Lineare Algebra

gehalten von Prof. Weitze-Schmithüsen 2020-2021

UNIVERSITÄT
DES
SAARLANDES

Hinweise

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

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

guenther@math.uni-sb.de

Inhaltsverzeichnis

Teil 1.

Lineare Algebra I

Kapitel I.

Grundlagen

1. Etwas Motivation

Wir sehen uns mit folgendem Rätsel konfrontiert: Gesucht ist eine dreistellige Zahl n mit den Eigenschaften

(i) Notiert man die Ziffern von n in umgekehrter Reihenfolge und addiert sie zu n, so ergibt dies 1110,

(ii) Die Quersumme von n ist 15.

Sagen wir, die Zahl die dadurch entsteht, die Zahl n „verkehrt herum“ aufzuschreiben, heiße $\bar{n}$. Basierend auf (i) können wir einfach Kandidaten raten, zum Beispiel $n_1 = 753$, $n_2 = 555$ oder $n_3 = 654$. Die Zahl n_1 erfüllt (i), aber erfüllt (ii) nicht. Die Zahlen n_2 und n_3 erfüllen beide Bedingungen. Hatten wir nur Glück beim Raten? Können wir das Problem nicht vielleicht systematisch angehen?

Die typischen Fragen, die sich ein Mathematiker zu einem solchen Rätsel stellen würde, sind:

(1) Gibt es überhaupt eine solche Zahl n?

(2) Wenn ja, wie viele gibt es?

Beide Fragen sind oft selbst ohne ein explizites Lösungsverfahren interessant.

Schreiben wir die gesuchte Zahl n als $n = abc$ mit $a, b, c \in \{0, \ldots, 9\}$, dann ist $\bar{n} = cba$, was wir verstehen wollen als

$$\begin{aligned} n &= 100 \cdot a + 10 \cdot b + c \\ \bar{n} &= 100 \cdot c + 10 \cdot b + a \end{aligned}$$

woraus wir mit den Bedingungen aus unserem Rätsel das Gleichunggsystem

$$\begin{aligned} 101 \cdot a + 20 \cdot b + 101 \cdot c &= 1110 \\ a + \quad b + \quad c &= 15 \end{aligned} \tag{I.1}$$

erhalten. Glauben wir für den Moment, dass es sich dabei um legale Umformungen für lineare Gleichungssysteme handelt, dann erhalten wir durch Subtraktion des 101-fachen der zweiten Gleichung von der ersten Gleichung und anschließender Normierung (da dadurch a und c aus der ersten Gleichung eliminiert werden) das äquivalente lineare Gleichungssystem

$$\begin{aligned} b \quad &= 5, \\ a + b + c &= 15. \end{aligned}$$

Eliminieren wir jetzt noch b aus der zweiten Gleichung, dann erhalten wir die bestimmende Gleichung $a = 10 - c$. Die allgemeine Lösung des linearen Gleichungssystem ist ein Zahlentripel (a, b, c) mit

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 10 - c \\ 5 \\ c \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ 0 \end{pmatrix} + c \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Bemerkenswert ist, dass das Tripel $(a, b, c) = (10, 5, 0)$ das lineare Gleichungssystem Gl. (I.1) löst.

Die Lösung des linearen Gleichungssystems Gl. (I.1) sollte uns ursprünglich bei der Lösung des Rätsels unterstützen. Unterwerfen die erhaltenen Lösungen der Bedingung an a, b, c, Ziffern zu sein, so erhalten wir die neun möglichen Lösungen 159, 258, ..., 951.

Entlang dieses Beispiels werden die folgenden grundlegenden Fragen aufgeworfen:

(i) *Wie beschreibt man ein lineares Gleichungssystem am besten?* Die wichtige Information in Gl. (I.1) sind die Koeffizienten, schematisch können wir das lineare Gleichungssystem schreiben als

$$\left(\begin{array}{ccc|c} 101 & 20 & 101 & 1110 \\ 1 & 1 & 1 & 15 \end{array}\right).$$

Die Weiterverfolgung dieser Darstellung führt in natürlicher Weise auf den Matrix-Vektor-Kalkül.

(ii) *Welche Rechenoperationen dürfen an Gleichungen durchgeführt werden?* Nach einiger Arbeit werden wir einsehen, dass die legalen Rechenoperationen

mit linearen Gleichungen die sogenannten *elementaren Zeilenumformungen* sind und dass es einen Algorithmus gibt, der lineare Gleichungssysteme mithilfe dieser Umformungen verlässlich löst – der sogenannte *Gauß-Algorithmus*.

(iii) *Wie beschreibt man die Lösungsmenge und welche Struktur hat sie?* Eine zentrale Rolle für die Beschreibung von Lösungsmengen linearer Gleichungssysteme sind sogenannte Fundamentallösungen; ausgezeichnete Lösungen des zugehörigen homogenen linearen Gleichungssystems. Ein Phänomen, das bei linearen Gleichungssystemen auftreten kann, ist, dass die Gleichungen nicht ausreichen, um alle Parameter festzulegen. Solche, die keinen durch die Gleichungen beschriebenen Bedingungen unterworfen sind, heißen *freie Parameter*.

Es wird sich herausstellen, dass der richtige Rahmen für die Theorie linearer Gleichungssysteme die Theorie der Vektorräume und ihrer strukturerhaltenden Abbildungen – lineare Abbildungen – ist. Beim Studium der Vektorräume stoßen wir auf die Begriffe „Spann“, „Lineare Unabhängigkeit“, „Basis“, „Dimension“ und „Basiswechsel“.

(iv) *Über welchem Zahlenbereich suchen wir Lösungen?* Abhängig vom zugrundeliegenden Zahlenbereich wird sich herausstellen, dass es auch auf der Menge der „legalen“ Rechenoperationen eine nützliche Struktur gibt; dass man mit „Rechenoperationen rechnen kann“. In natürlicher Weise stoßen wir hier auf die Begriffe „Gruppe“, „Ring“, „Körper“, die spezielle sogenannte algebraische Strukturen beschreiben.

Vom linearen Gleichungssystem zur linearen Abbildung Zum linearen Gleichungssystem Gl. (I.1) gehört in gewisser Weise die Abbildung

$$f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \qquad \begin{pmatrix} a \\ b \\ c \end{pmatrix} \longmapsto \begin{pmatrix} 101a + 20b + 101c \\ a + b + c \end{pmatrix}.$$

Das Tripel (a, b, c) ist eine Lösung des Gleichungssystems aus Gl. (I.1) genau dann, wenn $f(a, b, c) = (1110, 15)$ und das Tripel (a, b, c) ist eine Lösung des homogenen linearen Gleichungssystem (d. h. die rechte Seite ist $(0, 0)$) genau dann, wenn $f(a, b, c) = (0, 0)$.

Die Abbildung f ist eine sogenannte lineare Abbildung. Charakteristika für lineare Abbildungen sind ihr *Kern* und ihr *Bild*. Wir werden für ein Kriterium für die Lösbarkeit linearer Gleichungssysteme interessieren, das mit der zugehörigen linearen Abbildung zu tun hat; außerdem werden wir uns für die Dimension von Bild und Kern interessieren.

Weiterführende Fragen im Stile der vorherigen sind dann

(v) *Wie hängen lineare Gleichungssysteme, lineare Abbildungen und Matrizen zusammen?* Geeigneten linearen Abbildungen lassen sich Darstellungsmatrizen (bezüglich spezieller Basen) zuordnen. Versucht man diese Darstellungsmatrizen bezüglich unterschiedlicher Basen zu bestimmen, so stellt sich ein Zusammenhang mit dem Vorgang des Wechsels der Basen heraus und es gibt Kriterien, mithilfe derer man Darstellungsmatrizen derselben linearen Abbildung bezüglich unterschiedlicher Basen identifizieren kann.

(vi) *Was sind wichtige Kenngrößen linearer Abbildungen, insbesondere Selbstabbildungen?* Wichtige Kenngrößen linearer Abbildungen sind ihr Rang (die Dimension ihres Bildes) und ihr Defekt (die Dimension ihres Kerns). Für Selbstabbildungen gibt es zusätzliche Charakteristika, besonders entscheidend sind die Determinante, Eigenwerte und Eigenvektoren.

(vii) *Gibt es spezielle lineare Abbildungen, die mit der zusätzlichen Struktur auf* $\mathbb{R}$*-Vektorräumen und* $\mathbb{C}$*-Vektorräumen zusammenpassen?* Auf dem aus der Schule bekannten $\mathbb{R}^3$ gibt es zusätzlich zur Vektorraumstrukur ein Skalarprodukt und damit die Konzepte „Winkel“ und „Länge“. Insbesondere das wichtige Konzept von Orthogonalität macht $\mathbb{R}$-Vektorräume oder $\mathbb{C}$-Vektorräume mit einem Skalarprodukt interessant für Anwendungen. Tatsächlich gibt es spezielle lineare Abbildungen, die auch „diese zusätzliche Struktur respektieren“, nämlich sogenannte Isometrien.

2. Voraussetzungen aus der Mengenlehre und Aussagenlogik

Im Vorlesungsskript tauchen einige Symbole immer wieder auf. Die grundlegendsten dieser Symbole sind „$\Leftrightarrow$“ („ist äquivalent“), „$\Rightarrow$“ („daraus folgt“), „$:=$“ („wird definiert als“) und „$:\Leftrightarrow$“ („wird definiert durch die nachfolgende Eigenschaft“). Es folgt eine knappe Übersicht über Inhalte und Konzepte, die wir für diese Vorlesung voraussetzen möchten.

2.1. Naive Mengenlehre

Eine Menge besteht aus Objekten, welche auch Elemente der Menge genannt werden. Beispielsweise die aus der Schule bekannten Zahlbereiche – die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$, die ganzen Zahlen $\mathbb{Z}$, die rationalen Zahlen $\mathbb{Q}$

und die reellen Zahlen $\mathbb{R}$ – sind Mengen. Weitere Beispiele für Mengen sind

$$M_1 := \{1, 2, 7, 11\}, \qquad M_2 := \{\text{Saarbrücken}, \text{Neunkirchen}, \text{Bonn}, \text{Köln}\},$$
$$M_3 := \{\{1, 2\}, \{1, 7\}, \{1, 2, 7, 11\}\}.$$

Zwei für unsere Zwecke besonders wichtige Prinzipien für Mengen sind die *Extensionalität* und das *Aussonderungsaxiom*.

Extensionalität Zwei Mengen M_1 und M_2 sind genau dann gleich, wenn sie dieselben Elemente haben. Anders ausgedrückt: Es gilt $M_1 = M_2$ genau dann, wenn $x \in M_1 \Leftrightarrow x \in M_2$.

Aussonderungsaxiom Zu jeder Menge M_1 und jeder Aussage P über Elemente von M_1 gibt es genau 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

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

Ein Beispiel für diese Konstruktion sind die positiven reellen Zahlen $\mathbb{R}_{>0}$, denn $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$.

Schließlich erwähnen wir die *leere Menge*. Es gibt genau eine Menge, die keine Elemente enthält. Diese Menge heißt leere Menge und wird mit „$\varnothing$“ notiert.

2.2. Grundlagen der Aussagenlogik

Unter einer Aussage verstehen wir einen Satz, dem eindeutig ein Wahrheitswert (wahr oder falsch) zugeordnet werden kann.[1] Die folgenden Sätze sind Beispiele für Aussagen:

- Alle Quadrate sind rund.
- Saarbrücken ist die Hauptstadt des Saarlandes.
- $7 = 11$.
- $7 = 11$.
- Wenn 2 ungerade ist, dann ist 1 gleich 0.

Die Sätze

- Komm sofort her!

[1] In Wahrheit ist die Welt etwas komplizierter, aber diese naive Definition ist vorerst ausreichend.

- Meinst Du, dass es morgen regnet?

hingegen sind keine Aussagen.

Zu jeder Aussage A gibt es eine *Verneinung* $\neg A$: Ist A wahr, dann ist $\neg A$ falsch und ist A falsch, dann ist $\neg A$ wahr.

Aus zwei Aussagen A und B lassen sich neue Aussagen wie folgt bilden;

(1) Die *Konjunktion* $A \wedge B$ („A und B“). Die Aussage $A \wedge B$ ist wahr, wenn A und B wahr sind, und sonst falsch. Zur Konjunktion gehört also die Wahrheitstabelle

A	W	W	F	F
B	W	F	W	F
$A \wedge B$	W	F	F	F

(2) Die *Disjunktion* $A \vee B$ („A oder B“). Die Aussage $A \vee B$ ist falsch, falls A und B falsch sind, und sonst wahr. Zur Disjunktion gehört die Wahrheitstabelle

A	W	W	F	F
B	W	F	W	F
$A \vee B$	W	W	W	F

(3) Die *Implikation* $A \Rightarrow B$ („aus A folgt B“). Die Aussage $A \Rightarrow B$ ist falsch, falls A wahr und B falsch ist, und sonst wahr, d. h. die zugehörige Wahrheitstabelle ist

A	W	W	F	F
B	W	F	W	F
$A \Rightarrow B$	W	F	W	W

Für die Implikation $A \Rightarrow B$ heißt A auch die *Voraussetzung* und B die *Konklusion* oder *Folgerung*. Statt „Aus A folgt B“ wird oft „Wenn A, dann B“ verwendet. Die vierte Spalte der Wahrheitstabelle merkt man sich am besten als „Aus Falschem folgt Beliebiges“ (vornehm „Ex falso quodlibet“). Beispielsweise ist die Implikation „Wenn $2 = 5$ ist, dann ist 6 ungerade“ wahr, obwohl „6 ist ungerade“ falsch ist.

(4) Die *Äquivalenz* $A \Leftrightarrow B$ („A genau dann, wenn B“). Die Aussage $A \Leftrightarrow B$ ist wahr, falls A und B beide wahr sind oder beide falsch sind, und sonst falsch. Wir haben also

A	W	W	F	F
B	W	F	W	F
$A \Leftrightarrow B$	W	F	F	W

2.3. Einige Regeln

Es seien A, B und C Aussagen.

(i) Genau dann ist $\neg(\neg A)$ wahr, wenn A wahr ist.

(ii) Es gelten $A \wedge B \Leftrightarrow B \wedge A$, sowie $A \vee B \Leftrightarrow B \vee A$.

(iii) Es gelten $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$, sowie $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$.

(iv) Es gelten die Äquivalenzen $(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$, sowie $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$.

(v) Es gelten $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$, sowie $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$.

Die Regeln aus (v) heißen „Regeln von de Morgan“. Die Regeln aus (ii) bedeuten die Kommutativität von $\wedge$ und $\vee$, die Regeln aus (iii) bedeuten die Assoziativität von $\wedge$ und $\vee$ und die Regeln aus (iv) bedeuten die Distributivität von $\wedge$ über $\vee$ und umgekehrt.

Die Regeln können mithilfe von Wahrheitstabellen überprüft werden, exemplarisch demonstrieren wir dies für die Regeln von de Morgan:

A	W	W	F	F
B	W	F	W	F
$A \wedge B$	W	F	F	F
$\neg(A \wedge B)$	F	W	W	W
$\neg A$	F	F	W	W
$\neg B$	F	W	F	W
$\neg A \vee \neg B$	F	W	W	W

2.4. Aussagen über Mengen mittels Quantoren

Seien M eine Menge und $P(x)$ eine Eigenschaft, die von Elementen x abhängen darf. Mithilfe sogenannter Quantoren lassen sich dann Aussagen formulieren:

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

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

(iii) *Eindeutige Existenz* $\exists!$: „$\exists!\, x \in M : P(x)$“ bedeutet „Es existiert genau ein $x \in M$, für das $P(x)$ wahr ist“. Die Aussage ist wahr, falls es genau ein $x \in M$ gibt, für das die Eigenschaft $P(x)$ wahr ist, und sonst falsch.

Beispiel: Seien M die Menge der reellen Zahlen $\mathbb{R}$ und $P(x)$ die Eigenschaft „$x^2 < 42$“. Dann sind „$\forall\, x \in M : P(x)$“ falsch, „$\exists\, x \in M : P(x)$“ wahr und „$\exists!\, x \in M : P(x)$“ falsch.

2.5. Regeln für Negation und Quantoren

Auch Aussagen, die Quantoren enthalten, lassen sich verneinen. Die Regeln für die Verneinung sind dabei

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

Die Verneinung von Aussagen, die eine eindeutige Existenz enthalten, ist etwas komplizierter:

$$\neg(\exists!\, x \in M : P(x)) \iff (\forall\, x \in M : \neg P(x)) \vee (\exists\, x, y \in M : x \neq y \wedge P(x), P(y))$$

In Prosa: Es gibt nicht genau ein Element x von M, für das $P(x)$ gilt, genau dann, wenn es gar keins gibt oder mehr als eines.

3. Konstruktionen in Mengentheorie

Definition I.3.1 (Teilmenge): Seien M_1 und M_2 Mengen. Gilt für jedes $x \in M_1$, dass $x \in M_2$, so heißt M_1 eine *Teilmenge von* M_2. Wir schreiben $M_1 \subseteq M_2$.

In Zeichen liest sich die obige Definition so: „$M_1 \subseteq M_2 :\Leftrightarrow \forall\, x \in M_1 : x \in M_2$“

Notation I.3.2: (i) Wir schreiben „$M_1 \subsetneq M_2$“, falls M_1 eine Teilmenge von M_2, jedoch nicht gleich M_2 ist. Eine solche Teilmenge heißt *echte Teilmenge*.

(ii) Wir schreiben „$x \in M$“ fall x ein Element von M ist.

(iii) Wir schreiben „$x \notin M$“, falls x kein Element von M ist.

Definition I.3.3 (Konstruktion neuer Mengen): Seien M_1 und M_2 Mengen.

(i) Die Menge
$$M_1 \cap M_2 := \{x \mid x \in M_1 \text{ und } x \in M_2\}$$
heißt *Schnitt der Mengen* M_1 *und* M_2.

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

(iii) Die Menge

$$M_1 - M_2 := M_1 \setminus M_2 := \{x \mid x \in M_1 \text{ und } x \notin M_2\}$$

heißt *Differenzmenge*.

(iv) Die Menge

$$M_1 \times M_2 := \{(x, y) \mid x \in M_1 \text{ und } y \in M_2\}$$

heißt *kartesisches Produkt von M_1 und M_2*.

(v) Sei k eine natürliche Zahl. Dann heißt

$$M_1^k := \{(x_1, \ldots, x_k) \mid x_1 \in M_1 \text{ und } \ldots \text{ und } x_k \in M_1\}$$

die *k-te kartesische Potenz von M_1*.

(vi) Die Menge aller Teilmengen von M_1 heißt *Potenzmenge $\mathfrak{P}(M_1)$ von M_1*, d. h.

$$\mathfrak{P}(M_1) := \{M \mid M \subseteq M_1\}.$$

Beispiel I.3.4: Seien $M_1 := \{1, 2\}$, $M_2 := \{1, 2, 3\}$, $M_3 := \varnothing$, $M_4 := \{1, 7, a, b\}$.

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

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

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

(iv) Das kartesische Produkt von M_1 und M_4 ist die Menge

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

(v) Es ist $M_2 - M_4 = \{2, 3\}$.

(vi) Die Potenzmenge von M_2 ist

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

Notation I.3.5: Seien M_1 und M_2 Mengen.

(i) Wir schreiben $M_1 \supseteq M_2$, falls M_2 eine Teilmenge von M_1 ist, und nennen M_1 eine *Obermenge von M_2*.

(ii) Ist M_1 eine Teilmenge von M_2, dann heißt die Differenzmenge $M_2 - M_1$ auch das *Komplement von M_1 in M_2*. Auch die Schreibweisen M_1^{c} oder $C_{M_2}(M_1)$ sind gebräuchlich.

Bemerkung I.3.6: Seien M, M_1 und M_2 Mengen. Dann gilt:

(i) Die Menge M ist eine Teilmenge von M.

(ii) Es gilt $M_1 = M_2$ genau dann, wenn M_1 eine Teilmenge von M_2 ist und umgekehrt.

Beweis: (i) Wir haben zu prüfen, ob „$\forall\, x \in M : x \in M$“ gilt. Hier gibt es nichts zu tun, man nennt eine solche Aussage eine Tautologie.

(ii) Diese Aussage lässt sich leicht mithilfe der Extensionalität zeigen. □

Proposition I.3.7 (Regeln für Schnitt und Vereinigung): *Es seien M_1, M_2 und M_3 Mengen. Dann gilt:*

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

(ii) *Es ist* $M_1 \cap M_2 = M_2 \cap M_1$ *sowie* $M_1 \cup M_2 = M_2 \cup M_1$.

(iii) *Es ist*

$$M_1 \cap (M_2 \cup M_3) = (M_1 \cap M_2) \cup (M_1 \cap M_3)$$
$$M_1 \cup (M_2 \cap M_3) = (M_1 \cup M_2) \cap (M_1 \cup M_3)$$

Beweis: Die Aussagen lassen sich mithilfe des Extensionalitätsprinzips beweisen. Wir zeigen exemplarisch für (i), wie so ein Beweis funktioniert. Um zu zeigen, dass $(M_1 \cap M_2) \cap M_3 = M_1 \cap (M_2 \cap M_3)$, zeigen wir, dass x zu $(M_1 \cap M_2) \cap M_3$ gehört genau dann, wenn x zu $M_1 \cap (M_2 \cap M_3)$ gehört. Wir haben

$$\begin{aligned}
x \in (M_1 \cap M_2) \cap M_3 &\iff x \in M_1 \cap M_2 \wedge x \in M_3 && \text{(Definition I.3.3)}\\
&\iff (x \in M_1 \wedge x \in M_2) \wedge x \in M_3 && \text{(Definition I.3.3)}\\
&\iff x \in M_1 \wedge (x \in M_2 \wedge x \in M_3) && \text{(Abschnitt 2.3(iii))}\\
&\iff x \in M_1 \wedge x \in M_2 \cap M_3 && \text{(Definition I.3.3)}\\
&\iff x \in M_1 \cap (M_2 \cap M_3) && \text{(Definition I.3.3)}
\end{aligned}$$

was zu zeigen war. Die anderen Aussagen bleiben Ihnen als Übungsaufgabe auf dem ersten Übungsblatt überlassen. □

Beispiel I.3.8 (Regeln für das Komplement): Seien M_1, M_2 und M Mengen, sodass $M_1 \subseteq M$ und $M_2 \subseteq M$. Dann gilt:

(i) $M - (M - M_1) = M_1$,[2]

[2] In den alternativen Notationen für das Komplement also $C_M(C_M(M_1)) = M_1$ oder $(M_1^{\mathsf{c}})^{\mathsf{c}}) = M_1$.

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

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

(iv) $M - (M_1 \cup M_2) = M - M_1 \cap M - M_2$,

(v) $M - (M_1 \cap M_2) = M - M_1 \cup M - M_2$.

Beweis: (i) Wir haben die Äquivalenzen

$$\begin{aligned}
& x \in M - (M - M_1) \\
\iff & x \in M \wedge x \notin M - M_1 && \text{(Definition I.3.3)} \\
\iff & x \in M \wedge \neg(x \in M - M_1) && \text{(Notation I.3.2)} \\
\iff & x \in M \wedge \neg(x \in M \wedge x \notin M_1) && \text{(Definition I.3.3)} \\
\iff & x \in M \wedge (\neg(x \in M) \vee \neg(x \notin M_1)) && \text{(Abschnitt 2.3(i))} \\
\iff & (x \in M \wedge x \notin M) \vee (x \in M \wedge x \in M_1) && \text{(Proposition I.3.7)} \\
\iff & x \in M \wedge x \in M_1,
\end{aligned}$$

da $(x \in M \wedge x \notin M)$ stets falsch ist, wie man sich per Wahrheitstabelle klar machen kann. Wegen $M_1 \subseteq M$ ist die letzte Aussage schließlich äquivalent dazu, dass x zu M_1 gehört, was wir zeigen wollten.

(ii) Ein x gehört zu $M - M$ per Definition genau dann, wenn x zu M gehört und wenn x nicht zu M gehört. Das ist stets falsch, d. h. $M - M$ enthält keine Elemente und ist somit die leere Menge.

(iii) Mit den Regeln aus (ii) und (i) sehen wir $M - \varnothing = M - (M - M) = M$.

Die Beweise der Aussagen aus (iv) und (v) bleiben Ihnen zur Übung überlassen. □

Proposition I.3.9: *Es seien M_1, M_2 und M_3 Mengen. Dann gilt:*

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

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

(iii) *Wenn $M_1 \subseteq M_2$ 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 $M_1 \cup M_2 = M_2$.*

Die Aussagen lassen sich ähnlich wie die restlichen Aussagen in diesem Abschnitt zeigen. Einige dieser Aussagen dienen als Beispiele im folgenden Abschnitt, die restlichen bleiben Ihnen zur Übung überlassen.

4. Nützliche Beweisverfahren

4.1. Nachweis von Teilmengenbeziehungen

Seien A und B Mengen. Um zu zeigen, dass A eine Teilmenge von B ist, bietet es sich an die Äquivalenz

$$A \subseteq B \iff \forall\, a \in A : a \in B$$

zu verwenden (siehe Definition I.3.1). Beispielsweise lässt sich für Mengen M_1 und M_2 die Aussage „$M_1 \cap M_2 \subseteq M_1$“ damit leicht zeigen.

Sei dazu x ein Element von $M_1 \cap M_2$. Per Definition I.3.3 gehört x zu M_1 und zu M_2, insbesondere gehört x also zu M_1 und Definition I.3.1 liefert jetzt, dass $M_1 \cap M_2 \subseteq M_1$.

4.2. Gleichheit von Mengen

Seien A und B Mengen. Um zu zeigen, dass A und B gleich sind, verwenden wir die Äquivalenz

$$A = B \iff A \subseteq B \wedge B \subseteq A,$$

vergleiche Bemerkung I.3.6(ii). Illustrieren können wir dieses Verfahren für die folgende Aussage über Mengen M_1 und M_2: „$M_1 \subseteq M_2 \Rightarrow M_1 \cap M_2 = M_1$“.

Beweis: „$\subseteq$“: Wir wissen aus Abschnitt 4.1, dass $M_1 \cap M_2 \subseteq M_1$.

„$\supseteq$“: Für ein x aus M_1 haben wir

$$\begin{aligned} x \in M_1 &\Longrightarrow x \in M_2 && (M_1 \subseteq M_2) \\ &\Longrightarrow x \in M_1 \wedge x \in M_2 \\ &\Longrightarrow x \in M_1 \cap M_2 && (\text{Definition I.3.3}) \end{aligned}$$

d. h. $M_1 \subseteq M_1 \cap M_2$ und wir sind fertig. □

4.3. Äquivalenz von Aussagen

Für Äquivalenz zweier Aussagen A und B gibt es folgende Charakterisierung:

$$(A \iff B) \iff ((A \Longrightarrow B) \wedge (B \Longrightarrow A)).$$

Als Beispiel zeigen wir damit, dass für Mengen M_1 und M_2 gilt: „Genau dann ist $M_1 \subseteq M_2$, wenn $M_1 \cap M_2 = M_1$“.

Beweis: „$\Longrightarrow$“: Das ist genau das Beispiel aus Abschnitt 4.2.
„$\Longleftarrow$“: Für M_1 und M_2 gelte $M_1 \cap M_2 = M_1$. Dann haben wir

$$\begin{aligned} x \in M_1 &\Longrightarrow x \in M_1 \cap M_2 && (M_1 \cap M_2 = M_1) \\ &\Longrightarrow x \in M_1 \wedge x \in M_2 && (\text{Definition I.3.3}) \\ &\Longrightarrow x \in M_2, \end{aligned}$$

also ist M_1 enthalten in M_2. Aus der Gültigkeit von „$\Longrightarrow$“ und „$\Longleftarrow$“ folgt die behauptete Äquivalenz. □

In dieser Beweisstrategie heißt „$\Longrightarrow$“ auch *Hinrichtung* und „$\Longleftarrow$“ die *Rückrichtung* des Beweises.

4.4. Widerspruchsbeweis

Seien A und B Aussagen. Für die Implikation „$A \Rightarrow B$“ gilt die Charakterisierung

$$(A \Longrightarrow B) \Longleftrightarrow (\neg B \Longrightarrow \neg A).$$

Es gibt zwei Möglichkeiten, diese Äquivalenz in einem Beweis zu verwenden:

(i) Nehmen wir an, $\neg B$ wäre wahr und zeigen wir dann, dass das $\neg A$ impliziert, dann haben wir auch gezeigt, dass B aus A folgt. Dieses Beweisverfahren nennt man *Kontraposition*.

(ii) Nehmen wir an, $\neg B$ wäre wahr und zeigen, dass sich daraus ein Widerspruch ergibt, dann haben wir gezeigt, dass aus $\neg B$ etwas Falsches folgt. Das ist nach der obigen Äquivalenz gleichbedeutend damit, dass sich B aus etwas Wahrem folgern lässt. Wegen der Eigenschaften der Implikation ist das äquivalent dazu, dass B wahr ist. Dieses Beweisverfahren nennt man *Widerspruchsbeweis*.

Wir illustrieren die Strategien mit Beispielen. Zunächst zur Kontraposition: Seien a, b und c positive natürliche Zahlen. Dann gilt: Wenn c nicht ab teilt, dann teilt c weder a noch b.

In diesem Beispiel ist A die Aussage „c teilt nicht ab“ und B ist die Aussage „c teilt weder a noch b“. Wir zeigen „$\neg B \Rightarrow \neg A$“ wie folgt:

Beweis: Angenommen, c teilt a oder c teilt b. Dann gäbe es eine natürliche Zahl k, sodass $a = kc$, oder es gäbe eine natürliche Zahl ℓ, falls $b = \ell c$. Wir hätten also $ab = kcb$ oder $ab = a\ell c$, d. h. c würde ab teilen. □

Als zweites Beispiel zeigen wir den klassischen Beweis von Euklid, dass es unendlich viele Primzahlen gibt. Dazu erinnern wir zunächst an grundlegende Eigenschaften von Primzahlen:

(i) Sei p eine natürliche Zahl. Ist $p \geq 2$ und sind 1 und p die einzigen Teiler von p, dann heißt p eine *Primzahl*.

(ii) Jede natürliche Zahl lässt sich als Produkt von Primzahlen schreiben.

Eine Aussage B ist wahr genau dann, wenn B aus etwas Wahrem folgt, d. h.

$$(W \Longrightarrow B) \Longleftrightarrow B.$$

Wir setzen $A = W$ und $B =$ „Es gibt unendlich viele Primzahlen“ und zeigen $\neg B \Rightarrow F$ wie folgt:

Beweis: Angenommen, es gäbe nur endlich viele Primzahlen $a_1, \ldots, a_n$, wobei n eine natürliche Zahl ist. Die Zahl $N := a_1 \cdots a_n + 1$ wäre größer als jede der Primzahlen $a_1, \ldots, a_n$ und damit keine Primzahl. Als natürliche Zahl ließe sich N als Produkt von Primzahlen schreiben, d. h. wir hätten $N = p_1 \cdots p_r$ mit einer natürlichen Zahl r und Primzahlen $p_1, \ldots, p_r$. Dann gäbe es eine Primzahl p_i, die gleichzeitig N und $N - 1$ teilt – ein Widerspruch. □

4.5. Beweis durch vollständige Induktion

Seien $A(n)$ eine Aussage, die von n abhängt, und

$$S := \{n \in \mathbb{N} \mid A(n) \text{ ist wahr}\}.$$

Können wir zeigen, dass

(i) $\ldots n_0$ zu S gehört,

(ii) $\ldots$ für n aus S auch $n + 1$ zu S gehört,

dann gilt die Aussage $A(n)$ für jede natürliche Zahl n, die größer gleich n_0 ist.

Beispiel: Für eine natürliche Zahl n definieren wir $T(n) := 1 + 2 + \cdots + n$. Wir behaupten, dass $T(n) = n(n + 1)/2$.

Zunächst ist $T(1) = 1 = 1 \cdot 2/2 = 1$, d. h. die Aussage gilt für $n = 1$.

Ist jetzt n eine natürliche Zahl, für die $T(n) = n(n+1)/2$ gilt, dann berechnen wir unter Verwendung dessen, dass

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

Insgesamt haben wir damit gezeigt, dass $T(n) = n(n+1)/2$ für jede natürliche Zahl größer gleich Eins gilt.

Vollständige Induktion wird in der Analysis I intensiver eingeführt und geübt. Eine Beweisidee für den eben gezeigten Beweis wird dem kleinen Gauss zugeschrieben. Der Lehrer soll die Klasse mit der Rechenaufgabe, die natürlichen Zahlen bis 100 zu addieren, beschäftigt haben wollen und Gauß soll innerhalb kurzer Zeit durch geschicktes Zusammenzählen die korrekte Lösung gefunden haben:

$$\begin{array}{rrcr} 1 & 2 & \dots & 50 \\ + & + & \dots & + \\ 100 & 99 & \dots & 51 \\ \hline 101 & 101 & \dots & 101 \end{array}$$

d. h. $1 + 2 + \cdots + 99 + 100 = 50 \cdot 101 = 100/2 \cdot 101$. Vielleicht ist auf den ersten Blick verwunderlich, dass $n(n+1)/2$ nur natürliche Zahlen als Werte annimmt. Weil aber $n + 1$ der Nachfolger von n ist, muss eine der beiden Zahlen gerade sein.

5. Abbildungen

Notation: Sei $M = \{a_1, \dots, a_n\}$ eine Menge, die aus n verschiedenen Elementen besteht. Dann heißt n die *Anzahl der Elemente von* M. Wir schreiben dafür $\#(M) = n$ oder $|M| = n$. Ist $\#(M)$ eine natürliche Zahl, dann nennen wir M *endlich*, bzw. wir sagen, dass M *endlich viele Elemente hat*.

Definition I.5.1 (Abbildung/Funktion): Seien X und Y Mengen.

(i) Eine Vorschrift, die jedem Element x aus X ein Element y aus Y zuordnet, heißt eine *Abbildung* oder *Funktion*. Wir schreiben „$f\colon X \to Y$“ für eine Funktion von X nach Y und mit „$x \mapsto f(x) := y$“ notieren wir die Zuordnung auf Elementebene. Hierbei heißt X der *Definitionsbereich* und Y der *Wertebereich*.

(ii) Die Menge der Abbildungen von X nach Y bezeichnen wir mit $\mathrm{Abb}(X, Y)$, d. h.

$$\mathrm{Abb}(X, Y) := \{f \mid f \text{ ist eine Abbildung von } X \text{ nach } Y\}.$$

Bemerkung I.5.2: (i) Man kann äquivalent dazu Abbildungen mengentheoretisch beschreiben. Eine Abbildung $f\colon X \to Y$ ist gegeben durch eine Teilmenge $\Gamma_f \subseteq X \times Y$ 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) zu Γ_f gehört. Die Menge Γ_f heißt dann *mengentheoretischer Graph von f*.

(ii) Zwei Abbildungen $f\colon X \to Y$ und $g\colon X \to Y$ sind per Definition gleich genau dann, wenn sie auf jedem Element von x dasselbe tun, d. h. wenn für jedes $x \in X$ gilt, dass $f(x) = g(x)$. Offensichtlich ist das genau dann der Fall, wenn $\Gamma_f = \Gamma_g$.

(iii) Für $X = \varnothing$ und Y beliebig enthält $\mathrm{Abb}(X, Y)$ genau ein Element f, dessen Graph Γ_f die leere Menge ist, denn das kartesische Produkt einer Menge mit der leeren Menge ist die leere Menge, diese Menge hat genau eine Teilmenge (nämlich $\varnothing$) und $\varnothing$ hat die Eigenschaft eines Graphen.

Beispiel I.5.3 (für Abbildungen): (i) Für $X_1 := \{-1, 0, 1\}$ und $Y_1 := \{0, 1\}$ betrachten wir die Abbildungen

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

Diese beiden Abbildungen sind gleich.

(ii) Für $X_2 := \mathbb{R}$ und $Y_2 := \mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$ ist $f\colon X_2 \to Y_2$, $x \mapsto x^2$ 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) Sind

$$X_4 := \{s \mid s \text{ ist Student in dieser Vorlesung}\}$$
$$\text{und} \qquad Y_4 := \{t \mid t \text{ ist Datum eines Tages im Jahr}\},$$

dann ist $f_4\colon X_4 \to Y_4$, $s \mapsto$ Geburtsdatum von s eine Abbildung.

Definition I.5.4 (eine besondere Abbildung): Sei M eine Menge. Die Abbildung $\mathrm{id}_M\colon M \to M$, $x \mapsto x$ heißt *Identität auf M*.

Definition I.5.5 (Bild und Urbild): Seien X und Y Mengen und $f\colon X \to Y$ eine Abbildung.

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

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

Beispiel I.5.6: Für die Abbildungen aus Proposition I.5.3 haben wir das Folgende:

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

(ii) $f_2^{-1}(\{y \in \mathbb{R}_{\geq 0} \mid y \geq 1\} = \{x \in \mathbb{R} \mid x \geq 1 \vee x \leq -1\}$.

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

(iv) $f_4^{-1}(\{4.\,\text{Juni}\}) = \varnothing$.

Definition I.5.7 (Verkettung und Einschränkung): Seien X, Y und Z Mengen.

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

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

und nennen sie die *Verkettung* oder *Komposition von f und g*.

(ii) Für eine Abbildung $f\colon X \to Y$ und $A \subseteq X$ heißt die Abbildung

$$f|_A\colon A \longrightarrow Y, \qquad x \longmapsto f(x)$$

die *Einschränkung von f auf A*.

Bemerkung I.5.8 (Eigenschaften der Verkettung): (i) Die Verkettung von Funktionen ist assoziativ, d. h. 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 nichts beim Verketten. Genauer: Ist $f\colon X \to Y$ eine Abbildung, dann gilt $\mathrm{id}_Y \circ f = f$ und $f \circ \mathrm{id}_X = f$.

Definition I.5.9 (Injektiv, surjektiv, bijektiv): Seien X und Y Mengen und $f\colon X \to Y$ eine Abbildung.

(i) Folgt für irgendwelche x_1 und x_2 aus X mit $f(x_1) = f(x_2)$, dass $x_1 = x_2$, dann heißt f *injektiv*. Die Abbildung f ist injektiv genau dann, wenn für jedes $y \in Y$ gilt, dass $\#f^{-1}(\{y\}) \leq 1$.

(ii) Besitzt jedes $y \in Y$ ein Urbild, d. h. gibt es für jedes $y \in Y$ (wenigstens) ein $x \in X$ mit $f(x) = y$, dann heißt f *surjektiv*. Die Abbildung f ist surjektiv genau dann, wenn für jedes $y \in Y$ gilt, dass $\#f^{-1}(\{y\}) \geq 1$.

(iii) Ist f injektiv und surjektiv, dann heißt f *bijektiv*. Das ist genau dann der Fall, wenn jedes $y \in Y$ genau ein Urbild hat, d. h., falls für jedes $y \in Y$ gilt, dass $\#f^{-1}(\{y\}) = 1$.

Proposition I.5.10 (Injektivität, Surjektivität und Bijektivität): *Seien X, Y und Z Mengen und $f\colon X \to Y$, $g\colon Y \to Z$ Abbildungen. Dann haben wir die folgenden Aussagen:*

(i) *Sind f und g injektiv, dann ist auch $g \circ f$ injektiv.*

(ii) *Sind f und g surjektiv, dann ist auch $g \circ f$ surjektiv.*

(iii) *Die Abbildung f ist injektiv genau dann, wenn es $h\colon Y \to X$ gibt, sodass $h \circ f = \mathrm{id}_X$.*

(iv) *Die Abbildung f ist surjektiv genau dann, wenn es $h\colon Y \to X$ gibt, sodass $f \circ h = \mathrm{id}_Y$.*

(v) *Die Abbildung f ist bijektiv genau dann, wenn es $h\colon Y \to X$ gibt, sodass $h \circ f = \mathrm{id}_X$ und $f \circ h = \mathrm{id}_Y$. In diesem Fall ist h eindeutig.*

Definition I.5.11 (Umkehrabbildung): Seien X und Y Mengen und $f\colon X \to Y$ eine Abbildung. Ist $g\colon Y \to X$ eine Abbildung, sodass $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$, dann heißt g *Umkehrabbildung* oder auch *inverse Abbildung zu f*. In diesem Fall sagt man auch, f und g seien zueinander invers. Oft wird g mit f^{-1} bezeichnet.

Bemerkung I.5.12: Seien X und Y Mengen und $f\colon X \to Y$ eine Abbildung. Die Festsetzungen aus Definition I.5.5 erklären zur Abbildung f gehörige Funktionen

$$f\colon \mathfrak{P}(X) \longrightarrow \mathfrak{P}(Y), \qquad A \longmapsto f(A),$$
$$f^{-1}\colon \mathfrak{P}(Y) \longrightarrow \mathfrak{P}(X), \qquad B \longmapsto f^{-1}(B).$$

Es handelt sich hierbei um Missbrauch der Notation, denn $f\colon X \to Y$ und $f\colon \mathfrak{P}(X) \to \mathfrak{P}(Y)$ sind verschiedene Abbildungen, die mit demselben Symbol belegt sind und eine Umkehrabbildung $f^{-1}\colon Y \to X$ existiert potentiell gar nicht. Verwechslungen sind dennoch ausgeschlossen, denn aus dem Kontext heraus ist durch das Argument (d. h. das Element aus dem Definitionsbereich, das unter f abgebildet wird) immer klar, welche der beiden Funktionen gemeint ist.

Die Notation für die Urbildfunktion $f^{-1}\colon \mathfrak{P}(Y) \to \mathfrak{P}(X)$ hat allerdings etwas mit der Umkehrfunktion zu tun: Gibt es eine Umkehrfunktion zur Funktion f, dann kann diese wegen Definition I.5.9(iii) mit der Urbildfunktion identifiziert werden, indem man ein Element $x \in X$ beziehungsweise $y \in Y$ identifiziert mit der Einpunktmenge $\{x\} \in \mathfrak{P}(X)$ beziehungsweise $\{y\} \in \mathfrak{P}(Y)$.

Beispiel I.5.13: Sei k eine natürliche Zahl und betrachte die Mengen X^k sowie $\mathrm{Abb}(\{1,\dots,k\},X)$. Dann sind die Abbildungen

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

und

$$G\colon \mathrm{Abb}(\{1,\dots,k\},X)\longrightarrow X^k,\qquad f\longmapsto (f(1),\dots,f(k))$$

zueinander invers.

Beweis: Seien $(a_1,\dots,a_k)\in X^k$ und $f:=F(a_1,\dots,a_k)$ die zugehörige Abbildung, d. h. $f(i)=a_i$. Dann haben wir

$$G(F(a_1,\dots,a_k))=(f(1),\dots,f(k))=(a_1,\dots,a_k),$$

d. h. $G\circ F=\mathrm{id}_{X^k}$.

Ist auf der anderen Seite $f\in\mathrm{Abb}(\{1,\dots,k\},X)$, dann ist

$$F(G(f))=F[(f(1),\dots,f(k))]=(h\colon\{1,\dots,k\}\to X,\quad i\mapsto f(i))=f,$$

d. h. $F\circ G=\mathrm{id}_{\mathrm{Abb}(\{1,\dots,k\},X)}$. In der obigen Rechnung haben wir verwendet, dass Abbildungen α und β genau dann gleich sind, wenn $\Gamma_\alpha=\Gamma_\beta$ ist, um zu schließen, dass h genau f ist.

Via F und G erhalten wir so eine Identifikation der Mengen X^k und $\mathrm{Abb}(\{1,\dots,k\},X)$. □

Definition I.5.14 (Leeres Produkt): Für eine Menge X heißt $X^0:=\mathrm{Abb}(\varnothing,X)$ das *leere Produkt*. X^0 besteht nach Proposition I.5.2 aus einem Element.

Definition I.5.15 (Permutation): Sei X eine Menge. Ist $f\colon X\to X$ eine bijektive Abbildung, dann heißt f auch *Permutation*. Mit

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

bezeichnen wir die Menge der Permutationen von X.

Ist X eine Menge mit n Elementen, dann ist $\#\,\mathrm{Perm}(X)=n!:=\prod_{i=1}^{n} i$.

6. Relationen

Definition I.6.1 (Relation): Sei M eine Menge. Eine Teilmenge $R \subseteq M \times M$ heißt *zweistellige Relation* oder kurz *Relation auf M*. Statt $(x, y) \in R$ schreibt man auch xRy.[3]

Beispiel I.6.2: (i) Seien M eine Menge und $R_1 := \{(x, y) \in M^2 \mid x = y\}$. Die Relation R_1 heißt *Gleichheitsrelation*. Es gilt xR_1y genau dann, wenn $x = y$.

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

$$R_2 := \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}, \quad R_3 := \{(x, y) \in \mathbb{R}^2 \mid x < y\},$$
$$R_4 := \{(x, y) \in \mathbb{R}^2 \mid x \geq y\},$$
$$R_5 := \{(x, y) \in \mathbb{R}^2 \mid x > y\}, \quad R_6 := \{(x, y) \in \mathbb{R}^2 \mid x \neq y\}.$$

(iii) Ist M die Menge der Studenten dieser Vorlesung, dann ist

$$R_7 := \{(s_1, s_2) \in M^2 \mid s_1 \text{ und } s_2 \text{ haben dasselbe Geburtsdatum}\}$$

eine Relation auf M.

Definition I.6.3 (Eigenschaften von Relationen): Seien M eine Menge und R eine Relation auf M.

(i) Gilt für alle $x \in M$, dass xRx, dann heißt R *reflexiv*.

(ii) Gilt für alle $x, y \in M$ mit xRy, dass yRx, dann heißt R *symmetrisch*.

(iii) Gilt für alle $x, y \in M$ mit xRy und yRx, dass $x = y$, dann heißt R *antisymmetrisch*.

(iv) Gilt für alle $x, y, z \in M$ mit xRy und yRz, dass xRz, dann heißt R *transitiv*.

Beispiel I.6.4 (Eigenschaften der Beispielrelationen): Für die Relationen in Beispiel I.6.2 haben wir die folgende Tabelle:

	R_1	R_2	R_3	R_4	R_5	R_6	R_7
reflexiv	✓	✓	—	✓	—	—	✓
symmetrisch	✓	—	—	—	—	✓	✓
antisymmetrisch	✓	✓	✓	✓	✓	—	—
transitiv	✓	✓	✓	✓	✓	—	✓

[3] Oft werden Relationen mit Symbolen wie „$\sim$“ oder „$\leqslant$“ belegt, und dann sieht „xRy“ nicht mehr so seltsam aus.

Definition I.6.5 (Äquivalenz und Ordnungsrelationen): Seien M eine Menge und $R \subseteq M^2$ eine Relation.

(i) Ist R reflexiv, symmetrisch und transitiv, dann heißt R eine *Äquivalenzrelation.*

(ii) Ist R reflexiv, antisymmetrisch und transitiv, dann heißt R eine *Ordnungsrelation.*

Äquivalenzrelationen werden oft mit „$\sim$“ notiert, Ordnungsrelationen werden oft mit „$\leqslant$“ notiert.

Beispiel I.6.6 (Kongruenzrelation): Sei n eine natürliche Zahl. Die durch

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

gegebene Relation heißt *Kongruenz modulo n*. Statt aR_sb schreiben wir auch $a \equiv b \pmod{n}$ oder $a \equiv_n b$. Beispielsweise ist $1 \equiv 11 \pmod{5}$, jedoch ist $11 \not\equiv 21 \pmod{4}$.

Proposition I.6.7: *Sei n eine natürliche Zahl. Kongruenz modulo n ist eine Äquivalenzrelation.*

Beweis: Wir haben drei Punkte zu zeigen: Erstens, dass die Relation reflexiv ist; zweitens, dass die Relation symmetrisch ist und drittens, dass die Relation transitiv ist.

(1) Für irgendeine ganze Zahl x gilt $x - x = 0$, und weil 0 durch jede ganze Zahl teilbar ist, ist Kongruenz modulo n reflexiv.

(2) Seien x und y ganze Zahlen, deren Differenz durch n teilbar ist, d. h., es gibt eine ganze Zahl k, sodass $x - y = kn$. Wegen $y - x = -kn$ wird $y - x$ auch von n geteilt und damit gilt $y \equiv_n x$.

(3) Seien x, y und z ganze Zahlen, sodass $x \equiv_n y$ und $y \equiv_n z$. Das heißt es gibt ganze Zahlen k und ℓ, sodass $x - y = kn$ und $y - z = \ell n$. Dann ist

$$x - z = x - y + y - z = kn + \ell n = (k + \ell)n,$$

also gilt $x \equiv_n z$. □

Bemerkung I.6.8 (Restklassen): Sei n eine natürliche Zahl. Die Kongruenzrelation „$\equiv_n$“ definiert auf $\mathbb{Z}$ folgende Zerlegung in n Mengen: Für $0 \leq i \leq n - 1$ setzen wir

$$M_i := \{x \in \mathbb{Z} \mid x \equiv_n i\} = \{i + kn \mid k \in \mathbb{Z}\} =: [i].$$

Diese sind genau die Restklassen bezüglich „Teilen mit Rest durch n“.

Beispiel I.6.9 (Gerade und ungerade): Für $n = 2$ sind die Restklassen M_0 und M_1 genau die Mengen der geraden bzw. ungeraden Zahlen in $\mathbb{Z}$.

Auf einer Menge M definiert jede Äquivalenzrelation eine Zerlegung von M in Teilmengen – besser noch, in disjunkte nichtleere Teilmengen. Wir werden sehen, dass jede Zerlegung von M in disjunkte nichtleere Teilmengen tatsächlich von einer Äquivalenzrelation herrührt.

Definition I.6.10 (Äquivalenzklasse): Seien M eine Menge und „$\sim$“ eine Äquivalenzrelation auf M. Für $x \in M$ heißt

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

die *Äquivalenzklasse von x bezüglich „$\sim$“*. Ist aus dem Kontext klar, von welcher Äquivalenzrelation die Rede ist, schreiben wir auch kurz $[x]$ für die Äquivalenzklasse von x.

Proposition I.6.11 (Eigenschaften von Äquivalenzklassen): *Seien M eine Menge und „$\sim$“ eine Äquivalenzrelation auf M. Seien ferner x und y Elemente von M. Dann gilt:*

(i) *Das Element x gehört zu $[x]$.*

(ii) *Es gilt $x \sim y$ genau dann, wenn $[x] = [y]$.*

Beweis: (i) Weil „$\sim$“ als Äquivalenzrelation reflexiv ist, steht x in Relation zu x, d. h. x gehört zur Äquivalenzklasse von x.

(ii) „$\Longrightarrow$“: Wir nehmen an, dass x in Relation zu y steht. Für ein Element $z \in [x]$ gilt $z \sim x$ und wegen $x \sim y$ haben wir auch $z \sim y$, d. h. z liegt in $[y]$, womit wir $[x] \subseteq [y]$ erhalten. Durch Vertauschen der Rollen von x und y erhalten wir auch $[y] \subseteq [x]$.

„$\Longleftarrow$“: Wir nehmen an, dass $[x] = [y]$. Aus $[x] = [y]$ erhalten wir mit (i), dass $y \in [y] = [x]$, aber per Definition heißt das ja gerade „$x \sim y$“. □

Notation I.6.12 (Schnitt und Vereinigung über Indexmenge): Sei I eine Menge. Für jedes $i \in I$ sei eine Menge M_i gegeben. Dann heißen

$$\bigcap_{i \in I} M_i := \{x \mid \text{Für jedes } i \in I \text{ gilt } x \in M_i\},$$

$$\bigcup_{i \in I} M_i := \{x \mid \text{Es gibt } i \in I, \text{ sodass } x \in M_i\}$$

der *Schnitt* beziehungsweise die *Vereinigung der Mengen M_i*.

Satz 1: *Seien M eine Menge und „$\sim$“ eine Äquivalenzrelation auf M. Dann gilt:*

(i) *Alle Äquivalenzklassen sind nichtleer, d. h. für jedes $x \in M$ gilt $[x] \neq \varnothing$.*

(ii) *M ist die Vereinigung seiner Äquivalenzklassen, d. h. $M = \bigcup_{x \in M} [x]$.*

(iii) *Je zwei verschiedene Äquivalenzklassen sind disjunkt, d. h. für irgendwelche $x, y \in M$ gilt $[x] = [y]$ oder $[x] \cap [y] = \varnothing$.*

Beweis: (i) Sei x irgendein Element von M. Nach Proposition I.6.11(i) gehört x zu $[x]$.

(ii) „$\subseteq$“: Sei y ein Element von M. Dann gehört y nach (i) zu $[y]$, d. h. $y \in \bigcup_{x \in M} [x]$.
„$\supseteq$“: Sei y ein Element von $\bigcup_{x \in M} [x]$. Dann gibt es ein $x \in M$, sodass $y \in [x]$. Wegen $[x] \subseteq M$ folgt daraus $y \in M$.

(iii) Seien x und y Elemente von M mit $[x] \cap [y] \neq \varnothing$. Weil $[x] \cap [y]$ nichtleer ist, gibt es irgendein $z \in [x] \cap [y]$. Per Definition haben wir sowohl $x \sim z$ als auch $y \sim z$, wegen Symmetrie und Transitivität von „$\sim$“ ist dann $x \sim y$, d. h. $[x] = [y]$ nach Proposition I.6.11. □

Definition I.6.13 (Menge der Äquivalenzklassen): Seien M eine Menge und „$\sim$“ eine Äquivalenzrelation auf M.

(i) Die Menge $M/\sim := \{[x] \mid x \in M\}$ heißt *Menge der Äquivalenzklassen von M bezüglich „$\sim$“.*

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

Beispiel I.6.14 (Kongruenzrelation): Seien $M = \mathbb{Z}$ und $\sim$ die Kongruenzrelation modulo 5. Dann ist $M/\sim = \{[0], [1], [2], [3], [4]\}$ und

$$\pi \colon M \longrightarrow M/\sim, \qquad z \longmapsto [z]$$

ist die kanonische Projektion.

Definition I.6.15 (Partition): Seien M eine Menge und $P \subseteq \mathfrak{P}(M)$. Gilt

(i) Die leere Menge ist kein Element von P,

(ii) Die Vereinigung $\bigcup_{A \in P} A$ ist ganz M,

(iii) Für A und B aus P mit $A \neq B$ gilt $A \cap B = \varnothing$,

dann heißt P eine *Partition von* M.

Korollar I.6.16 (aus Satz 1): *Sind M eine Menge, $\sim$ eine Äquivalenzrelation auf M und $M/\sim$ die Menge der Äquivalenzklassen, dann ist $M/\sim$ eine Partition.*

Satz 2 (Partition definiert Äquivalenzrelation): *Seien M eine Menge und S eine Partition von M. Wir definieren mit S eine Relation „$\sim$“ auf M durch*

$$x \sim y :\Longleftrightarrow \exists\, A \in S : (x \in A \wedge y \in A),$$

d. h. x steht in Relation zu y, wenn beide im selben Element der Partition liegen. Es handelt sich bei „$\sim$“ um eine Äquivalenzrelation.

Beweis: Wieder haben wir die drei Eigenschaften einer Äquivalenzrelation nachzuweisen. Dazu verwenden wir die Eigenschaften (i), (ii) und (iii) einer Partition aus Definition I.6.15.

(1) Sei x ein Element von M. Da S eine Partition von M ist, gibt es nach (ii) ein Element A von S, sodass $x \in A$. Per Definition der Relation „$\sim$“ bedeutet das gerade $x \sim x$, d. h. „$\sim$“ ist reflexiv.

(2) Seien x und y Elemente von M sodass $x \sim y$. Per Definition der Relation gibt es dann $A \in S$, sodass x und y Elemente von A sind. Wiederum per Definition der Relation heißt das dann auch $y \sim x$.

(3) Seien x, y und z Elemente von M, sodass $x \sim y$ und $y \sim z$. Per Definition der Relation gibt es dann A und B in S, sodass x und y zu A und y und z zu B gehören. Insbesondere gehört y zu $A \cap B$. Nach Eigenschaft (iii) aus Definition I.6.15 müssen dann aber schon A und B übereinstimmen, d. h. x und z gehören beide zu A und damit gilt $x \sim z$. □

Bemerkung I.6.17: Die Konstruktionen aus Satz 1 und Satz 2 sind zueinander invers: Äquivalenzrelationen auf M und Partitionen von M entsprechen sich bijektiv.

7. Nachtrag und Ausblick

Eine formalere Einführung in die Mengenlehre findet man z. B. in {Referenz Dieser, Ebbinghaus }. Die Mengentheorie baut auf Axiomen auf, d. h. es werden Regeln definiert die gelrten sollen. Es gibt unterschiedliche Axiomensysteme,

am weitesten verbreitet ist das ZFC-Axiomensystem[4]. Zu den ZFC-Axiomen gehören beispielsweise

- *Extensionalitätsaxiom*: Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente haben;
- *Aussonderungsaxiom*: Prädikate P definieren Mengen, genauer: Sind P ein Prädikat und A eine Menge, dann gibt es eine Teilmenge $B \subseteq A$, die genau die Elemente x von A enthält, für die $P(x)$ wahr ist;
- *Leermengenaxiom*: Es gibt eine Menge ohne Elemente;
- *Auswahlaxiom*: Ist A eine Familie nichtleerer Mengen, dann gibt es eine Funktion $f\colon A \to \bigcup_{B \subseteq A} B$, die jedem Element B von A ein Element aus B zuordnet, also „ein Element von B auswählt“;
- *Fundierungsaxiom*: Jede nichtleere Menge A enthält ein Element B, sodass A und B disjunkt sind.

Das Auswahlaxiom folgt nicht aus den anderen Axiomen. Das ZF-Axiomensystem lässt das Auswahlaxiom weg. Manche Aussagen können deshalb darin nicht bewiesen werden (wie zum Beispiel die Aussage, die in einer Teilaufgabe zu Umkehrfunktionen mithilfe des Auswahlaxioms gezeigt werden sollte).

Die hier vorgestellte naive Mengenlehre birgt gewisse logische Schwierigkeiten, sogenannte *Antinomien*. Versucht man die Frage „Gibt es Mengen, die sich selbst enthalten?“ ist es naheliegend, die Menge aller Mengen, die sich selbst nicht enthalten, d. h. die Menge

$$M := \{A \text{ Menge} \mid A \notin A\}$$

zu betrachten. Gehört M nun zu M oder nicht? Wäre M kein Element von M, dann müsste M per Definition zu M gehören, was unmöglich ist. Wäre aber M ein Element von M, dann würde M per Definition nicht zu M gehören. So eine Menge M kann es also nicht geben! Das Fundierungsaxiom in ZFC löst dieses Problem.

Jedoch bleiben auch nach der Einführung des ZFC-Axiomensystem Schwierigkeiten bestehen. Der österreichische Mathematiker Kurt Gödel zeigte 1931 in seinen zwei Unvollständigkeitssätzen, dass es einerseits unbeweisbare Aussagen in jedem hinreichend komplexen Axiomensystem gibt, und dass andererseits hinreichend starke widerspruchsfreie Systeme ihre eigene Widerspruchsfreiheit

[4]Hierbei steht „Z“ für den deutschen Mathematiker Ernst Zermelo (1871-1953), das „F“ für den deutsch-israelischen Mathematiker Adolf Abraham Haleri Fraenkel (1891-1965) und das „C“ für das Auswahl-Axiom (englisch: Axiom of Choice).

nicht zeigen können. Und tatsächlich muss man sich noch nicht zu sehr anstrengen, um ein Beispiel für eine innerhalb von ZFC unbeweisbare Aussage zu geben. Für Interessierte sei auf den deutschen Wikipedia-Artikel zur Calkin-Algebra verwiesen.

Kapitel II.

Lineare Gleichungssysteme und reelle Vektorräume

Im Folgenden setzen wir die reellen Zahlen samt ihren Verknüpfungen (d. h. Addition und Multiplikation) und die zugehörigen Rechenregeln als bekannt voraus.

In dieser Vorlesung verstehen wir Elemente des $\mathbb{R}^n$, sofern nicht ausdrücklich anders angegeben, als „Spaltenvektoren“, das heißt

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}.$$

Um nicht zu liberal mit dem digitalen Papier umzugehen, markieren wir einen „Zeilenvektor“ mit einem „^{t}“, um zu verdeutlichen, wenn wir in Wahrheit einen Spaltenvektor meinen. So ist mit $(x_1, \ldots, x_n)^t$ ein Element des $\mathbb{R}^n$ gemeint. Diese Notation wird sich später als sinnvoll herausstellen.

1. Vom linearen Gleichungssystem zum Vektorraum

Beispiel II.1.1: Seien x_1, x_2 und x_3 „Variablen“ oder auch „Unbestimmte“. Wir wollen diejenigen Werte für diese Variablen finden, für die die Gleichungen

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

erfüllt sind, d. h. wir versuchen die Menge

$$\mathbb{L} = \{x = (x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x \text{ erfüllt die drei Gleichungen}\}$$

zu bestimmen. Dazu haben wir unterschiedliche Ansätze.

(i) Das lineare Gleichungssystem ist durch die Daten

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

bestimmt. Dabei heißt A die *Koeffizientenmatrix* und b die *rechte Seite*. Wir notieren das lineare Gleichungssystem als die Matrix $(A|b)$, die entsteht, wenn wir b als vierte Spalte neben A schreiben. Wir versuchen mithilfe von Matrizen- und Vektorrechnung die Lösungsmenge zu bestimmen.

(ii) *Reduktion auf homogenes lineares Gleichungssystem*: Sind $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^t$ und $\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)$ Lösungen des linearen Gleichungssystems Gl. (II.1), dann gilt

$$\begin{aligned} 2(\tilde{x}_1 - \hat{x}_1) + 6(\tilde{x}_2 - \hat{x}_2) + 4(\tilde{x}_3 - \hat{x}_3) &= \quad 8 - 8 = 0 \\ (\tilde{x}_2 - \hat{x}_2) - 2(\tilde{x}_3 - \hat{x}_3) &= \quad 6 - 6 = 0 \\ (\tilde{x}_1 - \hat{x}_1) + 4(\tilde{x}_2 - \hat{x}_2) &= 10 - 10 = 0 \end{aligned}$$

d. h. $v := \tilde{x} - \hat{x} = (\tilde{x}_1 - \hat{x}_1, \tilde{x}_2 - \hat{x}_2, \tilde{x}_3 - \hat{x}_3)^t$ ist eine Lösung des linearen Gleichungssystems

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

Das Gleichungssystem aus Gl. (II.2) heißt das zu Gl. (II.1) gehörige *homogene lineare Gleichungssystem*. Es hat dieselbe linke Seite wie das ursprüngliche lineare Gleichungssystem, jedoch die spezielle rechte Seite $(0, 0, 0)^t$.

Sind umgekehrt $\tilde{x}$ und $\hat{x}$ Elemente des $\mathbb{R}^3$, sodass $\hat{x}$ eine Lösung von Gl. (II.1) und $v := \tilde{x} - \hat{x}$ eine Lösung von Gl. (II.2) ist, dann ist auch $\tilde{x} = \hat{x} + v$ eine Lösung von Gl. (II.1).

Anders ausgedrückt: Bezeichnen wir mit $\mathbb{L}$ die Lösungsmenge von Gl. (II.1), mit $\mathbb{L}_h$ die Lösungsmenge von Gl. (II.2) und mit $\hat{x} \in \mathbb{R}^3$ eine „spezielle Lösung“ von II.1, dann gehört $\tilde{x} \in \mathbb{R}^3$ zu $\mathbb{L}$ genau dann, wenn $v := \tilde{x} - \hat{x}$ zu $\mathbb{L}_h$ gehört. Dies ist äquivalent zur Aussage: Ein $\tilde{x} \in \mathbb{R}^3$ gehört zu $\mathbb{L}$ genau dann, wenn es $v \in \mathbb{L}_h$ gibt, sodass $\tilde{x} = \hat{x} + v$. Wir können also schreiben

$$\mathbb{L} = \{\tilde{x} := \hat{x} + v \mid v \in \mathbb{L}_h\} =: \hat{x} + \mathbb{L}_h.$$

Um das Gleichungssystem Gl. (II.1) zu lösen genügt es also, eine Lösung von Gl. (II.1) und alles Lösungen von Gl. (II.2) zu kennen.

(iii) Mit ähnlichen Rechnungen wie in (ii) können wir etwas über die Struktur von $\mathbb{L}_h$ lernen. Seien dazu $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^t$ und $\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)^t$ Elemente des $\mathbb{R}^3$, die Gl. (II.2) lösen. Bezeichnet r eine reelle Zahl, dann gilt

$$\begin{aligned} 2(\tilde{x}_1 + \hat{x}_1) + 6(\tilde{x}_2 + \hat{x}_2) + 4(\tilde{x}_3 + \hat{x}_3) &= 0 \\ (\tilde{x}_2 + \hat{x}_2) - 2(\tilde{x}_3 + \hat{x}_3) &= 0 \\ (\tilde{x}_1 + \hat{x}_1) + 4(\tilde{x}_2 + \hat{x}_2) &= 0 \end{aligned}$$

sowie

$$\begin{aligned} 2(r \cdot \tilde{x}_1) + 6(r \cdot \tilde{x}_2) + 4(r \cdot \tilde{x}_3) &= 0 \\ (r \cdot \tilde{x}_2) - 2(r \cdot \tilde{x}_3) &= 0 \\ (r \cdot \tilde{x}_1) + 4(r \cdot \tilde{x}_2) &= 0 \end{aligned}$$

d. h. wir haben: Sind $\tilde{x}$ und $\hat{x}$ Elemente von $\mathbb{L}_h$, dann ist auch $\tilde{x} + \hat{x}$ ein Element von $\mathbb{L}_h$, und sind $\tilde{x} \in \mathbb{L}_h$ sowie r eine reelle Zahl, dann gehört auch $r\tilde{x}$ zu $\mathbb{L}_h$.

Bemerkung II.1.2: Nichtleere Teilmengen des $\mathbb{R}^n$, die die beiden Eigenschaften aus (iii) erfüllen, sind „Untervektorräume“ des $\mathbb{R}^n$. Im Folgenden geben wir die allgemeine Einführung von Vektorräumen.

2. Vektorräume

Das Ziel dieses Abschnittes ist die Verallgemeinerung des $\mathbb{R}^n$ zusammen mit seiner Addition und seiner Skalarmultiplikation zum allgemeinen Konzept des Vektorraums über einem Körpers.

Die Struktur des $\mathbb{R}^n$ wird durch die Addition, die Skalarmultiplikation und die Rechenregeln, die für diese Verknüpfungen gelten, bestimmt. Zur Erinnerung: Für Elemente $x = (x_1, \ldots, x_n)^t$ und $y = (y_1, \ldots, y_n)^t$ des $\mathbb{R}^n$ und eine reelle Zahl r sind

$$x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \qquad rx = \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}.$$

Für $x, y, z \in \mathbb{R}^n$ und eine reelle Zahl r sind folgende Regeln aus der Schule bekannt:

(i) Assoziativität: Es ist $(x + y) + z = x + (y + z)$,

(ii) Distributivität: Es ist $r(x + y) = rx + ry$,

(iii) Kommutativität: Es ist $x + y = y + x$.

Um diese Regeln verallgemeinern zu können, müssen wir die Verknüpfungen als Abbildungen fassen: Wir haben die Abbildungen

$$+\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, \qquad (x,y) \longmapsto x+y,$$
$$\cdot\colon \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, \qquad (r,x) \longmapsto rx$$

die den oben gesammelten Regeln genügen – und einige mehr.

Definition II.2.1: Seien V eine Menge und $\mathbf{0} = \mathbf{0}_V$ ein ausgezeichnetes Element von V, der sogenannte Nullvektor. Gibt es Abbildungen

$$+\colon V \times V \longrightarrow V, \qquad (v,w) \longmapsto +(v,w) =: v+w,$$
$$\cdot\colon \mathbb{R} \times V \longrightarrow V, \qquad (r,v) \longmapsto \cdot(r,v) =: r\cdot v =: rv,$$

genannt *Vektoraddition* bzw. *Addition* und *Skalarmultiplikation*, die die Regeln

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

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

(A3) Für alle $x \in V$ ist $x+\mathbf{0} = x = \mathbf{0}+x$,

(A4) Für alle $x \in V$ gibt es genau ein $y \in V$, sodass $x+y = \mathbf{0} = y+x$,

und

(S1) Für alle $x \in V$ ist $1x = x$,

(S2) Für alle $r,s \in \mathbb{R}$ und $x \in V$ gilt $(r+s)x = rx+sx$,

(S3) Für alle $r \in \mathbb{R}$ und $x,y \in V$ gilt $r(x+y) = rx+ry$,

(S4) Für alle $r,s \in \mathbb{R}$ und $x \in V$ gilt $r\cdot(sx) = (r\cdot s)x$,

erfüllen, dann heißt V ein *reeller Vektorraum*, *Vektorraum über* $\mathbb{R}$ oder $\mathbb{R}$-*Vektorraum*. Wir notieren den $\mathbb{R}$-Vektorraum V als Viertupel $(V,+,\cdot,\mathbf{0}_V)$.

Für das eindeutige Element y zu x mit $x+y = \mathbf{0} = y+x$ schreibt man auch $-x$ und nennt es *additives Inverses von* x.

Beispiel II.2.2 ($\mathbb{R}^n$ und $\mathbf{Abb}(M,\mathbb{R})$): (i) Der $\mathbb{R}^n$ zusammen mit den komponentenweisen Verknüpfungen (Addition und Skalarmultiplikation) und dem Nullvektor $\mathbf{0}_V = (0,\dots,0)^t$ ist ein $\mathbb{R}$-Vektorraum.

(ii) Sei M eine Menge. Dann wird

$$V := \mathbb{R}^M := \mathrm{Abb}(M,\mathbb{R}) = \{f\colon M \to \mathbb{R} \mid f \text{ ist Abbildung}\}$$

zusammen mit dem Nullvektor $\mathbf{0}_V := (f\colon M \to \mathbb{R},\ m \mapsto 0)$ und den Verknüpfungen

$$\oplus\colon V \times V \longrightarrow V, \qquad (f_1, f_2) \longmapsto (g\colon M \to \mathbb{R}, \quad m \mapsto f_1(m) + f_2(m)),$$

d. h. die Abbildung $g := f_1 \oplus f_2$ ist gegeben durch die Eigenschaft „Für alle $m \in M$ ist $g(m) = f_1(m) + f_2(m)$“, und

$$\odot\colon \mathbb{R} \times V \longrightarrow V, \qquad (r, f) \longmapsto (g\colon M \to \mathbb{R}, \quad m \mapsto rf(m))$$

zu einem Vektorraum über $\mathbb{R}$.

(iii) Seien M eine Menge und W ein beliebiger $\mathbb{R}$-Vektorraum. Dann wird

$$V := W^M := \mathrm{Abb}(M, V) = \{f\colon M \to W \mid f \text{ ist eine Abbildung}\}$$

zusammen mit dem Nullvektor $\mathbf{0}_V := (f\colon M \to W,\ m \mapsto \mathbf{0}_W)$ und den Verknüpfungen analog definiert zu den Verknüpfungen in (ii) zu einem Vektorraum über $\mathbb{R}$.

Beweis: Die Vektorraumaxiome aus Definition II.2.1 gelten jeweils, weil die entsprechenden Rechenregeln in $\mathbb{R}$ beziehungsweise in W gelten. Exemplarisch zeigen wir die Gültigkeit von (A3) für (iii). Wir haben also zu zeigen, dass für alle Abbildungen $f\colon M \to W$ gilt, dass $f \oplus \mathbf{0}_V = f = \mathbf{0}_V + f$. Diese Gleichheit von Abbildungen zeigen wir, indem wir zeigen, dass die Abbildungen auf allen Elementen von M dieselbe Wirkung haben. Für irgendein $m \in M$ ist

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

wegen der Definition von $\mathbf{0}_V$ und der Gültigkeit von (A3) für den $\mathbb{R}$-Vektorraum W. Genau so rechnet man nach, dass $\mathbf{0}_V \oplus f = f$.

Für die restlichen Regeln zeigen ähnliche Rechnungen deren Gültigkeit und bleiben deshalb zur Eigenübung. □

Proposition II.2.3: *Es sei $(V, +, \cdot, \mathbf{0}_V)$ ein Vektorraum über $\mathbb{R}$. Dann gilt:*

(i) *Für alle $v, w \in V$ gibt es genau ein $x \in V$, sodass $v + x = w$. Für dieses x schreiben wir $w - v$. Insbesondere ist $\mathbf{0}_V$ das einzige Element, das* (A3) *erfüllt.*

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

(iii) *Für alle $r \in \mathbb{R}$ und $v \in V$ ist $r(-v) = (-r)v$.*

(iv) *Für alle $r \in \mathbb{R}$ und $v, w \in V$ ist $r(v - w) = rv - rw$.*

(v) *Für alle* $r, s \in \mathbb{R}$ *und* $v \in V$ *ist* $(r - s)v = rv - sv$.

Beweis: Wir zeigen unter Verwendung von (i) exemplarisch, dass $0v = \mathbf{0}_V$. Aus (i) wissen wir, dass $\mathbf{0}_V$ das eindeutige Element ist, für das gilt: $0v + \mathbf{0}_V = \mathbf{0}_V$. Wegen (A3) gilt $0v + \mathbf{0}_V = \mathbf{0}_V$. Weil (S2) garantiert, dass $0v + 0v = (0+0)v = 0v$, liefert uns (i), dass $0v = \mathbf{0}_V$.

Die restlichen Aussagen werden auf dem vierten Übungsblatt und in der Präsenzübung gezeigt. □

Bemerkung II.2.4: Mit den Notationen aus Proposition II.2.3 gilt insbesondere:

(i) Für alle $v \in V$ ist $-v = (-1)v$,

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

Aussage (i) ist ein Spezialfall der Aussage (iii) aus Proposition II.2.3, das ist also klar. Für die zweite Aussage müssen wir uns an die Definition von $w - v$ erinnern, um zu sehen, was zu zeigen ist. Wegen

$$\begin{aligned} v + (w + (-v)) &= (v + w) + (-v) && \text{(A1)} \\ &= (w + v) + (-v) && \text{(A2)} \\ &= w + (v + (-v)) && \text{(A1)} \\ &= w + \mathbf{0}_V = w \end{aligned}$$

ist $w + (-v) = w - v$.

Beispiel II.2.5: Sei V der $\mathbb{R}$-Vektorraum $\mathbb{R}^2$. Wir betrachten die Teilmengen $U_1 := \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ und $U_2 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

Definition II.2.6 (Untervektorraum): Seien V ein Vektorraum und $U \subseteq V$ eine Teilmenge. Falls gilt:

(i) Der Nullvektor $\mathbf{0}_V$ gehört zu U,

(ii) Für v und w aus U gehört auch $v + w$ zu U,

(iii) Für $r \in \mathbb{R}$ und v aus U gehört auch rv zu U,

dann heißt U ein *Untervektorraum von* V.

