq n$, $j \neq i$) die Abbildung

$$V_i \longrightarrow W, \qquad v \longmapsto M(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

eine lineare Abbildung ist.

Einfach multilineare Abbildungen heißen lineare Abbildungen, zweifach multilineare Abbildungen heißen bilineare Abbildungen und ist speziell $W = K$, so sprechen wir von *Multilinearformen*.

Beispiel IX.2.2: (i) Die Determinante

$$\det \colon K^n \times \cdots \times K^n \longrightarrow K, \qquad (v_1, \ldots, v_n) \longmapsto \det(v_1 | \ldots | v_n)$$

ist eine n-fache Multilinearform.

(ii) Die skalare Multiplikation $K \times V \to V$, $(\lambda, v) \mapsto \lambda v$ ist eine bilineare Abbildung.

(iii) Für zwei K-Vektorräume V und W ist die Abbildung

$$\mathrm{Hom}(V,W) \times V \longrightarrow W, \qquad (\Phi, v) \longmapsto \Phi(v)$$

ist eine bilineare Abbildung.

(iv) Für natürliche Zahlen p,q,r,s ist die Matrizenmultiplikation

$$K^{p\times q} \times K^{q\times r} \times K^{r\times s} \longrightarrow K^{p\times s}, \qquad (A,B,C) \longmapsto A \cdot B \cdot C$$

eine dreifach multilineare Abbildung.

(v) Für einen beliebigen Ring R, z.B. $R = K[X]$, ist die Multiplikation $R \times R \to R$, $(f,g) \mapsto f \cdot g$ bilinear.

Proposition IX.2.3 (Multilinearformen sind durch Basen bestimmt): *In der Situation von Definition IX.2.1 seien $B_1, \ldots, B_n$ Basen von $V_1, \ldots, V_n$. Die Menge aller multilinearen Abbildungen von $V_1 \times \cdots \times V_n$ nach W bilden einen Vektorraum $\mathfrak{m} = \mathfrak{m}(V_1 \times \cdots \times V_n, W)$, der isomorph ist zu* $\mathrm{Abb}(B_1 \times \cdots \times B_n, W)$ *ist. Sind alle beteiligten Vektorräume endlichdimensional, dann auch $\mathfrak{m}$ mit*

$$\dim \mathfrak{m} = \dim V_1 \cdots \dim V_n \cdot \dim W.$$

Beweis: Die multilinearen Abbildungen M von $V_1 \times \cdots \times V_n$ nach W sind eindeutig durch die Bilder $M(b_1, \ldots, b_n)$ mit $b_1 \in B_1, \ldots, b_n \in B_n$ bestimmt. Genauer: Wir erhalten eine Abbildung

$$\Phi\colon \mathfrak{m} \longrightarrow \mathrm{Abb}(B_1 \times \cdots \times B_n, W), \qquad M \longmapsto ((b_1, \ldots, b_m) \mapsto M(b_1, \ldots, b_m)).$$

Φ ist linear. Weiterhin ist Φ injektiv; ist nämlich $\Phi(M_1) = \Phi(M_2)$ mit Multilinearformen $M_1, M_2 \in \mathfrak{m}$, und schreiben wir jeden Vektor im Tupel $(v_1, \ldots, v_n)$ mit $v_1 \in V_1, \ldots, v_n \in V_n$ als $v_i = \sum_{b_j \in B_j} \lambda^i_{b_j} b_j$, dann ist

$$\begin{aligned}
M_1(v_1, \ldots, v_n) &= M_1\Big(\sum_{b_j \in B_1} \lambda^1_{b_j} b_j, v_2, \ldots, v_n\Big) \\
&= \sum_{b_j \in B_1} \lambda^1_{b_j} M_1(b_j, v_2, \ldots, v_n) \\
&= \sum_{b_{j_1} \in B_1} \cdots \sum_{b_{j_n} \in B_n} \lambda^1_{b_{j_1}} \cdots \lambda^n_{b_{j_n}} M_1(b_{j_1}, \ldots, b_{j_n}) \\
&= \sum_{b_{j_1} \in B_1} \cdots \sum_{b_{j_n} \in B_n} \lambda^1_{b_{j_1}} \cdots \lambda^n_{b_{j_n}} M_2(b_{j_1}, \ldots, b_{j_n}) \\
&= M_2(v_1, \ldots, v_n).
\end{aligned}$$

Schließlich ist Φ surjektiv: Für ein $f \in \mathrm{Abb}(B_1 \times \cdots \times B_n, W)$ gibt die Zuordnung

$$\Big(\sum_{b_1 \in B_1} \lambda_{b_1}(b_1) b_1, \ldots, \sum_{b_n \in B_n} \lambda_{b_n}(b_n) b_n \Big) \longmapsto \prod_{i=1}^{n} \lambda_{b_i}(b_i) \cdot f(b_1, \ldots, b_n)$$

eine multilineare Abbildung.

Sind $V_1, \ldots, V_n$ endlichdimensional, so enthält $M := B_1 \times \cdots \times B_n$ genau $\dim V_1 \cdots \dim V_n$-viele Elemente, also haben wir die Behauptung über die Dimension auch gezeigt. □

Notation IX.2.4: Im Folgenden schreiben wir $[n] := \{1, \ldots, n\}$ und

$$\mathrm{pr}_i \colon K^n \longrightarrow K, \qquad (x_1, \ldots, x_n)^t \longmapsto x_i$$

für die sogenannte *i-te Projektionsabbildung*.

Beispiel IX.2.5 (n-fache Multilinearformen auf K^n): Es ist $\mathfrak{m}_K(\prod_{i=1}^n K^n, K)$ der Vektorraum der n-fachen Multilinearformen auf K^n. Zum Beispiel ist $\det \in \mathfrak{m}_K(\prod_{i=1}^n K^n, K)$; nach Proposition IX.2.3 ist

$$\dim \mathfrak{m}_K\Big(\prod_{i=1}^{n} K^n, K \Big) = n^n = \#\,\mathrm{Abb}([n],[n]).$$

Für jedes $k \in \mathrm{Abb}([n],[n])$ definieren wir die Multilinearform

$$M_k \colon K^n \times \cdots \times K^n \longrightarrow K, \qquad (v_1, \ldots, v_n) \longmapsto \prod_{i=1}^{n} \mathrm{pr}_{k(i)}(v_i).$$

Diese sind linear unabhängig, und da die Anzahl der Elemente gleich der Dimension von $\mathfrak{m}_K(\prod_{i=1}^n K^n, K)$ ist, ist $\{M_k \mid k \in \mathrm{Abb}([k],[k])\}$ eine Basis von $\mathfrak{m}_K(\prod_{i=1}^n K^n, K)$.

Die Leibnizformel für die Determinante in unserer Notation ist

$$\det(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} \mathrm{pr}_{\sigma(i)}(v_i),$$

d. h. $\det = \sum_{k \in \mathrm{Abb}([n],[n])} c_k M_k$ mit $c_k = \mathrm{sign}(\sigma)$, falls $k = \sigma$ bijektiv ist, und 0 sonst.

Definition IX.2.6 (Eigenschaften multilinearen Abbildungen): Seien V und W zwei K-Vektorräume. Eine multilineare Abbildung $M \colon \prod_{i=1}^n V \to W$ heißt

(i) *symmetrisch*, falls für alle $\sigma \in S_n$ und $v_1, \ldots, v_n \in V$ gilt

$$M(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = M(v_1, \ldots, v_n).$$

(ii) *schiefsymmetrisch*, falls für alle $\sigma \in S_n$ und $v_1, \ldots, v_n \in V$ gilt

$$M(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \operatorname{sign}(\sigma) M(v_1, \ldots, v_n).$$

(iii) *alternierend*, falls für alle $v_1, \ldots, v_n \in V$ mit $v_i = v_j$ für $i \neq j$ gilt $M(v_1, \ldots, v_n) = 0$.

Beispiel IX.2.7: (i) Reelle Skalarprodukte $\beta\colon V \times V \to \mathbb{R}$ sind bilinear und symmetrisch.

(ii) Die Determinante $\det\colon \prod_{i=1}^n K^n \to K$ ist multilinear, schiefsymmetrisch und alternierend.

(iii) Das Kreuzprodukt $\times\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$\left[\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \right] \longmapsto \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

ist bilinear, schiefsymmetrisch und alternierend.

(iv) Es bezeichnet $\mathbb{F}_2 := \mathbb{Z}/2\mathbb{Z}$ den Körper mit zwei Elementen. Die Abbildung

$$s\colon \mathbb{F}_2 \times \mathbb{F}_2 \longrightarrow \mathbb{F}_2, \qquad (a, b) \longmapsto ab$$

ist bilinear, symmetrisch und schiefsymmetrisch (da $-1 \equiv 1 \pmod 2$), allerdings nicht alternierend.

Proposition IX.2.8 (Schiefsymmetrisch vs alternierend): *Seien V und W zwei K-Vektorräume und $M\colon \prod_{i=1}^n V \to W$ eine multilineare Abbildung.*

(i) *Ist M alternierend, so ist M schiefsymmetrisch.*

(ii) *Ist $\operatorname{char}(K) \neq 2$, so gilt auch die Umkehrung.*

Beweis: Es seien $v_1, \ldots, v_n \in V$ und $1 \leq i < j \leq n$.

(i) Es sei M alternierend. Dann ist

$$\begin{aligned} 0 &= M(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\ &= M(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) \\ &\quad + M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) \\ &\quad + M(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\ &\quad + M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n), \end{aligned}$$

d. h.

$$-M(v_1,\dots,v_n) = M(v_1,\dots,v_{i-1},v_j,v_{i+1},\dots,v_{j-1},v_i,v_{j+1},\dots,v_n),$$

die Aussage gilt also für Transpositionen. Da Transpositionen die S_n erzeugen und sign ein Gruppenhomomorphismus ist, gilt die Aussage für alle Permutationen.

(ii) Sei M schiefsymmetrisch. Dann ist

$$\begin{aligned} M(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_{j-1},v_i,v_{j+1},\dots,v_n) \\ = -M(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_{j-1},v_i,v_{j+1},\dots,v_n) \end{aligned}$$

denn die Vertauschung des i-ten und j-ten Arguments bringt den Vorfaktor -1 mit sich. Wir haben also

$$0 = 2M(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_{j-1},v_i,v_{j+1},\dots,v_n),$$

können wir durch 2 teilen, dann ist M auch alternierend. □

3. Tensorprodukt

In diesem Abschnitt seien K ein Körper, sowie V_1, V_2 und W Vektorräume über K.

Bemerkung IX.3.1 („Erster Auftritt von τ“): Es seien $V_1 = K^n$, $V_2 = K^m$ und $T = K^{n\times m}$ sowie

$$\tau\colon V_1 \times V_2 \longrightarrow T, \qquad (v_1, v_2) \longmapsto v_1 v_2^t.$$

Dann gelten:

(i) τ ist bilinear,

(ii) Sind $\{e_1,\dots,e_n\}$ und $\{e'_1,\dots,e'_m\}$ die Standardbasen von K^n beziehungsweise K^m, dann gilt $\tau(e_i, e'_j) = E_{i,j}$.[1]

Beispiel IX.3.2: Für $V_1 = \mathbb{R}^3$ und $V_2 = \mathbb{R}^2$ ist $T = \mathbb{R}^{3\times 2}$ und zum Beispiel sind

$$\tau\left(\begin{pmatrix}3\\1\\2\end{pmatrix},\begin{pmatrix}2\\5\end{pmatrix}\right) = \begin{pmatrix}6&15\\2&5\\4&10\end{pmatrix}, \qquad \tau\left(\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}0&0\\1&0\\0&0\end{pmatrix}.$$

[1] Hier meint $E_{i,j}$ die Elementarmatrix die als (i,j)-ten Eintrag 1 und sonst nur 0-Einträge hat.

Proposition IX.3.3 (Universelle Eigenschaft von τ): *Die Abbildung $\tau\colon V_1 \times V_2 \to T$ hat folgende Eigenschaft: Ist W ein K-Vektorraum und $\beta\colon V_1 \times V_2 \to W$ eine bilineare Abbildung, dann gibt es genau ein $\Phi\colon T \to W$, sodass gilt:*

(i) Φ *ist linear,*

(ii) $\Phi \circ \tau = \beta$.

Beweis: Es bezeichnen wieder $\{e_1, \dots, e_n\}$ und $\{e'_1, \dots, e'_m\}$ die Standardbasen von K^n beziehungsweise K^m. Wir definieren $\Phi\colon T \to W$ als die eindeutige lineare Abbildung mit $\Phi(E_{i,j}) = \beta(e_i, e'_j)$. Nach Konstruktion gilt dann $\Phi \circ \tau = \beta$, denn für alle $v_1 = \sum_{i=1}^n \alpha_i e_i$ und $v_2 = \sum_{j=1}^m \beta_j e'_j$ ist

$$\begin{aligned}
\Phi(\tau(v_1, v_2)) &= \Phi\Big(\Big(\sum_{i=1}^n \alpha_i e_i\Big)\Big(\sum_{j=1}^m \beta_j e'_j\Big)^t\Big) \\
&= \Phi\Big(\sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j E_{i,j}\Big) \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j \Phi(E_{i,j}) \\
&= \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \beta(e_i, e'_j) = \beta\Big(\sum_{i=1}^n \alpha_i e_i, \sum_{j=1}^m \beta_j e'_j\Big) = \beta(v_1, v_2).
\end{aligned}$$

Die Abbildung Φ leistet also das Gewünschte. Da gelten muss

$$\Phi(E_{i,j}) = \Phi(\tau(e_i, e'_j) = \beta(e_i, e'_j),$$

muss Φ so gewählt werden und ist damit eindeutig bestimmt. □

Beispiel IX.3.4: Es seien $V_1 = \mathbb{R}^3 = V_2$ und $T = \mathbb{R}^{3\times 3}$. Weiter sei τ wie in Bemerkung IX.3.1. Wählen wir

$$\beta\colon \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}, \qquad (v_1, v_2) \longmapsto \langle v_1, v_2\rangle = v_1^t v_2,$$

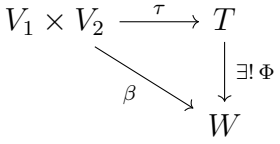
so erhalten wir wegen der universellen Eigenschaft das Diagramm

$$\begin{array}{ccc}
\mathbb{R}^3 \times \mathbb{R}^3 & \xrightarrow{\ \tau\ } & \mathbb{R}^{3\times 3} \\
& \searrow^{\beta} & \downarrow{\scriptstyle \Phi} \\
& & \mathbb{R}
\end{array}$$

Beachte, dass $\Phi(E_{i,j}) = \beta(e_i, e_j) = \delta_{i,j}$; für eine Matrix $A = (a_{i,j}) \in \mathbb{R}^{3\times 3}$ erhalten wir also

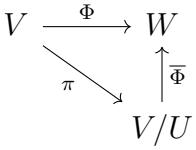
$$\Phi(A) = \Phi\Big(\sum_{i,j=1}^3 a_{i,j} E_{i,j}\Big) = \sum_{i=1}^n a_{i,i} = \mathrm{Spur}(A).$$

Definition IX.3.5 (Tensorprodukt): Ist T ein K-Vektorraum und $\tau\colon V_1 \times V_2 \to T$ eine bilineare Abbildung mit der Eigenschaft „Für jede bilineare Abbildung $\beta\colon V_1 \times V_2 \to W$ in einen K-Vektorraum W gibt es genau eine lineare Abbildung $\Phi\colon T \to W$ mit $\Phi \circ \tau = \beta$“, dann heißt (T, τ) *Tensorprodukt von V_1 und V_2 über K*.



Notation IX.3.6: In der Situation von Definition IX.3.5 schreiben wir für den Vektorraum im Tensorprodukt $V_1 \otimes_K V_2 := T$ und für $v_1 \in V_1$, $v_2 \in V_2$ schrieben wir $v_1 \otimes v_2 := \tau(v_1, v_2)$.

Erinnerung IX.3.7: Ist V ein K-Vektorraum und U ein Untervektorraum von V, dann gilt für den Quotientenraum V/U: Für jede lineare Abbildung $\Phi\colon V \to W$ mit $\mathrm{Kern}(\Phi) \supseteq U$ gibt es genau eine lineare Abbildung $\overline{\Phi}\colon V/U \to W$ mit $\overline{\Phi} \circ \pi = \Phi$, wobei wir mit $\pi\colon V \to V/U$ die kanonische Projektion bezeichnen. Wir sind also in der Situation



Erinnerung IX.3.8: Es sei M eine Menge.

(i) Für $f \in \mathrm{Abb}(M, K)$ heißt $\mathrm{Tr}(f) := \{m \in M \mid f(m) \neq 0\}$ der Träger von f,

(ii) Wir definieren $\mathrm{Abb}_0(M, K) := \{f\colon M \to K \mid \#\,\mathrm{Tr}(f) < \infty\}$; $\mathrm{Abb}_0(M, K)$ ist ein Untervektorraum von $\mathrm{Abb}(M, K)$.

(iii) Für $m \in M$ definieren wir die Abbildung

$$f_m\colon M \longrightarrow K, \qquad f_m(x) := \begin{cases} 1, & m = x, \\ 0, & \text{sonst.} \end{cases}$$

Die Menge $B = \{f_m \mid m \in M\}$ bildet eine Basis von $\mathrm{Abb}_0(M, K)$.

Beispiel IX.3.9: Ist $M = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, so erhalten wir eine Abbildung

$$\tau_1\colon \mathbb{R} \times \mathbb{R} \longrightarrow \mathrm{Abb}_0(\mathbb{R} \times \mathbb{R}, \mathbb{R}), \qquad m = (x, y) \longmapsto f_m = f_{(x,y)}.$$

Beachte, dass τ_1 *keine* bilineare Abbildung ist (zum Beispiel ist $\tau_1(0, 0)$ nicht der Nullvektor).

Satz 30: *Es seien V_1 und V_2 beliebige K-Vektorräume. Dann existiert ein Tensorprodukt von V_1 und V_2 über K.*

Beweis: Schritt 1 (Ein zu großer Kandidat für das Tensorprodukt): Es seien $F := \mathrm{Abb}_0(V_1 \times V_2, K)$ und

$$\tau_1 \colon V_1 \times V_2 \longrightarrow F, \qquad (v_1, v_2) \longmapsto f_{(v_1,v_2)}$$

mit $f_{(v_1,v_2)}$ wie in Erinnerung IX.3.8.

Gut an dieser Wahl ist, dass die Menge $\{f_{(v_1,v_2)} \mid (v_1, v_2) \in V_1 \times V_2\}$ eine Basis von F bildet, d. h. für jede Abbildung $\beta \colon V_1 \times V_2 \to W$ gibt es eine eindeutige lineare Abbildung mit $\Phi_\beta \circ \tau_1 = \beta$, nämlich die lineare Fortsetzung von $\Phi_\beta(f_{(v_1,v_2)}) := \beta(v_1, v_2)$. Schlecht an unserer Wahl ist, dass τ_1 keine bilineare Abbildung ist.

Schritt 2 (Verbesserung des Kandidaten): Wir wünschen uns, dass τ bilinear ist, d. h. dass für alle $v_1, v_1' \in V_1,\ v_2, v_2' \in V_2$ und $\alpha_1, \alpha_2 \in K$ gilt:

$$\tau(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') = \alpha_1\alpha_2\tau(v_1, v_2) + \alpha_1\tau(v_1, v_2') + \alpha_2\tau(v_1', v_2) + \tau(v_1', v_2').$$

Sei R der Untervektorraum von F der erzeugt wird von

$$\{f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')} - \alpha_1\alpha_2 f_{(v_1,v_2)} - \alpha_1 f(v_1, v_2') - \alpha_2 f_{(v_1',v_2)} - f_{(v_1',v_2')} \in F : \\ v_1, v_1' \in V_1, v_2, v_2' \in V_2, \alpha_1, \alpha_2 \in K\}.$$

Jetzt setzen wir $T := F/R$, nennen wieder $\pi \colon F \to T = F/R$ die kanonische Projektion und definieren $\tau := \pi \circ \tau_1$, d. h.

$$\tau \colon V_1 \times V_2 \longrightarrow T = F/R, \qquad (v_1, v_2) \longmapsto \pi(\tau_1(v_1, v_2)) = [f_{(v_1,v_2)}] = f_{(v_1,v_2)} + R.$$

Schritt 3 (Zeige, dass τ bilinear ist): Seien $v_1, v_1' \in V_1,\ v_2, v_2' \in V_2$ und $\alpha_1, \alpha_2 \in K$. Dann ist

$$\begin{aligned}\tau(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') &= [f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')}] \\ &= [\alpha_1\alpha_2 f_{(v_1,v_2)} + \alpha_1 f_{(v_1,v_2')} + \alpha_2 f_{(v_1',v_2)} + f_{(v_1',v_2')}] \\ &= \alpha_1\alpha_2\tau(v_1, v_2) + \alpha_1\tau(v_1, v_2') + \alpha_2\tau(v_1', v_2) + \tau(v_1', v_2'),\end{aligned}$$

d. h. τ ist tatsächlich bilinear.

Schritt 4 (Zeige, dass (T, τ) ein Tensorprodukt ist): Sei $\beta \colon V_1 \times V_2 \to W$ eine bilineare Abbildung, wir sind also in der Situation

$$\begin{array}{ccccc} V_1 \times V_2 & \xrightarrow{\tau_1} & F := \mathrm{Abb}_0(V_1 \times V_2, K) & \xrightarrow{\pi} & T := F/R \\ & \searrow^{\beta} & \downarrow{\scriptstyle \Phi_\beta} & \swarrow_{\overline{\Phi}_\beta} & \\ & & W & & \end{array}$$

Wie in Schritt 1 definieren wir Φ_β durch $\Phi_\beta(f_{(v_1,v_2)}) = \beta(v_1, v_2)$. Damit gilt bereits $\Phi_\beta \circ \tau_1 = \beta$. Weiter ist

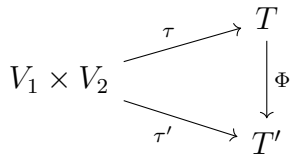
$$\begin{aligned}\Phi_\beta(f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')}) &= \beta(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') \\ &= \alpha_1\alpha_2\beta(v_1, v_2) + \alpha_1\beta(v_1, v_2') + \alpha_2\beta(v_1', v_2) + \beta(v_1', v_2') \\ &= \alpha_1\alpha_2\Phi_\beta(f_{(v_1,v_2)}) + \alpha_1\Phi_\beta(f_{(v_1,v_2')}) \\ &\quad + \alpha_2\Phi_\beta(f_{(v_1',v_2)}) + \Phi_\beta(f_{(v_1',v_2')}) \\ &= \Phi_\beta(\alpha_1\alpha_2 f_{(v_1,v_2)} + \alpha_1 f_{(v_1,v_2')} + \alpha_2 f_{(v_1',v_2)} + f_{(v_1',v_2')}),\end{aligned}$$

d. h. die Erzeuger von R liegen im Kern von Φ_β, also $R \subseteq \operatorname{Kern}(\Phi_\beta)$. Nach dem Homomorphiesatz gibt es demnach genau eine lineare Abbildung $\overline{\Phi}_\beta \colon T \to W$ mit $\overline{\Phi}_\beta \circ \pi = \Phi_\beta$ und schließlich gilt

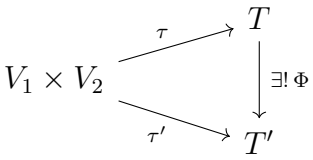
$$\overline{\Phi}_\beta \circ \tau = \overline{\Phi}_\beta \circ \pi \circ \tau_1 = \Phi_\beta \circ \tau_1 = \beta. \tag{IX.1}$$

$\overline{\Phi}_\beta$ muss auf diese Art und Weise konstruiert werden, damit Gl. (IX.1) gelten kann, also ist $\overline{\Phi}_\beta$ eindeutig und (T, τ) ist ein Tensorprodukt. □

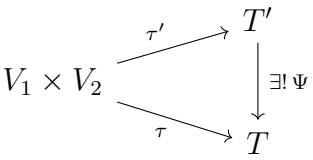
Proposition IX.3.10 (Eindeutigkeit des Tensorprodukts): *Das Tensorprodukt ist bis auf Isomorphie eindeutig, genauer gilt: Sind (T, τ) und (T', τ') Tensorprodukte von V_1 und V_2 über K, dann gibt es einen Isomorphismus $\Phi \colon T \to T'$ mit $\Phi \circ \tau = \tau'$. Wir sind also in der Situation*



Beweis: Da (T, τ) und (T', τ') Tensorprodukte sind, haben wir die kommutativen Diagramme

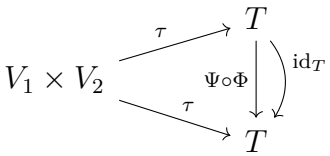


$\Phi \circ \tau = \tau'$



$\Psi \circ \tau' = \tau$

Wir haben also $(\Psi \circ \Phi) \circ \tau = \Psi \circ \tau' = \tau$, d. h. wegen der universellen Abbildungseigenschaft des Tensorproduktes haben wir das kommutative Diagramm



und wegen der Eindeutigkeit der linearen Abbildungen Φ und Ψ folgern wir $\mathrm{id}_T = \Psi \circ \Phi$, analog gehen wir für $\Phi \circ \Psi$ vor und erhalten $\Phi \circ \Psi = \mathrm{id}_T$. □

Proposition IX.3.11 (Konkretisierung mit Basen): *Es seien V_1 und V_2 zwei K-Vektorräume mit Basen $\{b_1, \ldots, b_n\}$ respektive $\{c_1, \ldots, c_m\}$. Dann ist die Menge*

$$D := \{b_i \otimes c_j \mid 1 \leq i \leq n, 1 \leq j \leq m\}$$

eine Basis von $T = V_1 \otimes_K V_2$.

Beweis: Sind wieder $F := \mathrm{Abb}_0(V_1 \times V_2, K)$ und

$$\tau_1 \colon V_1 \times V_2 \longrightarrow F, \qquad (v_1, v_2) \longmapsto f_{(v_1,v_2)} \colon (x, y) \mapsto \begin{cases} 1, & \text{falls } (x, y) = (v_1, v_2), \\ 0, & \text{sonst,} \end{cases}$$

dann ist $\{f_{(v_1,v_2)} = \tau_1((v_1, v_2)) \mid v_1 \in V_1, v_2 \in V_2\}$ eine Basis von F, d. h. die Menge $\{\tau((v_1, v_2)) = v_1 \otimes v_2 \mid v_1 \in V_1, v_2 \in V_2\}$ erzeugt T.

Sind $v_1 = \sum_{i=1}^n r_i b_i$ und $v_2 = \sum_{j=1}^m s_j c_j$, dann ist

$$v_1 \otimes v_2 = \tau\Big(\sum_{i=1}^n r_i b_i, \sum_{j=1}^m s_j c_j\Big) = \sum_{i=1}^n \sum_{j=1}^m r_i s_j \tau(b_i, c_j) = \sum_{i=1}^n \sum_{j=1}^m r_i s_j b_i \otimes c_j,$$

also wird T schon von D erzeugt.

Für die lineare Unabhängigkeit von D sei $\sum_{i=1}^n \sum_{j=1}^m r_{i,j} b_i \otimes c_j = 0$ mit Koeffizienten $r_{i,j} \in K$. Wegen der universellen Abbildungseigenschaft gilt für jede Bilinearform $\beta \colon V_1 \times V_2 \to K$ dann schon

$$\sum_{i=1}^n \sum_{j=1}^m r_{i,j} \beta(b_i, c_j) = 0,$$

da es genau ein lineares $\Phi \colon V_1 \otimes_K V_2 \to K$ mit $\beta = \tau \circ \Phi$ gibt, insbesondere also

$$\Phi\Big(\sum_{i=1}^n \sum_{j=1}^m r_{i,j} \tau(b_i, c_j)\Big) = 0 = \sum_{i=1}^n \sum_{j=1}^m r_{i,j} \beta(b_i, c_j).$$

Die Abbildung

$$\beta_{i,j} \colon V_1 \times V_2 \longrightarrow K, \qquad (b_k, c_\ell) \longmapsto \begin{cases} 1, & \text{falls } (k, \ell) = (i, j), \\ 0, & \text{sonst,} \end{cases}$$

ist bilinear, d. h. es gibt genau ein lineares $\Phi \colon V_1 \otimes_K V_2 \to K$ mit $\beta_{i,j} = \Phi \circ \tau$ und

$$\sum_{k=1}^n \sum_{\ell=1}^m r_{k,\ell} \beta_{i,j}(b_k, c_\ell) = r_{i,j} = 0 = \Phi\Big(\sum_{i=1}^n \sum_{j=1}^m r_{i,j} b_i \otimes c_j\Big),$$

d. h. $r_{i,j} = 0$ für $1 \leq i \leq n$, $1 \leq j \leq m$, was wir zeigen wollten. □

Korollar IX.3.12: *Seien V_1, V_2 zwei endlichdimensionale K-Vektorräume. Dann gilt*

$$\dim(V_1 \otimes_K V_2) = \dim(V_1) \cdot \dim(V_2).$$

Bemerkung IX.3.13: Sind $\Phi \in \mathrm{End}(V_1)$, $\Psi \in \mathrm{End}(V_2)$ und ist $h\colon V_1 \times V_2 \to W$ bilinear, dann ist

$$h \circ \Phi \times \Psi\colon V_1 \times V_2 \longrightarrow W, \qquad (v_1, v_2) \longmapsto h(\Phi(v_1), \Psi(v_2))$$

bilinear. Angewendet auf $W = V_1 \otimes_K V_2$ erhalten wir eine lineare Abbildung $\Psi \otimes \Phi\colon V_1 \otimes V_2 \to V_1 \otimes V_2$:

$$\begin{array}{ccc} V_1 \times V_2 & \xrightarrow{\ \tau\ } & V_1 \otimes_K V_2 \\ {\scriptstyle \Phi\times\Psi}\downarrow & \searrow^{\tau\circ(\Phi,\Psi)} & \downarrow{\scriptstyle \exists!\, J} \\ V_1 \times V_2 & \xrightarrow[\ \tau\ ]{} & V_1 \otimes_K V_2 \end{array}$$

wobei J leistet, dass $J \circ \tau = \tau \circ \Phi \times \Psi$; wir schreiben $\Phi \otimes \Psi := J$. Als Übungsaufgabe sei dem Leser überlassen zu zeigen, dass

$$(\Phi \otimes \Psi)(v_1 \otimes v_2) = \Phi(v_1) \otimes \Psi(v_2).$$

Sind wieder $\{b_1, \ldots, b_n\}$ eine Basis von V_1, $\{c_1, \ldots, c_m\}$ eine Basis von V_2,

$$D = (b_1 \otimes c_1, b_1 \otimes c_2, \cdots b_1 \otimes c_m, b_2 \otimes c_1, \ldots, b_n \otimes c_m),$$

und bezeichnen $(\beta_{k,i})_{k,i} = D_{B,B}(\Phi)$ und $(\gamma_{\ell,j})_{\ell,j} = D_{C,C}(\Psi)$ die Darstellungsmatrizen von Φ respektive Ψ bezüglich der Basen B respektive C, so gilt

$$J(b_i \otimes c_j) = \Phi(b_i) \otimes \Psi(c_j) = \Big(\sum_{k=1}^{n} \beta_{k,i} b_k \otimes \sum_{\ell=1}^{m} \gamma_{\ell,j} c_\ell\Big) = \sum_{k=1}^{n}\sum_{\ell=1}^{m} \beta_{k,i}\gamma_{\ell,j}(b_k \otimes c_\ell),$$

d. h.

$$D_{D,D}(J) = \begin{pmatrix} \beta_{1,1} D_{C,C}(\Psi) & \cdots & \beta_{1,n} D_{C,C}(\Psi) \\ \vdots & \ddots & \vdots \\ \beta_{n,1} D_{C,C}(\Psi) & \cdots & \beta_{n,n} D_{C,C}(\Psi) \end{pmatrix}.$$

Die Matrix $D_{D,D}(J)$ heißt *Kroneckerprodukt* von $D_{B,B}(\Phi)$ und $D_{C,C}(\Psi)$.

4. Tensorprodukte von Algebren und Moduln

In diesem Abschnitt seien stets K ein Körper, R ein Ring mit 1, A eine K-Algebra und V ein K-Vektorraum.

Im Folgenden wollen wir das Konzept des Vektorraums verallgemeinern und Multiplikation mit Skalaren aus Ringen statt Körpern zulassen, z. B. $\mathbb{Z} \cdot \mathbb{Q}$, $\mathbb{Z} \cdot \mathbb{Z}[\sqrt{2}]$, $K^{n\times n} \cdot K^n$, $K[X] \cdot K^{n\times n}$.

Definition IX.4.1: Ein *(Links-)Modul* über R ist eine abelsche Gruppe M zusammen mit einer (äußeren) Verknüpfung $\cdot\colon R \times M \to M$, sodass für alle $m, n \in M$ und $r, s \in R$ gelten:

(i) $1 \cdot m = m$,

(ii) $(r + s) \cdot m = r \cdot m + s \cdot m$, $r \cdot (m + n) = r \cdot m + r \cdot n$m

(iii) $(r \cdot s) \cdot m = r \cdot (s \cdot m)$.

Bemerkung IX.4.2: Für einen R-Modul M, $r \in R$ und das neutrale Element $0_M \in M$ gilt $r \cdot 0_M = 0_M$.

Beweis: Es ist wegen (ii) aus Definition IX.4.1

$$r \cdot 0_M = r \cdot (0_M + 0_M) = r \cdot 0_M + r \cdot 0_M,$$

d. h. $r \cdot 0_M = 0_M$. □

Beispiel IX.4.3: (i) Ist R ein Körper, dann ist M ein R-Modul genau dann, wenn M ein R-Vektorraum ist.

(ii) Ist $n \in \mathbb{N}$, dann ist $\prod_{i=1}^{n} R = R^n$ ein R-Modul mit komponentenweiser Addition und Skalarmultiplikation.

(iii) $M = \mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ ist ein $\mathbb{Z}$-Modul mit der Skalarmultiplikation

$$\cdot\colon \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z}, \qquad n \cdot \overline{0} := 0, \quad n \cdot \overline{1} := \begin{cases} \overline{0}, & \text{falls } 2 \mid n, \\ \overline{1}, & \text{sonst.} \end{cases}$$

(iv) Ist $(M, +)$ eine beliebige abelsche Gruppe, so ist M ein $\mathbb{Z}$-Modul mit der Skalarmultiplikation

$$\cdot\colon \mathbb{Z} \times M \longrightarrow M, \qquad k \cdot m \longmapsto \begin{cases} \sum_{i=1}^{k} m, & \text{falls } k \in \mathbb{N}_0, \\ \sum_{i=1}^{-k} (-m), & \text{falls } -k \in \mathbb{N}. \end{cases}$$

Definition IX.4.4: Eine Abbildung $\varphi\colon M \to M'$ zwischen R-Moduln heißt *R-Modulhomomorphismus* oder *R-linear*, falls $\varphi\colon (M,+) \to (M',+)$ ein Gruppenhomomorphismus ist und für alle $x \in M$ und $r \in R$ gilt $\varphi(rx) = r\varphi(x)$.

Bemerkung IX.4.5: Seien M, M' zwei R-Moduln und es bezeichne

$$\mathrm{Hom}_R(M,M') := \{\varphi\colon M \to M' \mid \varphi \text{ ist } R\text{-linear}\}.$$

Dann ist $\mathrm{Hom}_R(M,M')$ selbst ein R-Modul mit den (punktweisen) Verknüpfungen

$$(\varphi_1+\varphi_2)(x) := \varphi_1(x)+\varphi_2(x), \qquad (r\varphi_1)(x) := r\varphi_1(x)$$

für $\varphi_1, \varphi_2 \in \mathrm{Hom}_R(M,M')$, $r \in R$ und $x \in M$.

Definition IX.4.6: Es sei M ein R-Modul. Eine Teilmenge $U \subseteq M$ ist ein Untermodul von M, falls gelten:

(i) $(U,+)$ ist eine Untergruppe von $(M,+)$,

(ii) Für alle $r \in R$ und $u \in U$ ist $ru \in U$.

Das heißt: U ist R-Modul in eigenem Recht. Wir schreiben $U \leq M$.

Beispiel IX.4.7: (i) Es seien $M = \mathbb{Z}$, $R = \mathbb{Z}$, $k \in \mathbb{N}$ und $U := k\mathbb{Z}$. Dann ist U ein Untermodul von M.

(ii) Seien M und M' zwei R-Moduln und $\varphi \in \mathrm{Hom}_R(M,M')$. Dann sind die Mengen

$$\ker\varphi := \{m \in M \mid \varphi(m) = 0\} \subseteq M, \qquad \mathrm{Bild}\,\varphi := \{\varphi(m) \mid m \in M\} \subseteq M'$$

Untermoduln der jeweiligen Moduln.

Beweis: (i) Davon haben wir uns bereits in der Linearen Algebra I überzeugt.

(ii) Nach der Linearen Algebra I sind $\ker\varphi \subseteq M$ und $\mathrm{Bild}\,\varphi \subseteq M'$ Untergruppen. Sind $m \in \ker\varphi$ und $r \in R$, so gilt $\varphi(rm) = r\varphi(m) = 0$, d. h. $rm \in \ker\varphi$, und sind $r \in R$ und $m \in \mathrm{Bild}\,\varphi$, dann gilt $r\varphi(m) = \varphi(rm) \in \mathrm{Bild}\,\varphi$. □

Proposition IX.4.8: *Seien M ein R-Modul, $U \subseteq M$ ein Untermodul und es bezeichne*

$$M/U := \{m+U \mid m \in M\} = \{[m] \mid m \in M\}$$

mit $[m] = m+U := \{m+u \mid u \in M\}$.

(i) *M/U ist selbst ein R-Modul mit den Verknüpfungen*

$$(m_1 + U) + (m_2 + U) := (m_1 + m_2) + U, \qquad r \cdot (m + U) := (r \cdot m + U)$$

für m_1, m_2, $m \in M$, $r \in R$ und heißt Quotientenmodul.

(ii) *Die Abbildung $\pi\colon M \to M/U$, $m \mapsto m + U$ ist R-linear.*

(iii) *Der Quotientenmodul M/U erfüllt die folgende universelle Abbildungseigenschaft: Für jeden Modulhomomorphismus $\phi\colon M \to N$ mit $\ker\phi \subseteq U$ gibt es genau einen Modulhomomorphismus $\overline{\Phi}\colon M/U \to N$ mit $\overline{\Phi} \circ \pi = \phi$, d. h. das folgende Diagramm ist kommutativ:*

$$\begin{array}{ccc} M & \xrightarrow{\pi} & M/U \\ & \searrow^{\phi} & \downarrow{\scriptstyle \exists!\,\overline{\Phi}} \\ & & N \end{array}$$

Beweis: Für (i) weisen wir die Wohldefiniertheit nach: Für Elemente $u_1, u_2 \in U$, $m_1, m_2 \in M$ und $r \in R$ gelten

$$\begin{aligned} m_1 + u_1 + m_2 + u_2 = (m_1 + m_2) + (u_1 + u_2) &\in (m_1 + m_2) + U, \\ r \cdot (m_1 + u_1) = r \cdot m_1 + r \cdot u_1 &\in r \cdot m + U. \end{aligned}$$

Der Rest des Beweises funktioniert völlig analog zu Erinnerung IX.3.7. □

Es sei erwähnt, dass wir für die Wohldefiniertheit von der Kommutativität von U und M Gebrauch gemacht haben; für Gruppen müssen wir uns auf spezielle Untergruppen (sogenannte *Normalteiler*) einschränken, um den Quotienten eine Gruppenstruktur geben zu können.

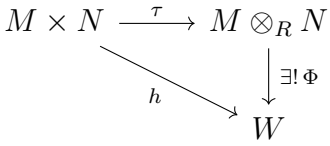
Beispiel IX.4.9: Es seien $M = \mathbb{Z}$, $U = 10\mathbb{Z}$ und $N = \mathbb{Z}/5\mathbb{Z}$. Ferner sei

$$\phi\colon \mathbb{Z} \longrightarrow \mathbb{Z}/5\mathbb{Z}, \qquad z \longmapsto \overline{z}.$$

Es ist $\ker\phi = 5\mathbb{Z} \supseteq 10\mathbb{Z}$, d. h. es gibt einen eindeutigen R-Modulhomomorphismus $\overline{\Phi}\colon \mathbb{Z}/10\mathbb{Z} \to \mathbb{Z}/5\mathbb{Z}$ mit $\overline{\Phi} \circ \pi = \phi$:

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\pi} & \mathbb{Z}/10\mathbb{Z} \\ & \searrow^{\phi} & \downarrow{\scriptstyle \exists!\,\overline{\Phi}} \\ & & \mathbb{Z}/5\mathbb{Z}. \end{array}$$

Proposition IX.4.10: *Die Begriffe „bilinear“ (vergleiche Definition IX.2.1) und „Tensorprodukt“ (vergleiche Definition IX.3.5) definieren wir für R-Moduln genau wie für K-Vektorräume. Sind M und N zwei R-Moduln, dann gibt es ein Tensorprodukt $(T = M \otimes_R N, \tau\colon M \times N \to M \otimes_R N)$, d. h. für jede bilineare Abbildung $h\colon M \times N \to W$ gibt es eine lineare Abbildung $\Phi\colon M \otimes_R N \to W$ mit $\Phi \circ \tau = h$; das folgende Diagramm ist also kommutativ:*



Beweis: Der Beweis für Satz 30 geht auch in dieser Situation durch. □

Beispiel IX.4.11: (i) Sei $h\colon \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \to W$ eine bilineare Abbildung. Dann ist

$$h(\overline{a}, \overline{b}) = h(\overline{1} \cdot \overline{a}, \overline{b}) = h(\overline{3} \cdot \overline{a}, \overline{b}) = h(\overline{a}, 3 \cdot \overline{b}) = h(\overline{a}, \overline{0}) = 0,$$

d. h. $h \equiv 0$. Damit ist $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z} = \{0\}$.

(ii) Sei M ein R-Modul. Dann ist $R \otimes_R M \cong M$. Genauer: M zusammen mit der Abbildung

$$\tau\colon R \times M \longrightarrow M, \qquad (r, m) \longmapsto rm$$

ist ein Tensorprodukt. Um das einzusehen halten wir zunächst fest, dass τ bilinear ist, und ist $h\colon M \times R \to W$ bilinear, dann tut $\Phi_h\colon M \to W$, $m \mapsto h(1, m)$ das Gewünschte.

Bemerkung IX.4.12: Sind M, N zwei R-Moduln, dann gilt: $M \otimes_R N$ wird erzeugt von $\{m \otimes n \mid m \in M, n \in N\}$ (wobei $m \otimes n := \tau(m, n)$), d. h. für $x \in M \otimes_R N$ gibt es $a_1, \ldots, a_k \in R$, $m_1, \ldots, m_k \in M$ und $n_1, \ldots, n_k \in N$, sodass

$$x = \sum_{i=1}^{k} a_i (m_i \otimes n_i).$$

Beweis: Im Beweis von Satz 30 haben wir das Tensorprodukt konstruiert als $M \otimes_R N := \mathrm{Abb}_0(M \times N, R)/U$ mit dem geeigneten Untermodul U und

$$\tau\colon M \times N \longrightarrow M \otimes_R N, \qquad (m, n) \longmapsto [f_{(m,n)}]$$

wobei $f_{(m,n)}(x, y) = \delta_{m,x}\delta_{n,y}$. Die $f_{(m,n)}$ erzeugen $\mathrm{Abb}_0(M \times N, R)$, also erzeugen die Restklassen $[f_{(m,n)}] = m \otimes n$ den Quotientenmodul. □

Definition IX.4.13: Eine *R-Algebra* A ist eine Menge mit Verknüpfungen

$$+\colon A \times A \longrightarrow A, \qquad \circ\colon A \times A \longrightarrow A, \qquad \cdot\colon R \times A \longrightarrow A$$

sodass gelten:

(i) $(A, +, \circ)$ ist ein Ring,

(ii) $(A, +, \cdot)$ ist ein R-Modul,

(iii) $\circ$ ist R-bilinear.

Manchmal nennt man so eine R-Algebra auch eine *assoziative R-Algebra*.

Proposition IX.4.14: *Seien $(A, +, \cdot, \circ)$ und $(B, +, \cdot, \circ)$ zwei R-Algebren und $A \otimes_R B$ das Tensorprodukt von beiden aufgefasst als R-Moduln. Dann wird $A \otimes_R B$ zur R-Algebra mit der Multiplikation*

$$\circ\colon A\otimes_R B \times A\otimes_R B \longrightarrow A\otimes_R B, \quad \Big(\sum_{i=1}^{m} a_i\otimes b_i, \sum_{j=1}^{n} a'_j\otimes b'_j\Big) \longmapsto \sum_{i=1}^{m}\sum_{j=1}^{m} a_i a'_j \otimes b_i b'_j$$

Beweis: (1) **„$\circ$“ ist wohldefiniert:** $A \otimes_R B$ ist Quotient von $\mathrm{Abb}_0(A \times B, R)$ und die zugehörige Äquivalenzrelation gibt die Gleichheit

$$(ra_1 + a_2) \otimes (sb_1 + b_2) = rs(a_1 \otimes b_1) + r(a_1 \otimes b_2) + s(a_2 \otimes b_1) + (a_2 \otimes b_2).$$

Seien $a \in A$ und $b \in B$. Dann ist

$$\begin{aligned}\Big((ra_1 + a_2) \otimes (sb_1 + b_2)\Big) \circ (a \otimes b) &= (ra_1 + a_2)a \otimes (sb_1 + b_2)b \\ &= ra_1a \otimes bsb_1 + aa_2 \otimes bsb_1 + ara_1 \otimes bb_2 + aa_2 \otimes bb_2 \\ &= (rsa_1 \otimes b_1 + ra_1 \otimes b_2 + sa_2 \otimes b_1 + a_2 \otimes b_2) \circ (a, b)\end{aligned}$$

(2) **„$\circ$“ ist assoziativ und es gilt Distributivität:** Übung. □

Definition IX.4.15 (R-Algebrenhomomorphismus): Seien A_1 und A_2 zwei R-Algebren. Eine Abbildung $\varphi\colon A_1 \to A_2$ heißt *R-Algebrenhomomorphismus*, falls für alle $a, b, \in A_1$ und $r \in R$ gelten

$$\varphi(a + b) = \varphi(a) + \varphi(b), \qquad \varphi(ab) = \varphi(a)\varphi(b), \qquad \varphi(ra) = r\varphi(a).$$

Sind A_1 und A_2 jeweils R-Algebren mit Eins, dann heißt φ ein *Homomorphismus von R-Algebren mit Eins*, falls zusätzlich gilt: $\varphi(1_{A_1}) = 1_{A_2}$.

5. Freie Moduln

Im ganzen Abschnitt sei R ein kommutativer Ring mit 1. Wir wollen in diesem Abschnitt der Frage nachgehen, welche Moduln Basen haben.

Definition IX.5.1: Sei M ein R-Modul.

(i) Für $X \subseteq M$ ist

$$\langle X \rangle := \bigcap \{U \mid U \subseteq M \text{ Untermodul mit } X \subseteq U\}$$

ein R-Modul und heißt das *Erzeugnis von* X, genauer: $\langle X \rangle$ ist der kleinste Untermodul von M, der X enthält.

(ii) Für $X \subseteq M$ ist $\langle X \rangle = \{\sum_{i=1}^{n} r_i x_i \mid n \in \mathbb{N}, r_i \in R, x_i \in X\}$.

(iii) Eine Teilmenge $B \subseteq M$ heißt *linear unabhängig*, falls für alle $b_i \in B$, $r_i \in R$ gilt:

$$\sum_{i=1}^{n} r_i b_i = 0 \Longrightarrow r_1 = \cdots = r_n = 0.$$

(iv) Eine Teilmenge $B \subseteq M$ heißt *Basis*, falls sich jedes $m \in M$ auf eindeutige Weise als Linearkombination $m = \sum_{i=1}^{n} r_i b_i$ mit $r_i \in R$ und $b_i \in B$ schreiben. Das ist genau dann der Fall, wenn B linear unabhängig ist und $\langle B \rangle = M$.

Sämtliche Aussagen zeigt man wie für Vektorräume.

Beispiel IX.5.2: Seien $n \in \mathbb{N}$ und $M = \mathbb{Z}/n\mathbb{Z}$ aufgefasst als $\mathbb{Z}$-Modul. Die Teilmenge $\{\overline{1}\} \subseteq \mathbb{Z}/n\mathbb{Z}$ ist linear abhängig, denn $n \cdot \overline{1} = \overline{0}$, aber $n \neq 0_{\mathbb{Z}}$. Insbesondere hat M keine Basis.

Definition IX.5.3: Ein R-Modul M heißt *frei*, falls M eine Basis besitzt.

Beispiel IX.5.4: (i) R^n ist ein freier R-Modul mit Basis $\{e_1, \ldots, e_n\}$ mit $e_i = (\delta_{i,j})_{1 \leq j \leq n}$.

(ii) Sei $\{0\} \subsetneq I \subsetneq R$ ein Ideal in R. Dann ist R/I nicht frei (als R-Modul).

(iii) Sei S eine beliebige Menge. Dann ist $\mathrm{Abb}_0(S, R)$ ein freier R-Modul.

Beweis: (i) Folgt aus (iii), da $R^n = \mathrm{Abb}(S, R)$ mit $S = \{1, \ldots, n\}$.

(ii) Seien $a \in R$ und $s \in I - \{0\}$. Dann ist $s[a] = [sa] = [0]$, da $sa \in I$. Damit ist $\{[a]\}$ nicht linear unabhängig, es kann also keine Basis geben.

(iii) Wähle $B := \{f_s \mid s \in S\}$ mit $f_s(x) = \delta_{s,x}$. Dann ist B ein Erzeugendensystem, denn für $f \in \mathrm{Abb}_0(S, R)$ mit $S' = \mathrm{Tr}(f) = \{s \in S \mid f(s) \neq 0\}$ ist

$$f = \sum_{s \in S'} f(s) f_s.$$

B ist ferner linear unabhängig, sind nämlich $r_1, \ldots, r_n \in R$ und $s_1, \ldots, s_n \in S$ mit

$$\sum_{i=1}^{n} r_i f_{s_i} = 0,$$

dann gilt für $j \in \{1, \ldots, n\}$: $0 = \sum_{i=1}^{n} r_i f_{s_i}(s_j) = r_j$, also $r_1 = \cdots = r_n = 0$. □

Proposition IX.5.5: *Ist M ein freier R-Modul mit Basis B, dann gilt: Für jede Abbildung $f\colon B \to M'$ in einen R-Modul M' gibt es genau einen R-Modulhomomorphismus $\Phi\colon M \to M'$ mit $\Phi|_B = f$.*

Beweis: Man zeigt dieses Aussage ganz genau wie für Vektorräume. □

Bemerkung IX.5.6 („Nicht-Beispiel“ für Basen): Sei M ein freier R-Modul mit Basis $B = \{b_1, \ldots, b_n\}$. Eine Teilmenge $\{c_1, \ldots, c_n\}$ von M mit n Elementen die linear unabhängig sind, muss keine Basis sein.

Ist zum Beispiel M der $\mathbb{Z}$-Modul $\mathbb{Z}$, dann ist M frei mit Basis $\{1\}$. Die Menge $\{2\}$ ist auch linear unabhängig in M, aber $\langle\{2\}\rangle = 2\mathbb{Z} \subsetneq \mathbb{Z}$.

Bemerkung IX.5.7: Wir brauchen Beispiel IX.5.4 und Proposition IX.5.5 im Beweis von Proposition IX.4.10.

Proposition IX.5.8 (Assoziatitvität des Tensorprodukts): *Es seien M_1, M_2 und M_3 drei R-Moduln.*

(i) *Für jede multilineare Abbildung $h\colon M_1 \times M_2 \times M_3 \to N$ gibt es genau eine lineare Abbildung $\Phi\colon M_1 \otimes_R (M_2 \otimes_R M_3) \to N$, sodass für alle $m_1 \in M_1$, $m_2 \in M_2$, $m_3 \in M_3$ gilt:*

$$\Phi(m_1 \otimes (m_2 \otimes m_3)) = h(m_1, m_2, m_3).$$

(ii) *Für jede multilineare Abbildung $h\colon M_1 \times M_2 \times M_3 \to N$ gibt es genau eine lineare Abbildung $\Psi\colon (M_1 \otimes_R M_2) \otimes_R M_3 \to N$, sodass für alle $m_1 \in M_1$, $m_2 \in M_2$, $m_3 \in M_3$ gilt:*

$$\Psi((m_1 \otimes m_2) \otimes m_3) = h(m_1, m_2, m_3).$$

(iii) *Es gibt einen eindeutigen Isomorphismus* $\Phi\colon M_1 \otimes_R (M_2 \otimes_R M_3) \to (M_1 \otimes_R M_2) \otimes_R M_3$ *mit*

$$\Phi(m_1 \otimes (m_2 \otimes m_3)) = \Phi((m_1 \otimes m_2) \otimes m_3)$$

für alle $m_1 \in M_1$, $m_2 \in M_2$, $m_3 \in M_3$.

Bemerkung IX.5.9: Seien X, Y, Z drei R-Moduln und $S = \{y_i \mid i \in I\}$ ein Erzeugendensystem für Y. Sei $h\colon X \times Y \to Z$ eine Abbildung mit folgenden Eigenschaften:

(i) h ist linear in der zweiten Komponente,

(ii) Für alle $y \in S$ gilt $h_y\colon X \to Z$, $x \mapsto h(x, y)$ ist linear.