Bemerkung II.2.7: Sei V ein Vektorraum.

(i) Die Teilmengen $\{\mathbf{0}_V\}$ und V sind Untervektorräume von V. Sie heißen *triviale Unterräume*.

(ii) Die leere Menge ist kein Untervektorraum von V.

(iii) Fordern wir von U in Proposition II.2.6, nichtleer zu sein, dann folgt (i) bereits aus (iii).

Proposition II.2.8 (Untervektorräume sind Vektorräume): *Sind V ein Vektorraum und $U \subseteq V$ ein Untervektorraum, dann gilt insbesondere: Die Verknüpfungen „$+$“ und „$\cdot$“ auf V schränken sich zu Verknüpfungen auf U ein, genauer: Die Abbildungen*

$$+|_{U\times U}\colon U \times U \longrightarrow U, \qquad (x,y) \longmapsto x+y,$$
$$\cdot|_{\mathbb{R}\times U}\colon \mathbb{R} \times U \longrightarrow U, \qquad (r,x) \longmapsto rx$$

sind wohldefiniert (d. h. ihre Bilder sind jeweils in U enthalten) und U wird mit diesen Verknüpfungen sowie dem Nullvektor $\mathbf{0}_V$ zu einem Vektorraum.

Beweis: Dass die Bilder der Einschränkungen der Verknüfungen auf V wirklich in U enthalten sind, folgt aus den Punkten (ii) und (iii) in Proposition II.2.6.

Da die Axiome (A1)-(A3) und (S1)-(S4) sogar für alle Elemente von V gelten, gelten sie erst recht für die Elemente von U. Bleibt zu zeigen, dass auch (A4) gilt, d. h. wir müssen zeigen: Für alle $x \in U$ gibt es genau ein $y \in U$ mit $x + y = \mathbf{0}_V = y + x$.

Sei also $x \in U$ gegeben. Weil V ein Vektorraum ist, gibt es genau ein Element $-x$ in V, das das Gewünschte leistet. Aus Proposition II.2.3 wissen wir, dass $-x = (-1)x$ und wegen Punkt (iii) aus Proposition II.2.6 wissen wir damit, dass $-x$ zu U gehört. □

Proposition II.2.9 (Direktes Produkt endlich vieler Vektorräume): *Seien n eine natürliche Zahl und $V_1, \ldots, V_n$ Vektorräume über $\mathbb{R}$. Ihr kartesisches Produkt*

$$W := V_1 \times \cdots \times V_n = \{(v_1, \ldots, v_n) \mid v_1 \in V_1, \ldots, v_n \in V_n\}$$

wird mit den Verknüpfungen, die für $v_1, v_1' \in V_1$, …, $v_n, v_n' \in V_n$ und $r \in \mathbb{R}$ durch

$$(v_1, \ldots, v_n) \oplus (v_1', \ldots, v_n') := (v_1 + v_1', \ldots, v_n + v_n'),$$
$$r \odot (v_1, \ldots, v_n) := (rv_1, \ldots, rv_n)$$

erklärt sind, und mit geeignet gewähltem Nullelement zu einem Vektorraum über $\mathbb{R}$. Dieser heißt das direkte Produkt der Vektorräume $V_1, \ldots, V_n$.

Beweis: Die Aussage lässt sich analog zu Aufgabe 4 auf Blatt 1 zeigen. □

Proposition II.2.10 (Beliebiger Schnitt von Untervektorräumen): *Seien I eine nichtleere Menge und für jedes $i \in I$ sei ein Untervektorraum U_i des Vektorraums V gegeben. Dann ist $W := \bigcap_{i\in I} U_i$ ebenfalls ein Untervektorraum von V.*

Beweis: Wir prüfen die Punkte aus Proposition II.2.6.

(i) Weil jeder der Untervektorräume U_i den Nullvektor $\mathbf{0}_V$ enthält, liegt $\mathbf{0}_V$ per Definition des Schnitts in W.

(ii) Seien v und w aus W gegeben. Per Definition von W gehören v und w dann zu jedem Vektorraum U_i, d. h. für jedes $i \in I$ gehört $v + w$ zu U_i. Wiederum per Definition des Schnitts folgt daraus $v + w \in W$.

(iii) Seien r eine reelle Zahl und $v \in W$ gegeben. Per Definition gehört v zu jedem Vektorraum U_i und deshalb gehört auch rv zu jedem Vektorraum U_i. Wieder erhalten wir $rv \in W$ und wir sind fertig. □

Proposition II.2.11 (Summe endlich vieler Untervektorräume): *Seien V ein Vektorraum über $\mathbb{R}$, n eine natürliche Zahl und $U_1, \ldots, U_n$ Untervektorräume von V. Dann ist*

$$W := \{v_1 + \cdots + v_n \mid v_1 \in U_1, \ldots, v_n \in U_n\}$$

ebenfalls ein Untervektorraum von V. Er wird die Summe von $U_1, \ldots, U_n$ *genannt.*

Beweis: Man überprüft ähnlich wie in Proposition II.2.10, dass die Punkte aus Proposition II.2.6 erfüllt sind. □

3. Der Vektorraum der Matrizen

Das Ziele dieses Abschnitts sind die Vorstellung der Matrizenrechnung und die Einsicht, dass die reellwertigen $p \times q$-Matrizen einen $\mathbb{R}$-Vektorraum bilden.

Notation: Seien a und b natürliche Zahlen, V ein Vektorraum über $\mathbb{R}$ und $f\colon \mathbb{N} \to V$ eine Funktion. Dann schreiben wir

$$\sum_{i=a}^{n} f(i) := \begin{cases} f(a) + f(a+1) + \cdots + f(b), & \text{falls } a \leq b, \\ 0, & \text{falls } a > b. \end{cases}$$

Seien allgemeiner I eine Indexmenge und $f\colon I \to X$ eine Funktion, die nur auf endlich vielen Elementen von I einen von $\mathbf{0}_V$ verschiedenen Wert annimmt. Es bezeichne $J := \{i \in I \mid f(i) \neq \mathbf{0}_V\} \subseteq I$. Weil J endlich ist, gibt es eine natürliche Zahl n sodass $J = \{j_1, \ldots, j_n\}$ mit $j_1, \ldots, j_n \in I$. Dann schreiben wir

$$\sum_{i \in I} f(i) := \sum_{i=1}^{n} f(j_i).$$

Definition II.3.1 (Matrizen): Seien p und q natürliche Zahlen. Eine Abbildung

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

heißt *reelle $p \times q$-Matrix*. Dabei heißt p die *Anzahl der Zeilen* und q die *Anzahl der Spalten*. Man schreibt $a_{i,j} := A(i,j) := A((i,j))$ und notiert die Matrix A suggestiv als Schema

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

Gilt $p = q$, dann heißt die Matrix A *quadratisch*. Mit

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

bezeichnen wir die Menge der reellen $p \times q$-Matrizen. Auch die Notationen $\mathrm{Mat}(p \times q, \mathbb{R})$, $\mathrm{Mat}(p, q)$ oder $M_{p \times q}(\mathbb{R})$ sind gebräuchlich. Im Folgenden identifizieren wir stets $\mathbb{R}^p$ mit $\mathbb{R}^{p \times 1}$.

Bemerkung II.3.2: In Definition II.3.1 kann man auch $p = 0$ oder $q = 0$ zulassen. Die Mengen $\mathbb{R}^{0 \times q} = \mathbb{R}^{p \times 0} = \mathrm{Abb}(\varnothing, \mathbb{R})$ bestehen dann aus einem Element.

Beispiel II.3.3 (Nullmatrix, Einheitsmatrix): Seien p und q natürliche Zahlen.

(i) Die Matrix $A \in \mathbb{R}^{p \times q}$ gegeben durch $A(i,j) = 0$ für alle $1 \leq i \leq p$ und $1 \leq j \leq q$ heißt *Nullmatrix*. Die Notationen $\mathbf{0}$, $\mathbf{0}_{p \times q}$, N und $N_{p \times q}$ sind gebräuchlich.

(ii) Die quadratische Matrix $A \in \mathbb{R}^{p \times p}$ gegeben durch

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

heißt *Einheitsmatrix* oder auch *Einsmatrix*. Die Notationen I_p, E_p oder $\mathbb{1}_p$ sind gebräuchlich. Die Funktion $\delta_{i,j}$ heißt *Kronecker-Delta*.

Definition II.3.4 (Summe und skalare Multiplikation): Seien p und q natürliche Zahlen, r eine reelle Zahl und $A, B \in \mathbb{R}^{p\times q}$. Dann definieren wir

(i) $A + B := C \in \mathbb{R}^{p\times q}$ mit $C(i,j) = A(i,j) + B(i,j)$,

(ii) $rA := D \in \mathbb{R}^{p\times q}$ mit $D(i,j) := rA(i,j)$

für alle $1 \leq i \leq p$ und $1 \leq j \leq q$.

Proposition II.3.5 ($\mathbb{R}^{p\times q}$ als Vektorraum): *Seien p und q natürliche Zahlen. Dann wird $\mathbb{R}^{p\times q}$ mit den Verknüpfungen aus Definition II.3.4 und der Nullmatrix $\mathbf{0}_{p\times q}$ zu einem Vektorraum über $\mathbb{R}$.*

Beweis: Die Vektorraumaxiome aus Definition II.2.1 lassen sich jeweils komponentenweise aus den entsprechenden Vektorraumaxiomen für $\mathbb{R}$ nachrechnen. Exemplarisch zeigen wir (A4), also dass es für jede Matrix $A \in \mathbb{R}^{p\times q}$ genau eine Matrix $B \in \mathbb{R}^{p\times q}$ gibt, sodass $A + B = \mathbf{0}_{p\times q} = B + A$.

Sei dazu $A \in \mathbb{R}^{p\times q}$ gegeben. Wie eingangs angedeutet definieren wir B durch $B(i,j) := -A(i,j)$ für alle $1 \leq i \leq p$ und $1 \leq j \leq q$ und erhalten dann $A + B = \mathbf{0}_{p\times q} = B + A$.

Jetzt müssen wir noch begründen, dass wir in Wahrheit keine andere Wahl hatten. Sei also $a = A(i,j) \in \mathbb{R}$. Da $\mathbb{R}$ ein $\mathbb{R}$-Vektorraum ist, gibt es genau ein $b \in \mathbb{R}$ mit $a + b = 0 = b + a$, und dieses b ist $-a$. Damit ist B eindeutig mit der geforderten Eigenschaft. □

Definition II.3.6 (Matrizenmultiplikation): Seien p, q und m natürliche Zahlen und $A \in \mathbb{R}^{p\times q}$ sowie $B \in \mathbb{R}^{q\times m}$. Die Matrix $C \in \mathbb{R}^{p\times m}$, die gegeben ist durch

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

heißt *Produkt von A und B* und wird mit $A \cdot B$ oder AB bezeichnet.

Definition II.3.7 (Transponierte einer Matrix): Seien p und q natürliche Zahlen sowie $A \in \mathbb{R}^{p\times q}$ eine Matrix. Die Matrix $B \in \mathbb{R}^{q\times p}$ definiert durch

$$B(i,j) = A(j,i), \qquad (1 \leq i \leq q, 1 \leq j \leq p)$$

heißt die *transponierte Matrix zu A* oder kurz *Transponierte von A*. Gebräuchliche Notationen sind A^t, A^T oder $A^\top$.

Beispiel II.3.8: Für die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

haben wir

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

Auch die Notation „$(x_1, \ldots, x_n)^t$“ hat jetzt ihre Rechtfertigung.

Proposition II.3.9 (Weitere Rechenregeln für Matrizen): *Im Folgenden seien A, B und C Matrizen, sodass die folgenden Ausdrücke definiert sind. Es gelten folgende weitere Rechenregeln (ergänzend zu Proposition II.3.5) für Matrizen:*

(i) *Assoziativität:* $A(BC) = (AB)C$,

(ii) *Distributivität:* $A(B + C) = AB + AC$,

(iii) *Distributivität:* $(A + B)C = AC + BC$,

(iv) *Für jede reelle Zahl r gilt* $A(rB) = (rA)B = r(AB)$,

(v) $(A + B)^t = A^t + B^t$ *und* $(AB)^t = B^t A^t$,

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

Beweis: Wir beweisen exemplarisch die Aussage in (i): Seien p, q, m und n natürliche Zahlen und $A \in \mathbb{R}^{p\times q}$, $B \in \mathbb{R}^{q\times m}$ sowie $C \in \mathbb{R}^{m\times n}$. Für irgendwelche Indizes $i \in \{1, \ldots, p\}$ und $\ell \in \{1, \ldots, n\}$ gilt

$$\begin{aligned}
(A \cdot (B \cdot C))(i, \ell) &= \sum_{j=1}^{q} A(i, j) \cdot (B \cdot C)(j, \ell) \\
&= \sum_{j=1}^{q} A(i, j) \cdot \Big(\sum_{k=1}^{m} B(j, k) \cdot C(k, \ell) \Big) \\
&= \sum_{j=1}^{q} \sum_{k=1}^{m} A(i, j) \cdot B(j, k) \cdot C(k, \ell) \\
&= \sum_{k=1}^{m} \Big(\sum_{j=1}^{q} A(i, j) \cdot B(j, k) \Big) \cdot C(k, \ell) \\
&= \sum_{k=1}^{m} (A \cdot B)(i, k) C(k, \ell) = ((A \cdot B) \cdot C)(i, \ell),
\end{aligned}$$

d. h. die Produkte stimmen eintragsweise überein und sind damit gleich. □

Bemerkung II.3.10 (... keine Kommutativität!): Achtung: Matrizenmultiplikation ist im Allgemeinen nicht kommutativ! Seien $A = (1, 2, 3) \in \mathbb{R}^{1\times 3}$ und $B = (1, -1, 1)^t \in \mathbb{R}^{3\times 1}$. Für die beiden Produkte $AB \in \mathbb{R}^{1\times 1}$ und $BA \in \mathbb{R}^{3\times 3}$ erhalten wir

$$AB = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = (2), \qquad BA = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 1 & 2 & 3 \end{pmatrix}.$$

Aber nicht nur die Dimension „stört“ bei der Kommutativität, auch Multiplikation quadratischer Matrizen ist üblicherweise nicht kommutativ.

Beispiel II.3.11: Wir betrachten die Matrix

$$A = \begin{pmatrix} 13 & 27 & 16 \\ -3 & 2,6 & 5 \end{pmatrix}.$$

Diese Matrix A lässt sich schreiben als Summe von Vielfachen „einfacherer“ Matrizen:

$$\begin{aligned} A = 13 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 27 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 16 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ - 3 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + 2,6 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + 5 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Das wollen im Folgenden allgemein aufschreiben.

Definition II.3.12 (Elementarmatrix): Seien n und m natürliche Zahlen. Für natürliche Zahlen i und j mit $1 \leq i \leq n$ und $1 \leq j \leq m$ definieren wir die Matrix

$$E_{i,j} \colon \{1, \ldots, n\} \times \{1, \ldots, m\} \longrightarrow \mathbb{R}, \qquad (k, \ell) \longmapsto \delta_{i,k}\delta_{j,\ell}.$$

Die Matrizen $E_{1,1}$, $E_{1,2}$, ..., $E_{1,n}, E_{2,1}, \ldots, E_{n,m}$ heißen *Elementarmatrizen.*

Die Matrix $E_{i,j}$ enthält genau eine Eins (nämlich an der Stelle (i, j)) und alle anderen Einträge sind Null.

Bemerkung II.3.13 (Matrizen als Linearkombinationen der $E_{i,j}$): Für eine beliebige Matrix $A = (a_{i,j}) \in \mathbb{R}^{n\times m}$ gilt

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

wie wir uns das in Proposition II.3.11 schon plausibel gemacht haben.

Beispiel II.3.14: (i) Für die Elementarmatrizen

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

finden wir

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

(ii) Betrachten wir eine beliebige 2×2-Matrix $A = \left(\begin{smallmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ und die Elementarmatrix $E_{3,1}$ von oben erhalten wir

$$E_{3,1} \cdot A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ a_{1,1} & a_{1,2} \end{pmatrix},$$

d. h. die erste Zeile von A steht in der dritten Zeile des Produkts.

(iii) Für eine beliebige 2×3-Matrix $A = \left(\begin{smallmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{smallmatrix}\right) \in \mathbb{R}^{2\times 3}$ und die Elementarmatrix $E_{3,1}$ aus (ii) finden wir

$$A \cdot E_{3,1} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_{1,3} & 0 \\ a_{2,3} & 0 \end{pmatrix},$$

d. h. die dritte Spalte von A steht in der ersten Spalte des Produkts.

Proposition II.3.15 (Multiplikation mit Elementarmatrizen): *Seien p, q und r natürliche Zahlen.*

(i) *Für Elementarmatrizen $E_{i,j} \in \mathbb{R}^{p\times q}$ und $E_{k,\ell} \in \mathbb{R}^{q\times r}$ gilt*

$$E_{i,j}E_{k,\ell} = \delta_{j,k}E_{i,\ell} \in \mathbb{R}^{p\times r}.$$

(ii) *Für $E_{i,j} \in \mathbb{R}^{p\times q}$ und $A \in \mathbb{R}^{q\times r}$ gilt*

$$E_{i,j}A = \sum_{k=1}^{r} A(j,k)E_{i,k} \in \mathbb{R}^{p\times r},$$

d. h. die j-te Zeile von A steht in der i-ten Zeile des Produkts und die restlichen Einträge sind Null.

(iii) *Für $A \in \mathbb{R}^{p\times q}$ und $E_{k,\ell} \in \mathbb{R}^{q\times r}$ gilt*

$$AE_{k,\ell} = \sum_{j=1}^{p} A(j,k)E_{j,\ell}$$

d. h. die k-te Spalte von A steht in der ℓ-ten Spalte des Produkts und die restlichen Einträge sind Null.

Beweis: (i) Es bezeichne M das Produkt $E_{i,j}E_{k,\ell} \in \mathbb{R}^{p\times r}$. Für Indizes (a,b) in $\{1,\dots,p\} \times \{1,\dots,r\}$ gilt dann

$$M(a,b) = \sum_{x=1}^{q} E_{i,j}(a,x)E_{k,\ell}(x,b)$$

und jeder einzelne Summand ist Null, es sei denn $i = a$, $j = x = k$ und $\ell = b$. Für diese Indexkombination erhalten wir 1. Das gibt die Behauptung.

(ii) Wie in Proposition II.3.13 schreiben wir A als Linearkombination von Elementarmatrizen und erhalten

$$\begin{aligned} E_{i,j}A &= E_{i,j} \cdot \Big(\sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k,l}E_{k,\ell} \Big) \\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{m} A(k,\ell)E_{i,j}E_{k,\ell} = \sum_{\ell=1}^{m} A(j,\ell)E_{i,\ell}. \end{aligned}$$

(iii) Diese Aussage lässt sich genau wie (ii) zeigen. □

Proposition II.3.16 (Blockmatrizen): *Seien m, p und q natürliche Zahlen und seien m_1, m_2, p_1, p_2, q_1 und q_2 natürliche Zahlen, sodass $m = m_1 + m_2$, $p = p_1 + p_2$ und $q = q_1 + q_2$. Ferner seien $A \in \mathbb{R}^{q_1\times p_1}$, $B \in \mathbb{R}^{q_1\times p_2}$, $C \in \mathbb{R}^{q_2\times p_1}$, $D \in \mathbb{R}^{q_2\times p_2}$, $E \in \mathbb{R}^{p_1\times m_1}$, $F \in \mathbb{R}^{p_1\times m_2}$, $G \in \mathbb{R}^{p_2\times m_1}$ und $H \in \mathbb{R}^{p_2\times m_2}$ Matrizen. Für die Matrizen $M_1 \in \mathbb{R}^{q\times p}$ und $M_2 \in \mathbb{R}^{p\times m}$, die entstehen durch*

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

gilt

$$M_1M_2 = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}.$$

Beweis: Die angegebene Gleichheit von Matrizen ist eintragsweise zu überprüfen. Zum Beispiel für $i \in \{1, \ldots, q_1\}$ und $j \in \{1, \ldots, m_1\}$ gilt

$$\begin{aligned}(M_1M_2)(i,j) &= \sum_{k=1}^{p} M_1(i,k)M_2(k,j) \\ &= \sum_{k=1}^{p_1} A(i,k)E(k,j) + \sum_{k=p_1+1}^{p} B(i,k)G(j,k) \\ &= (AE)(i,j) + (BG)(i,j) = (AE+BG)(i,j). \qquad \square\end{aligned}$$

4. Reguläre Matrizen

Ein lineares Gleichungssystem hat die Form $Ax = b$ mit einer Matrix $A \in \mathbb{R}^{n\times m}$ und Vektoren $x \in \mathbb{R}^m$ und $b \in \mathbb{R}^n$. Gibt es eine Matrix B mit $BA = I_n$, dann können wir umformen $x = Bb$.

Sind allgemeiner C und D Matrizen mit $CD = I_n$, dann gilt $Ax = b$ genau dann, wenn $DAx = Db$. Solche Matrizen C und D heißen *regulär* und helfen uns beim Lösen linearer Gleichungssysteme (genauer: Der Gauß-Algorithmus lässt sich mithilfe regulärer Matrizen formulieren). Aber auch in eigenem Recht sind reguläre Matrizen interessant.

Definition II.4.1 (Reguläre Matrix): Seien n eine natürliche Zahl und A eine reelle $n \times n$-Matrix. Gibt es eine Matrix $B \in \mathbb{R}^{n\times n}$ mit $AB = I_n = BA$, dann heißt A *regulär* oder *invertierbar*. Wir schreiben

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

für die Menge der regulären Matrizen in $\mathbb{R}^{n\times n}$. Hierbei steht das „Gl“ für „general linear group“.

Proposition II.4.2 (Inverse Matrix): *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$ eine reguläre Matrix. Dann gibt es genau eine Matrix $B \in \mathbb{R}^{n\times n}$, sodass $AB = I_n = BA$. Die Matrix B heißt die* inverse Matrix *oder einfach die* Inverse zu A *und wird mit A^{-1} notiert.*

Sind A und B reguläre $n \times n$-Matrizen, dann ist auch ihr Produkt AB regulär und die zugehörige Inverse ist $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis: Angenommen, wir hätten zwei inverse Matrizen B und B' zu A, d. h. $AB = BA = I_n = AB' = B'A$. Dann hätten wir

$$B = BI_n = B(AB') = (BA)B' = I_nB' = B',$$

und damit $B = B'$.

Dafür dass AB regulär ist, genügt es zu überprüfen, dass die angegebene Matrix eine Inverse (und damit die Inverse von AB) ist. Das können wir einfach nachrechnen:

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AI_nA^{-1} = AA^{-1} = I_n,$$

genau so für $(B^{-1}A^{-1})(AB)$. □

Beispiel II.4.3 (Spezielle reguläre Matrizen): Seien n eine natürliche Zahl, α, β sowie $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ reelle Zahlen. Es bezeichne

$$C := \begin{pmatrix} a_1 & \ldots & a_n \\ b_1 & \ldots & b_n \end{pmatrix} \in \mathbb{R}^{2\times n}.$$

(i) Setzen wir $A = \left(\begin{smallmatrix} 1 & \alpha \\ 0 & 1 \end{smallmatrix}\right)$ und $B = \left(\begin{smallmatrix} 1 & -\alpha \\ 0 & 1 \end{smallmatrix}\right)$, dann haben wir

$$AB = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} = BA,$$

d. h. A ist regulär und B ist ihre Inverse. Multiplizieren wir A von links an C, also

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_1 & \ldots & a_n \\ b_1 & \ldots & b_n \end{pmatrix} = \begin{pmatrix} a_1 + \alpha b_1 & \ldots & a_n + \alpha b_n \\ b_1 & \ldots & b_n \end{pmatrix},$$

dann bewirkt das die Addition des α-fachen der zweiten Zeile zur ersten Zeile von C. Multiplizieren wir A von rechts an C^t, also

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

dann bewirkt das die Addition des α-fachen der ersten Spalte zur zweiten Spalte von C^t.

(ii) Für die Matrix $V = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ finden wir

$$V^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

d. h. V ist regulär und ihre Inverse V^{-1} ist auch V. Man nennt solche Matrizen *selbstinvers*. Multiplizieren wir V von links mit C, dann erhalten wir

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & \ldots & a_n \\ b_1 & \ldots & b_n \end{pmatrix} = \begin{pmatrix} b_1 & \ldots & b_n \\ a_1 & \ldots & a_n \end{pmatrix},$$

also vertauscht V dann die Zeilen von C. Multiplizieren wir V von rechts mit C^t, dann erhalten wir

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

d. h. dann vertauscht V die Spalten von C^t.

(iii) Seien α und β beide verschieden von Null und setze $D := \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$ sowie $E := \left(\begin{smallmatrix} 1/\alpha & 0 \\ 0 & 1/\beta \end{smallmatrix}\right)$. Dann ist

$$DE = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1/\alpha & 0 \\ 0 & 1\beta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/\alpha & 0 \\ 0 & 1\beta \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = ED,$$

d. h. dann ist D regulär und E ist die zugehörige Inverse. Sind α und β nicht notwendigerweise von Null verschieden und multiplizieren wir D von links mit C, dann erhalten wir

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

d. h. die erste Zeile wird mit α und die zweite Zeile wird mit β multipliziert. Multiplizieren wir D von rechts mit C^t, also

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

dann wird die erste Spalte von C^t mit α und die zweite Spalte von C^t mit β multipliziert.

Definition II.4.4: Seien n eine natürliche Zahl $a, \alpha_1, \dots, \alpha_n$ relle Zahlen und $i, j \in \{1, \dots, n\}$ mit $i \neq j$. Wir definieren drei 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.4.5: *Die Matrizen aus Proposition II.4.4 sind invertierbar, d. h. liegen in* $\mathrm{Gl}_n(\mathbb{R})$.

Beweis: Aus der nachfolgenden Proposition folgt:

- $A_{i,j}^{\alpha}$ is regulär mit Inverser $A_{i,j}^{-\alpha}$,
- $V_{i,j}$ ist regulär mit Inverser $V_{i,j}$,
- $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ ist regulär mit Inverser $\operatorname{diag}(1/\alpha_1, \ldots, 1/\alpha_n)$. □

Proposition II.4.6 (Wirkung der elementaren Operationen): *Die Matrizen aus Proposition II.4.4 wirken bei Multiplikation von links auf eine Matrix* $M \in \mathbb{R}^{n\times m}$ *wie folgt:*

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

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

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

Beweis: Verwenden wir die Ergebnisse aus Proposition II.3.15, 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

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

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

(iii) Es ist

$$\operatorname{diag}(\alpha_1, \ldots, \alpha_n)M = \Big(\sum_{k=1}^{n} \alpha_k E_{k,k}\Big)\Big(\sum_{i=1}^{n}\sum_{j=1}^{m} M(i,j)E_{i,j}\Big) = \sum_{k=1}^{n}\sum_{j=1}^{m} \alpha_k M(k,j)E_{k,j},$$

wobei wir verwendet haben, dass $E_{k,k}E_{i,j} = \mathbf{0}$, falls $i \neq k$. An der Stelle (k,j) steht also der Einträge $\alpha_k M(k,j)$. □

Proposition II.4.7 (Wirkung der elementaren Operationen II): *Die Grundoperationen aus Proposition II.4.4 wirken bei Multiplikation von rechts auf eine Matrix $A \in \mathbb{R}^{m\times n}$ wie folgt:*

(i) *$MA^{\alpha}_{i,j}$ entsteht durch Addition des α-fachen der i-ten Spalte von M zur j-ten Spalte von M,*

(ii) *$MV_{i,j}$ entsteht durch Vertauschung der i-te und der j-te Spalte von M,*

(iii) *$M\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ entsteht durch Multiplikation jeweils die i-te Spalte von M mit α_i.*

Beweis: Wir zeigen exemplarisch (i). Wegen

$$(MA^{\alpha}_{i,j})^t = (A^{\alpha}_{i,j})^t M^t = A^{\alpha}_{j,i} M$$

liefert uns Proposition II.4.6 dass $A^{\alpha}_{j,i}M^t$ entsteht durch Addition des α-fachen i-ten Zeile zur j-ten Zeile, d. h. $MA^{\alpha}_{i,j}$ entsteht durch Addition des α-fachen der j-ten Spalte zur i-ten Spalte. □

Beispiel II.4.8: Sei $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$. Die Matrix A ist regulär genau dann, wenn $ad - bc$ verschieden von Null ist. In diesem Fall lässt sich die Inverse von A einfach angeben; es ist dann nämlich

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

Proposition II.4.9 (Spezielle Blockmatrizen): *Seien p und n natürliche Zahlen mit $p < n$ und $M \in \mathbb{R}^{n\times n}$ mit*

$$M = \begin{pmatrix} I_p & B \\ \mathbf{0}_{(n-p)\times p} & D \end{pmatrix},$$

wobei I_p die Einheitsmatrix in $\mathbb{R}^{p\times p}$, $\mathbf{0}$ die Nullmatrix in $\mathbb{R}^{(n-p)\times p}$, B irgendeine Matrix in $\mathbb{R}^{p\times(n-p)}$ und D irgendeine Matrix $\mathbb{R}^{(n-p)\times(n-p)}$ ist. Die Matrix M ist regulär genau dann, wenn D regulär ist. Die Inverse in diesem Fall ist

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

Beweis: „$\Longrightarrow$“: Angenommen M wäre regulär, d. h., angenommen es gäbe $M' \in \mathbb{R}^{n\times n}$ mit $MM' = M'M = I_n$. Diese Matrix könnten wir schreiben als $M' = \left(\begin{smallmatrix} E & F \\ G & H \end{smallmatrix}\right)$ mit Matrizen $E \in \mathbb{R}^{p\times p}$, $F \in \mathbb{R}^{p\times(n-p)}$, $G \in \mathbb{R}^{(n-p)\times p}$ und $H \in \mathbb{R}^{(n-p)\times(n-p)}$. Aus Proposition II.3.16 folgt

$$I_n = M'M = \begin{pmatrix} E & F \\ G & H \end{pmatrix}\begin{pmatrix} I_p & B \\ \mathbf{0} & D \end{pmatrix} = \begin{pmatrix} E & EB+FD \\ G & GB+HD \end{pmatrix}.$$

Durch Vergleich mit der Einheitsmatrix können wir ablesen, dass $E = I_{p\times p}$, dass $G = \mathbf{0}$, dass $EB + FD = \mathbf{0}$ und dass $GB + HD = I_{n-p}$. Außerdem haben wir

$$I_n = MM' = \begin{pmatrix} I_p & B \\ \mathbf{0} & D \end{pmatrix}\begin{pmatrix} E & G \\ H & F \end{pmatrix} = \begin{pmatrix} I_p & F+BH \\ \mathbf{0} & DH \end{pmatrix},$$

woraus wir ablesen, dass $I_{n-p} = DH$ und $\mathbf{0} = F + BH$.

Zusammengenommen folgt, dass D regulär mit Inverser $D^{-1} = H$ ist und dass $F = -BD^{-1}$.

„$\Longleftarrow$“: Hier rechen wir mithilfe von Proposition II.3.16 nach, dass die angegebene Blockmatrix tatsächlich die Inverse von M ist. □

5. Lineare Gleichungssysteme

In diesem Abschnitt wollen wir lineare Gleichungssysteme einführen, uns mit dem Lösungsalgorithmus von Gauß beschäftigen und die Struktur der Lösungsmenge verstehen.

Definition II.5.1 (Lineares Gleichungssystem): Seien n und m natürliche Zahlen.

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

$$\begin{aligned} a_{1,1}x_1 + \ldots + a_{1,m}x_m &= b_1 \\ \vdots \quad\quad \vdots \quad\quad\quad &\vdots \\ a_{n,1}x_1 + \ldots + a_{n,m}x_m &= b_n \end{aligned}$$

Hierbei sind die $a_{i,j}$ und die b_j für $1 \leq i \leq n$ und $1 \leq j \leq m$ reelle Zahlen. Die $x_1, \ldots, x_m$ heißen *Unbekannte*.

(ii) Das lineare Gleichungssystem hat folgende schematische Beschreibung:

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

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

(iii) Die Menge

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

heißt *Lösungsmenge* und $\mathbb{L}^h := \mathbb{L}^h(A,b) := \mathbb{L}(A, \mathbf{0})$ heißt zugehörige *homogene Lösungsmenge*.

Im Folgenden seien $A \in \mathbb{R}^{n\times m}$ eine Matrix und $b \in \mathbb{R}^n$ ein Vektor und wir betrachten das lineare Gleichungssystem mit der schematischen Beschreibung $(A|b)$.

Bemerkung II.5.2 (Lineares Gleichungssystem in Matrizenform): Für $x \in \mathbb{R}^m$ gilt $Ax = b$ genau dann, wenn x zu $\mathbb{L}(A,b)$ gehört, d. h. $\mathbb{L} = \{x \in \mathbb{R}^m \mid Ax = b\}$. Entsprechend gilt für die homogene Lösungsmenge $\mathbb{L}^h = \{x \in \mathbb{R}^m \mid Ax = \mathbf{0}\}$.

Proposition II.5.3 (Struktur der Lösungsmenge):

(i) *Die homogene Lösungsmenge* $\mathbb{L}^h$ *ist ein Untervektorraum von* $\mathbb{R}^m$.

(ii) *Ist* $x_{(s)} \in \mathbb{R}^m$ *mit* $Ax^{(s)} = b$, *dann gilt* $\mathbb{L} = x^{(s)} + \mathbb{L}^h = \{x^{(s)} + v \mid v \in \mathbb{L}^h\}$.

Beweis: (i) Wegen $A\mathbf{0} = \mathbf{0}$ gehört $\mathbf{0}$ zu $\mathbb{L}^h = \mathbb{L}(A, \mathbf{0})$. Sind x und y Elemente von $\mathbb{L}^h$, dann haben wir $A(x + y) = Ax + Ay = \mathbf{0} + \mathbf{0} = \mathbf{0}$, d. h. $\mathbb{L}^h$ ist stabil unter Summation. Sind schließlich $x \in \mathbb{L}^h$ und r eine reelle Zahl, dann ist $A(rx) = rAx = r\mathbf{0} = \mathbf{0}$, d. h. $\mathbb{L}^h$ ist ein Vektorraum.

(ii) „$\subseteq$“: Sei $x \in \mathbb{L}$, d. h. $Ax = b$. Dann ist $A(x - x^{(s)}) = Ax - Ax^{(s)} = \mathbf{0}$, was bedeutet, dass $v := x - x^{(s)} \in \mathbb{L}^h$. Wegen $x = x^{(s)} + x - x^{(s)} = x^{(s)} + v$ haben wir die gewünschte Darstellung.

„$\supseteq$“: Sei v ein Element der homogenen Lösungsmenge. Für $x := x^{(s)} + v$ finden wir $Ax = A(x^{(s)} + v) = Ax^{(s)} + Av = b + \mathbf{0} = b$, sodass x zu $\mathbb{L}$ gehört. □

Proposition II.5.4 (Lösungsstrategie für lineare Gleichungssysteme): *Ist* C *eine reguläre* $n \times n$*-Matrix, dann gilt für* $x \in \mathbb{R}^m$ *genau dann* $Ax = b$, *wenn* $CAx = Cb$. *Insbesondere gilt* $\mathbb{L}(A, b) = \mathbb{L}(CA, Cb)$.

Beweis: Für „$\Longrightarrow$“ multipliziere die linke Gleichung mit C von links und für „$\Longleftarrow$“ multipliziere die rechte Gleichung mit C^{-1} von links. □

Bemerkung II.5.5 (Elementare Zeilenumformungen): Wählt man in Proposition II.5.4 als reguläre Matrix C eine Additionsmatrix $A^{\alpha}_{i,j}$, eine Vertauschungsmatrix $V_{i,j}$ oder eine Diagonalmatrix $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ mit $\alpha_1 \cdots \alpha_n \neq 0$, dann erhält man die elementaren Zeilenumformungen

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

(ii) Vertauschung der i-ten und j-ten Gleichung,

(iii) Multiplikation der i-ten Gleichung mit $\alpha_i \neq 0$.

Insbesondere liefert Proposition II.5.4, dass elementare Zeilenumformungen die Lösungsmenge eines linearen Gleichungssystems nicht verändern.

Im Folgenden wollen wir elementare Zeilenumformungen verwenden, um lineare Gleichungssysteme in eine „schöne“ Form zu bringen.

Beispiel II.5.6 (Beispiel für Treppenform): Betrachte die Matrix

$$T = \begin{pmatrix} 0 & 1 & 0 & 13 & 0 & 7 \\ 0 & 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Das zugehörige lineare Gleichungssystem mit rechter Seite $b \in \mathbb{R}^4$ ist

$$\begin{aligned} x_2 \quad + 13x_4 \quad + 7x_6 &= b_1 \\ x_3 + \quad 2x_4 \quad + 5x_6 &= b_2 \\ x_5 + 3x_6 &= b_3 \\ 0 &= b_4 \end{aligned}$$

Wir haben $n = 4$ Gleichungen und die $m = 6$ Variablen $x_1, \ldots, x_6$. Für den Umstand, dass wir „drei Stufen“ in der Matrix T sehen können, wollen wir sagen, T habe „den Rang drei“. Die Spaltenindizes der „Stufen“ wollen wir später Stufenindizes nennen. In diesem Beispiel sind das $s_1 = 2$, $s_2 = 3$ und $s_3 = 5$.

Definition II.5.7 (Einheitsvektor): Sei n eine natürliche Zahl. Für $1 \leq i \leq n$ heißt $e_i = (\delta_{ij})_{1 \leq j \leq n}^t$ der *i-te Einheitsvektor*.

Der Vektor e_i ist also ein Spaltenvektor, der einen einzigen von Null verschiedenen Eintrag hat; und das ist eine Eins in der i-ten Zeile.

Bemerkung II.5.8: Für die Einheitsvektoren $e_1, \ldots, e_m$ im $\mathbb{R}^m$ haben wir:

(i) Ist $x = (x_1, \ldots, x_m)^t \in \mathbb{R}^m$, dann ist $x = \sum_{i=1}^m x_i e_i$.

(ii) Ist $A = (a_{i,j}) \in \mathbb{R}^{n \times m}$ eine Matrix, dann liefert Ae_j die j-te Spalte von A, d. h.

$$Ae_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix} = \sum_{k=1}^n a_{k,j} \tilde{e}_k,$$

wobei $\tilde{e}_1, \ldots, \tilde{e}_n$ die Einheitsvektoren des $\mathbb{R}^n$ sind.

Bemerkung II.5.9: Die Matrix T aus Beispiel II.5.6 hat folgende Eigenschaften:

(1) Für $1 \leq i \leq r$ steht eine Eins an der Stelle (i, s_i). Diese nennen wir „Treppenstufen“.

(2) In jeder Zeile $1 \leq i \leq r$ stehen links von den Treppenstufen nur Nullen.

(3) In den Spalten s_i mit Treppenstufen stehen an allen anderen Stellen (außer der Stufe) nur Nullen.

(4) Ab der $r+1$-ten Zeile sind alle Einträge Null.

Definition II.5.10 (Treppenform/Gauß-Normalform, Rang): Seien n und m natürliche Zahlen und $T = (t_{i,j}) \in \mathbb{R}^{n\times m}$ eine Matrix. Gibt es eine natürliche Zahl r und Indizes $s_1, \ldots, s_r$ mit $1 \leq s_1 < s_2 < \cdots < s_r \leq m$, sodass

(i) Für alle $1 \leq i \leq r$ gilt $t_{i,s_i} = 1$,

(ii) Für alle $1 \leq i \leq r$ und $1 \leq j \leq s_i$ ist $t_{i,j} = 0$,

(iii) Für alle $1 \leq i \leq r$ und $k \in \{1, \ldots, n\} - \{i\}$ ist $t_{k,s_i} = 0$,

(iv) Für $i \geq r+1$ und $1 \leq j \leq m$ ist $t_{i,j} = 0$,

dann hat T *Treppenform* beziehungsweise *Gauß-Normalform*. Dabei heißt r der *Rang von* T und $s_1, \ldots, s_r$ heißen *Spaltenindizes*.

Beispiel II.5.11: Ist T eine Matrix in Treppenform und ist b eine rechte Seite, wie sieht dann die Lösungsmenge von $(T|b)$ aus? Für das Gleichungssystem

$$\begin{aligned} x_2 \quad + 13x_4 \quad + 7x_6 &= b_1 \\ x_3 + \ 2x_4 \quad + 5x_6 &= b_2 \\ x_5 + 3x_6 &= b_3 \\ 0 &= b_4 \end{aligned}$$

sehen wir die „freien Variablen“ x_1, x_4 und x_6, die anderen sind festgelegt durch die Gleichungen

$$\begin{aligned} x_2 &= b_1 - 13x_4 - 7x_6 \\ x_3 &= b_2 - \ 2x_4 - 5x_6 \\ x_5 &= b_3 \qquad\quad - 3x_6 \end{aligned}$$

Setzen wir $x_1 = x_4 = x_6 = 0$, dann erhalten wir eine spezielle Lösung des linearen Gleichungssystems, indem wir die Werte für x_1, x_4 und x_6 in die obigen Gleichungen einsetzen um die zugehörigen Werte für x_2, x_3 und x_5 zu bestimmen. Diese spezielle Lösung ist $x^{(s)} = (0, b_1, b_2, 0, b_3, 0)^t$.

Proposition II.5.12 (Lösbarkeit und spezielle Lösung für Treppenform): *Seien n und m natürliche Zahlen, $b \in \mathbb{R}^n$ und $T = (t_{i,j}) \in \mathbb{R}^{n\times m}$ eine Matrix in Treppenform. Das lineare Gleichungssystem $(T|b)$ ist lösbar genau dann, wenn $b_{r+1} = \cdots = b_n = 0$. In diesem Fall ist $x^{(s)} := \sum_{i=1}^{r} b_i e_{s_i}$ eine spezielle Lösung des linearen Gleichungssystems $(T|b)$.*

Beweis: Es bezeichnen $e_1, \dots, e_m$ die Einheitsvektoren des $\mathbb{R}^m$ und $\tilde{e}_1, \dots, \tilde{e}_n$ die des $\mathbb{R}^n$. „$\Longrightarrow$“: Gäbe es ein $i \in \{r+1, \dots, n\}$ sodass $b_i \neq 0$, dann hätte die i-te Gleichung $0 = b_i$ keine Lösung.

„$\Longleftarrow$“: Wir rechnen nach, dass $x^{(s)}$ tatsächlich eine Lösung ist, d. h. dass $Tx^{(s)} = b$:

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

Bei der vorvorletzten Gleichheit haben wir verwendet, dass T in Treppenform ist und bei der vorletzten Gleichheit haben wir nur mit Nullen aufgefüllt. □

Bemerkung II.5.13: In Beispiel II.5.6 erhalten wir für das zugehörige lineare Gleichungssystem die Gleichungen

$$\begin{aligned} x_2 &= 0x_1 - 13x_4 - 7x_6 \\ x_3 &= 0x_1 - \ \ 2x_4 - 5x_6 \\ x_5 &= 0x_1 + \ \ 0x_4 - 3x_6 \end{aligned}$$

Wir erhalten drei besondere Lösungen duch spezielle Wahlen für die freien Parameter.

(i) Für $x_1 = 1$, $x_4 = 0$ und $x_6 = 0$ erhalten wir $x_2 = x_3 = x_5 = 0$ und die besondere Lösung $F^{(1)} = (1, 0, 0, 0, 0, 0)^t$.

(ii) Für $x_1 = 0$, $x_4 = 1$ und $x_6 = 0$ erhalten wir $x_2 = -13$, $x_3 = -2$ und $x_5 = 0$ und damit die besondere Lösung $F^{(4)} = (0, -13, -2, 1, 0, 0)^t$.

(iii) Für $x_1 = 0$, $x_4 = 0$ und $x_6 = 1$ erhalten wir $x_2 = -7$, $x_3 = -5$ und $x_5 = -3$ und damit die besondere Lösung $F^{(6)} = (0, -7, -5, 0, -3, 1)^t$.

Proposition II.5.14 (Fundamentallösungen): *Seien n und m natürliche Zahlen, $b \in \mathbb{R}^n$ und $T = (t_{i,j}) \in \mathbb{R}^{n \times m}$ eine Matrix in Treppenform.*

(i) *Sei $J := \{1, \dots, n\} - \{s_1, \dots, s_r\}$. Für $j \in J$ löst*

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

das zugehörige homogene lineare Gleichungssystem. Die $F^{(j)}$, $j \in J$, heißen Fundamentallösungen.

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

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

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

Beweis: (i) Für jedes $j \in J$ haben wir zu zeigen, dass $TF^{(j)} = \mathbf{0}$. Es gilt

$$TF^{(j)} = Te_j - \sum_{i=1}^{r} t_{i,j} Te_{s_j} = Te_j - \sum_{i=1}^{r} t_{i,j}\tilde{e}_j = Te_j - \sum_{i=1}^{n} t_{i,j}\tilde{e}_j = \mathbf{0}.$$

(ii) Die Inklusion „$\supseteq$“ folgt aus (i), weil $\mathbb{L}^h$ ein Untervektorraum von $\mathbb{R}^n$ ist.

Für „$\subseteq$“ sei $v = \sum_{i=1}^{m} v_i e_i$ eine homogene Lösung des linearen Gleichungssystem, d. h. $Tv = \mathbf{0}$. Sei $d := v - \sum_{j\in J} v_j F^{(j)}$. Wir zeigen $d = \mathbf{0}$, indem wir für $1 \le i \le m$ zeigen, dass $d_i = 0$. Dazu verwenden wir, dass $\{1, \dots, m\} = J \cup \{s_1, \dots, s_r\}$ und dass d zu $\mathbb{L}^h$ gehört (also $Td = \mathbf{0}$), da $\mathbb{L}^h$ ein Untervektorraum ist.

Gehört i zu J, dann hat $F^{(j)}$ in der i-ten Zeile den Eintrag δ_{ij} und damit ist $d_i = v_i - \sum_{j\in J} v_j \delta_{ij} = v_i - v_i = 0$.

Gehört i zur Menge $\{s_1, \dots, s_r\}$, dann ist $i = s_j$ für einen geeigneten Index $j \in \{1, \dots, r\}$. Die j-te Zeile von Td ist

$$\sum_{k=1}^{m} t_{j,k} d_k = 0 = 1 d_{s_j}$$

denn nach dem ersten Fall ist $d_k = 0$ für $k \in J$ und wir wissen, dass $t_{j,k} = 0$ für $k \notin J$ und $k \neq s_j$ und $t_{j,k} = 1$ für $k = s_j$. □

Bemerkung II.5.15 ((−1)-Trick): Die Fundamentallösungen aus Proposition II.5.14 erhalten wir wie folgt:

(i) Schreibe für $1 \le i \le r$ die i-te Zeile der Matrix T in Treppenform als s_i-te Zeile in eine neue Matrix $S \in \mathbb{R}^{m\times m}$, deren übrige Zeilen Null sind,

(ii) Die von Null verschiedenen Spalten der Matrix $I_m - S$ sind die Fundamentallösungen.

Im Fall von Beispiel II.5.6 erhalten wir so

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 13 & 0 & 7 \\ 0 & 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad I_6 - S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -13 & 0 & -7 \\ 0 & 0 & 0 & -2 & 0 & -5 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Satz 3 (Lösungsmenge für Treppenform): *Ist $T = (t_{i,j}) \in \mathbb{R}^{n\times m}$ in Treppenform vom Rang r mit Spaltenindizes $s_1, \ldots, s_r$ und $b \in \mathbb{R}^n$, dann gilt*

$$\mathbb{L}(T|b) = x^{(s)} + \Big\{ \sum_{j\in J} \lambda_j F^{(j)} : \lambda_j \in \mathbb{R} \Big\},$$

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

Der Beweis dieses Satzes ist das Zusammenschreiben der Beweise der Aussagen für Proposition II.5.3, Proposition II.5.12 und Proposition II.5.14.

Beispiel II.5.16 (Gauß-Algorithmus): Wir suchen die Lösungsmenge des folgenden Gleichungssystems:

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

$$\xrightarrow{\frac{1}{2}\text{I}} \left(\begin{array}{ccccc|c} 0 & 1 & 2 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 6 & -3 & -6 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right) \xrightarrow{\text{IV}-3\text{I}} \left(\begin{array}{ccccc|c} 0 & 1 & 2 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -6 & -6 & -12 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right)$$

$$\xrightarrow{\frac{1}{6}\text{III}} \left(\begin{array}{ccccc|c} 0 & 1 & 2 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right) \xrightarrow[\text{II}\leftrightarrows\text{III}]{\text{IV}-\text{III}} \left(\begin{array}{ccccc|c} 0 & 1 & 2 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

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

Mithilfe des -1-Tricks lässt sich jetzt die Lösungsmenge $\mathbb{L}$ bestimmen, die ist nämlich

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

Proposition II.5.17 (Gauß-Algorithmus): *Sei $A \in \mathbb{R}^{n\times m}$ gegeben. Dann gibt es eine reguläre Matrix $C \in \mathbb{R}^{n\times n}$, sodass CA Treppenform hat.*

Beweis: Wir zeigen dazu, dass sich A mit Zeilenumformungen in Treppenform bringen lässt. Die Behauptung folgt dann aus Bemerkung II.5.5. Das machen wir per vollständiger Induktion nach der Anzahl der Zeilen. Besteht A nur aus einer Zeile, dann ist A entweder die Nullzeile, oder nicht. Ist A die Nullzeile, dann ist $A = I_n A$ in Treppenform.

Ist A nicht die Nullzeile, dann setzen wir $s_1 := \min\{j \in \{1, \ldots, m\} \mid a_{1,j} \neq 0\}$ und wenden $\mathrm{Mult}_1(1/a_{1,s_1})$ auf A an, um Treppenform zu erhalten.

Für den Induktionsschritt von $n-1$ nach n müssen wir wieder Fälle unterscheiden. Ist $A = \mathbf{0}$, dann sind wir fertig. Ist A nicht die Nullmatrix, dann setzen wir

$$s_1 := \min\{j \mid \text{Es gibt } i \in \{1, \ldots, n\}, \text{sodass } a_{i,j} \neq 0\}, i_0 := \min\{i \mid a_{i,s_1} \neq 0\}.$$

Der Spaltenindex s_1 gehört zur ersten Spalte von Links, die keine Nullspalte ist, und i_0 ist der Index der obersten Zeile, in der in der Spalte s_1 keine Null steht. Dann ist A von der Form

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

Durch Anwendung von Vert_{1,i_0} und $\mathrm{Mult}_1(1/a_{i_0,s_1})$ bringen wir A in die Form

$$A_1 = \left(\begin{array}{ccc|c|c} 0 & \cdots & 0 & 1 & \tilde{z} \\ \hline 0 & \cdots & 0 & a_{2,s_1} & \\ \vdots & & \vdots & \vdots & \tilde{A}_1 \\ 0 & \cdots & 0 & a_{n,s_1} & \end{array}\right).$$

Setzen wir für $2 \leq i \leq n$ nun $\alpha_i := -a_{i,s_1}$ und wenden $\mathrm{Add}_{1,i}^{\alpha_i}$ auf A_1 an, dann erhalten wir die Matrix

$$A_2 = \left(\begin{array}{ccccc|c} 0 & \cdots & 0 & 1 & \tilde{z} \\ \hline & & & & \\ & & \mathbf{0} & & \tilde{A}_2 \\ & & & & \end{array}\right)$$

Die Matrix $\tilde{A}_2$ gehört dabei zu $\mathbb{R}^{(n-1)\times(n-1)}$ und kann nach Induktionvoraussetzung mithilfe elementarer Zeilenumformungen in Treppenform $\tilde{T}$ gebracht werden. Es bezeichne $\tilde{r}$ den Rang von $\tilde{T}$, $r := \tilde{r} + 1$ und $s_2, \ldots, s_r$ die Stufenzahlen von $\tilde{T}$.

Durch Anwendung der Zeilenumformungen, die $\tilde{A}_2$ in $\tilde{T}$ überführen, bringen wir A_2 in die Gestalt

$$A_2 = \left(\begin{array}{cccc|c} 0 & \cdots & 0 & 1 & \tilde{z} \\ \hline & & & & \\ & & \mathbf{0} & & \tilde{T} \\ & & & & \end{array}\right)$$

Die Matrix A_2 hat dabei die Eigenschaften (i), (ii) und (iv) aus Definition II.5.10 und die Spalten $s_2, \ldots, s_r$ von A_2 sind von der Form $(*, \delta_{2,j+1}, \ldots, \delta_{n-1,j+1})^t$ für $2 \leq j \leq r$. Durch Anwenden von $\mathrm{Add}_{j+1,1}(-a_{1,s_j})$ für $2 \leq j \leq r$ gelangen wir also zu Treppenform für die Matrix A. □

Im Folgenden wollen wir uns mit der Eindeutigkeit der Treppenform einer Matrix beschäftigen.

Bemerkung II.5.18: Ist A eine reguläre $n \times n$-Matrix, dann gilt für alle von Null verschiedenen Vektoren $v \in \mathbb{R}^n$, dass $Av \neq \mathbf{0}$. Wäre nämlich $Av = \mathbf{0}$, dann hätten wir $v = I_n v = A^{-1}Av = A^{-1}\mathbf{0} = \mathbf{0}$.

Proposition II.5.19 (Eindeutigkeit der Treppenform): *Seien T und $\tilde{T}$ Matrizen in $\mathbb{R}^{n\times m}$ in Treppenform. Ist $D \in \mathrm{Gl}_n(\mathbb{R})$ mit $\tilde{T} = DT$, dann gilt $\tilde{T} = T$.*

Beweis: Auch diese Aussage zeigen wir per vollständige Induktion nach der Anzahl der Zeilen n. Für $n = 1$ sind T und $\tilde{T}$ beides Matrizen, die nur aus einer Zeile bestehen. Die Matrizen T bzw. $\tilde{T}$ sind jeweils entweder die Nullzeile, oder es gibt jeweils irgendeinen Spaltenindex, in dem T bzw. $\tilde{T}$ eine Eins führt. Für $D = (d_{1,1})$ haben wir $\tilde{T} = d_{1,1}T$. Entsprechend sind entweder T und $\tilde{T}$ beide Nullzeilen, oder $d_{1,1} = 1$ und $T = \tilde{T}$.

Für den Induktionsschritt $n - 1$ nach n haben wir wieder Fälle zu unterscheiden. Ist $T = \mathbf{0}$, dann ist auch $\tilde{T} = DT = \mathbf{0}$; genauso falls $\tilde{T} = \mathbf{0}$. Sind nun T und $\tilde{T}$ nicht Null, dann bezeichne r den Rang von T mit Spaltenindizes $s_1, \ldots, s_r$ und $\tilde{r}$ den Rang von $\tilde{T}$ mit Spaltenindizes $\tilde{s}_1, \ldots, \tilde{s}_{\tilde{r}}$.

$$\tilde{T} = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & \tilde{T}_1 & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}, \quad T = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & T_1 & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}$$

Wir finden also in den Matrizen T und $\tilde{T}$ kleinere Matrizen $T_1 \in \mathbb{R}^{(n-1)\times(m-s_1)}$ und $\tilde{T}_1 \in \mathbb{R}^{(n-1)\times(m-\tilde{s}_1)}$, die Treppenform haben.

Wir schreiben die Matrizen als $T = (t^{(1)}|\dots|t^{(m)})$ und $\tilde{T} = (\tilde{t}^{(1)}|\dots|\tilde{t}^{(m)})$, wobei $t^{(1)}, \dots, t^{(m)} \in \mathbb{R}^n$ und $\tilde{t}^{(1)}, \dots, \tilde{t}^{(m)} \in \mathbb{R}^n$ die jeweiligen Spalten der Matrizen sind. Aus $\tilde{T} = DT$ folgt für die Spalten, dass $\tilde{t}^{(j)} = Dt^{(j)}$. Wegen Bemerkung II.5.18 muss $\tilde{s}_1 = s_1$ gelten, d. h. $e_1 = \tilde{t}^{(s_1)} = Dt^{(s_1)} = De_1$, sodass die erste Spalte von D der Vektor e_1 sein muss.

Die Matrix D hat also Blockgestalt mit Blöcken

$$D = \begin{pmatrix} 1 & D_{1,2} \\ \mathbf{0} & D_{2,2} \end{pmatrix}.$$

Nach (Proposition 4.9) ist $\hat{D} := D_{2,2} \in Gl_{n-1}(\mathbb{R})$ und nach Proposition 3.17 ist $\tilde{T}_1 = \hat{D}T_1$. Nach Induktionsvoraussetzung ist $\tilde{T}_1 = T_1$, insbesondere erhalten wir so dass $r = \tilde{r}$ und $s_2 = \tilde{s}_2, \dots s_r = \tilde{s}_r$. Für $k \in \{1, \dots, r\}$ ist die Spalte s_k von $\tilde{T}$ und T der Einheitsvektor e_k. Es muss also $De_k = Dt^{(s_k)} = \tilde{t}^{(s_k)} = e_k$ gelten. Damit hat D die Gestalt

$$D = \begin{pmatrix} I_r & * \\ \mathbf{0} & * \end{pmatrix},$$

d. h. $\tilde{T} = DT = T$, da in T alle Zeilen ab der $(r+1)$-ten Zeile Nullzeilen sind.□

Satz 4 (Gauß-Normalform): *Seien n und m natürliche Zahlen. Für jede Matrix $A \in \mathbb{R}^{n\times m}$ gibt es genau eine Matrix $T \in \mathbb{R}^{n\times m}$ in Stufenform mit der folgenden Eigenschaft: Es gibt $D \in \mathrm{Gl}_n(\mathbb{R})$, sodass $DA = T$.*

Die Matrix T heißt Gauß-Normalform von A *oder* Stufenform von A.

Beweis: Die Existenz von T folgt aus Proposition II.5.17. Zur Eindeutigkeit: Sind T_1 und T_2 Matrizen in Stufenform und sind $D_1, D_2 \in \mathrm{Gl}_n(\mathbb{R})$, sodass $A = D_1^{-1}T_1 = D_2^{-1}T_2$, dann ist $T_1 = D_1A = (D_1D_2^{-1})T_2$. Wegen $D_1D_2^{-1} \in \mathrm{Gl}_n(\mathbb{R})$ liefert Proposition II.5.19, dass $T_1 = T_2$. □

Definition II.5.20 (Rang): Seien n und m natürliche Zahlen. Für $A \in \mathbb{R}^{n\times m}$ heißt der Rang r der Stufenform zu A der *Rang von A*, in Zeichen $\mathrm{Rang}(A)$.

Fazit II.5.21: Wir erhalten das folgende Verfahren zum Lösen eines beliebigen linearen Gleichungssystems $Ax = b$, wobei $A \in \mathbb{R}^{n\times m}$ und $b \in \mathbb{R}^n$:

(i) Bestimme die eindeutige Stufenform zu A mit $CA = T$ für eine invertierbare Matrix $C \in \mathrm{Gl}_n(\mathbb{R})$ (siehe Proposition II.5.17).

(ii) Berechne die Lösungsmenge $\mathbb{L}$ von $Tx = Cb$ nach Satz 3. Nach Proposition II.5.4 ist $\mathbb{L}$ dann auch die Lösungsmenge des linearen Gleichungssystems $Ax = b$.

Korollar II.5.22 (Lösbarkeit vs. Rang): *Seien n und m natürliche Zahlen, A in $\mathbb{R}^{n\times m}$ und b in $\mathbb{R}^n$ gegeben. Für das lineare Gleichungssystem $Ax = b$ gilt:*

(i) *Das lineare Gleichungssystem ist lösbar dann und nur dann, wenn* $\mathrm{Rang}(A) = \mathrm{Rang}(A|b)$.

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

Weiter gilt:

(iii) *Genau dann ist für alle $c \in \mathbb{R}^n$ das lineare Gleichungssystem $Ax = c$ lösbar, wenn* $\mathrm{Rang}(A) = n$.

Beweis: (i) Nach Definition des Rangs bleibt dieser unverändert unter Multiplikation mit regulären Matrizen.

(ii) Sei $T = DA$ in Stufenform. Es ist $Ax = b$ eindeutig lösbar genau dann, wenn $Tx = Db$ eindeutig lösbar ist, und das ist genau dann der Fall, falls $Tx = Db$ lösbar ist und $m - \mathrm{Rang}(T) = 0 = m - \mathrm{Rang}(A)$ gilt.

(iii) Sei T die Stufenform zu A. Das lineare Gleichungssystem $Tx = c$ ist genau dann für alle $c \in \mathbb{R}^n$ lösbar, wenn T keine Nulllzeilen hat, d. h. wenn $n = \mathrm{Rang}(T) = \mathrm{Rang}(A)$. □

Korollar II.5.23 (Rang und invertierbare Matrizen): *Seien n eine natürliche Zahl und $A \in \mathbb{R}^{n\times n}$. Die folgenden Aussagen sind äquivalent:*

(i) *A ist regulär.*

(ii) *Der Rang von A ist n.*

(iii) *Es gibt eine Matrix $B \in \mathbb{R}^{n\times n}$ mit $AB = I_n$.*

Beweis: „(i) $\Longrightarrow$ (ii)“: Ist A regulär, dann gibt es eine Matrix $B \in \mathrm{Gl}_n(\mathbb{R})$ mit $BA = I_n$, d. h. I_n ist die Stufenform von A und $\mathrm{Rang}(A) = \mathrm{Rang}(I_n) = n$.

„(ii) $\Longrightarrow$ (i)“: Die Einheitsmatrix ist die einzige Stufenform mit Rang n. Ist also $\mathrm{Rang}(A) = n$, dann gibt es $C \in \mathrm{Gl}_n(\mathbb{R})$, sodass $CA = I_n$. Das heißt $A = C^{-1}$ ist invertierbar.

„(i) $\Longrightarrow$ (iii)“: Das ist klar.

„(iii) $\Longrightarrow$ (i)“: Ist $AB = I_n$, dann gilt für alle $c \in \mathbb{R}^n$, dass $A(Bc) = c$ und Korollar II.5.22 liefert, dass $\mathrm{Rang}(A) = n$. □

Beispiel II.5.24 (Invertierbarkeit von Matrizen): Ist die Matrix

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

invertierbar? Zur Bestimmung des Rangs müssen wir ohnehin die Treppenform von A berechnen und lösen dabei simultan die linearen Gleichungssysteme $(A|e_1)$, $(A|e_2)$ und $(A|e_3)$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{III}-\text{I}} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{\text{I}-\text{III}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$

Da A die Treppenform I_3 hat, gehört A zu $\mathrm{Gl}_3(\mathbb{R})$ und $A^{-1}e_i$ ist die eindeutige Lösung von $Ax = e_i$, d. h. durch das gleichzeitige Lösen der linearen Gleichungssysteme $(A|e_1)$, $(A|e_2)$ und $(A|e_3)$ haben wir die Inverse von A gleich mitbestimmt. Die ist nämlich

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

Bemerkung II.5.25 (Von Matrix zur linearen Abbildung): Seien $n, m \in \mathbb{N}$ und eine Matrix $A \in \mathbb{R}^{n\times m}$ gegeben. Die Abbildung $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$, $x \mapsto Ax$ hat folgende Eigenschaften:

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

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

(ii) Für jede reelle Zahl r und jedes $x \in \mathbb{R}^m$ gilt

$$\phi_A(rx) = A(rx) = rAx = r\phi_A(x).$$

Eine Abbildung von $\mathbb{R}^m$ nach $\mathbb{R}^n$ mit den Eigenschaften (i) und (ii) heißt *lineare Abbildung* oder auch $\mathbb{R}$-*Vektorraumhomomorphismus*.

Bemerkung II.5.26 (Von linearer Abbildung zur Matrix): Seien n und m natürliche Zahlen und $\phi \colon \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung. Für $1 \leq i \leq m$ setzen wir $a_i := \phi(e_i)$ und erklären eine Matrix $A \in \mathbb{R}^{n\times m}$ durch $A := (a_1|\ldots|a_m)$. Für diese Matrix A gilt mit der Notation aus der vorangegangenen Bemerkung, dass $\phi_A = \phi$.

Definition II.5.27 (Bild und Kern): Seien n und m natürliche Zahlen und $\phi\colon \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung. Dann heißen

$$\text{Bild}(\phi) := \{\phi(v) \mid v \in \mathbb{R}^m\}, \qquad \text{Kern}(\phi) := \{v \in \mathbb{R}^m \mid \phi(v) = \mathbf{0}\}$$

das *Bild* respektive der *Kern von* ϕ.

Bemerkung II.5.28: Seien n und m natürliche Zahlen. Ist $A \in \mathbb{R}^{n\times m}$, dann gilt für die zugehörige Abbildung ϕ_A:

$$\text{Kern}(\phi_A) = \{v \in \mathbb{R}^m \mid Av = \mathbf{0}\} = \mathbb{L}^h = \mathbb{L}(A|0).$$

Bemerkung II.5.29: Seien n und m natürliche Zahlen, $A \in \mathbb{R}^{n\times m}$ und $b \in \mathbb{R}^n$. Dann folgt aus Korollar II.5.22:

(i) Ein Vektor $b \in \mathbb{R}^n$ gehört zu $\text{Bild}(\phi_A)$ genau dann, wenn es $v \in \mathbb{R}^m$ mit $Av = b$ gibt, d. h. wenn $\text{Rang}(A) = \text{Rang}(A|b)$ ist.

(ii) Als Übungsaufgabe zeigen Sie, dass ϕ_A injektiv ist genau dann, wenn $\text{Kern}(\phi_A) = \{\mathbf{0}\}$. Nach der vorangegangenen Bemerkung ist $\text{Kern}(\phi_A) = \{\mathbf{0}\}$ genau dann, wenn $\mathbb{L}^h = \{\mathbf{0}\}$, was wiederum äquivalent zu $\text{Rang}(A) = m$ ist.

(iii) Per Definition ist ϕ_A surjektiv genau dann, wenn $\text{Bild}(\phi_A) = \mathbb{R}^n$. Das aber ist äquivalent zu $\text{Rang}(A) = n$.

Kapitel III.

Strukurmathematik: Gruppen, Ringe, Körper

Das Ziel dieses Kapitels ist die Verallgemeinerung auf andere Zahlbereiche als die bisher Bekannten. Darüber hinaus werden wir die strukturellen Eigenschaften dieser allgemeineren Zahlbereiche untersuchen.

1. Gruppen

In diesem Abschnitt wollen wir die Struktur von $(\mathbb{R}, +)$ untersuchen.

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

(i) Eine Abbildung $*\colon M \times M \to M$ heißt *Verknüpfung auf* M. Sind m_1 und m_2 aus M, dann schreiben wir üblicherweise $m_1 * m_2$ anstelle von $*(m_1, m_2)$.

(ii) Gilt für alle $a, b, c \in M$, dass $a * (b * c) = (a * b) * c$, dann heißt die Verknüpfung „$*$“ *assoziativ*.

(iii) Gilt für alle $a, b \in M$, dass $a * b = b * a$, dann heißt die Verknüpfung „$*$“ *kommutativ*.

(iv) Gibt es $e \in M$, sodass „$g * e = g = e * g$“ für alle $g \in M$ gilt, dann heißt e *neutrales Element*. In diesem Fall ist e eindeutig bestimmt. Wäre nämlich e' auch ein neutrales Element, dann wäre $e' = e' * e = e$.

(v) Seien „$*$“ eine assoziative Verknüpfung auf M mit neutralem Element e und g ein Element von M. Gibt es $h \in M$ mit $g * h = e = h * g$, dann heißt h *inverses Element zu* g oder einfach *Inverses zu* g. Auch Inverse sind eindeutig, ist nämlich h' ein weiteres Inverses zu g, dann haben wir $h' = h' * e = h' * (g * h) = (h' * g) * h = e * h = h$. Wir schreiben $g^{-1} := h$.

Definition III.1.2 (Gruppe): Seien G eine Menge und „$*$“ eine Verknüpfung auf G. Falls „$*$“ assoziativ ist, es ein neutrales Element $e \in G$ bezüglich „$*$“ gibt und es für jedes $g \in G$ ein Inverses bezüglich „$*$“ gibt, dann heißt $(G, *)$ eine *Gruppe*.

Ist „$*$“ zusätzlich kommutativ, so heißt $(G, *)$ *kommutative Gruppe* oder auch *abelsche Gruppe*.

Beispiel III.1.3: Das Folgende sind Gruppen:

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

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

(iii) $\mathrm{Gl}_n(\mathbb{R})$ zusammen mit der Matrizenmultiplikation. Diese Gruppe ist nicht abelsch.

Beispiel III.1.4 (Gruppe der Kongruenzklassen): Sei n eine natürliche Zahl. Auf $\mathbb{Z}$ wird eine Äquivalenzrelation „$\equiv_n$“ durch

$$a \equiv b \pmod{n} :\Longleftrightarrow n \mid a - b$$

erklärt (siehe Proposition I.6.7). Sei $\mathbb{Z}/n\mathbb{Z} := \{[0], [1], \ldots, [n-1]\}$. Auf $\mathbb{Z}/n\mathbb{Z}$ definiert $[a] + [b] := [a + b]$ eine ehrliche Verknüpfung, die $\mathbb{Z}/n\mathbb{Z}$ zu einer abelschen Gruppe macht.

Beweis: Die Verknüpfung „$+$“ ist wohldefiniert, d. h. hängt nur von den Klassen $[a]$ und $[b]$, jedoch nicht von den gewählten Vertretern a und b ab. Sind a' und b' ganze Zahlen mit $[a] = [a']$ und $[b] = [b']$, dann gilt $n \mid a - a'$ und $n \mid b - b'$, also $n \mid (a - a') + (b - b') = (a + b) - (a' + b')$ und das heißt gerade $[a + b] = [a' + b']$.

Die Verknüpfung „$+$“ ist außerdem assoziativ, da die Addition auf den ganzen Zahlen additiv ist. Das neutrale Element in $\mathbb{Z}/n\mathbb{Z}$ ist $[0]$ und für ein beliebiges $[a] \in \mathbb{Z}/n\mathbb{Z}$ ist $[-a]$ das inverse Element. □

Im Folgenden schreiben wir auch $\bar{a} := [a]$ für die Restklasse der ganzen Zahl a in $\mathbb{Z}/n\mathbb{Z}$, d. h. $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots \overline{n-1}\}$.

Definition III.1.5 (Untergruppe): Seien $(G, *)$ eine Gruppe und H eine Teilmenge von G. Falls gilt:

(i) Das neutrale Element e von G gehört zu H.

(ii) Für alle h_1, h_2 aus H gehört $h_1 * h_2$ zu H.

(iii) Für alle $h \in H$ gehört h^{-1} zu H,

dann heißt H *Untergruppe von* G.

Bemerkung III.1.6: In der Situation von Definition III.1.5 ist

$$*|_{H\times H}\colon H \times H \longrightarrow H, \qquad (h_1, h_2) \longmapsto h_1 * h_2$$

eine Verknüpfung auf H (denn durch Definition III.1.5(ii) wird sicher gestellt, dass das Bild von $*|_{H\times H}$ tatsächlich in H enthalten ist) und $(H, *)$ ist eine Gruppe in eigenem Recht.

Proposition III.1.7 (Untergruppenkriterium): *Es seien $(G, *)$ eine Gruppe und H eine Teilmenge von G. Genau dann ist H eine Untergruppe von G, wenn gilt:*

(1)′ *Die Teilmenge H ist nichtleer.*

(2)′ *Für alle h_1, h_2 in H gehört $h_1 * h_2^{-1}$ zu H.*

Beweis: Sei e das neutrale Element von $(G, *)$. „$\Longrightarrow$“: Ist H eine Untergruppe von G, dann ist nach (i) ...

„$\Longleftarrow$“: Da H nichtleer ist, gibt es irgendein Element $h \in H$. Nach (ii)' gehört $h * h^{-1} = e$ zu H, d. h. (i) gilt.

Aus (ii)' folgt weiter, dass für alle $h \in H$ auch $h^{-1} = e * h^{-1}$ zu H gehört, d. h. (iii) gilt.

Schließlich haben wir für alle $h_1, h_2 \in H$, dass h_2^{-1} zu H gehört und wegen (ii)' ist dann $h_1 * (h_2^{-1})^{-1} = h_1 * h_2$ ein Element von H. □

Proposition III.1.8: *Seien I eine nichtleere Menge, $(G, *)$ eine Gruppe und $(H_i)_{i\in I}$ eine Familie von Untergruppen von $(G, *)$. Dann ist $H := \bigcap_{i\in I} H_i$ eine Untergruppe von $(G, *)$.*

Beweis: Mit offensichtlichen Anpassungen greift der Beweis von Proposition II.2.10. □

Definition III.1.9 (Erzeugte Untergruppe, zyklische Gruppe): Seien $(G, *)$ eine Gruppe und M eine Teilmenge von G. Definiere

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

Dann heißt $\langle M\rangle := \bigcap_{H\in I} H$ das *Erzeugnis von M* oder die von M *erzeugte Untergruppe*.

Ist $M = \{g\}$, dann heißt $\langle M\rangle$ eine *zyklische Gruppe*. Wir schreiben auch $\langle g\rangle := \langle M\rangle = \langle\{g\}\rangle$.

Beispiel III.1.10: Sei $G = (\mathbb{Z}/6\mathbb{Z}, +)$. Dann ist $\langle [2] \rangle = \{[2], [4], [6]\}$, oder $\langle [1] \rangle = \mathbb{Z}/6\mathbb{Z}$. Was ist $\langle [5] \rangle$?

Definition III.1.11: Sei $(G, *)$ eine Gruppe.

(i) Die Anzahl der Elemente der Menge G heißt *Ordnung von* G, in Zeichen $\mathrm{ord}(G) := \#(G)$.

(ii) Für ein Element g von G heißt $\mathrm{ord}(g) := \#\langle g \rangle$ die *Ordnung von* g.

Beispiel III.1.12: Für $G = (\mathbb{Z}/6\mathbb{Z}, +)$ und $g = \bar{2}$ haben wir $\langle \bar{2} \rangle = \{\bar{0}, \bar{2}, \bar{4}\}$, d. h. $\mathrm{ord}(\bar{2}) = 3$.

Proposition III.1.13 (Ordnung als minimale Potenz): *Seien $(G, *)$ eine Gruppe mit neutralem Element e und g ein Element von g. Notiere $g^1 := g$ und für eine natürliche Zahl k notiere $g^k := g * g^{k-1}$. Gibt es eine natürliche Zahl k, sodass $g^k = e$, dann ist*

$$\mathrm{ord}(g) = \min\{k \in \mathbb{N} \mid g^k = e\}.$$

Der Beweis dieser Aussage bleibt Ihnen als Übungsaufgabe überlassen.