Dann ist h auch linear in der ersten Komponente, also bilinear.

Beweis: Ist $y \in Y$ beliebig, so gibt es $y_1, \ldots, y_k \in S$ und $r_1, \ldots, r_k \in R$ mit $y = \sum_{j=1}^{k} r_j y_j$. Für alle x, $x' \in X$ und $r \in R$ gilt dann

$$\begin{aligned} h_y(x + rx', y) &= h_y\Big(x + rx', \sum_{j=1}^{k} r_j y_j\Big) \\ &= \sum_{j=1}^{k} r_j h(x + rx', y_j) \\ &= \sum_{j=1}^{k} r_j h(x, y_j) + r_j r h(x', y_j) = h_y(x, y) + r h_y(x', y). \end{aligned}$$ □

Beweis (von Proposition IX.5.8): Wir schreiben im Folgenden kurz „$\otimes$“ statt „$\otimes_R$“.

(i) Als Fahrplan für den Beweis haben wir das folgende Diagramm:

$$\begin{array}{ccccc} M_1 \times M_2 \times M_3 & \longrightarrow & M_1 \times M_2 \otimes M_3 & \longrightarrow & M_1 \otimes (M_2 \otimes M_3) \\ & \searrow^{h} & \downarrow{\scriptstyle \beta \text{ bilinear}} & \swarrow_{\Phi} & \\ & \text{multilinear} & N & \text{linear} & \end{array}$$

Wegen der universellen Abbildungseigenschaft von $M_2 \otimes M_3$ erhalten wir für jedes $m_1 \in M_1$ eine lineare Abbildung $\Phi_{m_1}\colon M_2 \otimes M_3 \to N$ mit

$$\Phi_{m_1}(m_2 \otimes m_3) = h(m_1, m_2, m_3).$$

Damit können wir $\beta\colon M_1 \times (M_2 \otimes M_3) \to N$ durch $(m_1, x) \mapsto \Phi_{m_1}(x)$ definieren. β ist linear in der zweiten Komponente, da Φ_{m_1} eine lineare Abbildung ist. β ist außerdem linear in der ersten Komponente nach Bemerkung IX.5.9, denn für jeden Tensor $m_2 \otimes m_3$ und für alle m_1, $m_1' \in M_1$ und $r \in R$ gilt

$$\begin{aligned}\beta(m_1 + rm_1', m_2 \otimes m_3) &= \Phi_{m_1+rm_1'}(m_2 \otimes m_3)\\ &= h(m_1 + rm_1', m_2, m_3)\\ &= h(m_1, m_2, m_3) + rh(m_1', m_2, m_3)\\ &= \Phi_{m_1}(m_2 \otimes m_3) + r\Phi_{m_2}(m_2 \otimes m_3)\\ &= \beta(m_1, m_2 \otimes m_3) + r\beta(m_1', m_2 \otimes m_3).\end{aligned}$$

Wegen der universellen Abbildungseigenschaft von $M_1 \otimes (M_2 \otimes M_3)$ erhalten wir eine lineare Abbildung $\Phi\colon M_1 \otimes (M_2 \otimes M_3) \to N$ mit $\Phi(m_1 \otimes x) = \beta(m_1, x)$. Insbesondere gilt für alle $m_1 \in M_1$, $m_2 \in M_2$ und $m_3 \in M_3$

$$\Phi(m_1 \otimes (m_2 \otimes m_3)) = \beta(m_1, m_2 \otimes m_3) = h(m_1, m_2, m_3). \tag{IX.2}$$

Da die $m_1 \otimes (m_2 \otimes m_3)$ den R-Modul $M_1 \otimes (M_2 \otimes M_3)$ erzeugen, ist Φ durch Gl. (IX.2) bereits eindeutig bestimmt.

(ii) Zeigt man analog zu (i).

(iii) Wir betrachten das Diagramm

$$\begin{array}{ccc} M_1 \times M_2 \times M_3 & \xrightarrow{h_1} & M_1 \otimes (M_2 \otimes M_3) \\ & \searrow^{h_2} & \downarrow \Phi \ \curvearrowright \ \Psi \\ & & (M_1 \otimes M_2) \otimes M_3 \end{array}$$

mit den multilinearen Abbildungen

$$\begin{aligned} h_1\colon M_1 \times M_2 \times M_3 &\longrightarrow M_1 \otimes (M_2 \otimes M_3)\\ (m_1, m_2, m_3) &\longmapsto m_1 \otimes (m_2 \otimes m_3),\\ h_2\colon M_1 \times M_2 \times M_3 &\longrightarrow (M_1 \otimes M_2) \otimes M_3\\ (m_1, m_2, m_3) &\longmapsto (m_1 \otimes m_2) \otimes m_3. \end{aligned}$$

Wegen (i) und (ii) erhalten wir lineare Abbildungen

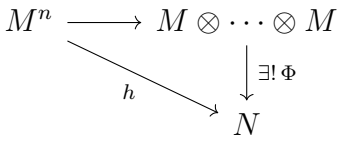
$$\Phi\colon M_1 \otimes (M_2 \otimes M_3) \longrightarrow (M_1 \otimes M_2) \otimes M_3$$
$$\text{und} \quad \Psi\colon (M_1 \otimes M_2) \otimes M_3 \longrightarrow M_1 \otimes (M_2 \otimes M_3)$$

mit $h_2 = \Phi \circ h_1$ und $h_1 = \Psi \circ h_2$, es sind also $\Psi \circ \Phi \circ h_1 = \Psi \circ h_2$. Aber id statt $\Psi \circ \Phi$ tut das genau so; wegen der Eindeutigkeit und (i) erhalten wir also $\Psi \circ \Phi = \mathrm{id}$. Analog gilt $\Phi \circ \Psi = \mathrm{id}$, d. h. Φ und Ψ sind Isomorphismen. □

6. Tensor-, symmetrische- und äußere Potenzen

Für diesen Abschnitt seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$ und M, N seien R-Moduln.

Das Tensorprodukt „klassifiziert“ bilineare Abbildungen. Wir wollen in diesem Abschnitt versuchen, eine ähnliche universelle Abbildungseigenschaft für symmetrische– und alternierende Abbildungen zu erarbeiten:



Definition IX.6.1: Wir bezeichnen

(i) $\mathrm{Mult}^n_M(N) := \mathfrak{m}_R(M \times \cdots \times M, N)$,

(ii) $\mathrm{Sym}^n_M(N) := \{h \in \mathrm{Mult}^n_M(N) \mid h \text{ ist symmetrisch}\}$,

(iii) $\mathrm{Alt}^n_M(N) := \{h \in \mathrm{Mult}^n_M(N) \mid h \text{ ist alternierend}\}$.

Bemerkung IX.6.2: Die Mengen $\mathrm{Mult}^n_M(N)$, $\mathrm{Sym}^n_M(N)$ und $\mathrm{Alt}^n_M(N)$ sind auf die gewohnte Weise (d. h. mit den punktweisen Verknüpfungen) R-Moduln.

$$(h_1 + h_2)(x_1, \ldots, x_n) := h_1(x_1, \ldots, x_n) + h_2(x_1, \ldots, x_n),$$
$$(\lambda h)(x_1, \ldots, x_n) := \lambda h(x_1, \ldots, x_n).$$

Beweis: Multilinearität, Symmetrie und Alterniertheit bleiben unter Summenbildung und Multiplikation mit Skalaren erhalten. □

Satz 31 (Tensor-, symmetrische- und äußere Potenz): *Es sei n eine positive ganze Zahl.*

(i) *Es gibt einen R-Modul $T^n(M)$ und ein multilineares $t\colon M^n \to T^n(M)$, sodass es für alle R-Moduln N und $h \in \mathrm{Mult}^n_M(N)$ genau eine lineare Abbildung $\Phi\colon T^n(M) \to N$ mit $\Phi \circ t = h$ gibt.*

(ii) *Es gibt einen R-Modul $S^n(M)$ und ein multilineares $s\colon M^n \to S^n(M)$, sodass es für alle R-Moduln N und $h \in \mathrm{Sym}^n_M(N)$ genau eine lineare Abbildung $\Phi\colon S^n(M) \to N$ mit $\Phi \circ t = h$ gibt.*

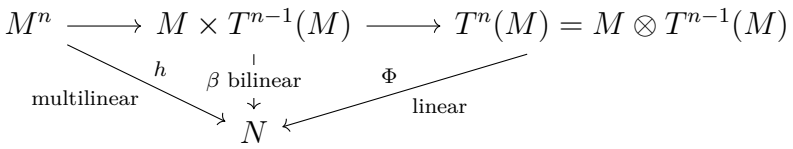
(iii) *Es gibt einen R-Modul $\bigwedge^n(M)$ und ein multilineares $a\colon M^n \to \bigwedge^n(M)$, sodass es für alle R-Moduln N und $h \in \mathrm{Alt}^n_M(N)$ genau eine lineare Abbildung $\Phi\colon \bigwedge^n(M) \to N$ mit $\Phi \circ a = h$ gibt.*

Definition IX.6.3: In der Situation von Satz 31 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 31): (i) Wir definieren $T^n(M)$ rekursiv wie folgt: Wir setzen $T^0(M) := R$, $T^1(M) := M$ und $T^n := M \otimes_R T^{n-1}(M)$. Aus Proposition IX.5.8 wissen wir $T^n(M) = \bigotimes_{i=1}^n M$. Wir weisen die universelle Eigenschaft per Induktion nach nach. Die Abbildung

$$t\colon M^n \longrightarrow T^n(M) = \bigotimes_{i=1}^{n} M, \qquad (m_1, \ldots, m_n) \longmapsto m_1 \otimes \cdots \otimes m_n$$

ist multilinear und wir haben das Diagramm



Wie im Beweis von Proposition IX.5.8 erhalten wir mit der Induktionsvoraussetzung und Bemerkung IX.5.9 erhalten wir eine bilineare Abbildung

$$\beta\colon M \times T^{n-1}(M) \longrightarrow N, \qquad \beta(m_1, m_2 \otimes \cdots \otimes m_n) = h(m_1, \ldots, m_n).$$

Wegen der universellen Abbildungseigenschaft von $M \otimes T^{n-1}(M)$ erhalten wir eine lineare Abbildung $\Phi\colon T^n(M) \to N$ mit $\Phi(m_1 \otimes \cdots \otimes m_n) = h(m_1, \ldots, m_n)$. Diese ist wieder eindeutig, da sie auf Erzeugern vorgegeben ist.

(ii) Für diese Aussage wollen wir verwenden, dass $T^n(M) = \bigotimes_{i=1}^n M$.

Die Abbildung $t\colon M^n \to T^n(M)$, $(m_1, \ldots, m_n) \mapsto m_1 \otimes \cdots \otimes m_n$ ist zwar multilinear, aber nicht symmetrisch. Dazu müsste nämlich für alle $\sigma \in S_n$ gelten, dass

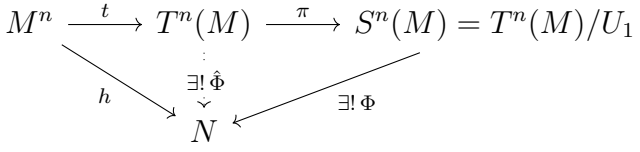
$$m_1 \otimes \cdots \otimes m_n = m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)}.$$

Diesen Mangel der Abbildung t wollen wir (wie üblich) durch Quotientenbildung beheben: Wir definieren

$$U_1 := \langle \{ m_1 \otimes \cdots \otimes m_n - m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} \mid m_1, \ldots, m_n \in M, \sigma \in S_n \} \rangle$$

und setzen $S^n(M) := T^n(M)/U_1$. Wie üblich bezeichne $\pi\colon T^n(M) \to S^n(M)$ die kanonische Projektion. In dieser Situation haben eine multilineare Abbildung $s\colon M^n \to S^n(M)$, nämlich $s := \pi \circ t$. Dieses s ist sogar symmetrisch. Es bleibt zu zeigen, dass wir auch hier eine universelle Abbildungseigenschaft haben.

Sei dazu $h\colon M^n \to N$ multilinear und symmetrisch. Wir wollen uns davon überzeugen, dass wir dann tatsächlich das folgende kommutative Diagramm haben:



Wegen (i) erhalten wir eine lineare Abbildung $\hat{\Phi}\colon T^n(M) \to N$ mit $\hat{\Phi} \circ t = h$ und für $\sigma \in S_n$ gilt

$$\hat{\Phi}(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) = h(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) = h(m_1, \ldots, m_n) = \hat{\Phi}(m_1, \ldots, m_n),$$

d. h. $U_1 \subseteq \operatorname{Kern} \hat{\Phi}$. Wegen des Homomorphiesatzes erhalten wir deshalb eine lineare Abbildung $\Phi\colon S^n(M) \to N$ mit $\Phi \circ \pi = \hat{\Phi}$, d. h. es ist

$$\Phi \circ s = \Phi \circ \pi \circ t = \hat{\Phi} \circ t = h.$$

Zusammengefasst: Das Diagramm ist kommutativ, da das linke Dreieck in diesem Diagramm wegen der universellen Abbildungseigenschaft von T^n kommutativ ist und das rechte Dreieck wegen des Homomorphiesatzes.

(iii) Wir gehen analog zu (ii) vor und wollen den Mangel von t wieder durch Quotientenbildung beheben. Dazu definieren wir den Untermodul

$$U_2 := \langle \{ m_1 \otimes \cdots \otimes m_n \mid m_1, \ldots, m_n \in M, m_i = m_j \text{ für } 1 \leq i, j \leq n \text{ mit } i \neq j \} \rangle$$

und $\bigwedge^n(M) := T^n(M)/U_2$. Die alternierende multilineare Abbildung

$$a\colon M^n \longrightarrow \bigwedge^n(M), \qquad (m_1, \ldots, m_n) \longmapsto [m_1 \otimes \cdots \otimes m_n]$$

leistet dann das Gewünschte. □

Bemerkung IX.6.4 (Trivialfälle):

(i) Für $n = 0$ ist $R = T^0(M) = S^0(M) = \bigwedge^0(M)$,

(ii) Für $n = 1$ ist $M = T^1(M) = S^1(M) = \bigwedge^1(M)$.

Bemerkung IX.6.5 (Eindeutigkeit): Ähnlich wie beim Tensorprodukt kann jeweils aus der universellen Abbildungseigenschaft gefolgert werden, dass $T^n(M)$, $S^n(M)$ und $\bigwedge^n(M)$ eindeutig bis auf Isomorphie sind.

Notation IX.6.6: Seien $m_1, \dots, m_n \in M$. In der Situation von Satz 31 schreiben wir

(i) $m_1 \otimes \cdots \otimes m_n := t(m_1, \dots, m_n)$,

(ii) $m_1 \odot \cdots \odot m_n := s(m_1, \dots, m_n)$,

(iii) $m_1 \wedge \cdots \wedge m_n := a(m_1, \dots, m_n)$.

Bemerkung IX.6.7 (Erzeuger und Rechenregeln): (i) $T^n(M)$ wird erzeugt von $\{m_1 \otimes \cdots \otimes m_n \mid m_1, \dots, m_n \in M\}$ (vergleiche Bemerkung IX.4.12), also wird $S^n(M)$ erzeugt von $\{m_1 \odot \cdots \odot m_n \mid m_1, \dots, m_n \in M\}$ und $\bigwedge^n(M)$ wird erzeugt von $\{m_1 \wedge \cdots \wedge m_n \mid m_1, \dots, m_n \in M\}$.

(ii) Für $m_1, \dots, m_n \in M$, $m_i' \in M$ und $r \in R$ gelten die Rechenregeln

$$\begin{aligned} m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i + rm_i' \otimes m_n & \\ = m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i \otimes m_{i+1} \otimes \cdots \otimes m_n & \\ + rm_1 \otimes \cdots \otimes m_{i-1} \otimes m_i' \otimes m_{i+1} \otimes \cdots \otimes m_n, & \end{aligned}$$

analog für „$\odot$“ und „$\wedge$“; für $\sigma \in S_n$ gelten

$$m_1 \odot \cdots \odot m_n = m_{\sigma(1)} \odot \cdots \odot m_{\sigma(n)}, \quad m_1 \wedge \cdots \wedge m_n = \operatorname{sgn}(\sigma) m_{\sigma(1)} \wedge \cdots \wedge m_{\sigma(n)},$$

und gilt $m_i = m_j$ mit $i \neq j$, so ist $m_1 \wedge \cdots \wedge m_n = \mathbf{0}$.

Proposition IX.6.8 (r-tes äußeres Produkt eines freien Moduls von Rang r): *Sei M ein freier R-Modul mit Basis $\{b_1, \dots, b_r\}$. Dann ist $\{b_1 \wedge \cdots \wedge b_r\}$ eine Basis von $\bigwedge^r(M)$.*

Lemma IX.6.9 (Determinante revised): *Sei M ein R-Modul und $n \in \mathbb{N}$. Die Abbildung*

$$h \colon M^n \longrightarrow T^n(M), \qquad (m_1, \dots, m_n) \longmapsto \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)})$$

ist eine alternierende multilineare Abbildung.

Beweis: Die Multilinearität ist klar, da das Tensorprodukt multilinear ist. Für die Alterniertheit seien $i, j \in \{1, \dots, n\}$ mit $i \neq j$ und $m_1, \dots, m_n \in M$ mit $m_i = m_j$. Für ein $\sigma \in S_n$ setzen wir $\tilde{\sigma} := (i, j) \circ \sigma$. Für dieses $\tilde{\sigma}$ gelten $\tilde{\sigma}(k) = (i, j) \circ \sigma(k) = \sigma(k)$, falls $k \notin \{\sigma^{-1}(i), \sigma^{-1}(j)\}$, $\tilde{\sigma}(\sigma^{-1}(i)) = j$ und $\tilde{\sigma}(\sigma^{-1}(j)) = i$; für alle $k \in \{1, \dots, n\}$ gilt also $m_{\sigma(k)} = m_{\tilde{\sigma}(k)}$. Damit ist

$$\begin{aligned} h(m_1, \dots, m_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)}) \\ &= \sum_{\sigma \in A_n} m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} + \sum_{\sigma \in A_n} (-1) m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} = \mathbf{0}, \end{aligned}$$

was wir zeigen wollten. □

Beweis (Proposition IX.6.8): Nach Bemerkung IX.6.7 ist $\{b_1 \wedge \cdots \wedge b_r\}$ ein Erzeugendensystem für $\bigwedge^r(M)$. Bleibt die lineare Unabhängigkeit zu zeigen, d. h. für alle $r \in R$ muss gelten: Ist $r \neq 0$, so ist $rb_1 \wedge \cdots \wedge b_r \neq 0$.

In der in Lemma IX.6.9 beschriebenen Situation haben wir das Diagramm

$$\begin{array}{ccc} M^r & \xrightarrow{a} & \bigwedge^r(M) \\ & \searrow^{h} & \downarrow \exists!\, \Phi \\ & & T^r(M) \end{array}$$

wobei wir wegen der universellen Abbildungseigenschaft die R-lineare Abbildung $\Phi\colon \bigwedge^r(M) \to T^r(M)$ mit $h = \Phi \circ a$ erhalten. Es ist

$$\Phi(rb_1 \wedge \cdots \wedge b_r) = r \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)} \otimes \cdots \otimes b_{\sigma(n)} \neq 0,$$

da $\{b_{\sigma(1)} \otimes \cdots \otimes b_{\sigma(n)} \mid \sigma \in S_n\}$ Teilmenge einer Basis und damit linear unabhängig ist (siehe Proposition IX.6.11). □

Beispiel IX.6.10 (Zweite Potenzen für $\mathbb{R}^3$): Es sei $M = \mathbb{R}^3$ mit der Standardbasis $\{e_1, e_2, e_3\}$. Dann gelten:

(i) $\{e_1 \otimes e_1, e_1 \otimes e_2, e_1 \otimes e_3, e_2 \otimes e_1, e_2 \otimes e_2, e_2 \otimes e_3, e_3 \otimes e_1, e_3 \otimes e_2, e_3 \otimes e_3\}$ ist Basis von $T^2(M)$,

(ii) $\{e_1 \odot e_1, e_1 \odot e_2, e_1 \odot e_3, e_2 \odot e_2, e_2 \odot e_3, e_3 \odot e_3\}$ ist Basis von $S^2(M)$,

(iii) $\{e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3\}$ ist Basis von $\bigwedge^2(M)$.

Beweis: Wegen Proposition IX.3.11 ist (i) klar.

Zu (iii): Aus Bemerkung IX.6.7 wissen wir, dass $\{e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3\}$ ein Erzeugendensystem für $\bigwedge^2(M)$ ist. Nun zur linearen Unabhängigkeit: Seien $a, b, c \in \mathbb{R}$ mit $ae_1 \wedge e_2 + be_1 \wedge e_3 + ce_2 \wedge e_3 = \mathbf{0}$. Dann gelten

$$\begin{aligned} \mathbf{0} &= (ae_1 \wedge e_2 + be_1 \wedge e_3 + ce_2 \wedge e_3) \wedge e_1 \\ &= ae_1 \wedge e_2 \wedge e_1 + be_1 \wedge e_3 \wedge e_1 + ce_2 \wedge e_3 \wedge e_1 = \mathbf{0} + \mathbf{0} + ce_1 \wedge e_2 \wedge e_3, \\ \mathbf{0} &= (ae_1 \wedge e_2 + be_1 \wedge e_3 + ce_2 \wedge e_3) \wedge e_2 \\ &= ae_1 \wedge e_2 \wedge e_2 + be_1 \wedge e_3 \wedge e_2 + ce_2 \wedge e_3 \wedge e_2 = \mathbf{0} + -be_1 \wedge e_2 \wedge e_3 + \mathbf{0}, \\ \mathbf{0} &= (ae_1 \wedge e_2 + be_1 \wedge e_3 + ce_2 \wedge e_3) \wedge e_3 \\ &= ae_1 \wedge e_2 \wedge e_3 + be_1 \wedge e_3 \wedge e_3 + ce_2 \wedge e_3 \wedge e_3 = ae_1 \wedge e_2 \wedge e_3 + \mathbf{0} + \mathbf{0}, \end{aligned}$$

also sind $a = b = c = 0$, was wir zeigen wollten.

Aussage (ii) zeigen wir später. □

Proposition IX.6.11: *Sei M ein freier R-Modul mit Basis $\{b_1, \dots, b_r\}$. Dann gelten:*

(i) *$T^n(M)$ hat die Basis $\{b_{i_1} \otimes \cdots \otimes b_{i_n} \mid i_j \in \{1, \dots, r\}, 1 \leq j \leq n\}$,*

(ii) *$S^n(M)$ hat die Basis $\{b_1^{\nu_1} \odot \cdots \odot b_r^{\nu_r} \mid \nu_1 + \cdots + \nu_r = n\}$,*

(iii) *$\bigwedge^n(M)$ hat die Basis $\{b_{i_1} \wedge \cdots \wedge b_{i_n} \mid 1 \leq i_1 < i_2 < \cdots < i_{n-1} < i_n \leq r\}$.*

Beweis: (i) Folgt aus Proposition III.3.11 (die dort verwendeten Argumente gehen auch für Moduln durch).

(ii) 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 $r_{(i_1,\dots,i_n)} \in R$ für $1 \leq i_1 < \cdots < i_n \leq r$ mit

$$\sum_{1 \leq i_1 < \cdots < i_n \leq r} r_{(i_1,\dots,i_n)} b_{i_1} \wedge \cdots \wedge b_{i_n} = \mathbf{0}.$$

Wir wollen für jeden n-Tupel $\boldsymbol{j} = (j_1, \dots, j_n)$ 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 $\sigma_{\boldsymbol{j}}(1) = j_1, \dots, \sigma_{\boldsymbol{j}}(n) = j_n$ und $\sigma_{\boldsymbol{j}}(n+1), \dots, \sigma_{\boldsymbol{j}}(r)$ die Werte in $\{1, \dots, r\} - \{j_1, \dots, j_n\}$ sind.

Jetzt ist

$$b_{j_1} \wedge \cdots \wedge b_{j_n} \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} = (-1)^{\ell} b_1 \wedge \cdots \wedge b_r \tag{IX.3}$$

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 $(i_1, \dots, i_n) \neq (j_1, \dots, j_n)$. Einsetzen in Gl. (IX.3) gibt

$$\begin{aligned} \mathbf{0} &= \Big(\sum_{1 \leq i_1 < \cdots < i_n \leq r} r_{(i_1,\dots,i_n)} b_{i_1} \wedge \cdots \wedge b_{i_n} \Big) \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} \\ &= (-1)^{\ell} \cdot r_{(i_1,\dots,i_n)} b_1 \wedge \cdots \wedge b_r, \end{aligned}$$

d. h. $r_{(i_1,\dots,i_n)} = 0$, also die lineare Unabhängigkeit. □

Proposition IX.6.12: *Sei V ein K-Vektorraum der Dimension d. Dann gelten:*

(i) *$T^n(M)$ ist ein K-Vektorraum der Dimension d^n,*

(ii) *$S^n(M)$ ist ein K-Vektorraum der Dimension $\binom{d+n-1}{n}$,*

(iii) *$\bigwedge^n(M)$ ist ein K-Vektorraum der Dimension $\binom{d}{n}$.*

Beweis: (i) Das haben wir bereits in Proposition III.6.11 gezeigt.

(ii) Wir zählen, wie viele Möglichkeiten wir haben, n Elemente aus einer d-elementigen Menge mit Zurücklegen und ohne Beachtung der Reihenfolge zu ziehen.

(iii) Wir zählen, wie viele Möglichkeiten wir haben, n Elemente aus einer d-elementigen Menge ohne Zurücklegen und ohne Beachtung der Reihenfolge zu ziehen. □

Proposition IX.6.13: *Seien M_1, M_2 zwei R-Moduln und $\varphi\colon M_1 \to M_2$ eine R-lineare Abbildung. Dann gibt es eindeutige R-lineare Abbildungen*

$$T^n(\varphi)\colon T^n(M_1) \longrightarrow T^n(M_2),$$
$$S^n(\varphi)\colon S^n(M_1) \longrightarrow S^n(M_2), \qquad \bigwedge^n(\varphi)\colon \bigwedge^n(M_1) \longrightarrow \bigwedge^n(M_2)$$

sodass für alle $v_1, \ldots, v_n \in M_1$ gelten:

(i) $T^n(\varphi)(v_1 \otimes \cdots \otimes v_n) = \varphi(v_1) \otimes \cdots \otimes \varphi(v_n)$,

(ii) $S^n(\varphi)(v_1 \odot \cdots \odot v_n) = \varphi(v_1) \odot \cdots \odot \varphi(v_n)$,

(iii) $\bigwedge^n(\varphi)(v_1 \wedge \cdots \wedge v_n) = \varphi(v_1) \wedge \cdots \wedge \varphi(v_n)$.

Beweis: Für den Beweis verwenden wir die universelle Abbildungseigenschaft der jeweiligen Potenz, exemplarisch zeigen wir die Behauptungen für $\bigwedge^n(\varphi)$.

Die Abbildung $(v_1, \ldots, v_n) \mapsto (\varphi(v_1) \wedge \cdots \wedge \varphi(v_n))$ ist multilinear und alternierend, d. h. wir haben das Diagramm

$$\begin{array}{ccc} M_1^n & \xrightarrow{a} & \bigwedge^n M_1 \\ & \searrow & \downarrow{\scriptstyle \exists!\, \bigwedge^n(\varphi)} \\ & & \bigwedge^n M_2 \end{array}$$

wobei wir aus der universellen Abbildungseigenschaft die eindeutige lineare Abbildung $\bigwedge^n(\varphi)\colon \bigwedge^n M_1 \to \bigwedge^n M_2$ erhalten, die das Gewünschte leistet. □

Proposition IX.6.14: *Seien $\varphi\colon M_1 \to M_2$ und $\varphi_2\colon M_2 \to M_3$ zwei R-lineare Abbildungen. Dann gelten*

(i) $T^n(\varphi_2 \circ \varphi_1) = T^n(\varphi_2) \circ T^n(\varphi_1)$,

(ii) $S^n(\varphi_2 \circ \varphi_1) = S^n(\varphi_2) \circ S^n(\varphi_1)$,

(iii) $\bigwedge^n(\varphi_2 \circ \varphi_1) = \bigwedge^n(\varphi_2) \circ \bigwedge^n(\varphi_1)$.

Beweis: Alle Aussagen zeigt man völlig gleich, wir zeigen exemplarisch (i). Wegen der Definition von $T^n(\cdot)$ gilt für alle $v_1, \ldots, v_n \in M_1$:

$$\begin{aligned}&(T^n(\varphi_2) \circ T^n(\varphi_1))(v_1 \otimes \cdots \otimes v_n)\\ &\quad = T^n(\varphi_2)(\varphi_1(v_1) \otimes \cdots \otimes \varphi_n(v_n)) = \varphi_2(\varphi_1(v_1)) \otimes \cdots \otimes \varphi_2(\varphi_1(v_n)). \quad \text{(IX.4)}\end{aligned}$$

Jetzt ist aber $T^n(\varphi_2 \circ \varphi_1)$ die einzige lineare Abbildung, die Gl. (IX.4) leistet, d. h. $T^n(\varphi_2 \circ \varphi_1) = T^n(\varphi_2) \circ T^n(\varphi_1)$. □

7. Äußere Potenzen und Determinante

In diesem Abschnitt wollen wir mithilfe von äußeren Produkten eine alternative Definition der Determinante geben. Dadurch verschaffen wir uns gleichzeitig eine Determinante über beliebigen kommutativen Ringen mit 1.

Im Folgenden seien $d, n, m \in \mathbb{N}_0$, R stets ein kommutativer Ring mit 1 und $\{e_1, \ldots, e_d\}$ bezeichne stets die Standardbasis des R^d.

Notation IX.7.1: Mit $R^{n\times m}$ bezeichnen wir die Menge der $(n \times m)$-Matrizen mit Einträgen in R, d. h. $R^{n\times m} := \mathrm{Abb}([n] \times [m], R)$. $R^{n\times m}$ ist eine R-Algebra mit der (bereits bekannten) Matrizenmultiplikation.

Bemerkung IX.7.2 („Es kann nur eine geben“): Wir haben das Diagramm

$$\begin{array}{ccc} \prod_{i=1}^d R^d & \xrightarrow{a} & \bigwedge^d R^d \\ & \searrow^{\Phi\circ a} & \downarrow \Phi \\ & & R \end{array}$$

Aus Proposition III.6.11 wissen wir, dass $\bigwedge^d R^d$ ein freier R-Modul mit Basis $e_1 \wedge \cdots \wedge e_d$ ist. Es gibt also nur eine R-lineare Abbildung $\Phi\colon \bigwedge^d R^d \to R$ mit $\Phi(e_1 \wedge \cdots \wedge e_d) = 1$. Wegen der universellen Abbildungseigenschaft gibt es deshalb genau eine multilineare Abbildung $h\colon \prod_{i=1}^d R^d \longrightarrow R$ mit $h(e_1, \ldots, e_d) = 1$. Dieses h ist unsere Determinate.

Definition IX.7.3 (Determinante):

(i) Die eindeutige multilineare und alternierende Abbildung $h\colon \prod_{i=1}^d R^d \to R$ mit $h(e_1, \ldots, e_d) = 1$ heißt *Determinante* und wird mit det notiert.

(ii) Sei $A = (a_{i,j}) \in R^{d\times d}$. Dann definieren wir $\det(A) := \det(a_1, \ldots, a_d)$, wobei $a_j := Ae_j$ die j-te Spalte von A bezeichnet.

(iii) Sei $\varphi \in \mathrm{End}(R^d)$. Dann definieren wir $\det(\varphi) := \det(\varphi(e_1), \ldots, \varphi(e_d))$.

Bemerkung IX.7.4 (Kleine Rechenregeln): In der Situation von Definition IX.7.3 gelten:

(i) $Ae_1 \wedge \cdots \wedge Ae_d = \det(A) e_1 \wedge \cdots \wedge e_d$,

(ii) $\varphi(e_1) \wedge \cdots \wedge \varphi(e_d) = \det(\varphi) e_1 \wedge \cdots \wedge e_d$.

Beweis: Zu (ii): Wir verwenden, dass $\{e_1 \wedge \cdots \wedge e_d\}$ eine Basis von $\bigwedge^d R^d$. Das bedeutet nämlich es muss $r \in R$ mit $\varphi(e_1) \wedge \cdots \wedge \varphi(e_d) = re_1 \wedge \cdots \wedge e_d$ geben. Für die eindeutige Abbildung $\Phi\colon \bigwedge^d R \to R$ aus Bemerkung IX.7.2 gilt

$$\begin{aligned} r\Phi(e_1 \wedge \cdots \wedge e_d) &= \Phi(re_1 \wedge \cdots \wedge e_d) \\ &= \Phi(\varphi(e_1) \wedge \cdots \wedge \varphi(e_d)) = \det(\varphi(e_1), \ldots, \varphi(e_d)) = \det(\varphi). \end{aligned}$$

Aussage (i) zeigt man analog. □

Korollar IX.7.5: *In der Situation von Definition IX.7.3 gilt: Die Abbildung $\bigwedge^d \varphi$ ist gegeben durch*

$$\bigwedge^d \varphi\colon \bigwedge^d R^d \longrightarrow \bigwedge^d R^d, \qquad w \longmapsto \det(\varphi) w.$$

Insbesondere gilt $\varphi(v_1) \wedge \cdots \wedge \varphi(v_d) = \det(\varphi) v_1 \wedge \cdots \wedge v_d$ für alle $v_1, \ldots, v_d \in R^d$.

Beweis: Folgt direkt aus Bemerkung IX.7.4. □

Proposition IX.7.6 (Multiplikativität): *Seien $\varphi_1, \varphi_2\colon R^d \to R^d$ zwei R-lineare Abbildungen und $A_1, A_2 \in R^{d\times d}$. Dann gelten:*

(i) $\det(\varphi_2 \circ \varphi_1) = \det(\varphi_2)\det(\varphi_1)$,

(ii) $\det(A_2 A_1) = \det(A_2)\det(A_1)$.

Beweis: Zu (i): Es gilt $\bigwedge^d(\varphi_2 \circ \varphi_1) = \bigwedge^d \varphi_2 \circ \bigwedge^d \varphi_1$ nach Proposition III.6.14, die Behauptung folgt dann aus Korollar III.7.5.

Zu (ii): Betrachte die R-linearen Abbildungen $v \mapsto A_1 v$, $v \mapsto A_2 v$ und verwende (i). □

Im Folgenden wollen wir $Ae_{j_1} \wedge \cdots \wedge Ae_{j_k}$ für $A \in R^{d\times d}$, $k < d$ und $j_1, \ldots, j_k \in \{1, \ldots, d\}$ untersuchen. Dazu wollen wir die $Ae_{j_1} \wedge \cdots \wedge Ae_{j_k}$ in der Standardbasis von $\bigwedge^k R^d$ ausdrücken, d. h.

$$Ae_{j_1} \wedge \cdots \wedge Ae_{j_k} = \sum_{1 \le i_1 < \cdots < i_k \le d} x_{(i_1,\ldots,i_k)} e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

Wir werden sehen, dass $x_{(i_1,\ldots,i_k)} = (-1)^\ell \det(A_{I,J})$, wobei $I = \{i_1, \ldots, i_k\}$, $J = \{j_1, \ldots, j_k\}$ und $A_{I,J}$ cie Teilmatrix von A mit Zeilen aus I und Spalten aus J ist.

Definition IX.7.7 (Index-Menge):

(i) $I_k^d := \{(i_1, \ldots, i_k) \mid 1 \leq i_1 < \cdots < i_k \leq d\}$,

(ii) Für $\boldsymbol{i} := (i_1, \ldots, i_k) \in I_k^d$ sei $\boldsymbol{i}^{\mathsf{c}} \in I_{d-k}^d$ das eindeutige $(d-k)$-Tupel for das gilt: $\{i_{k+1}, \ldots, i_d\} = \{1, \ldots, d\} - \{i_1, \ldots, i_k\}$ und $i_{k+1} < \cdots < i_d$.
Schreibe $I_{\boldsymbol{i}} := \{i_1, \ldots, i_k\}$ und $e_{\boldsymbol{i}} := e_{i_1} \wedge \cdots \wedge e_{i_k}$. Definiere $\sigma_{\boldsymbol{i}} \in S_d$ durch $\sigma(1) = i_1, \ldots, \sigma(d) := i_d$.
Für $\boldsymbol{i} \in I_k^d$ definieren wir $e_{\boldsymbol{i}} := e_{i_1} \wedge \cdots \wedge e_{i_k}$.

(iii) Wir bezeichnen $V_{\boldsymbol{i}} := \langle e_{i_1}, \ldots, e_{i_k} \rangle$.

Beispiel: Es seien $d = 4$ und $k = 2$. Dann sind zum Beispiel $\boldsymbol{i} = (1, 4) \in I_2^4$, $\boldsymbol{i}^{\mathsf{c}} = (2, 3)$,

$$\sigma_{\boldsymbol{i}} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix},$$

$e_{\boldsymbol{i}} = e_1 \wedge e_4$ und $V_{\boldsymbol{i}} = \langle e_1, e_4 \rangle$.

Definition IX.7.8 (Streichmatrix $\boldsymbol{A_{i,j}}$): Seien $A \in R^{d \times d}$ und $1 \leq k \leq d$. Weiter seien $\boldsymbol{i} = (i_1, \ldots, i_k)$ und $\boldsymbol{j} = (j_1, \ldots, j_k) \in I_k^d$. Dann definieren wir $A_{\boldsymbol{i},\boldsymbol{j}} := (a_{i_r, j_s})_{1 \leq r,s \leq k} \in R^{k \times k}$. Das heißt $A_{\boldsymbol{i},\boldsymbol{j}}$ ist die Teilmatrix von A mit den Zeilen $i_1, \ldots, i_k$ und Spalten $j_1, \ldots, j_k$.

Beispiel IX.7.9: Für die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

und $\boldsymbol{i} = (1, 3)$, $\boldsymbol{j} = (1, 4)$ ist

$$A_{\boldsymbol{i},\boldsymbol{j}} = \begin{pmatrix} 1 & 4 \\ 9 & 12 \end{pmatrix}.$$

Bemerkung IX.7.10 (Was tut $\boldsymbol{A_{i,j}}$?): Seien $A \in R^{d \times d}$ und $\varphi_A \colon v \mapsto Av$ die zugehörige lineare Abbildung. Weiter seien $\boldsymbol{i} = (i_1, \ldots, i_k), \boldsymbol{j} = (j_1, \ldots, j_k) \in I_k^d$ und $A_{\boldsymbol{i},\boldsymbol{j}}$ die zugehörige Streichmatrix. Schließlich sei

$$\mathrm{pr}_{\boldsymbol{i}} \colon R^d \longrightarrow V_{\boldsymbol{i}}, \qquad \sum_{i=1}^{d} x_i e_i \longmapsto \sum_{r=1}^{k} x_{i_r} e_{i_r}.$$

Dann ist $A_{\boldsymbol{i},\boldsymbol{j}}$ die Abbildungsmatrix von $\mathrm{pr}_{\boldsymbol{i}} \circ \varphi_A|_{V_{\boldsymbol{j}}} \colon V_{\boldsymbol{j}} \to V_{\boldsymbol{i}}$ bezüglich der Basen $e_{i_1}, \ldots, e_{i_k}$ beziehungsweise $e_{j_1}, \ldots, e_{j_k}$.

Proposition IX.7.11 (Abbildungsmatrix von $\bigwedge^k \varphi_A$): *Für $A = (a_{i,j}) \in R^{d\times d}$ und einen Multiindindex $\boldsymbol{j} = (j_1, \dots, j_k) \in I_k^d$ gilt:*

$$Ae_{j_1} \wedge \cdots \wedge Ae_{j_k} = \sum_{\boldsymbol{i}\in I_k^d} \det A_{\boldsymbol{i},\boldsymbol{j}} e_{i_1} \wedge \cdots \wedge e_{i_k}$$

Beispiel IX.7.12: Für die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

und den Multiindex $\boldsymbol{j} = (1, 4)$ gilt

$$\begin{aligned} Ae_1 \wedge Ae_4 &= \begin{pmatrix} 1 \\ 5 \\ 9 \\ 13 \end{pmatrix} \wedge \begin{pmatrix} 4 \\ 8 \\ 12 \\ 16 \end{pmatrix} \\ &= (8-20)e_1 \wedge e_2 + (12-36)e_1 \wedge e_3 + (16-52)e_1 \wedge e_4 \\ &\quad + (60-72)e_2 \wedge e_3 + (80-104)e_2 \wedge e_4 + (144-156)e_3 \wedge e_4. \end{aligned}$$

Beweis (von Proposition IX.7.11): Sei $\omega := Ae_{j_1} \wedge \cdots \wedge Ae_{j_k} \in \bigwedge^k R^d$. Dann können wir schreiben $\omega = \sum_{\boldsymbol{i}\in I_k^d} x_{\boldsymbol{i}} e_{\boldsymbol{i}}$, wobei die Koeffizienten $x_{\boldsymbol{i}}$ unbekannt sind. Wir wollen zeigen, dass $x_{\boldsymbol{i}} = \det A_{\boldsymbol{i},\boldsymbol{j}}$.

Dazu wollen wir für jedes $\boldsymbol{i} \in I_k^d$ das Produkt $\omega \wedge e_{\boldsymbol{i}^c} \in \bigwedge^d R^d$ bezüglich der Basis $\{e_1 \wedge \cdots \wedge e_d\}$ auf zwei verschiedene Weisen berechnen und vergleichen.

Erstens ist

$$\omega \wedge e_{\boldsymbol{i}^c} = x_{\boldsymbol{i}} e_{\boldsymbol{i}} \wedge e_{\boldsymbol{i}^c} = \operatorname{sgn}(\sigma) x_{\boldsymbol{i}} e_1 \wedge \cdots \wedge e_d.$$

Andererseits ist, da wir $Ae_{j_r} = \sum_{i=1}^d a_{i,j_r} e_i$ schreiben können,

$$\begin{aligned} \omega \wedge e_{\boldsymbol{i}^c} &= \Big(\sum_{i=1}^d a_{i,j_1} e_i \wedge \cdots \wedge \sum_{i=1}^d a_{i,j_k} e_i \Big) \wedge e_{\boldsymbol{i}^c} \\ &= \Big(\sum_{r=1}^k a_{i_r,j_1} e_{i_r} \wedge \cdots \wedge \sum_{r=1}^k a_{i_r,j_k} e_{i_r} \Big) \wedge e_{\boldsymbol{i}^c} \\ &= (\mathrm{pr}_{\boldsymbol{i}}(Ae_{j_1}) \wedge \cdots \wedge \mathrm{pr}_{\boldsymbol{i}}(Ae_{j_k})) \wedge e_{\boldsymbol{i}^c}, \end{aligned}$$

wobei $\mathrm{pr}_{\boldsymbol{i}}$ die Projektion aus Bemerkung IX.7.10 ist.

Wir haben das folgende Diagramm für $\mathrm{pr}_{\boldsymbol{i}} \circ \phi_A|_{V_{\boldsymbol{j}}} \in \mathrm{Hom}(V_{\boldsymbol{j}}, V_{\boldsymbol{i}})$:

$$\begin{array}{ccc} \bigwedge^k V_{\boldsymbol{j}} & \xrightarrow{\cong} & \bigwedge^k R^k \\ {\scriptstyle \bigwedge^k(\mathrm{pr}_{\boldsymbol{i}} \circ \phi_A|_{V_{\boldsymbol{j}}})}\downarrow & & \downarrow{\scriptstyle \bigwedge^k(\phi_{A_{\boldsymbol{i},\boldsymbol{j}}})} \\ \bigwedge^k V_{\boldsymbol{i}} & \xrightarrow[\cong]{} & \bigwedge^k R^k \end{array}$$

wobei $e_{j_1} \wedge \cdots \wedge e_{j_k}$ abgebildet wird auf $\det A_{\boldsymbol{i},\boldsymbol{j}} e_{i_1} \wedge \cdots \wedge e_{i_k}$ und $f_1 \wedge \cdots \wedge f_k$ abgebildet wird auf $A_{\boldsymbol{i},\boldsymbol{j}} f_1 \wedge \cdots \wedge A_{\boldsymbol{i},\boldsymbol{j}} f_k$ (hierbei bezeichnet $\{f_1, \ldots, f_k\}$ die Standardbasis von R^k, um Verwechslungen vorzubeugen).

Das heißt $(\mathrm{pr}_{\boldsymbol{i}}(Ae_{j_1}) \wedge \cdots \wedge \mathrm{pr}_{\boldsymbol{i}}(Ae_{j_k})) = \det A_{\boldsymbol{i},\boldsymbol{j}} e_{\boldsymbol{i}}$. Damit ist

$$\omega \wedge e_{\boldsymbol{i}^{\mathsf{c}}} = \det A_{\boldsymbol{i},\boldsymbol{j}} e_{\boldsymbol{i}} \wedge e_{\boldsymbol{i}^{\mathsf{c}}} = \mathrm{sgn}(\sigma) \det A_{\boldsymbol{i},\boldsymbol{j}} e_1 \wedge \cdots \wedge e_d. \qquad \square$$

Sind $A \in R^{d \times d}$ und $\boldsymbol{j} = (j_1, \ldots, j_k) \in J_k^d$ gegeben, so ist

$$\begin{aligned} \det Ae_1 \wedge \cdots \wedge e_d &= Ae_1 \wedge \cdots \wedge Ae_d \\ &= \mathrm{sgn}(\sigma_{\boldsymbol{j}}) Ae_{j_1} \wedge \cdots \wedge Ae_{j_d} \\ &= \mathrm{sgn}(\sigma_{\boldsymbol{j}})(Ae_{j_1} \wedge \cdots \wedge Ae_{j_k}) \wedge (Ae_{j_{k+1}} \wedge \cdots \wedge Ae_d) \\ &= \mathrm{sgn}(\sigma_{\boldsymbol{j}}) \Big(\sum_{\boldsymbol{i} \in I_k^d} \det A_{\boldsymbol{i},\boldsymbol{j}} e_{\boldsymbol{i}} \Big) \wedge \Big(\sum_{\boldsymbol{i}' \in I_{d-k}^d} \det A_{\boldsymbol{i}',\boldsymbol{j}^{\mathsf{c}}} e_{\boldsymbol{i}'} \Big) \\ &= \mathrm{sgn}(\sigma_{\boldsymbol{j}}) \sum_{\boldsymbol{i} \in I_k^d} \det A_{\boldsymbol{i},\boldsymbol{j}} \det A_{\boldsymbol{i}^{\mathsf{c}},\boldsymbol{j}^{\mathsf{c}}} e_{\boldsymbol{i}} \wedge e_{\boldsymbol{i}^{\mathsf{c}}} \\ &= \mathrm{sgn}(\sigma_{\boldsymbol{j}}) \sum_{\boldsymbol{i} \in I_k^d} \mathrm{sgn}(\sigma_{\boldsymbol{i}}) \det A_{\boldsymbol{i},\boldsymbol{j}} \det A_{\boldsymbol{i}^{\mathsf{c}},\boldsymbol{j}^{\mathsf{c}}} e_1 \wedge \cdots \wedge e_d, \end{aligned}$$

da die Dachprodukte der Summanden Null sind, falls $\boldsymbol{i}' \neq \boldsymbol{i}^{\mathsf{c}}$.

Satz 32 (Verallgemeinerte Laplace-Regel): *Für $A \in R^{d \times d}$, $\boldsymbol{j} \in J_k^d$ gilt*

$$\det A = \mathrm{sgn}(\sigma_{\boldsymbol{j}}) \sum_{\boldsymbol{i} \in I_k^d} \mathrm{sgn}(\sigma_{\boldsymbol{i}}) \det A_{\boldsymbol{i},\boldsymbol{j}} \det A_{\boldsymbol{i}^{\mathsf{c}},\boldsymbol{j}^{\mathsf{c}}} e_1 \wedge \cdots \wedge e_d.$$

Beispiel IX.7.13 (Anwendung der verallgemeinerten Laplace-Regel): Für die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \in \mathbb{R}^{4 \times 4},$$

$k = 2$ und $\boldsymbol{j} = (1, 4)$ ist

$$\sigma_{\boldsymbol{j}} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix} = (2, 4, 3),$$

d. h. $\operatorname{sgn}(\sigma_{\boldsymbol{j}}) = 1$. Für die restlichen Multiindizes haben wir die Tabelle

$\boldsymbol{i}$	$\sigma_{\boldsymbol{i}}$	$\operatorname{sgn}(\sigma_{\boldsymbol{i}})$	$\boldsymbol{i}^{\mathsf{c}}$
$(1, 2)$	id	1	$(3, 4)$
$(1, 3)$	$(2, 3)$	-1	$(2, 4)$
$(1, 4)$	$(2, 4, 3)$	1	$(2, 3)$
$(2, 3)$	$(1, 2, 3)$	1	$(1, 4)$
$(2, 4)$	$(1, 2, 4, 3)$	-1	$(1, 3)$
$(3, 4)$	$(1, 3)(2, 4)$	1	$(1, 2)$

Nach Satz 32 haben wir

$$\begin{aligned} \det A = {} & \det \begin{pmatrix} 1 & 4 \\ 5 & 8 \end{pmatrix} \det \begin{pmatrix} 10 & 11 \\ 14 & 15 \end{pmatrix} - \det \begin{pmatrix} 1 & 4 \\ 9 & 12 \end{pmatrix} \begin{pmatrix} 6 & 7 \\ 14 & 15 \end{pmatrix} \\ & + \det \begin{pmatrix} 1 & 4 \\ 13 & 16 \end{pmatrix} \det \begin{pmatrix} 6 & 7 \\ 10 & 11 \end{pmatrix} + \det \begin{pmatrix} 5 & 8 \\ 9 & 12 \end{pmatrix} \det \begin{pmatrix} 2 & 3 \\ 14 & 15 \end{pmatrix} \\ & - \det \begin{pmatrix} 5 & 8 \\ 13 & 16 \end{pmatrix} \det \begin{pmatrix} 2 & 3 \\ 10 & 11 \end{pmatrix} + \det \begin{pmatrix} 9 & 12 \\ 13 & 16 \end{pmatrix} \det \begin{pmatrix} 2 & 3 \\ 6 & 7 \end{pmatrix} \\ = {} & 0. \end{aligned}$$

8. Das charakteristische Polynom und die äußere Potenz

In diesem Abschnitt seien stets V ein K-Vektorraum der Dimension d, B eine Basis von V, $\varphi \in \operatorname{End}(V)$, $A = D_{B,B}(\varphi)$ die Darstellungsmatrix von φ bezüglich B und

$$\chi_A := \chi_\varphi := \mathrm{CP}_\varphi = \mathrm{CP}_A = \det(\varphi - X \operatorname{id}) = \det(A - XI_d) = \sum_{i=1}^{d} c_i X^i \in K[X]$$

das charakteristische Polynom von φ. Mithilfe der Leibnizregel können wir schreiben

$$\chi_A = \sum_{\sigma \in S_d} \operatorname{sgn}(\sigma) b_{1,\sigma(1)} \cdots b_{d,\sigma(d)} \tag{IX.5}$$

wobei $B := A - XI_d = (b_{i,j})_{1 \le i,j \le d}$.

Außerdem induziert φ eine lineare Abbildung

$$\bigwedge^d \varphi\colon \bigwedge^d V \longrightarrow \bigwedge^d V, \qquad v \longmapsto \det(\varphi)v$$

Definition IX.8.1: Sei $B \in K^{d\times d}$ eine Matrix. Dann heißt $\operatorname{Tr}(B) := \sum_{i=1}^d b_{i,j}$ die *Spur von B* (englisch: trace).

Bemerkung IX.8.2: Den Koeffizienten c_{d-1} im charakteristischen Polynom erhält man wie folgt: In Gl. (IX.5) trägt der Summand zu $\sigma = \mathrm{id}$ zu c_{d-1} bei, also: $(a_{1,1} - X)\cdots(a_{d,d} - X)$; d. h. es ist

$$c_{d-1} = (-1)^{d-1}(a_{1,1} + \cdots + a_{d,d}) = (-1)^{d-1}\operatorname{Tr}(A).$$

Insbesondere ist $\operatorname{Tr}(A)$ eine Ähnlichkeitsinvariante.

Definition IX.8.3: Wir nennen $\operatorname{Tr}(\varphi) := \operatorname{Tr}(D_{B,B}(\varphi)) = \operatorname{Tr}(A)$ die *Spur des Endomorphismus φ.*

Nach Bemerkung IX.8.2 ist die Spur eines Endomorphismus wohldefiniert, d. h. hängt nicht von der gewählten Basis ab.

Beispiel IX.8.4: Es ist $\operatorname{Tr}(\bigwedge^d \varphi) = \det A$, denn $\bigwedge^d \varphi\colon \bigwedge^d V \to \bigwedge^d V$ ist die Abbildung, die v abbildet auf $(\det A)v$ und $\bigwedge^d V$ ist eindimensional.

Bemerkung IX.8.5 (Einige Koeffizienten von χ_A):

(i) $c_d = (-1)^d = (-1)^d \operatorname{Tr}(\bigwedge^0 \varphi)$,

(ii) $c_{d-1} = (-1)^{d-1}\operatorname{Tr}(A) = (-1)^{d-1}\operatorname{Tr}(\varphi)$,

(iii) $c_0 = \chi_A(0) = \det(A) = \operatorname{Tr}(\bigwedge^d \varphi)$.