Definition III.1.14 (Nebenklasse): Seien $(G, *)$ eine Gruppe mit neutralem Element e und $H \subseteq G$ eine Untergruppe. Auf G erklärt

$$g_1 \sim g_2 :\Longleftrightarrow g_1 * g_2^{-1} \in H$$

eine Relation „$\sim$“. Es gilt:

(i) Die Relation „$\sim$“ ist sogar eine Äquivalenzrelation mit den Äquivalenzklassen

$$[g] = H * g := \{h * g \mid h \in H\} \qquad (g \in G).$$

Die Äquivalenzklasse $H * g$ heißt *Rechtsnebenklasse von g bezüglich H*.

(ii) Für $g \in G$ ist die Abbildung

$$F_g \colon H \longrightarrow H * g, \qquad h \longmapsto h * g$$

eine Bijektion. Insbesondere gilt: Ist H endlich, dann haben alle Rechtsnebenklassen gleich viele Elemente, und zwar so viele wie H.

Beweis: (i) Dass es sich bei „$\sim$“ um eine Äquivalenzrelation handelt, ist leicht nachzuprüfen.

(ii) Per Konstruktion ist F_g surjektiv. Zur Injektivität: Seien $h_1, h_2 \in H$ mit $F_g(h_1) = F_g(h_2)$, d. h. $h_1 * g = h_2 * g$. Dann ist

$$h_1 = h_1 * g * g^{-1} = h_2 * g * g^{-1} = h_2,$$

d. h. F_g ist injektiv. □

Satz 5 (von Lagrange): *Es sei $(G, *)$ eine endliche Gruppe (d. h. $\mathrm{ord}(G) < \infty$). Für jede Untergruppe $H \subseteq G$ ist die Ordnung $\mathrm{ord}(H)$ ein Teiler von $\mathrm{ord}(G)$.*

Beweis: Seien H eine Untergruppe von G und „$\sim$“ die Äquivalenzrelation aus Definition III.1.14. Aus Satz 1 wissen wir, dass G die disjunkte Vereinigung seiner Nebenklassen ist.

Nach Definition III.1.14 haben alle Nebenklassen gleich viele Elemente, nämlich so viele wie H, d. h. $\mathrm{ord}(G) = k \cdot \mathrm{ord}(H)$, wobei k die Anzahl der Nebenklassen ist. □

Korollar III.1.15: *Sei p eine Primzahl. Ist G eine Gruppe mit $\mathrm{ord}(G) = p$, dann ist G zyklisch.*

Beweis: Seien $g \in G - \{e_G\}$ und $U := \langle g \rangle \subseteq G$. Dann gehören e_G und g zu $\langle g \rangle$, d. h. $\mathrm{ord}(g) \geq 2$. Nach dem Satz von Lagrange muss dann aber U schon ganz G sein, d. h. $G = \langle g \rangle$ und G ist zyklisch. □

2. Homomorphismen

In diesem Abschnitt wollen wir strukturerhaltende Abbildungen zwischen Gruppen studieren.

Definition III.2.1 (Gruppenhomomorphismus): Seien $(G, *)$ und $(H, \circ)$ Gruppen mit neutralen Elementen e_G und e_H.

(i) Sei $\varphi\colon (G, *) \to (H, \circ)$ eine Abbildung. Gilt für alle g_1, g_2, dass

$$\varphi(g_1 * g_2) = \varphi(g_1) \circ \varphi(g_2),$$

dann heißt φ ein *Gruppenhomomorphismus* oder *Homomorphismus von Gruppen*. Wir schreiben $\varphi\colon (G, *) \to (H, \circ)$ um klar zu machen, mit welchen Verknüpfungen G und H ausgestattet sind.

(ii) Für die Menge der Gruppenhomomorphismen von G nach H schreiben wir $\mathrm{Hom}(G, H) := \{\varphi\colon (G, *) \to (H, \circ)$ ist Gruppenhomomorphismus$\}$.

Beispiel III.2.2: (i) Sind n und m natürliche Zahlen und $A \in \mathbb{R}^{n\times m}$, dann ist $\varphi_A\colon (\mathbb{R}^m,+) \to (\mathbb{R}^n,+)$, $x \mapsto Ax$ ein Gruppenhomomorphismus.

(ii) Die Abbildung $\varphi\colon (\mathbb{R},+) \to (\mathbb{R}-\{0\},\cdot)$, $x \mapsto e^x$ ist ein Gruppenhomomorphismus.

(iii) Seien $(G,*)$ eine beliebige Gruppe mit neutralem Element e und $g \in G$. Dann ist

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

ein Gruppenhomomorphismus. Hierbei ist $g^0 := e$, für natürliche Exponenten k ist $g^k := g * g^{k-1}$ wie gehabt und für negative Exponenten ist $g^k := (g^{-1})^{-k}$.

Beweis: (i) Das haben wir in Bemerkung II.5.25 bereits festgehalten.

(ii) In der Analysis I wird nachgewiesen, dass für alle $x_1, x_2 \in \mathbb{R}$ gilt:

$$\varphi(x_1+x_2) = e^{x_1+x_2} = e^{x_1} + e^{x_2} = \varphi(x_1) + \varphi(x_2).$$

(iii) Wir unterscheiden drei Fälle. Ist erstens $k_1 + k_2 \geq 0$ mit $k_1, k_2 \geq 0$, dann ist

$$\varphi_g(k_1) * \varphi_g(k_2) = g^{k_1} * g^{k_2} = g^{k_1+k_2} = \varphi_g(k_1+k_2).$$

Ist zweitens $k_1 + k_2 \geq 0$ mit $k_1 > 0$, $k_2 < 0$ oder $k_1 < 0$, $k_2 > 0$, dann ist $k_1 \geq -k_2$ respektive $k_2 \geq -k_1$. Ohne Einschränkung der Allgemeinheit dürfen wir annehmen, dass $k_1 > 0$ und $k_2 < 0$. Dann haben wir

$$\begin{aligned}\varphi_g(k_1) * \varphi_g(k_2) &= g^{k_1} * (g^{-1})^{-k_2} \\ &= g^{k_1-1} * g * g^{-1} * (g^{-1})^{-k_2-1} \\ &= g^{k_1-1} * (g^{-1})^{-k_2-1} = \cdots = g^{k_1-(-k_2)} = g^{k_1+k_2} = \varphi_g(k_1+k_2).\end{aligned}$$

Ist drittens $k_1 + k_2 < 0$, dann ist $\varphi_g(k_1+k_2) = (g^{-1})^{-(k_1+k_2)} = \varphi_{g^{-1}}(-(k_1+k_2))$ und wir können mit dem ersten oder zweiten Fall umformen:

$$\varphi_{g^{-1}}(-(k_1+k_2)) = \varphi_{g^{-1}}(-k_1) * \varphi_{g^{-1}}(-h_2). \qquad \square$$

Bemerkung III.2.3: (i) Aus Beispiel III.2.2 folgt insbesondere das additive Potenzgesetz in beliebigen abelschen Gruppen. Genauer: Ist $(G,*)$ eine abelsche Gruppe, dann gilt:

$$\forall\, g \in G \,\forall\, k_1, k_2 \in \mathbb{Z} : g^{k_1+k_2} = g^{k_1} * g^{k_2}.$$

(ii) In Beispiel III.2.2 folgt (ii) aus (iii).

Proposition III.2.4 (Verknüpfung): *Für zwei Homomorphismen von Gruppen* $\varphi_1\colon (G_1, *) \to (G_2, \bullet)$ *und* $\varphi_2\colon (G_2, \bullet) \to (G_3, \triangle)$ *ist die Verknüpfung*

$$\varphi_2 \circ \varphi_1\colon (G_1, *) \longrightarrow (G_3, \triangle)$$

ebenfalls ein Gruppenhomomorphismus.

Beweis: Für $a, b \in G_1$ gilt:

$$\begin{aligned}(\varphi_2 \circ \varphi_1)(a * b) &= \varphi_2(\varphi_1(a * b)) \\ &= \varphi_2(\varphi_1(a) \bullet \varphi_2(b)) \\ &= \varphi_2(\varphi_1(a)) \triangle \varphi_2(\varphi_1(b)) = (\varphi_2 \circ \varphi_1)(a) \triangle (\varphi_2 \circ \varphi_1)(b). \quad \square\end{aligned}$$

Proposition III.2.5 (Erste Rechengesetze): *Seien* $(G, *)$ *und* $(H, \bullet)$ *Gruppen und* $\varphi\colon (G, *) \to (H, \bullet)$ *ein Homomorphismus von Gruppen. Dann gilt:*

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

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

Beweis: (i) Wegen $\varphi(e_G) = \varphi(e_G * e_G) = \varphi(e_G) \bullet \varphi(e_G)$ ist

$$e_H = \varphi(e_G)^{-1} \bullet \varphi(e_G) = \varphi(e_G)^{-1} \bullet \varphi(e_G) \bullet \varphi(e_G) = \varphi(e_G).$$

(ii) Da φ ein Homomorphismus von Gruppen ist, gilt $\varphi(g^{-1}) \bullet \varphi(g) = \varphi(g^{-1} \bullet g) = \varphi(e_H)$. Aus (i) wissen wir, dass $\varphi(e_G) = e_H$. Genau so rechnet man nach, dass $\varphi(g) \bullet \varphi(g^{-1}) = e_H$, d. h. $\varphi(g^{-1})$ ist tatsächlich das Inverse $\varphi(g)^{-1}$ zu $\varphi(g)$. $\square$

Proposition III.2.6 (Bild und Urbild): *Seien* $(G, *)$ *und* $(H, \bullet)$ *Gruppen und* $\varphi\colon (G, *) \to (H, \bullet)$ *ein Homomorphismus von Gruppen. Ferner seien* U_1 *eine Untergruppe von* G *und* U_2 *eine Untergruppe von* H*. Dann gilt:*

(i) *Das Bild* $\mathrm{Bild}(U_1) = \varphi(U_1) = \{\varphi(g) \mid g \in U_1\}$ *ist eine Untergruppe von* H*.*

(ii) *Das Urbild* $\varphi^{-1}(U_2) = \{g \in G \mid \varphi(g) \in U_2\}$ *ist eine Untergruppe von* G*.*

Beweis: Wir zeigen exemplarisch (ii), (i) lässt sich auf ähnliche Art nachweisen. Um (ii) zu zeigen, gehen wir die Untergruppenkriterien durch. Zunächst liefert Proposition III.2.5(i), dass $\varphi(e_G) = e_H \in U_2$, sodass e_G zu $\varphi^{-1}(U_2)$ gehört.

Dann haben wir für alle h_1 und h_2 aus $\varphi^{-1}(U_2)$, dass

$$\varphi(h_1 * h_2^{-1}) = \varphi(h_1) \bullet \varphi(h_2^{-1}) = \varphi(h_1) \bullet \varphi(h_2)^{-1} \in U_2,$$

sodass $h_1 * h_2^{-1}$ zu $\varphi^{-1}(U_2)$ gehört. Damit ist alles gezeigt. $\square$

Definition III.2.7 (Kern): Seien $(G, *)$ und $(H, \bullet)$ Gruppen und $\varphi\colon (G, *) \to (H, \bullet)$ ein Homomorphismus von Gruppen. Die Menge

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

heißt *Kern von* φ.

Proposition III.2.8 (Eigenschafen des Kerns): *Seien* $(G, *)$ *und* $(H, \bullet)$ *Gruppen und* $\varphi\colon (G, *) \to (H, \bullet)$ *ein Homomorphismus von Gruppen. Dann gilt:*

(i) *Der Kern von* φ *ist eine Untergruppe von* $(G, *)$.

(ii) *Der Homomorphismus* φ *ist injektiv genau dann, wenn* $\mathrm{Kern}(\varphi) = \{e_G\}$.

Beweis: Die erste Aussage ist eine Konsequenz von Proposition III.2.6, die zweite Aussage zeigt man wie Aufgabe 2 von Blatt 8. □

Beispiel III.2.9: Für die Gruppenhomomorphismen aus Beispiel III.2.2 erhalten wir die folgenden Bilder und Kerne:

(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) $\mathrm{Kern}(\varphi) = \mathrm{ord}(g)\mathbb{Z} := \{\mathrm{ord}(g)k \mid k \in \mathbb{Z}\}$ falls $\mathrm{ord}(g) < \infty$ und $\mathrm{Kern}(\varphi) = \{0\}$, falls $\mathrm{ord}(g) = \infty$, $\mathrm{Bild}(\varphi) = \langle g \rangle$.

Definition III.2.10: Seien $(G, *)$ und $(H, \bullet)$ Gruppen und $\varphi\colon (G, *) \to (H, \bullet)$ ein Homomorphismus von Gruppen.

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

(ii) Gibt es einen Homomorphismus von Gruppen $\psi\colon (H, \bullet) \to (G, *)$ mit $\psi \circ \varphi = \mathrm{id}_G$ und $\varphi \circ \psi = \mathrm{id}_H$, dann heißt φ ein *Isomorphismus von Gruppen.*

(iii) Ist φ ein Isomorphismus und ein Endomorphismus, dann heißt φ ein Automorphismus.

Proposition III.2.11: *Seien* $(G, *)$ *und* $(H, \bullet)$ *Gruppen und* $\varphi\colon (G, *) \to (H, \bullet)$ *ein Homomorphismus von Gruppen. Genau dann ist* φ *ein Isomorphismus, wenn* φ *bijektiv ist.*

Beweis: „$\Longrightarrow$“: Folgt aus Proposition III.2.10 und Proposition I.5.10.

„$\Longleftarrow$“: Wir müssen zeigen, dass für einen bijektiven Gruppenhomomorphismus $\varphi\colon (G,*) \to (H,\bullet)$ die Umkehrabbildung $\varphi^{-1}\colon (H,\bullet) \to (G,*)$ ebenfalls ein Gruppenhomomorphismus ist. Im Folgenden schreiben wir ψ für φ^{-1}. Seien h_1 und h_2 Elemente von H. Dann gilt

$$\begin{aligned}\psi(h_1 \bullet h_2) &= \psi(\varphi(\psi(h_1)) \bullet \varphi(\psi(h_2))) \\ &= \psi(\varphi(\psi(h_1) * \psi(h_2))) = \psi(h_1) * \psi(h_2).\end{aligned}$$ □

Bemerkung III.2.12: Sei $(G,*)$ eine Gruppe. Die Menge

$$\mathrm{Aut}(G) := \{\varphi\colon G \to G \mid \varphi \text{ ist Automorphismus}\}$$

ist mit Komposition von Abbildungen eine Gruppe mit neutralem Element id_G.

3. Die symmetrische Gruppe

Definition III.3.1: Sei n eine natürliche Zahl. Die Menge

$$S_n := \mathrm{Perm}(\{1,\dots,n\}) = \{\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\} \mid \sigma \text{ ist bijektiv}\}$$

ist mit der Komposition von Abbildungen eine Gruppe.

Beweis: Wir wissen dass Komposition von Abbildungen assoziativ ist. Bezüglich Komposition ist $\sigma = \mathrm{id}$ das neutrale Element und für $\sigma \in S_n$ ist σ^{-1} das inverse Element. □

Beispiel III.3.2 (Die Gruppe S_3): Für $n = 3$ haben wir die folgenden bijektiven Abbildungen von $\{1,2,3\}$ nach $\{1,2,3\}$ aufgelistet als Wertetabellen:

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

Bemerkung III.3.3: Sei n eine natürliche Zahl. Die Ordnung von S_n ist $n!$.

Bemerkung III.3.4: Sei n eine natürliche Zahl. Wir notieren eine Permutation $\sigma \in S_n$ über ihre Wertetabelle wie folgt:

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

Definition III.3.5 (Träger): Seien M eine nichtleere Menge und $\sigma \in \mathrm{Perm}(M)$ eine Permutation. Die Menge

$$\mathrm{Tr}(\sigma) := \{x \in M \mid \sigma(x) \neq x\}$$

heißt *Träger von* σ.

Seien $\sigma_1, \sigma_2 \in \mathrm{Perm}(M)$ zwei Permutationen. Gilt $\mathrm{Tr}(\sigma_1) \cap \mathrm{Tr}(\sigma_2) = \varnothing$, dann heißen σ_1 und σ_2 *disjunkt*.

Definition III.3.6 (Zyklen und Transpositionen): Sei M eine nichtleere Menge.

(i) Für Elemente $x_1, \dots, x_k \in M$ definieren wir eine Permutation ζ wie folgt:

$$\zeta(x) := \begin{cases} x_{i+1}, & \text{falls } x = x_i \text{ mit } i \in \{1, \dots, k-1\}, \\ x_1, & \text{falls } x = x_k, \\ x, & \text{falls } x \notin \{x_1, \dots, x_k\}. \end{cases}$$

So eine Permutation heißt *k-Zyklus*. Wir schreiben für die oben definierte Abbildung $\zeta = (x_1 \dots, x_k)$.

Ist $k > 1$, so ist $\mathrm{Tr}(\zeta) = \{x_1, \dots, x_k\}$. Die Zahl k heißt *Länge des Zyklus*.

(ii) Ein 2-Zyklus $\sigma \in \mathrm{Perm}(M)$ heißt *Transposition*.

Beispiel III.3.7: (i) In Proposition III.3.2 sind τ_1, τ_2 und τ_3 Transpositionen.

(ii) Die Permutation

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

ist kein Zyklus, aber $\sigma = (12) \circ (345)$.

(iii) In der S_7 gilt

$$(25) \circ (53) \circ (37) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 5 & 7 & 4 & 3 & 6 & 2 \end{pmatrix} = (2537)$$

Bemerkung III.3.8 (Eigenschaften von *k*-Zyklen): (i) Ist ζ ein k-Zyklus, dann ist $\operatorname{ord}(\zeta) = \min\{n \in \mathbb{N} \mid \zeta^n = \mathrm{id}\} = k$.

(ii) Für einen k-Zyklus $\zeta = (x_1 \dots x_k)$ gilt

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

Insbesondere ist jeder Zyklus eine Verkettung von Transpositionen.

Satz 6 (Zykelzerlegung von Permutationen): *Sei M eine endliche Menge. Jede Permutation $\sigma \in \mathrm{Perm}(M)$ ist Verkettung disjunkter Zyklen. Genauer: Es gibt Zyklen $\zeta_1, \dots, \zeta_k$ mit $\sigma = \zeta_1 \circ \cdots \circ \zeta_k$, sodass die Träger $\mathrm{Tr}(\zeta_1)$, …, $\mathrm{Tr}(\zeta_r)$ paarweise disjunkt sind. Für $r = 0$ ist $\sigma = \mathrm{id}$.*

Beweis: Wir zeigen die Aussage per vollständige Induktion über $N = \#\,\mathrm{Tr}(\sigma)$. Für $N = 0$ ist $\sigma = \mathrm{id}$ und die Behauptung ist klar.

Die Behauptung gelte für alle σ' mit $\#\,\mathrm{Tr}(\sigma') \leq N$. Sei weiter σ eine Permutation mit $\#\,\mathrm{Tr}(\sigma) = N + 1$. Wähle ein $x_0 \in \mathrm{Tr}(\sigma)$ und setze

$$k := \min\{\ell \in \mathbb{N} \mid \sigma^\ell(x_0) = x_0\}.$$

Weil x_0 zum Träger von σ gehört, ist k jedenfalls größer als Eins. Setzen wir jetzt $M := \{x_0, \sigma(x_0), \dots \sigma^{k-1}(\sigma)\}$ und $\zeta_1 := (x_0 \dots \sigma^{k-1}(x_0))$, dann ist $\sigma|_{M_1} = \zeta_1|_{M_1}$ und $\zeta_1|_{M-M_1} = \mathrm{id}\,|_{M-M_1}$. Weil außerdem $\sigma^k(x_0) = x_0$ folgt für $x \in M_1$, dass sowohl $\sigma(x)$ als auch $\sigma^{-1}(x)$ zu M_1 gehören und außerdem folgt für $x \in M - M_1$, dass auch $\sigma(x)$ zu $M - M_1$ gehört.

Setze nun $\sigma_1 := \zeta_1^{-1} \circ \sigma$. Für $x \in M_1$ haben wir $\sigma_1(x) = \zeta_1^{-1}(\sigma(x)) = x$ und für $x \in M - M_1$ gilt, dass $\sigma_1(x) = \zeta_1^{-1}(\sigma(x)) = \sigma(x)$. Insbesondere haben wir $\mathrm{Tr}(\sigma_1) \subseteq \mathrm{Tr}(\sigma) - M$, d. h. $\#\,\mathrm{Tr}(\sigma_1) \leq \#\,\mathrm{Tr}(\sigma) - 1 = N$. Nach Induktionsvoraussetzung gibt es disjunkte Zyklen $\zeta_2, \dots, \zeta_r$, wobei $r \in \mathbb{N}_0$, mit $\sigma_1 = \zeta_2 \circ \cdots \circ \zeta_r$ und $\mathrm{Tr}(\zeta_2)$, …, $\mathrm{Tr}(\zeta_r) \subseteq \mathrm{Tr}(\sigma_1)$. Es folgt für $2 \leq i \leq r$, dass $\mathrm{Tr}(\zeta_1) \cap \mathrm{Tr}(\zeta_i) = \varnothing$. Insgesamt sehen wir $\sigma = \zeta_1 \circ \sigma_1 = \zeta_1 \circ \cdots \circ \zeta_r$ mit paarweise disjunkten Zyklen ζ_1, …, ζ_r. □

Korollar III.3.9: *Sei M eine endliche Menge. Jede Permutation σ von M ist Verkettung von Transpositionen, d. h., es gibt Transpositionen $\tau_1, \dots, \tau_m$ in $\mathrm{Perm}(M)$, sodass $\sigma = \tau_1 \circ \cdots \circ \tau_m$.*

Beweis: Das folgt aus Satz 6 und Proposition III.3.8. □

Bemerkung III.3.10: Wir notieren Permutationen als Verkettung von disjunkten Zyklen und lassen das Verkettungszeichen meist weg. Zum Beispiel so:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 5 & 1 & 4 & 2 & 6 & 3 \end{pmatrix} = (173)(25).$$

Definition III.3.11 (Signum einer Permutation): Seien n eine natürliche Zahl und $\sigma \in S_n$. Dann heißt

$$\operatorname{sgn}(\sigma) := \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

das *Signum von* σ.

Beispiel III.3.12: (i) Für $\sigma = (1234) \in S_4$ ist

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

(ii) Wir betrachten die Transposition $\sigma = (12) \in S_n$. Für jedes Indexpaar (i, j) mit $i < j$ und $\sigma(i) > \sigma(j)$ erhalten wir einen Faktor -1 im Signum und für σ ist das genau das Paar $(i, j) = (1, 2)$, d. h. $\operatorname{sgn}(\sigma) = -1$.

Bemerkung III.3.13: Da in Proposition III.3.11 im Zähler und im Nenner bis auf Reihenfolge und Vorzeichen die gleichen Faktoren stehen, gilt $\operatorname{sgn}(\sigma) \in \{\pm 1\}$.

Für $\pi \in S_n$ gilt

$$\operatorname{sgn}(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{\pi(j) - \pi(i)}$$

denn π vertauscht nur die Reihenfolge der Faktoren.

Proposition III.3.14 (Signum respektiert Vorzeichen): *Für Permutationen σ und τ in S_n gilt* $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau)$. *Mit anderen Worten: Das Signum* $\operatorname{sgn}\colon (S_n, \circ) \to (\{\pm 1\}, \cdot)$ *ist ein Gruppenhomomorphismus.*

Bemerke, dass $(\{\pm 1\}, \cdot)$ tatsächlich eine Gruppe ist.

Beweis: Wir haben

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

wie gewünscht. □

Proposition III.3.15 (Konjugationstrick): *Seien M eine endliche Menge, a, b, a' und b' Elemente von M mit $a \neq b$ und $a' \neq b'$ und $\pi \in \mathrm{Perm}(M)$ mit $\pi(a') = a$ und $\pi(b') = b$. Dann gilt $\pi^{-1} \circ (ab) \circ \pi = (a'b')$.*

Beweis: Für $x \in M$ gilt

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

□

Beispiel III.3.16: Gegeben seien die Transpositionen $\tau_1 = (12)$ und $\tau_2 = (35)$. Wähle zum Beispiel

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

Dann ist $\pi^{-1} \circ (12) \circ \pi = (35)$.

Satz 7 (über die Signumsfunktion):

(i) *Die Signumsfunktion* $\mathrm{sgn}\colon (S_n, \circ) \to (\{\pm 1\}, \cdot)$ *ist ein Gruppenhomomorphismus.*

(ii) *Ist τ eine Transposition in S_n, dann ist* $\mathrm{sgn}(\tau) = -1$.

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

$$\mathrm{sgn}(\zeta) = \begin{cases} -1, & \textit{falls } \ell \textit{ gerade}, \\ 1, & \textit{falls } \ell \textit{ ungerade}. \end{cases}$$

Beweis: (i) Das folgt aus Proposition III.3.14.

(ii) Nach dem Konjugationstrick gibt es $\pi \in S_n$, sodass $\tau = \pi^{-1} \circ (12) \circ \pi$. Das Signum von (12) kennen wir aus Proposition III.3.12, das ist nämlich -1. Weil das Signum ein Gruppenhomomorphismus ist, gilt

$$\mathrm{sgn}(\tau) = \mathrm{sgn}(\pi^{-1} \circ (12) \circ \pi) = \mathrm{sgn}(\pi)^{-1} \,\mathrm{sgn}(12)\,\mathrm{sgn}(\pi) = \mathrm{sgn}(12) = -1.$$

(iii) Sei $\zeta = (x_1, \ldots, x_\ell)$ ein ℓ-Zyklus mit $x_1, \ldots, x_\ell \in \{1, \ldots, n\}$. Nach Proposition III.3.8 ist

$$\zeta = (x_1 \ldots x_\ell) = (x_1 x_2) \circ \cdots \circ (x_{\ell-1} x_\ell),$$

sodass nach Teil (ii) gilt: $\mathrm{sgn}(\zeta) = (-1)^{\ell-1}$. Genau das haben wir behauptet. □

Bemerkung III.3.17: Mithilfe von Satz 7 lässt sich das Signum einer beliebigen Permutation $\sigma \in S_n$ berechnen. Zum Beispiel für

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 4 & 1 & 2 & 9 & 8 & 3 & 5 & 10 & 6 \end{pmatrix} = (173)(24)(5\,9\,10\,6\,8)$$

ist $\operatorname{sgn}(\sigma) = \operatorname{sgn}((173)(24)(5\,9\,10\,6\,8)) = 1 \cdot (-1) \cdot 1 = (-1)$.

4. Ringe

Definition III.4.1 (Ring): Sei R eine Menge mit zwei Verknüpfungen „$+$“ und „$\cdot$“. Falls gilt:

(i) $(R, +)$ ist eine abelsche Gruppe.

(ii) „$\cdot$“ ist assoziativ.

(iii) Es gelten die Distributivgesetze, d. h. für alle x, y, z aus R ist

$$x \cdot (y + z) = xy + xz, \qquad (y + z) \cdot x = yx + zx.$$

dann heißt R ein *Ring*.

Ist „$\cdot$“ kommutativ, dann heißt R ein *kommutativer Ring*.

Gibt es ein neutrales Element 1_R bezüglich „$\cdot$“, d. h. für alle $x \in R$ gilt $x \cdot 1_R = x = 1_R \cdot x$, dann heißt R ein *Ring mit Eins* oder *unitärer Ring*.

In der Situation von Proposition III.4.1 heißt „$+$“ *Addition* und „$\cdot$“ *Multiplikation*. Für das neutrale Element von $(R, +)$ schreiben wir 0_R und nennen es die *Null von* R, für das neutrale Element von $(R, \cdot)$ schreiben wir 1_R und nennen es die *Eins von* R. Für ein $x \in R$ notieren wir das Inverse bezüglich „$+$“ mit $-x$ und für $x, y \in R$ schreiben wir $x - y := x + (-y)$. Für ein bezüglich „$\cdot$“ invertierbares x notieren wir das Inverse als x^{-1}.

Wir vereinbaren, dass „$\cdot$“ stärker bindet als „$+$“, d. h. für x, y und z aus R schreiben wir $x \cdot y + z := (x \cdot y) + z$ (bekannt als „Punkt vor Strich“).

Ist R ein kommutativer Ring, sind $x, y \in R$ und ist y invertierbar bezüglich „$\cdot$“, dann schreiben wir $x/y := x \cdot y^{-1}$, insbesondere schreiben wir $1/y = y^{-1}$.

Beispiel III.4.2 (Erste Beispiel-Ringe): (i) Die klassischen Zahlbereiche $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ mit der gewöhnlichen Addition und Multiplikation sind kommutative Ringe mit Eins.

(ii) $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ mit den durch $[a] + [b] := [a + b]$, $[a] \cdot [b] := [ab]$ definierten Verknüpfungen ist ein kommutativer Ring mit Eins.

(iii) Sei n eine natürliche Zahl. Dann ist $\mathbb{R}^{n\times n}$ mit Matrizenaddition und Matrizenmultiplikation ein Ring mit Eins.

Proposition III.4.3 (Erste Eigenschaften): *Sei $(R,+,\cdot)$ ein Ring. Dann gilt:*

(i) *Für alle $x \in R$ ist $0_R \cdot x = 0_R = x \cdot 0_R$.*

(ii) *Für alle $x, y \in R$ ist $(-x)\cdot y = -(x\cdot y) = x\cdot(-y)$.*

Beweis: (i) Weil 0_R das neutrale Element bezüglich „+“ ist, gilt $0_R+0_R = 0_R$, und deshalb ist $0_R \cdot x = (0_R + 0_R)\cdot x = 0_R\cdot x + 0_R\cdot x$. Damit rechnen wir nach, dass

$$0_R = -(0_R\cdot x) + 0_R\cdot x = -(0_R\cdot x) + 0_R\cdot x + 0_R\cdot x = 0_R\cdot x.$$

(ii) Wir zeigen exemplarisch für $(-x)\cdot y$, dass es das additive Inverse von $x\cdot y$ ist:

$$(-x)\cdot y + x\cdot y = x\cdot y + (-x)\cdot y = (x-x)\cdot y = 0_R\cdot y = 0_R.$$

Bei der letzten Gleichheit haben wir dabei (i) verwendet. □

Definition III.4.4 (Ringhomomorphismus): Seien $(R,+_R,\cdot_R)$ und $(S,+_S,\cdot_S)$ Ringe und $\varphi\colon R\to S$ eine Abbildung.

(i) Gilt für alle $x, y \in R$, dass

$$\varphi(x +_R y) = \varphi(x) +_S \varphi(y), \qquad \varphi(x\cdot_R y) = \varphi(x)\cdot_S \varphi(y),$$

dann heißt φ ein *Homomorphismus von Ringen* oder *Ringhomomorphismus*.

(ii) Sind $(R,+_R,\cdot_R)$ und $(S,+_S,\cdot_S)$ sogar Ringe mit Eins und gilt zusätzlich $\varphi(1_R) = 1_S$, dann heißt φ ein *Homomorphismus von Ringen mit Eins*.
Wir schreiben $\varphi\colon (R,+_R,\cdot_R) \to (S,+_S,\cdot_S)$. Insbesondere ist φ ein Gruppenhomomorphismus von $(R,+_R)$ und $(S,+_S)$.

(iii) Die Menge $\mathrm{Kern}(\varphi) := \{r \in R \mid \varphi(r) = 0\}$ heißt *Kern von φ*.

(iv) Ist $R = S$, dann heißt φ ein *Ringendomorphismus*.

(v) Gibt es einen Homomorphismus von Ringen $\psi\colon (S,+_S,\cdot_S)\to(R,+_R,\cdot_R)$ mit $\varphi\circ\psi = \mathrm{id}_S$ und $\psi\circ\varphi = \mathrm{id}_R$, dann heißt φ ein *Ringisomorphismus*.

(vi) Ist φ sowohl ein Ringendomorphismus als auch ein Ringisomorphismus, dann heißt φ ein *Ringautomorphismus*.

(vii) Gibt es einen Isomorphismus $\varphi\colon R \to S$, dann heißen die Ringe R und S *isomorph.*

(viii) Wir schreiben

$$\begin{aligned}\operatorname{Hom}(R,S) &:= \operatorname{Hom}_{\mathrm{Ring}}(R,S)\\ &:= \{\varphi\colon R \to S \mid \varphi \text{ ist ein Ringhomomorphismus}\}\end{aligned}$$

für die Menge der Ringhomomorphismen von R nach S.

Im Folgenden schreiben wir einfach „$+$“ und „$\cdot$“ für die jeweiligen Ringverknüpfungen.

Bemerkung III.4.5 (Komposition): Seien $(R_1,+,\cdot)$, $(R_2,+,\cdot)$ und $(R_3,+,\cdot)$ Ringe und $\varphi_1\colon (R_1,+,\cdot) \to (R_2,+,\cdot)$ sowie $\varphi_2\colon (R_2,\cdot,+) \to (R_3,\cdot,+)$ Ringhomomorphismen. Dann ist auch die Komposition

$$\varphi_2 \circ \varphi_1\colon (R_1,+,\cdot) \longrightarrow (R_3,+,\cdot)$$

ein Homomorphismus von Ringen. Sind die Ringe R_1, R_2, R_3 sogar Ringe mit Eins und sind φ_1 und φ_2 Homomorphismen von Ringen mit Eins, dann ist auch $\varphi_2 \circ \varphi_1$ ein Homomorphismus von Ringen mit Eins.

Definition III.4.6 (Teilring): Seien $(R,+,\cdot)$ ein Ring und T eine Teilmenge von R. Gilt

(i) $0_R \in T$,

(ii) Für alle $t_1, t_2 \in T$ ist $t_1 + t_2 \in T$,

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

(iv) Für alle $t_1, t_2 \in T$ ist $t_1 \cdot t_2 \in T$,

dann heißt T ein *Teilring von R* oder *Unterring von R.*

Ist R sogar ein Ring mit Eins und gilt zusätzlich $1_R \in T$, dann heißt T ein *Teilring mit Eins.*

Bemerkung III.4.7 (Teilring als Ring): In Definition I.3.1 ist insbesondere $(T,+,\cdot)$ selbst wieder ein Ring.

Proposition III.4.8 („Additive Einsen-Abbildung“): *Sei $(R,+,\cdot)$ ein Ring mit Eins. Die Abbildung $\Phi\colon \mathbb{Z} \to R$ definiert durch*

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

ist ein Ringhomomorphismus von Ringen mit Eins.

Beweis: Nach Beispiel III.2.2 ist $\Phi\colon (\mathbb{Z},+) \to (R,+)$ ein Gruppenhomomorphismus. Per Definition gilt $\Phi(1) = 1_R$ und dass für alle $a, b \in \mathbb{Z}$ gilt, dass $\Phi(ab) = \Phi(a)\Phi(b)$, lässt sich leicht nachrechnen. □

Definition III.4.9 (Charakteristik): Seien $(R,+,\cdot)$ ein Ring mit Eins und Φ der Homomorphismus aus Proposition III.4.8. Dann heißt

$$\operatorname{char}(R) := \begin{cases} 0, & \text{falls für alle } k > 0 \text{ gilt: } \Phi(k) \neq 0_R, \\ \min\{k \in \mathbb{N} \mid \Phi(k) = 0\}, & \text{sonst} \end{cases}$$

die *Charakteristik von R*.

Ist die Charakteristik eines unitären Ringes R verschieden von Null, dann ist $\operatorname{char}(R) = \operatorname{ord}(1_R)$ in $(R,+)$. Es ist $\operatorname{Kern}(\Phi) = \langle \operatorname{char}(R) \rangle \subseteq \mathbb{Z}$.

Ist die Charakteristik eines Ringes Null, dann findet sich eine Kopie der ganzen Zahlen in diesem Ring. In Ringen mit endlicher Charakteristik findet sich keine Kopie der ganzen Zahlen, sondern eine Kopie von $\mathbb{Z}/p\mathbb{Z}$ für eine Primzahl p.

Definition III.4.10 (Einheitengruppe): Sei $(R,+,\cdot)$ ein Ring mit Eins. Die Menge

$$R^\times := \{r \in R \mid \text{Es gibt } s \in R \text{ mit } rs = sr = 1_R\}$$

bildet zusammen mit der Ringmultiplikation eine Gruppe mit neutralem Element 1_R und heißt *Einheitengruppe des Rings R*.

Definition III.4.11 (Polynome): Sei $(R,+,\cdot)$ ein Ring mit Eins.

(i) Ist $(a_i)_{i\in\mathbb{N}_0}$ eine Folge mit Einträgen a_i aus R und gibt es $N \in \mathbb{N}$, sodass für alle $i > N$ gilt $a_i = 0$, dann heißt $p = (a_i)_{i\in\mathbb{N}_0}$ ein *Polynom über R*. Wir schreiben $X := (\delta_{i1})_{i\in\mathbb{N}_0}$ und entsprechend $p = \sum_{i=0}^{N} a_i X^i$.

(ii) Die Zahl

$$\operatorname{Grad}(p) := \begin{cases} \max\{i \in \mathbb{N}_0 \mid a_i \neq 0\}, & \text{falls } p \neq (0,0,0,\dots) =: 0, \\ -\infty, & \text{falls } p = 0 \end{cases}$$

heißt der *Grad des Polynoms p*.

Wir schreiben $R[X] := \{p \mid p \text{ ist Polynom über } R\}$ für die Menge der Polynome über R und nennen $R[X]$ den *Polynomring über R*.

Bemerkung III.4.12 (Polynomring): Ist $(R, +, \cdot)$ ein Ring mit Eins, $R[X]$ die Menge der Polynome über R und gehören $p = \sum_{i=0}^{N_1} a_i X^i$, $q = \sum_{j=0}^{N_2} b_j X^j$ zu $R[X]$, dann definieren die Festsetzungen

$$p + q := \sum_{i=0}^{\max\{N_1,N_2\}} (a_i + b_i) \cdot X^i, \qquad p \cdot q := \sum_{i=0}^{N_1+N_2} c_i X^i \quad \left(c_i := \sum_{k=0}^{i} a_k b_{i-k}\right)$$

Verknüpfungen auf $R[X]$, die $R[X]$ zu einem Ring machen.

Definition III.4.13 (Körper): Sei $(R, +, \cdot)$ ein Ring mit Eins. Falls gilt:

(i) R ist kommutativ,

(ii) $0_R \neq 1_R$,

(iii) $R^\times = R - \{0\}$, d. h. für alle $r \in R - \{0\}$ gibt es $s \in R$ mit $rs = 1_R = sr$,

dann heißt R ein *Körper*.

Proposition III.4.14 (Der Körper $\mathbb{F}_p$): *Ist p eine Primzahl, so ist $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ ein Körper. Wir schreiben $\mathbb{F}_p := (\mathbb{Z}/p\mathbb{Z}, +, \cdot)$.*

Beweis: Als Erinnerung halten wir fest, dass $\mathbb{Z}/p\mathbb{Z}$ die Menge der Äquivalenzklassen von $\mathbb{Z}$ bezüglich „$\equiv_p$“ ist, d. h. $\mathbb{Z}/p\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{p-1}\}$. Sei $a \in \mathbb{Z}$ mit $\bar{a} \neq 0$. Nach Korollar III.1.15 ist $\mathbb{Z}/p\mathbb{Z}$ zyklisch und jedes von Null verschiedene Element erzeugt $(\mathbb{Z}/p\mathbb{Z}, +)$, d. h. $\langle\bar{1}\rangle = \langle\bar{a}\rangle = \mathbb{Z}/p\mathbb{Z}$. Aber das heißt es gibt eine ganze Zahl b, sodass $\bar{1} = \overline{ab} = \bar{a}\bar{b} = \bar{b}\bar{a}$. □

Beispiel III.4.15 (Die komplexen Zahlen): In der Analysis zeigt man, dass $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ zusammen mit den Verknüpfungen

$$(a_1, b_1)+(a_2, b_2) := (a_1+a_2, b_1+b_2), \quad (a_1, b_1)\cdot(a_2, b_2) := (a_1a_2-b_1b_2, a_1b_2+a_2b_1)$$

einen Körper bildet. Man schreibt üblicherweise $a + \mathrm{i}b := (a, b)$. Insbesondere gilt für $\mathrm{i} = (0, 1)$, dass $\mathrm{i}^2 = -1$.

Kapitel IV.

Vektorräume und Dimensionstheorie

In diesem Kapitel wollen wir das bereits bekannte Konzept des $\mathbb{R}$-Vektorraums verallgemeinern zum Vektorraum über einem beliebigen Körper und die Eigenschaften von Vektorräumen näher studieren.

1. Vektorräume

Definition IV.1.1 (Vektorraum): Seien K ein Körper und V eine Menge zusammen mit einer Verknüpfung $+\colon V \times V \to V$ und einer äußeren Verknüpfung $\cdot\colon K \times V \to V$. Falls gilt:

(i) $(V, +)$ ist eine abelsche Gruppe,

(ii) Für alle $v \in V$ ist $1_K \cdot v = v$,

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

(iv) Für alle $\lambda \in K$ und $v_1, v_2 \in V$ ist $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$,

(v) Für alle $\lambda_1, \lambda_2 \in K$ und $v \in V$ ist $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$,

dann heißt $(V, +, \cdot)$ ein *Vektorraum über K* oder *K-Vektorraum*.

Beispiel IV.1.2 (der K^n): Sei K ein Körper. Dann wird das n-fache kartesische Produkt K^n von K mit sich selbst, also $K^n := \{(x_1, \ldots, x_n)^t \mid x_1, \ldots, x_n \in K\}$, mit den Verknüpfungen erklärt durch

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

wobei $(x_1, \ldots, x_n)^t$, $(y_1, \ldots, y_n)^t \in K^n$ und $\lambda \in K$, zu einem Vektorraum über K.

Bemerkung IV.1.3: (i) Für $K = \mathbb{R}$ stimmt Definition IV.1.1 mit Definition II.2.1 überein.

(ii) Alle Aussagen in Kapitel II gelten genauso für Vektorräume, Untervektorräume, Matrizen und lineare Gleichungssysteme über einem beliebigen Körper K.

(iii) Den K-Vektorraum der $(n \times m)$-Matrizen über K notieren wir im Folgenden entsprechend als $K^{n\times m}$.

Definition IV.1.4 (Vektorraumhomomorphismus): Seien V und W Vektorräume über dem Körper K.

(i) Sei $\Phi\colon V \to W$ eine Abbildung. Gilt für alle $v_1, v_2 \in V$ und $\lambda \in K$, dass

$$\Phi(v_1 + v_2) = \Phi(v_1) + \Phi(v_2), \qquad \Phi(\lambda v_1) = \lambda\Phi(v_1),$$

dann heißt Φ ein *Homomorphismus von K-Vektorräumen* oder *K-Vektorraumhomomorphismus* oder *K-lineare Abbildung*. Ist der zugrundeliegende Körper aus dem Kontext klar, dann wird die Erwähnung des Körpers üblicherweise ausgelassen.

(ii) Sei $\Phi\colon V \to W$ ein Vektorraumhomomorphismus. Ist $V = W$, dann heißt Φ ein *Endomorphismus*. Gibt es einen Vektorraumhomomorphismus $\Psi\colon W \to V$ mit $\Psi \circ \Phi = \mathrm{id}_V$ und $\Phi \circ \Psi = \mathrm{id}_W$, dann heißt Φ ein *Isomorphismus*. Ist Φ sowohl ein Endomorphismus als auch ein Isomorphismus, dann heißt Φ ein Automorphismus.

(iii) Gibt es einen Isomorphismus $\Phi\colon V \to W$, dann heißen die Vektorräume V und W *isomorph*, geschrieben $V \cong W$.

Bemerkung IV.1.5: Seien K ein Körper und n, m natürliche Zahlen. Aus Kapitel II wissen wir insbesondere:

(i) Für $A \in K^{n\times m}$ ist $\Phi\colon K^m \to K^n$, $v \mapsto Av$ eine lineare Abbildung und jede lineare Abbildung ist von dieser Art mit $A = (\Phi(e_1)|\dots|\Phi(e_m))$. Die Matrix A heißt *Abbildungmatrix* bezüglich der Standardbasen $(e_1, \dots, e_m)$ von K^m und $(e'_1, \dots, e'_n)$ von K^n.

Weiterhin gilt:

(ii) Verkettungen linearer Abbildungen sind lineare Abbildungen.

(iii) Seien V und W Vektorräume über K und sei $\Phi\colon V \to W$ linear.

Ist $U_1 \subseteq V$ ein Untervektorraum, dann ist $\Phi(U_1) \subseteq W$ ein Untervektorraum.

Ist $U_2 \subseteq W$ ein Untervektorraum, dann ist das Urbild $\Phi^{-1}(U_2) \subseteq V$ ein Untervektorraum.

Der Kern $\operatorname{Kern}(\Phi) := \{v \in V \mid \Phi(v) = 0\} \subseteq V$ ist ein Untervektorraum.

(iv) Eine lineare Abbildung $\Phi\colon V \to W$ ist ein Isomorphismus genau dann, wenn sie bijektiv ist.

(v) Für zwei K-Vektorräume V und W ist

$$\operatorname{Hom}_K(V, W) := \{\Phi\colon V \to W \mid \Phi \text{ ist linear}\} \subseteq \operatorname{Abb}(V, W)$$

ein Untervektorraum von $\operatorname{Abb}(V, W)$. Wir schreiben auch kurz $\operatorname{Hom}(V, W)$, falls K aus dem Kontext klar ist.

2. Basen und lineare Unabhängigkeit

Ist K ein Körper und V der Vektorraum K^n, dann wissen wir bereits, dass sich ein Vektor $x = (x_1, \ldots, x_n)^t$ eindeutig als $x = \sum_{i=1}^n x_i e_i$ schreiben lässt.

Genau so etwas haben wir auch für den Vektorraum $V = \mathbb{L}(A, \mathbf{0})$ schon gesehen – jeder Vektor in V lässt sich in eindeutigerweise schreiben als Linearkombination der Fundamentallösungen $F^{(1)}, \ldots, F^{(k)}$.

Dieses Phänomen wollen wir mit dem Namen „Basis“ belegen und uns damit beschäftigen, wie man solche Basen erkennt, wann wir etwas über die Existenz einer solchen Basis sagen können und wir wollen uns fragen, ob die Mächtigkeit einer solchen Basis von der speziellen Basis abhängt oder nicht.

Definition IV.2.1 (Linearkombination): Seien K ein Körper und V ein K-Vektorraum. Seien weiter $M \subseteq V$ eine Teilmenge und v ein Element von V. Gibt es eine nichtnegative ganze Zahl n, Vektoren $v_1, \ldots, v_n \in M$ und $\lambda_1, \ldots, \lambda_n \in K$, sodass

$$v = \sum_{i=1}^{n} \lambda_i v_i,$$

dann heißt v eine *Linearkombination von M*.

Ist in der Situation der obigen Definition $n = 0$, dann ist $v = \mathbf{0}$.

Bemerkung IV.2.2: Seien K ein Körper und V ein K-Vektorraum.

(i) Der Nullvektor ist Linearkombination für jede Teilmenge $M \subseteq V$

(ii) Ist $M = \varnothing$, dann ist $\mathbf{0}$ die einzige Linearkombination von M.

Definition IV.2.3 (Linearkombinationen revised): Seien K ein Körper und M eine beliebige Menge.

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

(ii) Wir setzen $\mathrm{Abb}_0(M, K) := \{f \in \mathrm{Abb}(M, K) \mid \#\,\mathrm{Tr}(f) < \infty\}$.

(iii) Ist V ein K-Vektorraum, $M \subseteq V$ eine Teilmenge und $\lambda \in \mathrm{Abb}_0(M, K)$ eine Abbildung mit $\mathrm{Tr}(\lambda) = \{v_1, \ldots, v_n\}$, dann ist

$$\sum_{v \in M} \lambda(v)v = \sum_{v \in \mathrm{Tr}(\lambda)} \lambda(v)v = \sum_{i=1}^{n} \lambda_i v_i$$

eine Linearkombination von M.

Definition IV.2.4 (Basis): Seien K ein Körper, V ein K-Vektorraum und B eine Teilmenge von V. Gibt es für jedes $v \in V$ genau ein $\lambda \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{w \in B} \lambda(w)w$, dann heißt B eine *Basis von V*.

Im Folgenden wollen wir nach Kriterien suchen, um eine Teilmenge $B \subseteq V$ als Basis zu erkennen.

Definition IV.2.5 (Lineare Unabhängigkeit): Seien K ein Körper, V ein K-Vektorraum und $M \subseteq V$ eine Teilmenge. Falls gilt: „Ist $\lambda \in \mathrm{Abb}_0(M, K)$ mit $\sum_{v \in M} \lambda(v)v = 0$, dann ist für alle $v \in V$ schon $\lambda(v) = 0$“, dann heißt M *linear unanhängig*. Ist M nicht linear unabhängig, dann heißt M *linear abhängig*.

Gilt in der Situation der obigen Definition für alle $v \in V$ dass $\lambda(v) = 0$, dann ist $\lambda = \mathbf{0}_{\mathrm{Abb}_0(M,K)}$.

Bemerkung IV.2.6: Seien K ein Körper und V ein K-Vektorraum. Eine Teilmenge $M = \{v_1, \ldots, v_n\} \subseteq V$ ist linear unabhängig genau dann, wenn gilt: „Sind $\lambda_1, \ldots, \lambda_n \in K$ mit $\lambda_1 v_1 + \cdots + \lambda_n v_n = \mathbf{0}$, dann sind $\lambda_1 = \cdots = \lambda_n = 0$“.

Definition IV.2.7 (Lineare Hülle): Seien K ein Körper, V ein K-Vektorraum und $M \subseteq V$ eine Teilmenge.

(i) Die Menge

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

heißt *lineare Hülle*, *Spann von M* oder auch *Erzeugnis von M*. Gebäruchlich ist auch die Schreibweise $\langle M \rangle := \mathrm{Lin}(M)$ und für eine endliche Menge $\{v_1, \ldots, v_n\}$ schreibt man oft $\langle v_1, \ldots, v_n \rangle := \mathrm{Lin}(\{v_1, \ldots, v_n\})$.

(ii) Ist $V = \mathrm{Lin}(M)$, dann heißt M ein *Erzeugendensystem von* V.

Beispiel IV.2.8: (i) Für die Vektoren $v_1 = (1,0,0)^t$ und $v_2 = (0,1,0)^t$ in K^3 gilt $\langle v_1, v_2\rangle = \{(a,b,0)^t \mid a,b \in K\} \cong K^2$.

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

Proposition IV.2.9 (Eigenschaften der linearen Hülle): *Seien K ein Körper, V ein K-Vektorraum und $M \subseteq V$ eine Teilmenge.*

(i) *Es gilt* $M \subseteq \mathrm{Lin}(M)$.

(ii) *Die lineare Hülle von M ist ein Untervektorraum von V. Genauer: Mit $S = \{U \mid U$ ist Untervektorraum von V, $M \subseteq U\}$ ist* $\mathrm{Lin}(M) = \bigcap_{U\in S} U$, *d. h.* $\mathrm{Lin}(M)$ *ist der kleinste Untervektorraum von V, der M enthält.*

(iii) *Für $M' \subseteq V$ gilt: Ist $M \subseteq M'$, dann ist* $\mathrm{Lin}(M) \subseteq \mathrm{Lin}(M')$.

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

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

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

$$\mathrm{Lin}(U_1 \cup U_2) = U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}.$$

Beweis: (i) Sei v aus M gegben. Dann ist $v = 1 \cdot v$ eine Linearkombination von M, d. h. $v \in \mathrm{Lin}(M)$.

(ii) Zunächst zeigen wir, dass $\mathrm{Lin}(M)$ tatsächlich ein Untervektorraum ist. Weil die Null aus jeder Menge linearkombiniert werden kann, gehört sie auch zu $\mathrm{Lin}(M)$. Sind v und w Elemente aus $\mathrm{Lin}(M)$, dann gibt es nichtnegative ganze Zahlen n und m, Elemente $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ von M und Körperelemente $\lambda_1, \ldots, \lambda_n$ sowie $\mu_1, \ldots, \mu_m \in K$, sodass $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{j=1}^m \mu_j w_j$. Aber dann ist

$$v + w = \lambda_1 v_1 + \cdots + \lambda_n v_n + \mu_1 w_1 + \ldots \mu_m w_m$$

eine Linearkombination von M, d. h. $v + w \in \mathrm{Lin}(M)$. Ist jetzt $\alpha \in K$ und v wie oben, dann ist auch $\alpha v = \sum_{i=1}^n (\alpha\lambda_i) v_i \in \mathrm{Lin}(M)$.

Nun zur anderen Behauptung: Gehört v zu $\mathrm{Lin}(M)$, dann gibt es Elemente $v_1, \ldots, v_n$ von M und $\lambda_1, \ldots, \lambda_n$ aus K, sodass $v = \sum_{i=1}^n \lambda_i v_i$. Für jeden Untervektorraum U von V, der $\{v_1, \ldots, v_n\} \subseteq M$ enthält, gehört auch v zu U, d. h. $v \in \bigcap_{U\in S} U$.

Da andererseits $\mathrm{Lin}(M)$ ein Untervektorraum von V ist, der M enthält, taucht $\mathrm{Lin}(M)$ in der Indexmenge auf und somit ist der Schnitt $\bigcap_{U\in S} U$ eine Teilmenge von $\mathrm{Lin}(M)$.

(iii) Dies ist eine direkte Konsequenz der Definition.

(iv) „$\Longleftarrow$“ folgt aus (ii). „$\Longrightarrow$“: Aus (i) wissen wir, dass stets $M \subseteq \mathrm{Lin}(M)$. Weil aber M ein Untervektorraum von V ist, gehört M zur Indexmenge S, sodass $\mathrm{Lin}(M) \subseteq M$.

(v) Das folgt aus (ii) und (iv).

(vi) „$\subseteq$“: Da $U_1 + U_2$ ein Untervektorraum von V ist, der U_1 und U_2 enthält, gilt $\mathrm{Lin}(U_1 \cup U_2) \subseteq U_1 + U_2$.
„$\supseteq$“: Jedes Element in $U_1 + U_2$ ist eine spezielle Linearkombination von $U_1 \cup U_2$, weswegen $U_1 + U_2 \subseteq \mathrm{Lin}(U_1 \cup U_2)$. □

Satz 8 (Kriterium I für Basen): *Seien K ein Körper, V ein K-Vektorraum. Eine Teilmenge B von V ist eine Basis genau dann, wenn B linear unabhängig ist mit* $\mathrm{Lin}(B) = V$.

Beweis: „$\Longrightarrow$“: Das ist genau die Definition.
„$\Longleftarrow$“: Sei $v \in V$ gegeben. Wegen $\mathrm{Lin}(B) = V$ ist v eine Linearkombination von B. Bleibt zu zeigen, dass diese Linearkombination eindeutig ist. Angenommen, es gäbe $\lambda^{(1)}, \lambda^{(2)} \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{w\in B} \lambda^{(1)}(w)w = \sum_{w\in B} \lambda^{(2)}(w)w$. Dann wäre

$$\mathbf{0}_V = \sum_{w\in B} (\lambda^{(1)}(w) - \lambda^2(w))w = \sum_{w\in B} (\lambda^{(1)} - \lambda^{(2)})(w)w,$$

d. h. wegen der linearen Unabhängigkeit von B hätten wir für alle $w \in B$, dass $\lambda^{(1)}(w) = \lambda^{(2)}(w)$ und damit $\lambda^{(1)} = \lambda^{(2)}$ als Abbildungen. □

Beispiel IV.2.10: Seien $V = \mathbb{R}^2$ und

$$B = \left\{ b_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, b_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Ist B eine Basis des $\mathbb{R}^2$? Nach Satz 8 haben wir zwei Dinge zu prüfen: Die lineare Unabhängigkeit von B und ob B ganz $\mathbb{R}^2$ erzeugt.

Zur linearen Unabhängigkeit: Seien reelle Zahlen λ_1, λ_2 gegeben, sodass $\lambda_1 b_1 + \lambda_2 b_2 = \mathbf{0}$. Wir erhalten für die erste bzw. die zweite Koordinate die Gleichung $\lambda_1 + \lambda_2 = 0$ bzw. $\lambda_1 - \lambda_2 = 0$, welche äquivalent sind zum linearen Gleichungssystem

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Setzen wir $A := \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, dann wissen wir: B ist linear unabhängig genau dann, wenn das homogene Gleichungssystem $Ax = \mathbf{0}$ nur die Lösung $\{\mathbf{0}\}$ hat. Nach Korollar II.5.22 ist das genau dann der Fall, wenn der Rang von A gleich der Anzahl der Spalten von A (nämlich 2) ist.

Zur Erzeugenden-Eigenschaft: Die lineare Hülle von B ist $\mathbb{R}^2$ genau dann, wenn es für jedes $v \in \mathbb{R}^2$ reelle Zahlen λ_1 und λ_2 gibt, sodass $v = \lambda_1 b_1 + \lambda_2 b_2$. Das ist genau dann der Fall, wenn es für jedes $v \in \mathbb{R}^2$ reelle Zahlen λ_1, λ_2 gibt, sodass

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = v.$$

Wiederum nach Korollar II.5.22 ist das äquivalent dazu, dass der Rang von A gleich der Anzahl der Zeilen von A (nämlich 2) ist.

Es bleibt also, den Rang von A zu bestimmen:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \xrightarrow{\text{II}-\text{I}} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \xrightarrow{\frac{1}{2}\text{II}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \xrightarrow{\text{I}-\text{II}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Da der Rang von A tatsächlich 2 ist, ist B eine Basis von $\mathbb{R}^2$.

Proposition IV.2.11 (Kriterium für Basis im K^n): *Seien K ein Körper und n, m natürliche Zahlen. Ferner seien $v_1, \ldots, v_n \in K^n$ und $A = (v_1 | \ldots | v_m)$.*

(i) *Die Menge $\{v_1, \ldots, v_m\}$ ist linear unabhängig genau dann, wenn der Rang von A die Anzahl m der Spalten von A ist.*

(ii) *Die Menge $\{v_1, \ldots, v_m\}$ ist ein Erzeugendensystem von K^n genau dann, wenn der Rang von A die Anzahl n der Zeilen von A ist.*

Korollar IV.2.12 (Dimension des K^n): *Seien K ein Körper und n eine natürliche Zahl. Jedes Basis des K^n hat genau n Elemente.*

Beweis: Sei B eine Basis des K^n. Nach Proposition IV.2.11(i) ist die Mächtigkeit von B höchstens n. Insbesondere ist B endlich. Es gibt also eine natürliche Zahl m und Vektoren $v_1, \ldots, v_m$ aus V, sodass wir B schreiben können als $B = \{v_1, \ldots, v_m\}$. Setzen wir $A = (v_1 | \ldots | v_m)$, dann wissen wir nach Proposition IV.2.11 dass $m = \text{Rang}(A) = n$ gelten muss, und wir sind fertig. □

Satz 9 (Koordinatenabbildung): *Seien K ein Körper und V ein K-Vektorraum.*

(i) *Ist B eine Basis von V, dann ist $V \cong \text{Abb}_0(B, K)$.*

(ii) *Ist $B = \{v_1, \ldots, v_n\}$ endlich, dann ist $V \cong K^n$.*

Beweis: (i) Die Abbildung

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

ist linear wegen der Rechenregeln für endliche Summen und da die Vektorraumstruktur auf $\operatorname{Abb}_0(B,K)$ durch die punktweisen Verknüpfungen (punktweise Addition von Funktionen und punktweise Skalarmultiplikation für Funktionen) gegeben ist. Nach Voraussetzung gilt $\operatorname{Lin}(B) = V$, d. h. Λ ist surjektiv. Da B linear unabhängig ist und damit $\operatorname{Kern}(\Lambda) = \{\mathbf{0}\}$ gilt, ist Λ außerdem injektiv. Insgesamt ist Λ also ein Isomorphismus.

(ii) Nach (i) ist $V \cong \operatorname{Abb}_0(B,K) \cong \operatorname{Abb}_0(\{e_1,\dots,e_n\},K) \cong K^n$, wobei $\{e_1,\dots,e_n\}$ die Standardbasis des K^n ist. □

Definition IV.2.13: Seien K ein Körper, n eine natürliche Zahl, V ein Vektorraum über K und $B = \{v_1,\dots,v_n\}$ eine Basis von V. Dann hat jede Basis von V n Elemente und wir nennen $\dim(V) := n$ die *Dimension von* V. In diesem Fall heißt V *endlichdimensional*.

Beweis: (i) Durch Satz 9 können wir uns auf das Resultat aus Korollar IV.2.12 zurückziehen. □

Satz 10 (Kriterium II für Basen): *Seien K ein Körper, V ein K-Vektorraum und $B \subseteq V$ eine Teilmenge. Dann sind äquivalent:*

(i) *Die Menge B ist eine Basis.*

(ii) *Die Menge B ist eine bezüglich Inklusion maximale linear unabhängige Teilmenge von V, d. h. B ist linear unabhängig und ist M eine weitere Teilmenge von V mit $B \subsetneq M$, dann ist M nicht linear unabhängig.*

(iii) *Die Menge B ist ein bezüglich Inklusion minimales Erzeugendensysem von V, d. h.* $\operatorname{Lin}(B) = V$ *und ist M' eine echte Teilmenge von B, dann ist* $\operatorname{Lin}(M') \subsetneq V$.

Beweis: „(i) $\Longrightarrow$ (ii)“: Da B nach Voraussetzung eine Basis ist, ist B insbesondere linear unabhängig. Ist M eine Teilmenge von V mit $B \subsetneq M$, dann gibt es $v \in M - B$. Da B eine Basis ist, gibt es $\lambda \in \operatorname{Abb}_0(B,K)$, sodass $v = \sum_{w\in B} \lambda(w)w$. Nun erklären wir eine Abbildung $\lambda'\colon M \to K$ mit endlichem Träger durch

$$\lambda'(w) = \begin{cases} \lambda(w), & \text{falls } w \in B, \\ -1, & \text{falls } w = v, \\ 0, & \text{sonst,} \end{cases}$$

und finden $\sum_{w\in M}\lambda'(w)w = \sum_{w\in B}\lambda(w)w - 1v = \mathbf{0}$, d. h. M ist linear abhängig.

„(ii) $\Longrightarrow$ (iii)“: Sei B eine bezüglich Inklusion maximale linear unabhängige Teilmenge von V. Wir wollen zeigen, dass dann schon $\mathrm{Lin}(B) = V$ und dass B ein minimales Erzeugendensystem von V ist.

Um zu zeigen, dass $\mathrm{Lin}(B) = V$, sei $v \in V$ gegeben. Gehört v zu B, dann auch zu $\mathrm{Lin}(B)$. Gehört v nicht zu B, setze $M := B \cup \{v\}$. Da B maximal linear unabhängig ist und $B \subsetneq M$ gilt, muss M linear abhängig sein. Es gibt also $\mathbf{0} \neq \lambda \in \mathrm{Abb}_0(M, K)$, sodass $\sum_{w\in M}\lambda(w)w = \mathbf{0}$. Für dieses λ muss gelten, dass $\alpha := \lambda(v) \neq 0$, denn sonst wäre B bereits linear abhängig. Wir dürfen also durch α teilen und erhalten

$$\begin{aligned}\sum_{w\in M}\lambda(w)w = \mathbf{0} &\Longrightarrow \frac{1}{\alpha}\Big(\sum_{w\in M}\lambda(w)w\Big) = \mathbf{0}\\ &\Longrightarrow v = -\sum_{w\in B}\frac{\lambda(w)}{\alpha}w\\ &\Longrightarrow v \in \mathrm{Lin}(B).\end{aligned}$$

Gäbe es eine echte Teilemenge M' von B mit $\mathrm{Lin}(M') = B$, dann erhielten wir in „(i) $\Longrightarrow$ (ii)“, dass B linear abhängig sein müsste. Ein Widerspruch!

„(iii) $\Longrightarrow$ (i)“: Sei B ein minimales Erzeugendensystem von V. Wir wollen zeigen, dass B dann auch linear unabhängig sein muss.

Angenommen B wäre linear abhängig. Dann gäbe es $\mathbf{0} \neq \lambda \in \mathrm{Abb}_0(B, K)$, sodass $\sum_{w\in B}\lambda(w)w = \mathbf{0}$. Da $\lambda \neq \mathbf{0}$, gäbe es $v_0 \in B$ mit $\lambda(v_0) \neq 0$. Für dieses v_0 fänden wir dann wie in „(ii) $\Longrightarrow$ (iii)“, dass

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

sodass $\mathrm{Lin}(B - \{v_0\}) = V$ folgte. Ein Widerspruch! □

Korollar IV.2.14: *Seien K ein Körper, V ein K-Vektorraum und M eine nichtleere Teilmenge von V.*

(i) *Ist M linear abhängig, dann gibt es $v \in M$ mit $v \in \mathrm{Lin}(M - \{v\})$. Insbesondere gilt $\mathrm{Lin}(M) = \mathrm{Lin}(M - \{v\})$.*

(ii) *Ist M linear unabhängig und ist $v \in V - \mathrm{Lin}(M)$, dann ist auch $M \cup \{v\}$ linear unabhängig.*

Beweis: Aussage (i) folgt aus dem Beweis von „(iii) $\Longrightarrow$ (i)“ im Beweis von Satz 10, Aussage (ii) folgt aus dem Beweis von „(ii) $\Longrightarrow$ (iii)“ im Beweis von Satz 10. □

Satz 11 (Basisergänzungssatz): *Seien K ein Körper und V ein K-Vektorraum. Hat V ein endliches Erzeugendensystem. Dann gilt:*

(i) *Der Vektorraum V hat eine Basis.*

(ii) *Jedes Erzeugendensystem von V enthält eine Basis von V.*

(iii) *Jede linear unabhängige Teilmenge von V lässt sich durch Hinzunahme endlich vieler Elemente zu einer Basis ergänzen.*

Beweis: (i) Können wir (ii) zeigen, dann gibt es (i) gratis.

(ii) Wegen Proposition IV.2.14(i) und der Endlichkeit des Erzeugendensystems haben wir nach endlich vielen Rauswürfen ein bezüglich Inklusion minimales Erzeugendensystem von V. Nach Satz 10 ist das eine Basis von V.

(iii) Nach (ii) hat V eine endliche Basis. Nach Proposition IV.2.11 und Satz 10 ist jede linear unabhängige Teilmenge von V insbesondere endlich und mithilfe von Proposition IV.2.14(ii) erhalten wir durch Hinzunahme von endlich vielen Vektoren von V eine Basis von V. □

3. Lineare Fortsetzung und Abbildungsmatrix

Wegen der Rechenregeln für endliche Summen sind lineare Abbildungen dadurch eindeutig festgelegt, was sie auf einer Basis des Startvektorraumes tun. Wir werden sehen, dass uns das erlaubt, die Wirkung linearer Abbildungen zwischen endlichdimensionalen Vektorräumen durch Matrizen zu beschreiben.

Satz 12 (Fortsetzungssatz): *Seien K ein Körper, V und W zwei K-Vektorräume und B eine Basis von V.*

(i) *Jede lineare Abbildung $\phi\colon V \to W$ ist eindeutig durch ihre Einschränkung $f := \phi|_B\colon B \to W$ bestimmt.*

(ii) *Jede Abbildung $f\colon B \to W$ lässt sich auf genau eine Weise zu einer linearen Abbildung $\phi\colon V \to W$ fortsetzen, d. h. es gibt genau einen Vektorraumhomomorphismus $\phi\colon V \to W$, sodass $\phi|_B = f$. Dieser heißt* lineare Fortsetzung von f.

Beweis: (i) Sei v ein Element von V. Da B eine Basis von V ist, gibt es $\lambda_v \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{b\in B} \lambda_v(b)b$ und wir erhalten

$$\phi(v) = \phi\Big(\sum_{b\in B} \lambda_v(b)b\Big) = \sum_{b\in B} \lambda_v(b)\phi(b) = \sum_{b\in B} \lambda_v(b)f(b).$$

(ii) Für jedes Element v von V gibt es $\lambda_v \in \mathrm{Abb}_0(B,K)$ mit $v = \sum_{b\in B} \lambda_v(b)b$. Deshalb können wir $\phi\colon V \to W$ definieren durch

$$\Phi\colon V \longrightarrow W, \qquad v \longmapsto \sum_{b\in B} \lambda_v(b)b.$$

Wegen der Rechenregeln für endliche Summen liefert das eine lineare Abbildung. Die Eindeutigkeit dieser Festsetzung folgt aus (i). □

Korollar IV.3.1: *Seien K ein Körper, V ein K-Vektorraum und B eine Basis von V. Dann ist die Abbildung*

$$H\colon \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Abb}_0(B,W), \qquad \phi \longmapsto \phi|_B$$

ein Isomorphismus von K-Vektorräumen.

Beweis: Nach Satz 12 ist H bijektiv und wegen der punktweisen Verknüpfungen und der Rechenregeln für endliche Summen ist H linear. □

Wir haben in Satz 9 gesehen: Ist V ein Vektorraum über de Körper K mit Basis $B = \{b_1, \ldots, b_n\}$, dann erhalten wir einen Isomorphismus

$$\Lambda\colon K^n \longrightarrow V, \qquad (x_1, \ldots, x_n)^t \longmapsto \sum_{i=1}^{n} x_i b_i.$$

Für das, was folgt, wird die Reihenfolge der Elemente der Basis eine Rolle spielen, weswegen wir *geordnete Basen* $B = (b_1, \ldots, b_n)$ betrachten wollen.

Definition IV.3.2 (Koordinatenabbildung): Seien K ein Körper, V ein Vektorraum über K und $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V. Die Umkehrabbildung $D_B := \Lambda^{-1}$ zu Λ aus Satz 9, d. h.

$$D_B\colon V \longrightarrow K^n, \qquad v = \sum_{i=1}^{n} v_i b_i \longmapsto (v_1, \ldots, v_n)^t,$$

heißt *Koordinatenabbildung zu B*.

Sind n und m natürliche Zahlen und $A \in K^{n\times m}$ eine Matrix, dann schreiben wir $L_A\colon K^m \to K^n$, $x \mapsto Ax$.

Satz 13 (Darstellungsmatrix): *Seien K ein Körper, V und W Vektorräume über K mit geordneten Basen $B = (b_1, \dots, b_m)$ respektive $C = (c_1, \dots, c_n)$ und $\phi\colon V \to W$ eine lineare Abbildung. Dann gibt es genau eine Matrix $A \in K^{n\times m}$, sodass*

$$D_C \circ \phi = L_A \circ D_B.$$

Die Einträge der Matrix $A = (a_{ij})$ sind bestimmt durch $a_{ij} = \lambda_{ij}$, wobei $\phi(b_j) = \sum_{i=1}^{n} \lambda_{ij} c_i$ für $1 \le j \le m$. Wir schreiben $D_{C,B}(\phi) := A$ und nennen diese Matrix die Darstellungsmatrix von ϕ bezüglich B und C.

Beweis: Wir kennen die Situation für lineare Abbildungen $\psi\colon K^m \to K^m$ aus Bemerkung IV.1.5: Die Darstellungsmatrix von ψ bezüglich der Standardbasen ist die Matrix $A = (\psi(e_1)|\dots|\psi(e_m))$, denn Ae_i liefert die i-te Spalte von A, d. h. für $v = \sum_{i=1}^{m} v_i e_i$ haben wir deshalb

$$\psi(v) = \psi\Big(\sum_{i=1}^{m} v_i e_i\Big) = \sum_{i=1}^{m} v_i \psi(e_i) = \sum_{i=1}^{m} v_i A e_i = A\Big(\sum_{i=1}^{m} v_i e_i\Big) = A(v)$$

und somit $L_A = \psi$. Genau das wollen wir auch für $\phi\colon V \to W$ erreichen. Da wir aber auf der Ebene von V und W keine Matrizen zur Verfügung haben, müssen wir in die Koordinatenvektorräume K^m bzw. K^n übersetzen. Das machen wir bei fixierten geordneten Basen mittels Koordinatenabbildungen, um schließlich die Matrix $A \in K^{m\times n}$ zu finden, für die $L_A = D_C \circ \phi \circ D_B^{-1}$.

Dadurch, dass $D_B(b_i) = e_i$ für $1 \le i \le m$, dass $L_A(D_B(b_i))$ die i-te Spalte von A liefert und wir $D_C \circ \phi = L_A \circ D_B$ erreichen wollen, wissen wir, dass die i-te Spalte von A aus den Koordinaten von $\phi(b_i)$ bezüglich C bestehen muss. Aber genau das sind die Gleichungen, die wir für die Einträge der Matrix A bereits angegeben haben. □

Definition IV.3.3 (Basiswechselmatrix): Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum und $B = (b_1, \dots, b_n)$, $B' = (b'_1, \dots, b'_n)$ geordnete Basen von V. Dann heißt $D_{B',B} := D_{B',B}(\mathrm{id})$ *Basiswechselmatrix von B nach B'*.

Auch die Bezeichnung *Koordinatentransformationsmatrix von B nach B'* ist gebräuchlich. Das liegt daran, dass für alle $v \in V$ gilt, dass

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

Da die Einträge $\lambda_{i,j}$ der Basiswechselmatrix $D_{B',B}$ bestimmt sind durch die Gleichungen $b_j = \sum_{i=1}^{n} \lambda_{i,j} b'_i$ für $1 \le j \le n$, geben die Einträge dieser Matrix aber gleichzeitig an, wie die Basis B aus der Basis B' hervorgeht. Deshalb wird diese Matrix gelegentlich auch *Basiswechselmatrix von B' nach B genannt.*

Proposition IV.3.4 (Basiswechsel und Darstellungsmatrizen): *Sei K ein Körper.*

(i) *Seien $\phi\colon V_1 \to V_2$ und $\psi\colon V_2 \to V_3$ lineare Abbildungen zwischen endlichdimensionalen K-Vektorräumen und seien B_1, B_2 und B_3 geordnete Basen der jeweiligen Vektorräume. Dann gilt*

$$D_{B_3,B_1}(\psi \circ \phi) = D_{B_3,B_2}(\psi) D_{B_2,B_1}(\phi).$$

(ii) *Sei V ein endlichdimensionaler Vektorraum über K mit geordneten Basen B und B'. Dann ist die Basiswechselmatrix $D_{B',B}$ regulär und für die Inverse gilt $D_{B',B}^{-1} = D_{B,B'}$.*

(iii) *Für $V = K^n$ und die Standardbasis $E = (e_1, \ldots, e_n)$ von K^n und eine weitere geordnete Basis $B = (b_1, \ldots, b_n)$ gilt $D_{E,B} = (b_1 | \ldots | b_n)$.*

(iv) *Für geordnete Basen B und B' von K^n gilt*

$$D_{B',B} = D_{B',E} D_{E,B} = D_{E,B'}^{-1} D_{E,B}.$$

Beweis: (i) Die Situation können wir im Diagramm

$$\begin{array}{ccccc} V_1 & \xrightarrow{\phi} & V_2 & \xrightarrow{\psi} & V_3 \\ {\scriptstyle D_{B_1}}\downarrow & & {\scriptstyle D_{B_2}}\downarrow & & \downarrow{\scriptstyle D_{B_3}} \\ K^{n_1} & \xrightarrow[L_A]{} & K^{n_2} & \xrightarrow[L_B]{} & K^{n_3} \end{array}$$

einfangen, wobei $A = D_{B_2,B_1}(\phi)$, $B = D_{B_3,B_2}(\phi)$, $n_1 = \dim V_1$, $n_2 = \dim V_2$ und $n_3 = \dim V_3$. Für die Matrix $C := D_{B_3,B_1}(\psi \circ \phi)$ gilt jetzt

$$L_C = D_{B_3} \circ \psi \circ \phi \circ D_{B_1}^{-1} = L_B \circ L_A,$$

also folgt die Behauptung.

(ii) Das folgt aus (i) für $\psi = \phi = \mathrm{id}$ und $B_1 = B$, $B_2 = B'$, $B_3 = B$.

(iii) Mit $B' = E = (e_1, \ldots, e_n)$ erhalten wir die bestimmenden Gleichungen $b_j = \sum_{i=1}^{n} \lambda_{i,j} e_i$ für die Basiswechselmatrix, d. h. die $\lambda_{i,j}$ sind genau die Koordinaten von b_j bezüglich der Standardbasis.

(iv) Folgt aus (i) und (ii). □

Proposition IV.3.5: *Seien K ein Körper, V und W Vektorräume über K mit geordneten Basen $B = (b_1, \ldots, b_m)$ und $C = (c_1, \ldots, c_n)$ und sei $\Phi\colon V \to W$ eine lineare Abbildung. Ferner seien B' und C' weitere geordnete Basen. Dann gilt*

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

Beweis: Da im folgenden Diagramm jedes der Quadrate kommutiert, können wir die Situation mit dem folgenden kommutativen Diagramm abbilden:

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

Wegen $\mathrm{id} \circ \phi \circ \mathrm{id} = \phi$ beschreiben die äußeren Pfeile das Diagramm der Darstellungsmatrix $D_{C',B'}(\phi)$. Die Kompposition der Pfeile in der unteren Zeile liefert die Abbildung $x \mapsto D_{C',C} D_{C,B}(\phi) D_{B,B'} x$, was wir behauptet haben. □

4. Summen von Unterräumen und Faktorräume

Seien K ein Körper und V ein Vektorraum über K. Sind $U_1, \ldots, U_n$ Untervektorräume von V, dann haben wir bereits gesehen, dass

$$\sum_{i=1}^{n} U_i := U_1 + \cdots + U_n := \{x_1 + \cdots + x_n \mid x_1 \in U_1, \ldots, x_n \in U_n\} \subseteq V$$

auch ein Untervektorraum von V ist. Wir nennen $U_1 + \cdots + U_n$ die *Summe von* $U_1, \ldots, U_n$.

Definition IV.4.1 (Direkte Summe): Seien K ein Körper, V ein Vektorraum über K und $U_1, \ldots, U_n$ Untervektorräume von V. Falls gilt: „Wenn immer $u_1 \in U_1, \ldots, u_n \in U_n$ mit $u_1 + \cdots + u_n = \mathbf{0}$, dann sind $u_1 = \cdots = u_n = \mathbf{0}$“, dann heißt die Summe $U_1 + \cdots + U_n$ *direkt*. In diesem Fall schreiben wir $\bigoplus_{i=1}^n U_i := \sum_{i=1}^n U_i$.

Bemerkung IV.4.2: Seien K ein Körper, V ein K-Vektorraum und $U_1, \ldots U_n$ Untervektorräume von V. Ist $\sum_{i=1}^n U_i$ direkt, dann haben wir für $i, j \in \{1, \ldots, n\}$ mit $i \neq j$, dass $U_i \cap U_j = \{\mathbf{0}\}$. Das sieht man so: Für $v \in U_1 \cap U_2$ haben wir $v - v + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$, d. h. $v = \mathbf{0}$ per Definition der direkten Summe. Die Umkehrung dieser Aussage gilt nicht! Sind beispielsweise $V = \mathbb{R}^2$, $U_1 = \langle (1,0)^t \rangle$, $U_2 = \langle (0,1)^t \rangle$ und $U_3 = \langle (1,1)^t \rangle$, dann haben wir zwar $U_i \cap U_j = \{\mathbf{0}\}$ für $1 \leq i, j \leq 3$, $i \neq j$, aber

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

Satz 14 (über die direkte Summe): *Seien K ein Körper, V ein Vektorraum über K und $U_1, \ldots, U_n$ Untervektorräume von V.*

(i) *Seien $B_1, \dots, B_n$ Basen von $U_1, \dots, U_n$. Ist die Summe der U_i direkt, dann ist $B := B_1 \cup \dots \cup B_n$ eine Basis von $\bigoplus_{i=1}^n U_i$.*

(ii) *Seien $U_1, \dots, U_n$ endlichdimensional. Genau dann ist die Summe der U_i direkt, wenn $\dim(\sum_{i=1}^n U_i) = \sum_{i=1}^n \dim U_i$.*

Beweis: (i) Per Definition von $\bigoplus_{i=1}^n U_i$ ist B ein Erzeugendensystem. Bleibt also die lineare Unabhängigkeit zu zeigen. Sei dazu $\lambda \in \mathrm{Abb}_0(B, K)$ mit $\mathbf{0} = \sum_{b\in B} \lambda(b)b$. Setze $u_i := \sum_{b\in B_i} \lambda(b)b$. Dann ist $u_1 + \dots + u_n = \mathbf{0}$ und da die Summe direkt ist, erzwingt das $u_1 = \dots = u_n = \mathbf{0}$. Damit muss λ die Nullabbildung sein und B ist linear unabhängig.

(ii) „$\Longrightarrow$“ folgt aus (i). Zu „$\Longleftarrow$“: Da die U_i endlichdimensional sind, hat jeder der Vektorräume eine Basis, sagen wir $B_i \subseteq U_i$. Setze $B := \bigcup_{i=1}^n B_i$. Dann haben wir

$$\#B \leq \sum_{i=1}^n \#B_i = \sum_{i=1}^n \dim U_i = \dim\Big(\sum_{i=1}^n U_i\Big)$$

Da B ein Erzeugendensystem von $\sum_{i=1}^n U_i$ ist, gilt $\dim(\sum_{i=1}^n U_i) \leq \#B$. Nun liefert Satz 10, dass B eine Basis von $\sum_{i=1}^n U_i$ ist und die lineare Unabhängigkeit von B liefert die Direktheit der Summe. □

Definition IV.4.3 (Äquivalenz modulo Unterraum): Seien K ein Körper, V ein Vektorraum über K und $U \subseteq V$ ein Untervektorraum. Auf V wird durch

$$v_1 \sim v_2 :\Longleftrightarrow v_1 - v_2 \in U$$

eine Äquivalenzrelation, genannt *Äquivalenz modulo U*, erklärt. Für $v \in V$ bezeichnet

$$[v] := \{w \in V \mid v \sim w\} = \{w \in V \mid v - w \in U\} =: v + U$$

die Äquivalenzklasse von v und $V/U := V/\!\sim = \{[v] \mid v \in V\}$ bezeichnet die Menge der Äquivalenzklassen bezüglich Äquivalenz modulo U.

Proposition IV.4.4 (Quotient nach Unterraum): *Seien K ein Körper, V ein Vektorraum über K, $U \subseteq V$ ein Unterraum und V/U die Menge der Äquivalenzklassen bezüglich Äquivalenz modulo U. Auf V/U wird duch*

$$[v] + [w] := [v + w], \qquad \lambda[v] := [\lambda v]$$

eine K-Vektorraumstruktur erklärt. Zusammen mit dieser heißt V/U der Quotient von V nach U *oder* Faktorraum V/U.

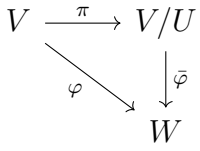
Beweis: Sobald wir uns davon überzeugt haben, dass die oben angegebenen Verknüpfungen wohldefiniert sind, sehen wir sofort dass V/U ein K-Vektorraum ist, da wir repräsentantenweise rechnen und wir wissen, dass V ein Vektorraum über K ist. Für die Wohldefiniertheit ist die Unabhängigkeit von der Wahl der Repräsentanten zu prüfen. □

Definition IV.4.5: Seien K ein Körper, V ein Vektorraum über K und $U \subseteq V$ ein Untervektorraum. Dann heißt

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

die *kanonische Projektion*. Die kanonische Projektion ist eine surjektive lineare Abbildung mit $\operatorname{Kern}(\pi) = U$.

Satz 15: *Es seien K ein Körper, V und W Vektorräume über K, $\varphi\colon V \to W$ eine lineare Abbildung und $U \subseteq V$ ein Untervektorraum mit $U \subseteq \operatorname{Kern}(\varphi)$. Dann gibt es genau eine lineare Abbildung $\bar{\varphi}\colon V/U \to W$, die das Diagramm*



kommutativ macht, d. h. $\bar{\varphi} \circ \pi = \varphi$. Sind sogar $U = \operatorname{Kern}(\varphi)$ und $W = \operatorname{Bild}(\varphi)$, dann ist $\bar{\varphi}$ injektiv (und per Konstruktion surjektiv), d. h. $V/\operatorname{Kern}(\varphi) \cong \operatorname{Bild}(\varphi)$ vermöge $\bar{\varphi}$.

Beweis: Wegen $\bar{\varphi} \circ \pi = \varphi$ haben wir keine andere Wahl, als zu definieren: $\bar{\varphi}([v]) := \varphi(v)$. Jetzt haben wir zu überprüfen, dass $\bar{\varphi}$ wohldefiniert ist, d. h. dass alle $w \in [v]$ unter φ dasselbe Bild haben.

Ist $U = \operatorname{Kern}(\varphi)$, dann ist $\bar{\varphi}$ injektiv. Wegen $\operatorname{Bild}(\bar{\varphi}) = \operatorname{Bild}(\varphi)$ ist $\bar{\varphi}$ auch surjektiv, d. h. $\bar{\varphi}\colon V/U \to \operatorname{Bild}(\varphi)$ ist ein Isomorphismus. □

Satz 16 (Basis des Faktorraums): *Seien K ein Körper, V ein Vektorraum über K und $U \subseteq V$ ein Untervektorraum. Sind $B' \subseteq U$ eine Basis und $B \subseteq V$ eine Basis von V, die B' enthält, dann ist*

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

eine Basis von V/U.

Beweis: Zunächst zeigen wir, dass $\mathrm{Lin}(C) = V/U$. Sei dazu $v \in V$ gegeben. Da B eine Basis von V ist, gibt es $\lambda \in \mathrm{Abb}_0(B, K)$ mit $v = \sum_{b\in B} \lambda(b)b$, d. h.

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

Nun zur linearen Unabhängigkeit: Sei $\lambda \in \mathrm{Abb}_0(B - B', K)$ gegeben, sodass $\sum_{b\in B-B'} \lambda([b])[b] = [0]$. Setze $u := \sum_{b\in B-B'} \lambda(b)b$. Dann ist $[u] = [0]$, d. h. u gehört zu U. Weil B' eine Basis von U ist, gibt es also $\lambda_u \in \mathrm{Abb}(B', K)$, sodass $u = \sum_{b\in B'} \lambda_u(b)b$ und damit ist

$$\sum_{b\in B-B'} \lambda(b)b - \sum_{b\in B'} \lambda_u(b)b = \mathbf{0}.$$

Da B eine Basis ist, muss nun λ die Nullabbildung sein. □

Satz 17 (Dimensionsformel): *Seien K ein Körper und V ein endlichdimensionaler Vektorraum über K mit* $\dim V = n$.

(i) *Ist $U \subseteq V$ ein Untervektorraum, dann ist* $\dim V/U = \dim V - \dim U$.

(ii) *Sind U_1 und $U_2 \subseteq V$ Untervektorräume, dann ist*

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

(iii) *Ist W ein weiterer K-Vektorraum und $\phi\colon V \to W$ linear, dann ist*

$$\dim V = \dim \mathrm{Kern}\,\phi + \dim \mathrm{Bild}\,\phi.$$

Beweis: (i) In Satz 16 haben wir gezeigt, wie man eine Basis von V/U erhalten kann. Insbesondere haben wir die Dimension von V/U bestimmt.

(ii) Aus Satz 15 kennen wir die Isomorphie $\mathrm{Bild}\,\phi \cong V/\mathrm{Kern}\,\phi$. In den Übungen werden Sie zeigen, dass das bedeutet, dass beide Vektorräume die selbe Dimension haben müssen. Wir können mit (i) deshalb folgern:

$$\dim \mathrm{Bild}\,\phi = \dim(V/\mathrm{Kern}\,\phi) = \dim V - \dim \mathrm{Kern}\,\phi.$$

(iii) Wir möchten den Homomorphiesatz anwenden, um die Behauptung zu zeigen. Dazu suchen wir uns eine geeignete surjektive lineare Abbildung mit dem richtigen Kern, nämlich

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

Ist $u_1 + u_2 \in U_1 + U_2$ vorgegeben, dann ist $\alpha(u_1, -u_2) = u_1 + u_2$, d. h. α ist surjektiv. Auch den Kern von α können wir leicht erkennen, der ist nämlich

$$\operatorname{Kern} \alpha = \{(u_1, u_2) \in U_1 \times U_2 \mid u_1 = u_2\}.$$

Wir haben somit einen Isomorphismus $U_1 \cap U_2 \to \operatorname{Kern} \alpha$, $u \mapsto (u, u)$. Schließlich können wir Dimensionen von $U_1 \times U_2$ und $U_1 + U_2$ miteinander in Verbindung bringen: Ist B_1 eine Basis von U_1 und ist B_2 eine Basis von U_2, dann erhalten wir durch $B := \{(b, 0) \mid b \in B_1\} \cup \{(0, b) \mid b \in B_2\}$ eine Basis von $U_1 \times U_2$, d. h. $\dim(U_1 \times U_2) = \dim(U_1 + U_2)$. Mit (iii) erhalten wir jetzt

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

Definition IV.4.6 (Rang und Kern): Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, W ein weiterer K-Vektorraum und $\phi\colon V \to W$ linear. Dann heißt $\operatorname{Rang} \phi := \dim \operatorname{Bild} \phi$ der *Rang von* ϕ.

Seien n und m natürliche Zahlen, K ein Körper und $A \in K^{n \times m}$ gegeben. Dann heißt $\operatorname{Kern} A := \{x \in K^m \mid Ax = \mathbf{0}\}$ der *Kern der Matrix* A.

Bemerkung IV.4.7: Sind K ein Körper, V ein endlichdimensionaler Vektorraum über K, W ein weiterer K-Vektorraum und $\phi\colon V \to W$ linear, dann liefert Satz 17 dass $\dim V = \dim \operatorname{Kern} \phi + \operatorname{Rang} \phi$.

Bemerkung IV.4.8: Seien K ein Körper, V und W endlichdimensionale K-Vektorräume mit geordneten Basen $B = (b_1, \ldots, b_m)$ von V und $C = (c_1, \ldots, c_n)$ von W und sei $\phi\colon V \to W$ linear. Bezeichnet $A := D_{C,B}(\phi)$, dann haben wir:

(i) Bezeichnet $\{e_1, \ldots, e_m\} \subseteq K^m$ die Standardbasis, dann ist das Bild von ϕ isomorph zu $\operatorname{Lin}(Ae_1, \ldots, Ae_m)$, d. h. der linearen Hülle der Spalten der Darstellungsmatrix von A.

(ii) Der Kern von ϕ ist isomorph zum Kern der Matrix A, d. h. beide Definitionen von „Kern“ passen zueinander.

Satz 18 (Rang): *Seien K ein Körper, V und W endlichdimensionale Vektorräume über K mit geordneten Basen $B = (b_1, \ldots, b_m)$ von V und $C = (c_1, \ldots, c_n)$ von W, $\phi\colon V \to W$ linear und $A := D_{C,B}(\phi)$.*

(i) *Es gilt* $\operatorname{Rang} A = \operatorname{Rang} \phi$.

(ii) *Bezeichnen $s_1, \dots, s_m$ die Spalten und $z_1, \dots, z_n$ die Zeilen von A, dann ist*

$$\operatorname{Rang} A = \dim \operatorname{Lin}(s_1, \dots, s_m) = \dim \operatorname{Lin}(z_1, \dots, z_n).$$

Beweis: (i) Wie wir wissen, ist $\operatorname{Rang} A = m - \dim \operatorname{Kern} A$. Nach Proposition IV.4.8 ist $\dim \operatorname{Kern} A = \dim \operatorname{Kern} \phi$, sodass $\operatorname{Rang} A = m - \dim \operatorname{Kern} \phi$. Wegen Proposition IV.4.7 ist das aber genau $\operatorname{Rang} \phi$.

(ii) Das Bild der linearen Abbildung $\phi_A \colon K^m \to K^n$, $x \mapsto Ax$ ist das Erzeugnis der Spalten von A, sodass (i) liefert: $\operatorname{Rang} A = \operatorname{Rang} \phi_A = \dim\langle s_1, \dots, s_m\rangle$.

Sei nun T die Treppenform von A. Die Treppenform von A entsteht aus A durch Zeilenoperationen, genauer: Es gibt elementare Zeilenumformungen $Z_1, \dots, Z_N$ (d. h. $Z_k = A^{\alpha}_{i,j}$, oder $Z_k = V_{i,j}$ oder $Z_k = \operatorname{diag}(\alpha_1, \dots, \alpha_n)$, wobei $\alpha, \alpha_1, \dots, \alpha_n \in K^\times$), sodass $T = Z_1 \cdots Z_N \cdot A$.

Setze $A_k := Z_{k+1} \cdots Z_N \cdot A$. Für $A_{k+1} = Z_k \cdot A_k$ ist der Spann der Zeilenvektoren von A_k der Spann der Zeilenvektoren von A_{k+1}, d. h. das Erzeugnis der Zeilen von A ist gleich dem Erzeugnis der Zeilenvektoren von T. Aber dann ist auch $\dim \operatorname{Lin}(z_1, \dots, z_n) = \operatorname{Rang} T = \operatorname{Rang} A$. □

Kapitel V.

Endomorphismen von Vektorräumen

1. Endomorphismen und Basiswechsel

Bemerkung V.1.1 (Basiswechsel): Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K mit geordneter Basis $B = (b_1, \ldots, b_n)$, $\phi\colon V \to V$ linear und $B' = (b'_1, \ldots, b'_n)$ eine weitere geordnete Basis von V. Setzen wir $A := D_{B,B}(\phi)$, $A' := D_{B',B'}(\phi)$ und $S := D_{B',B}$, dann liefert Proposition IV.3.5, dass

$$A' = D_{B',B} D_{B,B}(\phi) D_{B,B'} = SAS^{-1}.$$

Definition V.1.2 (Ähnlichkeit): Seien K ein Körper und V ein n-dimensionaler K-Vektorraum.

(i) Sind $A_1, A_2 \in K^{n\times n}$ gegeben und gibt es $S \in \mathrm{Gl}_n(K)$ mit $A_2 = SA_1S^{-1}$, dann heißen A_1 und A_2 *ähnlich*.

(ii) Zwei Matrizen $A_1, A_2 \in K^{n\times n}$ sind ähnlich genau dann, wenn sie Darstellungsmatrizen derselben linearen Abbildung ϕ sind.

Proposition V.1.3 (Rang als Ähnlichkeitsinvariante): *Seien K ein Körper, V ein n-dimensionaler K-Vektorraum und $A \in K^{n\times n}$. Der Rang von A ist eine* Ähnlichkeitsinvariante*, d. h. ist $B \in K^{n\times n}$ ähnlich zu A, dann gilt* $\operatorname{Rang} A = \operatorname{Rang} B$.

2. Eigenwerte und Eigenvektoren

Definition V.2.1 (Eigenvektoren, Eigenwerte): Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, $\phi\colon V \to V$ linear und $A \in K^{n\times n}$.

(i) Sei λ ein Element von K. Gibt es $v \in V - \{\mathbf{0}\}$ mit $\phi(v) = \lambda v$, dann heißt λ ein *Eigenwert von ϕ zum Eigenvektor v*.
Gibt es $x \in K^n - \{\mathbf{0}\}$ mit $Ax = \lambda x$, dann heißt λ ein *Eigenwert von A zum Eigenvektor x*.

(ii) Für $\lambda \in K$ heißt

$$\operatorname{Eig}(\phi, \lambda) := \{v \in V \mid \phi(v) = \lambda v\}$$
$$\operatorname{Eig}(A, \lambda) := \{x \in K^n \mid Ax = \lambda x\}$$

Eigenraum zu ϕ respektive *Eigenraum zu A*.

(iii) Die Menge der Eigenwerte

$$\operatorname{Spec}\phi := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } \phi\}$$
$$\operatorname{Spec} A := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } A\}$$

heißt *Spektrum von ϕ* respektive *Spektrum von A*.

Bemerkung V.2.2: Seien K ein Körper, V ein n-dimensionaler Vektorraum über K mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear.

(i) Ist $A = D_{B,B}(\phi)$, dann gilt $\operatorname{Spec}\phi = \operatorname{Spec} A$. Ferner ist $v \in V - \{\mathbf{0}\}$ ein Eigenvektor von ϕ zum Eigenwert λ genau dann, wenn $x = D_B(v)$ ein Eigenvektor von A zum Eigenwert λ ist.

(ii) Ein $\lambda \in K$ gehört zu $\operatorname{Spec}\phi$ respektive $\operatorname{Spec} A$ genau dann, wenn $\operatorname{Eig}(\phi, \lambda) \neq \{\mathbf{0}\}$ respektive $\operatorname{Eig}(A, \lambda) \neq \{\mathbf{0}\}$.

(iii) Wir haben die Äquivalenzen

$$Av = \lambda v \iff (A - \lambda I_n)v = \mathbf{0} \iff v \in \operatorname{Kern}(A - \lambda I_n),$$

d. h. $\operatorname{Eig}(A, \lambda) = \operatorname{Kern}(A - \lambda I_n)$. Analog ist $\operatorname{Eig}(\phi, \lambda) = \operatorname{Kern}(\phi - \lambda \operatorname{id}_V)$. Insbesondere sind Eigenräume von A beziehungsweise Eigenräume von ϕ Untervektorräume von K^n beziehungsweise V.

Beispiel V.2.3: (i) Seien $\lambda_1, \ldots, \lambda_n \in K$ und $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Dann ist $\operatorname{Spec}(A) = \{\lambda_1, \ldots, \lambda_n\}$, außerdem sind die Eigenräume leicht anzugeben: $\operatorname{Eig}(A, \lambda_i) = \operatorname{Lin}(e_i)$.

(ii) Sei $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Diagonalen. Bezüglich der Basis $B = \{(1,1)^t, (1,-1)^t\}$ von $\mathbb{R}^2$ hat ϕ die Darstellungsmatrix

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

d. h. $\operatorname{Spec}\phi = \{\pm 1\}$

Definition V.2.4 (Diagonalisierbarkeit): Seien K ein Körper, V ein K-Vektorraum mit geordneter Basis $B = (b_1, \dots, b_n)$, $\phi\colon V \to V$ linear und $A \in K^{n\times n}$.

(i) Gibt es eine Basis B' von V und Elemente $\lambda_1, \dots, \lambda_n$ von K, sodass $D_{B',B'}(\phi) = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, dann heißt ϕ *diagonalisierbar*.

(ii) Gibt es $S \in \mathrm{Gl}_n(K)$ und $\lambda_1, \dots, \lambda_n \in K$ mit $SAS^{-1} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, dann heißt A *diagonalisierbar*.

Ein Endomorphismus ϕ ist diagnonalisierbar genau dann, wenn seine Darstellungsmatrix $D_{B,B}(\phi)$ diagonalisierbar ist. Das liegt daran, wie Ähnlichkeit und Basiswechsel zusammenpassen.

3. Determinante

Wir erinnern an die Signumsfunktion $\operatorname{sgn}\colon S_n \to \{\pm 1\}$. Wir haben uns bereits davon überzeugt, dass folgende Rechenregeln gelten: Für einen k-Zyklus σ ist $\operatorname{sgn}(\sigma) = (-1)^{k+1}$; für $\sigma_1, \sigma_2 \in S_n$ ist $\operatorname{sgn}(\sigma_1 \circ \sigma_n) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$.

Definition V.3.1 (Determinante): Seien K ein Körper, n eine natürliche Zahl und $A \in K^{n\times n}$. Dann heißt

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

die *Determinante von A*.

Beispiel V.3.2: Seien $n = 2$ und $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in K^{2\times 2}$ gegeben. Wir haben $S_2 = \{\mathrm{id}, (12)\}$ mit $\operatorname{sgn}\mathrm{id} = 1$ und $\operatorname{sgn}(12) = -1$, sodass

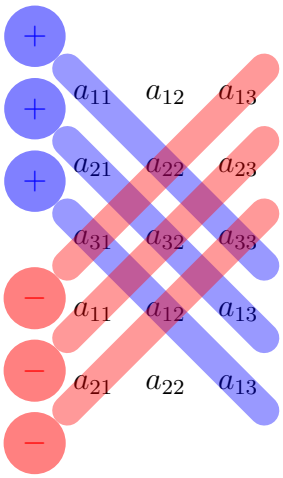
$$\det A = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = ad - bc.$$

Aus den Übungen ist bekannt, dass $\det A$ ein wichtiges Charakteristikum von A ist, das über die Invertierbarkeit von A Aufschluss gibt.

Beispiel V.3.3: Seien $n = 3$ und $A = (a_{i,j}) \in K^{3\times 3}$. Die symmetrische Gruppe vom Grad 3 ist $\{\mathrm{id}, (123), (132), (12), (13), (23)\}$ wobei die Transpositionen negatives Signum und die restlichen Permutationen positives Signum haben. Entsprechend ergibt sich

$$\begin{aligned}\det A = {} & a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ & - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2}.\end{aligned}$$

Für die obige Formel, die auch „Regel von Sarrus“ oder „Jägerzaunregel“ genannt wird, gibt es ein anschauliches Schema:



Eine Regel für $n \geq 4$ ist nicht praktikabel, da die Anzahl der Summanden explodiert. Stattdessen wird auf andere Sätze zur Berechnung von Determinanten zurückgegriffen.

Proposition V.3.4 (Eigenschaften der Determinante): *Seien K ein Körper, n eine natürliche Zahl und $D\colon \prod_{i=1}^{n} K^n \to K$, $(x_1, \ldots, x_n) \mapsto \det(x_1|\ldots|x_n)$.*

(i) *Sind $v_1, \ldots, v_n$ und v_i', $i \in \{1, \ldots, n\}$ in K^n, dann gilt*

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

(ii) *Sind $v_1, \ldots, v_n$ in K^n und $\lambda \in K$, dann ist*

$$D(v_1, \ldots, v_{i-1}, \lambda v_i, v_{i+1}, \ldots, v_n = \lambda D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n).$$

(iii) *Sind $v_1, \ldots, v_n$ in K^n und gibt es $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ und $v_i = v_j$, dann ist $D(v_1, \ldots, v_n) = 0$.*

(iv) *Bezeichnet $\{e_1, \ldots, e_n\}$ die Standardbasis, dann ist $D(e_1, \ldots, e_n) = 1$.*

Beweis: Seien $v_1, \ldots, v_n$ Elemente von K^n und $A = (v_1|\ldots|v_n)$.

(i) Seien i ein Element von $\{1, \ldots, n\}$, $v_i' = (t_1, \ldots, t_n)$ ein Vektor in K^n, $A' = (v_1|\ldots|v_{i-1}|v_i + v_i'|v_{i+1}|\ldots|v_n)$ und $A'' = (v_1|\ldots|v_{i-1}|v_i'|v_{i+1}|\ldots|v_n)$. Dann ist

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

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

(iii) Sei $v_k = v_\ell$ mit $k \neq \ell$ und $\sigma_0 := (k\ell) \in S_n$. Für $\sigma \in S_n$ sei $\sigma' := \sigma \circ \sigma_0$ (wir bemerken, dass $\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\sigma)$), ferner setze $A_n := \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$. Wir erhalten eine Bijektion $A_n \to S_n - A_n, \sigma \mapsto \sigma'$. Weiterhin gilt

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

Da $v_k = v_\ell$ erhalten wir $\prod_{i=1}^n a_{i,\sigma'^{-1}(i)} = \prod_{i=1}^n a_{i,\sigma^{-1}(i)}$, eingesetzt in die Leibniz-Formel gibt das

$$\begin{aligned} \det(A) &= \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} + \sum_{\sigma \in S_n - A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} \\ &= \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} + \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma') \prod_{i=1}^n a_{i,\sigma'(i)} \\ &= \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} - \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} = 0. \end{aligned}$$

(iv) Sei $A = (e_1|\dots|e_n)$. Wegen $a_{i,j} = \delta_{i,j}$ leistet nur id einen Beitrag in $\det A$, d. h. $\det A = \prod_{i=1}^n a_{i,i} = 1$. □

Beispiel V.3.5: Seien K ein Körper und $A = (v_1|v_2|v_3) \in K^{3\times 3}$ gegeben.

(i) Für $A' = (v_1 + \lambda v_2|v_2|v_3)$, $\lambda \in K$, gilt $\det A' = \det A$, denn

$$\det A' = \det(v_1|v_2|v_3) + \lambda \det(v_2|v_2|v_3) = \det A,$$

wobei wir für das erste Gleichheitszeichen die Eigenschaften (i) und (ii) aus Proposition V.3.4 und für das zweite Gleichheitszeichen die Eigenschaft (iv) aus der gleichen Proposition verwendet haben.

(ii) Für $A' = (v_3|v_2|v_1)$ gilt $-\det(v_1|v_2|v_3)$, denn

$$\begin{aligned} 0 &= \det(v_1 + v_3|v_2|v_1 + v_3) \\ &= \det(v_1|v_2|v_1 + v_3) + \det(v_3|v_2|v_1 + v_3) \\ &= \det(v_1|v_2|v_1) + \det(v_1|v_2|v_3) + \det(v_3|v_2|v_1) + \det(v_3|v_2|v_3) \\ &= \det(v_1|v_2|v_3) + \det(v_3|v_2|v_1), \end{aligned}$$

sodass $\det A = \det(v_1|v_2|v_3) = -\det(v_3|v_2|v_1) = -\det A'$.

(iii) Für $A' := (v_1|\lambda v_2|v_3)$ mit $\lambda \in K$ gilt nach Eigenschaft (ii) der Determinante, dass $\det A' = \lambda \det A$.

Zu den elementaren Zeilenumformungen gehörten die Matrizen

(i) $A^{\alpha}_{k,\ell} = I_n + \alpha E_{k,\ell}$ (Additionsmatrizen),

(ii) $V_{k,\ell} = I_n - E_{k,k} - E_{\ell,\ell} + E_{k,\ell} + E_{\ell,k}$ (Vertauschungsmatrizen),

(iii) $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ mit $\alpha_1 \cdots \alpha_n \neq 0$,

die wir im folgenden *spezielle Matrizen* nennen wollen.

Korollar V.3.6 (Determinante und spezielle Matrizen): *Seien K ein Körper, n eine natürliche Zahl und $A \in K^{n\times n}$.*

(i) *Für $A' := AA^{\alpha}_{k,\ell}$ ist* $\det A' = \det A$.

(ii) *Für $A' := AV_{k,\ell}$ ist* $\det A' = -\det A$.

(iii) *Für $A' := A \operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \alpha_1 \cdots \alpha_n \det A$.*

(iv) *Insbesondere haben wir für die speziellen Matrizen:* $\det A^{\alpha}_{k,\ell} = 1$, $\det V_{k,\ell} = -1$, $\det \operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \alpha_1 \cdots \alpha_n$.

(v) *Ist X eine spezielle Matrix, dann ist* $\det(AX) = \det(A)\det(X)$.

Beweis: Die gleichen Rechnungen wie in Proposition V.3.5 zeigen die Aussagen. □

Bemerkung V.3.7: Für die Determinante gelten also folgende Rechenregeln:

(i) Entsteht A' aus A durch Addition der k-ten Spalte zur ℓ-ten Spalte ($k \neq \ell$), dann ist $\det A' = \det A$.

(ii) Entsteht A' aus A durch Vertauschung der k-ten und ℓ-ten Spalte, dann gilt $\det A' = -\det A$.

(iii) Entsteht A' aus A durch Multiplikation einer Spalte mit λ, dann ist $\det A' = \lambda \det A$.

(iv) Enthält A eine Nullspalte, dann ist $\det A = 0$.

Bemerkung V.3.8 (Determinante von Treppenformen): Seien K ein Körper, n eine natürliche Zahl und $T \in K^{n\times n}$ in Treppenform.

(i) Genau dann ist T regulär, wenn $T = I_n$.

(ii) Gehört T nicht zu $\mathrm{Gl}_n(K)$, dann gehört auch T^t nicht zu $\mathrm{Gl}_n(K)$. Ist nämlich T' die Treppenform von T^t, d. h. $T^t = X_1 \cdots X_n T'$ mit speziellen Matrizen $X_1, \ldots, X_n$, dann hat T' eine Nullzeile, d. h. T'^t hat eine Nullspalte. Wegen Proposition V.3.6 gilt dann

$$\det T = \det(T'^t X_n^t \cdots X_1^t) = \det T'^t \det X_1^t \cdots \det X_n^t = 0.$$

Satz 19 (Eigenschaften der Determinante): *Seien K ein Körper, n eine natürliche Zahl und A, A_1, A_2 in $K^{n\times n}$.*

(i) *Genau dann ist* $\det A \neq 0$, *wenn* $A \in \mathrm{Gl}_n(K)$.

(ii) *Es gilt* $\det A = \det A^t$.

(iii) *Es gilt* $\det(A_1 A_2) = \det A_1 \det A_2$.

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

(v) *Die Determinante ist eine Ähnlichkeitsinvariante, d. h. ist* $S \in \mathrm{Gl}_n(K)$, *dann gilt* $\det(SAS^{-1}) = \det A$.

Beweis: (i) Angenommen, A gehörte nicht zu $\mathrm{Gl}_n(K)$. Es gäbe eine Treppenform T' und spezielle Matrizen $X_1, \ldots, X_k$, sodass $A^t = X_1 \cdots X_k T'$. Es wäre dann $A = T'^t X_k^t \cdots X_1^t$, wobei $\det T'^t = 0$. Außerdem wären $X_1^t, \ldots, X_k^t$ ebenfalls Vertauschungsmatrizen. Wegen Proposition V.3.8 und Proposition V.3.6 wäre dann $\det A = 0$.

Angenommen, A gehörte zu $\mathrm{Gl}_n(K)$. Wir haben uns bereits überlegt, dass dann die Einheitsmatrix die Treppenform von A wäre, es gäbe also spezielle Matrizen $X_1, \ldots, X_k$, sodass $A = X_1 \cdots X_k I_n = I_n X_1 \cdots X_k$. Nach Proposition V.3.6 ist also $\det A = \det X_1 \cdots \det X_k \neq 0$.

(ii) Angenommen, A wäre nicht invertierbar. Dann wäre auch A^t nicht invertierbar, und nach (i) hätten wir $\det A = 0 = \det A^t$.

Angenommen, A wäre invertierbar. Dann gäbe es spezielle Matrizen $X_1, \ldots, X_\ell$, sodass $A = X_1 \cdots X_\ell$ und es wäre $A^t = X_\ell^t \cdots X_1^\ell$, d. h. nach Proposition V.3.6 wäre

$$\det A = \det X_1 \cdots \det X_\ell = \det(X_1^t) \cdots \det(X_\ell^t) = \det A^t.$$

(iii) Sind A_1 und A_2 invertierbar, dann sind sowohl A_1 als auch A_2 Produkte spezieller Matrizen und die Behauptung folgt aus Proposition V.3.6.

Ist A_1 invertierbar, aber A_2 nicht, dann ist $\mathrm{Rang}(A_2) \leq n-1$, d. h. nach der Dimensionsformel ist $\dim \mathrm{Kern}\, A_2 \geq 1$. Also gibt es $v \in K^n - \{\mathbf{0}\}$ mit $A_2 v = \mathbf{0}$ und dann ist erst recht $A_1 A_2 v = \mathbf{0}$, was $\dim \mathrm{Kern}\, A_1 A_2 \geq 1$ erzwingt. Damit ist $A_1 A_2$ nicht invertierbar, nach (i) also

$$0 \det(A_1 A_2) = \det A_1 \cdot 0 = \det A_1 \det A_2.$$

Ist A_1 nicht invertierbar, aber A_2 schon, dann ist

$$\det(A_1A_2) = \det(A_2^tA_1^t) = \det A_2^t \det A_1^t = \det A_2 \det A_1.$$

(iv) Ist A invertierbar, dann haben wir

$$\det A^{-1} \det A = \det(A^{-1}A) = \det I_n = 1,$$

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

(v) Wegen (iv) ist $\det(SAS^{-1}) = \det S \det A(\det S)^{-1} = \det A$. □

Korollar V.3.9 (Zeilenoperationen und Determinante): *Seien $A \in K^{n\times n}$ eine Matrix und $z_1, \ldots, z_n$ die Zeilen von A (also $A = (z_1|\ldots|z_n)^t$). Für Zeilenoperationen gelten die analogen Aussagen* (i)-(iii) *aus Proposition V.3.4, das heißt:*

(i) *Für $w \in K^{1\times n}$ gilt*

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

(ii) *Für $\lambda \in K$ gilt* $\det((z_1|\ldots|z_{i-1}|\lambda z_i|z_{i+1}|\ldots|z_n)^t = \lambda \det A$.

(iii) *Gibt es $i \neq j$ mit $z_i = z_j$, dann ist* $\det A = 0$

Damit gelten auch die analogen Aussagen zu denen aus Proposition V.3.6:

(i) $\det(A^{\alpha}_{k,\ell}A) = \det A$,

(ii) $\det(V_{k,\ell}A) = -\det A$,

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

Bemerkung V.3.10 (Rechenregeln für Determinante): Für die Determinante gelten also die folgenden Rechenregeln:

(i) Entsteht A' aus A durch Addition der k-ten Zeile zur ℓ-ten Zeile ($k \neq \ell$), dann ist $\det A' = \det A$.

(ii) Entsteht A' aus A durch Vertauschen der k-ten und der ℓ-ten Zeile ($k \neq \ell$), dann ist $\det A' = -\det A$.

(iii) Entsteht A' aus A durch Multipilikation einer Zeile mit $\lambda \in K$, dann gilt $\det A' = \lambda \det A$.

(iv) Enthält A eine Nullzeile, dann gilt $\det A = 0$.

Definition V.3.11 (Determinante eines Endomorphismus): Seien K ein Körper, V ein n-dimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear. Dann heißt $\det \phi := \det D_{B,B}(\phi)$ die *Determinante von* ϕ.

Die Determinante eines Endomorphismus ist wohldefiniert, da wir bereits gezeigt haben, dass die Determinante eine Ähnlichkeitsinvariante ist, d. h., eine Darstellungsmatrix von ϕ bezüglich einer anderen Basis hat dieselbe Determinante.

Korollar V.3.12 (Eigenwerte und Determinante): *Seien K ein Körper, V ein n-dimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear. Ferner sei λ ein Element von K. Genau dann ist λ ein Eigenwert von A, wenn* $\det(A - \lambda I_n) = 0$.

Definition V.3.13 (Charakteristisches Polynom): Seien K ein Körper, V ein n-dimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$ und $\phi\colon V \to V$ linear. Dann heißt

$$\mathrm{CP}_A := \det(A - XI_n) \in K[X]$$

das *charakteristische Polynom von A*.

Beispiel V.3.14: Für $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ist

$$\mathrm{CP}_A = \det \begin{pmatrix} 1 - X & 1 \\ 0 & 1 - X \end{pmatrix} = (1 - X)^2,$$

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

Bemerkung V.3.15: Die Eigenwerte der Matrix A sind genau die Nullstellen des charakteristischen Polynoms.

Satz 20: *Sei $A \in K^{n \times n}$ mit* $\operatorname{Spec} A = \{\lambda_1, \ldots, \lambda_k\}$. *Dann gilt:*

(i) *Genau dann ist $\lambda \in K$ ein Eigenwert von A, wenn* $\det(A - \lambda I_n) = 0$.

(ii) *Die Summe der Eigenräume ist direkt, d. h.*

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

(iii) *Es gilt $k \leq n$, d. h. A hat höchstens n Eigenwerte.*

(iv) *Die Matrix A ist diagonalisierbar genau dann, wenn*

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

(v) *Für Endomorphismen $\phi\colon V \to V$ wie oben gilt: ϕ hat höchstens $n = \dim V$ Eigenwerte und ist diagonalisierbar genau dann, wenn V die direkte Summe der Eigenräume von ϕ ist.*

Beweis: (i) Das haben wir in Proposition V.3.12 bereits festgehalten.

(ii) Wir zeigen die Aussage per Induktion nach der Anzahl der Eigenwerte. Für $k = 0$ und $k = 1$ ist die Aussage richtig.

Die Aussage gelte jetzt für $k-1$. Sei $\mathbf{0} = u_1 + \cdots + u_k$ mit $u_i \in \mathrm{Eig}(A, \lambda_i)$. Dann haben wir einerseits, dass $\mathbf{0} = \phi(\mathbf{0}) = \sum_{i=1}^{k} \lambda_i u_i$ und andererseits, dass $\mathbf{0} = \lambda_k(\sum_{i=1}^{k} u_i)$. Damit ist

$$\mathbf{0} = (\lambda_1 - \lambda_k)u_1 + \cdots + (\lambda_{k-1} - \lambda_k)u_{k-1} + (\lambda_k - \lambda_k)u_k$$

und da $\lambda_i \neq \lambda_k$ für $i \neq k$ folgt aus der Induktionsvoraussetzung, dass die Vektoren $u_1, \ldots, u_{k-1}$ alle Null sind. Wegen $u_k = -\sum_{i=1}^{k-1} u_i$ muss dann aber auch $u_k = \mathbf{0}$ gelten und die Summe ist direkt.

(iii) Wären $\lambda_1, \ldots, \lambda_{n+1}$ paarweise verschiedene Eigenwerte von A, dann wäre

$$\dim K^n \geq \dim \bigoplus_{i=1}^{n+1} \mathrm{Eig}(A, \lambda_i) \geq n + 1,$$

denn per Definition ist $\mathrm{Eig}(A, \lambda_i) = \mathrm{Kern}(A - \lambda_i I_n) \supsetneq \{\mathbf{0}\}$. Das kann aber nicht sein.

(iv) Das folgt aus der Definition der Diagonalisierbarkeit und (ii).

(v) Folgt aus (iii) und (iv). □

4. Die Regel von Laplace

Definition V.4.1 (Streichmatrix): Seien K ein Körper, n eine natürliche Zahl und A in $K^{n\times n}$. Für $1 \leq i, j \leq n$ bezeichnet $A_{i,j} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

Beispiel V.4.2: Seien $n = 3$ und

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

Für $(i, j) = (1, 2)$ und $(i, j) = (2, 2)$ und $(i, j) = (3, 2)$ haben wir die Streichmatrizen

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

Satz 21 (Entwicklungssatz von Laplace): *Es seien K ein Körper, n eine natürliche Zahl und A in $K^{n\times n}$. Ferner sei $k \in \{1, \ldots, n\}$.*

(i) *Die Laplace-Entwicklung nach der k-ten Zeile ist*

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

(ii) *Die Laplace-Entwickung nach der k-ten Spalte ist*

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

Beispiel V.4.3: Für die Matrix A aus Proposition V.4.2 und $k = 2$ haben wir Folgendes für die Entwicklung nach der zweiten Spalte:

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

Proposition V.4.4 (Blockmatrizen): *Seien $k \in \{1, \ldots, n\}$, $X \in K^{n\times n}$, $Y \in K^{(n-k)\times k}$ und $Z \in K^{(n-k)\times(n-k)}$. Dann ist*

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

Bemerkung V.4.5 (Laplace-Entwicklung nach erster Zeile): Schreibe $A \in K^{n\times n}$ als

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

mit $a_{1,1}, \ldots, a_{1,n} \in K$ und $s_1, \ldots, s_n \in K^{n-1}$. Dann gilt

$$\begin{aligned}
\det A &= \det \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ s_1 & s_2 & \cdots & s_n \end{pmatrix} + \cdots + \det \begin{pmatrix} 0 & \cdots & 0 & a_{1,n} \\ s_1 & \cdots & s_{n-1} & s_n \end{pmatrix} \\
&= \det \begin{pmatrix} a_{1,1} & \mathbf{0} \\ s_1 & A_{1,1} \end{pmatrix} - \det \begin{pmatrix} a_{1,2} & \mathbf{0} \\ s_2 & A_{1,2} \end{pmatrix} \\
&\qquad + \det \begin{pmatrix} a_{1,3} & \mathbf{0} \\ s_3 & A_{1,3} \end{pmatrix} + \cdots + (-1)^{n+1} \begin{pmatrix} a_{1,n} & \mathbf{0} \\ s_n & A_{1,n} \end{pmatrix} \\
&= \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det A_{1,j}.
\end{aligned}$$

Beweis (von Satz 21): (i) Schreibe $A = (z_1 | \ldots | z_n)^t$ als Vektor der Zeilenvektoren von A. Durch $(i-1)$ Zeilenvertauschungen können wir erreichen, dass die i-te Zeile an die Stelle der ersten Zeile rückt, d. h.

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

sodass die vorangegangene Bemerkung die Behauptung liefert.

(ii) Durch Transposition können wir uns auf den ersten Fall zurückziehen. □

Teil 2.

Lineare Algebra II

Kapitel VI.

Die Jordan-Normalform

1. Motivation

Seien K ein Körper und V ein n-dimensionaler K-Vektorraum. Wir möchten Endomorphismen $\phi\colon V \to V$, $v \mapsto \phi(v)$ genauer studieren. Aus der Linearen Algebra I wissen wir bereits:

(i) Bezüglich einer Basis B von V ist ϕ von der Form $x \mapsto Ax$ mit einer geeigneten Matrix $A \in K^{n\times n}$. Genauer: $D_B(\phi(v)) = AD_B(v)$ mit $A = D_{B,B}(\phi)$.

(ii) Zwei Matrizen A, $\bar{A} \in K^{n\times n}$ sind Darstellungsmatrizen desselben Endomorphismus ϕ genau dann, wenn A und $\bar{A}$ ähnlich sind, d. h. es gibt $S \in \mathrm{Gl}_n(K)$, sodass $\bar{A} = SAS^{-1}$.

(iii) Der Endomorphismus ϕ ist diagonalisierbar genau dann, wenn es eine Basis B von V gibt, sodass $D_{B,B}(\phi) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ mit $\lambda_1, \ldots, \lambda_n \in K$. Das ist genau dann der Fall, wenn $V = \bigoplus_{\lambda\in\mathrm{Spec}(\phi)} \mathrm{Eig}(\phi, \lambda)$.

Leider ist ϕ im Allgemeinen nicht diagonalisierbar. Unser Ziel im Folgenden wird sein, für jeden Endomorphismus ϕ eine „besonders schöne“ Darstellungsmatrix J_ϕ zu finden. Das ist äquivalent zur Beschaffung einer „besonders schönen“ Matrix J_A, die ähnlich zur gegebenen Matrix A ist. Dabei möchten wir haben, dass J_A der eindeutige Repräsentant mit diesen Eigenschaften der Ähnlichkeitsklasse von A ist, d. h. A soll ähnlich zu einer Matrix A' sein genau dann, wenn $J_A = J_{A'}$.

Es wird sich zeigen, dass sich dieses Problem in der angegebenen Allgemeinheit nur über bestimmten Körpern, wie dem Körper $\mathbb{C}$, lösen lässt. Wir brauchen Körper, die „algebraisch abgeschlossen“ sind, d. h. jedes $p \in K[X]$ gibt es $\alpha \in K$ mit $f(\alpha) = 0$.

Im Folgenden bezeichnen wir für eine gegebene Matrix A das charakteristische Polynom von A mit χ_A; genau so für einen Endomorphismus ϕ. Zur Erinnerung:

$\chi_A = \det(A - \lambda I_n)$ bzw. $\chi_\phi = \det(\phi - \lambda \,\mathrm{id})$.

Unser Ziel für diesen Abschnitt wird das folgende Resultat sein:

Satz 22: *Seien K ein algebraisch abgeschlossener Körper und V ein K-Vektorraum von Dimension n. Dann gibt es eine geordnete Basis B von V, sodass*

$$J_\phi := D_{B,B}(\phi) = \begin{pmatrix} \lambda_1 & & & & & & & & \\ 1 & \ddots & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & 1 & \lambda_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_k & & & \\ & & & & & 1 & \ddots & & \\ & & & & & & \ddots & \ddots & \\ & & & & & & & 1 & \lambda_k \end{pmatrix}$$

Hierbei sind alle nicht eingetragenen Werte Null, $\lambda_1, \ldots, \lambda_r$ sind die Eigenwerte von ϕ (nicht notwendigerweise verschieden) und die Matrix J_ϕ ist eindeutig bist auf Vertauschen der Kästchen und heißt Jordan-Normalform.

2. Der Satz von Cayley-Hamilton

Das Ziel dieses Abschnitts wird zu zeigen, dass für $A \in K^{n\times n}$ mit charakteristischen Polynom $\chi_A \in K[X]$ gilt: $\chi_A(A) = 0$. Dazu sind einige Vorbemerkungen angebracht.

Bemerkung: Seien K ein Körper und V ein n-dimensionaler K-Vektorraum.

(i) Man kann quadratische Matrizen in Polynome einsetzen. Genauer: Sind $p \in K[X]$ mit $p = \sum_{i=0}^{d} a_i X^i$ und $A \in K^{n\times n}$, dann ergibt der Ausdruck

$$p(A) = a_d A^d + a_{d-1} A^{d-1} + \cdots + a_1 A + a_0 I_n$$

Sinn, denn wir können Matrizen addieren, potenzieren und mit Skalaren multiplizieren.

(ii) Wir haben $Ap(A) = p(A)A$, denn

$$\begin{aligned} Ap(A) &= A(a_d A^d + \cdots + a_1 A + a_0 I_n) \\ &= a_d A^{d+1} + \cdots + a_1 A^2 + a_0 A = (a_d A^d + \cdots + a_1 A + a_0 I_n)A = p(A)A. \end{aligned}$$

(iii) Sind $p_1, p_2 \in K[X]$, dann sind $(p_1 + p_2)(A) = p_1(A) + p_2(A)$ und $(p_1 p_2)(A) = p_1(A)p_2(A)$. Das kann man einfach nachrechnen.

(iv) Man kann genauso Endomorphismen in Polynome einsetzen. Ist genauer $p \in K[X]$ mit $p = \sum_{i=0}^{d} a_i X^i$ und $\phi \in \mathrm{End}(V)$, dann ergibt

$$p(\phi) = a_d \phi^d + \cdots + a_1 \phi + a_0 \,\mathrm{id}_V$$

Sinn. Hier steht ϕ^k für die k-fache Komposition von ϕ mit sich selbst und id_V bezeichnet die Identität von V.

(v) Sind ϕ ein Endomorphismus von V, $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V, $A := D_{B,B}(\phi)$ und $p \in K[X]$, dann gilt $D_{B,B}(p(\phi)) = p(A)$.

Beispiel VI.2.1 (Endomorphismus im Polynom): Seien $V = \mathbb{R}^2$ und $\phi\colon V \to V$, $(x, y) \mapsto (-x, y)$ die „Spiegelung an der y-Achse“. Weiter sei $p_1 = X^2 + 1$. Dann ist $p_1(\phi) = \phi^2 + \mathrm{id}_V = 2\,\mathrm{id}_V$.

Für das Polynom $p_2 = X^2 - 1$ ist $p_2(\phi) = \phi^2 - \mathrm{id}_V = \mathrm{id}_V - \mathrm{id}_V = \mathbf{0}$ die Nullabbildung.

Beispiel VI.2.2 (Zaubertrick): Sei

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

Das charakteristische Polynom von A ist $\chi_A = X^4 + 2X^3 - X^2 - 2X + 1$. Um $\chi_A(A)$ auszuwerten, müssen wir die Potenzen A^2, A^3 und A^4 ausrechnen. Diese sind

$$A^2 = \begin{pmatrix} 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & -5 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 5 \end{pmatrix},$$

$$A^3 = \begin{pmatrix} 0 & -1 & 2 & -5 \\ 0 & 2 & -5 & 12 \\ 0 & 1 & 0 & 0 \\ 1 & -2 & 5 & -10 \end{pmatrix}, \quad A^4 = \begin{pmatrix} -1 & 2 & -5 & 10 \\ 2 & -5 & 12 & -25 \\ 1 & 0 & 0 & 2 \\ -2 & 5 & 10 & 20 \end{pmatrix}.$$

Damit berechnen wir

$$\begin{aligned}\chi_A(A) &= \begin{pmatrix} -1 & 2 & -5 & 10 \\ 2 & -5 & 12 & -25 \\ 1 & 0 & 0 & 2 \\ -2 & 5 & -10 & 20 \end{pmatrix} + \begin{pmatrix} 0 & -2 & 4 & -10 \\ 0 & 4 & -10 & 24 \\ 0 & 2 & 0 & 0 \\ 2 & -4 & 10 & -20 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 5 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -5 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 2 \\ -2 & 0 & 0 & -4 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

Beispiel VI.2.3 (Charakteristisches Polynom für spezielle Matrizen): Es sei

$$A := \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 1 & 0 & 0 & \beta \\ 0 & 1 & 0 & \gamma \\ 0 & 0 & 1 & \delta \end{pmatrix}.$$

Dann ist

$$\begin{aligned}\chi_A &= \det(A - XI_4) \\ &= \det \begin{pmatrix} -X & 0 & 0 & \alpha \\ 1 & -X & 0 & \beta \\ 0 & 1 & -X & \gamma \\ 0 & 0 & 1 & \delta - X \end{pmatrix} \\ &= (-\alpha) \det \begin{pmatrix} 1 & -X & 0 \\ 0 & 1 & -X \\ 0 & 0 & 1 \end{pmatrix} + \beta \det \begin{pmatrix} -X & 0 & 0 \\ 0 & 1 & -X \\ 0 & 0 & 1 \end{pmatrix} \\ &\quad - \gamma \det \begin{pmatrix} -X & 0 & 0 \\ 1 & -X & 0 \\ 0 & 0 & 1 \end{pmatrix} + (\delta - X) \det \begin{pmatrix} -X & 0 & 0 \\ 1 & -X & 0 \\ 0 & 1 & -X \end{pmatrix} \\ &= -\alpha - \beta X - \gamma X^2 - \delta X^3 + X^4.\end{aligned}$$

Bemerkung VI.2.4: Mit (e_1, e_2, e_3, e_4) bezeichnen wir wie gewöhnlich die Standardbasis in K^4. Ferner sei A die Matrix aus Proposition VI.2.3, d. h.

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

Wir halten fest:

(i) $Ae_1 = e_2$, $Ae_2 = e_3$, $Ae_3 = e_4$,

(ii) $Ae_4 = \alpha e_1 + \beta e_2 + \gamma e_3 + \delta e_4$,

(iii) $\chi_A = -\alpha - \beta X - \gamma X^2 - \delta X^3 + X^4$.

Damit erhalten wir:

(iv) $\chi_A(A)e_1 = (-\alpha - \beta A - \gamma A^2 - \delta A^3 + A^4)e_1 = -\alpha e_1 - \beta e_2 - \gamma e_3 - \delta e_4 + A^4 e_4$ nach (i), sodass nach (ii) gilt: $Ae_1 = \mathbf{0}_{K^n}$.

(v) $\chi_A(A)e_2 = \chi_A(A)Ae_1 = A\chi_A e_1 = A\mathbf{0}_{K^n} = \mathbf{0}_{K^n}$.

Analog zeigt man, dass $\chi_A(A)e_3 = \mathbf{0}_{K^n}$ und $\chi_A(A)e_4 = \mathbf{0}_{K^n}$, sodass tatsächlich $\chi_A(A) = \mathbf{0}_{K^{n\times n}}$.

Proposition VI.2.5 (Cayley-Hamilton für spezielle Matrizen): *Sei*

$$A = \begin{pmatrix} & & & \alpha_0 \\ 1 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & 1 & \alpha_{n-1} \end{pmatrix} \in K^{n\times n}.$$

Dann gilt

(i) $\chi_A(X) = (-1)^{n+1}(\alpha_0 + \alpha_1 X + \cdots + \alpha_{n-1}X^{n-1} - X^n)$,

(ii) $\chi_A(A) = \mathbf{0}_{K^{n\times n}}$.

Beweis: Mit den Argumenten aus Proposition VI.2.4 lässt sich die Aussage zeigen. □

Um Proposition VI.2.5 weiter zu verwenden, suchen wir uns zu einem gegebenen Endomorphismus ϕ einen ϕ-invarianten Unterraum (d. h. $\phi(U) \subseteq U$) und eine Basis in U, sodass $D_{B,B}(\phi|_U)$ die Form der Matrix in Proposition VI.2.5 besitzt.

Definition VI.2.6 (ϕ-invariante Unterräume): Seien K ein Körper, V ein K-Vektorraum, $\phi \in \mathrm{End}(V)$ und U ein Untervektorraum von V. Gilt für alle $u \in U$, dass $\phi(u) \in U$, dann heißt U *ϕ-invariant*.

Proposition VI.2.7 (Minimaler ϕ-invarianter Unterraum): *Seien K ein Körper, V ein K-Vektorraum mit $\dim V = m$, $\phi \in \mathrm{End}(V)$ und $v \in V$. Sei weiter n minimal mit der Eigenschaft, dass $\{v, \phi(v), \ldots, \phi^n(v)\}$ linear abhängig ist, d. h. $\phi^n(v) = \sum_{i=0}^{n-1} \alpha_i \phi^i(v)$. Dann gilt:*

(i) *$U := \langle v, \phi(v), \ldots, \phi^{n-1}(v)\rangle$ ist ein ϕ-invarianter Untervektorraum von V, der v enthält.*

(ii) *U ist minimal bezüglich Inklusion mit der Eigenschaft aus* (i).

(iii) *$B = (v, \phi(v), \ldots, \phi^{n-1}(v))$ ist eine geordnete Basis von U und*

$$D_{B,B}(\phi|_U) = \begin{pmatrix} & & & \alpha_0 \\ 1 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & 1 & \alpha_{k-1} \end{pmatrix}$$

Beweis: (i) Wir bemerken, dass $\phi(\phi^{n-1}(v)) = \phi^n(v) = \sum_{i=0}^{n-1} \alpha_i \phi^i(v)$ zu U gehört. Ist jetzt $u \in U$, dann gibt es $c_0, \ldots, c_{n-1} \in K$ mit $u = \sum_{i=0}^{n-1} c_i \phi^i(v)$ und

$$\phi(u) = c_0\phi(v) + c_1\phi^2(v) + \cdots + c_{n-2}\phi^{n-1}(v) + c_{n-1}\phi^n(v)$$

gehört wieder zu U, d. h. U ist ϕ-invariant.

(ii) Sind W ein ϕ-invarianter Unterraum von V und $v \in W$, dann müssen auch die $\phi^i(v)$ für natürliche Exponenten i zu W gehören, d. h. $U \subseteq W$.

(iii) Per Wahl von n ist B eine Basis von U. Setzen wir $b_i := \phi^{i-1}(v)$ für $1 \leq i \leq n$, dann erhalten wir für $1 \leq i \leq n-1$, dass $\phi(b_i) = b_{i+1}$, und für $i = n$ ist $\phi(b_n)\phi^n(v) = \sum_{i=0}^{n-1} \alpha_i b_i$, was die Behauptung über die Darstellungsmatrix zeigt. □

Satz 23 (Cayley-Hamilton): *Seien n eine natürliche Zahl, K ein Körper und V ein n-dimensionaler K-Vektorraum.*

(i) *Für $A \in K^{n\times n}$ gilt $\chi_A(A) = \mathbf{0}_{K^{n\times n}}$.*

(ii) *Für $\phi \in \mathrm{End}(V)$ ist $\chi_\phi(\phi) = \mathbf{0}_{\mathrm{End}(V)}$.*

Beweis: Wir zeigen zuerst (ii). Seien χ_ϕ das charakteristische Polynom von ϕ und $\psi := p(\phi) \in \mathrm{End}(V)$. Wir wollen zeigen, dass $\psi = 0_{\mathrm{End}(V)}$. Dazu genügt es zu zeigen, dass für alle $v \in V$ gilt: $\psi(v) = \mathbf{0}_V$.

Sei also $v \in V$ gegeben. Zunächst wählen wir eine „schöne Basis“. Sei dazu k wie in Proposition VI.2.7 minimal, sodass die Menge $B' := \{v, \phi(v), \ldots, \phi^{k-1}(v)\}$ linear unabhängig ist, d. h. $U := \langle v, \phi(v), \ldots, \phi^{k-1}(v)\rangle$ ist der minimale ϕ-invariante Unterraum von V, der v enthält. Ergänze B' zu einer Basis B von V. Nach Proposition VI.2.7 ist

$$A := D_{B,B}(\phi) = \begin{pmatrix} A' & * \\ \mathbf{0} & C \end{pmatrix}, \qquad A' = \begin{pmatrix} & & & \alpha_0 \\ 1 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & 1 & \alpha_{n-1} \end{pmatrix}$$

Nun berechnen wir das charakteristische Polynom χ_A. Aus der Vorlesung „Lineare Algebra I“ wissen wir, dass die Determinante einer Blockmatrix der Gestalt von A das Produkt der Determinanten der quadratischen Blöcke A' und C ist – auch $\lambda I_n - A$ ist eine Blockmatrix von der gleichen Gestalt wie A, d. h. auch für das charakteristische Polynom gilt das. Entsprechend ist $\chi_A = \chi_{A'} \cdot \chi_C$. Setzen wir A in χ_A ein, dann erhalten wir die Blockmatrix

$$\chi_A(A) = \begin{pmatrix} \chi_{A'}(A') & * \\ \mathbf{0} & \chi_C(C) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & * \\ \mathbf{0} & \chi_C(C) \end{pmatrix}.$$

denn $A^k = \left(\begin{smallmatrix} (A')^k & * \\ \mathbf{0} & C^k \end{smallmatrix}\right)$. Wir haben also gezeigt: $\psi|_U = \mathbf{0}$, insbesondere ist $\psi(v) = \mathbf{0}$.

Aussage (i) folgt jetzt aus (ii) wie folgt: Zur gegebenen Matrix A definieren wir die lineare Abbildung $\phi\colon x \mapsto Ax$, sodass $A = D_{E,E}(\phi)$. Insbesondere gilt $\chi_A = \chi_\phi$, d. h. $\chi_A(A) = D_{E,E}(\chi_A(\phi)) = D_{E,E}(\chi_\phi(\phi)) = \mathbf{0}_{K^{n\times n}}$. □

Bemerkung VI.2.8: Sei A eine Matrix in $K^{n\times n}$. Bei der Definition des charakteristischen Polynoms $\chi_A = \det(A - XI_n)$ berechnen wir eigentlich die Determinante einer Matrix über dem Ring $R := K[X]$. Hierfür benötigen wir allgemeiner Determinanten von Matrizen über Ringen R, also $\det\colon R^{n\times n} \to R$. Diese kann genau so wie über Körpern definiert werden und es gelten auch die Regel von Laplace und die Rechenregeln für elementare Zeilen- und Spaltenoperationen.

Bemerkung VI.2.9 (Alternativer Beweis für Cayley-Hamilton): Es seien K ein Körper und $A \in K^{n\times n}$ gegeben. Wir wollen zeigen, dass $\chi_A(A) = \mathbf{0}$ in $K^{n\times n}$.

Wegen $\chi_A(X) = \det(A - XI_n)$ ist $\chi_A(A) = \det(A - AI_n) = \det(A - A) = 0$. Das ist Quatsch!

$$A - XI_n = \begin{pmatrix} a_{1,1} - X & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - X & \dots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - X \end{pmatrix} \in (K[X])^{n\times n}.$$

Ist $B \in K^{n\times n}$, dann gilt *nicht*: $\chi_A(B) = A - BI_n = A - B \in K^{n\times n}$.

Definition VI.2.10: Seien K ein Körper und $A \in K^{n\times n}$. Gibt es eine natürliche Zahl k, sodass $A^k = \mathbf{0}$, dann heißt A *nilpotent*.

Beispiel VI.2.11: Die folgende Matrix A ist nicht die Nullmatrix, aber nilpotent:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Proposition VI.2.12 (Spektrum nilpotenter Matrizen): *Seien K ein Körper und $A \in K^{n\times n}$ eine nilpotente Matrix.*

(i) *Es ist* $\operatorname{Spec} A \subseteq \{0\}$.

(ii) *Ist K algebraisch abgeschlossen, dann ist* $\operatorname{Spec} A = \{0\}$.

Beweis: (i) Da A nilpotent ist, gibt es eine natürliche Zahl k, sodass $A^k = 0$. Ist jetzt $\lambda \in K$ ein Eigenwert von A, dann gibt es $v \in V - \{0\}$ mit $Av = \lambda v$. Vom ersten Tutoriumsblatt ist bekannt, dass λ^k ein Eigenwert von A^k ist, d. h. $A^k v = \mathbf{0}v = \lambda^k v$, sodass $\lambda^k = 0$, was $\lambda = 0$ impliziert.

(ii) Da K algebraisch abgeschlossen ist, hat χ_A eine Nullstelle und diese muss nach (i) Null sein. □

3. Der Polynomring über einem Körper

Aus dem Satz von Cayley-Hamilton wissen wir für eine Matrix $A \in K^{n\times n}$, dass $\chi_A(A) = \mathbf{0}$. Gibt es außer Vielfachen von χ_A noch weitere Polynome, die das auch erfüllen? Wir werden das im folgenden Abschnitt systematisch untersuchen.

Es wird sich herausstellen, dass sich der Polynomring $K[X]$ in seinen Eigenschaften im Wesentlichen wie der Ring der ganzen Zahlen verhält, und dass es ein nicht-triviales Polynom kleinsten Grades gibt, sodass jedes Polynom, das die Matrix A annulliert, ein Vielfaches dieses kleinsten Polynomes ist.

Erinnerung VI.3.1: Sei K ein Körper. Dann heißt

$$K[X] := \left\{ \sum_{i=1}^{n} a_i X^i : n \in \mathbb{N}_0, a_i \in K \right\}$$

mit der Addition und Multiplikation von Polynomen der *Polynomring über K*.

Definition VI.3.2 (K-Algebra): Seien K ein Körper und A eine Menge mit drei Abbildungen

$$+ \colon A \times A \longrightarrow A, \qquad \bullet \colon A \times A \longrightarrow A, \qquad \cdot \colon K \times A \longrightarrow A.$$

Falls gilt

(i) $(A, +, \bullet)$ ist ein Ring,

(ii) $(A, +, \cdot)$ ist ein K-Vektorraum,

(iii) „$\bullet$“ ist bilinear, d. h. für alle $x, y, z \in A$ und $\lambda \in K$ gilt

$$(x + y) \bullet z = x \bullet z + y \bullet z,$$
$$x \bullet (y + z) = x \bullet y + x \bullet z, \qquad (\lambda x) \bullet y = \lambda(x \bullet y) = x \bullet (\lambda y),$$

dann heißt A eine *K-Algebra*.

Ist $(A, +, \bullet)$ ein kommutativer bzw. unitärer Ring, dann heißt A eine *kommutative K-Algebra* bzw. *unitäre K-Algebra* oder *K-Algebra mit Eins* (d. h. es gibt $1 \in A$, sodass für alle $a \in A$ gilt: $1 \bullet a = a = a \bullet 1$).

Seien A_1 und A_2 zwei K-Algebren und $\phi \colon A_1 \to A_2$ eine Abbildung. Ist ϕ ein Homomorphismus von K-Vektorräumen und ein Homomorphismus von (unitären) Ringen, dann heißt ϕ ein *Homomorphismus (unitärer) K-Algebren*.

Ist $\phi \colon A_1 \to A_2$ ein Homomorphismus (unitärer) K-Algebren und gibt es einen Homomorphismus (unitärer) K-Algebren $\psi \colon K_2 \to K_1$ mit $\phi \circ \psi = \mathrm{id}_{A_2}$ und $\psi \circ \phi = \mathrm{id}_{A_1}$, dann heißt ϕ ein *Isomorphismus (unitärer) K-Algebren*.

Wie bei Gruppen, Vektorräumen und Ringen zeigt man, dass bijektive Homomorphismen genau die Isomorphismen sind.

Beispiel VI.3.3 (Erste K-Algebren): Sei K ein Körper.

- Der Polynomring $K[X]$ mit Addition, Multiplikation und skalarer Multiplikation ist eine K-Algebra.
- Der Ring der $n \times n$- Matrizen $K^{n\times n}$ ist eine K-Algebra mit Addition, skalarer Multiplikation und Multiplikation von Matrizen.
- Für einen K-Vektorraum V ist $\mathrm{End}_K(V)$ zusammen mit Addition, skalarer Multiplikation und Verknüpfung von Endomorphismen eine K-Algebra.

Sei jetzt V ein endlichdimensionaler K-Vektorraum mit geordneter Basis B. Dann ist

$$D_{B,B}\colon \mathrm{End}_K(V) \longrightarrow K^{n\times n}, \qquad \phi \longmapsto D_{B,B}(\phi)$$

ein Isomorphismus unitärer K-Algebren.

Definition VI.3.4 (Einsetzen in Polynome): Seien K ein Körper und $(A, +, \bullet, \cdot)$ eine unitäre K-Algebra. Dann kann man die Elemente aus A in Polynome aus $K[X]$ einsetzen. Genauer: Für $p = \sum_{i=0}^{n} c_i X^i \in K[X]$ und $a \in A$ setzen wir

$$p(a) := \sum_{i=0}^{n} c_i a^i$$

wobei a^i die i-fache Multiplikation von a mit sich selbst bezeichnet und wir die Konvention $a^0 := 1_A$ verwenden.

Definition VI.3.5 (Einsetzungsmorphismus): Seien K ein Körper, $(A, +, \bullet, \cdot)$ eine unitäre K-Algebra und $a \in A$. Dann heißt

$$\varphi_a\colon K[X] \longrightarrow A, \qquad p \longmapsto p(a)$$

der *Einsetzungshomomorphismus von a*. Tatsächlich ist der Einsetzungshomomorphismus ein Homomorphismus unitärer K-Algebren, d. h. es gilt für alle $p_1, p_2 \in K[X]$ und $\lambda \in K$:

$$(p_1+p_2)(a) = p_1(a)+p_2(a), \qquad (p_1p_2)(a) = p_1(a)\bullet p_2(a), \qquad (\lambda p)(a) = \lambda p(a).$$

Definition VI.3.6: Seien K ein Körper und $p = \sum_{i=0}^{n} c_i X^i$ in $K[X]$. Die Zahl

$$\deg p := \begin{cases} n, & \text{falls } c_n \neq 0, \\ -\infty, & \text{falls } p = 0 \end{cases}$$

heißt der *Grad des Polynoms p*. Hierbei ist „$-\infty$“ ein Symbol, das nicht zu $\mathbb{N}_0$ gehört. Für das Symbol „$-\infty$“ definieren wir folgende Rechenregeln:

(i) Für alle $k \in \mathbb{N}_0$ ist $\max\{k, -\infty\} = k$,

(ii) $\max\{-\infty, -\infty\} = -\infty$,

(iii) Für alle $k \in \mathbb{N}_0$ ist $k + (-\infty) := -\infty =: (-\infty) + k$,

(iv) $(-\infty) + (-\infty) := -\infty$,

(v) Für alle $k \in \mathbb{N}_0$ ist $-\infty < k$.

Ist p nicht das Nullpolynom, und ist $c_n \neq 0$, dann heißt c_n der *Leitkoeffizient von p*. Ist 1 der Leitkoeffizient von p, dann heißt p *normiert*.

Bemerkung VI.3.7: Seien K ein Körper und $p_1, p_2 \in K[X]$. Dann gilt:

(i) $\deg(p_1 + p_2) \leq \max\{\deg(p_1), \deg(p_2)\}$,

(ii) $\deg(p_1 p_2) = \deg(p_1) + \deg(p_2)$,

(iii) Die Einheitengruppe $K[X]^\times$ (d. h. die multiplikativ invertierbaren Polynome) lässt sich folgendermaßen beschreiben:

$$\begin{aligned} K[X]^\times &= \{f \in K[X] \mid \text{Es gibt } g \in K[X] \text{ mit } gf = fg = 1\} \\ &= \{f \in K[X] \mid \deg(f) = 0\} \cong K^\times. \end{aligned}$$

Die beiden Gruppen sind isomorph vermöge $\lambda \mapsto \lambda X^0$.

(iv) Es gibt keine *Nullteiler* in $K[X]$, d. h. es gibt kein $f \in K[X] - \{0\}$, sodass es $g \in K[X] - \{0\}$ gibt mit $gf = 0$.

(v) Für unitäre kommutative Ringe R kann man genauso den Polynomring $R[X]$ definieren. In (ii) gilt dann allerdings nur noch „$\leq$“.

Proposition VI.3.8 (Polynomdivision mit Rest): *Seien K ein Körper und p_1, p_2 in $K[X]$ mit $p_2 \neq 0$. Dann gibt es Polynome $h, r \in K[X]$ mit $\deg(r) < \deg(p_2)$ und $p_1 = hp_2 + r$.*

Definition VI.3.9 (Verschwindungsideal): Seien K ein Körper, n eine natürliche Zahl und V ein K-Vektorraum.

(i) Für $A \in K^{n\times n}$ heißt die Menge $I(A) := \{f \in K[X] \mid f(A) = \mathbf{0}\}$ *Verschwindungsideal von A*.

(ii) Für $\phi \in \mathrm{End}(V)$ heißt die Menge $I(\phi) := \{f \in K[X] \mid f(\phi) = \mathbf{0}\}$ *Verschwindungsideal von ϕ*.

Bemerkung VI.3.10 (Eigenschaften des Verschwindungsideals): (i) Die Verschwindungsideale sind nicht leer, denn das Nullpolynom gehört jedenfalls immer dazu.

(ii) Nach dem Satz von Cayley-Hamilton gehört zu $I(A)$ auch das charakteristische Polynom χ_A.

(iii) Sind p_1 und p_2 in $I(A)$, dann gehört wegen Definition VI.3.5 auch die Summe $p_1 + p_2$ zu $I(A)$, denn $(p_1 + p_2(A) = p_1(A) + p_2(A) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.

(iv) Sind h in $K[X]$ und p in $I(A)$, dann gehört wegen Definition VI.3.5 auch hp zu $I(A)$, denn $(hp)(A) = h(A)p(A) = h(A)\mathbf{0} = \mathbf{0}$.

Geleitet von den Beobachtungen der vorangegangenen Bemerkung wollen wir Teilmengen von Ringen mit den Eigenschaften (i), (iii) und (iv) Ideale nennen.