Satz 33 (Alle Koeffizienten von χ_A): *Sei $\chi_\varphi = \sum_{i=0}^d c_i X^i \in K[X]$ das charakteristische Polynom von $\varphi \in \operatorname{End}(V)$. Dann gilt*

$$c_k = (-1)^k \operatorname{Tr}(\bigwedge^{d-k} \varphi)$$

für die induzierten linearen Abbildungen $\bigwedge^{d-k}\varphi \in \operatorname{End}(\bigwedge^{d-k} V)$.

Lemma IX.8.6 (Spur von $\bigwedge^k \varphi$): *Es gilt*

$$\operatorname{Tr}(\bigwedge^k \varphi) = \sum_{\boldsymbol{j}\in I_k^d} \det A_{\boldsymbol{j},\boldsymbol{j}}.$$

Beweis: Aus Proposition IX.7.11 wissen wir, dass

$$Ab_{j_1} \wedge \cdots \wedge Ab_{j_k} = \sum_{\boldsymbol{i} \in I_k^d} \det A_{\boldsymbol{i},\boldsymbol{j}} b_{i_1} \wedge \cdots \wedge b_{i_k}.$$

Jetzt ist

$$\operatorname{Tr}(\bigwedge^k \varphi) = \sum_{\boldsymbol{j} \in I_k^d} \text{Koeffizient von } b_{\boldsymbol{j}} \text{ in } \bigwedge^k \varphi(b_{\boldsymbol{j}}),$$

und nach Proposition IX.7.11 ist der Koeffizient gerade $\det A_{\boldsymbol{j},\boldsymbol{j}}$. □

Beweis (von Satz 33): Es sei $B := (A - XI_d) = (b_{i,j})$ wie zuvor. Wegen der Leibniz-Formel ist

$$\chi_\varphi = \det B = \sum_{\sigma \in S_d} \operatorname{sgn}(\sigma) b_{1,\sigma(1)} \cdots b_{d,\sigma(d)}.$$

Setze $p_\sigma := \operatorname{sgn}(\sigma) b_{1,\sigma(1)} \cdots b_{d,\sigma(d)}$.

(1) p_σ trägt zum Koeffizienten von X^k genau dann bei, wenn es einen Tupel $\boldsymbol{i} = (i_1, \ldots, i_k) \in I_k^d$ gibt, sodass $\sigma(i_1) = i_1, \ldots, \sigma(i_k) = i_k$. Für jedes solche $\boldsymbol{i}$ ist der Beitrag

$$(-1)^k \operatorname{sgn}(\sigma) \prod_{t=k+1}^{n} a_{i_t,\sigma(i_t)}.$$

(2) Für jede Permutation σ wie in (1) erklären wir $\sigma' \in \operatorname{Sym}\{i_{k+1}, \ldots, i_d\}$ durch $\sigma' := \sigma|_{\{i_{k+1},\ldots,i_d\}}$. Da $\sigma|_{\{i_1,\ldots,i_k\}} = \mathrm{id}$, ist $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma')$. Im Folgenden schreiben wir $I := \{i_1, \ldots, i_k\}$ und $I^{\mathsf{c}} := \{i_{k+1}, \ldots, i_d\}$.

(3) Addiert man für festes $\boldsymbol{i} \in I_k^d$ für alle σ wie in (1) die Koeffizienten auf, so erhält man die Summe

$$\sum_{\sigma' \in \operatorname{Sym}(I^{\mathsf{c}})} (-1)^k \operatorname{sgn}(\sigma') a_{i_{k+1},\sigma'(k+1)} \cdots a_{i_d,\sigma'(d)} = (-1)^k \det A_{\boldsymbol{i}^{\mathsf{c}},\boldsymbol{i}^{\mathsf{c}}}.$$

(4) Insgesamt ergibt sich

$$c_k = \sum_{\boldsymbol{i} \in I_k^d} (-1)^k \det A_{\boldsymbol{i}^{\mathsf{c}},\boldsymbol{i}^{\mathsf{c}}} = \sum_{\boldsymbol{i} \in I_{d-k}^d} (-1)^k \det A_{\boldsymbol{i},\boldsymbol{i}} = (-1)^k \operatorname{Tr}(\bigwedge^{d-k} \varphi)$$

nach Lemma IX.8.6. □

9. Tensor-, symmetrische- und äußere Algebra

In diesem Abschnitt bezeichne R stets einen kommutativen Ring mit Eins und M einen R-Modul.

Definition IX.9.1 (Unendliche Summe von R-Moduln): Sei I eine Menge und für jedes $i \in I$ sei M_i ein R-Modul. Dann heißt

$$\begin{aligned}\bigoplus_{i\in I} M_i &:= \{(m_i)_{i\in I} \mid m_i \in M_i \text{ für alle } i \in I \text{ und } m_i \neq 0 \text{ nur für endlich viele } i\} \\ &= \left\{ f \in \operatorname{Abb}_0\left(I, \bigcup_{i\in I} M_i\right) : f(i) \in M_i \right\}\end{aligned}$$

die *unendliche Summe der R-Moduln* $(M_i)_{i\in I}$. $\bigoplus_{i\in I} M_i$ wird mit komponentenweiser Addition und skalarer Multiplikation zu einem R-Modul, d. h. für $m = (m_i)_{i\in I}$ und $m' = (m'_i)_{i\in I} \in \bigoplus_{i\in I} M_i$ sowie $r \in R$ sind

$$m + m' := (m_i + m'_i)_{i\in I}, \qquad r \cdot m := (rm_i)_{i\in I}.$$

Im Folgenden schreiben wir auch $\sum_{i\in I} m_i$ für $(m_i)_{i\in I}$.

Definition IX.9.2 (Die Hauptakteure): Wir definieren die R-Moduln

(i) $T(M) := \bigoplus_{n\in\mathbb{N}} T^n(M)$,
(ii) $S(M) := \bigoplus_{n\in\mathbb{N}} S^n(M)$,
(iii) $\bigwedge(M) := \bigoplus_{n\in\mathbb{N}} \bigwedge^n(M)$.

Bemerkung IX.9.3 (Erzeuger): Als R-Modul werden

(i) $T(M)$ von $\{m_1 \otimes \cdots \otimes m_n \mid n \in \mathbb{N}, m_i \in M\}$,
(ii) $S(M)$ von $\{m_1 \odot \cdots \odot m_n \mid n \in \mathbb{N}, m_i \in M\}$ und
(iii) $\bigwedge(M)$ von $\{m_1 \wedge \cdots \wedge m_n \mid n \in \mathbb{N}, m_i \in M\}$

erzeugt.

Proposition IX.9.4 (gemischte Tensoren): *Für natürliche Zahlen k und ℓ erhalten wir eine bilineare Abbildung*

$$\begin{aligned}&h_{k,\ell}\colon T^k(M) \times T^\ell(M) \longrightarrow T^{k+\ell}(M), \\ &\quad h_{k,\ell}(m_1 \otimes \cdots \otimes m_k, m'_1 \otimes \cdots \otimes m'_\ell) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell.\end{aligned}$$

Beweis: Für den Beweis wollen wir die universelle Abbildungseigenschaft von $T^k(M)$ verwenden.

(1) Wir erhalten für jedes $m' = (m'_1, \ldots, m'_\ell) \in M^\ell$ eine lineare Abbildung $\varphi_{m'} \colon T^k(M) \to T^{k+\ell}(M)$ mit

$$\varphi_{m'}(m_1 \otimes \cdots \otimes m_k) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell$$

wie folgt: Die Abbildung

$$(m_1, \ldots, m_k) \longmapsto m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell$$

ist multilinear, d. h. wegen der universellen Abbildungseigenschaft von $T^k(M)$ erhalten wir die lineare Abbildung $\varphi_{m'} \colon T^k(M) \to T^{k+\ell}(M)$.

(2) Aus (1) erhalten wir die Abbildung

$$M^\ell \longrightarrow \mathrm{Hom}(T^k(M), T^{k+\ell}(M)), \qquad m' \longmapsto \varphi_{m'}.$$

Diese ist multilinear und definiert somit die folgende lineare Abbildung:

$$\varphi \colon T^\ell(M) \to \mathrm{Hom}(T^k(M), T^{k+\ell}(M)), \qquad \varphi(m'_1 \otimes \cdots \otimes m'_\ell) = \varphi_{m'}.$$

(3) Aus (2) erhalten wir die Abbildung

$$h_{k,\ell} \colon T^k(M) \times T^\ell(M) \longrightarrow T^{k+\ell}(M), \qquad (m, m') \longmapsto h_{k,\ell}(m, m') := \varphi_{m'}(m).$$

Diese ist linear in der ersten Komponente, da $\varphi_{m'}$ linear für jedes $m' \in T^\ell(M)$ ist, und lineare in der zweiten Komponente, da φ linear ist. □

Definition IX.9.5 (Multiplikation auf der Tensoralgebra): Wir definieren auf $T(M)$ eine Multiplikation „$\otimes$“ wie folgt: Für Elemente $m = \sum_{i \in I} m_i$ und $m' = \sum_{j \in J} m'_j \in T(M)$ setzen wir

$$m \otimes m' := \sum_{i \in I} \sum_{j \in J} h_{i,j}(m_i, m'_j),$$

insbesondere ist also für $m = m_1 \otimes \cdots \otimes m_k$ und $m'_1 \otimes \cdots \otimes m'_\ell \in T(M)$

$$m \otimes m' = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell.$$

Bemerkung IX.9.6: (i) $T(M)$ wird mit „$\otimes$“ zu einer R-Algebra.

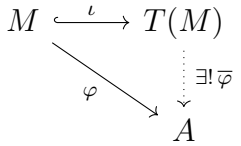
(ii) Auf analoge Art und Weise erhalten wir Multiplikationen „$\odot$“ auf $S(M)$ beziehungsweise „$\wedge$“ auf $\bigwedge(M)$, die diese zu R-Algebren machen.

Beweis: (i) Man rechnet nach, dass „$\otimes$“ assoziativ und distributiv bezüglich $+$ ist; offensichtlich ist „$\otimes$“ R-bilinear.

(ii) In Definition IX.9.5 müssen wir zusätzlich „symmetrisch“ respektive „alternierend“ fordern. $\square$

Definition IX.9.7 (Tensor-, symmetrische- und äußere Algebra): $T(M)$ beziehungsweise $S(M)$ beziehungsweise $\bigwedge(M)$ heißen mit der R-Algebrenstruktur aus (Bemerkung III.9.6) *Tensoralgebra* beziehungsweise *symmetrische Algebra* beziehungsweise *äußere Algebra.*

Satz 34 (Universelle Abbildungseigenschaft für $T(M)$): *Die Tensoralgebra $T(M)$ erfüllt die folgende universelle Abbildungseigenschaft: Ist A eine R-Algebra mit Eins und $\varphi\colon M \to A$ eine R-lineare Abbildung, dann gibt es genau einen R-Algebrenhomomorphismus $\overline{\varphi}\colon T(M) \to A$ mit $\overline{\varphi} \circ \iota = \varphi$ und $\overline{\varphi}(1) = 1$, d. h. das Diagramm*



ist kommutativ. Hierbei ist $\iota\colon M \hookrightarrow T(M)$, $m \mapsto (0, m, 0, \dots)$ die Einbettung.

Beweis: (1) Wir definieren $\overline{\varphi}^{(0)}\colon R = T^0(M) \to A$ durch $r \mapsto r1_A$ und für jedes $n \in \mathbb{N}$ definieren wir die R-lineare Abbildung $\overline{\varphi}^{(n)}\colon T^n(M) \to A$ wie folgt: Die Abbildung

$$M^n \longrightarrow A, \qquad (m_1, \dots, m_n) \longmapsto \varphi(m_1) \cdots \varphi(m_n)$$

ist multilinear und induziert die R-lineare Abbildung

$$\overline{\varphi}^{(n)}\colon T^n(M) \longrightarrow A, \qquad m_1 \otimes \cdots \otimes m_n \longmapsto \varphi(m_1) \cdots \varphi(m_n).$$

(2) Die Abbildung

$$\overline{\varphi}\colon T(M) \longrightarrow A, \qquad \sum_{i \in \mathbb{N}_0} m_i \longmapsto \sum_{i \in \mathbb{N}_0} \overline{\varphi}^i(m_i)$$

ist R-linear.

(3) Es bleibt zu zeigen, dass $\overline{\varphi}$ ein R-Algebrenhomomorphismus ist. Für $\sum_{i\in\mathbb{N}_0} m_i$ und $\sum_{j\in\mathbb{N}_0} m'_j$ aus $T(M)$ gilt:

$$\begin{aligned}\overline{\varphi}(m\otimes m') &= \overline{\varphi}\Big(\sum_{i\in\mathbb{N}_0}\sum_{j\in\mathbb{N}_0} m_i\otimes m'_j\Big)\\ &= \sum_{i\in\mathbb{N}_0}\sum_{j\in\mathbb{N}_0}\overline{\varphi}(m_i\otimes m'_j)\\ &= \sum_{i\in\mathbb{N}_0}\sum_{j\in\mathbb{N}_0}\overline{\varphi}(m_i)\cdot\overline{\varphi}(m'_j)\\ &= \Big(\sum_{i\in\mathbb{N}_0}\overline{\varphi}(m_i)\Big)\Big(\sum_{j\in\mathbb{N}_0}\overline{\varphi}(m'_j)\Big) = \overline{\varphi}(m)\cdot\overline{\varphi}(m').\end{aligned}$$

(4) Für Elemente $m_i \in T^i(M)$ und $m'_j \in T^j(M)$ können wir schreiben $m_i = \sum_k m_{k,1}\otimes\cdots\otimes m_{k,i}$ und $\sum_l m'_{l,1}\otimes\dots m'_{l,j}$ und berechnen

$$\begin{aligned}\overline{\varphi}(m_i\otimes m'_j) &= \overline{\varphi}^{(i+j)}\Big(\sum_k\sum_l m_{k,1}\otimes\cdots\otimes m_{k,i}\otimes m'_{l,1}\otimes\cdots\otimes m'_{l,j}\Big)\\ &= \sum_k\sum_l\overline{\varphi}^{(i+j)}(m_{k,1}\otimes\cdots\otimes m_{k,i}\otimes m'_{l,1}\otimes\cdots\otimes m'_{l,j})\\ &= \sum_k\sum_l\varphi(m_{k,1})\cdots\varphi(m_{k,i})\cdot\varphi(m'_{l,1})\cdots\varphi(m'_{l,j})\\ &= \overline{\varphi}(m_i)\cdot\overline{\varphi}(m'_j).\end{aligned}$$

Die Eindeutigkeit von $\overline{\varphi}$ folgt, da die Abbildung auf Erzeugern festgelegt ist. □

Satz 35 (Universelle Abbildungseigenschaft für $S(M)$, $\bigwedge(M)$): *Die symmetrische Algebra $S(M)$ beziehungsweise die äußere Algebra $\bigwedge(M)$ erfüllt folgende universelle Abbildungseigenschaft: Ist A eine R-Algebra mit Eins und $\varphi\colon M\to A$ ein R-Modulhomomorphismus mit $\varphi(x)\varphi(y)=\varphi(y)\varphi(x)$ beziehungsweise $\varphi(x)\varphi(x)=\mathbf{0}$ für alle $x,y\in M$, dann gibt es genau einen R-Algebrenhomomorphismus $\overline{\varphi}\colon S(M)\to A$ beziehungsweise $\overline{\varphi}\colon\bigwedge(M)\to A$ mit $\overline{\varphi}\circ\iota=\varphi$ und $\overline{\varphi}(1)=1$, wobei $\iota\colon M\hookrightarrow S(M)$ beziehungsweise $\iota\colon M\hookrightarrow\bigwedge(M)$ die natürliche Einbettung $m\mapsto(0,m,0,\dots)$ ist.*

Beweis: Analog zum Beweis von Satz 34. □

Bemerkung IX.9.8 (Funktorialität von T, S und $\bigwedge$): Seien M_1 und M_2 zwei R-Moduln und $f\colon M_1\to M_2$ eine R-lineare Abbildung. Dann gibt es eindeutige Algebrenhomomorphismen

$$T(f)\colon T(M_1)\longrightarrow T(M_2),$$
$$S(f)\colon S(M_1)\longrightarrow S(M_2),\qquad \bigwedge(f)\colon\bigwedge(M_1)\longrightarrow\bigwedge(M_2)$$

mit

$$T(f)(m_1 \otimes \cdots \otimes m_k) = f(m_1) \otimes \cdots \otimes f(m_k),$$
$$S(f)(m_1 \odot \cdots \odot m_k) = f(m_1) \odot \cdots \odot f(m_k),$$
$$\bigwedge(f)(m_1 \wedge \cdots \wedge m_k) = f(m_1) \wedge \cdots \wedge f(m_k).$$

Es gilt für zwei R-lineare Abbildungen $f_1 \colon M_1 \to M_2$ und $f_2 \colon M_2 \to M_3$, dass

$$T(f_2 \circ f_1) = T(f_2) \circ T(f_1),$$
$$S(f_2 \circ f_1) = S(f_2) \circ S(f_1), \qquad \bigwedge(f_2 \circ f_1) = \bigwedge(f_2) \circ \bigwedge(f_1).$$

Beweis: Folgt aus Satz 34 und Satz 35. □

10. Die symmetrische Algebra und der Polynomring

In diesem Abschnitt sei R stets ein kommutativer Ring mit Eins.

Definition IX.10.1 (Polynomringe über Ringen):

(i) $R[X] := \{\sum_{i=0}^{k} a_i X^i \mid k \in \mathbb{N}_0, a_0, \ldots, a_k \in R\}$ heißt *Polynomring über* R,

(ii) Für $n \geq 2$ wird $R[X_1, \ldots, X_n]$ iterativ definiert durch

$$R[X_1, \ldots, X_n] := (R[X_1, \ldots, X_{n-1}])[X_n]$$

und heißt *Polynomring in* n *Unbestimmten*,

(iii) Für $n = 0$ definieren wir $R[X_1, \ldots, X_n] := R$.

Bemerkung IX.10.2 (Algebrenstruktur von Polynomringen): Die Addition, die Multiplikation und die skalare Multiplikation von Polynomen in $R[X]$ wird genau wie in $K[X]$ definiert. $R[X]$ wird damit zur kommutativen R-Algebra mit Eins.

Beispiel IX.10.3: (i) In $R[X_1, X_2, X_3]$, $f = 2X_1X_2X_3 + X_1^2 + X_2X_3^2$ und $g = X_1X_3 - X_2$ ist

$$\begin{aligned} f \cdot g &= (X_1^2 + 2X_1X_2X_3 + X_2X_3^2) \cdot (-X_2 + X_1X_3) \\ &= -X_1^2X_2 + (X_1^3 - 2X_1X_2^2)X_3 + (2X_1X_2X_1 - X_2^2)X_3^2 + X_2X_1X_3^3 \\ &= -X_1^2X_2 + (X_1^3 - 2X_1X_2^2)X_3 + (2X_1^2X_2 - X_2^2)X_3^2 + X_2X_1X_3^3 \end{aligned}$$

(ii) Sind $R = \mathbb{Z}/6\mathbb{Z}$, $f = 2X^2 + 1$ und $g = 3X - 1 \in R[X]$, so ist

$$f \cdot g = -2X^2 + 3X - 1,$$

insbesondere gilt $\deg(f \cdot g) < \deg(f) + \deg(g)$.

Bemerkung IX.10.4 (Ring- vs. Algebrenhomomorphismen): Seien A_1 und A_2 zwei R-Algebren mit Eins und $\varphi\colon A_1 \to A_2$ ein Ringhomomorphismus. φ ist ein Homomorphismus von R-Algebren mit Eins genau dann, wenn für alle $r \in R$ gilt $\varphi(r1_{A_1}) = r1_{A_2}$.

Proposition IX.10.5 (Universelle Eigenschaft von $R[X]$ als Algebra): *Ist A eine kommutative R-Algebra mit Eins und sind $a_1, \ldots, a_n \in A$, dann gibt es genau einen Homomorphismus von R-Algebren mit Eins $\varphi\colon R[X_1, \ldots, X_n] \to A$ mit $\varphi(X_1) = a_1, \ldots, \varphi(X_n) = a_n$.*

Proposition IX.10.6 (Universelle Eigenschaft von $R[X]$ als Ring): *Für den Polynomring $R[X]$ gilt: Sind R' ein kommutativer Ring mit Eins und $\overline{\varphi}\colon R \to R'$ ein Homomorphismus von Ringen mit Eins und $r' \in R'$, dann gibt es genau einen Ringhomomorphismus $\varphi\colon R[X] \to R'$ mit $\varphi|_R = \overline{\varphi}$ und $\varphi(X) = r'$.*

Beweis: Für $f = \sum_{i=0}^{n} a_i X^i$ definieren wir

$$\varphi\Big(\sum_{i=0}^{n} a_i X^i\Big) := \sum_{i=0}^{n} \tilde{\varphi}(a_i) r'^i;$$

leicht kann man sich davon überzeugen, dass das einen Ringhomomorphismus erklärt. Weiter sind $\varphi(X) = \tilde{\varphi}(1)r' = r'$ und $\varphi(r) = \varphi(rX^0) = \tilde{\varphi}(r)$ für alle $r \in R$. Damit ist φ eindeutig festgelegt, da es auf den Erzeugern $1, X$ festgelegt ist. □

Beweis (von Proposition IX.10.5): Zum Beweis dieser Aussage verwenden wir Proposition IX.10.6 und Induktion nach n.

Für den Induktionsanfang $n = 0$ tut es $\varphi\colon R \to A$, $r \mapsto r1_A$.

Zum Induktionsschritt von n nach $n+1$: Per Induktionsvoraussetzung gibt es einen R-Algebrenhomomorphismus $\tilde{\varphi}\colon R[X_1, \ldots, X_n] \to A$ mit $\tilde{\varphi}(1_R) = 1_A$ und $\tilde{\varphi}(X_i) = a_i$ für $1 \le i \le n$, insbesondere $\tilde{\varphi}(r1_R) = r1_A$. Nach Proposition IX.10.6 gibt es genau einen Ringhomomorphismus

$$\varphi\colon R[X_1, \ldots, X_{n+1}] = R[X_1, \ldots, X_n][X_{n+1}] \longrightarrow A$$

mit $\varphi|_{R[X_1,\ldots,X_n]} = \tilde{\varphi}$ und $\varphi(X_{n+1}) = a_{n+1}$. Dieser Ringhomomorphismus φ ist nach Bemerkung IX.10.4 sogar ein R-Algebrenhomomorphismus. Die Eindeutigkeit folgt, da φ auf Erzeugern von $R[X_1, \ldots, X_{n+1}]$ festgelegt ist. □

Satz 36 (Symmetrische Algebra als Polynomring): *Sei M ein freier R-Modul mit Basis $\{b_1, \ldots, b_d\}$. Dann gilt $S(M) \cong R[X_1, \ldots, X_d]$.*

Beweis: (1) Da $S(M)$ ein kommutativer Ring mit Eins ist, gibt es nach Proposition IX.10.5 genau einen Ringhomomorphismus

$$\varphi\colon R[X_1, \ldots, X_d] \longrightarrow S(M)$$

mit $\varphi(X_i) = b_i$ für $1 \leq i \leq d$ und $\varphi(1) = 1$.

(2) Da M ein freier Modul mit Basis $\{b_1, \ldots, b_d\}$ ist, gibt es nach Proposition IX.5.5 einen R-Modulhomomorphismus

$$\psi_1\colon M \longrightarrow R[X_1, \ldots, X_d]$$

mit $\psi_1(b_i) = X_i$ für $1 \leq i \leq d$ und $\psi_1(1) = 1$. Satz 35 liefert uns jetzt einen R-Algebrenhomomorphismus $\Psi\colon S(M) \to R[X_1, \ldots, X_d]$ mit $\psi|_M = \psi_1$, insbesondere also $\psi(b_i) = X_i$ für $1 \leq i \leq d$ und $\psi(1) = 1$.

(3) Für die Verkettungen $\psi \circ \varphi\colon R[X_1, \ldots, X_d] \to R[X_1, \ldots, X_d]$ beziehungsweise $\varphi \circ \psi\colon S(M) \to S(M)$ gilt: $\psi \circ \varphi$ beziehungsweise $\varphi \circ \psi$ sind Homomorphismen von R-Algebren mit Eins und $(\psi \circ \varphi)(X_i) = X_i$ sowie $(\varphi \circ \psi)(b_i) = b_i$ für $1 \leq i \leq d$. Die Eindeutigkeit in Satz 35 beziehungsweise Proposition IX.10.5 liefert jetzt $\psi \circ \varphi = \mathrm{id}_{R[X_1,\ldots,X_d]}$ und $\varphi \circ \psi = \mathrm{id}_{S(M)}$. □

Bemerkung IX.10.7: Für φ und ψ aus dem Beweis von Satz 36 gelten:

$$\varphi(X_{i_1} \cdots X_{i_k}) = b_{i_1} \odot \cdots \odot b_{i_k}, \qquad \psi(b_{i_1} \odot \cdots \odot b_{i_k}) = X_{i_1} \cdots X_{i_k}.$$

Kapitel X.

Kategorientheorie

In diesem Abschnitt wollen wir die Struktur von Gruppen, Ringen und Körpern verallgemeinern und universelle Abbildungseigenschaften beschreiben. Die Hauptakteure werden Objekte und ihre Morphismen sein.

1. Kategorien

Definition X.1.1: Eine *Kategorie* $\mathcal{C}$ besteht aus einer Klasse von Objekten $\mathrm{Obj}(\mathcal{C})$ und zu je zwei Objekten $A, B \in \mathrm{Obj}(\mathcal{C})$ aus einer Menge $\mathrm{Mor}_{\mathcal{C}}(A, B)$, für die folgende Eigenschaften erfüllt sind:

(i) Für $A \in \mathrm{Obj}(\mathcal{C})$ gibt es $\mathrm{id}_A \in \mathrm{Mor}_{\mathcal{C}}(A, A)$,

(ii) Für $A, B, C \in \mathrm{Obj}(\mathcal{C})$ gibt es eine Abbildung

$$\circ\colon \mathrm{Mor}_{\mathcal{C}}(B, C) \times \mathrm{Mor}_{\mathcal{C}}(A, B) \longrightarrow \mathrm{Mor}_{\mathcal{C}}(A, C), \qquad (g, f) \longmapsto g \circ f$$

mit $f \circ \mathrm{id}_A = f$ beziehungsweise $\mathrm{id}_B \circ f = f$ für alle $f \in \mathrm{Mor}(B, C)$ beziehungsweise für alle $f \in \mathrm{Mor}(A, B)$,

(iii) Für $A, B, C, D \in \mathrm{Obj}(\mathcal{C})$ und alle $f \in \mathrm{Mor}_{\mathcal{C}}(A, B)$, $g \in \mathrm{Mor}_{\mathcal{C}}(B, C)$ und $h \in \mathrm{Mor}_{\mathcal{C}}(C, D)$ gilt

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

Die Elemente aus $\mathrm{Mor}_{\mathcal{C}}(A, B)$ heißen *Morphismen von A nach B*, wir schreiben auch kurz $\mathrm{Mor}(A, B) := \mathrm{Mor}_{\mathcal{C}}(A, B)$, wenn klar ist, von welcher Kategorie die Rede ist. Für $f \in \mathrm{Mor}(A, B)$ schreiben wir kurz $f\colon A \to B$.