Definition VI.3.11: Seien R ein unitärer kommutativer Ring und $\varnothing \neq I \subseteq R$ eine Teilmenge. Falls gilt:

(i) Für alle $a, b \in I$ ist $a + b \in I$,

(ii) Für alle $a \in I$ und $r \in R$ ist $ra \in I$,

dann heißt I ein *Ideal*.

Proposition VI.3.12 (Konstruktion von Idealen): *Seien R ein kommutativer unitärer Ring und $a_1, \ldots, a_k$ Elemente von R. Dann ist*

$$I := Ra_1 + \cdots + Ra_k := (a_1, \ldots, a_k) := \{r_1a_1 + \cdots + r_ka_k \mid r_1, \ldots, r_k \in R\}$$

ein Ideal. Insbesondere erhalten wir für $a \in R$, dass $Ra := \{ra \mid r \in R\} = (a)$ ein Ideal ist. Dieses heißt das von a erzeugte Hauptideal.

Beweis: (i) Seien $a = r_1a_1 + \cdots + r_ka_k$ und $\bar{a} = \bar{r}_1a_1 + \cdots + \bar{r}_ka_k$ Elemente von I. Dann ist

$$a + \bar{a} = r_1a_1 + \cdots + r_ka_k + \bar{r}_1a_1 + \cdots + \bar{r}_ka_k = (r_1 + \bar{r}_1)a_1 + \cdots + (r_k + \bar{r}_k)a_k,$$

d. h. auch $a + \bar{a}$ gehört zu I.

(ii) Für $a = r_1a_1 + \cdots + r_ka_k$ in I und $r \in R$ ist $ra = (rr_1)a_1 + \cdots + (rr_k)a_k$ in I. □

Beispiel VI.3.13: Sei R ein kommutativer unitärer Ring. Dann gibt es zwei sogenannte *triviale Ideale*; zum Einen ist $N := (0) := \{0\}$ ein Ideal von R, das sogenannte *Nullideal*, und zum Anderen ist R selbst ein Ideal in R.

Definition VI.3.14: Sei R ein kommutativer unitärer Ring.

(i) Falls für alle r_1, r_2 in R gilt: „Ist $r_1 r_2 = 0$, dann ist $r_1 = 0$ oder $r_2 = 0$“, dann heißt R *nullteilerfrei*.

(ii) Sei $I \subseteq R$ ein Ideal. Gibt es $m \in I$, sodass $I = Rm = (m)$, dann heißt I ein *Hauptideal*.

(iii) Sei R zusätzlich nullteilerfrei. Ist jedes Ideal von R ein Hauptideal, dann heißt R ein *Hauptidealring*.

Satz 24 (Polynomring als Hauptidealring): *Seien K ein Körper und $K[X]$ der Polynomring über K. Dann ist $K[X]$ ein Hauptidealring, d. h. für jedes Ideal $I \subseteq K[X]$ gibt es ein Polynom $p_0 \in I$, sodass $I = (p_0) = \{hp_0 \mid h \in K[X]\}$.*

Beweis: Wir wissen bereits, dass der Polynomring nullteilerfrei und kommutativ ist. Sei nun $I \subseteq K[X]$ ein Ideal. Ist I das Nullideal, dann ist I auch ein Hauptideal. Sonst gibt es $p_0 \in I - \{0\}$, sodass $\deg(p_0) \leq \deg(p)$ für alle $p \in I - \{0\}$. Wir zeigen jetzt, dass $I = \{hp_0 \mid h \in K[X]\}$.

Die Inklusion „$\supseteq$“ ist dabei klar, da p_0 zu I gehört und damit auch alle Vielfachen von p_0.

Für „$\subseteq$“ sei nun $p \in I$ gegeben. Da wir in $K[X]$ Polynomdivision durchführen können, gibt es Polynome h und r, sodass $p = hp_0 + r$ mit $\deg(r) < \deg(p_0)$. Wegen $r = p - hp_0$ folgt aus der Minimalität des Grades von p_0 schon, dass r das Nullpolynom sein muss. Also ist $p = hp_0$ wie gewünscht. □

Der obige Beweis greift für alle nullteilerfreien kommutativen unitären Ringe, in denen es das Konzept von Teilbarkeit mit Rest gibt. Solche Ringe heißen *euklidische Ringe*. Das heißt: Alle euklidischen Ringe sind Hauptidealringe. Insbesondere ist der Ring der ganzen Zahlen $\mathbb{Z}$ ein Hauptidealring.

Proposition VI.3.15 (Eindeutigkeit des Idealerzeugers): *Seien K ein Körper, $K[X]$ der Polynomring über K und $I \subseteq K[X]$ ein Ideal. Dann ist der Erzeuger p_0 aus dem Beweis von Satz 24 eindeutig bis auf einen skalaren, von Null verschiedenen Faktor. Genauer: Ist $I = K[X]f = K[X]g$ mit $f, g \in K[X]$, dann gibt es $\lambda \in K^\times$ mit $g = \lambda f$.*

Beweis: Für das Nullideal $I = (0)$ gilt die Aussage. Sei also jetzt $I \neq (0)$ mit $I = K[X]f = K[X]g$. Dann gibt es $h, h' \in K[X]$ mit $g = h'f$ und $f = hg$, d. h. $f = hg = hh'f$, sodass $(1 - hh')f = 0$. Weil f nach Voraussetzung nicht das Nullpolynom ist und $K[X]$ nullteilerfrei ist, muss $1 - hh' = 0$ sein, sodass $1 = hh'$. Damit gehören h und h' zu $K[X]^\times$ und $h' = \lambda X^0$ mit $\lambda \in K^\times$ wie gewünscht. □

Für den Nachweis haben wir nur benötigt, dass $K[X]$ keine Nullteiler hat und damit gezeigt, dass zwei Erzeuger eines Hauptideals (das vom Nullideal verschieden ist) bis auf Mulitplikation mit einer Einheit gleich sind.

Definition VI.3.16 (Teiler in $K[X]$): Seien K ein Körper und $f, g \in K[X]$.

(i) Gibt es $h \in K[X]$ mit $g = hf$, dann heißt f ein *Teiler von* g. In diesem Fall schreiben wir „$f \mid g$“.

(ii) Gilt für alle $h \in K[X]$ mit $h \mid f$ und $h \mid g$, dass $h \in K[X]^\times$, dann heißen f und g *teilerfremd*.

(iii) Sei $f \in K[X] - (K^\times \cup \{0\})$. Gilt für alle h_1 und h_2 in $K[X]$ mit $f = h_1 h_2$, dass $h_1 \in K[X]^\times$ oder $h_2 \in K[X]^\times$, dann heißt f *irreduzibel*.

Ein Polynom heißt irreduzibel, falls es keine *echten* Teiler hat – durch Einheiten kann man immer teilen.

Lemma VI.3.17 (Lemma von Bézout): *Seien K ein Körper und $f, g \in K[X]$. Die Polynome f und g sind teilerfremd genau dann, wenn es h_1 und $h_2 \in K[X]$ mit $1 = h_1 f + h_2 g$ gibt.*

Beweis: „$\Longleftarrow$“: Sei $h \in K[X]$ mit $h \mid f$ und $h \mid g$, d. h. es gibt f' und g' in $K[X]$, sodass $f = hf'$ und $g = hg'$. Dann ist

$$1 = h_1 h f' + h_2 h g' = h(h_1 f' + h_2 g'),$$

sodass h eine Einheit sein muss.

„$\Longrightarrow$“: Sei $I = (f, g) = K[X]f + K[X]g$. Nach Proposition VI.3.12 ist I ein Ideal in $K[X]$. Wegen Satz 24 ist I außerdem ein Hauptideal, d. h. $I = (p_0)$ für ein $p_0 \in K[X]$. Weil f und g zu I gehören, ist p_0 ein Teiler von f und ein Teiler von g und da f und g teilerfremd sind, ist p_0 eine Einheit. Es gibt also $p_0' \in K[X]^\times$ mit $p_0' p_0 = 1$. Insbesondere gehört 1 zu I, d. h. es gibt h_1 und $h_2 \in K[X]$, sodass $1 = h_1 f + h_2 g$ wie gewünscht. □

Die Implikation „$\Longleftarrow$“ gilt in allen kommutativen unitären Ringen. Die Implikation „$\Longrightarrow$“ hat nur verwendet, dass $K[X]$ ein Hauptidealring ist. Somit gilt das Lemma von Bézout in allen Hauptidealringen, insbesondere auch im Ring der ganzen Zahlen $\mathbb{Z}$.

Proposition VI.3.18: *Seien K ein Körper und $f \in K[X]$.*

(i) *Genau dann ist $a \in K$ eine Nullstelle von f, wenn $(X - a)$ ein Teiler von f ist.*

(ii) *Gibt es* $a_1, \ldots, a_n \in K$, *sodass* $f = \prod_{i=1}^{n}(X - a_i)$ *und ist* $h \in K[X]$ *ein normierter Teiler von* f, *dann ist* $h = \prod_{j=1}^{k}(X - a_{i_j})$, *wobei* $i_1, \ldots, i_k$ *in* $\{1, \ldots, n\}$ *mit* $1 \leq i_1 < \cdots < i_k \leq n$.

Beweis: (i) „$\Longleftarrow$“: Angenommen, $f = g \cdot (X - a)$ für ein $g \in K[X]$. Dann ist $f(a) = 0 \cdot g(a) = 0$, da Einsetzen ein Algebrenhomomorphismus ist.

„$\Longrightarrow$“: Polynomdivision ergibt, dass $f = g \cdot (X - a) + r$ mit $\deg(r) \leq 0$. Aber das heißt $r = cX^0$ mit $c \in K$ und wegen $0 = f(a) = g(a) \cdot 0 + r(a) = c$ muss sogar $c = 0$ gelten. Damit ist $(X - a)$ ein Teiler von f, wie gewünscht.

(ii) Wir werden später (voraussichtlich in Kapitel V) sehen, dass sich jedes Polynom in $K[X]$ eindeutig (bis auf Reihenfolge und Multiplikation mit Einheiten) als Produkt irreduzibler Polynome schreiben lässt. Das heißt: Auch in $K[X]$ gibt es (wie im Ring der ganzen Zahlen $\mathbb{Z}$) eine eindeutige Primfaktorzerlegung.

Damit erhalten wir: Unser Polynom f lässt sich schreiben als $f'h$ mit f' aus $K[X]$ und f' sowie h lassen sich „eindeutig“ in irreduzible Faktoren zerlegen, sagen wir $f' = f'_1 \cdots f'_s$ und $h = h_1 \cdot h_k$. Deshalb ist

$$f = f'_1 \cdots f'_s \cdot h_1 \cdots h_k = (X - a_1) \cdots (X - a_n)$$

und die Behauptung folgt dann aus der Eindeutigkeit der Primfaktorzerlegung. □

Korollar VI.3.19: *Seien* K *ein Körper,* $a \in K$ *und* $n \in \mathbb{N}_0$. *Ferner seien* f *und* g *Polynome in* $K[X]$ *mit* $f = (X - a)^n$, *und* $g(a) \neq 0$. *Dann sind* f *und* g *teilerfremd.*

Definition VI.3.20: Seien K ein Körper und f ein Polynom in $K[X]$.

(i) Die Menge $\mathrm{Nst}(f) := \{a \in K \mid f(a) = 0\}$ heißt *Nullstellenmenge von* f.

(ii) Ist $g \in K[X]$ ein Teiler von f, dann ist $\mathrm{Nst}(g) \subseteq \mathrm{Nst}(f)$.

4. Das Minimalpolynom

In diesem Abschnitt wollen wir das kleinste Polynom kennenlernen, das einen Endomorphismus bzw. eine Matrix annulliert. Dieses Polynom heißt *Minimalpolynom*. Wir werden feststellen, dass die Nullstellenmenge des Minimalpolynoms mit der des charakteristischen Polynoms übereinstimmt, also gleich dem Spektrum des Endomorphismus bzw. der Matrix ist.

Bemerkung VI.4.1: Seien n eine natürliche Zahl, K ein Körper, V ein K-Vektorraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus bzw. $A \in K^{n\times n}$. Da $K[X]$ ein Hauptidealring ist, haben die Verschwindungsideale $I(\phi)$ respektive $I(A)$ eindeutige normierte Erzeuger kleinstens Grades m_ϕ respektive m_A.

Definition VI.4.2 (Minimalpolynom): Seien n eine natürliche Zahl, K ein Körper, V ein K-Vektorraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus bzw. $A \in K^{n\times n}$. Der Erzeuger m_ϕ respektive m_A des Verschwindungsideals $I(\phi)$ respektive $I(A)$ heißt *Minimalpolynom von* ϕ respektive *Minimalpolynom von* A.

Insbesondere ist das Minimalpolynom dasjenige Polynom f kleinsten nichtnegativen Grades, das ϕ respektive A annulliert, d. h. das $f(\phi) = \mathbf{0}$ respektive $f(A) = \mathbf{0}$ leistet.

Bemerkung VI.4.3: Aus dem Satz von Cayley-Hamilton folgt, dass m_ϕ ein Teiler des charakteristischen Polynoms χ_ϕ ist und genauso, dass m_A ein Teiler des charakteristischen Polynoms χ_A ist.

Proposition VI.4.4: *Seien n eine natürliche Zahl, K ein Körper, V ein K-Vektorraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus bzw. $A \in K^{n\times n}$. Dann gilt:*

(i) $\mathrm{Nst}(m_\phi) = \mathrm{Nst}(\chi_\phi) = \mathrm{Spec}(\phi)$,

(ii) $\mathrm{Nst}(m_A) = \mathrm{Nst}(\chi_A) = \mathrm{Spec}(A)$.

Beweis: Wir zeigen (ii), (i) geht völlig analog. Wie in Proposition VI.3.20 bemerkt, haben wir $\mathrm{Nst}(m_A) \subseteq \mathrm{Nst}(\chi_A)$. Bleibt also „$\supseteq$“ zu zeigen.

Sei dazu $\lambda \in \mathrm{Nst}(\chi_A)$. Dann ist λ ein Eigenwert von A, d. h. $m_A(\lambda)$ ist ein Eigenwert von $m_A(A) = \mathbf{0}$. Wir haben also $m_A(\lambda) = 0$, d. h. $\lambda \in \mathrm{Nst}(m_A)$. □

Beispiel VI.4.5: Es sei K ein Körper.

(i) Es seien $\lambda_1, \ldots, \lambda_n \in K$ und $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Das charakteristische Polynom von A ist $\chi_A = \prod_{i=1}^n (X - \lambda_i)$, aber das Minimalpolynom von A ist $m_A = \prod_{\lambda\in\mathrm{Spec}(A)} (X - \lambda)$. Insbesondere haben wir $\deg m_A = \#\,\mathrm{Spec}(A)$ und $\deg \chi_A = n$.

(ii) Seien $a \in K$ und

$$A = \begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix} \in K^{4\times 4}.$$

Da A eine untere Dreiecksmatrix ist, ist $\chi_A = \det(A - XI_4) = (X-a)^4$. Wir haben

$$A - aI_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

und da wir die Potenzen von $(A - aI_4)$ schon ausgerechnet haben, wissen wir bereits, dass erst $(A - aI_4)^4 = \mathbf{0}$. In diesem Fall ist also das Minimalpolynom gleich dem charakteristischen Polynom.

(iii) Betrachte $\Phi\colon K^{n\times n} \to K^{n\times n}$, $A \mapsto A^t$. Wir wissen, dass $(A^t)^t = A$, sodass $\Phi^2 = \mathrm{id}_{K^{n\times n}}$ gilt. Das Polynom $X^2 - 1$ annulliert also Φ. Ein Polynom kleineren nicht-negativen Grades kann nicht annullieren, sodass $m_\Phi = X^2 - 1$. Da $\dim K^{n\times n} = n^2$ ist auch $\deg \chi_\Phi = n^2$, die Eigenwerte von Φ sind ± 1. Überlegen Sie sich, wie das charakteristische Polynom von Φ aussieht!

Um uns das Leben leichter zu machen, möchten wir im Folgenden das charakteristische Polynom in teilerfremde Faktoren zerlegen und diese zuerst bearbeiten.

Notation VI.4.6: Seien K ein Körper, n eine natürliche Zahl und $A \in K^{n\times n}$. Mit $\phi_A\colon K^n \to K^n$ bezeichnen wir die zugehörige lineare Abbildung und wir setzen $\mathrm{Bild}(A) := \mathrm{Bild}(\phi_A)$ sowie $\mathrm{Kern}(A) := \mathrm{Kern}(\phi_A)$.

Lemma VI.4.7 (Zerlegungslemma für Matrizen): *Seien K ein Körper, n eine natürliche Zahl, $A \in K^{n\times n}$ und g_1, g_2 sowie g Polynome in $K[X]$.*

(i) *Sind g_1 und g_2 teilerfremd, dann gilt*

$$\mathrm{Kern}(g_1(A)) \cap \mathrm{Kern}(g_2(A)) = \{\mathbf{0}\}, \qquad \mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A)) = K^n.$$

(ii) *Ist $(g_1g_2)(A) = \mathbf{0}$, also $g_1g_2 \in I(A)$, dann gilt*

$$\mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_1(A)), \qquad \mathrm{Bild}(g_1(A)) \subseteq \mathrm{Kern}(g_2(A)).$$

(iii) *Sind g_1 und g_2 teilerfremd und $g_1g_2 \in I(A)$, dann haben wir*

$$K^n = \mathrm{Kern}(g_1(A)) \oplus \mathrm{Kern}(g_2(A))$$

sowie $\mathrm{Kern}(g_1(A)) = \mathrm{Bild}(g_2(A))$ *und* $\mathrm{Kern}(g_2(A)) = \mathrm{Bild}(g_1(A))$.

(iv) *Die Räume* $\mathrm{Kern}(g(A))$ *und* $\mathrm{Bild}(g(A))$ *sind A-invariante Untervektorräume des K^n.*

Beweis: (i) Aus dem Lemma von Bézout wissen wir, dass es Polynome h_1 und h_2 mit $1 = h_1g_1 + h_2g_2$ gibt. Setzen wir A in diese Darstellung ein, so erhalten wir

$$I_n = h_1(A)g_1(A) + h_2(A)g_2(A) = g_1(A)h_1(A) + g_2(A)h_2(A).$$

Für $v \in \mathrm{Kern}(g_1(A)) \cap \mathrm{Kern}(g_2(A))$ erhalten wir mit der ersten Gleichheit aus der obigen Gleichung, dass

$$v = h_1(A)g_1(A)v + h_2(A)g_2(A)v = h_1(A)\mathbf{0} + h_2(A)\mathbf{0} = \mathbf{0}$$

und für $v \in K^n$ ist wegen der zweiten Gleichheit in der ersten Gleichung, dass $v = g_1(A)h_1(A)v + g_2(A)h_2(A)v$, d. h. v gehört zu $\mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A))$.

(ii) Es sei $v = g_2(A)w$ ein Element von $\mathrm{Bild}(g_2(A))$. Dann ist

$$g_1(A)v = g_1(A)g_2(A)w = (g_1g_2)(A)w = \mathbf{0},$$

sodass v zu $\mathrm{Kern}(g_1(A))$ gehört und deshalb $\mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_1(A))$. Die zweite Aussage zeigt man genau mit Vertauschung der Indizes.

(iii) Aus (i) wissen wir, dass $K^n = \mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A))$. Aus (ii) wissen wir außerdem, dass $\mathrm{Bild}(g_i(A)) \subseteq \mathrm{Kern}(g_j(A))$ für $i \in \{1, 2\}$ und $j \in \{1, 2\} - \{i\}$. Wieder wegen (i) ist $K^n = \mathrm{Kern}(g_2(A)) \oplus \mathrm{Kern}(g_1(A))$, d. h. wir haben $\mathrm{Bild}(g_1(A)) + \mathrm{Bild}(g_2(A)) \subseteq \mathrm{Kern}(g_2(A)) \oplus \mathrm{Kern}(g_1(A))$. Aus Dimensionsgründen folgern wir jetzt

$$\mathrm{Bild}(g_1(A)) = \mathrm{Kern}(g_2(A)), \qquad \mathrm{Bild}(g_2(A)) = \mathrm{Kern}(g_1(A)).$$

(iv) Gehört v zu $\mathrm{Kern}(g(A))$, dann ist $g(A)v = \mathbf{0}$. Da A mit $g(A)$ vertauscht ist deshalb auch $g(A)Av = Ag(A)v = A\mathbf{0} = \mathbf{0}$, d. h. Av gehört zu $\mathrm{Kern}(g(A))$.

Gehört v zu $\mathrm{Bild}(g(A))$, dann gibt es irgendein $w \in K^n$ mit $v = g(A)w$. Wieder ist $Av = Ag(A)w = g(A)Aw$, und das ist ein Element von $\mathrm{Bild}(g(A))$. □

Korollar VI.4.8 (Zerlegungslemma für Endomorphismen): *Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $\phi\colon V \to V$ eine lineare Abbildung und $f = g_1g_2$ ein Polynom über K mit $f \in I(\phi)$ und teilerfremden g_1, g_2. Dann gilt*

$$V = \mathrm{Kern}(g_1(\phi)) \oplus \mathrm{Kern}(g_2(\phi))$$

sowie $\mathrm{Kern}(g_1(\phi)) = \mathrm{Bild}(g_2(\phi))$ *und* $\mathrm{Kern}(g_2(\phi)) = \mathrm{Bild}(g_1(\phi))$. *Hierbei sind* $\mathrm{Kern}(g_1(\phi))$ *und* $\mathrm{Kern}(g_2(\phi))$ *zwei ϕ-invariante Unterräume von V.*

Korollar VI.4.9: *Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $A \in K^{n\times n}$ eine Matrix und $\phi\colon V \to V$ ein Endomorphismus. Zerfällt χ_A beziehungsweise χ_ϕ in Linearfaktoren, d. h. $\chi_A = \prod_{i=1}^{r}(X-\lambda_i)^{e_i}$ beziehungsweise $\chi_\phi = \prod_{i=1}^{r}(X-\lambda_i)^{e_i}$, dann gilt*

$$K^n = \bigoplus_{i=1}^{r} \operatorname{Kern}\left((A-\lambda_i I_n)^{e_i}\right) \quad \textit{bzw.} \quad V = \bigoplus_{i=1}^{r} \operatorname{Kern}\left((\phi-\lambda_i \operatorname{id}_V)^{e_i}\right).$$

Der Raum $\operatorname{Kern}((A-\lambda_i I_n)^{e_i})$ *beziehungsweise* $\operatorname{Kern}((\phi-\lambda_i \operatorname{id}_V)^{e_i})$ *heißt* Hauptraum zum Eigenwert λ_i.

5. Nilpotente Endomorphismen

Erinnerung VI.5.1: Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $\phi\colon V \to V$ ein Endomorphismus und $A \in K^{n\times n}$ eine Matrix. Der Endomorphismus ϕ beziehungsweise die Matrix A heißt nilpotent, falls es einen natürlichen Exponenten k gibt, sodass $\phi^k = \mathbf{0}$ beziehungsweise $A^k = \mathbf{0}$. Insbesondere wird ϕ beziehungsweise A annulliert vom Polynom $f = X^k$.

Bemerkung VI.5.2: Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $\phi\colon V \to V$ ein nilpotenter Endomorphismus und $A \in K^{n\times n}$ eine nilpotente Matrix.

(i) Es gibt einen natürlichen Exponenten k, sodass $m_\phi = X^k$ beziehungsweise $m_A = X^k$.

(ii) Es ist $\operatorname{Spec}(\phi) = \{0\}$ beziehungsweise $\operatorname{Spec}(A) = \{0\}$.

(iii) Es ist $\chi_\phi = X^n$ beziehungsweise $\chi_A = X^n$.

Definition VI.5.3 (Zyklischer Unterraum): Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n, $\phi\colon V \to V$ ein Endomorphismus und $U \subseteq V$ ein Untervektorraum. Ist U ein ϕ-invarianter Unterraum und gibt es $u_0 \in U$ sowie eine natürliche Zahl k, sodass $U = \operatorname{Lin}(u_0, \phi(u_0), \ldots, \phi^k(u_0))$, dann heißt U ein *ϕ-zyklischer Unterraum.*

Proposition VI.5.4: *Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler K-Vektorraum und $\phi\colon V \to V$ ein nilpotenter Endomorphismus mit Minimalpolynom von Grad $d := \deg(m_\phi)$. Dann hat V einen d-dimensionalen ϕ-zyklischen Unterraum.*

Beweis: Da ϕ^{d-1} noch nicht die Nullabbildung ist, gibt es einen Vektor $u_0 \in V$ mit $\phi^{d-1}(u_0) \neq \mathbf{0}$. Setze $U := \operatorname{Lin}(u_0, \phi(u_0), \ldots, \phi^{d-1}(u_0))$. Wegen $\phi^d(u_0) = \mathbf{0}$ ist U ein ϕ-invarianter Unterraum, außerdem ist offensichtlich $\dim U \leq d$.

Sei $f := \chi_{\phi|_U}$. Weil $\phi|_U$ nilpotent ist, ist $f = X^{\dim U}$. Ferner liefert der Satz von Cayley-Hamilton, dass $f(\phi|_U) = \mathbf{0}$. Weil aber $\phi^{d-1}(u_0) \neq \mathbf{0}$ ist, ist auch $(\phi|_U)^{d-1}$ nicht die Nullabbildung, was $\dim U \geq d$ liefert. Wir erhalten also $\dim U = d$ wie gewünscht. □

Bemerkung VI.5.5: In der in Proposition VI.5.4 beschriebenen Situation ist $B = (u_0, \phi(u_0), \ldots, \phi^{d-1}(u_0))$ eine Basis von U und die Darstellungsmatrix von $\phi|_U$ bezüglich B ist

$$J := D_{B,B}(\phi|_U) = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} =: J_d.$$

Proposition VI.5.6: *Sei J_d die Matrix aus Proposition VI.5.5, d. h. $J_d = \sum_{i=2}^{d} E_{i,i-1}$. Dann gilt $J_d^\ell = \sum_{i=\ell+1}^{d} E_{i,i-\ell}$, d. h. J_d^ℓ hat Einsen auf der ℓ-ten unteren Nebendiagonalen und sonst Nullen. Insbesondere gilt:*

$$\operatorname{Rang}(J_d^\ell) = \begin{cases} d - \ell, & \text{falls } d - \ell \geq 0, \\ 0, & \text{sonst.} \end{cases}$$

Beweis: Wir zeigen die Behauptung per vollständiger Induktion nach ℓ. Für $\ell = 1$ stimmt die Behauptung per Definition der Matrix J_d.

Die Aussage gelte nun für ℓ. Dann haben wir

$$J_d^{\ell+1} = \Big(\sum_{i=\ell+1}^{d} E_{i,i-1} \Big) \Big(\sum_{i=2}^{d} E_{i,i-1} \Big) = \sum_{i=\ell+1}^{d} \sum_{j=2}^{d} E_{i,i-\ell} E_{j,j-1} = \sum_{i=\ell+2}^{d} E_{i,i-\ell-1}$$

wie gewünscht. Hierbei haben wir verwendet, dass $E_{i,i-1} E_{j,j-1} = \mathbf{0}$ genau dann, wenn $i - \ell \neq j$. □

Definition VI.5.7 (Invariantes Komplement): Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler Vektorraum über K, $\phi\colon V \to V$ ein Endomorphismus und $U \subseteq V$ ein ϕ-invarianter Unterraum von V. Ist W ein weiterer ϕ-invarianter Unterraum von V mit $V = U \oplus W$, dann heißt W ein *ϕ-invariantes Komplement zu U*.

Proposition VI.5.8: *Seien K ein Körper, n eine natürliche Zahl, V ein Vektorraum über K der Dimension n, $\phi\colon V \to V$ ein nilpotenter Endomorphismus, d der Grad des Minimalpolynoms von ϕ und $u_0 \in V$ mit $\phi^{d-1}(u_0) \neq \mathbf{0}$. Dann hat der ϕ-zyklische Unterraum $U = \mathrm{Lin}(u_0, \phi(u_0), \ldots, \phi^{d-1}(u_0))$ ein ϕ-invariantes Komplement.*

Beweis: Wir schreiben

$$\mathfrak{M} := \{W' \mid W' \subseteq V \text{ ist } \phi\text{-invarianter Unterraum mit } W' \cap U = \{\mathbf{0}\}\}$$

und halten fest, dass $\mathfrak{M}$ wegen $\{\mathbf{0}\} \in \mathfrak{M}$ nichtleer ist. Wir wählen ein Element W von $\mathfrak{M}$ mit maximaler Dimension.

Wir wollen zeigen, dass W das gesuchte ϕ-invariante Komplement ist, d. h. dass $V = U \oplus W$. Das erreichen wir, indem wir „$U \oplus W \subsetneq V$“ zu einem Widerspruch führen.

Angenommen, $U \oplus W$ wäre ein echter Unterraum von V. Dann fänden wir ein $v' \in V - (U \oplus W)$. Wegen $\phi^d = \mathbf{0}$ wäre insbesondere $\phi^d(v') = \mathbf{0}$, was zu $U \oplus W$ gehörte. Wegen $v \notin U \oplus W$ hätten wir außerdem $\phi^0(v) = v \notin U \oplus W$. Sei also nun $\ell \in \{1, \ldots, d\}$ minimal mit der Eigenschaft, dass $\phi^\ell(v') \in U \oplus W$. Wir schreiben $v := \phi^{\ell-1}(v')$.

Wegen $\phi(v) \in U \oplus W$ fänden wir eindeutige $u \in U$ und $w \in W$, sodass $\phi(v) = u' + w$ und da wir eine Basis von U kennen, könnten wir dieses u' mit geeigneten $\alpha_0, \ldots, \alpha_{d-1} \in K$ schreiben als $u' = \sum_{i=0}^{d-1} \alpha_i \phi^i(u_0)$. Anwendung von ϕ^{d-1} auf $\phi(v)$ ergäbe

$$\mathbf{0} = \alpha_0 \phi^{d-1}(u_0) + \phi^{d-1}(w),$$

was wegen $U \cap W = \{\mathbf{0}\}$ zunächst erzwänge, dass $\alpha_0 \phi^{d-1}(u_0) = \mathbf{0}$ sowie $\phi^{d-1}(w) = \mathbf{0}$, und schließlich wegen $\phi^{d-1}(u_0) \neq \mathbf{0}$, dass $\alpha_0 = 0$ gelten müsste. Mit $u := \sum_{i=1}^{d-1} \alpha_i \phi^{i-1}(u_0)$ erhielten wir so, dass $\phi(v) = \phi(u) + w$.

Das aber erlaubte es uns, w zu schreiben als $w = \phi(v) - \phi(u) = \phi(v - u)$, d. h. $\phi(v - u)$ gehörte zu W. Außerdem gehörte $v - u$ nicht zu W, denn andernfalls müsste schon v zu $U \oplus W$ gehört haben. Wir könnten also W echt vergrößern zu $W + \mathrm{Lin}(v - u)$ (d. h. $W \subsetneq W + \mathrm{Lin}(v - u)$).

Wäre jetzt w' ein Element von $(W + \mathrm{Lin}(v - u)) \cap U$, dann gäbe es $w \in W$ und ein $c \in K$, sodass $w' = w + c(v - u)$ ein Element von U wäre. Wäre $c \neq 0$, dann erhielten wir, dass $v = c^{-1}(w' + cu - w)$ zu $U \oplus W$ gehörte, was wir ausgeschlossen haben. Es müsste also $c = 0$ gelten, sodass $w' = w$ in Wahrheit zu W gehörte, was wegen $W \cap U = \{\mathbf{0}\}$ zur Folge hätte, dass $w' = \mathbf{0}$.

Insgesamt hätten wir ein Element W' von $\mathfrak{M}$ enthalten, das W echt enthielte, was der Wahl von W widerspräche. Es muss also $V = U \oplus W$ gelten und W ist das gewünschte ϕ-invariante Komplement von U. □

Korollar VI.5.9: *Seien K ein Körper, n eine natürliche Zahl, V ein K-Vektorraum der Dimension n und $\phi\colon V \to V$ ein nilpotenter Endomorphismus. Dann ist V die direkte Summe ϕ-zyklischer Unterräume.*

Die Aussage zeigt man per vollständiger Induktion mit der Aussage aus Proposition VI.5.8.

Korollar VI.5.10: *Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler Vektorraum über K und $\phi\colon V \to V$ ein nilpotenter Endomorphismus. Dann gibt es eine Basis B von V, sodass*

$$D_{B,B}(\phi) = \begin{pmatrix} J_{d_1} & & \\ & \ddots & \\ & & J_{d_s} \end{pmatrix}$$

wobei $d_1 \geq d_2 \geq \cdots \geq d_s \geq 1$. Die Matrizen J_d heißen Jordan-Kästchen.

Proposition VI.5.11 (Kenngrößen für nilpotente Endomorphismen): *Seien K ein Körper, n eine natürliche Zahl, V ein n-dimensionaler Vektorraum über K und $\phi\colon V \to V$ ein nilpotenter Endomorphismus. Für die Darstellungsmatrix $J = D_{B,B}(\phi)$ aus Proposition VI.5.10 gilt:*

(i) *Die Summe der $d_1, \ldots, d_s$ ist n,*

(ii) *Seien m_k die Anzahl der Jordan-Kästchen der Länge k und $r_k := \operatorname{Rang}(\phi^k)$. Dann gilt für alle k:*

$$m_k = r_{k-1} - 2r_k + r_{k+1}.$$

(iii) *Das größte Jordan-Kästchen hat die Größe* $\deg(m_\phi)$.

(iv) *Die Anzahl der Jordan-Kästchen ist* $\dim \operatorname{Eig}(\phi, 0)$.

Beweis: (i) Das ist klar.

(ii) Für das Jordan-Kästchen J_{d_i} haben wir uns bereits überlegt, dass $\operatorname{Rang}(J_{d_i}^k) = \max\{0, d_i - k\}$. Für r_k und alle nichtnegativen ganzen Zahlen k erhalten wir also die Charakterisierung

$$r_k = \operatorname{Rang}(\phi^k) = \operatorname{Rang}(J_k) = \sum_{i=1}^{s} \operatorname{Rang}(J_{d_i}^k) = \sum_{d=k+1}^{n} m_d(d-k).$$

Damit gilt für alle natürlichen Zahlen k, dass

$$r_{k-1} - r_k = \sum_{d=k}^{n} m_d(d-(k-1)) - \sum_{d=k+1}^{n} m_d(d-k) = m_k + \sum_{d=k+1}^{n} m_d = \sum_{d=k}^{n} m_d, \tag{VI.1}$$

d. h. für alle natürlichen Zahlen k erhalten wir

$$m_k = \sum_{d=k}^{n} m_d - \sum_{d=k+1}^{n} m_d = r_{k-1} - r_k - (r_k - r_{k+1}) = r_{k-1} - 2r_k + r_{k+1}$$

wie gewünscht.

(iii) Sei nun $d := \deg(m_\phi)$, d. h. $\phi^d = \mathbf{0}$. Aus (ii) folgern wir für alle $k \geq d+1$, dass $m_k = 0$ und dass $m_d = r_{d-1} = \mathrm{Rang}(\phi^{d-1}) > 0$.

(iv) Wegen Gl. (VI.1) ist

$$\begin{aligned}\sum_{d=1}^{n} m_d &= r_0 - r_1 \\ &= \mathrm{Rang}(\phi^0) - \mathrm{Rang}(\phi) \\ &= n - (n - \dim(\mathrm{Kern}(\phi)) = \dim \mathrm{Kern}(\phi) = \dim \mathrm{Eig}(\phi, 0),\end{aligned}$$

was wir zeigen wollten. □

6. Jordansche Normalform

Wir erinnern daran, dass ein Endomorphismus ϕ eines endlichdimensionalen K-Vektorraums V diagonalisierbar ist genau dann, wenn V in die direkte Summe der Eigenräume zerfällt, d. h. wenn $V = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} \mathrm{Eig}(\phi, \lambda)$.

Jeder dieser Eigenräume $\mathrm{Eig}(\phi, \lambda)$ ist ϕ-invariant. Für eine (allgemeinere) ϕ-invariante Zerlegung $V = \bigoplus_{i=1}^{r} H_i$ brauchen wir jeweils für jeden H_i ein ϕ-invariantes Komplement.

Beispiel VI.6.1: Seien K ein Körper,

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

und $\phi\colon K^3 \to K^3$, $x \mapsto Ax$ die zugehörige lineare Abbildung. Das charakteristische Polynom von ϕ ist $\chi_\phi = X^3$, $\mathrm{Spec}(\phi) = \{0\}$ und $\mathrm{Eig}(\phi, 0) = \mathrm{Lin}(e_3)$. Für irgendein $x = (x_1, x_2, x_3)^t$ in K^3 gilt

$$\phi(x) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ x_1 \\ x_2 \end{pmatrix}, \qquad \phi^2(x) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ x_1 \end{pmatrix},$$

d. h. $\phi^2(x) \in \mathrm{Eig}(\phi, 0)$, sodass $\mathrm{Eig}(\phi, 0)$ kein ϕ-invariantes Komplement haben kann.

Im Folgenden wollen wir zu einem Eigenraum $\mathrm{Eig}(\phi, \lambda)$ den kleinsten Untervektorraum H_λ finden, der $\mathrm{Eig}(\phi, \lambda)$ enthält, der selbst ϕ-invariant ist und der ein ϕ-invariantes Komplement besitzt.

Proposition VI.6.2: *Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, ϕ ein Endomorphismus von V, $U, W \subseteq V$ ϕ-invariante Untervektorräume mit $V = U \oplus W$ und $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V) = \mathrm{Eig}(\phi, \lambda) \subseteq U$. Dann gilt bereits für alle natürlichen Zahlen k, dass*

$$\mathrm{Kern}\left((\phi - \lambda \,\mathrm{id}_V)^k\right) \subseteq U.$$

Als Vorbereitung für den Beweis dieser Aussage machen wir eine Bemerkung.

Bemerkung VI.6.3: Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, ϕ ein Endomorphismus von V und $U \subseteq V$ ein Untervektorraum. Der Unterraum U ist ϕ-invariant genau dann, wenn U invariant unter $(\phi - \lambda \,\mathrm{id}_V)$ ist.

Beweis: „$\Longrightarrow$“: Seien U ein ϕ-invarianter Unterraum von V und $v \in U$. Dann ist auch $(\phi - \lambda \,\mathrm{id}_V)(v) = \phi(v) - \lambda v$ ein Element von U, wie gewünscht. Die Implikation „$\Longleftarrow$“ zeigt man genauso. □

Beweis: Wir zeigen die Aussage per vollständiger Induktion auf k. Für $k = 1$ gilt die Aussage per Voraussetzung.

Die Aussage gelte nun für eine natürliche Zahl k. Sei v ein Element von $\mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^{k+1} \subseteq V$. Als Element von $V = U \oplus W$ lässt sich v auf eindeutige Weise schreiben als $v = u + w$ mit $u \in U$, $w \in W$. Wir haben also

$$\mathbf{0} = (\phi - \lambda \,\mathrm{id}_V)^{k+1}(v) = (\phi - \lambda \,\mathrm{id}_V)^{k+1}(u) + (\phi - \lambda \,\mathrm{id}_V)^{k+1}(w)$$

mit $(\phi - \lambda \,\mathrm{id}_V)^{k+1}(u) \in U$ und $(\phi - \lambda \,\mathrm{id}_V)^{k+1}(w) \in W$. Da die Summe direkt ist, muss $(\phi - \lambda \,\mathrm{id}_V)^{k+1}(w) = \mathbf{0}$ gelten, sodass $(\phi - \lambda \,\mathrm{id}_V)^k(w)$ im Kern von $\phi - \lambda \,\mathrm{id}_V$ gehört, der in U liegt. Da aber auch $(\phi - \lambda \,\mathrm{id}_V)^k(w)$ zu W gehört, muss auch $(\phi - \lambda \,\mathrm{id}_V)^k(w) = \mathbf{0}$ gelten. Dann aber ist $w \in \mathrm{Kern}(\phi - \lambda \,\mathrm{id}_V)^k \subseteq U$ nach Induktionsvoraussetzung, d. h. $w = \mathbf{0}$ wegen der Direktheit der Summe $V = U \oplus W$. Damit gehört v zu U. □

Bemerkung VI.6.4: Die Kette

$$\mathrm{Kern}\left(\phi - \lambda \,\mathrm{id}_V\right) \subseteq \mathrm{Kern}\left((\phi - \lambda \,\mathrm{id}_V)^2\right) \subseteq \mathrm{Kern}\left((\phi - \lambda \,\mathrm{id}_V)^3\right) \subseteq \ldots$$

wird aus Dimensionsgründen stationär, d. h. es gibt eine natürliche Zahl k, sodass für alle natürlichen Zahlen $\ell \geq k$ gilt:

$$\mathrm{Kern}((\phi - \lambda \,\mathrm{id}_V)^k) = \mathrm{Kern}((\phi - \lambda \,\mathrm{id}_V)^\ell).$$

Definition VI.6.5: Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, ϕ ein Endomorphismus von V und λ ein Element von K. Dann heißt

$$H_\lambda := H(\phi, \lambda) := \bigcup_{k=1}^{\infty} \mathrm{Kern}((\phi - \lambda \,\mathrm{id}_V)^k)$$

der *Hauptraum von ϕ zu λ*.

Definition VI.6.6 (Algebraische Vielfachheit): Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, ϕ ein Endomorphismus von V und λ ein Element von K. Dann heißt

$$\mu_a(\phi, \lambda) := \max\{e \in \mathbb{N}_0 \mid (X - \lambda)^e \mid \chi_\phi\}$$

die *algebraische Vielfachheit von λ für ϕ*. Ferner heißt

$$\delta := \max\{e \in \mathbb{N}_0 \mid (X - \lambda)^e \mid m_\phi\}$$

heißt *Exponent im Minimalpolynom m_ϕ von $(X - \lambda)$*.

Proposition VI.6.7: *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, ϕ ein Endomorphismus von V und λ ein Element von K.*

(i) *Sei e der Exponent von $(X - \lambda)$ im Minimalypolynom von ϕ. Dann ist*

$$H_\lambda = H(\phi, \lambda) = \mathrm{Kern}\left((\phi - \lambda \,\mathrm{id}_V)^e\right).$$

(ii) *Der Hauptraum $H(\phi, \lambda)$ ist ϕ-invariant und enthält* $\mathrm{Eig}(\phi, \lambda)$.

(iii) *Der Hauptraum $H(\phi, \lambda)$ hat ein ϕ-invariantes Komplement.*

(iv) *Der Hauptraum $H(\phi, \lambda)$ ist der kleinste Unterraum von V, der* (ii) *und* (iii) *erfüllt.*

(v) *Die Dimension des Hauptraums $H(\phi, \lambda)$ ist $\mu_a(\phi, \lambda)$.*

Beweis: Wir schreiben das charakteristische Polynom $\chi_\phi = (X-\lambda)^e g$ mit einem Polynom $g \in K[X]$. Nach (Korollar I.3.19) sind $(X - \lambda)^e$ und g teilerfremd, außerdem ist ihr Produkt (das charakteristische Polynom) ein Element des Verschwindungsideals von ϕ. Für $W := \mathrm{Kern}(g(\phi))$ und $H := \mathrm{Kern}((\phi-\lambda \,\mathrm{id}_V)^e)$ gibt (Korollar I.4.8), dass $V = W \oplus H$ mit ϕ-invarianten Unterräumen W und H.

(i) Nach Definition des Hauptraums $H(\phi,\lambda)$ ist die Inklusion „$\supseteq$“ klar. Für „$\subseteq$“ beachten wir, dass $\mathrm{Eig}(\phi,\lambda)\subseteq H$. Nach Proposition VI.6.2 enthält H also H_λ.

(ii) Direkte Konsequenz aus (i).

(iii) Direkte Konsequenz aus $V=W\oplus H$, wie in (i) gezeigt.

(iv) Das folgt aus Proposition VI.6.2.

(v) Wir haben das charakteristische Polynom χ_ϕ geschrieben als Produkt teilerfremder Polynome $\chi_\phi=(X-\lambda)^e g$, andererseits erhalten wir aus der Zerlegung $V=U\oplus H$, dass $\chi_\phi=\chi_{\phi|_W}\chi_{\phi|_H}$. Wegen $H=\mathrm{Kern}((\phi-\lambda\,\mathrm{id}_V)^e)$ wird $\phi|_H$ annulliert von $(X-\lambda)^e$. Damit ist das Minimalpolynom ein Teiler von $(X-\lambda)^e$ und $\chi_{\phi|_H}=(X-\lambda)^n$ für eine geeignete natürliche Zahl n.

Da g und $(X-\lambda)^e$ teilerfremd ist, ist $(X-\lambda)$ kein Teiler von g und außerdem ist $(X-\lambda)$ auch kein Teiler von $\chi_{\phi|_W}$, da $\mathrm{Eig}(\phi,\lambda)$ ganz in H enthalten ist und somit λ nicht zum Spektrum $\mathrm{Spec}(\phi|_W)$ gehören kann. Aus der Gleichheit der Produktdarstellungen für das charakteristische Polynom und der Eindeutigen Primfaktorzerlegung in $K[X]$ gilt darum $n=e$. □

Proposition VI.6.8 (Summe der Haupträume ist direkt): *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, ϕ ein Endomorphismus über K und $\{\lambda_1,\dots,\lambda_k\}=\mathrm{Spec}(\phi)$. Dann gilt:*

$$\sum_{i=1}^{k} H(\phi,\lambda_i)=\bigoplus_{i=1}^{k} H(\phi,\lambda_i).$$

Beweis: Wir zeigen die Aussage per Induktion nach k. Für $k=0$ und $k=1$ ist nichts zu zeigen. Die Aussage gelte nun für eine natürliche Zahl k und sei $\mathbf{0}=v_1+\cdots+v_k$ mit $v_i\in H(\phi,\lambda_i)$. Wähle eine natürliche Zahl e mit $(\phi-\lambda_k\,\mathrm{id})^e v_k=\mathbf{0}$. Dann erhalten wir

$$\mathbf{0}=\sum_{i=1}^{k}(\phi-\lambda_k\,\mathrm{id}_V)^e(v_i)=\sum_{i=1}^{k-1}(\phi-\lambda_k\,\mathrm{id}_V)^e(v_i).$$

Weil jeder $H(\phi,\lambda_i)$ ein ϕ-invarianter Unterraum von V ist, ist nach (Bemerkung I.6.3) jeder der $H(\phi,\lambda_i)$ auch $(\phi-\lambda_k\,\mathrm{id}_V)$-invariant, d. h. $(\phi-\lambda_k\,\mathrm{id}_V)(v_i)$ gehört zu $H(\phi,\lambda_i)$. Per Induktionsvoraussetzung ist deshalb $(\phi-\lambda\,\mathrm{id}_V)(v_i)=\mathbf{0}$ für $1\le i\le k-1$. Da für $1\le i\le k-1$ jedes v_i zu $H(\phi,\lambda_i)$ gehört, gibt es Exponenten f_i mit $(\phi-\lambda_i\,\mathrm{id}_V)^{f_i}(v_i)=0$. Die Polynome $(X-\lambda_k)^e$ und

$(X-\lambda_i)^{f_i}$ sind teilerfremd, sodass das Lemma von Bézout die Existenz von Polynomen g_i und h_i liefert, die

$$1 = g_i(X-\lambda_k)^e + h_i(X-\lambda_i)^{f_i}$$

leisten. Damit ist $\mathrm{id}_V = g_i(\phi)(\phi-\lambda_k)^e + h_i(\phi)(\phi-\lambda_i\,\mathrm{id}_V)^{f_i}$, d. h.

$$v_i = \mathrm{id}_V(v_i) = (g_i(\phi)(\phi-\lambda_k\,\mathrm{id}_V)^e(v_i) + h(\phi)(\phi-\lambda_i)^{f_i}(v_i) = \mathbf{0},$$

d. h. $v_i = \mathbf{0}$ für $1 \leq i \leq k-1$. Aber dann muss auch $v_k = \mathbf{0}$ gewesen sein und wir sind fertig. □

Korollar VI.6.9 (aus VI.6.7): *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, ϕ ein Endomorphismus von V und $\lambda \in K$. Ferner sei δ_λ der Exponent von $(X-\lambda)$ im Minimalpolynom m_ϕ*

(i) *Es ist*

$$\begin{aligned}\delta_\lambda &= \min\{e \in \mathbb{N} \mid H(\phi,\lambda) = \mathrm{Kern}\left((\phi-\lambda\,\mathrm{id}_V)^e\right)\}\\ &= \min\{e \in \mathbb{N} \mid \mathrm{Kern}(\phi-\lambda\,\mathrm{id}_V)^e = \mathrm{Kern}(\phi-\lambda\,\mathrm{id}_V)^{e+1}\}.\end{aligned}$$

(ii) *Sei H_λ der Hauptraum $H(\phi,\lambda)$ und betrachte $\phi|_{H_\lambda}$. Dann ist $m_{\phi|_{H_\lambda}} = (X-\lambda)^{\delta_\lambda}$ und $\chi_{\phi|_{H_\lambda}} = (X-\lambda)^{\alpha_\lambda}$, wobei α_λ die algebraische Vielfachheit von λ, d. h. die Dimension von H_λ ist.*

Bemerkung VI.6.10 (Invariante Komplemente von Haupträumen): Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, ϕ ein Endomorphismus von V und $\lambda \in K$. Des Weiteren sei $V = H(\phi,\lambda) \oplus W$ eine Zerlegung von V in ϕ-invariante Unterräume. Dann haben wir:

(i) Der Eigenwert λ gehört nicht zu $\mathrm{Spec}(\phi|_W)$.

(ii) Ist $\lambda' \in K$ mit $\lambda' \neq \lambda$, dann ist $H(\phi,\lambda') \subseteq W$.

Beweis: (i) Das ist klar, da $\mathrm{Eig}(\phi,\lambda)$ per Definition des Hauptraums in $H(\phi,\lambda)$ enthalten ist.

(ii) Wir wollen zeigen, dass $\mathrm{Eig}(\phi,\lambda')$ ein Teilraum von W ist. Dann folgt die Behauptung aus VI.6.7(iv). Sei dazu $v \in \mathrm{Eig}(\phi,\lambda')$, d. h. $\phi(v) = \lambda' v$, und schreibe $v = u + w$ mit $u \in H(\phi,\lambda)$ und $w \in W$. Anwendung von ϕ liefert

$$\phi(u) + \phi(w) = \phi(v) = \lambda' v = \lambda' u + \lambda' w$$

wobei sowohl $\phi(u) \in H(\phi,\lambda)$ als auch $\phi(w) \in W$, sodass wegen der Direktheit der Summe folgt, dass $\lambda' u = \phi(u)$ und $\lambda' w = \phi(w)$. Wegen $H_\lambda \cap H_{\lambda'} = \{\mathbf{0}\}$ muss $u = \mathbf{0}$ folgen, d. h. $v = w$ und somit $\mathrm{Eig}(\phi,\lambda') \subseteq W$, wie gewünscht. □

Seien K ein Körper und f ein Polynom über K. Gibt es $a, \lambda_1, \ldots, \lambda_n \in K$ mit $f = a(X - \lambda_1) \cdots (X - \lambda_n)$, dann sagen wir, f zerfalle in Linearfaktoren.

Proposition VI.6.11: *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K und ϕ ein Endomorphismus von V. Dann sind äquivalent:*

(i) *Der Vektorraum V zerfällt in die direkte Summe der Haupträume, d. h. $V = \bigoplus_{\lambda \in \operatorname{Spec}(\phi)} H(\phi, \lambda)$.*

(ii) *Das charakteristische Polynom χ_ϕ zerfällt in Linearfaktoren.*

(iii) *Das Minimalpolynom m_ϕ zerfällt in Linearfaktoren.*

Beweis: Für $\lambda \in K$ schreiben wir $H_\lambda := H(\phi, \lambda)$.

„(i) $\Longrightarrow$ (ii)“: Aus Proposition VI.6.9 wissen wir, dass $\chi_{\phi|_{H_\lambda}} = (X - \lambda)^{\dim H_\lambda}$. Nach Voraussetzung und weil die Haupträume ϕ-invariant sind, erhalten wir $\chi = \prod_{\lambda \in \operatorname{Spec}(\phi)} \chi_{\phi|_{H_\lambda}}$.

„(ii) $\Longrightarrow$ (iii)“: Weil das Minimalpolynom das charakteristische Polynom teilt, liefert (Proposition I.3.17), dass auch m_ϕ in Linearfaktoren zerfällt.

„(iii) $\Longrightarrow$ (i)“: Per (Propisition I.4.4) entspricht die Nullstellenmenge von m_ϕ genau dem Spektrum von ϕ, sagen wir $\operatorname{Spec}(\phi) = \{\lambda_1, \ldots, \lambda_k\}$. Nun zeigen wir die Aussage durch vollständige Induktion über $k = \#\operatorname{Spec}(\phi)$.

Für $k = 0$ ist wegen $\deg(m_\phi)$ und damit $V = \{\mathbf{0}\}$ nichts zu tun. Für $k = 1$ gibt es $\lambda \in K$ und einen natürlichen Exponenten δ, sodass $m_\phi = (X - \lambda)^\delta$. Per Definition ist $(\phi - \lambda \operatorname{id}_V)^\delta = \mathbf{0}$, d. h. $V = \operatorname{Kern}(\phi - \lambda \operatorname{id}_V)^\delta = H_\lambda$.

Die Aussage gelte jetzt für $k - 1$. Wir wollen zeigen, dass die Aussage dann auch für k gilt. Hat das Spektrum Mächtigkeit k, dann können wir schreiben

$$m_\phi = \Big(\prod_{i=1}^{k-1} (X - \lambda_i)^{\delta_i} \Big) (X - \lambda_k)^{\delta_k}.$$

Schreiben wir $f := \prod_{i=1}^{k-1} (X - \lambda_i)^{\delta_i}$, dann liefert das Zerlegungslemma, dass V zerfällt in die direkte Summe

$$V = \operatorname{Kern}(\phi - \lambda_k \operatorname{id}_V)^{\delta_k} \oplus \operatorname{Kern}(f(\phi)) =: U \oplus W.$$

Aus Proposition VI.6.9 und Proposition VI.6.10 erhalten wir $\operatorname{Spec}(\phi|_W) = \{\lambda_1, \ldots, \lambda_{k-1}\}$. Nach Induktionsvoraussetzung ist $W = \bigoplus_{i=1}^{k-1} H(\phi|_U, \lambda_i) = \bigoplus_{i=1}^{k-1} H(\phi, \lambda_i)$, was wegen $V = U \oplus W$ die Behauptung liefert. □

Bemerkung VI.6.12: (i) Sind die Aussagen in Proposition VI.6.11 erfüllt, dann gilt insbesondere: Bezeichnet α_λ die algebraische Vielfachheit des Eigenwerts λ (also auch $\alpha_\lambda = \dim H(\phi, \lambda)$), dann ist $\chi_\phi = \prod_{\lambda \in \operatorname{Spec}(\phi)} (X - \lambda)^{\alpha_\lambda}$. Ferner ist $m_\phi = \prod_{\lambda \in \operatorname{Spec}(\phi)} (X - \lambda)^{\delta_\lambda}$ mit $\delta_\lambda \leq \alpha_\lambda$.

(ii) Ist K algebraisch abgeschlossen, d. h. jedes Polynom nichtkonstante Polynom hat eine Nullstelle, dann gelten die Aussagen aus Proposition VI.6.11 für alle Endomorphismen endlichdimensionaler K-Vektorräume. Insbesondere gilt das für den Körper der komplexen Zahlen $\mathbb{C}$.

Bemerkung VI.6.13 (Nicht-Beispiel für Jordannormalform): (i) Seien $K = \mathbb{R}$ und $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto Ax$ mit $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Dann ist

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

sodass $A^2 + I_2 = \mathbf{0}$. Insbesondere ist dann auch $\phi^2 + \mathrm{id}_V = \mathbf{0}$, und weil es kein Polynom vom Grad 1 gibt, das ϕ annulliert, ist auch $m_\phi = X^2 + 1$. Weil m_ϕ keine Nullstellen hat, ist auch das Spektrum von ϕ leer und damit ist $\bigoplus_{\lambda \in \mathrm{Spec}(\phi)} H_\lambda = \{\mathbf{0}\} \subsetneq \mathbb{R}^2$.

(ii) Betrachten wir ϕ als lineare Abbildung über $K = \mathbb{C}$, dann ist $m_\phi = (X + \mathrm{i})(X - \mathrm{i})$, d. h. $\mathrm{Spec}(\phi) = \{\mathrm{i}, -\mathrm{i}\}$ und $H(\phi, \mathrm{i}) = \mathrm{Eig}(\phi, \mathrm{i}) = \mathrm{Lin}((\mathrm{i}, 1)^t)$ und $H(\phi, -\mathrm{i}) = \mathrm{Eig}(\phi, -\mathrm{i}) = \mathrm{Lin}((\mathrm{i}, -1)^t)$.

Bemerkung VI.6.14: Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K und ϕ ein Endomorphismus von V.

(i) Ist λ ein Eigenwert von ϕ, dann ist $\phi - \lambda\,\mathrm{id}_V$ nilpotent auf $H(\phi, \lambda)$.

(ii) Für jede Basis B von V gilt $D_{B,B}(\phi) = D_{B,B}(\phi - \lambda\,\mathrm{id}_V) + \lambda I_n$.

Definition VI.6.15 (Jordankästchen): Seien K ein Körper, $\lambda \in K$ und d eine natürliche Zahl. Dann heißt

$$J_d(\lambda) = \lambda I_d + J_d(0) = \dots$$

Jordankästchen der Größe d zum Eigenwert λ.

Satz 25 (über die Jordan-Normalform): *Seien K ein Körper, V ein K-Vektorraum der Dimension n und ϕ ein Endomorphismus von V mit* $\mathrm{Spec}(\phi) = \{\lambda_1, \dots, \lambda_\ell\}$. *Zerfällt das charakteristische Polynom in Linearfaktoren, dann gibt es eine Basis B von V mit*

$$D_{B,B}(\phi) = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_\ell \end{pmatrix},$$

wobei für $1 \le i \le \ell$ gilt: Es gibt natürliche Zahlen $d_{1,i}, \dots, d_{k_i,i}$ mit $d_{1,i} \ge \dots \ge d_{k_i,i}$, sodass D_i auf der Diagonalen die Jordankästchen $J_{d_{1,i}}(\lambda_i), \dots, J_{d_{k_i,i}}(\lambda_i)$ stehen hat.

Die Matrix D_i heißt Jordan-Block *zum Eigenwert λ_i und die Matrizen $J_{d_{j,i}}(\lambda_i)$ heißen* Jordan-Kästchen zum Eigenwert λ_i.

Beweis: Wir erhalten die Aussage aus Proposition VI.6.14, Proposition VI.6.11 und Proposition VI.5.10. □

Proposition VI.6.16 (Kenngrößen): *Mit der Notation aus Satz 25 gilt für* $\lambda \in \mathrm{Spec}(\phi)$*:*

(i) *Die Mächtigkeit von* $\mathrm{Spec}(\phi)$ *entspricht der Anzahl der Jordan-Blöcke. Die Basisvektoren in* B *zum Jordan-Block* D_i *bilden eine Basis* B_i *des Hauptraums* $H(\phi, \lambda_i)$ *mit* $D_{B_i,B_i}(\phi|_{H_{\lambda_i}}) = D_i$*:*

(ii) *Die Größe des Jordan-Blocks* D_i *ist die Dimension des Hauptraums* $H(\phi, \lambda_i)$*, welche der algebraischen Vielfachheit* $\mu_a(\phi, \lambda_i) = \alpha_\lambda$ *entspricht.*

(iii) *Die Anzahl der Jordan-Kästchen im Jordan-Block* D_i *entspricht der Dimension des Eigenraums* $\mathrm{Eig}(\phi, \lambda_i)$*. Diese nennt man auch* geometrische Vielfachheit $\mu_G(\phi, \lambda_i)$ von λ_i.

(iv) *Die Größe des größten Jordan-Kästchens im Jordan-Block* D_i *ist der Exponent* δ_{λ_i} *von* $(X - \lambda_i)$ *im Minimalpolynom* m_ϕ*.*

(v) *Seien* $m_d(\lambda)$ *die Anzahl der Jordan-Kästchen der Länge* d *im Jordan-Block* D_λ *und für* $k \in \mathbb{N}$ *bezeichne* $r_k = \mathrm{Rang}(\phi|_{H_\lambda} - \lambda \,\mathrm{id}\,|_{H_\lambda})^k$*. Dann gilt*

$$m_d(\lambda) = r_{d-1} - 2r_d + r_{d+1}.$$

Weiter gilt für alle $k \in \mathbb{N}_0$*, dass* $r_{k-1} - r_k = \sum_{d=k}^{n} m_d$*.*

Proposition VI.6.17 (Eindeutigkeit): *Die Jordan-Normalform in Satz 25 ist eindeutig bis auf Vertauschung der Jordan-Blöcke.*

Beweis: Die Kenngrößen aus Proposition VI.6.16 sind eindeutig durch ϕ gegeben und bestimmen die Matrix J vollständig. □

Korollar VI.6.18: *Seien* K *ein Körper,* $A \in K^{n\times n}$ *mit* $\mathrm{Spec}(A) = \{\lambda_1, \ldots, \lambda_k\}$ *und* χ_A *zerfalle in Linearfaktoren. Dann gibt es eine Matrix* $B \in \mathrm{Gl}_n(K)$*, sodass* $J(A) := BAB^{-1}$ *von der in Satz 25 beschriebenen Form ist. Die Jordan-Normalform ist (durch unsere Konventionen) eindeutig und es gelten die analogen Aussagen zu Proposition VI.6.16.*

Beispiel VI.6.19 (Jordan-Normalform): Betrachte

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

Seien $V = K^{12}$ und $\phi\colon V \to V$, $x \mapsto Jx$. Über diese Abbildung können wir fast alle wichtigen Informationen aus der Jordan-Normalform ablesen:

(i) $\operatorname{Spec}(\phi) = \{2, -3\}$,

(ii) $\chi_\phi = (X-2)^{10}(X+3)^2$,

(iii) $m_\phi = (X-2)^4(X+3)^2$,

(iv) $\dim(H(\phi, 2)) = 10$, $\dim(H(\phi, -3)) = 2$,

(v) $H(\phi, 2) = \langle e_1, e_2, \ldots, e_{10}\rangle$, $H(\phi, -3) = \langle e_{11}, e_{12}\rangle$,

(vi) $\dim(\operatorname{Eig}(\phi, 2)) = 3$, $\dim(\operatorname{Eig}(\phi, -3)) = 1$.

Genauer ist $\operatorname{Eig}(\phi, \lambda) = \langle e_4, e_8, e_{10}\rangle$, $\operatorname{Eig}(\phi, -3) = \langle e_{12}\rangle$. Insbesondere sind die algebraischen Vielfachheiten $\alpha_2 = \mu_a(\phi, 2) = 10$, $\alpha_{-3} = \mu_a(\phi, -3) = 2$, die geometrischen Vielfachheiten sind $\mu_g(\phi, 2) = \dim \operatorname{Eig}(\phi, 2) = 3$, $\mu_g(\phi, -3) = \dim(\operatorname{Eig}(\phi, -3)) = 1$ und die Exponenten im Minimalpolynom sind $\delta_2 = 4$ (die Größe des größten Jordan-Kästchens) sowie $\delta_{-3} = 2$.

Korollar VI.6.20 (Diagonalisierbarkeit): *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K, $\phi\colon V \to V$ ein Endomorphismus und $A \in K^{\dim V}$ eine Matrix. Ferner zerfalle χ_ϕ beziehungsweise χ_A in Linearfaktoren.*

(i) *Genau dann ist ϕ bzw. A diagonalisierbar, wenn das Minimalpolynom in paarweise verschiedene Linearfaktoren zerfällt, d. h. wenn für jeden Eigenwert λ der Exponent δ_λ im Minimalpolynom Eins ist.*

(ii) *Ist ϕ beziehungsweise A diagonalisierbar und ist U ein ϕ-invarianter bzw. A-invarianter Unterraum, dann ist $\phi|_U$ bzw. $A|_U$ diagonalisierbar.*

Beweis: Die erste Aussage ist klar. Die zweite Aussage zeigen wir für Endomorphismen, für Matrizen funktioniert derselbe Beweis. Seien $m := m_\phi$ und $m_1 := m_{\phi|_U}$. Dann gibt es ein Polynom h mit $m = hm_1$. Da m nach (i) in paarweise verschiedene Linearfaktoren zerfällt, zerfällt auch m_1 in paarweise verschiedene Linearfaktoren, sodass $\phi|_U$ auch diagonalisierbar ist. □

Proposition VI.6.21 (Jordan-Zerlegung): *Seien K ein Körper, $A \in K^{n\times n}$ eine Matrix und χ_A das charakteristische Polynom von A, das in Linearfaktoren zerfalle. Dann gibt es eine eindeutige Zerlegung $A = D + N$ mit Matrizen $D, N \in K^{n\times n}$, sodass gilt:*

(i) *D ist diagonalisierbar,*

(ii) *N ist nilpotent,*

(iii) *$DA = AD$ sowie $NA = AN$.*

Diese Zerlegung nennen wir Jordan-Zerlegung von A. *Andere in der Literatur gebräuchliche Namen sind* Jordan-Chevalley-Zerlegung *oder* Dunford-Zerlegung.

Beweis: *Existenz der Zerlegung*: Ist $A = J$ in Jordan-Normalform, dann seien $D = \operatorname{diag}(\alpha_{1,1}, \ldots, \alpha_{n,n})$ und $N := A - D$. Diese zwei Matrizen erfüllen die gewünschten Eigenschaften. Ist A nicht in Jordan-Normalform, dan gibt es eine Matrix $S \in \mathrm{Gl}_n(K)$, sodass $J := SAS^{-1}$ in Jordan-Normalform ist. Zerlege $J = D' + N'$ wie oben beschrieben und erhalte $A = S^{-1}N'S + S^{-1}D'S$. Setze $N := S^{-1}N'S$ und $D := S^{-1}D'S$. Die Matrix D ist diagonalisierbar, da sie konjugiert zu einer Diagonalmatrix ist, die Matrix N ist nilpotent, da N' nilpotent ist und $N^k = S^{-1}N'^kS$ gilt, und schließlich ist

$$DA = S^{-1}D'SS^{-1}A'S = S^{-1}D'A'S = S^{-1}A'D'S = S^{-1}A'SS^{-1}D'S = AD,$$

genau so erhält man $NA = AN$.

Eindeutigkeit der Zerlegung: Sei $A = D + N$ eine Zerlegung mit Matrizen D und N wie in der Aussage beschrieben. Weil D und N mit A kommutieren, ist insbesondere $(A - \lambda I_n)D = D(A - \lambda I_n)$. Insbesondere gilt für alle natürlichen Zahlen k, dass $(A - \lambda I_n)^k D = D(A - \lambda I_n)^k$.

Im nächsten Schritt überlegen wir uns, dass $H(A, \lambda)$ ein D-invarianter Unterraum ist. Sei dazu $v \in H(A, \lambda)$ gegeben, d. h. es gibt eine natürliche Zahl k, sodass $(A - \lambda I_n)^k v = \mathbf{0}$. Dann ist $(A - \lambda I_n)^k Dv = D(A - \lambda I_n)^k v = \mathbf{0}$, d. h. auch Dv gehört zu $\operatorname{Kern}(A - \lambda I_n)^k \subseteq H(A, \lambda)$.

Schreibe $\phi_D\colon D \to D$, $v \mapsto Av$. Nach Proposition VI.6.20(ii) ist $\phi_{D|_{H(A,\lambda)}}$ diagonalisierbar. Können wir zeigen, dass $\operatorname{Spec}(\phi_{D|_{H(A,\lambda)}})$ genau $\{\lambda\}$ ist, dann ist $\phi_{D|_{H(A,\lambda)}} = \lambda \operatorname{id}_{H(A,\lambda)}$, d. h. ϕ_D ist auf ganz K^n bestimmt und damit auch D sowie insbesondere $N = A - D$.

Sei nun $v \in H(A,\lambda)$ ein Eigenvektor von ϕ_D zum Eigenwert α. Dann ist $Av = (D+N)v = Dv + Nv = \alpha v + Nv$, sodass $(A - \alpha I_n)v = Nv$ und nach dem vorher gezeigten haben wir für alle natürlichen Zahlen k, dass $(A - \alpha I_n)^k v = N^k v$. Weil N nilpotent ist, gibt es einen Exponenten k, sodass $v \in \operatorname{Kern}(A - \alpha I_n)^k$, d. h. es gilt $v \in H(A,\alpha)$. Da sich Haupträume trivial schneiden, aber $v \neq \mathbf{0}$ ist, gilt $H(A,\alpha) = H(A,\lambda)$. □

Die Jordan-Normalform erlaubt es uns, unter allen Darstellungsmatrizen eines Endomorphismus eine besonders schöne bzw. einfache bzw. praktische Darstellungsmatrix auszuwählen, die alle wichtigen Daten des Endomorphismus kodieren, die nicht von der Basis abhängen. Insbesondere kann man die Jordan-Normalform verwenden, um einfacher Aussagen über Endomorphismen zu zeigen.

Da die Jordan-Normalform bis auf Reihenfolge der Blöcke eindeutig ist, ist die Jordan-Normalform eine Möglichkeit zu entscheiden, ob zwei Matrizen die Darstellungsmatrizen desselben Endomorphismus sind.

Mithilfe der Jordan-Normalform kann zum Beispiel entscheiden werden, ob zwei Matrizen A, B ähnlich sind (d. h. ob es $S \in \mathrm{Gl}_n(K)$ mit $A = S^{-1}BS$ gibt). Genauer: Sind K ein algebraisch abgeschlossener Körper und n eine natürliche Zahl, dann hat jede Ähnlichkeitsklasse in $K^{n\times n}$ einen Vertreter in Jordan-Normalform. Aussagen über Ähnlichkeitsinvarianten lassen sich besonders komfortabel über die Jordan-Normalform studieren.

Kapitel VII.

Multilineare Algebra - Teil 1

1. Multilineare Abbildungen

Definition VII.1.1 (Multilineare Abbildung): Seien K ein Körper und $V_1, \ldots, V_n$ und W Vektorräume über K. Ferner sei $M\colon V_1 \times \cdots \times V_n \to W$ eine Abbildung. Ist für jedes $i \in \{1, \ldots, n\}$ und beliebige $v_j \in V_j$ mit $j \in \{1, \ldots, n\} - \{i\}$ die Abbildung $V_i \mapsto W$, $v \mapsto M(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$ linear, dann heißt M eine *n-fach multilineare Abbildung* oder kurz *multilineare Abbildung*.

Eine 1-multilineare Abbildung eine lineare Abbildung, eine 2-mulitilineare Abbildung heißt *bilinear* und ist $W = K$, dann spricht man von *Multilinearformen*.

Bemerkung VII.1.2: Seien K ein Körper und V_1, V_2 und W Vektorräume über K. Genau dann ist eine Abbildung $\beta\colon V_1 \times V_2 \to W$ bilinear, wenn für alle $v_1, v_1' \in V$, $v_2, v_2' \in V_2$ und $\lambda \in K$ gilt:

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

Ist nur ein Argument einer multilinearen Funktion der Nullvektor, dann ist das Bild des entsprechenden Tupels unter der multilinearen Abbildung die Null im Bild.

Beispiel VII.1.3: (i) Für einen Körper K ist

$$\det\colon K^n \times \cdots \times K^n \longrightarrow K, \qquad (v_1, \ldots, v_n) \longmapsto \det(v_1|\ldots|v_n)$$

eine n-fache Multilinearform.

(ii) Die skalare Multiplikation $K \times V \to V$, (λ, v) ist eine bilineare Abbildung.

(iii) Auf K^3 ist das *Kreuzprodukt*

$$K^3 \times K^3 \longrightarrow K^3, \qquad \left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \right) \longmapsto \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

eine bilineare Abbildung.

(iv) Für zwei K-Vektorräume V und W ist die Einsetzungabbildung

$$\mathrm{Hom}(V, W) \times V \longrightarrow W, \qquad (\phi, v) \longmapsto \phi(v)$$

ist eine bilineare Abbildung.

(v) Seien p, q, r und s natürliche Zahlen. Dann ist

$$K^{p\times q} \times K^{q\times r} \times K^{r\times s} \longrightarrow K^{p\times s}, \qquad (A, B, C) \longmapsto ABC$$

eine dreifach multilineare Abbildung.

Definition VII.1.4 (Vertauschungseigenschaften): Seien K ein Körper, n eine natürliche Zahl und V und W Vektorräume über K. Ferner sei $M\colon V^n \to W$ eine multilineare Abbildung.

(i) Gilt für alle $\sigma \in S_n$ und $v_1, \ldots, v_n \in V$, dass

$$M(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = M(v_1, \ldots, v_n),$$

dann heißt M *symmetrisch*.

(ii) Gilt für alle $\sigma \in S_n$ und $v_1, \ldots, v_n \in V$, dass

$$M(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \mathrm{sgn}(\sigma) M(v_1, \ldots, v_n),$$

dann heißt M *schiefsymmetrisch*.

(iii) Gilt für alle $v_1, \ldots, v_n \in V$, für die es $i \neq j$ mit $v_i = v_j$ gibt, dass $M(v_1, \ldots, v_n) = \mathbf{0}$ gilt, dann heißt M *alternierend*.

Proposition VII.1.5 (Schiefsymmetrisch vs. alternierend): *Seien K ein Körper, V und W Vektorräume über K, n eine natürliche Zahl und $M\colon V^n \to W$ eine multilineare Abbildung.*

(i) *Ist M alternierend, dann ist M schiefsymmetrisch.*

(ii) *Ist M schiefsymmetrisch und ist* $\mathrm{char}(K) \neq 2$*, dann ist M auch alternierend.*

Beweis: (i) Sei M alternierend. Aus der Linearen Algebra I ist bekannt, dass die symmetrische Gruppe S_n von Transpositionen erzeugt wird, d. h. für jedes $\sigma \in S_n$ gibt es Transpositionen $\tau_1, \ldots, \tau_k$ mit $\sigma = \tau_1 \circ \cdots \circ \tau_k$ und $\operatorname{sgn}(\sigma) = \prod_{i=1}^{k} \operatorname{sgn}(\tau_i) = (-1)^k$. Es genügt also, Schiefsymmetrie für Transpositionen nachzuweisen.

Seien also $1 \leq i < j \leq n$ und $\tau = (ij)$. Für alle $v_1, \ldots, v_n$ aus V gilt:

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

sodass $M(v_1, \ldots, v_n) = (-1)M(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)$.

(ii) Seien M schiefsymmetrisch und $1 \leq i < j \leq n$. Für alle $v_1, \ldots, v_n$ in V ist

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

sodass $2M(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) = \mathbf{0}$, woraus wegen $\operatorname{char}(K) \neq 2$ die Behauptung folgt. □

Beispiel VII.1.6: (i) Die Determinante $\det\colon \prod_{i=1}^{n} K^n \to K$ ist alternierend und somit auch schiefsymmetrisch.

(ii) Das Kreuzprodukt $\times\colon K^3 \times K^3 \to K^3$ ist alternierend und schiefsymmetrisch.

(iii) Mit $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ wird der Körper mit zwei Elementen bezeichnet. Das Produkt

$$s\colon \mathbb{F}_2 \times \mathbb{F}_2 \longrightarrow \mathbb{F}_2, \qquad (a, b) \longmapsto ab$$

ist symmetrisch, da die Multiplikation auf $\mathbb{F}_2$ kommutativ ist. Außerdem ist das Produkt schiefsymmetrisch, weil $[-1] = [1]$, jedoch nicht alternierend, da $s(1, 1) = 1 \neq 0$.

2. Bilinearformen

Im Folgenden bezeichnet wir für einen K-Vektorraum V mit

$$\mathrm{Bl}(V,K) := \{\beta\colon V \times V \longrightarrow K \mid \beta \text{ Bilinearform}\}$$

den K-Vektorraum der K-Bilinearformen auf V. Es handelt sich um einen Untervektorraum von $\mathrm{Abb}(V \times V, K)$, d. h. $\mathrm{Bl}(V,K)$ wird zu einem Vektorraum mit den punktweisen Verknüpfungen.

Bemerkung VII.2.1: Seien K ein Körper, V ein Vektorraum über K und $\beta \in \mathrm{Bl}(V,K)$.

(i) Genau dann ist β symmetrisch, wenn für alle $v_1, v_2 \in V$ gilt, dass $\beta(v_1,v_2) = \beta(v_2,v_1)$.

(ii) Genau dann ist β schiefsymmetrisch, wenn für alle $v_1, v_2 \in V$ gilt, dass $\beta(v_1,v_2) = -\beta(v_2,v_1)$.

(iii) Genau dann ist β alternierend, wenn für alle $v \in V$ gilt, dass $\beta(v,v) = 0$.

Beispiel VII.2.2 (Einheitsform): Seien K ein Körper und n eine natürliche Zahl. Dann erklärt

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

eine symmetrische Bilinearform auf K^n.

Definition VII.2.3 (Skalarprodukte über den reellen Zahlen): Seien $K = \mathbb{R}$ und $\beta\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine Bilinearform.

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

(ii) Ist β eine positiv definite, symmetrische Bilinearform, dann heißt β ein *Skalarprodukt*.

Beispiel VII.2.4 (für Skalarprodukte): (i) Für $K = \mathbb{R}$ ist die Einheitsform aus Proposition VII.2.2 ein Skalarprodukt auf $\mathbb{R}^n$ und heißt *Standardskalarprodukt* oder auch *euklidisches Skalarprodukt*.

(ii) Es bezeichne $V = C([0,1])$ den Vektorraum der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$. Auf V wird durch

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

ein Skalarprodukt erklärt.

Bemerkung VII.2.5: Seien K ein Körper und $A = (a_{i,j}) \in K^{n\times n}$. Durch

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

wird eine Bilinearform erklärt. Es bezeichne $\{e_1, \ldots, e_n\}$ die Standardbasis des K^n. Für die Bilinearform β_A gilt $\beta_A(e_i, e_j) = e_i^t A e_j = a_{i,j}$. Insbesondere folgt für zwei $n \times n$-Matrizen A, A' mit $A \neq A'$, dass $\beta_A \neq \beta_{A'}$. Wir erhalten also einen injektiven Homomorphismus von K-Vektorräumen

$$K^{n\times n} \hookrightarrow \mathrm{Bl}(K^n, K), \qquad A \longmapsto \beta_A.$$

Genau dann ist die Bilinearform β_A symmetrisch, wenn $A = A^t$ gilt und genau dann ist β_A schiefsymmetrisch, wenn $A = -A^t$ ist. Für $A = I_n$ ist β_A genau die Einheitsform aus Proposition VII.2.2 (vergleiche Übungsblätter 9 und 10 zur Linearen Algebra I aus dem letzten Semester).

Proposition VII.2.6: *Seien K ein Körper und $\beta \colon K^n \times K^n \to K$ eine Bilinearform. Dann gibt es eine Matrix $A = (a_{i,j}) \in K^{n\times n}$, sodass $\beta = \beta_A$. Die Einträge der Matrix A sind bestimmt durch $a_{i,j} = \beta(e_i, e_j)$.*

Die Abbildung $K^{n\times n} \hookrightarrow \mathrm{Bl}(K^n, K)$, $A \mapsto \beta_A$ aus der vorangegangenen Bemerkung ist also sogar ein Isomorphismus.

Beweis: Seien $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$ in K^n. Dann ist

$$\beta(x, y) = \beta\Big(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\Big) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \beta(e_i, e_j) = x^t A y,$$

d. h. die Wirkung von β ist eindeutig durch die Werte $\beta(e_i, e_j)$, $1 \leq i, j \leq n$, festgelegt. Die Matrix A heißt *Gram-Matrix zu β bezüglich der Standardbasis.*□

In der Linearen Algebra I haben wir uns davon überzeugt, dass wir durch Wahl einer (geordneten) Basis $B = (b_1, \ldots, b_n)$ in einem K-Vektorraum V der Dimension n einen Isomorphismus $V \to K^n$ erhalten, via

$$v = \sum_{i=1}^n v_i b_i \longmapsto D_B(v) = \sum_{i=1}^n v_i e_i = (v_1, \ldots, v_n)^t.$$

Proposition VII.2.7 (Gram-Matrix): *Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$, $\beta \colon V \times V \to K$ eine Bilinearform und $G := (g_{i,j})_{1\leq i,j\leq n} \in K^{n\times n}$ die Matrix mit den Einträgen $g_{i,j} := \beta(b_i, b_j)$. Dann gilt für alle $v, w \in V$:*

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

Die Matrix G heißt Gram-Matrix von β bezüglich B.

Die Aussage zeigt man genau so wie die Behauptung aus Proposition VII.2.6, die Eindeutigkeit folgt wie in Proposition VII.2.5.

Proposition VII.2.8: *Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$, $\beta\colon V \times V \to K$ eine Bilinearform und sei $B' = (b'_1, \ldots, b'_n)$ eine weitere geordnete Basis von V. Ferner bezeichnen G die Gram-Matrix von β bezüglich B und G' die Gram-Matrix von β bezüglich B'. Dann gilt*

$$G' = D_{B,B'}^t G D_{B,B'}$$

wobei $D_{B,B'}$ bestimmt ist durch $D_B(v) = D_{B,B'} D_{B'}(v)$.

Beweis: Für Elemente v, w von V ist

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

d. h. $G' = D_{B,B'}^t G D_{B,B'}$ wegen der Eindeutigkeit der Gram-Matrix. □

Definition VII.2.9 (Orthogonalität): Seien K ein Körper, V ein Vektorraum über K und $\beta\colon V \times V \to K$ eine symmetrische Bilinearform.

(i) Zwei Elemente v, w von V mit $\beta(v,w) = 0$ heißen *orthogonal*.

(ii) Für eine Teilmenge $M \subseteq V$ heißt

$$M^\perp := \{v \in V \mid \text{Für alle } w \in M \text{ ist } \beta(w,v) = 0\}$$

das *orthogonale Komplement von M in V*. Es handelt sich wegen der Bilinearität von β um einen Untervektorraum von V. Außerdem gilt $M \subseteq (M^\perp)^\perp$.

(iii) Seien U_1 und U_2 Untervektorräume von V. Gilt für alle $v \in U_1$ und alle $w \in U_2$, dass $\beta(v,w) = 0$, dann schreiben wir $U_1 \perp U_2$. Insbesondere ist $U \perp U^\perp$.

Beweis: Wir wollen zeigen, dass in Situation von (ii) tatsächlich $M \subseteq (M^\perp)^\perp$. Sei dazu $v \in M$ gegeben. Für alle $w \in M^\perp$ ist $0 = \beta(v,w) = \beta(w,v)$, sodass $v \in (M^\perp)^\perp$. □

Definition VII.2.10 (Orthogonalsystem und Orthogonalbasis): Seien K ein Körper, V ein Vektorraum über K und $\beta\colon V \times V \to K$ eine symmetrische Bilinearform. Sind $v_1, \ldots, v_k$ Elemente von V, sodass für alle $1 \leq i, j \leq k$ mit $i \neq j$ gilt $\beta(v_i, v_j) = 0$, dann heißt $(v_1, \ldots, v_k)$ ein *Orthogonalsystem bezüglich β*. Gilt

sogar $\beta(v_i, v_j) = \delta_{ij}$, dann heißt $(v_1, \ldots, v_k)$ ein *Orthonormalsystem bezüglich* β. Ist β aus dem Kontext klar, dann lassen wir den Zusatz „bezüglich β“ auch weg.

Ist $(v_1, \ldots, v_k)$ eine Basis und gleichzeitig ein Orthogonalsystem beziehungsweise ein Orthonormalsystem, dann heißt $(v_1, \ldots, v_k)$ eine *Orthogonalbasis* beziehungsweise eine *Orthonormalbasis*.

Auch für Bilinearformen $\beta\colon V \times V \to K$ die nicht symmetrisch sind, lässt sich das Konzept von Orthogonalität erklären – allerdings müssen wir dann unterscheiden zwischen Linksorthogonalität und Rechtsorthogonalität, d. h. wir erhalten Mengen ${}^{\perp}M$ und $M^{\perp}$. Darauf wollen wir aber im Rahmen dieser Vorlesung nicht weiter eingehen.

Definition VII.2.11 (Anisotrop): Seien K ein Körper, V ein Vektorraum über K und $\beta\colon V \times V \to K$ eine symmetrische Bilinearform auf V. Gibt es ein $v \in V - \{\mathbf{0}\}$ mit $\beta(v, v) = 0$, dann heißt β *isotrop*. Ist β nicht isotrop, dann heißt β *anisotrop*.

Satz 26 (Fourierformel): *Seien K ein Körper, V ein Vektorraum über K, $\beta\colon V \times V \to K$ eine symmetrische anisotrope Bilinearform und $(v_1, \ldots, v_k)$ ein Orthogonalsystem bezüglich β, $v_1, \ldots, v_k \neq \mathbf{0}$. Dann gilt:*

(i) *Ist $v \in \operatorname{Lin}(v_1, \ldots, v_k)$, d. h. gibt es $\lambda_1, \ldots, \lambda_k \in K$ mit $v = \sum_{i=1}^{k} \lambda_i v_i$, dann gilt für $1 \leq i \leq k$:*

$$\lambda_i = \frac{\beta(v, v_i)}{\beta(v_i, v_i)}.$$

(ii) *Die Vektoren $v_1, \ldots, v_k$ sind linear unabhängig.*

Beweis: Zu (i): Wir haben

$$\beta(v, v_i) = \beta\Big(\sum_{j=1}^{k} \lambda_j v_j, v_i\Big) = \sum_{j=1}^{k} \lambda_j \beta(v_j, v_i) = \lambda_i \beta(v_i, v_i),$$

und da β anisotrop ist, dürfen wir durch $\beta(v_i, v_i)$ teilen, was die Behauptung liefert. Aussage (ii) ist nun eine direkte Konsequenz. □

Satz 27 (Orthogonalisierungsverfahren nach Gram, Schmidt): *Seien K ein Körper, V ein endlichdimensionaler Vektorraum über K mit Basis $(v_1, \ldots, v_n)$ und $\beta\colon V \times V \to K$ eine symmetrische anisotrope Bilinearform. Rekursiv definieren wir Vektoren $w_1, \ldots, w_n$ durch*

$$w_1 := v_1, \qquad w_\ell := v_\ell - \sum_{i=1}^{\ell-1} \frac{\beta(w_i, v_\ell)}{\beta(w_i, w_i)} w_i.$$

Dann gilt:

(i) *Das Tupel* $(w_1, \ldots, w_n)$ *ist eine Orthogonalbasis von* V.

(ii) *Für jedes* $1 \leq \ell \leq n$ *gilt* $\mathrm{Lin}(w_1, \ldots, w_\ell) = \mathrm{Lin}(v_1, \ldots, v_\ell)$.

Beweis: Aussage (ii) folgt induktiv aus der Definition der w_ℓ, denn der Vektor $\sum_{i=1}^{\ell-1} \beta(w_i, v_\ell)/\beta(w_i, w_i) w_i$ gehört zu $\mathrm{Lin}(w_1, \ldots, w_{\ell-1}) = \mathrm{Lin}(v_1, \ldots, v_{\ell-1})$.

Zu Aussage (i): Wir zeigen per Induktion über $1 \leq \ell \leq n$ für $1 \leq i, j \leq \ell$ mit $i \neq j$, dass $\beta(w_i, w_j) = 0$. Für den Induktionsanfang ist nichts zu zeigen. Die Aussage gelte nun für $\ell - 1$. Per Induktionsvoraussetzung gilt also für $1 \leq i, j \leq \ell - 1$ und $i \neq j$, dass $\beta(w_i, w_j) = 0$. Für beliebiges $1 \leq j \leq \ell - 1$ ist

$$\begin{aligned}
\beta(w_j, w_\ell) &= \beta\Big(w_j, v_\ell - \sum_{i=1}^{\ell-1} \frac{\beta(w_i, v_\ell)}{\beta(w_i, w_i)} w_i\Big) \\
&= \beta(w_j, v_\ell) - \sum_{i=1}^{\ell-1} \frac{\beta(w_i, v_\ell)}{\beta(w_i, w_i)} \beta(w_j, w_i) \\
&= \beta(w_j, v_\ell) - \frac{\beta(w_j, v_\ell)}{\beta(w_j, w_j)} \beta(w_j, w_j) = \beta(w_j, v_\ell) - \beta(w_j, v_\ell) = 0.
\end{aligned}$$

Insbesondere liefert Satz 26 zusammen mit (ii), dass $(w_1, \ldots, w_n)$ eine Orthogonalbasis bildet. □

Definition VII.2.12 (Orthogonale Projektion): Seien K ein Körper, V ein Vektorraum über K, β eine symmetrische anisotrope Bilinearform.

(i) Sind U_1 und U_2 orthogonale Unterräume von V, dann ist $U_1 + U_2 = U_1 \oplus U_2$, d. h. die Summe ist direkt.

(ii) Ist V ein endlichdimensionaler Vektorraum, dann ist $V = U \oplus U^\perp$. Die Abbildung

$$\pi \colon V = U \oplus U^\perp \longrightarrow U, \qquad v = u + u' \longmapsto u$$

heißt *orthogonale Projektion*.

Proposition VII.2.13 (Satz des Phytagoras): *Seien K ein Körper, V ein Vektorraum über K und β eine symmetrische Bilinearform. Sind v und w Elemente von V mit $\beta(v, w) = 0$, dann ist*

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

Beweis: Mit der Bilinearität von β rechnen wir nach:

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

□

3. Linearformen und der Dualraum

In diesem Abschnitt möchten wir Linearformen, d. h. lineare Abbildungen $f\colon V \to K$ von einem K-Vektorraum V in den Grundkörper K, systematisch untersuchen und den Dualraum $V^* := \mathrm{Hom}(V, K)$ einführen. Dabei erhalten wir folgende zentrale Ergebnisse: Erstens ist V ein endlichdimensionaler Vektorraum, dann ist V isomorph zu seinem Dualraum. Jedoch gibt es nicht einen eindeutigen Isomorphismus $V \to V^*$, für jede Basis von V oder jede „geeignete“ Bilinearform erhalten wir einen Solchen. Zweitens gibt es einen natürlichen Morphismus von V nach V^{**}, den sogenannten *Bidualraum von V*.

Definition VII.3.1: Seien K ein Körper und V ein endlichdimensionaler K-Vektorraum.

(i) Eine lineare Abbildung $f\colon V \to K$ heißt *Linearform* oder *lineares Funktional*.

(ii) Die Menge $V^* := \mathrm{Hom}_K(V, K) = \mathrm{Hom}(V, K) = \{f\colon V \to K \text{ linear}\}$ heißt *Dualraum von V*.

Als Untervektorraum von $\mathrm{Abb}(V, K)$ ist $\mathrm{Hom}(V, K)$ ein K-Vektorraum (natürlich mit den punktweisen Verknüpfungen).

Beispiel VII.3.2 (für Linearformen): (i) Sei V der K-Vektorraum K^3. Dann ist

$$f\colon V \longrightarrow K, \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 + 2x_2 + 3x_3$$

eine Linearform auf V wegen der Rechenregeln für Matrizenmultiplikation. Allgemein gilt $(K^n)^* = K^{1\times n}$.

(ii) Sei V der Vektorraum der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$. Dann ist

$$I\colon C([0,1]) \longrightarrow \mathbb{R}, \qquad h \longmapsto \int_0^1 h(x)\,dx$$

eine Linearform. Für ein festes $x_0 \in [0,1]$ ist außerdem auch

$$f_{x_0} := \left.\frac{d}{dx}\right|_{x=x_0} \colon C([0,1]) \longrightarrow \mathbb{R}, \qquad h \longmapsto h'(x_0)$$

eine Linearform auf V. Genauso ist Auswertung in x_0, d. h. $A_{x_0}\colon h \mapsto h(x_0)$ eine Linearform auf V.

(iii) Ist V ein K-Vektorraum, ist $\beta\colon V \times V \to K$ eine Bilinearform und ist $w \in V$ ein Element, dann ist $L_w\colon V \to K$, $v \mapsto \beta(v,w)$ eine Linearform.

Bemerkung VII.3.3 (Beschreibung des Dualraums V^*): Seien K ein Körper und V ein K-Vektorraum mit Basis B.

(i) Nach dem Fortsetzungssatz aus der Linearen Algebra I ist die Abbildung

$$R_B\colon V^* = \mathrm{Hom}(V,K) \longrightarrow \mathrm{Abb}(B,K), \qquad f \longmapsto f|_B$$

ein Isomorphismus. Insbesondere gilt $V^* \cong \mathrm{Abb}(B,K)$.

(ii) In der Linearen Algebra I haben wir die Koordinatenabbildung

$$D_B\colon V \longrightarrow \mathrm{Abb}_0(B,K) = \{h\colon B \to K \mid \#\{b \in B \mid h(b) \neq 0\} < \infty\}$$

als Isomorphismus von K-Vektorräumen kennengelernt. Das heißt wir erhalten einen injektiven Homomorphismus

$$\Theta_b := R_B^{-1} \circ D_B\colon V \longrightarrow V^*, \qquad v = \sum_{b \in B} \lambda(b) b \mapsto f$$

wobei f diejenige Abbildung mit $f(b) = \lambda(b)$ ist.

Bemerkung VII.3.4 (Duale Basis): Seien K ein Körper und V ein endlichdimensionaler K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_n)$. Die Homomorphismen aus Proposition VII.3.3 lassen sich in diesem Fall wie folgt angeben:

(i) $R_B\colon V^* \to K^{1\times n}$, $f \mapsto (f(b_1), \ldots, f(b_n))$,

(ii) $\Theta_B\colon V \to V^*$, $v = \sum_{i=1}^n \lambda_i b_i \mapsto f$ mit $f(b_i) = \lambda_i$.

In dieser Situation ist Θ_B sogar ein Isomorphismus.

(iii) Für $1 \leq i \leq n$, setze $b_i^* := \Theta_B(b_i)$. Das Tupel $B^* := (b_1^*, \ldots, b_n^*)$ ist eine geordnete Basis von V^*, die sogenannte *zu B duale Basis*. Es gilt insbesondere $b_i^*(b_j) = \delta_{ij}$.

(iv) Für $B^* = (b_1^*, \ldots, b_n^*)$ ist $b_i^*(v)$ die i-te Koordinate von $D_B(v)$, genauer: Für $v = \sum_{i=1}^n \lambda_i b_i$ ist

$$b_i^*(v) = b_i^*\Big(\sum_{j=1}^n \lambda_j b_j\Big) = \sum_{j=1}^n \lambda_j b_i^*(b_j) = \lambda_i b_i^*(b_i) = \lambda_i.$$

(v) Für $B^* = (b_1^*, \ldots, b_n^*)$ gilt: Ist $f = \sum_{i=1}^n f_i b_i^*$ eine Linearform auf V, dann ist $D_{B^*}(f) = (f(b_1), \ldots, f(b_n))$. Für $v = \sum_{i=1}^n \lambda_i b_i$ ist nämlich

$$f(v) = f\Big(\sum_{i=1}^n \lambda_i b_i\Big) = \sum_{i=1}^n \lambda_i f(b_i) = \sum_{i=1}^n b_i^*(v) f(b_i) = \sum_{i=1}^n f(b_i) b_i^*(v),$$

d. h. $f = \sum_{i=1}^n f(b_i) b_i^*$ wie gewünscht.

Beispiel VII.3.5: Es sei $V = P_2 = \{p \in \mathbb{R}[X] \mid \deg(p) \leq 2\}$ zusammen mit der Standardbasis $B = (b_0, b_1, b_2) = (1, X, X^2)$. Für ein $p = a_0 + a_1 X + a_2 X^2$ aus V ist $b_i^*(p) = a_i$, was uns die b_i^* als Abbildungen beschreibt.

Wollen wir das Funktional $A_2 \colon P_2 \to \mathbb{R}$, $p \mapsto p(2)$ in der dualen Basis ausdrücken, dann müssen wir nach der vorangegangenen Bemerkung die Werte von A_2 auf b_0, b_1, b_2 bestimmen. Es sind $A_2(b_0) = 1$, $A_2(b_1) = 2$, $A_2(b_2) = 4$, d. h. $D_{B^*}(A_2) = (1, 2, 4)^t$, also $A_2 = b_0^* + 2b_1^* + 4b_2^*$.

Definition VII.3.6 (Duale Abbildung): Seien K ein Körper, V_1 und V_2 zwei K-Vektorräume und $\phi \colon V_1 \to V_2$ eine lineare Abbildung. Dann ist

$$\phi^* \colon V_2^* \longrightarrow V_1^*, \qquad f \longmapsto f \circ \phi$$

eine lineare Abbildung und heißt die *zu ϕ duale Abbildung* oder einfach *duale Abbildung zu ϕ*.

Beweis: Seien f und g beliebige Elemente von V_2^* und λ irgendein Element von K. Wir wollen überprüfen, ob ϕ^* linear ist, d. h. ob $\phi^*(f+g) = \phi^*(f) + \phi^*(g)$ und $\phi^*(\lambda f) = \lambda \phi^*(f)$. Diese Gleichheit von Abbildungen von V_1 nach K rechnen wir auf einem beliebigen Element von V_1 nach. Für so ein v ist

$$\begin{aligned}\phi^*(f+g)(v) &= ((f+g) \circ \phi)(v) \\ &= (f+g)(\phi(v)) \\ &= f(\phi(v)) + g(\phi(v)) + \phi^*(f)(v) + \phi^*(g)(v) = (\phi^*(f) + \phi^*(g))(v),\end{aligned}$$

sowie

$$\phi^*(\lambda f)(v) = ((\lambda f) \circ \phi)(v) = \lambda f(\phi(v)) = \lambda \phi^*(f)(v). \qquad \square$$

Beispiel VII.3.7: Seien $V = P_3 = \{p \in \mathbb{R}[X] \mid \deg(p) \leq 3\}$, $B = (1, X, X^2, X^3)$ und $\phi \colon V \to V$ die Ableitung, d. h.

$$\phi(a_3 X^3 + a_2 X^2 + a_1 X + a_0) \longmapsto 3a_3 X^2 + 2a_2 X + a_1.$$

Die Darstellungsmatrix von ϕ bezüglich B können wir einfach ausrechnen und erhalten

$$D_{B,B}(\phi) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Wie können wir jetzt die duale Abbildung $\phi^*\colon V^* \to V^*$ beschreiben? Die duale Abbildung ϕ^* schickt ein Funktional auf die Präkomposition mit ϕ, d. h. die Abbildung, die ein Polynom p auf $f(\phi(p)) = f(p')$ schickt. Für die duale Basis $b_0^*, \ldots, b_3^*$ erhalten wir

$$\begin{aligned} \phi^*(b_0^*) &= (p \mapsto b_0^*(p') = a_1) = b_1^* \\ \phi^*(b_1^*) &= (p \mapsto b_1^*(p')) = 2a_2) = 2b_2^* \\ \phi^*(b_2^*) &= (p \mapsto b_2^*(p') = 3a_3) = 3b_3^* \\ \phi^*(b_3^*) &= (p \mapsto b_3^*(p') = 0) = 0 \end{aligned}$$

d. h. die Abbildungsmatrix von ϕ^* bezüglich B^* ist gegeben durch

$$D_{B^*,B^*}(\phi^*) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

Proposition VII.3.8 (Abbildungsmatrix für duale Abbildung): *Seien V_1 und V_2 endlichdimensionale Vektorräume über K mit Basen $B = (b_1, \ldots, b_n)$ und $C = (c_1, \ldots, c_m)$, $\phi\colon V_1 \to V_2$ eine lineare Abbildung und $A := D_{C,B}(\phi)$. Dann ist die Darstellungsmatrix $D_{B^*,C^*}(\phi^*)$ der dualen Abbildung $\phi^*\colon V_2^* \to V_1^*$ gegeben durch $D_{B^*,C^*}(\phi^*) = A^t$.*

Beweis: Seien $g \in V_2^*$ und für $1 \leq i \leq n$ bezeichne $\alpha_i := g(c_i)$. Wir haben das kommutative Diagramm

$$\begin{array}{ccccc} V_1 & \xrightarrow{\phi} & V_2 & \xrightarrow{g} & K \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} & & \downarrow{\scriptstyle \mathrm{id}} \\ K^n & \xrightarrow[x \mapsto Ax]{} & K^m & \xrightarrow[y \mapsto (\alpha_1, \ldots, \alpha_m)y]{} & K \end{array}$$

Insbesondere gilt $(g \circ \phi)(v) = (\alpha_1, \ldots, \alpha_m)AD_B(v)$. Das Diagramm für die

Abbildungsmatrix der dualen Abbildung ist

$$\begin{array}{ccc} V_2^* & \xrightarrow{\phi^*} & V_1^* \\ {\scriptstyle D_{C^*}}\downarrow & & \downarrow{\scriptstyle D_B^*} \\ K^m & \xrightarrow[D_{B^*,C^*}(\phi^*)]{} & K^n \end{array}$$

und aus dem vorangegangenen Diagramm erhalten wir

$$(g \circ \phi(b_1)| \dots |g \circ \phi(b_n)) = (\alpha_1, \dots, \alpha_m)AI = (\alpha_1, \dots, \alpha_m)A,$$

sodass $(g \circ \phi(b_1), \dots, g \circ \phi(b_n))^t = A^t(\alpha_1, \dots, \alpha_m)^t$, d. h. $y \mapsto A^t y$ macht das Diagramm für die Abbildungsmatrix $D_{B^*,C^*}(\phi^*)$ kommutativ und die Behauptung folgt. □

Definition VII.3.9 (Einsetzungshomomorphismus): Seien K ein Körper, V ein K-Vektorraum und $v \in V$ gegeben. Die Abbildung

$$A_v \colon V^* \longrightarrow K, \qquad f \longmapsto f(v)$$

nennen wir *Einsetzungshomomorphismus*. Der Einsetzungshomomorphismus A_v ist eine Linearform, d. h. A_v gehört zu $(V^*)^*$.

Proposition VII.3.10 (Einbettung in Bidualraum): *Seien K ein Körper und V ein K-Vektorraum. Setze*

$$\psi \colon V \longrightarrow V^{**}, \qquad v \longmapsto A_v,$$

wobei A_v den Einsetzungshomomorphismus aus Proposition VII.3.9 bezeichnet. Dann gilt:

(i) *Die Abbildung ψ ist linear,*

(ii) *Die Abbildung ist injektiv.*

Beweis: (i) Seien v_1 und v_2 in V und $\lambda \in K$. Dann gilt

$$\begin{aligned} \psi(v_1 + v_2) &= A_{v_1+v_2} \\ &= (f \mapsto f(v_1 + v_2) = f(v_1) + f(v_2)) = A_{v_1} + A_{v_2} = \psi(v_1) + \psi(v_2). \end{aligned}$$

Analog rechnet man nach, dass $\psi(\lambda v) = \lambda\psi(v)$.

(ii) Sei $v \in \mathrm{Kern}(\psi)$, d. h. $A_v = \mathbf{0}$. Dann gilt für alle $f \in V^*$, dass $f(v) = \mathbf{0}$. Es gibt eine Basis B von V. (Wir werden am Ende der Vorlesung zeigen, dass das auch für unendlichdimensionale Vektorräume gilt; dazu benötigt man das sogenannte Lemma von Zorn). Für $b \in B$, definiere b^* durch lineare Fortsetzung von $b \mapsto 1$ und $b' \mapsto 0$ für $b' \in B - \{b\}$. Schreibe $v = \sum_{b \in B} \lambda(b) b$. Für jedes $b \in B$ ist $b^*(v) = 0 = \lambda(b)$, d. h. v muss der Nullvektor sein. □

Der K-Vektorraum V kann via ψ aufgefasst werden als Untervektorraum von V^{**} und falls V endlichdimensional ist, dann werden V und V^{**} sogar kanonisch identifiziert, denn wegen $\dim V^{**} = \dim V^* = \dim V$ ist ψ dann ein Isomorphismus.

Proposition VII.3.11 (Kanonische Einbettung in Dualraum via Bilinearform): *Seien K ein Körper, V ein Vektorraum über K und $\beta \colon V \times V \to K$ eine anisotrope Bilinearform. Für ein festes $w \in V$ ist $L_w \colon V \to K$, $v \mapsto \beta(v, w)$ eine Linearform. Für*

$$\Theta_\beta \colon V \longrightarrow V^*, \qquad w \longmapsto L_w$$

gilt:

(i) *Die Abbildung Θ_β ist linear.*

(ii) *Die Abbildung Θ_β ist injektiv.*

(iii) *Ist V endlichdimensional, dann ist Θ_β ein Isomorphismus.*

Beweis: (i) Das ist klar wegen der Bilinearität von β.

(ii) Ist $w \in V$ mit $L_w = \mathbf{0}$, dann gilt für alle $v \in V$, dass $L_w(v) = \mathbf{0}$, d. h. für alle $v \in V$ ist $\beta(v, w) = 0$. Insbesondere ist dann auch $\beta(w, w) = 0$, sodass die Anisotropie $w = \mathbf{0}$ erzwingt.

(iii) Das folgt aus $\dim V^* = \dim V$. □

Bei Anwendungen von Proposition VII.3.11 ist β häufig ein Skalarprodukt über $\mathbb{R}$. Die Aussage Proposition VII.3.11 ist ein Spezialfall des Satzes von Riesz.

Einen Vektorraum V über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ zusammen mit einem Skalarprodukt heißt Prä-Hilbertraum. Auf V wird durch

$$d(v, w) := \langle v - w, v - w \rangle^{1/2}$$

eine Metrik erklärt. Ist (V, d) ein vollständiger metrischer Raum, dann heißt V *Hilbertraum.*

Satz (von Riesz): *Seien V ein Hilbertraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $\Theta_{\langle \cdot, \cdot \rangle}$ die Abbildung aus VII.3.11. Dann ist*

$$\Theta_{\langle \cdot, \cdot \rangle} \colon V \hookrightarrow (V^*)^{\mathrm{top}}$$

ein Isomorphismus. Hierbei bezeichnet $(V^*)^{\mathrm{top}} = \{f \in V^* \mid f \text{ ist stetig}\}$ den topologischen Dualraum von V.

Beachte, dass die Aussage des Satzes von Riesz auch für unendlichdimensionale Hilberträume gilt.

Bemerkung VII.3.12: Seien V ein endlichdimensionaler Vektorraum über K und $\beta \colon V \times V \to K$ eine anisotrope Bilinearform.

(i) Ist B eine Basis von V, G die Gram-Matrix von β bezüglich B und B^* die duale Basis zu B von V^*, dann gilt $D_{B^*,B}(\Theta_\beta) = G$. Insbesondere gilt für für ein $w \in V$, dass $D_{B^*}(L_w) = GD_B(w)$.

(ii) Sei nun $\beta = \langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. Via Θ_β kann $\langle \cdot, \cdot \rangle$ auf V^* übertragen werden durch

$$\langle f, g \rangle_{V^*} := \langle \Theta_\beta^{-1}(f), \Theta_\beta^{-1}(g) \rangle.$$

Insbesondere gilt für alle $v, w \in V$, dass $\langle \Theta_\beta(v), \Theta_\beta(w) \rangle_{V^*} = \langle v, w \rangle$, d. h. Θ_β ist eine *Isometrie* (siehe später).

(iii) Das Kreuzprodukt kann mithilfe des Satzes von Riesz koordinatenfrei definiert werden.

Aussagen (i) und (iii) werden Sie sich auf einem kommenden Übungsblatt genauer ansehen dürfen.

Satz 28 (über den Dualraum): *Seien K ein Körper und V ein K-Vektorraum.*

(i) *Jede Basis B von V definiert durch $\Theta_B \colon V \hookrightarrow V^*$, $v = \sum_{b \in B} \lambda(b) b \mapsto \sum_{\lambda \in B} \lambda(b) b^*$ eine Identifikation von V mit einem Untervektorraum des Dualraums V^*. Ist V endlichdimensional, dann ist Θ_B ein Isomorphismus.*

(ii) *Jede anisotrope Bilinearform $\beta \colon V \times V \to K$ definiert durch $\Theta_\beta \colon V \hookrightarrow V^*$, $w \mapsto L_w = \beta(\cdot, w)$ eine Identifikation von V mit einem Untervektorraum von V^*. Ist V endlichdimensional, dann ist Θ_β ein Isomorphismus.*

(iii) *Durch $V \hookrightarrow V^{**}$, $v \mapsto A_v = (f \mapsto f(v))$ erhalten wir eine kanonische Einbettung von V in V^{**}. Ist V endlichdimensional, dann handelt es sich um einen Isomorphismus.*

Der Dualraum verdankt seinen Namen dem lateinischen Wort „dual“, welches „zweifach“, oder „zwei betreffend“ heißt. Hier ist dual am besten zu verstehen als „gegenläufig“, oder „Pfeile drehen sich um“ – gemeint sind Abbildungspfeile.

Zum Beispiel in der Analysis oder der Geometrie haben Dualräume vielfältige Anwendungen. Ein prominentes Beispiel sind Tangentialräume und Kotangentialräume. Ist beispielsweise X eine Fläche im $\mathbb{R}^3$ und p ein Punkt der Fläche X, dann heften wir einen $\mathbb{R}$-Vektorraum T_pX, den Tangentialraum an X in p, an diese Fläche an. Diesen Vektorraum können wir uns vorstellen als den Graph der linearen Abbildung, die die Fläche X in p „am besten approximiert“.[1] Auf natürliche Weise ist dieser Tangentialraum in den umgebenden $\mathbb{R}^3$ eingebettet und erbt so das Standardskalarprodukt.

Ist $f\colon X \to \mathbb{R}$ eine stetig differenzierbare Funktion, und gehört v zu T_pX, dann heißt $D_vf(p)$ die Richtungsableitung in Richtung p. Wir erhalten eine Abbildung

$$Df(p)\colon T_pX \longrightarrow \mathbb{R}, \qquad v \longmapsto D_vf(p).$$

Diese Abbildung $Df(p)$ gehört zum Dualraum von T_pX, dem sogenannten Kotangentialraum, der üblicherweise mit T_p^*X bezeichnet wird. Nach dem Satz von Riesz gibt es genau einen Vektor $w \in T_pX$, sodass für alle $v \in T_pX$ gilt: $D_vf(p) = \langle v, w\rangle$. Dieser spezielle Vektor heißt Gradient von f in p.

4. Tensorprodukte

Seien V_1, V_2 und W Vektorräume über dem selben Körper K. In diesem Abschnitt möchten wir „die Mutter aller bilinearen Abbildungen $V_1 \times V_2 \to W$“ kennen lernen. Die führt zum Begriff des Tensorprodukts $V_1 \otimes_K V_2$.

Dieses Konzept hat viele Anwendungen in der Analysis, der Geometrie und wird insbesondere in vielen Ingenieurswissenschaften häufig gebraucht.

Für einen K-Vektorraum V und einen Untervektorraum U haben wir in der Linearen Algebra I den besonderen Vektorraum V/U kennengelernt. Dieser hieß „Quotientenvektorraum“ oder auch „Faktorraum“ und war definiert als die Menge der Äquivalenzklassen $V/U = \{[v]_\sim \mid v \in V\}$ bezüglich der auf V erklärten Äquivalenzrelation

$$v \sim w :\Longrightarrow v - w \in U,$$

welchen wir per $[v]+[w] := [v+w]$ und $\lambda[v] := [\lambda v]$ eine K-Vektorraumstruktur aufgeprägt haben. Für ein $v \in V$ ist die Äquivalenzklasse $[v]$ genau die Menge

[1] Ist X beispielsweise der Graph einer differenzierbaren Funktion, dann ist der Tangentialraum in einem Punkt der Graph der totalen Ableitung dieser differenzierbaren Funktion an einer geeigneten Stelle.

$v+U = \{v+u \mid u \in U\}$. Die Restklassenabbildung $\pi\colon V \to V/U$, $v \mapsto [v]$ heißt *kanonische Projektion* und ist per Konstruktion der K-Vektorraumstruktur von V/U trivialerweise eine lineare Abbildung.

Außerdem haben wir den Homomorphiesatz in der Linearen Algebra I kennengelernt: Für K-Vektorräume V und W, eine lineare Abbildung $\phi\colon V \to W$ und einen Untervektorraum $U \subseteq V$ mit $U \subseteq \mathrm{Kern}(\phi)$ haben wir das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U \\ & \searrow^{\phi} & \downarrow{\scriptstyle \bar{\phi}} \\ & & W \end{array}$$

Das heißt es gibt genau eine lineare Abbildung $\bar{\phi}\colon V/U \to W$ mit $\bar{\phi} \circ \pi = \phi$.

Bemerkung VII.4.1: Seien K ein Körper, n und m natürliche Zahlen, $V_1 := K^n$, $V_2 := K^m$ und $T := K^{n\times m}$. Ferner sei

$$\tau\colon V_1 \times V_2 \longrightarrow T, \qquad (x, y) \longmapsto xy^t.$$

Dann gilt:

(i) Die Abbildung τ is bilinear.

(ii) Für die Standardbasen $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_m)$ die Standardbasen von K^n beziehungsweise K^m gilt $\tau(e_i, e'_j) = E_{i,j}$, wobei $E_{i,j}$ die Elementarmatrix in $K^{n\times m}$ bezeichnet, die durch $E_{i,j} = (\delta_{i,k}\delta_{j,\ell})_{1\le k,\ell\le n}$ definiert ist.

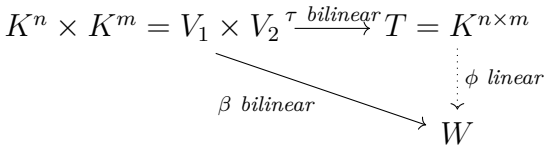
Beispiel VII.4.2: Seien $V_1 = \mathbb{R}^2$, $V_2 = \mathbb{R}^3$ und $T = \mathbb{R}^{2\times 3}$. Für die Abbildung τ aus Proposition VII.4.1 haben wir beispielsweise

$$\tau\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 0 & -2 \end{pmatrix}$$

$$\tau\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Proposition VII.4.3 (Universelle Abbildungseigenschaft): *Die Abbildung τ aus Proposition VII.4.1 hat folgende Eigenschaft: Für jeden K-Vektorraum W und*

jede bilineare Abbildung $\beta\colon V_1 \times V_2 \to W$ gibt es genau eine lineare Abbildung $\phi\colon T \to W$ mit $\phi \circ \tau = \beta$, d. h. wir haben das kommutative Diagramm



Beweis: Wieder bezeichne $(e_1, \ldots, e_n)$ beziehungsweise $(e'_1, \ldots, e'_m)$ die Standardbasis von K^n beziehungsweise K^m. Die Menge der Elementarmatrizen $\{E_{i,j} \mid 1 \le i \le n, 1 \le j \le n\}$ bildet eine Basis von $T = K^{n\times m}$.

Nach dem Fortsetzungssatz für lineare Abbildungen gibt es genau eine lineare Abbildung $\phi\colon T \to W$ mit $\phi(E_{i,j}) = \beta(e_i, e'_j)$ und für dieses ϕ gilt tatsächlich $\phi \circ \tau = \beta$. Seien dazu $v = \sum_{i=1}^n \lambda_i e_i$ in V_1 und $w = \sum_{j=1}^m \lambda'_j e'_j$ in V_2. Dann ist

$$\begin{aligned}
\phi(\tau(v,w)) &= \phi\Big(\tau\Big(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^m \lambda'_j e'_j\Big)\Big) \\
&= \phi\Big(\sum_{i=1}^n \sum_{j=1}^m \lambda_i \lambda_j \tau(e_i, e'_j)\Big) \\
&= \sum_{i=1}^n \sum_{j=1}^m \lambda_i \lambda'_j \phi(\tau(e_i, e'_j)) \\
&= \sum_{i=1}^n \sum_{j=1}^m \lambda_i \lambda'_j \beta(e_i, e_{j'}) = \beta\Big(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^m \lambda'_j e'_j\Big) = \beta(v,w).
\end{aligned}$$

Außerdem ist ϕ eindeutig, da insbesondere $\phi \circ \tau(e_i, e'_j) = \beta(e_i, e'_j)$ gelten muss.□

Beispiel VII.4.4: Seien $V_1 = V_2 = K^n$ und $\tau\colon K^n \times K^n \to K^{n\times n}$, $(x,y) \mapsto xy^t$. Ferner sei $\beta\colon K^n \times K^n \to K$ die Einheitsform, d. h. $\beta(x,y) = x^t y$. Was ist die Abbildung $\phi\colon K^{n\times n} \to K$ aus der vorangegangenen Proposition? Auf den Elementarmatrizen $E_{i,j}$ muss ϕ folgendes machen:

$$\phi(E_{i,j}) = \beta(e_i, e_j) = e_i^t e_j = \delta_{i,j}.$$

Für eine Matrix $A = (a_{i,j}) \in K^{n\times n}$ ist also

$$\phi(A) = \phi\Big(\sum_{i=1}^n \sum_{j=1}^n a_{i,j} E_{i,j}\Big) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}\delta_{i,j} = \sum_{i=1}^n a_{i,i},$$

d. h. $\phi(A)$ gibt die Summe der Diagonaleinträge von A zurück.

Definition VII.4.5: Seien K ein Körper, n eine natürliche Zahl und $A = (a_{i,j})$ in $K^{n\times n}$ eine Matrix. Dann heißt

$$\operatorname{tr}(A) := \sum_{i=1}^{n} a_{i,i} \in K$$

die *Spur von* A. Die Abbildung $\operatorname{tr}\colon K^{n\times n} \to K$, $A \mapsto \operatorname{tr}(A)$ ist eine lineare Abbildung, d. h. insbesondere $\operatorname{tr} \in (K^{n\times n})^*$.

Das Tupel (T, τ) aus Proposition VII.4.3 nennt man auch ein *Tensorprodukt von* V_1 *und* V_2. Dadurch, dass wir uns bereits für beliebige natürliche Zahlen n und m davon überzeugt haben, dass K^n und K^m ein Tensorprodukte haben, haben wir allgemeiner für endlichdimensionale Vektorräume die Existenz von Tensorprodukten etabliert. Es ist allerdings ein natürliches Bedürfnis, solche Tensorprodukte auch für unendlichdimensionale Vektorräume zu haben, wie zum Beispiel für Polynomringe. Damit wollen wir uns im Folgenden beschäftigen.

Definition VII.4.6 (Tensorprodukt): Seien K ein Körper, V_1, V_2 und T Vektorräume über K und $\tau\colon V_1 \times V_2 \to T$ eine bilineare Abbildung. Gibt es für jeden K-Vektorraum W und jede bilineare Abbildung $\beta\colon V_1 \times V_2 \to W$ genau eine lineare Abbildung $\phi\colon T \to W$ mit $\phi \circ \tau = \beta$, dann heißt das Tupel (T, τ) ein *Tensorprodukt von* V_1 *und* V_2 *über* K.

Notation VII.4.7: In der Situation von Proposition VII.4.6 schreiben wir für das Tensorprodukt $V_1 \otimes_K V_2 := V_1 \otimes V_2 := T$ und für $v_1 \in V_1$, $v_2 \in V_2$ schreiben wir $v_1 \otimes v_2 := \tau(v_1, v_2)$.

Wir werden zeigen: Je zwei K-Vektorräume V und W haben ein Tensorprodukt, und Tensorprodukte sind eindeutig bis auf Isomorphie.

Satz 29 (Existenz von Tensorprodukten): *Seien* V_1 *und* V_2 *Vektorräume über* K. *Dann existiert ein Tensorprodukt* (T, τ) *von* V_1 *und* V_2 *über* K.

Für endlichdimensionale Vektorräume V_1 und V_2 liefert Proposition VII.4.1 das Gewünschte. In dieser Situation haben wir $\{(e_i, e'_j) \mid 1 \le i \le n, 1 \le j \le m\}$ auf die Basis $\{E_{i,j} \mid 1 \le i \le n, 1 \le j \le m\}$ von $K^{n\times m}$ abgebildet.

Erinnerung VII.4.8 (Abbildungen mit endlichem Träger): Seien M eine Menge und K ein Körper.

(i) Für eine Abbildung $f\colon M \to K$ heißt $\operatorname{Tr}(f) := \{m \in M \mid f(m) \neq 0\}$ der *Träger von* f.

(ii) Wir schreiben

$$\mathrm{Abb}_0(M,K) := \{f\colon M \to K \mid \#\,\mathrm{Tr}(f) < \infty\} \subseteq \mathrm{Abb}(M,K)$$

für den Untervektorraum der Abbildungen mit endlichem Träger.

(iii) Für $m \in M$ bezeichne $f_m\colon M \to K$ die Abbildung mit

$$f_m(m') = \begin{cases} 1, & \text{falls } m = m', \\ 0, & \text{sonst.} \end{cases}$$

Die Menge $\{f_m \mid m \in M\}$ bildet eine Basis von $\mathrm{Abb}_0(M,K)$.

Beispielsweise für $M = \mathbb{Z}$ ist $\mathrm{Abb}_0(M,K)$ die Menge der endlichen K-wertigen Folgen indiziert über $\mathbb{Z}$. Für $M = \mathbb{R}^2$ und $(0,0) \in \mathbb{R}^2$ ist

$$f_{(0,0)}\colon \mathbb{R}\times\mathbb{R}, \qquad (x,y) \longmapsto \begin{cases} 1, & \text{falls } x = 0 \text{ und } y = 0, \\ 0, & \text{sonst.} \end{cases}$$

Beweis (von Satz 29): Im ersten Schritt möchten wir uns einen ersten (zu großen) Kandidaten für unser Tensorprodukt einführen. Wir setzen

$$F := \mathrm{Abb}_0(V_1 \times V_2, K), \qquad \tau_1\colon V_1 \times V_2 \longrightarrow F, \quad (v_1,v_2) \longmapsto f_{(v_1,v_2)}$$

mit $f_{(v_1,v_2)}$ wie in Proposition VII.4.8. Bis jetzt gut an diesen Kandidaten für T und τ ist, dass wir eine Basis $\{f_{(v_1,v_2)} \mid (v_1,v_2) \in V_1 \times V_2\}$ von F kennen, d. h. für jeden K-Vektorraum W und jede Abbildung $\beta\colon V_1 \times V_2 \to W$ erhalten wir per Fortsetzungssatz eine eindeutige lineare Abbildung $\hat{\phi}\colon F \to W$ mit $\hat{\phi} \circ \tau_1 = \beta$. (Dazu definieren wir $\hat{\phi}$ auf der Basis durch $\hat{\phi}(f_{(v_1,v_2)}) = \beta(v_1,v_2)$.) Schlecht ist, dass τ_1 weit entfernt davon ist, bilinear zu sein.

Im zweiten Schritt möchten wir unseren Kandidaten verbessern. Für alle $v_1, v_1' \in V_1$, v_2, $v_2' \in V_2$ und $\alpha_1, \alpha_2 \in K$ brauchen wir, dass

$$\tau(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') = \alpha_1\alpha_2\tau(v_1,v_2) + \alpha_1\tau(v_1,v_2') + \alpha_2\tau(v_1',v_2) + \tau(v_1',v_2').$$

Wir wollen im Folgenden durch geeignete Quotientenbildung „erzwingen“, dass genau das gilt, was wir uns wünschen.[2] Sei R der Untervektorraum von V, der von den Elementen

$$\{f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')} - \alpha_1\alpha_2 f_{(v_1,v_2)} - \alpha_1 f_{(v_1,v_2')} - \alpha_2 f_{(v_1',v_2)} - f_{(v_1',v_2')}\}$$

[2] Gelegentlich nennt man so etwas auch „Pippi-Langstrumpf-Mathematik“.

erzeugt wird. Wir setzen $T := F/R$, bezeichnen mit π die kanonische Projektion $\pi\colon F \to T = F/R$ und definieren $\tau\colon V_1 \times V_2 \to T = F/R$ durch $\tau := \pi \circ \tau_1$, also $(v_1, v_2) \mapsto [f_{(v_1,v_2)}] = f_{(v_1,v_2)} + R$. Dieses τ ist tatsächlich bilinear: Für $v_1, v_1' \in V$, $v_2, v_2' \in V'$ und $\alpha_1, \alpha_2 \in K$ gilt per Definition von $T = F/R$ und per Definition der Vektorraumstruktur auf T, dass

$$\begin{aligned}
&\tau(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') \\
&\quad = [f_{(\alpha v_1 + v_1', \alpha_2 v_2 + v_2')}] \\
&\quad = [\alpha_1 f_{(v_1, v_2')} + \alpha_2 f_{(v_1', v_2)} + f_{(v_1', v_2')} + \alpha_1\alpha_2 f_{(v_1,v_2)}] \\
&\quad = \alpha_1\alpha_2 [f_{(v_1,v_2)}] + \alpha_1 [f_{(v_1,v_2')}] + \alpha_2 [f_{(v_1',v_2)}] + [f_{(v_1',v_2')}] \\
&\quad = \alpha_1\alpha_2 \tau(v_1, v_2) + \alpha_1 \tau(v_1, v_2') + \alpha_2 \tau(v_1', v_2) + \tau(v_1', v_2').
\end{aligned}$$

Durch die Definition von T und τ haben wir die folgende Erweiterung des Diagramms von oben:

$$\begin{array}{ccccc}
V_1 \times V_2 & \xrightarrow{\tau_1} & F & \xrightarrow{\pi} & T = F/R \\
& \searrow^{\beta} & \downarrow \exists!\, \hat{\phi} & \swarrow_{\phi} & \\
& & W & &
\end{array}$$

Können wir zeigen, dass $R \subseteq \operatorname{Kern}(\hat{\phi})$, dann folgt aus dem Homomorphiesatz, dass es eine lineare Abbildung $\phi\colon T \to W$ gibt, die das rechte Dreieck im Diagramm kommutativ macht, d. h. die $\phi \circ \pi = \hat{\phi}$ leistet. Insgesamt ergibt das

$$\phi \circ \tau = \phi \circ \pi \circ \tau_1 = \hat{\phi} \circ \tau_1 = \beta,$$

d. h. wenn $R \subseteq \operatorname{Kern}(\hat{\phi})$, dann ist (T, τ) ein Tensorprodukt von V_1 und V_2. Um zu zeigen, dass R im Kern von $\hat{\phi}$ liegt, genügt es, das für die Erzeuger nachzurechnen. Für $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$ und $\alpha_1, \alpha_2 \in K$ gilt

$$\begin{aligned}
&\hat{\phi}(f_{(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2')}) \\
&\quad = \beta(\alpha_1 v_1 + v_1', \alpha_2 v_2 + v_2') \\
&\quad = \alpha_1\alpha_2 \beta(v_1, v_2) + \alpha_1 \beta(v_1, v_2') + \alpha_2 \beta(v_1', v_2) + \beta(v_1', v_2') \\
&\quad = \alpha_1\alpha_2 \hat{\phi}(f_{(v_1,v_2)}) + \alpha_1 \hat{\phi}(f_{(v_1,v_2')}) + \alpha_2 \hat{\phi}(f_{(v_1',v_2)}) + \hat{\phi}(f_{(v_1',v_2')})
\end{aligned}$$

was mit der Linearität von $\hat{\phi}$ die Behauptung liefert. Schließich ist ϕ eindeutig, weil $\{[f_{(v_1,v_2)}] \mid v_1 \in V_1, v_2 \in V_2\}$ ein Erzeugendensystem für T ist. □

Da wir $v_1 \otimes v_2 = \tau(v_1, v_2) = \pi(f_{(v_1,v_2)})$ vereinbart haben, und die Menge $\{f_{(v_1,v_2)} \mid v_1 \in V_1, v_2 \in V_2\}$ eine Basis von F ist, ist $\{v_1 \otimes v_2 \mid v_1 \in V_1, v_2 \in V_2\}$

ein Erzeugendensystem von $T = V_1 \otimes_K V_2$. Das bedeutet: Nicht jedes Element von $V_1 \otimes V_2$ ist von der Form $v_1 \otimes v_2$; ein beliebiges Element von $V_1 \otimes V_2$ können wir aber als endliche Summe von Elementen der Form $v_1 \otimes v_2$ schreiben.

Sind $\sum_{i=1}^n r_i v_i \in V_1$ und $\sum_{j=1}^m s_j w_j \in V_2$, dann ist

$$\begin{aligned}\Big(\sum_{i=1}^n r_i v_i\Big) \otimes \Big(\sum_{j=1}^m s_j w_j\Big) &= \tau\Big(\sum_{i=1}^n r_i v_i, \sum_{j=1}^m s_j w_j\Big) \\ &= \sum_{i=1}^n \sum_{j=1}^m r_i s_j \tau(v_i, w_j) = \sum_{i=1}^n \sum_{j=1}^m r_i s_j v_i \otimes w_j.\end{aligned}$$

Proposition VII.4.9: *Seien K ein Körper, V_1 und V_2 Vektorräume über K. Ein Tensorprodukt (T, τ) von V_1 und V_2 ist eindeutig bis auf eindeutige Isomorphie, d. h. sind (T, τ) und (T', τ') Tensorprodukte von V_1 und V_2 über K, dann gibt es einen eindeutigen Isomorphismus $\phi\colon T \to T'$ mit $\tau' = \phi \circ \tau$.*

Beweis: Wir befinden uns in der Situation

$$\begin{array}{ccc} V_1 \times V_2 & \xrightarrow{\ \tau\ } & T \\ & \searrow^{\tau'} & \\ & & T' \end{array}$$

und sowohl $\tau\colon V_1 \times V_2 \to T$ als auch $\tau'\colon V_1 \times V_2 \to T'$ sind bilinear. Über die universelle Abbildungseigenschaft von (T, τ) erhalten wir also eine eindeutige lineare Abbildung $\phi\colon T \to T'$ mit $\phi \circ \tau = \tau'$ und wegen der universellen Abbildungseigenschaft von (T', τ') gibt es eine eindeutige lineare Abbildung $\psi\colon T' \to T$ mit $\psi \circ \tau' = \tau$.

Wir haben also $(\psi \circ \phi) \circ \tau = \psi \circ \tau' = \tau$ und außerdem ist $\mathrm{id} \circ \tau = \tau$. Wegen der Eindeutigkeitsaussage in der universellen Abbildungseigenschaft für (T, τ) erhalten wir $\psi \circ \phi = \mathrm{id}$. Analog erhalten wir, dass $\phi \circ \psi = \mathrm{id}$. Damit ist ϕ ein eindeutiger Isomorphismus mit Inverser ψ. □

Proposition VII.4.10: *Seien K ein Körper, V_1 und V_2 zwei Vektorräume über K und B eine Basis von V_1 sowie C eine Basis von V_2. Dann ist*

$$D := \{b \otimes c \mid b \in B, c \in C\} \subseteq V_1 \otimes V_2$$

eine Basis von $V_1 \otimes V_2$.

Beweis: Wir haben zu zeigen, dass D das Tensorprodukt $V_1 \otimes V_2$ erzeugt und, dass D linear unabhängig ist. Wir kennen bereits ein Erzeugendensystem für

$V_1 \otimes V_2$, nämlich $\{v_1 \otimes v_2 \mid v_1 \in V_1, v_2 \in V_2\}$. Wir zeigen für die Erzeugendeneigenschaft von D, dass D dieses Erzeugendensystem erzeugt. Seien dazu $v_1 \in V_1$ und $v_2 \in V_2$. Dann gibt es eindeutige Linearkombinationen $v_1 = \sum_{i=1}^{n} r_i b_i$ und $v_2 = \sum_{j=1}^{m} s_j c_j$ und

$$v_1 \otimes v_2 = \Big(\sum_{i=1}^{n} r_i b_i\Big) \otimes \Big(\sum_{j=1}^{m} s_j c_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} r_i s_j b_i \otimes c_j,$$

d. h. $v_1 \otimes v_2$ gehört zu $\mathrm{Lin}(D)$ wie gewünscht.

Nun zu linearen Unabhängigkeit: Sei $\mathbf{0} = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j} b_i \otimes c_j$ eine Nulldarstellung. Wegen der universellen Abbildungseigenschaft von $V_1 \otimes V_2$ gilt für jede Bilinearform $\beta\colon V_1 \times V_2 \to K$ und die zugehörige lineare Abbildung $\phi_\beta\colon V_1 \otimes V_2 \to K$, dass

$$0 = \phi_\beta(\mathbf{0}) = \phi_\beta\Big(\sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j} b_i \otimes c_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j}\beta(b_i, c_j).$$

Seien $1 \le k \le n$ und $1 \le \ell \le m$. Insbesondere für die Bilinearform erklärt durch

$$\beta_{(k,\ell)}\colon B \times C \longrightarrow K, \qquad (b_i, c_j) \longmapsto \begin{cases} 1, & \text{falls } (i,j) = (k,\ell), \\ 0, & \text{sonst.} \end{cases}$$

gilt die obige Gleichung, d. h. $0 = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{i,j}\beta_{(k,\ell)}(b_i, c_j) = r_{k,\ell}$ und somit ist D linear unabhängig. □

Korollar VII.4.11: *Seien K ein Körper und V_1 und V_2 endlichdimensionale Vektorräume über K. Dann gilt*

$$\dim(V_1 \otimes V_2) = (\dim V_1)(\dim V_2).$$

Bemerkung VII.4.12: Seien K ein Körper, V_1, V_2, V_1', V_2' Vektorräume über K und $\phi\colon V_1 \to V_2$, $\psi\colon V_1' \to V_2'$ lineare Abbildungen. Ferner bezeichne $\phi \times \psi$ die Abbildung $V_1 \times V_1' \to V_2 \times V_2'$, $(v_1, v_1) \mapsto (\phi(v_1), \psi(v_1))$. Dann haben wir das Diagramm

$$\begin{array}{ccc} V_1 \times V_1' & \xrightarrow{\tau_1} & V_1 \otimes V_1' \\ \downarrow{\scriptstyle \phi\times\psi} & \searrow^{\tau_2\circ(\phi\times\psi)} & \downarrow{\scriptstyle \phi\otimes\psi} \\ V_2 \times V_2' & \xrightarrow[\tau_2]{} & V_2 \otimes V_2' \end{array}$$

und die Abbildung längs des roten Pfeils ist bilinear. Das heißt für den roten Pfeil erhalten wir eine eindeutige lineare Abbildung von $V_1 \otimes V_1'$ nach $V_2 \otimes V_2'$,

die wir suggestiv mit $\phi \otimes \psi$ bezeichnen wollen, die für alle $v_1 \in V_1$ und $v_1' \in V_1'$ erfüllt:

$$(\phi \otimes \psi)(v_1 \otimes v_1') = \phi(v_1) \otimes \psi(v_1').$$

Bemerkung VII.4.13 (Abbildungsmatrix von $\phi \otimes \psi$): Seien alle Vektorräume in Proposition VII.4.12 endlichdimensional mit Basen $B = (b_1, \ldots, b_m)$ von V_1, $C = (c_1, \ldots, c_n)$ von V_2, $B' = (b_1', \ldots, b_{m'}')$ von V_1' und $C' = (c_1, \ldots, c_{n'}')$ von V_2'. Ferner bezeichne $D_{C,B}(\phi) = (\beta_{i,j}) \in K^{n \times m}$ und $D_{C',B'}(\psi) = (\gamma_{k,\ell}) \in K^{n' \times m'}$ die Darstellungsmatrizen. Für die Basen

$$\begin{aligned} D_1 &:= (b_1 \otimes b_1', \ldots, b_1 \otimes b_{m'}', b_2 \otimes b_1', \ldots, b_2 \otimes b_{m'}', \ldots, b_m \otimes b_1', \ldots, b_m \otimes b_{m'}'), \\ D_2 &:= (c_1 \otimes c_1', \ldots, c_1 \otimes c_{n'}', \ldots, c_n \otimes c_1', \ldots, c_n \otimes c_{n'}') \end{aligned}$$

und die Darstellungsmatrix $A := D_{D_2,D_1}(\phi \otimes \psi)$ gilt

$$A = \begin{pmatrix} \beta_{1,1} D_{C',B'}(\psi) & \beta_{1,2} D_{C',B'}(\psi) & \cdots & \beta_{1,m} D_{C',B'}(\psi) \\ \beta_{2,1} D_{C',B'}(\psi) & \beta_{2,2} D_{C',B'}(\psi) & \ldots & \beta_{2,m} D_{C',B'}(\psi) \\ \vdots & \vdots & & \vdots \\ \beta_{n,1} D_{C',B'}(\psi) & \beta_{n,2} D_{C',B'}(\psi) & \cdots & \beta_{n,m} D_{C',B'}(\psi) \end{pmatrix}.$$

Man nennt A auch das Kroneckerprodukt von $D_{C,B}(\phi)$ und $D_{C',B'}(\psi)$.

Beweis: Für $1 \leq i \leq m$ und $1 \leq j \leq m'$ gilt

$$\begin{aligned} (\phi \otimes \psi)(b_i \otimes b_j') &= \phi(b_i) \otimes \psi(b_j') \\ &= \Big(\sum_{i=1}^{n} \beta_{k,i} c_k \Big) \otimes \Big(\sum_{\ell=1}^{n'} \gamma_{\ell,j} c_\ell' \Big) = \sum_{k=1}^{n} \sum_{\ell=1}^{n'} \beta_{k,i} \gamma_{\ell,j} c_k \otimes c_\ell'. \end{aligned}$$

Die Zeilen von A sind indiziert durch die Basiselemente $c_k \otimes c_\ell'$ und die Spalten von A sind indiziert durch die Basiselemente $b_i \otimes b_j'$. □

Proposition VII.4.14: *Seien K ein Körper, $(A, +, \circ, \cdot)$ und $(B, +, \circ, \cdot)$ zwei Algebren über K und $A \otimes_K B = A \otimes B$ das Tensorprodukt von A und B als K-Vektorräume. Dann wird das Tensorprodukt $A \otimes B$ zu einer K-Algebra durch*

$$\begin{aligned} \bullet \colon A \otimes B \times A \otimes B &\longrightarrow A \otimes B, \\ \Big(\sum_{i=1}^{m} a_i \otimes b_i, \sum_{j=1}^{n} a_j' \otimes b_j' \Big) &\longmapsto \sum_{i=1}^{m} \sum_{j=1}^{n} (a_i \circ a_j') \otimes (b_i \circ b_j'). \end{aligned}$$

Insbesondere gilt für $a, a' \in A$, $b, b' \in B$ und die reinen Tensoren $a \otimes b$, $a' \otimes b'$, dass $(a \otimes b) \bullet (a' \otimes b') = (a \circ a') \otimes (b \circ b')$.

Beweis: WIr haben zu zeigen, dass die Multiplikation „$\circ$“ auf $A \otimes B$ wohldefiniert ist. Dass diese Verknüpfung $A \otimes B$ zu einer K-Algebra macht, ist eine Standardrechnung, wobei man auf die Algebreneigenschaften von A und B zurückführt.

Wir erinnern uns daran, dass $A \otimes B$ per Konstruktion der Vektorraum F/R ist, wobei $F = \mathrm{Abb}_0(A \times B)$, für den $\{f_{(a,b)} \mid a \in A, b \in B\}$ eine Basis bildet, und $R \subseteq F$ der Unterraum ist, der von Elementen der Form $f_{(ra_1+a_2,sb_1+b_2)} - rsf_{(a_1,b_1)} - rf_{(a_1,b_2)} - sf_{(a_2,b_1)} - f_{(a_2,b_2)}$ erzeugt wird. Seien $r, s \in K$, $a, a_1, a_2 \in A$ und $b, b_1, b_2 \in B$. Dann haben wir

$$\begin{aligned}
&(ra_1 + a_2) \otimes (sb_1 + b_2) \bullet (a \otimes b) \\
&\quad = (ra_1 + a_2) \circ a \otimes (sb_1 + b_2) \circ b \\
&\quad = rsa_1 \circ a \otimes b_1 \circ b + ra_1 \circ a \otimes b_2 \circ b + a_2 \circ a \otimes sb_1 \circ b + a_2 \circ a \otimes b_2 \circ b \\
&\quad = (rsa_1 \otimes b_1 + ra_1 \otimes b_2 + a_2 \otimes sb_1 + a_2 \otimes b_2) \bullet (a \otimes b). \qquad \square
\end{aligned}$$

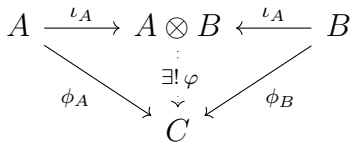
Proposition VII.4.15: *Seien K ein Körper, A und B kommutative unitäre K-Algebren.*

(i) *Die Abbildungen*

$$\begin{aligned}
\iota_A \colon A &\longrightarrow A \otimes B, \qquad a \longmapsto a \otimes 1 \\
\iota_B \colon B &\longrightarrow A \otimes B, \qquad b \longmapsto 1 \otimes b
\end{aligned}$$

sind K-Algebren-Homomorphismen.

(ii) *Das Tensorprodukt $A \otimes B$ hat die folgende universelle Abbildungseigenschaft: Für jede kommutative unitäre K-Algebra C und K-Algebren-Homomorphismen $\phi_A \colon A \to C$, $\phi_B \colon B \to C$ gibt es genau einen K-Algebren-Homomorphismus $\varphi \colon A \otimes B \to C$, der das folgende Diagramm kommutativ macht:*



Beweis: (i) Für a_1, a_2 aus A gilt

$$\iota_A(a_1 \circ a_2) = a_1 \circ a_2 \otimes 1 = (a_1 \otimes 1) \bullet (a_2 \otimes 1) = \iota_A(a_1) \bullet \iota_A(a_2),$$

ferner ist $\iota_A(1) = 1 \otimes 1$ die Eins von $A \otimes B$. Die weiteren Eigenschaften folgen aus der Bilinearität des Tensorprodukts.

(ii) Die Abbildung

$$\beta\colon A \times B \longrightarrow C, \qquad (a,b) \longmapsto \phi_A(a) \circ \phi_A(b)$$

ist bilinear, was uns einen eindeutigen Homomorphismus $\varphi\colon A \otimes B \to C$ von K-Vektorräumen liefert, der $\varphi(a \otimes b) = \phi_A(a) \circ \phi_B(b)$ erfüllt. Dieses φ ist sogar ein K-Algebren-Homomorphismus, denn für $a_1, a_2 \in A$ und $b_1, b_2 \in B$ gilt

$$\begin{aligned}
\varphi((a_1 \otimes b_1) \bullet (a_2 \otimes b_2)) &= \varphi(a_1 \circ a_2 \otimes b_1 \circ b_2) \\
&= \phi_A(a_1 \circ a_2) \circ \phi_A(b_1 \circ b_2) \\
&= \phi_A(a_1) \circ \phi_A(a_2) \circ \phi_B(b_1) \circ \phi_B(b_2) \\
&= \phi_A(a_1) \circ \phi_B(b_1) \circ \phi_A(a_2) \circ \phi_B(b_2) \\
&= \varphi(a_1 \otimes b_1) \circ \varphi(a_2 \otimes b_2). \qquad \square
\end{aligned}$$

Möchte man auch nicht-kommutative K-Algebren betrachten, so fordert man bei der universellen Abbildungseigenschaft, dass für $a \in A$ und $b \in B$ gilt: $\phi_A(a) \circ \phi_B(b) = \phi_B(b) \circ \phi_A(a)$.

Kapitel VIII.

Euklidische und unitäre Vektorräume

Für eine quadratische Gleichung der Form $ax^2 + bx + c = 0$ mit $a \neq 0$ sind die Lösungen, wenn sie existieren, gegeben durch

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ax}}{2a}.$$

Die Zahl $D := b^2 - 4ac$ heißt Diskriminante. An der Diskriminante können wir ablesen, wie viele Lösungen es gibt: Ist $D > 0$, dann gibt es zwei Lösungen; ist $D = 0$, dann gibt es eine Lösung und ist $D < 0$, dann gibt es keine Lösung.

1. Skalarprodukte

In diesem Abschnitt möchten wir Skalarprodukte erneut betrachten, die Definition auf $\mathbb{C}$-Vektorräume erweitern und verstehen, welche zusätzlichen Informationen diese zur Vektorraumstruktur liefern. Zum Beispiel erlaubt ein Skalarprodukt die Messung von Längen. Von immenser Wichtigkeit ist auch das Konzept der Orthogonalität, was mit gewissen Optimalitätseigenschaften einhergeht.

Sowohl in der reinen Mathematik als auch in der angewandten Mathematik und in anderen Wissenschaften spielen Skalarprodukte an vielen Stellen eine prominente Rolle.

Erinnerung: Seien V ein $\mathbb{R}$-Vektorraum und $\beta\colon V \times V \to \mathbb{R}$ eine Bilinearform. Gilt für alle $v, w \in V$, dass $\beta(v, w) = \beta(w, v)$, so heißt β symmetrisch. Gilt für alle $v \in V - \{\mathbf{0}\}$, dass $\beta(v, v) > 0$, dann heißt β positiv definit. Eine symmetrische positiv definite Bilinearform auf V heißt Skalarprodukt auf V. Üblicherweise schreibt man $\langle v, w\rangle := \beta(v, w)$.

Insbesondere ermöglicht die positive Definitheit die gewünschte Definition der induzierten Norm.

Für die obige Definition möchten wir ein Analogon auf $\mathbb{C}$-Vektorräumen geben. Um von positiver Definitheit sprechen zu können, müssen wir aber die Eigenschaft, symmetrisch zu sein, durch etwas ersetzen, was sicherstellt, dass $\beta(v,v)$ für jedes $v \in V$ reell ist.

Erinnerung: Die Abbildung

$$c\colon \mathbb{C} \longrightarrow \mathbb{C}, \qquad z = a + b\mathrm{i} \longmapsto \bar{z} := a - b\mathrm{i}$$

heißt *komplexe Konjugation*. Es handelt sich um einen Körperhomomorphismus. Der komplexe Betrag $|\cdot|\colon \mathbb{C} \to \mathbb{C}$ ist definiert durch $|z| := (z\bar{z})^{1/2}$. Es bezeichne $h\colon \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, $(z, z') \mapsto z\bar{z}'$. Für $\lambda \in \mathbb{C}$ und $z_1, z_2, z \in \mathbb{C}$ gilt

$$\begin{aligned} h(z_1 + \lambda z_2, z) &= (z_1 + \lambda z_2)\bar{z} = z_1\bar{z} + \lambda z_2\bar{z} = h(z_1 z) + \lambda h(z_2, z), \\ h(z, z_1 + \lambda z_2) &= z\overline{(z_1 + \lambda z_2)} = z\bar{z}_1 + \bar{\lambda} z\bar{z}_2 = h(z, z_1) + \bar{\lambda} h(z, z_2), \\ h(z', z) &= z'\bar{z} = \overline{\bar{z}'z} = \overline{h(z, z')}. \end{aligned}$$

Definition VIII.1.1: Seien V ein $\mathbb{C}$-Vektorraum und $s\colon V \times V \to \mathbb{C}$ eine Abbildung. Gilt für alle $\lambda \in \mathbb{C}$, $u_1, u_2, u, v_1, v_2, v \in V$, dass

(i) $s(\lambda u_1 + u_2, v) = \lambda s(u_1, v) + s(u_2, v)$,

(ii) $s(u, \lambda v_1 + v_2) = \bar{\lambda} s(u, v_1) + s(u, v_2)$,

dann heißt s eine „Sesquilinearform“.[1]

Ist $h\colon V \times V \to \mathbb{C}$ eine Sesquilinearform und gilt für alle $v, w \in V$, dass $h(v,w) = \overline{h(w,v)}$, dann heißt h eine *hermitesche Form*.[2]

Beispiel VIII.1.2: (i) Die Abbildung

$$h\colon \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \qquad x^t\bar{y} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_n \end{pmatrix} = \sum_{i=1}^{n} x_i \bar{y}_i$$

ist eine hermitesche Form.

[1] „Sesqui“ ist das lateinische Wort für „anderthalb“, eine Sesquilinearform ist also „anderthalbfach linear“.

[2] Die hermitsche Form verdankt ihren Namen dem französischen Mathematiker Charles Hermite (1822-1901).

(ii) Ist $A \in \mathbb{C}^{n\times n}$ und bezeichnet

$$h_A\colon \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \qquad (x,y) \longmapsto x^t A\bar{y},$$

dann ist h_A eine Sesquilinearform. Genau dann ist h_A hermitesch, wenn $\bar{A}^t = A$.

Definition VIII.1.3 (Adjungierte Matrix): Für $A \in \mathbb{C}^{n\times n}$ heißt $A^* := A^H := \bar{A}^t$ die *adjungierte Matrix zu* A. Gilt $A = A^*$, dann heißt A *hermitesch*.

Proposition VIII.1.4:

(i) *Ist* $h\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ *eine Sesquilinearform, dann ist* $h = h_G$ *für die Matrix* $G = (g_{i,j})$ *mit* $g_{i,j} = h(e_i, e_j)$.

(ii) *Ist* V *ein endlichdimensionaler Vektorraum über* $\mathbb{C}$ *mit geordneter Basis* $B = (b_1, \ldots, b_n)$, *ist* $h\colon V \times V \to \mathbb{C}$ *eine Sesquilinearform und ist die Matrix* $G = (g_{i,j})$ *definiert durch* $g_{i,j} = h(b_i, b_j)$, *dann gilt für alle* v, w *in* V, *dass* $h(v,w) = D_B(v)^t G \overline{D_B(w)}$.

(iii) *Ist* $B' = (b'_1, \ldots, b'_n)$ *eine weitere Basis von* V *und ist* G' *die Matrix* $G' = (g'_{i,j})$ *definiert durch* $g_{i,j} = h(b'_i, b'_j)$, *dann gilt* $G' = D^t_{B,B'} G \bar{D}_{B,B'}$.