Beispiel X.1.2: (i) Mit Set bezeichnet man die Kategorie, deren Objekte die Mengen und deren Morphismen die Abbildungen zwischen Mengen sind,

(ii) Mit Gr bezeichnet man die Kategorie, deren Objekte die Gruppen und deren Morphismen die Gruppenhomomorphismen sind,

(iii) Mit $\mathbb{K}$-Vr bezeichnet man die Kategorie, deren Objekte die $\mathbb{K}$-Vektorräume und deren Morphismen die linearen Abbildungen sind,

(iv) Mit Met bezeichnet man die Kategorie, deren Objekte die metrischen Räume und deren Morphismen die stetigen Abbildungen sind,

(v) Mit Top bezeichnet man die Kategorie, deren Objekte die topologischen Räume und deren Morphismen die stetigen Abbildungen sind,

(vi) Mit kRi bezeichnet man die Kategorie, deren Objekte die kommutativen Ringe mit Eins und deren Morphismen die Homomorphismen von Ringen mit Eins sind,

(vii) Mit R-Mod bezeichnet man die Kategorie, deren Objekte die R-Moduln und deren Morphismen die linearen Abbildungen sind,

(viii) Mit R-Alg bezeichnet man die Kategorie, deren Objekte die R-Algebren und deren Morphismen die Homomorphismen von Algebren sind,

(ix) Mit Fields bezeichnet man die Kategorie, deren Objekte Körper und deren Morphismen die Körperhomomorphismen sind.

In (iii) bezeichnet $\mathbb{K}$ den Körper der reellen- oder komplexen Zahlen und in (vi), (vii) bezeichnet R einen kommutativen Ring mit Eins. Offensichtlich handelt es sich in allen Fällen um Kategorien.

Definition X.1.3: Seien $A, B \in \mathrm{Obj}(\mathcal{C})$. Ein Morphismus $f \in \mathrm{Mor}(A, B)$ heißt *Isomorphismus*, falls es $g \in \mathrm{Mor}(B, A)$ mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$ gibt.

Bemerkung X.1.4: In den Beispielen (i)-(iii) und (vi)-(ix) ist $f \in \mathrm{Mor}(A, B)$ ein Isomorphismus genau dann, wenn f ein bijektiver Morphismus ist. In (iv) und (v) gilt das nicht!

Lemma X.1.5: *Für einen kommutativen Ring mit Eins definieren wir*

$$\mathrm{Gl}_n(R) := \{A \in R^{n\times n} \mid \textit{Es gibt } B \in R^{n\times n} \textit{ mit } AB = BA = I_n\},$$

wobei I_n die $n \times n$-Einheitsmatrix bezeichnet. Ist $f\colon R_1 \to R_2$ ein Ringhomomorphismus und ist $A = (a_{i,j})_{i,j} \in \mathrm{Gl}_n(R)$, so ist ${}_fA := (f(a_{i,j}))_{i,j} \in \mathrm{Gl}_n(R)$.

Beweis: Sei $B = (b_{i,j})_{i,j} \in \mathrm{Gl}_n(R)$ mit $AB = BA = I_n$. Für alle $1 \leq i, k \leq n$ gilt dann $\sum_{j=1}^{n} a_{i,j}b_{j,k} = \delta_{i,k}$, und da f ein Ringhomomorphismus ist, gilt

$$\sum_{j=1}^{k} f(a_{i,j})f(b_{j,k}) = \sum_{j=1}^{n} f(a_{i,j} \cdot b_{j,k}) = f\Big(\sum_{j=1}^{n} a_{i,j} \cdot b_{j,k}\Big) = \delta_{i,k},$$

d. h. ${}_fA \cdot_f B = I_n$. Mit der gleichen Rechnung überzeugen wir uns davon, dass ${}_fB \cdot_f A = I_n$, also ist ${}_fA \in \mathrm{Gl}_n(R)$. □

Bemerkung X.1.6: Wir haben eine Zuordnung

$$\mathrm{Obj}(\mathsf{kRi}) \longrightarrow \mathrm{Obj}(\mathsf{Gr}), \qquad R \longmapsto \mathrm{Gl}_n(R),$$

welche wir nach Lemma X.1.5 auf die Morphismen erweitern können:

$$(f\colon R_1 \to R_2) \longmapsto \left\{ \mathrm{Gl}_n(R_1) \to \mathrm{Gl}_n(R_2), \quad A \mapsto_f A \right\}$$

Als Übung bleibt dem Leser überlassen nachzurechnen, dass $A \mapsto_f A$ ein Homomorphismus von Gruppen ist. Außerdem ist diese Erweiterung verträglich mit der Verknüpfung von Morphismen: Für $f\colon R_1 \to R_2$ und $g\colon R_2 \to R_3$ sind ${}_g({}_fA) =_{g\circ f} A$ und ${}_{\mathrm{id}}A = A$.

Definition X.1.7: Seien $\mathcal{C}_1$ und $\mathcal{C}_2$ Kategorien.

(i) Ein *kovarianter Funktor* $F\colon \mathcal{C}_1 \to \mathcal{C}_2$ besteht aus

- Einer Abbildung $\mathrm{Obj}(\mathcal{C}_1) \to \mathrm{Obj}(\mathcal{C}_2)$, $A \mapsto F(A)$,
- Für $A, B \in \mathrm{Obj}(\mathcal{C}_1)$ einer Abbildung

$$F = F_{A,B}\colon \mathrm{Mor}_{\mathcal{C}_1}(A,B) \longrightarrow \mathrm{Mor}_{\mathcal{C}_2}(F(A),F(B)),$$

sodass für alle $A, B, C \in \mathrm{Obj}(\mathcal{C}_1)$ und $f \in \mathrm{Mor}(A,B)$, $g \in \mathrm{Mor}(B,C)$ gilt
(1) $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$,
(2) $F(g \circ f) = F(g) \circ F(f)$.

(ii) Die Definition eines *kontravarianten Funktors* ist analog zu der eines kovarianten Funktors, die Abbildungen zwischen den Morphismen unterscheiden sich wie folgt:

$$F = F_{A,B}\colon \mathrm{Mor}_{\mathcal{C}_1}(A,B) \longrightarrow \mathrm{Mor}_{\mathcal{C}_2}(F(B),F(A)$$

und in (2) gilt $F(g \circ f) = F(f) \circ F(g)$.

Beispiel X.1.8 (Einige Funktoren):

(i) **Der Funktor** Gl_n**:** Sei n eine natürliche Zahl. Dann ist die Zuordnung

$$\begin{aligned} \mathrm{Gl}_n\colon \mathsf{kRi} &\longrightarrow \mathsf{Gr} \\ R &\longmapsto \mathrm{Gl}_n(R) && \text{(für Objekte)} \\ (f\colon R_1 \to R_2) &\longmapsto \begin{pmatrix} \mathrm{Gl}_n(R_1) \to \mathrm{Gl}_n(R_2) \\ A \mapsto {}_fA \end{pmatrix} && \text{(für Morphismen)} \end{aligned}$$

ein Funktor.

(ii) **Wedge-Funktor $\bigwedge^n$:** Seien n eine natürliche Zahl und R ein Objekt in kRi. Dann ist die Zuordnung

$$\begin{aligned} \bigwedge^n \colon R\text{-}\mathsf{Mod} &\longrightarrow R\text{-}\mathsf{Mod} \\ M &\longmapsto \bigwedge^n(M) \quad \text{(für Objekte)} \\ f &\longmapsto \bigwedge^n(f) \quad \text{(für Morphismen)} \end{aligned}$$

ein Funktor.

(iii) **n-Tupel Funktor A_n:** Seien n eine natürliche Zahl und $[n] = \{1, \dots, n\}$. Die Zuordnung

$$\begin{aligned} A_n \colon \mathbb{R}\text{-}\mathsf{VR} &\longrightarrow \mathsf{Set} \\ V &\longmapsto \mathrm{Abb}([n], V) && \text{(für Objekte)} \\ (f \colon V \to W) &\longmapsto \begin{pmatrix} \mathrm{Abb}([n], V) \to \mathrm{Abb}([n], W) \\ t \mapsto f \circ t \end{pmatrix} && \text{(für Morphismen)} \end{aligned}$$

ist ein Funktor.

(iv) **Funktor multiplikative Gruppe:** Die Zuordnung

$$\begin{aligned} G_m \colon \mathsf{kRi} &\longrightarrow \mathsf{Gr} \\ R &\longmapsto R^\times = \{r \in R \mid \text{Es gibt } s \in R \text{ mit } sr = rs = 1\} \\ (f \colon R_1 \to R_2) &\longmapsto (f|_{R^\times} \colon R_1^\times \to R_2^\times) \end{aligned}$$

ist ein Funktor.

(v) **Vergiss-Funktor:** Die Zuordnung

$$\begin{aligned} V \colon \mathsf{Gr} &\longrightarrow \mathsf{Set} \\ G &\longmapsto G \\ (f \colon G_1 \to G_2) &\longmapsto (f \colon G_1 \to G_2) \end{aligned}$$

ist ein Funktor.

(vi) **Dualisierungsfunktor:** Die Zuordnung

$$\begin{aligned} D \colon K\text{-}\mathsf{Vr}^{\mathrm{fin}} &\longrightarrow K\text{-}\mathsf{Vr}^{\mathrm{fin}} \\ V &\longmapsto V^* \\ (\varphi \colon V_1 \to V_2) &\longmapsto (\varphi^* \colon V_2^* \to V_1^*, \quad g \mapsto g \circ \varphi) \end{aligned}$$

ist ein Funktor (siehe Übungsblatt).

Die Funktoren aus (i) - (v) sind kovariante Funktoren, der Funktor aus (vi) ist kontravariant.

Definition X.1.9: Seien $\mathcal{C}$ eine Kategorie und A_0 ein Objekt in $\mathcal{C}$. Dann sind die Zuordnungen

$$\begin{aligned} \hom_{A_0}\colon \mathcal{C} &\longrightarrow \mathsf{Set} \\ C &\longmapsto \mathrm{Mor}_{\mathcal{C}}(A_0, C) && \text{(für Objekte)} \\ (f\colon C_1 \to C_2) &\longmapsto \begin{pmatrix} \mathrm{Mor}_{\mathcal{C}}(A_0, C_1) \to \mathrm{Mor}_{\mathcal{C}}(A_0, C_2) \\ h \mapsto f \circ h \end{pmatrix} && \text{(für Morphismen)} \end{aligned}$$

sowie

$$\begin{aligned} {}_{A_0}\hom\colon \mathcal{C} &\longrightarrow \mathsf{Set} \\ C &\longmapsto \mathrm{Mor}_{\mathcal{C}}(C, A_0) && \text{(für Objekte)} \\ (f\colon C_1 \to C_2) &\longmapsto \begin{pmatrix} \mathrm{Mor}_{\mathcal{C}}(C_2, A_0) \to \mathrm{Mor}_{\mathcal{C}}(C_1, A_0) \\ h \mapsto h \circ f \end{pmatrix} && \text{(für Morphismen)} \end{aligned}$$

Funktoren. Der Funktor $\hom_{A_0}$ ist kovariant, der Funktor ${}_{A_0}\hom$ ist kontravariant.

Beweis: Wir zeigen, dass $\hom_{A_0}$ ein Funktor ist, für ${}_{A_0}\hom$ funktioniert der Beweis analog. Zunächst ist

$$\hom_{A_0}(\mathrm{id}_C) = \begin{pmatrix} \mathrm{Mor}_{\mathcal{C}}(C, A_0) \longrightarrow \mathrm{Mor}_{\mathcal{C}}(A_0, C) \\ h \longmapsto \mathrm{id} \circ h = h \end{pmatrix} = \mathrm{id}_{\mathrm{Mor}_{\mathcal{C}}(A_0, C)},$$

weiter gilt für $f \in \mathrm{Mor}_{\mathcal{C}}(A, B)$, $g \in \mathrm{Mor}_{\mathcal{C}}(B, C)$ und $g \circ f \in \mathrm{Mor}_{\mathcal{C}}(A, B)$, dass

$$\begin{aligned} \hom_{A_0}(g \circ f) &= (h \longmapsto g \circ f \circ h), \\ \hom_{A_0}(g) \circ \hom_{A_0}(f) &= (h \longmapsto f \circ h \longmapsto g \circ f \circ h), \end{aligned}$$

d. h. $\hom_{A_0}(g) \circ \hom_{A_0}(f) = \hom_{A_0}(g \circ f)$. □

2. Universelle Objekte

In diesem Abschnitt wollen wir „universelle Abbildungseigenschaften“ im Kontext von Kategorientheorie beschreiben. Dazu führen wir „Morphismen“ zwischen Funktoren, sogenannte *natürliche Transformationen*, ein.

Beispiel X.2.1: Für einen kommutativen Ring mit Eins R haben wir eine Abbildung

$$\det_R\colon \mathrm{Gl}_n(R) \longrightarrow R^\times, \qquad A \longmapsto \det A.$$

Die Abbildung $\det_R$ ist verträglich mit Ringhomomorphismen wie folgt: Für einen Ringhomomorphismus $f\colon R_1 \to R_2$ gilt

$$\begin{aligned}\det({}^f A) = \det(f(a_{i,j})) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) f(a_{1,\sigma(1)}) \cdots f(a_{n,\sigma(n)}) \\ &= f\Big(\sum_{\sigma\in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}\Big) = f(\det(A)),\end{aligned}$$

d. h. wir haben das folgende kommutative Diagramm:

$$\begin{array}{ccc} \mathrm{Gl}_n(R_1) & \xrightarrow{\mathrm{Gl}_n(f)} & \mathrm{Gl}_n(R_2) \\ \Big\downarrow{\scriptstyle \det_{R_1}} & & \Big\downarrow{\scriptstyle \det_{R_2}} \\ G_m(R_1) & \xrightarrow[G_m(f)]{} & G_m(R_2) \end{array}$$

Definition X.2.2 (natürliche Transformation): Seien $F, G\colon \mathcal{C}_1 \to \mathcal{C}_2$ zwei kovariante Funktoren. Eine *natürliche Transformation* $\alpha\colon F \to G$ ist eine Familie von Morphismen $\{\alpha_C\colon F(C) \to G(C) \mid C \in \mathcal{C}_1\}$ in die Kategorie $\mathcal{C}_2$, für die gilt: Für je zwei Objekte A, B in $\mathcal{C}_1$ und einen Morphismus $f \in \mathrm{Mor}_{\mathcal{C}_1}(A, B)$ kommutiert das folgende Diagramm:

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ \Big\downarrow{\scriptstyle \alpha_A} & & \Big\downarrow{\scriptstyle \alpha_B} \\ G(A) & \xrightarrow[G(f)]{} & G(B) \end{array}$$

Sind die Morphismen in der Familie $\{\alpha_C\colon F(C) \to G(C) \mid C \in \mathcal{C}_1\}$ alles Isomorphismen, so heißt die natürliche Transformation α eine *funktorielle Äquivalenz*.

Beispiel X.2.3 (Determinante als natürliche Transformation): Die Determinante $\det\colon \mathrm{Gl}_n \to G_m$ ist eine natürliche Transformation (vergleiche Beispiel X.2.1).

Im Folgenden bezeichne K-$\mathsf{VR}^{\mathrm{fin}}$ die Kategorie der endlichdimensionalen K-Vektorräume.

Beispiel X.2.4 (Der Bidualfunktor): Sei $D\colon K\text{-}\mathsf{VR}^{\mathrm{fin}} \to K\text{-}\mathsf{VR}^{\mathrm{fin}}$ der Dualisierungsfunktor aus Beispiel X.1.8. Die Komposition $D \circ D\colon K\text{-}\mathsf{VR}^{\mathrm{fin}} \to K\text{-}\mathsf{VR}^{\mathrm{fin}}$ ist ebenfalls ein Funktor (siehe Blatt 11, Aufgabe 1) mit $(D \circ D)(V) = V^{**}$. Wir erhalten eine funktorielle Äquivalenz $\alpha\colon \mathrm{id} \to D \circ D$ wie folgt:

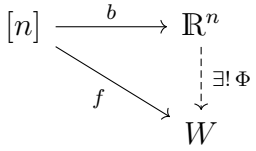
$$\alpha_V\colon V \longrightarrow V^{**}, \qquad v \longmapsto \varphi_v : \begin{pmatrix} V^* \to K \\ \beta \mapsto \beta(v) \end{pmatrix},$$

siehe Blatt 12.

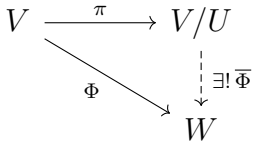
Erinnerung: (i) Satz 12, Lineare Algebra I (Fortsetzungssatz): Es bezeichne e_i den i-ten kanonischen Basisvektor von $\mathbb{R}^n$ und

$$b\colon [n] \longrightarrow \mathbb{R}^n, \qquad i \longmapsto e_i.$$

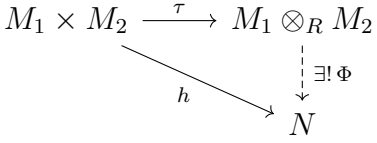
Für alle Abbildungen $f\colon [n] \to W$ in einen $\mathbb{R}$-Vektorraum W gibt es genau eine lineare Abbildung $\Phi\colon \mathbb{R}^n \to W$ mit $\Phi \circ b = f$, d. h. wir haben das kommutative Diagramm



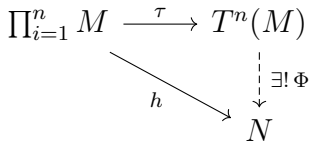
(ii) Satz 10, Lineare Algebra I (Homomorphiesatz): Es seien V ein K-Vektorraum, U ein Untervektorraum von V und $\pi\colon V \to V/U$ die kanonische Projektion. Für alle K-Vektorräume W und $\Phi \in \mathrm{Hom}(V, W)$ mit $U \subseteq \mathrm{Kern}(\Phi)$ gibt es genau ein $\overline{\Phi} \in \mathrm{Hom}(V/U, W)$ mit $\overline{\Phi} \circ \pi = \Phi$, d. h. wir haben das kommutative Diagramm



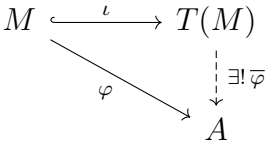
(iii) Proposition IX.4.10, Lineare Algebra II (Tensorprodukt): Seien M_1 und M_2 zwei R-Moduln und $\tau\colon M_1 \times M_2 \to M_1 \otimes M_2$ die Tensorabbildung. Für alle R-Moduln N und alle $h \in \mathrm{Mult}(M_1 \otimes M_2, N)$ gibt es genau ein $\Phi\colon \mathrm{Hom}(M_1 \otimes_R M_2, N)$ mit $\Phi \circ \tau = h$, d. h. wir haben das kommutative Diagramm



(iv) Satz 31, Lineare Algebra II (Tensorpotenz): Seien M ein R-Modul, n eine natürliche Zahl und $\tau\colon M^n \to T^n(M)$ die natürliche Abbildung in die n-te Tensorpotenz. Für alle R-Moduln N und alle $h \in \mathrm{Mult}^n_M(N)$ gibt es genau ein $\Phi \in \mathrm{Hom}(T^n(M), N)$ mit $\Phi \circ \tau = h$ (analog für $S^n(M)$ und $\bigwedge^n(M)$), d. h. wir haben das kommutative Diagramm



(v) Satz 34, Lineare Algebra II (Tensoralgebra): Seien M ein R-Modul, $T(M)$ die zugehörige Tensoralgebra und $\iota\colon M \hookrightarrow T(M)$ die natürliche Einbettung. Für alle R-Algebren A mit Eins und $\varphi \in \mathrm{Hom}_{R\text{-Mod}}(M, A)$ gibt es genau ein $\overline{\varphi} \in \mathrm{Hom}_{R\text{-Alg}}(T(M), A)$ mit $\overline{\varphi} \circ \iota = \varphi$, d. h. wir haben das kommutative Diagramm



Definition X.2.5 (Universelles Objekt): Seien $F\colon \mathcal{C} \to \mathsf{Set}$ ein kovarianter Funktor und C ein Objekt von $\mathcal{C}$. F heißt *darstellbar mit darstellendem Objekt* $\mathcal{C}$, falls es eine funktorielle Äquivalenz $\alpha\colon F \to \mathrm{hom}_C$ gibt. C heißt dann auch *universelles Objekt*.

Satz 37 (Universelle Abbildungseigenschaft des Fortsetzungssatzes): *Sei*

$$\begin{aligned} F\colon \mathbb{R}\text{-}\mathsf{VR} &\longrightarrow \mathsf{Set}, \\ V &\longmapsto \mathrm{Abb}([n], V), \\ (V_1 \xrightarrow{\alpha} V_2) &\longmapsto \begin{pmatrix} \mathrm{Abb}([n], V_1) \to \mathrm{Abb}([n], V_2) \\ f \mapsto \varphi \circ f \end{pmatrix} \end{aligned}$$

Dann ist F darstellbar mit universellem Objekt $\mathbb{R}^n$.

Beweis: Wir definieren die natürliche Transformation $\alpha\colon F \to \mathrm{hom}_{\mathbb{R}^n}$ wie folgt: Für ein Objekt W in $\mathbb{R}$-VR sei

$$\alpha_W\colon F(W) = \mathrm{Abb}([n], W) \longrightarrow \mathrm{hom}_{\mathbb{R}^n}(W) = \mathrm{Hom}(\mathbb{R}^n, W), \qquad f \longmapsto \Phi = \Phi_f$$

wobei Φ_f die eindeutige Fortsetzung von f auf $\mathbb{R}^n$ ist.

(1) Das so definierte α ist eine natürliche Transformation: Für $\mathbb{R}$-Vektorräume W_1, W_2 und $\varphi\colon W_1 \to W_2$ haben wir das Diagramm

$$\begin{array}{ccc} F(W_1) = \mathrm{Abb}([n], W_1) & \xrightarrow{F(\varphi)} & F(W_2) = \mathrm{Abb}([n], W_2) \\ \downarrow{\scriptstyle \alpha_{W_1}} & & \downarrow{\scriptstyle \alpha_{W_2}} \\ \hom_{\mathbb{R}^n}(W_1) = \mathrm{Hom}(\mathbb{R}^n, W_1) & \xrightarrow[\hom_{\mathbb{R}^n}(\varphi)]{} & \hom_{\mathbb{R}^n}(W_2) = \mathrm{Hom}(\mathbb{R}^n, W_2) \end{array}$$

Bleibt zu zeigen, dass $\varphi \circ \Phi_f = \Phi_{\varphi \circ f}$. Für den i-ten Basisvektor e_i gilt

$$\Phi_{\varphi\circ f}(e_i) = (\varphi \circ f)(e_i) = \varphi(f(i)) = \varphi(\Phi_f(e_i)) = (\varphi \circ \Phi_f)(e_i),$$

d. h. $\Phi_{\varphi\circ f} = \varphi \circ \Phi_f$ nach der Eindeutigkeit der linearen Fortsetzung; das obige Diagramm kommutiert also.

(2) Das so definierte α ist eine funktorielle Äquivalenz, denn: α_V ist nach dem Fortsetzungssatz ein Isomorphismus (d. h. eine bijektive Abbildung) mit Umkehrabbildung

$$\mathrm{Hom}(\mathbb{R}^n, W) \longrightarrow \mathrm{Abb}([n], W), \qquad \Phi \longmapsto \Phi \circ b. \qquad \square$$

Zu den anderen universellen Abbildungseigenschaften gibt es ähnliche Funktoren, die darstellbar sind. Darauf wird in den Übungen weiter eingegangen werden.

Kapitel XI.

Etwas mehr Strukturmathematik

1. Gruppenaktionen

$\mathrm{Sym}(X)$ „operiert“ auf X, $\mathrm{Gl}_n(K)$ „operiert“ auf K^n und $\mathrm{Aut}(V)$ „operiert“ auf V. Wir wollen das im Folgenden verallgemeinern zu Operationen von Gruppen auf Mengen.

Definition XI.1.1: Seien $(G, \star)$ eine Gruppe mit neutralem Element 1_G und X eine Menge. Eine Abbildung

$$\alpha\colon G \times X \longrightarrow X$$

heißt *Gruppenaktion von G auf X*, falls gelten:

(i) Für alle $x \in X$ gilt $\alpha(1_G, x) = x$,

(ii) Für alle $g_1, g_2 \in G$ und $x \in X$ ist $\alpha(g_1 \star g_2, x) = \alpha(g_1, \alpha(g_2, x))$.

Im Folgenden schreiben wir $g \cdot x := \alpha(g, x)$.

Bemerkung XI.1.2: In der Situation von Definition XI.1.1 gilt für alle $x, y \in X$ und $g \in G$: Ist $gx = y$, so ist $x = g^{-1}y$.

Das sieht man so: Es ist $g^{-1}g = g^{-1} \cdot (gx) = (g^{-1} \star g) \cdot x = (1_G) \cdot x = x$.

Beispiel XI.1.3 (Erste Beispiele von Gruppenaktionen): (i) Sei K ein Körper und n eine natürliche Zahl. Dann operiert $\mathrm{Gl}_n(K)$ auf K^n vermöge

$$\alpha\colon \mathrm{Gl}_n(K) \times K^n \longrightarrow K^n, \qquad (A, v) \mapsto Av,$$

denn für alle $v \in K^n$ ist $I_n v = v$ und für alle Matrizen $A_1, A_2 \in \mathrm{Gl}_n(K)$ ist $(A_1 A_2) \cdot v = A_1 \cdot (A_2 v)$.

(ii) Seien X eine Menge und $\mathrm{Sym}(X) = \{f\colon X \to X \mid f \text{ ist bijektiv}\}$. $\mathrm{Sym}(X)$ wird zur Gruppe durch Komposition von Abbildungen und operiert auf X vermöge

$$\alpha\colon \mathrm{Sym}(X) \times X \longrightarrow X, \qquad (f, x) \longmapsto f(x).$$

(iii) Sei $X = \mathbb{R}^2$ und $G = (\mathbb{R}, +)$. Dann operiert G auf X durch

$$\alpha\colon G \times X \longrightarrow X, \qquad (\varphi, v = (x, y)^t) \longmapsto \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Wegen

$$\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi_2 & -\sin\varphi_2 \\ \sin\varphi_2 & \cos\varphi_2 \end{pmatrix} = \begin{pmatrix} \cos(\varphi_1+\varphi_2) & -\sin(\varphi_1+\varphi_2) \\ \sin(\varphi_1+\varphi_2) & \cos(\varphi_1+\varphi_2) \end{pmatrix}$$

erklärt das wirklich eine Gruppenaktion.