Bemerkung VIII.1.5: Seien V ein $\mathbb{C}$-Vektorraum und $h\colon V \times V \to \mathbb{C}$ eine hermitesche Form auf V, dann gilt für alle v in V, dass $h(v,v) = \overline{h(v,v)}$. Insbesondere ist $h(v,v)$ eine relle Zahl.

Definition VIII.1.6: Seien V ein $\mathbb{C}$-Vektorraum und $h\colon V \times V \to \mathbb{C}$ eine hermitesche Form auf V.

(i) Gilt für alle $v \in V - \{0\}$, dass $h(v,v) > 0$, dann heißt h *positiv definit*.

(ii) Eine positiv definite hermitesche Sesquilinearform $h\colon V \times V \to \mathbb{C}$ heißt *Skalarprodukt auf* V. In diesem Fall schreibt man für gewöhnlich $\langle v, w\rangle$ für $h(v,w)$.

Beispiel VIII.1.7 (Standard-Skalarprodukt): Auf $\mathbb{C}^n$ ist die Form

$$h\colon \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \qquad (x,y) \longmapsto x^t\bar{y} = \sum_{i=1}^{n} x_i\bar{y}_i$$

ein Skalarprodukt, das sogenannte *Standard-Skalarprodukt auf* $\mathbb{C}^n$.

Üblicherweise schreibt man $\mathbb{K}$ stellvertretend für den Körper der reellen Zahlen $\mathbb{R}$ oder den Körper der komplexen Zahlen $\mathbb{C}$.

Definition VIII.1.8 (Prähilbertraum, euklidischer- oder unitärer Vektorraum): Ein $\mathbb{K}$-Vektorraum mit Skalaprodukt heißt *Prähilbertraum*. Ein $\mathbb{R}$-Vektorraum mit Skalarprodukt heißt *euklidischer Vektorraum*. Ein $\mathbb{C}$-Vektorraum mit Skalaprodukt heißt *unitärer Vektorraum*.

Manchmal wird in der Literatur für euklidische- oder unitäre Vektorräume endlichdimensionalität verlangt.

Definition VIII.1.9 (Norm, Länge, Abstand): Sei $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum.

(i) Für $v \in V$ heißt $\|v\| := \langle v, v\rangle^{1/2}$ die *Norm* oder *Länge von* v.

(ii) Für $v, w \in V$ heißt $d(v,w) := \|v - w\|$ *Abstand von* v *und* w. Die Abbildung
$$d\colon V \times V \longrightarrow \mathbb{R}_{\geq 0}, \qquad (v,w) \longmapsto d(v,w)$$
heißt *Metrik zu* $\langle\cdot,\cdot\rangle$.

Satz 30 (Cauchy-Schwarz'sche Ungleichung): *Sei* $(V, \langle\cdot,\cdot\rangle)$ *ein Prähilbertraum.*

(i) *Für* $v, w \in V$ *gilt* $|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$*, und somit* $|\langle v, w\rangle| \leq \|v\|\|w\|$.

(ii) *Gleichheit in (i) gilt genau dann, wenn* v *und* w *linear abhängig sind.*

Beweis: (i) Für $w = \mathbf{0}$ stimmen beide Behauptungen. Wir dürfen also annehmen, dass $w \neq \mathbf{0}$. Wir betrachten zunächst den Fall „$\langle v, w\rangle \in \mathbb{R}$“ und definieren $f\colon \mathbb{R} \to \mathbb{R}$ durch $\lambda \mapsto \|v + \lambda w\|^2$. Dann ist

$$\begin{aligned} f(\lambda) = \langle v + \lambda w, v + \lambda w\rangle &= \langle v, v\rangle + \lambda\langle w, v\rangle + \lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle \\ &= \langle v, v\rangle + 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle \end{aligned}$$

Die Zahlen $a := \langle v, v\rangle$, $b := 2\langle v, w\rangle$ und $c := \langle w, w\rangle$ sind reelle Zahlen, außerdem ist $a \neq 0$ per Voraussetzung. Das heißt, f ist in Wahrheit eine quadratische Funktion – mit Schulwissen erkennt man, dass der zugehörige Graph eine „nach oben geöffnete Parabel“ ist. Per Definition der Funktion f gilt für alle $\lambda \in \mathbb{R}$, dass $f(\lambda) \geq 0$, d. h. f hat höchstens eine Nullstelle, was äquivalent zu „$D = b^2 - 4ac \leq 0$“ ist. Ausgeschrieben erhalten wir also

$$4|\langle v, w\rangle|^2 - 4\langle w, w\rangle\langle v, v\rangle \leq 0,$$

wobei die Betragsstriche um $\langle v, w\rangle$ keine Rolle spielen, da $\langle v, w\rangle$ nach Voraussetzung eine reelle Zahl ist. Umstellen liefert die Behauptung.

Nun betrachten wir den allgemeinen Fall „$\langle v, w\rangle \in \mathbb{C}^\times$“. Die komplexe Zahl $\alpha := \langle v, w\rangle / |\langle v, w\rangle|$ hat Betrag Eins. Wir setzen $v' := \alpha^{-1} v$ und erhalten daraus

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

Nun können wir schreiben $|\langle v, w\rangle|^2 = |\langle \alpha v', w\rangle|^2 = |\alpha|^2 |\langle v', w\rangle|^2$. Weil $\langle v', w\rangle$ nach der obigen Gleichung eine reelle Zahl ist, liefert der erste Schritt, dass

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

wie gewünscht.

(ii) Im ersten Schritt von (i) gilt Gleichheit in der Cauchy-Schwarz'schen Ungleichung genau dann, wenn $D = 0$ oder $w = \mathbf{0}$. Die Diskriminante ist Null genau dann, wenn f eine Nullstelle λ hat, was wegen der positiven Definitheit bedeutet, dass $v = -\lambda w$. In beiden Fällen sind v und w linear abhängig.

Im zweiten Schritt von (i) gilt Gleichheit in der Cauchy-Schwarz'schen Ungleichung genau dann, wenn v' und w linear abhängig sind. Wegen der Definition von v' als Vielfaches von v gilt Gleichheit also genau dann, wenn v und w linear abhängig sind. □

Korollar VIII.1.10 (Winkeldefinition): *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum. Dann gilt:*

(i) *Für alle $v, w \in V - \{0\}$ ist $-1 \leq \langle v, w\rangle / (\|v\|\|w\|) \leq 1$, d. h. es gibt genau ein $\alpha \in [0, \pi]$ mit $\cos(\alpha) = \langle v, w\rangle / (\|v\|\|w\|)$. Wir schreiben $\measuredangle(v, w) := \alpha$.*

(ii) *Zwei Vektoren $v, w \in V$ sind orthogonal genau dann, wenn $\langle v, w\rangle = 0$, also wenn $\alpha = \measuredangle(v, w) = \pi/2$.*

Proposition VIII.1.11: *Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und*

$$\|\cdot\| \colon V \longrightarrow \mathbb{R}, \qquad v \longmapsto \|v\| := \langle v, v\rangle^{1/2}.$$

Die Abbildung $\|\cdot\|$ hat folgende Eigenschaften:

(i) *Für alle $v \in V$ ist $\|v\| \geq 0$ und $\|v\| = 0$ gilt genau dann, wenn $v = \mathbf{0}$.*[3]

(ii) *Für alle $v \in V$ und $\lambda \in \mathbb{K}$ ist $\|\lambda v\| = |\lambda| \|v\|$.*[4]

[3] Positive Definitheit
[4] Homogenität

(iii) *Für alle* $v, w \in V$ *gilt* $\|v + w\| \leq \|v\| + \|w\|$.[5]

Beweis: Aussagen (i) und (ii) kann man direkt nachrechnen. Zu (iii): Es gilt

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

Wegen der Monotonie der Wurzel folgt die Behauptung. □

Definition VIII.1.12: Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $N\colon V \to \mathbb{R}$ mit den Eigenschaften (i), (ii) und (iii) aus Proposition VIII.1.11 heißt *Norm auf* V. Das Tupel $(V, \|\cdot\|)$ heißt *normierter Vektorraum über* $\mathbb{K}$ oder kurz *normierter Vektorraum*.

Proposition VIII.1.13: *Seien* $(V, \|\cdot\|)$ *ein normierter Vektorraum und*

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

Die Abbildung d *hat folgende Eigenschaften:*

(i) *Für alle* $v, w \in V$ *ist* $d(v, w) \geq 0$ *und* $d(v, w) = 0$ *genau dann, wenn* $v = w$.[6]

(ii) *Für alle* $v, w \in V$ *ist* $d(v, w) = d(w, v)$.

(iii) *Für alle* $u, v, w \in V$ *ist* $d(u, v) + d(v, w) \geq d(u, w)$.[7]

Beweis: Aussage (i) folgt aus der positiven Definitheit der Norm und Aussage (ii) folgt aus der Homogenität der Norm. Für (iii) müssen wir nur einmal die Dreiecksungleichung für die Norm verwenden; für $u, v, w \in V$ ist nämlich

$$d(u, v) + d(v, w) = \|u - v\| + \|v - w\| \geq \|u - v + v - w\| = \|u - w\| = d(u, w)$$

wie gewünscht. □

Definition VIII.1.14: Sei X eine Menge. Eine Abbildung $d\colon X \times X \to \mathbb{R}$ mit den Eigenschaften (i), (ii), (iii) aus Proposition VIII.1.13, heißt *Metrik auf* X.

Achtung: Nicht jede Norm kommt von einem Skalarprodukt und nicht jede Metrik wird von einer Norm induziert!

[5] Dreiecksungleichung
[6] Positive Definitheit
[7] Dreiecksungleichung

2. Orthogonale und unitäre Endomorphismen

In diesem Abschnitt möchten wir die strukturerhaltenden Abbildungen eines Prähilberträums $(V, \langle\cdot,\cdot\rangle)$ kennenlernen, das heißt wir suchen Abbildungen $\phi\colon V \to V$, die die Vektorraumstruktur erhalten (solche nennt man linear) die gleichzeitig das Skalarprodukt erhalten, d. h. für alle $v, w \in V$ soll gelten: $\langle v, w\rangle = \langle\phi(v), \phi(w)\rangle$ – solche Abbildungen erhalten dann auch die Längen.

Definition VIII.2.1: Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und $\phi\colon V \to V$ ein Endomorphismus. Ist $\mathbb{K} = \mathbb{R}$ und gilt für alle $v, w \in V$, dass $\langle\phi(v), \phi(w)\rangle = \langle v, w\rangle$, dann heißt ϕ *orthogonal bezüglich* $\langle\cdot,\cdot\rangle$ oder kurz *orthogonal*. Ist $\mathbb{K} = \mathbb{C}$ und gilt für alle $v, w \in V$, dass $\langle\phi(v), \phi(w)\rangle = \langle v, w\rangle$, dann heißt ϕ *unitär bezüglich* $\langle\cdot,\cdot\rangle$ oder kurz *unitär*.

Bemerkung VIII.2.2: Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und $\phi \in \mathrm{End}(V)$ orthogonal beziehungsweise unitär.

(i) Für alle $v \in V$ gilt $\|\phi(v)\| = \|v\|$. Abbildungen mit dieser Eigenschaft nennt man *Isometrie*.

(ii) Seien $v, w \in V$. Genau dann gilt $\langle v, w\rangle = 0$, wenn $\langle\phi(v), \phi(w)\rangle = 0$. Das heißt, orthogonale respektive unitäre Abbildungen erhalten die Orthogonalität.

(iii) Isometrien sind injektiv. Ist nämlich $v \in V$ mit $\phi(v) = \mathbf{0}$, dann ist $\|\phi(v)\| = 0$, d. h. $\|v\| = 0$ und wegen den Eigenschaften der Norm heißt das $v = \mathbf{0}$.

(iv) Ist ϕ regulär, dann ist auch ϕ^{-1} orthogonal beziehungsweise unitär.

(v) Kompositionen orthogonaler beziehungsweise unitärer Abbildungen sind orthogonal beziehungsweise unitär. Genauer: Ist auch $\psi\colon V \to V$ orthogonal beziehungsweise unitär, dann ist $\phi \circ \psi$ und $\psi \circ \phi$ orthogonal beziehungsweise unitär.

Bemerkung VIII.2.3: Seien nun $V = \mathbb{K}^n$, $A \in \mathbb{K}^{n\times n}$ und $f_A = \phi\colon \mathbb{K}^n \to \mathbb{K}^n$, $x \mapsto Ax$ und $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt auf $\mathbb{K}^n$. Genau dann ist f_A orthogonal beziehungsweise unitär, wenn für alle $x, y \in \mathbb{K}^n$ gilt:

$$x^t A^t \bar{A}\bar{y} = (Ax)^t\overline{Ay} = \langle Ax, Ay\rangle = \langle x, y\rangle = x^t\bar{y}.$$

Das ist genau dann der Fall, wenn $A^t\bar{A} = I_n$, was äquivalent ist zu $A^*A = I_n$.

Definition VIII.2.4 (Orthogonale und unitäre Matrizen):

(i) Eine Matrix $A \in \mathbb{R}^{n\times n}$ mit $A^tA = I_n$ heißt *orthogonal*.

(ii) Eine Matrix $A \in \mathbb{C}^{n\times n}$ mit $A^*A = I_n$ heißt *unitär*.

Insbesondere sind orthogonale und unitäre Matrizen regulär. Ist A orthogonal, dann ist $A^{-1} = A^t$ und ist A unitär, dann ist $A^{-1} = A^*$.

Proposition VIII.2.5 (Charakterisierung orthogonaler und unitärer Matrizen): *Die folgenden Aussagen sind äquivalent:*

(i) $A \in \mathbb{K}^{n\times n}$ *is orthogonal beziehungsweise unitär,*

(ii) *Die Spaltenvektoren von A bilden eine Orthonormalbasis des $\mathbb{K}^n$ bezüglich des Standardskalarprodukts,*

(iii) *Die Zeilenvektoren von A bilden eine Orthonormalbasis des $\mathbb{K}^n$ bezüglich des Standardskalarprodukts.*

Seien $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler Prähilbertraum, G die Gram-Matrix des Skalarprodukts bezüglich der geordneten Basis B. Genau dann ist ein Endomorphismus $\phi \in \mathrm{End}(V)$ orthogonal beziehungsweise unitär, wenn

$$D_{B,B}(\phi)^t G \overline{D_{B,B}(\phi)} = G.$$

Ist B eine Orthonormalbasis, dann ist $G = I_n$ und das heißt, dass ϕ orthogonal beziehungsweise unitär ist genau dann, wenn $D_{B,B}(\phi)$ orthogonal beziehungsweise unitär ist. Insbesondere folgt für eine orthogonale oder unitäre Matrix S und eine Matrix $B \in \mathbb{K}^n$: Genau dann ist B orthogonal beziehungsweise unitär, wenn SBS^{-1} orthogonal beziehungsweise unitär ist.

Ist $\phi \in \mathrm{End}(V)$ orthogonal beziehungsweise unitär und ist $\lambda \in \mathrm{Spec}(\phi)$, dann ist $|\lambda| = 1$. Insbesondere gilt für orthogonale ϕ beziehungsweise A, dass $\mathrm{Spec}(\phi) \subseteq \{\pm 1\}$ *beziehungsweise* $\mathrm{Spec}(A) \subseteq \{\pm 1\}$.

Beweis: Wir verwenden die Äquivalenz „$AA^* = I_n \Leftrightarrow A^*A = I_n \Leftrightarrow A^t\bar{A} = I_n$“ aus dem vorangegangenen Beweis. Die dritte Gleichheit gilt genau dann, wenn die Spalten von A eine Orthonormalbasis des $\mathbb{K}^n$ bezüglich des Standardskalarprodukts bilden. Die zweite Gleichheit ist äquivalent dazu, dass A orthogonal beziehungsweise unitär ist. Schließlich gilt die erste Gleichheit genau dann, wenn die Zeilen von A eine Orthonormalbasis des $\mathbb{K}^n$ bezüglich des Standardskalarprodukts bilden.

Wir erinnern uns, dass die Abbildungsmatrix charakteriert ist durch „Für alle $v \in V$ ist $D_B(\phi(v)) = D_{B,B}(\phi)D_B(v)$“. Der Endomorphismus ϕ is orthogonal beziehungsweise unitär genau dann, wenn für alle $v, w \in V$ gilt: $\langle\phi(v), \phi(w)\rangle =$

$\langle v, w\rangle$. Wir haben also die folgenden Äquivalenzen für alle $v, w \in V$:

$$\begin{aligned} D_B(\phi(v))^t G\overline{D_B(\phi(w))} &= D_B(v) G\overline{D_B(w)} \\ \iff (D_{B,B}(\phi)D_B(v))^t G\overline{D_{B,B}(\phi)D_B(w)} &= D_B(v) G\overline{D_B(w)} \\ \iff D_B(v)^t D_{B,B}(\phi)^t G\overline{D_{B,B}(\phi)D_B(w)} &= D_B(v) G\overline{D_B(w)} \\ \iff D_{B,B}(\phi)^t G\overline{D_{B,B}(\phi)}^t &= G. \end{aligned}$$

Wir zeigen die Aussage für orthogonale beziehungsweise unitäre Endomorphismen. Sei ϕ ein solcher und $\lambda \in \mathrm{Spec}(\phi)$ ein Eigenwert. Dann gibt es $v \in V - \{\mathbf{0}\}$, sodass $\phi(v) = \lambda v$, d. h. $\|\phi(v)\| = \|\lambda v\|$, woraus wir wegen $\|v\| \neq 0$ folgern können, dass $|\lambda| = \|\phi(v)\|/\|v\| = 1$. □

Bemerkung VIII.2.6: Sei B eine geordnete Basis des $\mathbb{K}^n$, dann haben wir folgende Äquivalenzen:

$$\begin{aligned} B \text{ ist eine Orthonormalbasis} &\iff D_{E,B} \text{ ist orthogonal beziehungsweise unitär} \\ &\iff D_{B,E} \text{ ist orthogonal beziehungsweise unitär} \end{aligned}$$

Das folgt aus Proposition 2.5 (i), Bemerkung 2.2 (iv) und Bemerkung 2.3.

Definition VIII.2.7 (Orthogonale und unitäre Gruppe): Sei n eine natürliche Zahl. Dann ist

$$\mathrm{O}(n) := \{A \in \mathbb{R}^{n\times n} \mid A \text{ orthogonal}\}$$

eine Untergruppe der $\mathrm{Gl}_n(\mathbb{R})$ und heißt *orthogonale Gruppe*. Die Teilmenge

$$\mathrm{SO}(n) := \{A \in \mathbb{R}^{n\times n} \mid A \in \mathrm{O}(n), \det(A) = 1\} \subseteq \mathrm{O}(n)$$

ist eine Untergruppe der speziellen linearen Gruppe $\mathrm{Sl}_n(\mathbb{R})$ und heißt *spezielle orthogonale Gruppe*.

Die Menge

$$\mathrm{U}(n) := \{A \in \mathbb{C}^{n\times n} \mid A \text{ unitär}\}$$

ist eine Untergruppe der $\mathrm{Gl}_n(\mathbb{C})$ und heißt *unitäre Gruppe*. Die Teilmenge

$$\mathrm{SU}(n) := \{A \in \mathbb{C}^{n\times n} \mid A \in \mathrm{U}(n), \det(A) = 1\} \subseteq \mathrm{U}(n)$$

ist eine Untergruppe der speziellen linearen Gruppe $\mathrm{Sl}_n(\mathbb{C})$ und heißt *spezielle unitäre Gruppe*.

Beispiel VIII.2.8: (i) Für jede reelle Zahl φ ist die Matrix

$$A := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

orthogonal, denn wir haben

$$AAt = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Aufgefasst als Darstellungsmatrix einer linearen Abbildung bezüglich der Standardbasen ist A eine Drehung um den Winkel φ.

(ii) Auch die Matrix

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

ist orthogonal. Aufgefasst als Darstellungsmatrix einer linearen Abbildung $\phi_2 \colon \mathbb{R}^2 \to \mathbb{R}^2$ bezüglich der Standardbasen beschreibt sie eine „Spiegelung an der x-Achse“.

(iii) Für reelle Zahlen α_1, α_2 ist

$$\begin{pmatrix} \exp(\mathrm{i}\alpha_1) & 0 \\ 0 & \exp(\mathrm{i}\alpha_2) \end{pmatrix}$$

eine unitäre Matrix.

(iv) Für die Basis $B = ((\begin{smallmatrix}1\\0\end{smallmatrix}), (\begin{smallmatrix}1\\1\end{smallmatrix}))$ des $\mathbb{R}^n$ haben wir die Basiswechselmatrizen $D_{E,B} = (\begin{smallmatrix}1&1\\0&1\end{smallmatrix})$ und $D_{B,E} = D_{E,B}^{-1} = (\begin{smallmatrix}1&-1\\0&1\end{smallmatrix})$. Für $A := D_{B,B}(\phi_2)$ mit ϕ_2 wie in (ii) gilt

$$A = D_{B,E} D_{E,E}(\phi_2) D_{E,B} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

und $AA^t = (\begin{smallmatrix}5&-2\\-2&1\end{smallmatrix})$ ist jedenfalls nicht die Einheitsmatrix, d. h. A ist nicht orthogonal. Das ganze geht schief, weil B keine Orthonormalbasis bezüglich des Standardskalarprodukts ist.

(v) Die Eigenschaft eines Endomorphismus, orthogonal zu sein, hängt ehrlich vom gewählten Skalarprodukt ab. Für das Skalarprodukt $\langle\langle \cdot, \cdot \rangle\rangle$, das durch die Gram-Matrix $G = (\begin{smallmatrix}2&1\\1&2\end{smallmatrix})$ gegeben ist, gilt

$$\left\langle\left\langle \phi_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \phi_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\rangle\right\rangle = \left\langle\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}\right\rangle\right\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -1,$$

aber $\langle\langle e_1, e_2 \rangle\rangle = 1$, d. h. ϕ_2 ist nicht orthogonal bezüglich $\langle\langle \cdot, \cdot \rangle\rangle$.

3. Normale Matrizen

Sei $A \in \mathbb{C}^{n\times n}$ gegeben. Wir erinnern uns daran: Ist $AA^* = I_n$, dann heißt A *unitär*. Ist $A = A^*$, dann heißt A *hermitesch*. Die beiden Eigenschaften sehen ähnlich aus, kommen von sehr verschiedenen Eigenschaften her. Das Ziel des Abschnitts wird es sein, zu zeigen, dass hermitesche und unitäre Matrizen diagonalisierbar sind, und dass es Orthonormalbasen aus Eigenvektoren gibt.

Definition VIII.3.1 (Normale Matrix): Sei $A \in \mathbb{C}^{n\times n}$ gegeben. Gilt $AA^* = A^*A$, dann heißt A *normal*.

Insbesondere sind hermitesche und unitäre Matrizen normal.

Bemerkung VIII.3.2: Seien K ein Körper und $A \in K^{n\times n}$. Genau dann ist A diagonalisierbar, wenn es $B \in \mathrm{Gl}_n(K)$ gibt, sodass $A = B\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)B^{-1}$ und das ist äquivalent dazu, dass A eine Basis aus Eigenvektoren hat.

Genauer: Ist $B = (b_1|\ldots|b_n)$, dann ist $(b_1, \ldots, b_n)$ eine Basis aus Eigenvektoren, denn $b_i = Be_i$ für die Standardbasis $e_1, \ldots, e_n$.

In der speziellen Situation, dass $K = \mathbb{K}$ ist und wir das Standardskalarprodukt zur Verfügung haben, gilt: Die Matrix A ist diagonalisierbar durch eine orthogonale beziehungsweise unitäre Matrix (d. h. es gibt $S \in \mathrm{O}(n)$ beziehungsweise $S \in \mathrm{U}(n)$ mit $A = S\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)S^{-1}$) genau dann, wenn es eine Orthonormalbasis von $\mathbb{K}^n$ gibt, die aus Eigenvektoren von A besteht.

Satz 31 (Spektralsatz für normale Matrizen): *Sei $A \in \mathbb{C}^{n\times n}$ gegeben. Genau dann ist A normal, wenn es einen unitären Basiswechsel $S \in \mathrm{U}(n)$ und Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ gibt, sodass $A = S\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)S^{-1}$.*

Der Beweis dieses wichtigen Satzes erfordert einiges an Vorbereitung.

Notation VIII.3.3: Seien K ein Körper, $A \in K^{n\times n}$ und $U \subseteq K^n$ ein Unterraum. Dann schreiben wir $AU := \{Ax \mid x \in U\}$.

Proposition VIII.3.4 (Invariante Unterräume für kommutierende Matrizen): *Seien K ein Körper und $A, B \in K^{n\times n}$ mit $AB = BA$. Ist $\lambda \in \mathrm{Spec}(A)$, dann ist $B\,\mathrm{Eig}(A, \lambda) \subseteq \mathrm{Eig}(A, \lambda)$, d. h. $\mathrm{Eig}(A, \lambda)$ ist B-invariant.*

Im speziellen Fall $K = \mathbb{C}$ ist ferner $B^*\,\mathrm{Eig}(A, \lambda)^\perp \subseteq \mathrm{Eig}(A, \lambda)^\perp$ bezüglich des Standardskalarprodukts auf $\mathbb{C}^n$.

Insbesondere erhalten wir für eine normale Matrix A eine A-invariante Zerlegung $\mathbb{C}^n = \mathrm{Eig}(A, \lambda) \oplus \mathrm{Eig}(A, \lambda)^\perp$.

Beweis: (i) Sei x ein Eigenvektor von A zum Eigenwerte λ. Dann gilt $ABx = BAX = B\lambda x = \lambda Bx$, d. h. auch Bx ist ein Eigenvektor von A zum Eigenwert λ.

(ii) Seien $x \in \mathrm{Eig}(A,\lambda)^\perp$ und $y \in \mathrm{Eig}(A,\lambda)$. Dann haben wir

$$\langle B^*x, y\rangle = (B^*x^*)^t\bar{y} = x^t\bar{B}\bar{y} = x^t\overline{By} = \langle x, By\rangle = 0,$$

da $By \in \mathrm{Eig}(A,\lambda)$ nach (i) und $x \in \mathrm{Eig}(A,\lambda)^\perp$. Damit gehört B^*x zu $\mathrm{Eig}(A,\lambda)$.

(iii) zeigt man genau wie die anderen Aussagen unter Verwendung von $AA = AA$ und $AA^* = A^*A$. □

Proposition VIII.3.5 (Von A und A^* erzeugte $\mathbb{C}$-Algebra): *Sei $A \in \mathbb{C}^{n\times n}$ eine normale Matrix. Die Menge*

$$\mathfrak{A} := \mathbb{C}[A, A^*] := \left\{ \sum_{i=0}^{m}\sum_{j=0}^{m}\sum_{j=1}^{k} b_{i,j}A^i(A^*)^j : m,k \in \mathbb{N}, b_{i,j} \in \mathbb{C} \right\}$$

heißt die von A und A^* erzeugte Unteralgebra von $C^{n\times n}$ und es gilt:

(i) *$\mathfrak{A}$ ist eine Unteralgebra von $\mathbb{C}^{n\times n}$,*

(ii) *$\mathfrak{A}$ ist kommutativ.*

(iii) *$\mathfrak{A}$ ist abgeschlossen unter Adjunktion, d. h. für $B \in \mathfrak{A}$ ist auch $B^* \in \mathfrak{A}$.*

Beweis: Zu (i): Dass $\mathfrak{A}$ ein Untervektorraum von $\mathbb{C}^{n\times n}$ ist, ist klar. Bleibt zu zeigen, dass $\mathfrak{A}$ unter Multiplikation von Matrizen abgeschlossen ist. Seien dazu $M_1 = \sum_{i=0}^{m}\sum_{j=0}^{n} b_{i,j}A^i(A^*)^j$ und $M_2 = \sum_{r=0}^{m'}\sum_{s=0}^{k'} b'_{r,s}A^r(A^*)^s$ Elemente von $\mathfrak{A}$. Dann gehört auch

$$M_1M_2 = \sum_{i=0}^{m}\sum_{j=0}^{k}\sum_{r=0}^{m'}\sum_{s=0}^{k'} b_{i,j}b'_{r,s}A^i(A^*)^jA^r(A^*)^s = \sum_{i=0}^{m}\sum_{j=0}^{k}\sum_{r=0}^{m'}\sum_{s=0}^{k'} A^{i+r}(A^*)^{j+s}$$

zu $\mathfrak{A}$ und die letzte Gleichheit gilt, da A normal ist. Damit haben wir auch sofort eingesehen, dass $M_1M_2 = M_2M_1$, d. h. auch (ii) ist gezeigt.

Zu (iii): Die Adjungierte des Elements $M = \sum_{i=0}^{m}\sum_{j=0}^{k} b_{i,j}A^i(A^*)^j$ von $\mathfrak{A}$ ist $M^* = \sum_{i=0}^{m}\sum_{j=0}^{k} \bar{b}_{i,j}A^j(A^*)^i$ und gehört auch zu $\mathfrak{A}$. □

Satz 32 (Simultane Orthonormalbasis): *Sei $\mathfrak{A} \subseteq \mathbb{C}^{n\times n}$ eine kommutative Unteralgebra, die unter Adjunktion abgeschlossen ist. Dann gibt es eine Orthonormalbasis $B = \{b_1, \ldots, b_n\}$ des $\mathbb{C}^n$ bezüglich des Standardskalarprodukts von simultanen Eigenvektoren, d. h. für jedes $A \in \mathfrak{A}$ sind $b_1, \ldots, b_n$ Eigenvektoren von A.*

Beweis: Wir zeigen die Aussage per Induktion nach der Dimension n. Für $n = 1$ ist nichts zu zeigen, die kanonische Basis $B = \{1\}$ leistet offensichtlich das Gewünschte.

Die Aussage gelte nun für eine natürliche Zahl $n-1$ größergleich Eins. Ist $\mathfrak{A} = \mathbb{C}I_n = \{\operatorname{diag}(\lambda, \ldots, \lambda) \mid \lambda \in \mathbb{C}\}$ oder $\mathfrak{A} = \{\mathbf{0}\}$, dann gilt die Behauptung mit der Standardbasis.

Ist $\mathfrak{A}$ keine der beiden oben genannten Algebren, dann gibt es $A_0 \in \mathfrak{A} - \mathbb{C}I_n$. Weil $\mathbb{C}$ algebraisch abgeschlossen ist, hat A_0 einen Eigenwert λ und $\mathbb{C}^n$ zerfällt in $\mathbb{C}^n = \operatorname{Eig}(A, \lambda) \oplus \operatorname{Eig}(A, \lambda)^\perp$. Weil irgendeine andere Matrix $B \in \mathfrak{A}$ mit A kommutiert, ist wegen Proposition VIII.3.4 auch $B \operatorname{Eig}(A, \lambda) \subseteq \operatorname{Eig}(A, \lambda)$ sowie $B^* \operatorname{Eig}(A, \lambda)^\perp \subseteq \operatorname{Eig}(A, \lambda)^\perp$, d. h. sowohl $\operatorname{Eig}(A, \lambda)$ als auch $\operatorname{Eig}(A, \lambda)^\perp$ sind invariant unter allen Elementen von $\mathfrak{A}$.

Da $\dim \operatorname{Eig}(A, \lambda) > 0$ können wir die Induktionsvoraussetzung auf $\operatorname{Eig}(A, \lambda)$ und $\operatorname{Eig}(A, \lambda)^\perp$ anwenden und erhalten Orthonormalbasen B_1 von $\operatorname{Eig}(A, \lambda)$ sowie B_2 von $\operatorname{Eig}(A, \lambda)^\perp$ aus simultanen Eigenvektoren für $\mathfrak{A}$. Die Vereinigung $B := B_1 \cup B_2$ liefert eine Orthonormalbasis von $\mathbb{C}^n$ aus simultanen Eigenvektoren von $\mathfrak{A}$. □

Beweis (von Satz 31): „$\Longrightarrow$“ folgt aus Proposition 2.5 und Satz 32.

„$\Longleftarrow$“: Seien S eine unitäre $n \times n$-Matrix und $\lambda_1, \ldots, \lambda_n$ komplexe Zahlen, sodass $S^*AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Dann ist

$$\begin{aligned} S^*AA^*S &= S^*ASS^*A^*S \\ &= \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \operatorname{diag}(\bar{\lambda}_1, \ldots, \bar{\lambda}_n) = (S^*A^*S)(S^*AS) = S^*A^*AS, \end{aligned}$$

d. h. A ist in der Tat normal. □

4. Die adjungierte Matrix

Definition VIII.4.1: Sei $A \in \mathbb{C}^{n\times m}$ eine Matrix. Dann heißt $A^* \in \mathbb{C}^{m\times n}$, die eintragsweise definiert ist durch $(A^*)_{i,j} = \bar{A}_{j,i}$, die *Adjungierte zu A*.

Für $x \in \mathbb{C}^m$, $y \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times m}$ und die Standardskalarprodukte auf $\mathbb{C}^n$ und $\mathbb{C}^m$ gilt

$$\langle Ax, y\rangle = x^t A^t \bar{y} = x^t \bar{A}^* \bar{y} = x^t \overline{A^* y} = \langle x, A^* y\rangle. \tag{VIII.1}$$

Zu einer linearen Abbildung $\phi\colon V \to W$ wollen wir eine „Adjungierte“ $\phi^*\colon W \to V$ definieren, die die Eigenschaft Gl. (VIII.1) erfüllt. Dazu müssen V und W jeweils mit Skalarprodukten ausgestattet sein. Fur unendlichdimensionale Prähilberträume V und W muss es so eine Adjungierte nicht geben, im endlichdimensionalen Fall ist aber alles in Ordnung.

Sind die Vektorräume endlichdimensional und sind B beziehungsweise C Orthonormalbasen von V beziehungsweise W, dann gilt für die Darstellungsmatrizen:

$$D_{B,C}(\phi^*) = (D_{C,B}(\phi))^*.$$

Für das was folgt, wird der Satz von Riesz für endlichdimensionale Prähilberträume eine entscheidende Rolle spielen. Wir wollen uns deshalb nochmal an seine Aussage erinnern:

Erinnerung: Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und

$$\Theta\colon V \longrightarrow V^*, \qquad w \longmapsto L_w\colon v \mapsto \langle v, w\rangle).$$

Dann ist Θ eine injektivee antilineare Abbildung und falls V endlichdimensional ist, dann ist Θ eine antilineare Bijektion von Vektorräumen.

In Proposition VII.3.11 haben wir die Aussage für anisotrope Bilinearformen eingeführt, der Beweis geht aber genau so für Skalarprodukte über $\mathbb{C}$ durch.

Durch $\langle f, g\rangle := \langle\Theta^{-1}(g), \Theta^{-1}(f)\rangle$ wird V^* zu einem Prähilbertraum und Θ ist dann verträglich mit diesem Skalarprodukt.

Lemma VIII.4.2: *Seien $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und $v_1, v_2 \in V$. Gilt für alle $v \in V$, dass $\langle v_1, v\rangle = \langle v_2, v\rangle$, dann ist $v_1 = v_2$.*

Beweis: Für Vektoren v_1, v_2 mit der gegebenen Eigenschaft ist $L_{v_1} = L_{v_2}$, und wegen der Injektivität von Θ ist $v_1 = v_2$. □

Proposition VIII.4.3 (Eindeutigkeit der Adjungierten): *Sei $\phi\colon V \to W$ eine lineare Abbildung zwischen Prähilberträumen über $\mathbb{K}$. Für jedes $w \in W$ gibt es höchstens ein $\phi^*(w) \in V$, sodass für alle $v \in V$ gilt:*

$$\langle\phi(v), w\rangle = \langle v, \phi^*(w)\rangle. \tag{VIII.2}$$

Beweis: Das ist eine direkte Konsequenz aus Proposition VIII.4.2. □

Definition VIII.4.4 (Die Adjungierte): Seien V und W Prähilberträume über $\mathbb{K}$ und $\phi\colon V \to W$ eine lineare Abbildung. Gibt es für jedes $w \in W$ ein $\phi^*(w) \in V$ mit der Eigenschaft Gl. (VIII.2), dann heißt die Abbildung

$$\phi^*\colon W \longrightarrow V, \qquad w \longmapsto \phi^*(w)$$

die *Adjungierte Abbildung* oder auch *adjungierte Abbildung zu ϕ.*

Proposition VIII.4.5 (Linearität der Adjungierten): *Seien V und W Prähilberträume über $\mathbb{K}$ und $\phi\colon V \to W$ linear. Falls die adjungierte Abbildung $\phi^*\colon W \to V$ existiert, dann ist sie linear.*

Beweis: Für alle $w_1, w_2 \in W$, $v \in V$ und $\alpha \in \mathbb{K}$ gilt:

$$\begin{aligned}\langle v, \phi^*(\alpha w_1 + w_2)\rangle &= \langle \phi(v), \alpha w_1 + w_2\rangle \\ &= \bar{\alpha}\langle \phi(v), w_1\rangle + \langle \phi(v), w_2\rangle \\ &= \bar{\alpha}\langle v, \phi^*(w_1)\rangle + \langle v, \phi^*(w_2)\rangle = \langle v, \alpha\phi^*(w_1) + \phi^*(w_2)\rangle.\end{aligned}$$

Aus Proposition VIII.4.2 folgt jetzt $\phi^*(\alpha w_1 + w_2) = \alpha\phi^*(w_1) + \phi^*(w_2)$. □

Bemerkung VIII.4.6: Es sei $V = \mathbb{K}$. Die Abbildung

$$m\colon V \longrightarrow \mathrm{Hom}(\mathbb{K}, V), \qquad v \longmapsto (c \mapsto cv)$$

ist ein kanonischer Isomorphismus von $\mathbb{K}$-Vektorräumen mit Umkehrabbildung $m^{-1}\colon \mathrm{Hom}(\mathbb{K}, V) \to V$, $f \mapsto f(1)$. Betrachte die Abbildung

$$\Theta\colon V \cong \mathrm{Hom}(\mathbb{K}, V) \longrightarrow \mathrm{Hom}(V, \mathbb{K}) = V^*, \qquad v \longmapsto m(v) \longmapsto m(v)^*,$$

wobei wir $\mathbb{K}$ mit dem Standardskalarprodukt ausstatten. In dieser Situation ist genau $m(v)^* = L_v$ und Θ is die Abbildung aus dem Satz von Riesz, denn für alle $v' \in V$ ist

$$m(v)^*(v') = \langle m(v)^*(v'), 1\rangle = \langle v', m(v)(1)\rangle = \langle v', v\rangle$$

Definition VIII.4.7 (Selbstadjungiert, Normal): Seien V ein $\mathbb{K}$-Prähilbertraum und $\phi \in \mathrm{End}(V)$ ein Endomorphismus, dessen Adjungierte $\phi^*\colon V \to V$ existiere. Gilt $\phi^* = \phi$, dann heißt ϕ *selbstadjungiert*. Ist $\phi \circ \phi^* = \phi^* \circ \phi$, dann heißt ϕ *normal*.

Bemerkung VIII.4.8: Seien V ein endlichdimensionaler $\mathbb{K}$-Prähilbertraum und $\phi \in \mathrm{End}(V)$. Genau dann ist ϕ orthogonal beziehungsweise unitär, wenn ϕ^* existiert und $\phi^* = \phi^{-1}$ gilt.

Beweis: „$\Longrightarrow$“: Für alle $v, w \in V$ ist

$$\langle \phi(v), w\rangle = \langle \phi(v), \phi(\phi^{-1}(w))\rangle = \langle v, \phi^{-1}(w)\rangle,$$

da ϕ per Voraussetzung orthogonal beziehungsweise unitär ist.

„$\Longleftarrow$“: Für alle $v, w \in V$ haben wir $\langle \phi(v), \phi(w)\rangle = \langle v, \phi^*(\phi(w))\rangle = \langle v, w\rangle$, sodass ϕ orthogonal beziehungsweise unitär ist. □

Erinnerung VIII.4.9: Seien V ein $\mathbb{K}$-Prähilbertraum und B eine Basis von V. Genau dann gilt für alle $v, w \in V$, dass $\langle v, w\rangle = \langle D_B(v), D_B(w)\rangle$, wenn B eine Orthonormalbasis ist, denn beim Übersetzen in den $\mathbb{K}^n$ „ersetzen“ wir B durch die Standardbasis und statten $\mathbb{K}^n$ immer mit dem Standardskalarprodukt aus.

Proposition VIII.4.10 (Adjungierte im Endlichdimensionalen): *Seien V und W endlichdimensionale Prähilberträume über $\mathbb{K}$ und $\phi\colon V \to W$ linear. Dann gilt:*

(i) *Die Adjungierte $\phi^*\colon W \to V$ existiert.*

(ii) *Sind B respektive C Orthonormalbasen von V respektive W, dann ist gilt $D_{B,C}(\phi^*) = D_{C,B}(\phi)^*$, d. h. adjungierte Matrix und adjungierte Abbildung passen zusammen.*

Beweis: Wir sind in der Situation

$$\begin{array}{ccccc} V & \xrightarrow{\phi} & W & \xrightarrow{\psi} & V \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_C} & & \downarrow{\scriptstyle D_B} \\ \mathbb{K}^m & \xrightarrow{x\mapsto Ax} & \mathbb{K}^n & \xrightarrow[x\mapsto A^*x]{} & \mathbb{K}^m \end{array}$$

mit $A := D_{B,C}(\phi)$. Es bezeichne $\langle\cdot,\cdot\rangle_E$ beziehungsweise $\langle\cdot,\cdot\rangle_{E'}$ das Standardskalarprodukt auf $\mathbb{K}^m$ beziehungsweise $\mathbb{K}^n$ und definiere $\psi\colon W \to V$ durch $w \mapsto D_B^{-1}(A^* D_C(w))$. Für $v \in V$ und $w \in W$ haben wir dann

$$\begin{aligned} \langle \phi(v), w\rangle &= \langle D_C(\phi(v)), D_C(w)\rangle_{E'} \\ &= \langle A D_B(v), D_C(w)\rangle_{E'} \\ &= \langle D_B(v), A^* D_C(w)\rangle_E = \langle D_B(v), D_B(\psi(w))\rangle_E = \langle v, \psi(w)\rangle, \end{aligned}$$

d. h. ψ ist die Adjungierte von ϕ und insbesondere gilt Behauptung (ii), was den Beweis beschließt. □

Korollar VIII.4.11: *Seien V ein endlichdimensionaler $\mathbb{K}$-Prähilbertraum, $\phi \in \operatorname{End}(V)$ und B eine Orthonormalbasis von V. Dann gilt:*

(i) *Genau dann ist ϕ selbstadjungiert, wenn $D_{B,B}(\phi) = D_{B,B}(\phi)^*$, d. h. wenn $D_{B,B}(\phi)$ symmetrisch beziehungsweise hermitesch ist.*

(ii) *Genau dann ist ϕ normal, wenn $D_{B,B}(\phi)$ normal ist.*

(iii) *Genau dann ist ϕ orthogonal beziehungsweise unitär, wenn $D_{B,B}(\phi)$ orthogonal beziehungsweise unitär ist.*

5. Anwendungen des Spektralsatzes

Bemerkung VIII.5.1: Sei $A \in \mathbb{R}^{n\times n}$ normal, d. h. $A^tA = AA^t$. Sind alle komplexen Eigenwerte von A reell, dann gibt es eine Orthonormalbasis des $\mathbb{R}^n$ aus Eigenvektoren von A.

Beweis: Sei $\mathfrak{A}$ die von A und A^t erzeugte $\mathbb{R}$-Unteralgebra von $\mathbb{R}^{n\times n}$. Nach 32 gibt es eine Orthonormalbasis des $\mathbb{C}^n$ aus simultanen Eigenvektoren für $\mathfrak{A}$. Es gibt also eine unitäre Matrix $S \in \mathrm{U}(n)$, sodass

$$\mathfrak{A}^S := \{SBS^{-1} \mid B \in \mathfrak{A}\} \subseteq \{\mathrm{diag}(\alpha_1, \ldots, \alpha_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{C}\}.$$

Weil aber $\mathfrak{A}^S$ erzeugt wird von SAS^{-1} und SA^tS^{-1}, die nur reelle Diagonaleinträge haben, da alle Eigenwerte von A per Voraussetzung reell sind, ist in Wahrheit $\mathfrak{A}^S \subseteq \{\mathrm{diag}(\alpha_1, \ldots, \alpha_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$.

Nun liefert der Beweis von Satz 31, der über $\mathbb{R}$ genau so durchgeht, die gewünschte Orthonormalbasis aus simultanen Eigenvektoren. □

Beispiel VIII.5.2: Seien φ eine reelle Zahl und

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \in \mathrm{O}(2)$$

die Drehung um φ (in mathematisch positiver Richtung). Das charakteristische Polynom von A ist

$$\chi_A = \det \begin{pmatrix} (\cos\varphi) - X & -\sin\varphi \\ \sin\varphi & (\cos\varphi) - X \end{pmatrix} = X^2 - 2(\cos\varphi)X + 1,$$

die zugehörigen Nullstellen sind $X_{1,2} = \cos\varphi \pm \sqrt{\cos^2\varphi - 1}$. Für $\varphi \notin \mathbb{Z}\pi$ sind diese Nullstellen nicht reell, d. h. für $\varphi \in \mathbb{R} - \mathbb{Z}\pi$ ist A nicht diagonalisierbar.

Satz 33 (Spektralsatz für orthogonale und unitäre Endomorphismen):

(i) *Seien $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler unitärer Raum und $\phi \in \mathrm{End}(V)$ unitär. Dann gibt es eine Orthonormalbasis von V aus Eigenvektoren von ϕ. Für die Eigenwerte λ_i von ϕ gilt $|\lambda_i| = 1$.*

(ii) *Seien $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler euklidischer Raum und $\phi \in \mathrm{End}(V)$ orthogonal. Dann gibt es eine Orthonormalbasis B von V, sodass*

$$D_{B,B}(\phi) = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, A_1, \ldots, A_k)$$

mit Drehmatrizen $A_i = \left(\begin{smallmatrix} \cos(\alpha_i) & -\sin(\alpha_i) \\ \sin(\alpha_i) & \cos(\alpha_i) \end{smallmatrix}\right)$ und reellen Zahlen α_i.

Beispiel VIII.5.3 (Der Spektralsatz in Dimension 2): Seien $V = \mathbb{R}^2$ ausgestattet mit dem Standardskalarprodukt und $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch $x \mapsto Ax$ für eine orthogonale Matrix $A \in \mathrm{O}(2) \subseteq \mathrm{U}(n)$. Aufgefasse als komplexe Matrix liefert 31 erhalten wir eine Orthonormalbasis $B = (b_1, b_2)$ des $\mathbb{C}^2$ bestehend aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \lambda_2 \in \mathbb{C}$, die beide komplexen Betrag 1 haben.

Bezeichne $\phi'\colon \mathbb{C}^2 \to \mathbb{C}^2$ die Abbildung $x \mapsto Ax$. Bezüglich B ist die Darstellungsmatrix von ϕ' sehr einfach, nämlich $D_{B,B}(\phi') = \mathrm{diag}(\lambda_1, \lambda_2)$, und außerdem ist $\phi'|_{\mathbb{R}^2} = \phi$, wobei wir $\mathbb{R}^2$ mit einem Unterraum des $\mathbb{C}^2$ identifizieren.

Gehört λ_1 zu $\mathbb{C}-\mathbb{R}$, dann ist $A\bar{b}_1 = \overline{Ab_1} = \overline{\lambda_1 b_1} = \bar{\lambda}_1 \bar{b}_1$, d. h. $\bar{b}_1$ ist Eigenvektor zum Eigenwert $\bar{\lambda}_1 \neq \lambda_1$. Weil A nur zwei Eigenwerte hat, muss das heißen, dass $\lambda_2 = \bar{\lambda}_1$; wir dürfen also annehmen dass $\bar{b}_1 = b_2$. Wir definieren

$$c_1 := \mathrm{Re}(b_1) = \frac{1}{2}(b_1 + \bar{b}_1) = \frac{1}{2}(b_1 + b_2), \qquad c_2 := \mathrm{Im}(b_1) = \frac{1}{2\mathrm{i}}(b_1 - \bar{b}_1) = \frac{1}{2\mathrm{i}}(b_1 - b_2).$$

Dann ist $C = (\sqrt{2}c_1, \sqrt{2}c_2)$ eine Orthonormalbasis des $\mathbb{C}^2$ und

$$D_{B,C} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix}, \qquad D_{C,B} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix}.$$

Außerdem ist C sogar eine Orthonormalbasis des $\mathbb{R}^2$, d. h. die Einträge von c_1 und c_2 sind reell. Wir haben

$$\begin{aligned} D_{C,C}(\phi') &= D_{C,B} D_{B,B}(\phi') D_{B,C} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \bar{\lambda}_1 \end{pmatrix} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \lambda_1 & \bar{\lambda}_1 \\ \mathrm{i}\lambda_1 & -\mathrm{i}\bar{\lambda}_1 \end{pmatrix} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix} \\ &= \begin{pmatrix} 2\,\mathrm{Re}(\lambda_1) & 2\,\mathrm{Im}(\lambda_1) \\ -2\,\mathrm{Im}(\lambda_1) & 2\,\mathrm{Re}(\lambda_1) \end{pmatrix}. \end{aligned}$$

Wegen $|\lambda_1| = 1$ ist auch $|\lambda_1|^2 = \mathrm{Re}(\lambda_1)^2 + \mathrm{Im}(\lambda_1)^2 = 1$, d. h. es gibt $\varphi \in \mathbb{R}$ mit $\cos(\varphi) = \mathrm{Re}(\lambda_1)$ und $\sin(\varphi) = -\mathrm{Im}(\varphi)$. Wir können $D_{C,C}(\phi')$ also schreiben als

$$D_{C,C}(\phi') = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Da C eine Orthonormalbasis des $\mathbb{R}^2$ ist, gilt außerdem $D_{C,C}(\phi) = D_{C,C}(\phi')$.

Gehört λ_1 zu $\mathbb{R}$, dann ist auch λ_2 eine reelle Zahl, da $\det A = \lambda_1\lambda_2$ und wegen $|\lambda_1| = |\lambda_2| = 1$ haben wir $\lambda_1, \lambda_2 \in \{\pm 1\}$. Nach Proposition VIII.5.1 gibt

es eine Orthonormalbasis C des $\mathbb{R}^2$, sodass

$$D_{C,C}(\phi) \in \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}.$$

Beweis: (i) Folgt direkt aus dem Spektralsatz Satz 31.

(ii) Wir wählen eine Orthonormalbasis von V und identifizieren V mit $\mathbb{R}^n$ vermöge D_B, wir dürfen also ohne Einschränkung annehmen, dass $V = \mathbb{R}^n$ versehen mit dem Standardskalarprodukt und dass $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ die Abbildung $x \mapsto Ax$ mit $A \in \mathrm{O}(n)$ ist. Nach Satz 31 gibt es eine Orthonormalbasis $B' = (b'_1, \dots, b'_n)$ des $\mathbb{C}^n$ bestehend aus Eigenvektoren von A, sodass b_i Eigenvektor von A zum Eigenwert λ_i ist.

Wir sortieren $b'_1, \dots, b'_n$ so um, dass $\lambda_1 = \cdots = \lambda_r = 1$, $\lambda_{r+1} = \cdots = \lambda_{r+s} = -1$ und $\lambda_{s+1}, \dots, \lambda_n \in \mathbb{C} - \mathbb{R}$. Wie im vorangegangenen Beispiel definieren wir $\phi'\colon \mathbb{C}^n \to \mathbb{C}^n$ durch $x \mapsto Ax$. Wie im vorangegangenen Beispiel erhalten wir: Ist v ein Eigenvektor von ϕ' zum Eigenwert λ, dann ist $\bar{v}$ ein Eigenvektor von ϕ' zum Eigenwert $\bar{\lambda}$.

Durch Umsortierung können wir erreichen, dass $\lambda_{s+1} = \bar{\lambda}_{s+2}, \dots, \lambda_{n-1} = \bar{\lambda}_n$ und $b_{s+1} = \bar{b}'_{s+2}, \dots, b'_{n-1} = \bar{b}'_n$.

Verfahren wir für $U := \mathrm{Lin}(b'_{s+2i-1}, b'_{s+2i})$ wie im vorangegangenen Beispiel, d. h. setzen wir

$$c_{s+2i-1} := \sqrt{2}\,\mathrm{Re}(b'_{s+2i-1}), \qquad c_{s+2i} := \sqrt{2}\,\mathrm{Im}(b'_{s+2i-1}),$$

dann ist (c_{s+2i-1}, c_{s+2i}) eine Orthonormalbasis von U und durch Zusammensetzung all dieser Basen zu $C := (b_1, \dots, b_s, c_{sn}, \dots, c_{n-1}, c_n)$ erhalten wir eine Orthonormalbasis des $\mathbb{R}^n$, sodass $D_{B,B}(\phi)$ die angegebene Form hat. □

Korollar VIII.5.4 (Spektralsatz für Matrizen): (i) *Ist $A \in \mathrm{U}(n)$, dann gibt es eine Orthonormalbasis des $\mathbb{C}^n$ bestehend aus Eigenvektoren von A und es gibt eine unitäre Matrix $S \in \mathrm{U}(n)$, sodass $SAS^{-1} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{C}$.*

(ii) *Ist $A \in \mathrm{O}(n)$, dann gibt es eine orthogonale Matrix $S \in \mathrm{O}(n)$, sodass $S^{-1}AS$ die in Satz 33(ii) angegebene Form hat.*

Satz 34 (Hauptachsentransformationssatz):

(i) *Seien $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler euklidischer beziehungsweise unitärer Raum und $\phi \in \mathrm{End}(V)$ selbstadjungiert. Dann gibt es eine Orthonormalbasis B von V bestehend aus Eigenvektoren von ϕ und alle Eigenwerte sind reell.*

(ii) *Ist $A \in \mathbb{C}^{n\times n}$ hermitesch, dann gibt es eine unitäre Matrix $S \in \mathrm{U}(n)$, sodass $S^{-1}AS = S^*AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ mit reellen Zahlen $\lambda_1, \ldots, \lambda_n$.*

(iii) *Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, dann gibt es eine orthogonale Matrix $S \in \mathrm{O}(n)$, sodass $S^{-1}AS = S^tAS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ mit reellen Zahlen $\lambda_1, \ldots, \lambda_n$.*

Proposition VIII.5.5 (Eigenwerte hermitescher Matrizen): *Eigenwerte hermitescher Matrizen sind reell.*

Beweis: Seien $A \in \mathbb{C}^{n\times n}$ hermitesch und $\lambda \in \mathbb{C}$ Eigenwert von A zum Eigenvektor $vin\mathbb{C}^n - \{\mathbf{0}\}$. Dann haben wir

$$\lambda\|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\|v\|^2,$$

d. h. $\lambda = \bar{\lambda}$ wie gewünscht. □

Beweis (von Satz 34): In Satz 31, Proposition VIII.5.5 und Proposition VIII.5.1 haben wir die ganze Arbeit bereits geleistet. □

Seien V ein $\mathbb{K}$-Vektorraum mit Basis B und $h\colon V \times V \to \mathbb{K}$ eine Bilinearform beziehungsweise Sesquilinearform. Im Folgenden schreiben wir $G_B(h)$ für die Gram-Matrix von h bezüglich B.

Ist B' eine weitere Basis von V, dann gilt nach (Proposition II.1.4) folgende Gleichung:

$$G_{B'}(h) = D^t_{B,B'} G_B(h) \bar{D}_{B,B'}.$$

Definition VIII.5.6 (Positiv definite Matrix): (i) Sei $G \in \mathbb{K}^{n\times n}$ eine symmetrische beziehungsweise hermitesche Matrix. Gilt für alle $x \in \mathbb{K}^n - \{\mathbf{0}\}$, dass $x^tG\bar{x} > 0$, dann heißt G *positiv definit*. Gilt für alle $x \in \mathbb{K}^n - \{\mathbf{0}\}$, dass $x^tG\bar{x} < 0$, so heißt G *negativ definit*.

Die symmetrische beziehungsweise hermitesche Matrix ist also positiv beziehungsweise negativ definit genau dann, wenn die zugehörige Bilinearform beziehungsweise Sesquilinearform h_G positiv beziehungsweise negativ definit ist.

(ii) Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit Basis B, $h\colon V \times V \to \mathbb{K}$ eine symmetrische beziehungsweise hermitesche Form und $G = G_B(h)$ die Gram-Matrix von h bezüglich B. Ist $G_B(h)$ positiv definit beziehungswiese negativ definit, so heißt h *positiv definit* beziehungsweise *negativ definit*.

(iii) Seien $G \in \mathbb{K}^{n\times n}$ symmetrisch beziehungsweise hermitesch und $A \in \mathrm{Gl}_n(\mathbb{K})$. Genau dann ist G positiv definit, wenn $A^tG\bar{A}$ positiv definit ist, was äquivalent zur positiven Definitheit von AGA^* ist.

Das sieht man so: Nach (Proposition II.1.4) ist $A^tG\bar{A}$ die Gram-Matrix von h_G bezüglich der Basis $B = (b_1, \ldots, b_n)$ mit $A = (b_1|\ldots|b_n)$. Die zweite Äquivalenz folgt nach Ersetzung von A durch A^t.

Bemerkung VIII.5.7 (Determinanten positiv definiter Matrizen): Sei $G \in \mathbb{K}^{n\times n}$ symmetrisch beziehungsweise hermitesch und positiv definit und $h_G\colon \mathbb{K}^n\times\mathbb{K}^n \to \mathbb{K}$ das zugehörige Skalarprodukt. Wir wählen eine Orthonormalbasis B des $\mathbb{K}^n$ bezüglich h_G – das dürfen wir machen, da das Gram-Schmidt'sche Orthogonalisierungsverfahren auch über Skalarprodukte über $\mathbb{C}$ funktioniert. Dann ist G die Gram-Matrix von h_G bezüglich der Standardbasis des $\mathbb{K}^n$ und I_n ist die Gram-Matrix von h_G bezüglich B.

Aus (Proposition II.1.4) folgt mit dem Basiswechsel $A := D_{B,E} \in \mathrm{Gl}_n(\mathbb{K})$, dass $G = A^tI_n\bar{A} = A^t\bar{A}$. Insbesondere erhalten wir

$$\det G = \det A^t \det \bar{A} = \det A \cdot \overline{\det A} = \det A^2 > 0.$$

Beispiel VIII.5.8 (Positiv definite Diagonalmatrizen): Seien $\lambda_1, \ldots, \lambda_n$ reelle Zahlen. Genau dann ist $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ positiv definit, wenn $\lambda_1, \ldots, \lambda_n > 0$.

Korollar VIII.5.9: *Seien G eine symmetrische beziehungsweise hermitesche Matrix. Genau dann ist G positiv definit, wenn alle Eigenwerte von G größer als Null sind.*

Beweis: Nach dem Hauptachsentransformationssatz gibt es eine orthogonale Matrix $S \in \mathrm{O}(n)$ beziehungsweise eine unitäre Matrix $T \in \mathrm{U}(n)$, sodass $S^tGS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ beziehungsweise $TGT^* = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ mit $\mathrm{Spec}\, G = \{\lambda_1, \ldots, \lambda_n\}$.

Per VIII.5.6 ist G positiv definit genau dann, wenn $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ positiv definit ist. Nach VIII.5.8 ist das äquivalent zu $\lambda_1, \ldots, \lambda_n > 0$. □

Satz 35 (Hurwitz-Kriterium): *Sei $G = (g_{i,j}) \in \mathbb{K}^{n\times n}$ eine symmetrische beziehungsweise hermitesche Matrix. Für $1 \le k \le n$ heißt*

$$G_k := \begin{pmatrix} g_{1,1} & \cdots & g_{1,k} \\ \vdots & & \vdots \\ g_{k,1} & \cdots & g_{k,k} \end{pmatrix}$$

der k-te Hauptminor*. Genau dann ist G positiv definit, wenn die Determinanten $\det(G_1)$, …, $\det(G_n)$ der Hauptminoren positiv sind.*

Beweis: „$\Longrightarrow$“: Es bezeichne $E = (e_1, \ldots, e_n)$ die Standardbasis des $\mathbb{K}^n$, $h\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $(x, y) \mapsto x^t G \bar{y}$ bezeichne das zu G gehörige Skalarprodukt und für $1 \leq k \leq n$ bezeichne h_k die Einschränkung von h auf den Teilraum $\mathrm{Lin}(e_1, \ldots, e_k) \times \mathrm{Lin}(e_1, \ldots, e_k)$. Ferner bezeichnen wir mit G_k die Gram-Matrix von h_k bezüglich der Standardbasis.

Da h positiv definit per Voraussetzung ist, sind auch die Einschränkungen h_k positiv definit und aus VIII.5.7 erhalten wir $\det(G_k) > 0$ für alle k.

„$\Longleftarrow$“: Wir zeigen die Behauptung per Induktion nach n. Für $n = 1$ ist nichts zu zeigen. Die Aussage gelte nun für $n - 1$. Per Induktionsvoraussetzung sind die Einschränkungen G_k für $1 \leq k \leq n - 1$ positiv definit, und so auch die h_k.

Wegen des Verfahren von Gram-Schmidt gibt es eine Basis $B' = (v_1, \ldots, v_{n-1})$ von $\mathrm{Lin}(e_1, \ldots, e_{n-1})$, die bezüglich h_{n-1} orthonormal ist. Wir erhalten eine Orthogonalbasis des $\mathbb{K}^n$ durch $v_n := e_n - \sum_{i=1}^{n} h(v_j, e_n) v_j$, da v_n nicht zu $\mathrm{Lin}(v_1, \ldots, v_{n-1}) = \mathrm{Lin}(e_1, \ldots, e_{n-1})$ gehört und außerdem für $i \in \{1, \ldots, n-1\}$ gilt, dass

$$h(v_n, v_i) = h(e_1, v_i) - \sum_{j=1}^{n-1} h(v_j, e_n) h(v_j, v_i) = h(e_n, v_i) - h(v_i, e_n) = 0.$$

Die Gram-Matrix von h bezüglich B ist $G' = \mathrm{diag}(1, \ldots, 1, c)$ mit einer reellen Zahl c. Wegen $G' = D_{E,B}^t G \bar{D}_{E,B}$ erhalten wir $c = \det G' = |\det D_{E,B}| \det G > 0$, d. h. G' ist eine positiv definite Matrix und damit auch h. □

Satz 36 (Trägheitssatz von Sylvester): *Seien h eine symmetrische Bilinearform beziehungsweise eine hermitesche Sesquilinearform auf einem euklidischen beziehungsweise unitären endlichdimensionalen Vektorraum V. Dann gilt:*

(i) *Es gibt eine Basis B von V, sodass*

$$G_B(h) = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0 \ldots, 0)$$

mit k vielen Einsen, ℓ vielen Minus-Einsen und m vielen Nullen auf der Diagonalen.

(ii) *Die Anzahlen k, ℓ und m hängen nicht von der gewählten Basis ab. Genauer: Ist G eine Gram-Matrix von h, dann ist k die Anzahl der positiven Eigenwerte von G (mit Vielfachheiten gezählt), ℓ die Anzahl der negativen Eigenwerte von G (mit Vielfachheiten gezählt) und m ist* $\dim V - \mathrm{Rang}\, G$.

Beweis: Wir geben alle Argumente für den Fall $\mathbb{K} = \mathbb{C}$, für den reellen Fall funktioniert der Beweis genau so.

(i) Seien B eine beliebige Basis von V und $G := D_B(h)$ die Gram-Matrix von h bezüglich b. Dann ist G hermitesch. Nach dem Hauptachsentransformationssatz Satz 34 gibt es eine unitäre Matrix $S = (s_{i,j}) \in \mathrm{U}(n)$ und reelle Zahlen $\lambda_1, \ldots, \lambda_n$, sodass $S^*GS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Da S regulär ist, ist auch $\bar{S}$ regulär.

Nun gibt es eine Basis B' von V mit $D_{B,B'} = \bar{S}$; genauer: Die Vektoren $b'_i := \sum_{j=1}^n \bar{s}_{j,i} b_j$ bilden eine solche Basis. Bezüglich dieser neuen Basis B' haben wir

$$D_{B'}(h) = D^t_{B,B'} D_B(h) \bar{D}_{B,B'} = S^* D_B(h) S = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

es gilt also $h(b'_i, b'_j) = \lambda_i \delta_{i,j}$. Jetzt definieren wir $\bar{B} = (\bar{b}_1, \ldots, \bar{b}_n)$ durch

$$\bar{b}_i := \begin{cases} |\lambda_i|^{-1/2}, & \text{falls } \lambda_i \neq 0, \\ b'_i, & \text{sonst.} \end{cases}$$

und erhalten $h(\bar{b}_i, \bar{b}_j) = \mathrm{sign}(\lambda_i)\delta_{i,j}$ wie gewünscht.

(ii) Seien $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ Basen so, dass $D_B(h)$ k-viele Einsen, ℓ-viele Minus-Einsen und m-viele Nullen auf der Diagonalen trägt und so, dass $D_{B'}(h)$ k'-viele Einsen, ℓ'-viele Minus-Einsen und m'-viele Nullen auf der Diagonalen trägt. Wir setzen

$$\begin{aligned} V_+ &:= \mathrm{Lin}(b_1, \ldots, b_k), \qquad V'_+ := \mathrm{Lin}(b'_1, \ldots, b'_{k'}), \\ V_- &:= \mathrm{Lin}(b_{k+1}, \ldots, b_{k+\ell}), \qquad V'_- := \mathrm{Lin}(b'_{k'+1}, \ldots, b'_{k'+\ell'}), \\ V_0 &:= \mathrm{Lin}(b_{k+\ell+1}, \ldots, b_n), \qquad V'_0 := \mathrm{Lin}(b'_{k'+\ell'+1}, \ldots, b'_n) \end{aligned}$$

und erhalten Zerlegungen $V = V_+ \oplus V_- \oplus V_0 = V'_+ \oplus V'_- \oplus V'_0$ mit $V_+ \perp V_-$, $V_+ \perp V_0$ und $V_- \perp V_0$ und den gleichen Eigenschaften für V'_+, V'_- und V'_0. Ferner sind $h|_{V+}$ und $h|_{V'_+}$ positiv definit, $h|_{V_-}$ und $h|_{V'_-}$ negativ definit und $h|V_0$ und $h|_{V'_0}$ konstant Null. Wir wollen zeigen, dass dann bereits alle Summanden übereinstimmen müssen.

Angenommen es wäre $k' = \dim V'_+ > k = \dim V_+$. Dann wäre $V'_+ \cap (V_- \oplus V_0)$ nicht-trivial, d. h. es gäbe $\mathbf{0} \neq u \in V'_+ \cap (V_- \oplus V_0)$. Dieses u können wir auf eindeutige Weise schreiben als $u = u_- + u_0$ mit $u_- \in V_-$ und $u_0 \in V_0$ und erhalten

$$h(u,u) = h(u_-, u_-) + h(u_-, u_0) + h(u_0, u_-) + h(u_0, u_0) = h(u_-, u_-) < 0$$

im Widerspruch zu $u \in V'_+$. Aus Symmetriegründen gilt auch $k \leq k'$, sodass $k = k'$. Schließlich erhalten wir analog, dass auch $\ell = \ell'$ und $m = m'$. □

6. Singulärwertzerlegung

Für eine gegebene Matrix $A \in \mathbb{K}^{n\times m}$ suchen wir für die lineare Abbildung $\varphi_A\colon \mathbb{K}^m \to \mathbb{K}^n$, $x \mapsto Ax$ geeignete Orthonormalbasen des $\mathbb{K}^m$ und $\mathbb{K}^n$ bezüglich des Standardskalarprodukts, die die Abbildungsmatrix von φ_A besonders schön machen. Wir zeigen die Aussagen dieses Abschnittes nur für $\mathbb{K} = \mathbb{R}$, über den komplexen Zahlen gehen die Argumente aber genau so durch.

Satz 37 (Singulärwertzerlegung): *Sei $A \in \mathbb{R}^{n\times m}$ gegeben.*

(i) *Ist $m \geq n$, dann gibt es Matrizen $U_1 \in \mathrm{O}(n)$ und $U_2 \in \mathrm{O}(m)$ sowie nicht-negative reelle Zahlen $\mu_1, \ldots, \mu_n$, sodass*

$$U_1 A U_2 = \begin{pmatrix} \mu_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \mu_n & 0 & \cdots & 0 \end{pmatrix}$$

Die Zahlen $\mu_1, \ldots, \mu_n$ sind durch A eindeutig bestimmt und heißen Singulärwerte *von A. Die Quadrate $\mu_1^2, \ldots, \mu_n^2$ sind die Eigenwerte von AA^t (mit Vielfachheit angegeben).*

(ii) *Ist $m \leq n$, dann gibt es Matrizen $U_1 \in \mathrm{O}(n)$ und $U_2 \in \mathrm{O}(m)$ sowie nicht-negative reelle Zahlen $\mu_1, \ldots, \mu_m$, sodass*

$$U_1 A U_2 = \begin{pmatrix} \mu_1 & & & \\ & \ddots & & \\ & & \mu_m & \\ 0 & \cdots & 0 & \\ \vdots & & \vdots & \\ 0 & \cdots & 0 & \end{pmatrix}$$

Die Zahlen $\mu_1, \ldots, \mu_m$ sind eindeutig durch A bestimmt und heißen Singulärwerte *von A. Die Quadrate $\mu_1^2, \ldots, \mu_m^2$ sind die Eigenwerte von A^tA.*

Lemma VIII.6.1 (Die Symmetrisierung A^tA): *Sei $A \in \mathbb{R}^{n\times m}$ gegeben.*

(i) *Die Matrix A^tA symmetrisch und für $\lambda \in \mathrm{Spec}(A^tA)$ gilt $\lambda \geq 0$.*

(ii) *Es ist* $\mathrm{Kern}(A^tA) = \mathrm{Kern}(A)$ *und somit auch* $\mathrm{Rang}(A^tA) = \mathrm{Rang}(A)$.

Beweis: (i) Die Symmetrie ist klar, denn $(A^tA)^t = A^tA$. Sei nun λ ein Eigenwert von A zum Eigenvektor $v \neq \mathbf{0}$. Dann haben wir

$$\|Av\|^2 = \langle Av, Av\rangle = \langle v, A^tAv\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle = \lambda|v|^2$$

(ii) Die Inklusion $\mathrm{Kern}(A) \subseteq \mathrm{Kern}(A^tA)$ ist klar. Für die andere Inklusion sei $v \in \mathrm{Kern}(A^tA)$ gegeben, d. h. $A^tAv = \mathbf{0}$. Dann ist

$$0 = \langle A^tAv, v\rangle = \langle Av, Av\rangle = \|Av\|^2,$$

sodass $Av = \mathbf{0}$ wegen den Eigenschaften der Norm, also $v \in \mathrm{Kern}(A)$ wie gewünscht. □

Beweis: Wir zeigen nur Aussage (ii), die Aussage (i) erhält man dann durch Transposition. Da A^tA symmetrisch ist, liefert der Hauptachsentransformationssatz die Existenz einer Matrix $S \in \mathrm{O}(m)$ und nicht-negativer reeller Zahlen $\lambda_1, \ldots, \lambda_m$, sodass $S^tA^tAS = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$. Ohne Einschränkung dürfen wir annehmen, dass die Eigenwerte sortiert sind, sagen wir $\lambda_1, \ldots, \lambda_p > 0$ und $\lambda_{p+1}, \ldots, \lambda_m = 0$ mit $p = \mathrm{Rang}(A^tA) = \mathrm{Rang}(A)$.

Es bezeichne die Standardbasis $E = (e_1, \ldots, e_m)$ des $\mathbb{R}^m$ und $v_i := Se_i \in \mathbb{R}^m$, $1 \leq i \leq m$. Dann ist $S = (v_1|\ldots|v_m)$, der Vektor v_i ist Eigenvektor von A^tA zum Eigenwert λ_i und $(v_1, \ldots, v_m)$ ist eine Orthonormalbasis, da S orthogonal ist. Es ist also

$$\langle Av_i, Av_j\rangle = e_i^tS^tA^tASe_j = e_i^t\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)e_j = \lambda_i\delta_{ij}.$$

Insbesondere ist $Av_i \neq \mathbf{0}$ für $1 \leq i \leq p$ und $Av_i = \mathbf{0}$ für $i > p$. Für $1 \leq i \leq n$ definieren wir $\mu_i := \lambda_i^{1/2}$ und für $1 \leq i \leq p$ setzen wir $w_i := \mu_i^{-1}Av_i$. Dann ist $\langle w_i, w_j\rangle = (\mu_i\mu_j)^{-1}\langle Av_i, Av_j\rangle = \delta_{ij}$, d. h. wir haben ein Orthonormalsystem im $\mathbb{R}^n$ erhalten. Dieses ergänzen wir zu einer Orthonormalbasis $(w_1, \ldots, w_n)$ des $\mathbb{R}^n$ und es gilt insbesondere $Av_i = \mu_iw_i$ für $1 \leq i \leq n$. Nun ist $U_1' := (w_1|\ldots|w_n)$ orthogonal und wir setzen $U_1 = U_1'^{-1}$.

Sind $(e_1, \ldots, e_m)$ die Standardbasis des $\mathbb{R}^m$ und $(e_1', \ldots, e_n')$ die Standardbasis des $\mathbb{R}^n$, dann ist für $1 \leq i \leq m$:

$$U_1AU_2e_i = U_1Av_i = \mu_iU_1w_i = \mu_ie_i',$$

d. h. U_1AU_2 hat die gewünschte Form. Nun bleibt zu zeigen, dass Σ eindeutig ist. Für $U_2^tA^tU_1^tU_1AU_2$ erhalten wir

$$\begin{pmatrix} \mu_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \mu_m & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_m \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} \mu_1^2 & & 0 \\ & \ddots & \\ 0 & & \mu_m^2 \end{pmatrix},$$

das heißt $\mu_1^2, \ldots, \mu_m^2$ sind die Eigenwerte von A^tA mit Vielfachheit gezählt und damit eindeutig. □

Korollar VIII.6.2: *Die Singulärwertzerlegung geht analog über $\mathbb{C}$ mit unitären Matrizen und nicht-negativen reellen Zahlen $\mu_1, \ldots, \mu_n$.*

Beispiel VIII.6.3 (Trick 17): (i) Seien

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad \Sigma^+ := \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{pmatrix},$$

dann sind

$$\Sigma\Sigma^+ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \Sigma^+\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(ii) Für die Matrizen

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \Sigma^+ := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

erhalten wir

$$\Sigma\Sigma^+ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \Sigma^+\Sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Proposition VIII.6.4 (Konstruktion einer Pseudo-Inversen): *Zu $A \in \mathbb{R}^{n\times m}$ mit Singulärwertzerlegung $A = U_1\Sigma U_2$ wie in Satz 37 definiere Σ^+ und A^+ aus $\mathbb{R}^{m\times n}$ durch*

$$(\Sigma^+)_{i,j} = \begin{cases} \mu_i^{-1}, & \text{falls } 1 \le i = j \le p, \\ 0 & \text{sonst,} \end{cases}$$

und $A^+ := U_2^t\Sigma^+U_1^t$. Dann gilt:

(i) *AA^+ ist ähnlich zu $\left(\begin{smallmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{smallmatrix}\right) \in \mathbb{R}^{n\times n}$.*

(ii) *A^+A ist ähnlich zu $\left(\begin{smallmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{smallmatrix}\right) \in \mathbb{R}^{m\times m}$.*

(iii) *$AA^+A = A$ und $A^+AA^+ = A^+$.*

(iv) *AA^+ und A^+A sind symmetrisch.*

Bemerkung VIII.6.5: Eine Matrix A^+ mit den Eigenschaften aus Proposition VIII.6.4 heißt Pseudo-Inverse. Eine Matrix A^+ mit den Eigenschaften aus (iii) heißt meist *Moore-Penrose-Inverse* oder auch allgemeiner *partielle Isometrie*. Die Pseudo-Inverse hat vor allem praktische Relevanz in der Numerik bei der Lösung linearer Gleichungssysteme. Genauer kann man die Pseudo-Inverse verwenden, um für lineare Gleichungssysteme, die nicht lösbar sind, eine Lösung anzugeben, die in einem speziellen Sinne eine möglichst gute Näherung darstellt.