(iv) Seien $G = (\mathbb{Z}/2\mathbb{Z}, +)$ und $X = (\mathbb{Z}/2\mathbb{Z})^3$. G operiert auf X vermöge

$$\alpha\colon G \times X \longrightarrow X, \qquad (\overline{a}, \begin{pmatrix} \overline{x} \\ \overline{y} \\ \overline{z} \end{pmatrix}) \longmapsto \begin{pmatrix} \overline{x+a} \\ \overline{y+a} \\ \overline{z+a} \end{pmatrix}.$$

Proposition XI.1.4: *Sei $\alpha\colon G \times X \to X$, $(g, x) \mapsto gx$ eine Gruppenaktion. Diese definiert eine Äquivalenzrelation auf X wie folgt: $x_1 \sim x_2$ falls es $g \in G$ mit $x_2 = g \cdot x_1$ gibt.*

Beweis: Wegen der Definition der Gruppenaktion gilt $x = 1_G \cdot x$ für alle $x \in X$, also $x \sim x$ für alle $x \in X$ und „$\sim$“ ist reflexiv.

Sind $x_1, x_2 \in X$ mit $x_1 \sim x_2$, so gibt es $g \in G$ mit $x_2 = g \cdot x_1$. Nach Bemerkung XI.1.2 ist also $g^{-1}x_2 = x_1$ und damit $x_2 \sim x_1$.

Sind $x_1, x_2, x_3 \in X$ mit $x_1 \sim x_2$ und $x_2 \sim x_3$, so gibt es $g_1, g_2 \in G$ mit $x_2 = g_1x_1$ und $x_3 = g_2x_2$. Es ist also $x_3 = g_2x_2 = g_2 \cdot (g_1x_1) = (g_2g_1) \cdot x_1$ und damit $x_1 \sim x_3$. □

Definition XI.1.5: Sei $\alpha\colon G \times X \to X$, $(g, x) \mapsto g \cdot x$ eine Gruppenaktion und „$\sim$“ die nach Proposition XI.1.4 zugehörige Äquivalenzrelation.

(i) Die Äquivalenzklasse $G \cdot x := [x]_\sim := \{g \cdot x \mid g \in G\}$ heißt *Bahn von x unter G*.

(ii) Die Menge $\mathrm{Stab}(x) := \{g \in G \mid g \cdot x = x\}$ heißt *Stabilisator von x in G* oder auch *Fixgruppe von x*.

(iii) Die Aktion α heißt *transitiv*, falls es nur eine Bahn gibt (oder äquivalenterweise falls es für beliebige $x, y \in X$ ein $g \in G$ mit $y = g \cdot x$ gibt).

(iv) α heißt *treu*, falls es für alle $g \in G - \{1_G\}$ ein $x \in X$ mit $g \cdot x \neq x$ gibt.

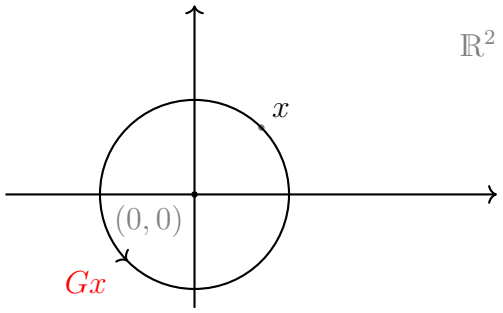
Beispiel XI.1.6: In Beispiel XI.1.3 erhalten wir folgende Bahnen und Fixgruppen:

(i) Es gilt $x_1 \sim x_2$ für alle $x_1, x_2 \in K - \{0\}$ (Basisergänzung und Fortsetzung), d. h. es gibt zwei Bahnen: $K^n - \{0\}$ und $\{0\}$. Es ist $\mathrm{Stab}(\mathbf{0}) = \mathrm{Gl}_n(K)$ und zum Beispiel

$$\mathrm{Stab}(e_1) = \{(e_1|*|\cdots|*) \in \mathrm{Gl}_n(K)\}.$$

(ii) $\mathrm{Sym}(X)$ operiert transitiv, denn für alle $x, y \in X$ erhalte Bijektion $x \mapsto y$, $y \mapsto x$, sonst id. Es ist $\mathrm{Stab}(x) \cong \mathrm{Sym}(X - \{x\})$.

(iii) Die Bahnen sind Kreise um $\mathbf{0} = (0,0)^t$, es sind $\mathrm{Stab}(x) = 2\pi\mathbb{Z}$, falls $x \neq \mathbf{0}$ und $\mathrm{Stab}(\mathbf{0}) = \mathbb{R}$.



(iv) Wir erhalten vier Bahnen mit je zwei Elementen. Für jedes $x \in X$ ist $\mathrm{Stab}(x) = \{\mathbf{0}\}$, d. h. die Aktion ist treu. Tatsächlich ist sie sogar fixpunktfrei.

Proposition XI.1.7 (Gruppenaktionen als Gruppenhomomorphismen): *Seien $(G, \star)$ eine Gruppe und X eine Menge.*

(i) *Jeder Gruppenhomomorphismus $\hat{\alpha}\colon G \to \mathrm{Sym}(X)$ definiert eine Gruppenaktion wie folgt:*

$$\alpha\colon G \times X \longrightarrow X, \qquad (g, x) \longmapsto (\hat{\alpha}(g))(x).$$

(ii) *Jede Gruppenaktion $\alpha\colon G \times X \to X$, $(g, x) \mapsto g.x$ definiert einen Gruppenhomomorphismus $\hat{\alpha}\colon G \to \mathrm{Sym}(X)$ wie folgt:*

$$\hat{\alpha}(g) = (x \longmapsto g.x).$$

Beweis: (i) Die Abbildung α ist eine Gruppenaktion, denn es gelten

$$(1_G, x) \longmapsto (\hat{\alpha}(1_G))(x) = \mathrm{id}(x) = x,$$

sowie

$$(g_1 \star g_2, x) \longmapsto (\hat{\alpha}(g_1 \star g_2))(x) = (\hat{\alpha}(g_1) \circ \hat{\alpha}(g_2))(x) = \alpha(g_1, \alpha(g_2, x)).$$

(ii) Die Abbildung $\hat{\alpha}$ ist ein Gruppenhomomorphismus, denn für alle g_1 und g_2 aus G gilt

$$\begin{aligned}\hat{\alpha}(g_1 \star g_2) &= (x \longmapsto (g_1 \star g_2).x) \\ &= (x \longmapsto g_1.(g_2.x)) = (x \longmapsto g_1.x) \circ (x \longmapsto g_2.x).\end{aligned}$$ □

Bemerkung XI.1.8: Die beiden Konstruktionen in Proposition XI.1.7 sind invers zueinander, d. h. Gruppenaktionen können auch als Homomorphismen von G nach $\mathrm{Sym}(X)$ aufgefasst werden. Häufig werden Gruppenaktionen so definiert.

Bemerkung XI.1.9 (Treue): Seien $\alpha\colon G \times X \to X$ eine Gruppenaktion und $\hat{\alpha}\colon G \to \mathrm{Sym}(X)$ der zugehörige Gruppenhomomorphismus aus Proposition XI.1.7. Die Aktion α ist treu genau dann, wenn $\hat{\alpha}$ injektiv ist.

Satz 38 (Bahnformel): *Seien $\alpha\colon G \times X \to X$, $(g, x) \mapsto g.x$ eine Gruppenaktion und G eine endliche Gruppe. Dann gilt für alle $x \in X$:*

$$\#G = \#Gx \cdot \#\,\mathrm{Stab}(x).$$

Beweis: Auf G erkären wir eine Äquivalenzrelation durch

$$g \sim g' :\Longleftrightarrow g.x = g'.x.$$

Es bezeichne im Folgenden $G/\!\sim$ die Menge der Äquivalenzklassen bezüglich dieser Äquivalenzrelation.

Zunächst erhalten wir eine Bijektion

$$\Phi\colon G/\!\sim \longrightarrow Gx, \qquad [g] \longmapsto g.x;$$

diese Abbildung ist wohldefiniert, da für g und g' aus $[g]$ gilt $g.x = g'.x$, außerdem ist sie per Konstruktion injektiv und sie ist surjektiv, da $g.x$ das Urbild $[g]$ hat.

Weiter enthält jede Äquivalenzklasse $\#\operatorname{Stab}(x)$ viele Elemente, denn $g'.x = g.x$ gilt genau dann, wenn $g^{-1} \star g'.x = x$, d. h. genau dann wenn $g^{-1} \star g' \in \operatorname{Stab}(x)$, also ist die Abbildung

$$\{g' \in G \mid g'.x = g.x\} \longrightarrow \operatorname{Stab}(x), \qquad g' \longmapsto g^{-1} \star g'$$

eine Bijektion.

Insgesamt erhalten wir

$$\#G = \#G/\!\sim \cdot \#\operatorname{Stab}(x) = \#Gx \cdot \#\operatorname{Stab}(x). \qquad \square$$

Beispiel XI.1.10 (Induzierte Aktion auf $\mathfrak{P}(X)$): Sei $\alpha\colon G \times X \to X, (g, x) \mapsto g.x$ eine Gruppenaktion. Dies definiert eine Gruppenaktion von G auf $\mathfrak{P}(X)$ wie folgt:

$$G \times \mathfrak{P}(X) \longrightarrow \mathfrak{P}(X), \qquad (g, A) \longmapsto gA = \{g.a \mid a \in A\}.$$

Die Abbildung $A \mapsto g.A$, $a \mapsto g.a$ ist eine Bijektion mit Umkehrabbildung $g.A \to A$, $b \mapsto g^{-1}.b$.

Insbesondere gilt: Ist A endlich, so haben alle Mengen in der Bahn von A die gleiche Anzahl an Elementen.

Beispiel XI.1.11 (Aktion durch Linksmultiplikation): Sei $(G, \star)$ eine Gruppe. G operiert auf sich selbst wie folgt:

$$\alpha\colon G \times G \longrightarrow G, \qquad (g_1, g_2) \longmapsto g_1.g_2 := g_1 \star g_2.$$

Diese Aktion ist transitiv und treu.

Satz 39 (von Cayley): *Jede Gruppe bettet sich durch $\hat{\alpha}\colon G \hookrightarrow \operatorname{Sym}(G)$ in $\operatorname{Sym}(G)$ ein. Insbesondere: Ist G endlich mit $\#G = n$, so erhalten wir eine Einbettung*

$$\hat{\alpha}\colon G \hookrightarrow \operatorname{Sym}(G) \cong S_n.$$

Das heißt: Jede endliche Gruppe ist isomorph zu einer Untergruppe einer symmetrischen Gruppe S_n.

Beweis: Sei α_G die Aktion von G auf G durch Linksmultiplikation wie in Beispiel XI.1.11 und $\hat{\alpha}_G$ der induzierte Gruppenhomomorphismus (siehe Proposition XI.1.7). Bemerkung XI.1.9 liefert dann die Behauptung. $\square$

2. Teilbarkeit in Ringen

Ziel dieses Abschnittes soll die Verallgemeinerung des von den ganzen Zahlen bekannten Teilbarkeitsbegriffes auf kommutative Ringe mit Eins sein. Auf dem Weg dorthin werden wir Teiler, größte gemeinsame Teiler und Primfaktorzerlegungen untersuchen und uns fragen, welche Ringe den passenden Rahmen für diese Begriffe darstellen.

Im Folgenden sei stets $(R, +, \cdot)$ ein kommutativer Ring mit Eins mit $\#R > 1$.

Erinnerung XI.2.1: (i) Es seien a und b Elemente von R. Wir sagen, a *teile* b, falls es $c \in R$ mit $b = ca$ gibt. Wir schreiben $a \mid b$ für „a teilt b“.

(ii) Ein Element a von R (beziehungsweise $R - \{0\}$) heißt *(echter) Nullteiler*, falls es $c \in R - \{0\}$ mit $ac = 0$ gibt.

(iii) Der Ring R heißt *nullteilerfrei*, falls R keine echten Nullteiler hat.

(iv) Ein Element a von R heißt *invertierbar*, falls es $b \in R$ mit $ab = ba = 1$ gibt.

(v) Die Teilmenge $R^\times := \{r \in R \mid r \text{ ist invertierbar}\} \subseteq R$ heißt *multiplikative Gruppe* oder *Einheitengruppe*.

Bemerkung XI.2.2 (Kürzungsregel):

(i) Der Ring R ist nullteilerfrei genau dann, wenn für alle $a, b \in R$, $a \neq 0$, gilt: Das c aus Erinnerung XI.2.1 (i) ist eindeutig (d. h. gilt $ca = c'a$, so ist $c = c'$),

(ii) Teiler von Einheiten sind Einheiten.

Beweis: (i) Seien $a \neq 0$ und $c, c' \in R$. Ist $ca = c'a$, so ist $(c - c')a = 0$, d. h. es gilt $c = c'$ oder a ist ein echter Nullteiler.

(ii) Seien $a \in R^\times$ und $b \in R$ mit $ab = 1$. Für alle $t \in R$ mit $t \mid a$ gibt es $c \in R$ mit $a = ct$, d. h. $1 = ab = ctb = tcb$, d. h. $t \in R^\times$. □

Beispiel XI.2.3: (i) Sei $R = (\mathbb{Z}/n\mathbb{Z}, +, \cdot)$. Aus der Linearen Algebra I (Satz 6) wissen wir

$$R^\times = \{\overline{a} \in \mathbb{Z}/n\mathbb{Z} \mid \mathrm{ggT}(a, n) = 1\},$$

Elemente $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ mit $\mathrm{ggT}(a, n) > 1$ sind Nullteiler. Zum Beispiel in $\mathbb{Z}/6\mathbb{Z}$ gilt $\overline{5} \cdot \overline{4} = \overline{20} = \overline{8} = \overline{2} \cdot \overline{4}$, die Kürzungsregel gilt also nicht.

(ii) Sei $R = \mathbb{Q}[X]$. Die Einheitengruppe ist

$$R^\times = \{\text{Polynome von Grad } 0\}.$$

Nullteiler in diesem Ring ist nur das Nullpolynom.

Ist „$a \mid b$“ eine Ordnungrelation? Offensichtlich ist Teilbarkeit reflexiv, denn $a = 1a$ für alle $a \in R$, damit $a \mid a$ für alle $a \in R$.

Sind $a, b, c \in R$ mit $a \mid b$ und $b \mid c$, so gibt es $t_1, t_2 \in R$, sodass $c = t_2 b$ und $b = t_1 a$, d. h. $c = t_2 t_1 a$ und damit $a \mid c$, also ist Teilbarkeit transitiv.

Sind $a, b \in R$ sodass $a \mid b$ und $b \mid a$, dann gibt es $t_1, t_2 \in R$ mit $b = t_1 a$ und $a = t_2 b$, also $1a = t_2 b = t_2 t_1 a$. Ist R nullteilerfrei, so können wir folgern dass $1 = t_2 t_1$ und damit $t_1 t_2 \in R^\times$, also ist Teilbarkeit für nullteilerfreie Ringe fast antisymmetrisch.

Proposition XI.2.4 (Teilbarkeit ist fast antisymmetrisch): *Sei R nullteilerfrei. Für alle $a, b \in R$ gilt genau dann $a \mid b$ und $b \mid a$, wenn es $t \in R^\times$ mit $b = ta$ gibt.*

Beweis: „$\Rightarrow$“: Das haben wir uns eben klar gemacht.
„$\Leftarrow$“: Gilt $b = ta$, so ist $t^{-1}b = a$ und somit $a \mid b$ und $b \mid a$. □

Definition XI.2.5 (Assoziiertheit):

(i) Zwei Elemente $a, b \in R$ heißen *assoziiert*, falls es $t \in R^\times$ mit $b = ta$ gibt. In diesem Fall schreiben wir $a \sim b$.

(ii) $aR^\times := \{at \mid t \in R^\times\}$ heißt *Assoziiertheitsklasse von a*.

Im Folgenden sei R stets ein nullteilerfreier Ring.

Bemerkung XI.2.6: Assoziiertheit ist eine Äquivalenzrelation auf einem nullteilerfreien Ring und die Assoziiertheitsklassen $aR^\times$ sind genau die Äquivalenzklassen.

Proposition XI.2.7 (Teilbarkeit auf Assoziiertheitsklassen): *Für alle $a, b, a', b' \in R$ mit $a \sim a'$ und $b \sim b'$ gilt $a \mid b$ genau dann, wenn $a' \mid b'$.*

Beweis: Da a und a' sowie b und b' assoziiert sind, gibt es Einheiten t_1, t_2, sodass $a' = t_1 a$ und $b' = t_2 b$. Wegen $a \mid b$ gibt es $c \in R$ mit $b = ca$ und wir können ersetzen

$$b' = t_2 b = t_2 ca = t_2 c t_1^{-1} a',$$

d. h. $a' \mid b'$. □

Bemerkung XI.2.8: Teilbarkeit ist eine Ordnungsrelation auf der Menge der Assoziiertheitsklassen eines nullteilerfreien Rings. Für $a, b \in R$ mit $a \mid b$ schreiben wir auch $aR^\times \leqslant bR^\times$.

Definition XI.2.9 (Größter gemeinsamer Teiler): Seien $a, b \in R$. Ein Element $g \in R$ heißt *größter gemeinsamer Teiler von a und b*, falls gelten:

(i) $g \mid a$ und $g \mid b$,

(ii) Für $g' \in R$ mit $g' \mid a$ und $g' \mid b$ gilt schon $g' \mid g$ (oder äquivalent: $gR^\times$ ist die größte Assoziiertheitsklasse unter allen Assoziiertheitsklassen von gemeinsamen Teilern).

Bemerkung XI.2.10: (i) Sind g_1 und g_2 größte gemeinsame Teiler von a und b, so gilt $g_1 \sim g_2$ nach Proposition V.2.4.

(ii) Ist g ein größter gemeinsamer Teiler von a und b, so schreiben wir auch $\mathrm{ggT}(a, b) = gR^\times$ oder $\mathrm{ggT}(a, b) = g$.

(iii) Zwei Elemente $a, b \in R$ heißen *teilerfremd*, falls für alle $h \in R$ mit $h \mid a$ und $h \mid b$ schon $h \in R^\times$ gilt. Existiert ein größter gemeinsamer Teiler, so ist das äquivalent zu $\mathrm{ggT}(a, b) = 1$ bzw. $\mathrm{ggT}(a, b) = R^\times$.

Beispiel XI.2.11 (Ein paar ggTs): Seien $a \in R$ und $u \in R^\times$.

(i) Es ist $\mathrm{ggT}(a, u) = R^\times$ (vergleiche Bemerkung V.2.2 (ii)).

(ii) Es ist $\mathrm{ggT}(a, 0) = a$.

(iii) Ist $R = \mathbb{Z}[X]$, so ist $\mathrm{ggT}(2, X) = 1$, denn zunächst gelten $1 \mid 2$ und $1 \mid X$; ist andererseits g' ein Teiler von 2, so muss $\deg(g') = 0$ sein und wegen $g' \mid X$ muss $g' \in \{\pm 1\}$ gelten.

Erinnerung: In der Linearen Algebra I (III.4.15) und in der Linearen Algebra II (I.3.16) tauchte bereits das Lemma von Bézout auf; dessen Aussage war: Für $R = \mathbb{Z}$ und $R = K[X]$ für einen Körper K gilt: Sind $f, g \in R$ mit $\mathrm{ggT}(f, g) = 1$, so gibt es $a, b \in R$ mit $1 = af + bg$.

In welchen Ringen gilt eigentlich das Lemma von Bézout?

Erinnerung XI.2.12: (i) Eine Teilmenge $I \subseteq R$ heißt *Ideal*, falls für alle $a, b \in I$ und $r \in R$ gelten: $a + b \in I$, $ra \in I$.

(ii) Sind $a_1, \ldots, a_n \in R$, dann ist

$$I := \langle a_1, \ldots, a_n \rangle := Ra_1 + \cdots + Ra_n = \{r_1 a_1 + \cdots + r_n a_n \mid r_1, \ldots, r_n \in R\}$$

ein Ideal in R, nämlich das von $a_1, \ldots, a_n$ *erzeugte Ideal*.

(iii) I heißt *Hauptideal*, falls es $g \in R$ mit $I = \langle g \rangle = Rg$ gibt.

(iv) R heißt *Hauptidealring*, falls jedes Ideal in R ein Hauptideal ist.

(v) $\mathbb{Z}$ und $K[X]$ (für einen Körper K) sind Hauptidealringe.

Bemerkung XI.2.13: Seien $a, b \in R$ und d ein gemeinsamer Teiler von a und b, d. h. $d \mid a$ und $d \mid b$. Dann gilt $\langle a, b \rangle = Ra + Rb \subseteq \langle d \rangle = Rd$.

Beweis: Gelten $d \mid a$ und $d \mid b$, dann gibt es $r_1, r_2 \in R$ mit $a = r_1 d$ und $b = r_2 d$, d. h. für alle $x, y \in R$ gilt $xa + by = (xr_1 + yr_2)d \in \langle d \rangle$. □

Proposition XI.2.14 (ggT vs Hauptideal): *Seien $a, b \in R$, sodass $\langle a, b \rangle$ ein Hauptideal ist. Dann gilt $\langle a, b \rangle = Rg$ genau dann, wenn $g = \mathrm{ggT}(a, b)$.*

Beweis: Siehe Übungsblatt 2, Aufgabe 2 (Lineare Algebra II). □

Korollar XI.2.15: *In Hauptidealen gilt das Lemma von Bézout. Mehr noch: Ist R ein Hauptidealring, so gibt es für alle $a, b \in R$ ein $g \in R$ mit $g = \mathrm{ggT}(a, b)$. Für solch ein g gilt dann $\langle a, b \rangle = Rg$.*

Beweis: Das ist eine direkte Konsequenz von Proposition V.2.14. □

Beispiel XI.2.16 (Nicht-Beispiele): (i) Sei $R = \mathbb{Z}[X]$. In Beispiel V.2.11 haben wir uns überlegt, dass $\mathrm{ggT}(2, X) = 1$, aber $\langle 2, X \rangle \neq 1$ (denn alle Polynome in $\langle 2, X \rangle$ haben geraden konstanten Term). $\mathbb{Z}[X]$ ist also *kein* Hauptidealring; das Lemma von Bezóut gilt nicht.

(ii) Sei

$$R = \mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{-3} \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{C}.$$

Hierbei meint $\sqrt{-3} := \mathrm{i}\sqrt{3}$. $(R, +, \cdot)$ ist ein Ring (als Teilring von $\mathbb{C}$). Die Elemente $x = 4$ und $y = 2 + 2\sqrt{-3}$ haben keinen größten gemeinsamen Teiler.

3. Euklidische Ringe

Im Folgenden sei stets R ein kommutativer Ring mit Eins.

In diesem Abschnitt suchen wir nach einer geeigneten Verallgemeinerung vom Teilen mit Rest.

Sind $a, b \in \mathbb{Z}$, $b \neq 0$, dann gibt es $c \in \mathbb{Z}$ mit $|a - bc| < |b|$ und sind $f, g \in K[X]$, $g \neq 0$, dann gibt es $h \in K[X]$ mit $\deg(f - gh) < \deg(g)$.

Definition XI.3.1 (Euklidischer Ring): Sei R ein nullteilerfreier Ring. R heißt *euklidisch*, falls es $\varphi\colon R \to \mathbb{N}_0$ mit

(i) $\varphi(r) = 0$ genau dann, wenn $r = 0$,

(ii) Für alle $a, b \in R$, $b \neq 0$, gibt es $c \in R$, sodass $\varphi(a - bc) < \varphi(b)$

gibt. φ heißt *Gradfunktion.*

Beispiel XI.3.2 (Erste euklidische Ringe): (i) $\mathbb{Z}$ mit $\varphi\colon \mathbb{Z} \to \mathbb{N}_0$, $z \mapsto |z|$ ist euklidischer Ring,

(ii) Sei K ein Körper. Der Polynomring $\mathbb{K}[X]$ ist euklidisch mit

$$\varphi\colon K[X] \longrightarrow \mathbb{N}_0, \qquad \begin{cases} 0, & \text{falls } f = 0, \\ \deg(f) + 1, & \text{sonst.} \end{cases}$$

(iii) Der Ring $R = \mathbb{Z}[\mathrm{i}] = \{a + b\mathrm{i} \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$ ist ein euklidischer Ring mit der Gradfunktion

$$\varphi\colon \mathbb{Z}[i] \longrightarrow \mathbb{N}_0, \qquad a + b\mathrm{i} \longmapsto a^2 + b^2.$$

Das dürfen Sie auf dem kommenden Übungsblatt überprüfen.

Proposition XI.3.3: *Ist R ein euklidischer Ring, so ist R ein Hauptidealring.*

Beweis: Es bezeichne $\varphi\colon R \to \mathbb{N}_0$ die Gradfunktion von R (vergleiche Definition XI.3.1). Ferner seien $\{0\} \neq I \subseteq R$ ein Ideal und

$$M := \{\varphi(a) \mid a \in I - \{0\}\} \subseteq \mathbb{N}.$$

Als Teilmenge der natürlichen Zahlen hat M ein Minimum m. Wähle ein Ringelement $a_0 \in \varphi^{-1}(m)$; wegen der Eigenschaften von φ ist $a_0 \neq 0$.

Ist $a \in I$ ein beliebiges Element, so gibt es $c \in R$ mit $\varphi(a - ca_0) < \varphi(a_0)$, d. h. $\varphi(a - ca_0) = 0$ wegen der Minimalität von $m = \varphi(a_0)$. Wegen den Eigenschaften von φ heißt das $a - ca_0 = 0$, d. h. $a \in Ra_0$. Offensichtlich gilt auch $Ra_0 \subseteq I$, also ist $I = Ra_0$ und damit ist I ein Hauptideal. □

Korollar XI.3.4 (In euklidischen Ringen gibt es ggTs): *Seien R euklidisch und $a, b \in R$. Dann gibt es $g \in R$ mit $g = \mathrm{ggT}(a, b)$.*

Für zwei Ringelemente $a, b \in R$ wollen wir jetzt Koeffizienten $x, y \in R$ mit $\mathrm{ggT}(a, b) = xa + yb$ finden.

Beispiel XI.3.5 (Iteriertes Teilen mit Rest): Gegeben seien die ganzen Zahlen $a = 93$ und $b = 42$.

$$\begin{aligned} 93 &= 1 \cdot 93 + 0 \cdot 42 \\ 42 &= 0 \cdot 93 + 1 \cdot 42 \\ 9 &= 1 \cdot 93 - 2 \cdot 42 \\ 6 &= -4 \cdot 93 + 9 \cdot 42 \\ 3 &= 5 \cdot 93 - 11 \cdot 42 \\ 0 &= -14 \cdot 93 + 31 \cdot 42 \end{aligned}$$

Das Ergebnis ist $\mathrm{ggT}(93, 42) = 3$ und $3 = 5 \cdot 93 - 11 \cdot 42$.

Proposition XI.3.6 (Euklidischer Algorithmus): *Seien R ein euklidischer Ring und $a, b \in R - \{0\}$. Berechne r_i, q_i, x_i und y_i iterativ wie folgt:*

$$r_{-1} = a, x_{-1} = 1, y_{-1} = 0, \qquad r_0 = b, x_0 = 0, y_0 = 1.$$

Für $i \geq 0$: Sind alle Werte bis i bereits bestimmt und ist $r_i \neq 0$, wähle q_{i+1} mit $\varphi(r_{i-1} - q_{i+1}r) < \varphi(r_i)$ und setze

$$r_{i+1} := r_{i-1} - q_{i+1}r_i, \qquad x_{i+1} := x_{i-1} - q_{i+1}x_i, \qquad y_{i+1} = y_{i-1} - q_{i+1}y_i.$$

Dann gelten die folgenden Aussagen:

(i) *Es gibt $n \in \mathbb{N}$ mit $r_n = 0$,*

(ii) *Es gilt $r_n = \mathrm{ggT}(a, b)$,*

(iii) *Es gilt $\mathrm{ggT}(a, b) = x_n \cdot a + y_n \cdot b$.*

Beweis: (i) Folgt aus den Tatsachen, dass $0 \leq \varphi(r_{i+1}) < \varphi(r_i)$ für alle i und dass φ in die natürlichen Zahlen abbildet.

(iii) Wir zeigen per vollständiger Induktion, dass $r_i = x_i a + y_i b$; dann folgt die Behauptung aus (i). Für $i = -1$ und $i = 0$ ist die Behauptung wahr. Für den Induktionsschritt von i nach $i + 1$ berechnen wir

$$x_{i+1}a + y_{i+1}b = (x_{i-1} - q_{i+1}x_i)a + (y_{i-1} - q_{i+1}y_i b) = r_{i-1} - q_{i+1}r_i = r_{i+1}.$$

(ii) Sei g ein größter gemeinsamer Teiler von a und b. Aus dem Beweis von (iii) folgt dass $r_n = x_n a + y_n b$, d. h. g teilt r_n.

Per vollständiger Induktion zeigen wir $r_n \mid a$ und $r_n \mid b$. Für den Induktionsanfang beobachten wir $r_{n+1} = 0$, also $r_{n-1} = q_{n+1}r_n$ und damit $r_n \mid r_{n-1}$ sowie $r_n \mid r_n$.

Für den Induktionsschritt von $i + 1$ nach i: Wegen $r_n \mid r_{i+1}$ und $r_n \mid r_i$ gilt $r_n \mid r_{i-1} = r_{i+1} + q_{i+1}r_i$, also die Behauptung. Insbesondere erhalten wir $r_n \mid q$ und damit insgesamt $g \sim r_n$. □

4. Primelemente in Ringen

Für diesen Abschnitt sei wieder R ein kommutativer Ring mit Eins. Ziel dieses Abschnittes soll eine Verallgemeinerung des Satzes über die eindeutige Primfaktorzerlegung natürlicher Zahlen sein.

Erinnerung: Aus der Schule ist bekannt: Eine natürliche Zahl n heißt *prim*, falls für jede Produktdarstellung $n = ab$ mit natürlichen Zahlen a, b schon gilt $a = 1$ oder $b = 1$.

Weiter bekannt ist: Jede natürliche Zahl n besitzt eine (bis auf Reihenfolge) eindeutige Darstellung als Produkt von Primfaktoren $n = \prod_{i=1}^{n} p_i$ mit Primzahlen p_i für $1 \leq i \leq n$, die sogenannte *Primfaktorzerlegung von* n.

Proposition XI.4.1 (Alternative Primalitätsbedingung in $\mathbb{N}$): *Sei $p \in \mathbb{N} - \{1\}$. Genau dann ist p prim, wenn für alle $a, b \in \mathbb{N}$ gilt: Wenn $p \mid ab$, dann $p \mid a$ oder $p \mid b$.*

Beweis: „⇒“: Seien p prim und $a, b \in \mathbb{N}$ mit $p \mid ab$. Sowohl a als auch b haben Primfaktorzerlegungen $a = \prod_{i=1}^{n} p_i$, $b = \prod_{j=1}^{m} q_i$ und es ist

$$ab = \Big(\prod_{i=1}^{n} p_i\Big)\Big(\prod_{j=1}^{m} q_j\Big).$$

Wegen der Eindeutigkeit der Primfaktorzerlegung ist $p \in \{p_1, \ldots, p_n, q_1, \ldots, q_m\}$, also $p \mid a$ oder $p \mid b$.

„⇐“: Seien $a, b \in \mathbb{N}$ mit $p = ab$, insbesondere gilt also $p \mid a$ oder $p \mid b$ nach Voraussetzung. Ohne Einschränkung gelte $p \mid a$, d. h. $a = tp$ mit $t \in \mathbb{N}$. Dann wäre $p = ab = tpb$, also $tb = 1$. Dann aber müssen t und b multiplikativ invertierbar in $\mathbb{N}$ sein, also $t = b = 1$ und p ist eine Primzahl in bekanntem Sinne. □

Die Hinrichtung im obigen Beweis braucht den Satz über die eindeutige Primfaktorzerlegung, die Rückrichtung benötigt die Kürzungsregel.

Definition XI.4.2 (Prim und irreduzibel): Sei $m \in R - (R^\times \cup \{0\})$.

(i) m heißt irreduzibel, falls für alle $a, b \in R$ mit $m = ab$ gilt $a \in R^\times$ oder $b \in R^\times$.

(ii) m heißt prim, falls für alle $a, b \in R$ mit $m \mid ab$ gilt $m \mid a$ oder $m \mid b$.

Bemerkung XI.4.3 (Prim vs irreduzibel): Seien R ein nullteilerfreier Ring und $m \in R - (R^\times \cup \{0\})$. Ist m prim, so ist m irreduzibel. Das beweist man so wie im Beweis von Proposition XI.4.1. Die Umkehrung gilt im Allgemeinen nicht, siehe Blatt 13.

Proposition XI.4.4 (Prim vs irreduzibel in HIR): *Seien R nullteilerfreier Hauptidealring und $m \in R - (R^\times \cup \{0\})$. Dann ist m irreduzibel genau dann, wenn m prim ist.*

Beweis: Wir müssen nur „$\Leftarrow$“ zeigen wegen (Bemerkung VI.4.4).

Seien $a, b \in R$ mit $m \mid ab$. Wir haben zu zeigen, dass $m \mid a$ oder $m \mid b$. Sei $g = \mathrm{ggT}(a, m)$; da m irreduzibel ist, ist $g \in R^\times$ oder $g \sim m$. Falls $g \sim m$ sind wir fertig, denn dann gilt $m \mid a$ wegen $g \mid a$. Ist $g \in R^\times$, so ist $\langle g, m\rangle = \langle g\rangle = R$, d. h. wir finden $x, y \in R$ mit $1 = xa + ym$ und können also schreiben $b = abx + bmy$. Da m beide Summanden in dieser Darstellung teilt, teilt m auch b. □

Definition XI.4.5: Ein Ideal $I \subsetneq R$ heißt *Primideal*, falls für alle $x, y \in R$ mit $xy \in I$ gilt: $x \in I$ oder $y \in I$.

Bemerkung XI.4.6 (Prim vs Primideal): Für $a \in R$ gilt: Genau dann ist a prim, wenn Ra ein Primideal ist.

Satz 40: *Sei R ein nullteilerfreier Hauptidealring und $r \in R - (R^\times \cup \{0\})$. Dann gelten:*

(i) *Es gibt prime $p_1, \ldots, p_n \in R$ mit $r = \prod_{i=1}^n p_i$,*

(ii) *Die Zerlegung ist eindeutig bis auf Reihenfolge und Assoziiertheit.*

Beweis (von Satz 40): Für den Beweis verwenden wir (Proposition VI.4.4).

(i) *Zur Existenz*: Es sei

$$S = \{r \in R \mid r \notin R^\times, r \neq 0, r \text{ hat keine solche Zerlegung}\}$$

die Menge aller Störenfriede. Angenommen, $S \neq \varnothing$ und $x \in S$. Dieses x wäre verschieden von Null, keine Einheit und nicht irreduzibel (denn sonst hätte es eine Darstellung als Produkt von irreduziblen Elementen), es gäbe also $x_1, y_1 \in R - R^\times$ mit $x = x_1 y_1$. Da $x \in S$ wären nicht beide Faktoren irreduzibel, ohne Einschränkung sei x_1 nicht irreduzibel. Außerdem wäre $Rx \subsetneq Rx_1$ (da $x \not\sim x_1 - y_1$ wäre ja keine Einheit).

Da x_1 nicht irreduzibel wäre, gäbe es $x_2, y_2 \in R - R^\times$ mit $x_1 = x_2 y_2$. Ohne Einschränkung wäre wieder x_2 nicht irreduzibel; weiter erhielten wir wieder $Rx \subsetneq Rx_1 \subsetneq Rx_2$.

Induktiv erhielten wir eine Folge $x, x_1, x_2, x_3, \ldots$ in S und eine aufsteigende Folge von Idealen $Rx \subsetneq Rx_1 \subsetneq Rx_2 \subsetneq Rx_3 \subsetneq \ldots$ in R. Nun bezeichne $I := \bigcup_{i \in \mathbb{N}} Rx_i$ die Vereinigung all dieser Ideale; I wäre selbst ein Ideal und da R ein Hauptidealring ist, gäbe es $a \in I$ mit $I = Ra$. Außerdem gäbe es ein $i \in \mathbb{N}$ mit $a \in Rx_i$, d. h. für alle $k \geq i$ gälte $Rx_k = Rx_i$. Dies ist ein Widerspruch.

(ii) *Zur Eindeutigkeit*: Es sei $x \in R - (R^\times \cup \{0\})$ mit Primfaktorzerlegungen $x = p_1 \cdots p_n = q_1 \cdots q_m$ mit Primelementen $p_1, \ldots, p_n$ und $q_1, \ldots, q_m$. Der Primfaktor p_1 muss das Produkt $q_1 \cdots q_m$ teilen, und da p_1 prim ist, gilt $p \mid q_i$ für ein $1 \leq i \leq m$. Ohne Einschränkung gelte $p_1 \mid q_1$, dann ist $q_1 = \varepsilon_1 p_1$ mit $\varepsilon_1 \in R^\times$, da q_1 irreduzibel ist, d. h. $p_1 \cdots p_n = \varepsilon_1 p_1 \cdot q_2 \cdots q_m$ und damit $p_2 \cdots p_n = \varepsilon_1 q_2 \cdots q_m$. Induktiv erhalten wir $n = m$ und $p_i \sim q_i$ für $1 \leq i \leq n$, also die Behauptung. □

Kapitel XII.

Unendlichdimensionale Vektorräume und Zornsches Lemma

In diesem Kapitel sei stets K ein Körper.

1. Motivation

Ziel dieses Abschnittes wird sein einzusehen, dass jeder K-Vektorraum eine Basis besitzt.

Erinnerung XII.1.1: Aus der Linearen Algebra I wissen wir: Jede maximale linear unabhängige Teilmenge B von V ist eine Basis. Dazu zeigt man, dass B auch ein Erzeugendensystem ist: Für irgendein $v \in V$ ist $B \cup \{v\}$ linear abhängig, d. h. es gibt $c_1, \ldots, c_k \in K$ und $b_1, \ldots, b_k \in B \cup \{v\}$ mit $\sum_{i=1}^{n} c_i b_i = 0$. Da B linear unabhängig ist, können wir ohne Einschränkung annehmen, dass $b_1 = v$, $c_1 \neq 0$ und $b_2, \ldots, b_n \neq v$, d. h.

$$v = b_1 = -\sum_{i=2}^{n} \frac{c_i}{c_1} b_i$$

und damit $v \in \langle B \rangle$.

Dieser Beweis benutzt nicht die Dimension von V. Was ist also das Problem für unendlichdimensionale Vektorräume? Wir müssen uns fragen: Gibt es eine maximale linear unabhängige Teilmenge von V?

Wir betrachten also $\mathfrak{P}(V)$ mit „$\subseteq$“ als Ordnungsrelation beziehungsweise genauer betrachten wir $\zeta := \{S \subseteq V \mid S \text{ ist linear unabhängig}\} \subseteq \mathfrak{P}(V)$ und untersuchen, ob ζ ein bezüglich „$\subseteq$“ maximales Element hat.

2. Das Zornsche Lemma

Erinnerung XII.2.1: Seien X eine Menge und „$\leqslant$“ eine Relation auf X.

(i) „$\leqslant$“ ist eine Ordnungsrelation, falls „$\leqslant$“ reflexiv, antisymmetrisch und transitiv ist. Das Paar $(X, \leqslant)$ heißt dann auch *geordnete Menge*.

(ii) „$\leqslant$“ heißt *Totalordnung*, falls für alle $x, y \in X$ gilt: $x \leqslant y$ oder $y \leqslant x$.

(iii) Eine Teilmenge $\varnothing \neq K \subseteq X$ heißt *Kette*, falls „$\leqslant$“ eine Totalordnung auf K ist.

Wir schreiben $x < y$, falls $x \leqslant y$ und $y \neq x$, $x \geqslant y$, falls $y \leqslant x$, und $x > y$, falls $y < x$.

Definition XII.2.2 (Maximal/größtes/minimal/kleinstes): Sei $(X, \leqslant)$ eine geordnete Menge.

(i) $x_0 \in X$ heißt *maximal*, falls für alle $x \in X$ mit $x_0 \leqslant x$ schon $x = x_0$ gilt.

(ii) $x_0 \in X$ heißt *minimal*, falls für alle $x \in X$ mit $x_0 \geqslant x$ schon $x = x_0$ gilt.

(iii) $x_0 \in X$ heißt *größtes Element*, falls für alle $x \in X$ schon $x \leqslant x_0$ gilt.

(iv) $x_0 \in X$ heißt *kleinstes Element*, falls für alle $x \in X$ schon $x \geqslant x_0$ gilt.

Bemerkung XII.2.3: X kann mehrere maximale– beziehungsweise minimale Elemente enthalten, aber nur ein größtes– beziehungsweise kleinstes.

Definition XII.2.4 (Obere Schranke, induktiv geordnet): Sei $(X, \leqslant)$ eine geordnete Menge.

(i) Für $S \subseteq X$ heißt $x_0 \in X$ *obere Schranke von* S, falls für alle $s \in S$ gilt: $s \leqslant x_0$. Eine obere Schranke $x_0 \in X$ heißt *kleinste obere Schranke*, falls für alle oberen Schranken $\tilde{x}_0 \in X$ von S gilt: $x_0 \leq \tilde{x}_0$.

(ii) X heißt *induktiv geordnet*, falls jede Kette in X eine obere Schranke hat. X heißt *strikt induktiv geordnet*, falls jede Kette in X eine kleinste obere Schrank hat.

Beispiel XII.2.5: (i) $(\mathbb{N}, \leqslant)$ ist nicht induktiv geordnet, denn zum Beispiel $\mathbb{N}$ ist eine Kette und hat keine obere Schranke.

(ii) Sei M eine Menge und $X = \mathfrak{P}(M)$ mit der Ordnungsrelation „$\subseteq$“. Für irgendeine Kette K in X erhalten wir durch Vereinigung aller Mengen in K eine kleinste obere Schranke von K, d. h. $(X, \subseteq)$ ist strikt induktiv geordnet. Außerdem sind $\varnothing$ das kleinste Element und M das größte Element.

(iii) Es sei $X = \mathfrak{P}(M) - \{\varnothing\}$. Weiterhin ist $(X, \subseteq)$ eine geordnete Menge. Alle einelementigen Teilmengen von M sind minimale Elemente.

Definition XII.2.6: Es sei $(X, \leqslant)$ eine geordnete Menge. Die Ordnungsrelation „$\leqslant$“ heißt *Wohlordnung*, falls jede Teilmenge $\varnothing \neq X' \subseteq X$ ein kleinstes Element enthält.

Beispiel XII.2.7: (i) $(\mathbb{N}, \leq)$ ist wohlgeordnet,

(ii) $(\mathfrak{P}(\mathbb{N}), \subseteq)$ ist nicht wohlgeordnet, zum Beispiel hat $\{\{1\}, \{2\}\}$ kein kleinstes Element.

Wir betrachten die folgenden Aussagen

(i) *Auswahlaxiom*: Seien $\varnothing \neq M$ eine Menge, $\varnothing \neq I$ eine Indexmenge und $(M_i)_{i \in I}$ eine Familie von nicht-leeren Teilmengen von M. Dann gibt es eine Abbildung $f\colon I \to M$ mit $f(i) \in M_i$ für alle $i \in I$. Eine solche Funktion f heißt *Auswahlfunktion*.

(ii) *Zornsches Lemma*: Sei $(M, \leqslant)$ eine (partiell) geordnete Menge, $M \neq \varnothing$. Falls $(M, \leqslant)$ induktiv geordnet ist, so gibt es ein maximales Element in M.

(iii) *Wohlordnungssatz*: Jede Menge M besitzt eine totale Ordnung, bezüglich der sie wohlgeordnet ist.

Satz 41: *Das Auswahlaxiom, das Zornsche Lemma und das Wohlordnungsaxiom sind äquivalent zueinander.*

Teile von Satz 41 werden nächstes Mal bewiesen. Das Auswahlaxiom ist ein eigenes Axiom, das nicht aus den ZF-Axiomen folgt. Manche Mathematiker akzeptieren das Auswahlaxiom nicht, und können so weniger Aussagen zeigen, zum Beispiel nicht das Zornsche Lemma, den Wohlordnungssatz oder die Existenz von Basen in Vektorräumen. Nimmt man das Auswahlaxiom zum ZF-Axiomensystem dazu, so spricht man von den ZFC-Axiomen (hierbei steht „C“ für „choice“).

3. Anwendung auf Vektorräume

Satz 42 (Existenz von Basen): *Sei K ein Körper. Jeder K-Vektorraum V hat eine Basis.*

Satz 42 folgt direkt aus dem nachfolgenden Satz:

Satz 43 (Basisergänzungssatz): *Sei $M \subseteq V$ eine linear unabhängige Menge. Dann gibt es eine Basis B von V mit $M \subseteq B$.*

Beweis (von Satz 43): Wir setzen

$$X := \{S \subseteq V \mid M \subseteq S \text{ und } S \text{ ist linear unabhängig}\}.$$

Es ist $X \neq \varnothing$, da $M \in X$, ferner ist $(X, \subseteq)$ partiell geordnet. Bleibt zu zeigen, dass $(X, \subseteq)$ induktiv geordnet ist um das Zornsche Lemma auf X anwenden zu können – dieses ergäbe dann, dass es ein maximales Element B in X gäbe. Wegen Erinnerung XII.1.1 wäre dieses B auch eine Basis.

Sei K eine Kette in X und setze $S' := \bigcup_{S \in K} S$. Wir wollen zeigen, dass $S' \in X$, denn dann ist S' eine obere Schranke.

Zunächst ist S' linear unabhängig: Ist $\{v_1, \ldots, v_k\}$ eine Teilmenge von S', so gibt es Indizes j_i, $1 \leq i \leq k$, mit $v_i \in S_{j_i}$ für $1 \leq i \leq k$. Da K total geordnet ist, gibt es i_0 mit $S_{j_i} \subseteq S_{i_0}$ für $1 \leq i \leq k$, also $\{v_1, \ldots, v_k\} \subseteq S_{i_0}$. Da S_{i_0} linear unabhängig ist, ist es auch $\{v_1, \ldots, v_k\}$.

Schließlich ist offensichtlich $M \subseteq S'$, also $S' \in X$, was den Beweis beschließt.□

4. Beweis des Zornschen Lemmas

Wir setzen nun das Auswahlaxiom voraus und wollen das Zornsche Lemma beweisen.

Proposition XII.4.1 (Ein Fixpunktsatz): *Sei $(X, \leqslant)$ eine partiell geordnete Menge mit $X \neq \varnothing$ für die gilt:*

(i) *X hat ein kleinstes Element $x_{\min}$,*

(ii) *X ist strikt induktiv geordnet.*

Ferner sei $F\colon X \to X$ eine Abbildung mit folgender Eigenschaft: Für jedes $x \in X$ ist $F(x) \geq x$. Dann hat F einen Fixpunkt.

Beweis (Lemma von Zorn): Sei $(X, \leqslant)$ partiell geordnete Menge, $X \neq \varnothing$ und X sei induktiv geordnet. Wir wollen zeigen, dass X dann maximale Elemente hat.

Zunächst wollen wir die Behauptung zeigen, falls X strikt induktiv geordnet ist. Ohne Einschränkung können wir annehmen, dass X ein kleinstes Element besitzt; hätte X kein kleinstes Element, so wählten wir $x_0 \in X$ und ersetzten X durch $X_{\geqslant x_0} := \{x \in X \mid x \geqslant x_0\}$.

Angenommen, X enthielte kein maximales Element. Für beliebiges $x \in X$ definierten wir $M_x := \{x' \in X \mid x' > x\} \subseteq X$ und nach unserer Annahme wäre jedes M_x nicht leer. Mit dem Auswahlaxiom erhielten wir eine Abbildung $F\colon X \to X$ mit $F(x) \in M_x$, d. h. für alle $x \in X$ hätten wir $F(x) > x$, F hätte also einen Fixpunkt; ein Widerspruch zu Proposition XII.4.1.

Nun wollen wir die allgemeine Situation auf den Fall von streng induktiv geordneten Mengen zurückführen. Zu unserer Menge X definieren wir die Menge $H := \{K \subseteq X \mid K \text{ ist eine Kette}\}$. Dieses H ist partiell geordnet mit Ordnungsrelation „$\subseteq$“. H ist außerdem strikt induktiv, denn für eine Kette $\{K_i \mid i \in I\}$ in H ist $\bigcup_{i\in I} K_i$ die kleinste obere Schranke. Unsere vorherigen Überlegungen liefern uns jetzt, dass H ein maximales Element K_0 hat, welches eine obere Schranke k_0 hat. Dieses k_0 ist unser Kandidat für ein maximales Element in X.

Zunächst bemerken wir, dass $k_0 \in K_0$, denn sonst wäre $K \cup k_0$ eine größere Kette als K_0. Für alle $x \in X$ mit $x \geqslant k_0$ gilt: Ist $k_0 \leqslant x$, dann ist x auch obere Schranke von K_0, d. h. $x \in K_0$ (wie für k_0), also $x \leqslant k_0$ und damit $x = k_0$. Damit ist k_0 maximal. □

Beweis (von Proposition XII.4.1): Wir nennen $S \subseteq X$ *zulässig*, falls gelten:

(i) $x_{\min} \in S$,

(ii) $F(S) \subseteq S$,

(iii) Für jedes Kette K in S liegt auch die kleinste obere Schranke m_K von K in S.

Es gibt solche Mengen, da X selbst natürlich zulässig ist. Wir setzen

$$S_0 := \bigcap \{S \mid S \subseteq X \text{ zulässig}\}.$$

Dieses S_0 erfüllt die Forderungen (i), (ii) und (iii), ist also zulässig, und ist außerdem die kleinste zulässige Teilmenge. Können wir zeigen, dass S_0 total geordnet ist, so ist ihre kleinste obere Schranke m_{S_0} ein Fixpunkt (da S_0 zulässig ist, liegt m_{S_0} in S_0 und wegen (ii) ist entsprechend $F(m_{S_0}) \in S_0$, d. h. $F(m_{S_0}) \leqslant m_{S_0}$, sodass wir insgesamt erhalten $m_{S_0} \leqslant F(m_{S_0}) \leqslant m_{S_0}$).

Wir nennen $e \in S_0$ *extremal*, falls für alle $s \in S_0$ mit $s < e$ gilt, dass $F(s) \leqslant e$. Zum Beispiel $x_{\min}$ ist extremal. Für extremales e setze

$$S_e := \{s \in S_0 \mid s \leqslant e \text{ oder } s \leqslant F(e)\}.$$

Wir behaupten, dass S_e zulässig ist: Zunächst gilt $x_{\min} \in S_e$.

Für $s \in S_e$ können auftreten $s < e$, in diesem Fall ist $F(s) \leqslant e$ da s extremal ist, d. h. $F(s) \in S_e$; $s = e$, dann gilt $F(s) = F(e) \in S_e$; $s > e$, dann ist $F(s) \leqslant s \leqslant F(e)$, also $F(s) \in S_e$.

Ist schließlich K eine Kette in S_e und m_K die kleinste obere Schranke, dann ist $m_K \in S_0$, da S_0 zulässig ist. Nun haben wir zwei Fälle zu unterscheiden: Gilt für alle $s \in K$ gilt $s \leqslant e$, so ist $m_K \leqslant e$ und damit $m_K \in S_e$. Gibt es $s \in K$ mit $s \not\leqslant e$, dann ist wegen $s \in S_e$ schon $s \geqslant F(e)$, d. h. $F(e) \leqslant s \leqslant m_K$ und somit $m_K \in S_e$.

Außerdem behaupten wir: Jedes Element $e \in S_0$ ist extremal.

Wir setzen $E := \{e \in S_0 \mid e \text{ ist extremal}\}$ und überprüfen, dass E zulässig ist. Offensichtlich ist $x_{\min} \in E$.

Für $e \in E$ wollen wir zeigen, dass $F(e) \in E$. Sei $s \in S_0$ mit $s < F(e)$, $s \neq e$. Dann ist $s \not\geqslant F(e)$ und da $s \in S_e$ ist, ist $s \leqslant e$. Weil e extremal ist, ist $F(s) \leqslant e \leqslant F(e)$, also $F(e) \in E$. Für $s = e$ ist $F(s) = F(e)$, insbesondere $F(s) \leqslant F(e)$ und $F(s) \in E$.

Sei nun K eine Kette in E und m_K die kleinste obere Schranke von K. Wir haben zu zeigen, dass $m_K \in E$, d. h. dass m_K extremal ist. Sei $s \in S_0$ mit $s < m_K$. Wir haben Fälle zu unterscheiden: Sind wir in der Situation, dass für alle $k \in K$ gilt $F(k) \leqslant s$, wäre wegen $k \leqslant F(k)$ dann s obere Schranke von K, d. h. $m_K \leqslant s$ – ein Widerspruch. Sind wir in der Situation, dass es $k \in K$ gibt mit $F(k) \not\leqslant s$, dann hätten wir $S_0 = S_k$ und da $s \in S_0$ wäre $s \leqslant k$. Falls $s < k$ ist, dann ist da k extremal ist auch $F(s) \leqslant k \leqslant m_K$. Falls $s = k$ ist, dann ist $k = s < m_K$ nach der Voraussetzung an s, d. h. es gibt $k' \in K$ mit $k = s < k'$, da m_K die kleinste obere Schranke ist, und wir können mit der vorherigen Überlegung schließen.

Aus unseren beiden Behauptungen können wir folgern, dass S_0 total geordnet ist: Sind nämlich $x, y \in S_0$, dann ist x extremal nach der zweiten Behauptung, $y \in S_x$ nach der ersten Behauptung, also $y \leqslant x$ oder $x \leqslant F(x) \leqslant y$. □