Kapitel IX.

Etwas mehr Strukturmathematik

1. Gruppenoperationen

Seien G eine Gruppe und X eine Menge. Wir wollen zunächst an Beispielen untersuchen, was Gruppenoperationen können sollten. Die Paradebeispiele für Gruppenoperationen sind die folgenden beiden Abbildungen:

(i) $G = \mathrm{Gl}_n(K)$, $X = K^n$ und

$$\mathrm{Gl}_n(K) \times K^n \longrightarrow K^n, \qquad (A, x) \longmapsto Ax,$$

(ii) $G = S_n$, $X = \{1, \ldots, n\}$ und

$$S_n \times \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}, \qquad (\sigma, i) \longmapsto \sigma(i).$$

Wie passen die angegebenen Abbildungen mit der Gruppenstruktur auf den jeweiligen Gruppen zusammen? Wir haben $(I_n, x) \mapsto x$ respektive $(\mathrm{id}, i) \mapsto i$ und für $A, B \in \mathrm{Gl}_n(K)$ respektive $\sigma_1, \sigma_2 \in S_n$ ist $(AB, x) \mapsto (AB)x = A(Bx)$ respektive $(\sigma_1 \circ \sigma_2, i) \mapsto (\sigma_1 \circ \sigma_2)(i) = \sigma_1(\sigma_2(i))$. Von diesen Beispielen abstrahierend erhalten wir

Definition IX.1.1 (Gruppenaktion): Seien $(G, *)$ eine Gruppe mit neutralem Element 1_G, X eine nichtleere Menge und $\alpha\colon G \times X \to X$ eine Abbildung. Falls gilt

(i) Für alle $x \in X$ ist $\alpha(1_G, x) = x$.

(ii) Für alle $g_1, g_2 \in G$ und $x \in X$ ist $\alpha(g_1 * g_2, x) = \alpha(g_1, \alpha(g_2, x))$

so heißt α eine *Gruppenaktion von G auf X*. Gelegentlich werden Gruppenaktionen auch Gruppenoperationen genannt.

Notation IX.1.2: Seien $(G, *)$ eine Gruppe, X eine nichtleere Menge und $\alpha\colon G\times X \to X$ eine Gruppenoperation. Dann schreiben wir $g\cdot x := \alpha(g,x)$. Die beiden Bedingungen an die Gruppenaktion lesen sich dann als

(i) $1_G \cdot x = x$,

(ii) $(g_1 * g_2)\cdot x = g_1 \cdot (g_2 \cdot x)$.

Bemerkung IX.1.3: In der Sitation von Proposition IX.1.1 gilt: Für alle $x, y \in X$ und $g \in G$ ist $g\cdot x = y$ genau dann, wenn $x = g^{-1}\cdot y$.

Beispiel IX.1.4: (i) Seien n eine natürliche Zahl, K ein Körper und

$$\alpha\colon \mathrm{Gl}_n(K)\times K^n \longrightarrow K^n, \qquad (A,x)\mapsto Ax.$$

Dann ist α eine Gruppenoperation, wie wir uns vorher schon überlegt haben. Man nennt α auch *Aktion durch Matrizenmultiplikation*.

(ii) Seien X eine nichtleere Menge und $G = \mathrm{Perm}(X) = \{f\colon X \to X \text{ bijektiv}\}$. Dann ist

$$\alpha\colon \mathrm{Perm}(X)\times X \longrightarrow X, \qquad (\sigma, x)\mapsto \sigma(x)$$

eine Gruppenoperation. Man nennt diese Gruppenoperation auch *Aktion durch Permutation*.

(iii) Seien $X = \mathbb{R}^2$, $G = (\mathbb{R}, +)$ und

$$\alpha\colon \mathbb{R}\times\mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad \left(t, \begin{pmatrix} x \\ y \end{pmatrix}\right) \longmapsto \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

Dann ist auch dieses α eine Gruppenoperation.

(iv) Seien $X = (\mathbb{Z}/2\mathbb{Z})^3$, $G = (\mathbb{Z}/2\mathbb{Z}, +)$ und

$$\alpha\colon \mathbb{Z}/2\mathbb{Z}\times(\mathbb{Z}/2\mathbb{Z})^3 \longrightarrow \mathbb{Z}/2\mathbb{Z}^3, \qquad \left(\bar{a}, \begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix}\right) \longmapsto \begin{pmatrix} \bar{x}+\bar{a} \\ \bar{y}+\bar{a} \\ \bar{z}+\bar{a} \end{pmatrix}$$

(v) Seien $X = \{1,2,3,4\}$, $G = \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ und definiere $\alpha\colon G\times X \to X$ durch

α	1	2	3	4
$(0,0)$	1	2	3	4
$(1,0)$	2	1	4	3
$(0,1)$	3	4	1	2
$(1,1)$	4	3	2	1

Proposition IX.1.5: *Seien X eine nichtleere Menge, $(G, *)$ eine Gruppe und $\alpha\colon G \times X \to X$ eine Gruppenaktion. Dann erklärt*

$$x_1 \sim x_2 :\Longleftrightarrow \exists\, g \in G : g \cdot x_1 = x_2$$

eine Äquivalenzrelation auf G.

Beweis: Die Reflexivität ist klar, denn für jede $x \in G$ ist $x = 1_G \cdot x$ per Definition der Gruppenaktion und damit $x \sim x$.

Zur Symmetrie seien x_1 und x_2 in X mit $x_1 \sim x_2$. Das heißt es gibt $g \in G$, sodass $x_2 = g \cdot x_1$ und aus Proposition IX.1.3 wissen wir, dass $x_1 = g^{-1} \cdot x_2$, also ist auch $x_2 \sim x_1$.

Zur Transitivität seien x_1, x_2 und x_3 in X mit $x_1 \sim x_2$ und $x_2 \sim x_3$. Per Definition gibt es also g_1 und g_2 aus G mit $x_2 = g_1 \cdot x$ und $x_3 = g_2 \cdot x_2$. Es ist also $x_3 = g_2 \cdot x_2 = g_2 \cdot (g_1 \cdot x_1) = (g_2 * g_1) \cdot x_1$ und damit $x_1 \sim x_3$. □

Definition IX.1.6 (Bahn): Seien X eine nichtleere Menge, $(G, *)$ eine Gruppe, $\alpha\colon G \times X \to X$ eine Gruppenaktion und „$\sim$“ die Äquivalenzrelation, die durch α auf X erklärt wird.

(i) Die Äquivalenzklasse $G \cdot x := [x]_\sim := \{g \cdot x \mid g \in G\}$ heißt *Bahn von x*.

(ii) Gibt es nur eine Bahn, dann heißt die Operation α *transitiv*. Genau dann ist α transitiv, wenn es für alle $x, y \in X$ ein $g \in G$ gibt, sodass $y = g \cdot x$.

(iii) Seien $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$ jeweils k paarweise verschiedene Elemente von X und $(x_1, \ldots, x_k)$, $(y_1, \ldots, y_k)$ die zugehörigen Tupel. Gibt es für je zwei solche Tupel ein $g \in G$ mit $(y_1, \ldots, y_k) = (g \cdot x_1, \ldots, g \cdot x_k)$, dann heißt die Operation k-transitiv.

(iv) Für $x \in X$ heißt $\mathrm{Stab}(x) := \{g \in G \mid g \cdot x = x\}$ der *Stabilisator von x* oder *Fixgruppe von x* oder auch *Isotropiegruppe von x*.

(v) Für eine Teilmenge $S \subseteq X$ heißt

$$\mathrm{Fix}(S) := \{g \in G \mid \text{Für alle } x \in S \text{ ist } g \cdot x = x\}$$

die *Fixgruppe von S* oder auch *punktweiser Stabilisator von S*.

(vi) Gibt es für jedes $g \in G - \{1_G\}$ ein $x \in X$ mit $g \cdot x \neq x$, dann heißt α *treu*.

(vii) Gilt für alle $x \in X$ und $g \in G - \{1_G\}$, dass $g \cdot x \neq x$, dann heißt die Aktion *fixpunktfrei*.

Bemerkung IX.1.7: Aus Proposition IX.1.5 folgt: Die Bahnen bilden eine Partition der Menge X, d. h. X ist die disjunkte Vereinigung seiner Bahnen. Insbesondere gilt: Ist X eine endliche Menge, die aus den k Bahnen $B_1, \ldots, B_k$ besteht und wählen wir aus jeder Bahn B_i ein Element x_i – man nennt $\{x_1, \ldots, x_k\}$ ein vollständiges Repräsentantensystem oder Vertretersystem der Bahnen – dann gilt $\#X = \sum_{i=1}^{k} \#G \cdot x_i$.

Beispiel IX.1.8: In Proposition IX.1.4 erhalten wir folgende Bahnen und Stabilisatorgruppen:

(i) Für alle $A \in \mathrm{Gl}_n(K)$ gilt $A \cdot \mathbf{0} = \mathbf{0}$, das heißt $G \cdot \mathbf{0} = \{\mathbf{0}\}$. Ist beispielsweise $x = e_1$ der erste Standardbasisvektor, dann liefert $A \cdot e_1$ die erste Spalte der Matrix A. Weil wir wissen, dass wir jeden von Null verschiedenen Vektor $x \in K^n - \{0\}$ in einer Basis des K^n wiederfinden können (das sagt der Basisergänzungssatz), finden wir eine Matrix, deren erste Spalte genau der Vektor x ist. Das heißt, es gibt genau zwei Bahnen unter dieser Operation, nämlich $\{\mathbf{0}\}$ und $K^n - \{\mathbf{0}\}$.

Beispielsweise für $x = \mathbf{0}$ ist $\mathrm{Stab}(\mathbf{0}) = \mathrm{Gl}_n(K)$ und für $x = e_1$ ist

$$\mathrm{Stab}(e_1) = \left\{ \begin{pmatrix} 1 & * & \cdots & * \\ 0 & * & & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} \in \mathrm{Gl}_n(K) \right\}.$$

Die Aktion ist treu, aber nicht fixpunktfrei.

(ii) Die Permutationsgruppe $\mathrm{Perm}(X)$ operiert transitiv, denn für $x, y \in X$ definiert

$$f \colon X \longrightarrow X, \qquad z \longmapsto \begin{cases} x, & \text{falls } z = y, \\ y, & \text{falls } z = x, \\ z, & \text{sonst,} \end{cases}$$

eine bijektive Abbildung von X nach X, die x nach y verschiebt. Für irgendein $x \in X$ ist $\mathrm{Stab}(x) = \mathrm{Perm}(X - \{x\})$. Ferner ist die Gruppenaktion α treu, und k-transitiv für jede natürliche Zahl k, aber α ist nicht fixpunktfrei.

(iii) Ist $x \in \mathbb{R}^2 - \{\mathbf{0}\}$, dann ist $G \cdot x = \partial B(0, \|x\|) = \{y \in \mathbb{R}^2 \mid \|y\| = \|x\|\}$ und ist $x = \mathbf{0}$, dann ist $G \cdot \mathbf{0} = \{\mathbf{0}\}$. Für ein $x \in \mathbb{R}^2 - \{\mathbf{0}\}$ ist $\mathrm{Stab}(x) = 2\pi\mathbb{Z} = \{2\pi k \mid k \in \mathbb{Z}\}$ wegen der 2π-Periodizität des Sinus und Kosinus und für $x = \mathbf{0}$ ist $\mathrm{Stab}(x) = G = \mathbb{R}$. Die Operation α ist nicht transitiv und außerdem nicht treu, denn für $g = 2\pi$ gilt für alle $x \in X$, dass $g \cdot x = x$.

(iv) Diese Aktion hat 4 Bahnen, die jeweils diagonal gegenüberliegene Punkte zusammenfasst und außerdem ist die Aktion fixpunktfrei.

(v) Aus der Definition der Aktion können wir ablesen, dass die Aktion transitiv ist. Außerdem kann man bei genauer Betrachtung der Definition erkennen, dass es keine Fixpunkte gibt und die Operation deshalb treu ist.

Proposition IX.1.9 (Gruppenaktionen als Gruppenhomomorphismen): *Seien* $(G, *)$ *eine Gruppe,* X *eine nichtleere Menge und* $\mathrm{Sym}(X) := \mathrm{Perm}(X)$ *die Gruppe der bijektiven Selbstabbildungen zusammen mit der Komposition von Abbildungen.*

(i) *Ist* $g \in G$ *fest und* $\alpha\colon G \times X \to X$ *eine Gruppenaktion, dann ist* $\alpha(g, \cdot)\colon X \to X$ *ein Element von* $\mathrm{Sym}(X)$ *und*

$$\hat{\alpha}\colon G \longrightarrow \mathrm{Sym}(X), \qquad g \longmapsto (x \mapsto g \cdot x)$$

ist ein Homomorphismus von Gruppen.

(ii) *Ist* $\hat{\alpha}\colon G \to \mathrm{Sym}(X)$ *ein Gruppenhomomorphismus, dann erklärt*

$$\alpha\colon G \times X \longrightarrow X, \quad (g, x) \longmapsto (\hat{\alpha}(g))(x)$$

eine Gruppenoperation auf X*.*

Beweis: (i) Die Abbildung $x \mapsto g \cdot x$ ist eine Permutation, denn $x \mapsto g^{-1} \cdot x$ ist die zugehörige Umkehrabbildung. Bleibt zu zeigen, dass $\hat{\alpha}$ ein Gruppenhomomorphismus ist. Für $g_1, g_2 \in G$ und $x \in X$ gilt

$$\begin{aligned}(\hat{\alpha}(g_1 * g_2))(x) &= (g_1 * g_2) \cdot x \\ &= g_1 \cdot (g_2 \cdot x) \\ &= (\hat{\alpha}(g_1))(g_2 \cdot x) = (\hat{\alpha}(g_1))(\hat{\alpha}(g_2)(x)) = (\hat{\alpha}(g_1) \cdot \hat{\alpha}(g_2))(x)\end{aligned}$$

wie gewünscht.

(ii) Wir müssen die Axiome einer Gruppenaktion nachweisen. Erstens gilt für ein $x \in X$, dass $\alpha(1, x) = (\hat{\alpha}(1))(x) = \mathrm{id}(x) = x$, und zweitens haben wir für $g_1, g_2 \in G$:

$$\begin{aligned}\alpha(g_1 * g_2, x) &= (\hat{\alpha}(g_1 * g_2))(x) \\ &= (\hat{\alpha}(g_1) \circ \hat{\alpha}(g_2))(x) = \hat{\alpha}(g_1)(\alpha(g_2, x)) = \alpha(g_1, \alpha(g_2, x)),\end{aligned}$$

sodass α eine Gruppenaktion ist. □

Bemerkung IX.1.10: Die beiden Konstruktionen aus Proposition IX.1.9 sind invers zueinander. Etwas salopp gesagt ist eine Gruppenaktionen von G auf X nichts anderes als ein Gruppenhomomorphismus von G nach $\mathrm{Sym}(X)$ (mit Komposition von Abbildungen).

Besitzt die Menge X eine Zusatzstruktur (beispielsweise eine Vektorraumstruktur oder die Struktur eines metrischen Raums), dann werden für $\mathrm{Sym}(X)$ oft nicht alle bijektiven Mengenabbildungen von X nach X zugelassen, sondern nur die strukturerhaltenden Selbstabbildungen.

Bemerkung IX.1.11 (Treue von Gruppenaktionen): Seien G eine Gruppe, X eine Menge, $\alpha\colon G \times X \to X$ eine Gruppenaktion und $\hat{\alpha}\colon G \to \mathrm{Sym}(X)$ der zugehörige Homomorphismus aus Proposition IX.1.9. Genau dann ist α eine treue Aktion, wenn $\hat{\alpha}$ injektiv ist.

Satz 38 (Bahnformel): *Seien G eine endliche Gruppe, X eine Menge und $\alpha\colon G \times X \to X$, $(g, x) \mapsto g \cdot x$ eine Gruppenaktion. Für jedes $x \in X$ gilt*

$$\#G = \#Gx \# \operatorname{Stab}(x).$$

Beweis: Für ein $x \in X$ wird auf G durch

$$g \sim g' :\Longleftrightarrow gx = g'x \Longleftrightarrow g^{-1}g' \in \operatorname{Stab}(x)$$

eine Äquivalenzrelation erklärt. Es bezeichne $G/\!\sim$ die Menge der Äquivalenzklassen bezüglich dieser Äquivalenzrelation. Dann erhalten wir die wohldefinierte Abbildung

$$\varphi\colon G/\!\sim \longrightarrow Gx, \qquad [g] \longmapsto g \cdot x$$

denn per Konstruktion gilt $g_1 \sim g_2$ genau dann, wenn $g_1 \cdot x = g_2 \cdot x$. Außerdem ist φ surjektiv per Definition der Bahn und schließlich ist φ injektiv, denn es gilt $g_1 \cdot x = g_2 \cdot x$ genau dann, wenn $[g_1] = [g_2]$. Insbesondere ist $\#G/\!\sim = \#Gx$.

Für ein festes $g \in G$ und ein Gruppenelement g' gilt per Vorüberlegung, dass $g' \in [g]$ genau dann, wenn $g^{-1}g' \in \operatorname{Stab}(x)$. Wir haben also Abbildungen $\psi\colon [g] \to \operatorname{Stab}(x)$, $g' \mapsto g^{-1}g'$ und $\psi'\colon \operatorname{Stab}(x) \to [g]$, $g' \mapsto g \cdot g'$, die Invers zueinander sind. Insbesondere ist $\#\operatorname{Stab}(x) = \#[g]$.

Da G die disjunkte Vereinigung der Äquivalenzklassen ist und jede Äquivalenzklasse $\#\operatorname{Stab}(x)$-viele Elemente hat, haben wir

$$\#G = \#G/\!\sim \# \operatorname{Stab}(x) = \#Gx \# \operatorname{Stab}(x). \qquad \square$$

Korollar IX.1.12: *Seien G eine endliche Gruppe, X eine endliche Menge und $\alpha\colon G \times X \to X$ eine Gruppenaktion.*

(i) *Für jedes $x \in X$ ist $\#Gx$ ein Teiler der Ordnung von G.*

(ii) *Falls es nur endlich viele Bahnen gibt und $\{x_1, \ldots, x_k\}$ ein Vertretersystem der Bahnen ist, dann gilt*

$$\#X = \#Gx_1 + \cdots + \#Gx_k = \sum_{i=1}^{k} \frac{\#G}{\#\operatorname{Stab}(x_i)}.$$

Beispiel IX.1.13: Sei $(G, *)$ eine Gruppe. Dann operiert G auf sich selbst durch $g_1 \cdot g_2 := g_1 * g_2$, d. h., die Abbildung

$$\alpha\colon G \times G \longrightarrow G, \qquad (g_1, g_2) \longmapsto g_1 * g_2$$

ist eine Gruppenaktion. Diese Operation ist treu (da die Eins bewegt wird) und transitiv (weil wir invertieren können).

Satz 39 (von Cayley): *Für jede Gruppe $(G, *)$ gibt es eine Einbettung (d. h. einen injektiven Gruppenhomomorphismus) $\hat{\alpha}\colon G \hookrightarrow \operatorname{Sym}(G)$. Insbesondere ist jede endliche Gruppe isomorph zu einer Untergruppe einer symmetrischen Gruppe, genauer: Ist $\#G = n$, dann ist $\operatorname{Sym}(G) \cong S_n$ und G kann als Untergruppe von S_n aufgefasst werden.*

Beweis: Sei G eine Gruppe. Nach Proposition IX.1.13 operiert G treu auf sich selbst durch Linksmultiplikation, was heißt, dass der zugehörige Homomorphismus $\hat{\alpha}\colon G \to \operatorname{Sym}(G)$ injektiv ist. □

Der Satz von Cayley lässt uns das Studium endlicher Gruppen als das Studium der Untergruppen der Permutationsgruppen interpretieren.

Beispiel IX.1.14: Seien K und L Körper mit $K \subseteq L$ (man sagt, L sei eine Körpererweiterung von K) und $G := \operatorname{Aut}(L) = \{\varphi\colon L \to L \text{ Körperautomorphismus}\}$. Dann operiert G auf L durch

$$\alpha\colon G \times L \longrightarrow L, \qquad (\varphi, \ell) \longmapsto \varphi(\ell).$$

Dann heißt $\operatorname{Aut}(L|K) := \operatorname{Fix}(K) = \{\varphi \in \operatorname{Aut}(L) \mid \text{Für alle } c \in K \text{ ist } \varphi(c) = c\}$ die *Automorphismengruppe* der Körpererweiterung $L|K$. Für sogenannte galois'sche Erweiterungen schreibt man auch $\operatorname{Gal}(L|K) := \operatorname{Aut}(L|K)$.

In der Vorlesung „Algebra“ wird die Beziehung zwischen Gruppen und Körpererweiterungen intensiv studiert und gipfelt in folgendem unglaublichen Satz: Ist $K \subseteq L$ eine endliche galois'sche Körpererweiterung, dann haben wir eine Bijektion

$$\{\text{Zwischenkörper } K \subseteq E \subseteq L\} \longleftrightarrow \{\text{Untergruppen von } G = \operatorname{Gal}(L|K)\},$$
$$E \longmapsto \operatorname{Gal}(L|E).$$

Diese Korrespondenz ermöglicht unter anderem einen Beweis dafür, dass es keine allgemeine Lösungsformel für Gleichungen des Grades mindestens 5 geben kann.

Ein ungelöstes Problem ist folgende Frage: „Sei $K = \mathbb{Q}$. Gibt es für jede endliche Gruppe G eine endliche galois'sche Körpererweiterung $\mathbb{Q} \subseteq L$, sodass $\mathrm{Gal}(L|\mathbb{Q}) = G$?“ Ein bisher nicht erfolgreicher Lösungsansatz von Emmy Noether verwendet als Grundidee den Satz von Cayley.

2. Teilbarkeit in Ringen

Das Ziel dieses Abschnittes ist das systematische Studium der Konzepte „Teilbarkeit“ und „größter gemeinsamer Teiler“ um beispielsweise zu erkennen, auf welche Ringe sich das wichtige Lemma von Bézout verallgemeinert.

Definition IX.2.1: Sei $(R, +, \cdot)$ ein kommutativer unitärer Ring.

(i) Seien $a, b \in R$. Gibt es $c \in R$ mit $b = ca$, dann sagen wir a *teile* b und schreiben $a \mid b$.

(ii) Sei $a \in R - \{0\}$. Gibt es $c \in R - \{0\}$ mit $ac = 0$, dann heißt a ein *echter Nullteiler*.

(iii) Gibt es in R keine echten Nullteiler, dann heißt R *nullteilerfrei*. Das heißt: Sind $a, b \in R$ mit $ab = 0$, dann ist $a = 0$ oder $b = 0$. Ein nullteilerfreier kommutativer unitärer Ring heißt auch Integritätsring oder Integritätsbereich.

(iv) Sei $a \in R$. Gibt es $b \in R$ mit $ab = 1$, dann heißt a *invertierbar*. Die Menge $R^\times := \{u \in R \mid u \text{ ist invertierbar}\}$ heißt *Einheitengruppe* oder *multiplikative Gruppe von* R und ist eine Gruppe mit Multiplikation.

Bemerkung IX.2.2 (Kürzungsregel, Einheitenteiler): Sei R ein kommutativer unitärer Ring.

(i) Genau dann ist R nullteilerfrei, wenn für alle $a, c, c' \in R$ mit $a \neq 0$ gilt: Wenn $ac = ac'$, dann ist $c = c'$.

(ii) Teiler von Einheiten sind selbst Einheiten.

Beweis: (i) Ist $ac = ac'$, dann ist $a(c - c') = 0$, d. h. wegen der Nullteilerfreiheit von R ist $c - c' = 0$ und damit $c = c'$.

(ii) Seien $a \in R$, $u \in R^\times$ und a teile u. Dann gibt es $c \in R$ mit $u = ac$ und $u' \in R$ mit $uu' = 1$. Aber dann erhalten wir $a(cu') = uu' = 1$, d. h. a ist eine Einheit. □

Bemerkung IX.2.3 (Teilbarkeit ist fast eine Ordnungsrelation): Seien R ein kommutativer unitärer nullteilerfreier Ring und $a, b, c \in R$. Dann ist Teilbarkeit fast eine Ordnungsrelation auf R:

(i) Jedes $a \in R$ teilt sich selbst, denn $a = 1a$.

(ii) Sind $a, b, c \in R$ mit $a \mid b$ und $b \mid c$, dann gibt es $a', b' \in R$ mit $b = aa'$ und $c = bb'$, d. h. $c = bb' = aa'b'$ und damit teilt a auch c.

(iii) Sind $a, b \in R$ mit $a \mid b$ und $b \mid a$, dann gibt es a' und b' aus R mit $b = a'a$ und $a = b'b$, d. h. $b = a'a = a'b'b$, sodass $1 = a'b'$, d. h. $a'b'$ gehört zu $R^\times$.

Uns stört, dass wir in (iii) keine Gleichheit, sondern nur Gleichheit bis auf Multiplikation mit einer Einheit haben. Der passende Ausweg aus dieser misslichen Lage sind Äquivalenzrelationen: Wir sollten diejenigen Elemente zusammenfassen, die sich nur durch Multiplikation mit einer Einheit unterscheiden.

Proposition IX.2.4 (Assoziiertheit und Teilbarkeit): *Seien R ein kommutativer unitärer nullteilerfreier Ring und $a, b \in R$. Es gilt $a \mid b$ und $b \mid a$ genau dann, wenn es $u \in R^\times$ mit $b = ua$ gibt. Solche Elemente a, b nennen wir* zueinander assoziiert. *Assoziiertheit ist eine Äquivalenzrelation auf R.*

Sind $a, b \in R$ und $a', b' \in R$ so, dass a und a' sowie b und b' assoziiert sind, dann gilt $a \mid b$ genau dann, wenn $a' \mid b'$.

Wir nennen $R^\times a := \{ua \mid u \in R^\times\} = [a]_\sim$ die Assoziiertheitsklasse von a. *Teilbarkeit ist eine Ordnungrelation auf der Menge der Assoziiertheitsklassen.*

Beweis: Ist $a \mid b$ und $b \mid a$, dann liefert Proposition IX.2.3(iii) ein $u \in R^\times$ mit $b = ua$. Ist $b = ua$, dann gelten sowohl $a \mid b$ als auch $b \mid a$ wegen $u^{-1}b = a$. Dass Assoziiertheit eine Äquivalenzrelation ist, ist klar.

Sind $a, b \in R$ und $a', b' \in R$ so, dass a und a' sowie b und b' assoziiert sind, dann wissen wir, dass $a' \mid a$, $a \mid b$ und $b \mid b'$. Wegen Proposition IX.2.3 folgt daraus $a' \mid b'$.

Dass Teilbarkeit eine Ordnungsrelation auf der Menge der Assozziertheitsklassen ist, ist klar – genau das war das Ziel. □

Definition IX.2.5 (Größter gemeinsamer Teiler): Seien R ein kommutativer unitärer nullteilerfreier Ring und $a, b, g \in R$. Gilt $g \mid a$, $g \mid b$ und gilt für alle $g' \in R$ mit $g' \mid a$ und $g' \mid b$, dass $g \mid g'$, dann heißt g ein *größter gemeinsamer Teiler von a und b.*

Bemerkung IX.2.6: Seien $a, b, g, g' \in R$ und g ein größter gemeinsamer Teiler von a, b. Genau dann ist g' assoziiert zu g, wenn g' auch ein größter gemeinsamer

Teiler von a und b ist, d. h. größte gemeinsame Teiler sind nur bis auf Assoziiertheit eindeutig. Wir schreiben in diesem Fall $g = \mathrm{ggT}(a, b)$ beziehungsweise $R^\times g = \mathrm{ggT}(a, b)$. Wenn $\mathrm{ggT}(a, b) = 1$ beziehungsweise $\mathrm{ggT}(a, b) = R^\times$, dann heißen a und b teilerfremd.

Erinnerung IX.2.7: Seien R ein Ring und $I \subseteq R$ eine nichtleere Teilmenge.

(i) Gilt für alle $a, b \in I$ und $r \in R$, dass $a + b \in I$ und $ra \in I$, so heißt I ein *Ideal in R.*

(ii) Gibt es $a_1, \ldots, a_n$ mit $I = \{r_1 a_1 + \cdots + r_n a_n \mid r_1, \ldots, r_n\}$, dann schreiben wir $I := \langle a_1, \ldots, a_n \rangle$.

(iii) Gibt es $g \in R$ mit $I = \langle g \rangle$, dann heißt I ein *Hauptideal.*

(iv) Sind alle Ideale in R Hauptideale, dann ist R ein Hauptidealring.

(v) Die Ringe $\mathbb{Z}$ und $K[X]$ (für einen Körper K) sind Hauptidealringe.

Bemerkung IX.2.8: Seien R ein Integritätsring und $a, b, d \in R$. Es gilt $d \mid a$ und $d \mid b$ genau dann, wenn $\langle a, b \rangle \subseteq \langle d \rangle$.

Beweis: Es gilt $\langle a, b \rangle \subseteq \langle d \rangle$ genau dann, wenn $a, b \in \langle d \rangle$, was per Definition gerade bedeutet, dass $d \mid a$ und $d \mid b$. □

Satz 40 (Verallgemeinerter Bézout): *Seien R ein Integritätsring, $a, b, g \in R$ und $\langle a, b \rangle$ ein Hauptideal. Es gilt genau dann $g = \mathrm{ggT}(a, b)$, wenn $\langle g \rangle = \langle a, b \rangle$.*

Beweis: „$\Longrightarrow$“: Ist $g = \mathrm{ggT}(a, b)$, dann haben wir per Proposition IX.2.8, dass $\langle a, b \rangle \subseteq \langle g \rangle$. Weil $\langle a, b \rangle$ per Voraussetzung ein Hauptideal ist, gibt es $g' \in R$ mit $\langle g' \rangle = \langle a, b \rangle$. Wieder mit Proposition IX.2.8 erhalten wir $g' \mid a$ und $g' \mid b$, sodass g' ein Teiler von g sein muss, was gerade $g' \in \langle g \rangle$ bedeutet. Schließlich erhalten wir $\langle g \rangle \subseteq \langle g' \rangle = \langle a, b \rangle$.

„$\Longleftarrow$“: Ist $\langle g \rangle = \langle a, b \rangle$, dann gibt Proposition IX.2.8, dass $g \mid a$ und $g \mid b$. Für $g' \in R$ mit $g' \mid a$ und $g' \mid b$ liefert Proposition IX.2.8 $\langle g' \rangle \supseteq \langle a, b \rangle = \langle g \rangle$, d. h. $g' \mid g$. □

Korollar IX.2.9: *Aus Satz 40 folgt insbesondere:*

(i) *Ist R ein Hauptidealring, dann gibt es für je zwei Elemente $a, b \in R$ einen größten gemeinsamen Teiler.*

(ii) *Ist R ein Hauptidealring und sind $a, b, g \in R$ mit $g = \mathrm{ggT}(a, b)$, dann gibt es $k, \ell \in R$, sodass $g = ka + \ell b$. Insbesondere gilt das Lemma von Bézout in beliebigen Hauptidealringen.*

Achtung! In beliebigen Integritätsringen muss es keine größten gemeinsamen Teiler geben!

3. Euklidische Ringe

Das Hauptwerkzeug für den Beweis, dass der Polynomring über einem Körper $K[X]$ und der Ring der ganzen Zahlen $\mathbb{Z}$ Hauptidealringe sind, war die Teilbarkeit mit Rest. In diesem Abschnitt wollen wir die Teilbarkeit mit Rest allgemeiner fassen und die Ringe – sogenannte euklidische Ringe – studieren, in denen es so etwas gibt.

Definition IX.3.1 (Euklidischer Ring): Sei R ein Integritätsring. Gibt es auf R eine Funktion $\varphi\colon R \to \mathbb{N}_0$ mit den Eigenschaften

(i) Es gilt $\varphi(r) = 0$ genau dann, wenn $r = 0$,

(ii) Für alle $a \in R$ und $b \in R - \{0\}$ gibt es $c \in R$, sodass $\varphi(a - bc) < \varphi(b)$,

dann heißt φ eine Gradfunktion. Das Paar (R, φ) heißt *euklidischer Ring*. Sind keine Verwechslungen zu befürchten, so sprechen wir vom *euklidischen Ring* R.

Beispiel IX.3.2 (Erste euklidische Ringe): (i) Sei R der Ring der ganzen Zahlen $\mathbb{Z}$ und $\varphi\colon \mathbb{Z} \to \mathbb{N}_0$, $z \mapsto |z|$. MIt dieser Gradfunktion ist $\mathbb{Z}$ ein euklidischer Ring.

(ii) Seien K ein Körper, $R = K[X]$ und φ die Funktion

$$\varphi\colon K[X] \longrightarrow \mathbb{N}_0, \qquad f \longmapsto \begin{cases} 0, & \text{falls } f = 0, \\ \deg(f) + 1, & \text{sonst.} \end{cases}$$

Dann ist $K[X]$ ein euklidischer Ring.

Proposition IX.3.3 (Euklidische Ringe sind Hauptidealringe): *Ist R ein euklidischer Ring, dann ist R ein Hauptidealring.*

Beweis: Wie damals schon angemerkt geht der Beweis von Satz 24 auch in dieser allgemeineren Situation durch. □

Korollar IX.3.4 (Existenz des ggT): *Seien R ein euklidischer Ring und $a, b \in R$. Dann gibt es $g \in R$ mit $g = \mathrm{ggT}(a, b)$.*

Beweis: Das folgt sofort aus Proposition IX.2.9. □

Beispiel IX.3.5 (Iteriertes Teilen mit Rest): Seien R der Ring der ganzen Zahlen $\mathbb{Z}$ und a, b die ganzen Zahlen $a = 93$, $b = 42$.

$$\begin{aligned} 93 &= 1 \cdot 93 + 0 \cdot 42 \\ 42 &= 0 \cdot 93 + 1 \cdot 42 \\ 9 &= 1 \cdot 93 - 2 \cdot 42 \\ 6 &= -4 \cdot 93 + 9 \cdot 42 \\ 3 &= 5 \cdot 93 - 42 \\ 0 &= -14 \cdot 93 - 11 \cdot 42 \end{aligned}$$

Unser größter gemeinsamer Teiler ist $\mathrm{ggT}(93, 42) = 3$ und $3 = 5 \cdot 93 - 11 \cdot 42$.

Proposition IX.3.6 (Euklidischer Algorithmus): *Seien R ein euklidischer Ring und $a, b \in R - \{0\}$. Wir berechnen r_i, q_i, x_i und y_i iterativ wie folgt: Es seien*

$$r_{-1} = a, \qquad x_{-1} = 1, \qquad y_{-1} = 0, \qquad r_0 = b, \qquad x_0 = 0, \qquad y_0 = 1.$$

Seien für $i \geq 0$ alle Werte r_i, x_i, y_i und q_i bereits bestimmt. Wähle dann q_{i+1} mit $\varphi(r_{i-1} - q_{i+1}r_i) < \varphi(r_i)$ und definiere

$$r_{i+1} := r_{i-1} - q_{i+1}r_i, \qquad x_{i+1} := x_{i-1} - q_{i+1}x_i, \qquad y_{i+1} := y_{i-1} - q_{i+1}y_i.$$

Dann gilt:

(i) *Es gibt eine nichtnegative ganze Zahl n mit $r_{n+1} = 0$.*

(ii) *Für $-1 \leq i \leq n+1$ ist $r_i = x_i a + y_i b$.*

(iii) *Es ist $r_n = \mathrm{ggT}(a, b)$ und somit $\mathrm{ggT}(a, b) = x_n a + y_n b$.*

Beweis: (i) Wegen $0 \leq \varphi(r_{i+1}) < \varphi(r_i)$ muss das Verfahren terminieren.

(ii) Per vollständiger Induktion zeigen wir, dass $r_i = x_i a + y_i b$. Der Induktionsanfang für $i \in \{-1, 0\}$ ist klar. Die Aussage gelte nun für eine natürliche Zahl i. Dann haben wir

$$\begin{aligned} x_{i+1}a + y_{i+1}b &= (x_{i-1} - q_{i+1})a + (y_{i-1} - a_{i+1}y_i)b \\ &= x_{i-1}a + y_{i-1}b - q_{i+1}(x_i a + y_i b) = r_{i-1} - q_{i+1}r_i = r_{i+1}, \end{aligned}$$

und der Induktionsschluss ist vollzogen.

(iii) Ist g ein gemeinsamer Teiler von a und b, dann ist g wegen (ii) ein Teiler von r_n. Nun zeigen wir per Induktion für $-1 \leq i \leq n$, dass $r_n \mid r_i$. Der Induktionsanfang $i = n$ ist klar. Für $i = n-1$ haben wir $0 = r_{n+1} = r_{n-1} - q_{n+1}r_n$, also $r_{n-1} = q_{n+1}r_n$, sodass r_n ein Teiler von r_{n-1} ist.

Die Aussage gelte nun für i. Per Induktionsvoraussetzung haben wir, dass r_n ein Teiler von r_i und von r_{i+1} ist. Unter Verwendung von $r_{i+1} = r_{i-1} - q_{i+1}r_i$ erhalten wir $r_{i-1} = r_{i+1} + q_{i+1}r_i$, sodass r_n ein Teiler von r_{i-1} ist. Wegen $r_{-1} = a$ und $r_0 = b$ folgt die Behauptung. □

4. Primelemente in Ringen

In diesem Abschnitt wollen wir das Konzept Primalität verallgemeinern. Wir werden sehen, dass es in euklidischen Ringen eine eindeutige Primfaktorzerlegung gibt. Wie aus der Schule bekannt heißt eine natürliche Zahl prim, falls sie keine Teiler außer Eins und sich selbst hat, d. h., ist n eine natürliche Zahl und gilt für alle natürlichen Zahlen a, b mit $ab = n$, dass $a = 1$ oder $b = 1$, dann heißt n prim. Ferner ist aus der Schule bekannt, dass sich jede natürliche Zahl auf eindeutige Weise als Produkt endlich vieler Primzahlen schreiben lässt.

Proposition IX.4.1: *Sei p eine natürliche Zahl größer als Eins. Genau dann ist p prim, wenn für alle natürlichen Zahlen a, b mit $p \mid ab$ gilt, dass $p \mid a$ oder $p \mid b$.*

Beweis: „$\Longrightarrow$“: Seien a und b natürliche Zahlen mit $p \mid ab$, d. h. es gibt $k \in \mathbb{N}$ mit $ab = kp$. Wie aus der Schule bekannt, lassen sich a, b und k in eindeutige Weise in Primfaktoren zerlegen, sagen wir $a = \prod_{i=1}^{n} p_i$, $b = \prod_{i=1}^{m} q$ und $k = \prod_{i=1}^{r} \ell_i$ mit Primzahlen $p_1, \ldots, p_n$, $q_1, \ldots, q_m$ und $\ell_1, \ldots, \ell_r$. Dann ist

$$ab = \Big(\prod_{i=1}^{n} p_i\Big)\Big(\prod_{i=1}^{m} q_j\Big) = \Big(\prod_{i=1}^{r} \ell_i\Big)p.$$

Wegen der Eindeutigkeit der Primfaktorzerlegung ist $p \in \{p_1, \ldots, p_n, q_1, \ldots, q_m\}$, also $p \mid a$ oder $p \mid b$.

„$\Longleftarrow$“: Seien a, b natürliche Zahlen mit $p = ab$. Per Voraussetzung ist $p \mid a$ oder $p \mid b$, ohne Einschränkung dürfen wir annehmen, dass $p \mid a$. Es gibt also eine natürliche Zahl t, sodass $a = tp$, d. h. $p = ab = tbp$. Mit der Kürzungsregel dürfen wir $1 = tb$ folgern, sodass $t, b \in \mathbb{N}^\times$ gelten muss, also $t = b = 1$. Damit ist p prim. □

Für die Implikation „$\Longrightarrow$“ haben wir die Primfaktorzerlegung verwendet, für „$\Longleftarrow$“ haben wir nur die Kürzungsregel gebraucht.

Definition IX.4.2: Seien R ein Integritätsring und $m \in R - (R^\times \cup \{0\})$.

(i) Gilt für alle $a, b \in R$ mit $m = ab$, dass $a \in R^\times$ oder $b \in R^\times$, so heißt m *irreduzibel*.

(ii) Gilt für alle $a, b \in R$ mit $m \mid ab$, dass $m \mid a$ oder $m \mid b$, dann heißt m *prim*.

Bemerkung IX.4.3: Seien R in Integritätsring und $m \in R - (R^\times \cup \{0\})$. Ist m prim, dann ist m auch irreduzibel. Die Umkehrung gilt im Allgemeinen nicht!

Proposition IX.4.4 (Primalität vs. Irreduzibelität in Hauptidealringen): *Seien R ein Hauptidealring und $m \in R - (R^\times \cup \{0\})$. Genau dann ist m prim, wenn m irreduzibel ist.*

Beweis: „$\Longrightarrow$“: Siehe Proposition IX.4.3.

„$\Longleftarrow$“: Seien a und b Elemente von R mit $m \mid ab$ und sei $g := \mathrm{ggT}(a, m)$. Insbesondere ist dann $a = r_1 g$ und $m = r_2 g$ mit r_1 und r_2 aus R. Wegen der Irreduzibelität von m ist r_2 oder g eine Einheit. Ist r_2 eine Einheit, dann haben wir $a = r_1 r_2^{-1} m$, d. h. $m \mid a$. Ist andernfalls g eine Einheit, dann liefert das Lemma von Bézout, dass es $x, y \in R$ mit $g = xa + ym$ gibt und wir erhalten $b = g^{-1}gb$, sodass $b = g^{-1}(xab + ymb)$, also wird b von m geteilt. □

Definition IX.4.5: Seien R ein Integritätsring und $I \subsetneq R$ ein Ideal. Gilt für alle $x, y \in R$ mit $xy \in I$, dass $x \in I$ oder $y \in I$, so heißt I ein *Primideal*.

Bemerkung IX.4.6: Seien R ein kommutativer unitärer Ring und $0 \neq m$ ein Element von R, das keine Einheit ist. Genau dann ist m prim, wenn $\langle m \rangle$ ein Primideal ist. Ferner ist $\langle 0 \rangle$ ein Primideal genau dann, wenn R nullteilerfrei ist.

Satz 41: *Seien R ein Integritätsbereich und $r \in R - (R^\times \cup \{0\})$. Ist R zusätzlich ein Hauptidealring, dann gilt:*

(i) *Es gibt Primelemente $p_1, \ldots, p_n \in R$ mit $r = p_1 \cdots p_n$.*

(ii) *Die Zerlegung ist eindeutig bis auf die Reihenfolge der Faktoren und Assoziiertheit.*

Proposition IX.4.7 (Aufsteigende Idealketten): *Seien R ein Integritätsbereich und $i_1, I_2, \ldots$ Ideale in R mit $I_1 \subseteq I_2 \subseteq \ldots$. Dann gilt:*

(i) *Die Vereinigung* $I := \bigcup_{n\in\mathbb{N}} I_n$ *ist wieder ein Ideal.*

(ii) *Ist* R *ein Hauptidealring, dann wird die Idealkette stationär, d. h. es gibt einen Index* n*, sodass für alle* $k \geq n$ *gilt:* $I_k = I_n$*.*

Beweis: (i) Seien a und b Elemente von I. Dann gibt es Indizes k_1 und k_2, sodass $a \in I_{k_1}$ und $b \in I_{k_2}$. Wir dürfen annehmen, dass $k_1 \leq k_2$, d. h. a und b liegen in I_{k_2}. Weil I_{k_2} ein Ideal ist, gehört auch $a + b$ zu I_{k_2}, und wir haben $a + b \in I$.

Sind a ein Element von I und $r \in R$, dann gibt es einen Index k, sodass $a \in I_k$, und weil I ein Ideal ist, gehört auch ra zu I_k. Aber dann ist auch $ra \in I$.

(ii) Da R ein Hauptidealring ist, ist I ein Hauptideal. Das heißt es gibt $a \in R$, sodass $I = \langle a\rangle$. Außerdem gibt es einen Index N, sodass $a \in I_N$. Da dann aber auch $\langle a\rangle \subseteq I_N$, muss $I_k = I_N$ für alle $k \geq N$ gelten. □

Integritätsringe, in denen jede aufsteigende Kette von Idealen stationär werden, heißen Noethersch. Die Forderung, dass alle aufsteigenden Ketten von Idealen in einem Ring stationär werden, ist äquivalent dazu, dass alle Ideale dieses Rings endlich erzeugt sind.

Der Polynomring $K[X_1, \ldots, X_n]$, ein zentrales Objekt des Interesses in der algebraischen Geometrie, ist kein Hauptidealring, aber Noethersch.

Beweis (von Satz 41): (i) Es bezeichne

$$S := \{r \in R - (R^\times \cup \{0\}) \mid \text{Es gibt keine Primfaktorzerlegung von } r\}$$

die Menge der Störenfriede. Angenommen, S wäre nicht leer. Dann gäbe es $x \in S$, und dieses x wäre reduzibel (da wir uns in einem Hauptidealring befinden, wäre ein irreduzibles Element von S auch prim und hätte deshalb eine Zerlegung in Primelemente). Es gäbe also $x_1, y_1 \in R - R^\times$, sodass $x = x_1y_1$. Es müsste $\langle x\rangle \subsetneq \langle x_1\rangle$ gelten, denn sonst wären x und x_1 zueinander assoziiert, d. h. y_1 wäre eine Einheit.

Wären x_1 und y_1 irreduzibel, dann wären beide insbesondere prim, was $x \in S$ widerspräche. Ohne Einschränkung dürfen wir also annehmen, x_1 wäre irreduzibel. Wir fänden also wieder $x_2, y_2 \in R - R^\times$ mit $x_1 = x_2y_2$ und $\langle x_1\rangle \subsetneq \langle x_2\rangle$. Induktiv erhielten wir eine nicht stationär werdende aufsteigende Kette von Idealen in R, was nicht sein kann. Die Menge S muss also leer sein, und jedes Element von $R - (R^\times \cup \{0\})$ hat eine Primfaktorzerlegung.

(ii) Sei $x \in R - (R^\times \cup \{0\})$ und angenommen, es gäbe zwei Zerlegungen $x = p_1 \cdots p_n = q_1 \cdots q_m$ mit Primelementen $p_1, \ldots, p_n, q_1, \ldots, q_m$. Dann müsste

p_1 das Produkt $q_1 \cdots q_m$ teilen. Ohne Einschränkung dürften wir annehmen, dass $p_1 \mid q_1$, d. h. es gäbe $\varepsilon_1 \in R$ mit $q_1 = \varepsilon_1 p_1$. Wegen der Irreduzibelität von q_1 folgte $\varepsilon_1 \in R^\times$, d. h. p_1 wäre assoziiert zu q_1 und wir hätten $p_1 \cdots p_n = \varepsilon_1 p_1 \cdot q_2 \cdots q_m$. Per Kürzungsregel erhielten wir $p_2 \cdot p_n = q_2 \cdots q_m$. Induktiv erhielten wir $n = m$ und für $1 \leq i \leq n$, dass $p_i \sim q_i$. □

Der zweite Teil des Beweises benötigt nur, dass R ein Integritätsbereich ist. Ein Integritätsbereich R, in dem es für jedes $r \in R - (R^\times \cup \{0\})$ eine Primfaktorzerlegung gibt (die dann auch eindeutig bis auf Reihenfolge und Assoziiertheit ist), heißt faktorieller Ring.

5. Moduln

In diesem Abschnitt möchten wir das Konzept eines Vektorraums über einem Körper verallgemeinern zu dem eines Moduls über einem Ring.

Definition IX.5.1: Seien R ein kommutativer Ring mit Eins, $(M, +)$ eine abelsche Gruppe und

$$\cdot \colon R \times M \longrightarrow M$$

eine Abbildung. Falls gilt:

(i) Für alle $m \in M$ ist $1 \cdot m = m$.

(ii) Für alle $r, s \in R$ und $m \in M$ ist $(r + s) \cdot m = r \cdot m + s \cdot m$, $r \cdot (m + n) = r \cdot m + r \cdot n$.

(iii) Für alle $r, s \in R$ und $m \in M$ ist $(r \cdot s) \cdot m = r \cdot (s \cdot m)$,

dann heißen M ein R-Modul und „$\cdot$“ eine *Skalarmultiplikation*

Proposition IX.5.2 (Multiplikation mit 0_M): *Seien R ein kommutativer Ring mit Eins, M ein R-Modul und $r \in R$. Dann ist $r \cdot 0_M = 0_M$. Hierbei bezeichnet natürlich 0_M das neutral Element von M bezüglich „$+$“.*

Beispiel IX.5.3: Sei R ein kommutativer Ring mit Eins.

(i) Ist R ein Körper, dann fallen die Begriffe R-Modul und R-Vektorraum zusammen.

(ii) Für eine natürliche Zahl n wird $R^n := \prod_{i=1}^n R$ mit komponentenweiser Addition und skalarer Multiplikation zu einem Modul über R.

(iii) Der Ring $\mathbb{Z}/n\mathbb{Z}$ wird mit

$$\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, \qquad (a, \bar{b}) \longmapsto \overline{ab}$$

zu einem $\mathbb{Z}$-Modul.

(iv) Eine beliebige abelsche Gruppe $(M, +)$ wird zu einem $\mathbb{Z}$-Modul durch die folgende skalare Multiplikation:

$$\cdot\colon \mathbb{Z} \times M \longrightarrow M, \qquad (k, m) \longmapsto \begin{cases} \sum_{i=1}^{k} m, & \text{falls } k \in \mathbb{N}_0, \\ \sum_{i=1}^{-k} (-m), & \text{falls } -k \in \mathbb{N}. \end{cases}$$

(v) Ist $I \subseteq R$ ein Ideal, dann ist I ein R-Modul mit den Einschränkungen der Addition und Skalarmultiplikation.

Definition IX.5.4 (Modulhomomorphismus): Seien R ein kommutativer Ring mit Eins, M_1 und M_2 Moduln über R und $\varphi\colon M_1 \to M_2$ eine Abbildung. Ist $\varphi\colon (M_1, +) \to (M_2, +)$ ein Gruppenhomomorphismus und gilt für alle $r \in R$, $m_1 \in M_1$, dass $\varphi(rm) = r\varphi(m)$, dann heißt φ ein *R-Modul-Homomorphismus* oder kurz *R-linear*.

Bemerkung IX.5.5: Seien R ein kommutativer unitärer Ring und M_1, M_2 Moduln über R. Wir schreiben

$$\mathrm{Hom}_R(M_1, M_2) := \{\varphi\colon M_1 \to M_2 \text{ linear}\}.$$

Zusammen mit der punktweisen Addition und der punktweisen Skalarmultiplikation bildet $\mathrm{Hom}_R(M_1, M_2)$ selbst einen Modul über R.

Definition IX.5.6 (Untermoduln): Seien R ein kommutativer Ring mit Eins, M ein R-Modul und U eine Teilmenge von M. Ist $(U, +)$ eine Untergruppe von $(M, +)$ und ist für alle $r \in R$ und $m \in U$ auch rm in U, dann heißt U ein *Untermodul*. Insbesondere ist U in diesem Fall selbst ein R-Modul.

Beispiel IX.5.7: (i) Seien $M = \mathbb{Z} = R$ und n eine nichtnegative ganze Zahl. Dann ist $n\mathbb{Z}$ ein Untermodul von M.

(ii) Sind M_1 und M_2 Moduln über dem kommutativen unitären Ring R und $\varphi\colon M_1 \to M_2$ linear, dann sind jeweils

$$\mathrm{Kern}(\varphi) := \{m \in M_1 \mid \varphi(m) = 0_{M_2}\} \subseteq M_1,$$
$$\mathrm{Bild}(\varphi) := \{\varphi(m_1) \mid m_1 \in M_1\} \subseteq M_2$$

Untermoduln.

Bemerkung IX.5.8: Analog zur entsprechenden Aussage für K-Vektorräume haben wir: Sind M ein Modul über dem kommutativen unitären Ring R und U ein Untermodul von M, dann ist M/U[1] ein R-Modul.

Es gilt genau wie für Vektorräume der Homomorphiesatz, d. h. für einen weiteren R-Modul N und eine lineare Abbildung $\varphi\colon M \to N$ mit $U \subseteq \operatorname{Kern}(\varphi)$ gibt es genau eine R-lineare Abbildung $\bar{\varphi}\colon M/U \to N$, sodass $\bar{\varphi}\circ\pi = \varphi$. Hierbei bezeichnet $\pi\colon M \to M/U$ wie immer die kanonische Projektion.

Bemerkung IX.5.9 (Tensorprodukt): Die Begriffe „multilineare Abbildung“, „bilineare Abbildung“ (siehe Definition II.1.1) und „Tensorprodukt“ (siehe Proposition VII.4.6) werden genau so wie für Vektorräume definiert. Der Satz von der Existenz des Tensorprodukts (Satz 29) geht genau so für R-Moduln durch. Genauer: Sind M_1 und M_2 Moduln über R, dann gibt es einen R-Modul $T = M_1 \otimes_R M_2$ und eine bilineare Abbildung $\tau\colon M_1 \times M_2 \to T = M_1 \otimes_R M_2$, sodass gilt: Für jede bilineare Abbildung $h\colon M_1 \times M_2 \to N$ in einen anderen R-Modul N, gibt es genau eine lineare Abbildung $\phi\colon T \to N$ mit $\phi \circ \tau = h$.

Beispiel IX.5.10 (Was so schief gehen kann): Wir betrachten die $\mathbb{Z}$-Moduln $M_1 = \mathbb{Z}/2\mathbb{Z}$ und $M_2 = \mathbb{Z}/3\mathbb{Z}$ sowie eine bilineare Abbildung $h\colon M_1 \times M_2 \to N$ in einen beliebigen anderen $\mathbb{Z}$-Modul N. Für $\bar{a} \in \mathbb{Z}/2\mathbb{Z}$ und $\bar{b} \in \mathbb{Z}/3\mathbb{Z}$ haben wir

$$h(\bar{a},\bar{b}) = h(\bar{3}\cdot\bar{a},\bar{b}) = h(3\bar{a},\bar{b}) = 3h(\bar{a},\bar{b}) = h(\bar{a},3\bar{b}) = h(\bar{a},\bar{0}) = 0_N.$$

Das Tensorprodukt $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z}$ ist also der Nullmodul.

Definition IX.5.11: Sei R ein kommutativer unitärer Ring.

(i) Sei $(A, +, \circ)$ ein weiterer Ring. Gibt es eine Abbildung

$$\cdot\colon R \times A \longrightarrow A, \qquad (r,a) \longmapsto ra,$$

sodass $(A, +, \cdot)$ ein R-Modul und „$\circ$“ R-bilinear ist, dann heißt A eine *R-Algebra*.

(ii) Seien A_1 und A_2 zwei Algebren über R und $\varphi\colon A_1 \to A_2$ eine Abbildung. Falls für alle $a, b \in A_1$ und $r \in R$ gilt, dass

$$\varphi(a+b) = \varphi(a) + \varphi(b), \qquad \varphi(a\circ b) = \varphi(a)\circ\varphi(b), \qquad \varphi(ra) = r\varphi(a),$$

dann heißt φ ein *R-Algebrenhomomorpshismus*.

[1] M/U ist die Menge der Äquivalenzklassen bezüglich folgender Äquivalenzrelation auf M: „$m \sim n :\Leftrightarrow m - n \in U$“.

6. Freie Moduln

Definition IX.6.1: Seien R ein kommutativer Ring mit 1 und M ein R-Modul.

(i) Sei X eine Teilmenge von M. Dann ist

$$\mathrm{Lin}(X) := \langle X\rangle := \bigcap(U \mid U \subseteq M \text{ Untermodul mit } X \subseteq U)$$

ein R-Untermodul von X, genannt *Erzeugnis von* X. Genau wie für Vektorräume haben wir $\mathrm{Lin}(X) = \{\sum_{i=1}^{n} r_i m_i \mid n \in \mathbb{N}_0, r_i \in R, m_i \in X\}$.

(ii) Gibt es in $\mathrm{Lin}(X)$ nur die triviale Nulldarstellung, d. h. ist für eine natürliche Zahl n, $r_1, \ldots, r_n \in R$ und $x_1, \ldots, x_n \in X$ mit $\sum_{i=1}^{n} r_i x_i = \mathbf{0}$ immer schon $r_1 = \cdots = r_n = 0$, dann heißt X *linear unabhängig*.

(iii) Ist B eine linear unabhängige Teilmenge von M mit $\mathrm{Lin}(B) = M$, dann heißt B eine *Basis von* M.

Alle Aussagen in der obigen Definition lassen sich zeigen wie für Vektorräume.

Beispiel IX.6.2 (Modul ohne Basis): Seien n eine natürliche Zahl, R der Ring der ganzen Zahlen $\mathbb{Z}$ und $M = \mathbb{Z}/n\mathbb{Z}$ aufgefasst als $\mathbb{Z}$-Modul. Ist $X = \{\bar{a}\}$ eine Teilmenge von $M - \{\bar{0}\}$, dann ist X bereits linear abhängig, da $n\bar{a} = \overline{na} = \bar{0}$.

Definition IX.6.3 (Freier Modul): Seien R ein kommutatativer unitärer Ring und M ein R-Modul. Hat M eine Basis, dann heißt M *frei*.

Beispiel IX.6.4 (Zwei Moduln oder einer?): (i) Seien n eine natürliche Zahl und R ein kommutativer unitärer Ring. Dann ist R^n ein freier Modul mit Basis $(e_1, \ldots, e_n)$, wobei $e_i := (\delta_{ij})_{1 \le j \le n}$.

(ii) Sei S eine beliebige Menge. Dann ist $\mathrm{Abb}_0(S, R)$ ein freier R-Modul.

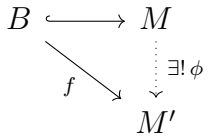
Beweis: Eigentlich müssen wir nur (ii) zeigen, denn (i) folgt dann mit der speziellen Wahl $S = \{1, \ldots, n\}$. Also zu (ii): Für $s \in S$ setze

$$f_s \colon S \longrightarrow R, \qquad t \longmapsto \begin{cases} 1, & \text{falls } t = s, \\ 0, & \text{sonst} \end{cases}$$

und definiere $B := \{f_s \mid s \in S\}$. Dann ist dieses B eine Basis von $\mathrm{Abb}_0(S, R)$.□

Bemerkung IX.6.5: Für eine natürliche Zahl n ist $\mathbb{Z}/n\mathbb{Z}$ aufgefasst als $\mathbb{Z}$-Modul nicht frei; als $\mathbb{Z}/n\mathbb{Z}$-Modul ist er aber frei.

Proposition IX.6.6 (Fortsetzungssatz für Moduln): *Seien R ein kommutativer unitärer Ring, M ein freier R-Modul mit Basis B und M' ein weiterer R-Modul. Dann gibt es für jede Abbildung $f\colon B \to M'$ genau eine R-lineare Abbildung $\phi\colon M \to M'$, die das folgende Diagramm kommutativ macht:*



Der Beweis geht genau wie für Vektorräume. Man benötigt Proposition IX.6.6 für $\mathrm{Abb}_0(M_1 \times M_2, R)$ im Beweis zur Existenz des Tensorprodukts $M_1 \otimes_R M_2$.

7. Kategorientheorie

In diesem Abschnitt möchten wir das Studium algebraischer Strukturen samt ihrer strukturerhaltenden Abbildungen systematisieren.

Definition IX.7.1 (Kategorie): Sei $\mathrm{Obj}(\mathbf{C})$ eine Klasse[2] und es gebe für zwei Elemente A und B von $\mathrm{Obj}(\mathbf{C})$ eine Menge $\mathrm{Mor}_{\mathbf{C}}(A, B)$, deren Elemente wir Morphismen nennen. Falls gilt:

(i) Für Elemente A, B, C von $\mathrm{Obj}(C)$ gibt es eine Abbildung

$$\circ\colon \mathrm{Mor}_{\mathbf{C}}(B, C) \times \mathrm{Mor}_{\mathbf{C}}(A, B) \longrightarrow \mathrm{Mor}_C(A, C), \qquad (g, f) \longmapsto g \circ f.$$

(ii) Für jedes $A \in \mathrm{Obj}(\mathbf{C})$ gibt es eine Abbildung id_A, sodass gilt: Für alle $f \in \mathrm{Mor}_{\mathbf{C}}(A, B)$ ist $f \circ \mathrm{id}_A = f$ und $\mathrm{id}_B \circ f = f$.

(iii) Für Elemente A, B, C, D von $\mathrm{Obj}(\mathbf{C})$ und Morphismen $f \in \mathrm{Mor}_{\mathbf{C}}(A, B)$, $g \in \mathrm{Mor}_{\mathbf{C}}(B, C)$ und $h \in \mathrm{Mor}_{\mathbf{C}}(C, D)$ ist $h \circ (g \circ f) = (h \circ g) \circ f$.

dann heißt $\mathbf{C}$ eine *Kategorie*. Ist $\mathbf{C}$ aus dem Kontext klar, dann schreiben wir auch kurz $\mathrm{Mor}(A, B)$ statt $\mathrm{Mor}_{\mathbf{C}}(A, B)$. Für $f \in \mathrm{Mor}(A, B)$ schreiben wir auch $f\colon A \to B$.

Beispiel IX.7.2 (Viele Kategorien): Das Folgende sind alles Kategorien:

- Die Kategorie der Mengen Set mit Mengen als Objekte und gewöhnlichen Mengenabbildungen als Morphismen.

[2] Eine saubere Einführung des Begriffs „Klasse“ würden Rahmen dieser Vorlesung deutlich sprengen, weshalb wir an dieser Stelle darauf verzichten.

- Die Kategorie der Gruppen Gr mit Gruppen als Objekte und Gruppenhomomorphismen als Morphismen.

Analog werden die nachfolgend aufgelisteten Gebilde Kategorien. Wir schreiben K-Vr für die Kategorie der Vektorräume über einem Körper K, K-Vr^{fin} für die Kategorie der endlichdimensionalen K-Vektorräume, Ri für die Kategorie der Ringe, kuRi für die Kategorie der kommutativen unitären Ringe, R-Mod für die Kategorie der Moduln über dem Ring R, R-Alg die Kategorie der Algebren über einen Ring R und Fields für die Kategorie der Körper.

Sei nun X eine Menge zusammen mit einer Ordnungsrelation „$\leqslant$“. Dann erhalten wir eine Kategorie $\mathbf{C}_1$, wenn wir $\mathrm{Obj}(\mathbf{C}_1) = X$ setzen, und für $x, y \in X$ definieren

$$\mathrm{Mor}_{\mathbf{C}_1}(x,y) := \begin{cases} \{(x,y)\}, & \text{falls } x \leqslant y, \\ \varnothing, & \text{sonst.} \end{cases}$$

Insbesondere ist $\#\,\mathrm{Mor}_{\mathbf{C}_1}(x,y) \leq 1$.

Wir erhalten eine Kategorie $\mathbf{C}_2$ durch $\mathrm{Obj}(\mathbf{C}_2) := \{(X,d) \text{ metrischer Raum}\}$ zusammen mit stetigen Abbildungen als Morphismen, genauer: Für metrische Räume (X_1,d_1) und (X_2,d_2) ist

$$\mathrm{Mor}_{\mathbf{C}_2}((X_1,d_1),(X_2,d_2)) := \{f\colon X_1 \to X_2 \text{ stetig}\}.$$

Definition IX.7.3 (Isomorphismus): Seien $\mathbf{C}$ eine Kategorie, A und B Objekte von $\mathbf{C}$ und $f\colon A \to B$ ein Morphismus. Gibt es einen Morphismus $g\colon B \to A$ mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$, dann heißt f ein *Isomorphismus*.

Beispiel IX.7.4 (für Isomorphismen): Für alle bis auf eine Kategorie aus Beispiel IX.7.2 sind Morphismen $f\colon A \to B$ Isomorphismen, wenn sie bijektiv als Mengenabbildungen sind.

In der Kategorie der metrischen Räume mit stetigen Abbildungen ist das falsch! Seien $X_1 = (0,\infty) \times (-\pi,\pi]$ und $X_2 = \mathbb{R}^2 - \{0\}$ jeweils mit der Einschränkung der euklidischen Metrik. Dann ist die Polarkoordinaten-Abbildung

$$f\colon X_1 \longrightarrow X_2, \qquad (r,\vartheta) \longmapsto (r\cos\vartheta, r\sin\vartheta)^t$$

bijektiv und stetig, aber kein Isomorphismus, da die Umkehrabbildung nicht stetig ist.

Beispiel IX.7.5: Seien n eine natürliche Zahl, R ein kommutativer unitärer Ring und $\mathrm{Gl}_n(R) := \{A \in R^{n\times n} \mid \text{Es gibt } B \in R^{n\times n} \text{ mit } AB = BA = I_n\}$.

Für einen Homomorphismus kommutativer unitärer Ringe $f\colon R_1 \to R_2$ erhalten wir einen Homomorpismus $f'\colon \mathrm{Gl}_n(R_1) \to \mathrm{Gl}_n(R_2)$ per $A = (a_{i,j})_{i,j} \mapsto f'(A) := (f(a_{i,j}))_{i,j}$. Für diese Zuordnung haben wir außerdem $(f_1 \circ f_2)' = f_1' \circ f_2'$.

Definition IX.7.6 (Funktor): Seien $\mathbf{C}_1$ und $\mathbf{C}_2$ zwei Kategorien. Haben wir eine Abbildung $\mathrm{Obj}(\mathbf{C}_1) \to \mathrm{Obj}(\mathbf{C}_2)$, $A \mapsto F(A)$ und für je zwei Objekte $A, B \in \mathrm{Obj}(\mathbf{C}_1)$ eine Abbildung

$$F = F_{A,B}\colon \mathrm{Mor}_{\mathbf{C}_1}(A,B) \longrightarrow \mathrm{Mor}_{\mathbf{C}_2}(F(A),F(B)),$$

sodass für je drei Objekte $A, B, C \in \mathrm{Obj}(\mathbf{C}_1)$ und Morphismen $f \in \mathrm{Mor}(A,B)$ und $g \in \mathrm{Mor}(B,C)$ gilt, dass $F(g \circ f) = F(g) \circ F(f)$, $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$, dann heißt F ein *kovarianter Funktor*. Wir schreiben in dieser Situation oft $F\colon \mathbf{C}_1 \to \mathbf{C}_2$.

Haben wir eine Abbildung F ähnlich wie oben mit

$$F_{A,B}\colon \mathrm{Mor}_{\mathbf{C}_1}(A,B) \to \mathrm{Mor}_{\mathbf{C}_2}(F(B),F(A))$$

und sodass $F(g \circ f) = F(f) \circ F(g)$, dann heißt F ein *kontravarianter Funktor*.

Beispiel IX.7.7 (einige Funktoren): (i) Für eine natürliche Zahl n ist die Zuordnung $\mathrm{Gl}_n\colon \mathsf{kuRi} \to \mathsf{Grp}$ aus Beispiel IX.7.5 ein kovarianter Funktor.

(ii) Die Zuordnung $G_m\colon \mathsf{kuRi} \to \mathsf{Grp}$, die den Ring R auf seine Einheitengruppe $R^\times$ und einen Ringhomomorhismus $f\colon R_1 \to R_2$ auf die Einschränkung $f|_{R_1^\times}\colon R_1^\times \to R_2^\times$ schickt, ist ein kovarianter Funktor.

(iii) Die Zuornung $V\colon \mathsf{Grp} \to \mathsf{Set}$, die die Gruppe G auf die zugrundeliegende Menge G abbildet und den Gruppenhomomorphismus $\varphi\colon G \to H$ schickt auf die Mengenabbildung $\varphi\colon G \to H$, heißt *Vergissfunktor*. Der Vergissfunktor (in diesem Fall für Gruppen) ist ein kovarianter Funktor.

(iv) Die Zuordnung $D\colon K\text{-}\mathsf{VR}^{\mathrm{fin}} \to K\text{-}\mathsf{VR}^{\mathrm{fin}}$ mit $V \mapsto V^*$, und der einen Morphismus $f\colon V \to W$ auf die duale Abbildung $f^*\colon W \to V$ schickt, heißt *Dualisierungsfunktor* und ist ein kontravarianter Funktor („dreht die Pfeile um“), wovon sie sich auf dem kommenden Übungsblatt überzeugen werden.

(v) Sei $\mathbf{C}_2$ die Kategorie der metrischen Räume aus Beispiel IX.7.2. Die Zuordnung $\mathbf{C}_2 \to \mathsf{Grp}$, die auf Objektebene $(X,d) \mapsto \mathrm{Mor}(X,\mathbb{R})$ leistet und einen Morphismus $f\colon (X_1,d_1) \to (X_2,d_2)$ schickt auf $f^*\colon \mathrm{Mor}(X_2,\mathbb{R}) \to \mathrm{Mor}(X_1,\mathbb{R})$, $h \mapsto h \circ f$, ist ein kontravarianter Funktor.

(vi) Seien M_1 und M_2 Moduln über dem kommutativen unitären Ring R. Dann ist die Zuordnung $F\colon R\text{-}\mathsf{Mod} \to \mathsf{Set}$ mit

$$N \longmapsto \mathrm{BL}(M_1 \times M_2, N) := \{h\colon M_1 \times M_2 \to N \text{ bilinear}\}$$
$$(f\colon N_1 \to N_2) \longmapsto (f_*\colon \mathrm{BL}(M_1 \times M_2, N_1) \to \mathrm{BL}(M_1 \times M_2, N_2), \quad B \mapsto f \circ B)$$

ein kovarianter Funktor.

Definition IX.7.8 (Der Hom-Funktor): Seien $\mathbf{C}$ eine Kategorie und $A_0 \in \mathrm{Obj}(\mathbf{C})$ ein Objekt. Dann haben wir folgende beiden Funktoren:

(i) Die Zuordnung $\hom_{A_0}\colon \mathbf{C} \to \mathsf{Set}$ mit

$$C \longmapsto \mathrm{Mor}_{\mathbf{C}}(A_0, C),$$
$$f\colon C_1 \to C_2 \longmapsto \mathrm{Mor}_{\mathbf{C}}(A_0, \mathbf{C}_1) \to (\mathrm{Mor}_{\mathbf{C}}(A_0, \mathbf{C}_2), \quad h \mapsto f \circ h)$$

ist ein kovarianter Funktor.

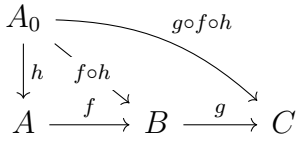
(ii) Die Zuordnung ${}_{A_0}\hom\colon \mathbf{C} \to \mathsf{Set}$ mit

$$C \longmapsto \mathrm{Mor}_{\mathbf{C}}(C, A_0),$$
$$f\colon C_1 \to C_2 \longmapsto (\mathrm{Mor}_{\mathbf{C}}(C_2, A_0) \to \mathrm{Mor}_{\mathbf{C}}(C_1, A_0), \quad h \mapsto h \circ f)$$

ist ein kontravarianter Funktor.

Beweis: Wir zeigen exemplarisch, dass $\hom_{A_0}\colon \mathbf{C} \to \mathsf{Set}$ ein Funktor ist. Für $A \in \mathrm{Obj}(\mathbf{C})$ ist $\hom_{A_0}(\mathrm{id}_A) = (f_*\colon h \mapsto \mathrm{id}_A \circ h = h) = \mathrm{id}_{\mathrm{Mor}(A_0,A)}$.

Für A, B und C aus $\mathrm{Obj}(\mathbf{C})$, $f \in \mathrm{Mor}(A, B)$ und $g \in \mathrm{Mor}(B, C)$ sind wir in der Situation



Für $h \in \mathrm{Mor}(A_0, A) = \hom_{A_0}(A)$ gilt

$$\begin{aligned}\hom_{A_0}(g \circ f)(h) &= (g \circ f) \circ h \\ &= g \circ (f \circ h) \\ &= \hom_{A_0}(g)(f \circ h) \\ &= \hom_{A_0}(g)(\hom_{A_0}(f)(h)) = (\hom_{A_0}(g) \circ \hom_{A_0}(f))(h),\end{aligned}$$

sodass $\hom_{A_0}(g \circ f) = \hom_{A_0}(g) \circ \hom_{A_0}(f)$. Dass ${}_{A_0}\hom$ ebenfalls ein Funktor ist, zeigt man ganz genau so. □

Definition IX.7.9 (Natürliche Transformation): Seien $F, G\colon \mathbf{C}_1 \to \mathbf{C}_2$ zwei kovariante Funktoren sowie $(\alpha_C\colon F(C) \to G(C))_{C \in \mathrm{Obj}(\mathbf{C}_1}$ eine Familie von Morphismen. Kommutiert für jedes $f \in \mathrm{Mor}_{\mathbf{C}_1}(A, B)$ das Diagramm

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ {\scriptstyle \alpha_A}\downarrow & & \downarrow{\scriptstyle \alpha_\beta} \\ G(A) & \xrightarrow[G(f)]{} & G(B) \end{array}$$

dann heißt $\alpha\colon F \to G$ eine *natürliche Transformation*. Ist für jedes Objekt A von $\mathbf{C}_1$ der Morphismus α_A ein Isomorphismus, dann heißt das α eine *funktorielle Äquivalenz*.

Beispiel IX.7.10: Für die beiden Funktoren $F = \mathrm{Gl}_n, G = R^\times\colon \mathsf{kuRi} \to \mathsf{Grps}$ ist $\alpha\colon F \to G$ mit $\alpha_R\colon \mathrm{Gl}_n(R) \to R^\times$, $A \mapsto \det A$ eine natürliche Transformation. Für einen Homomorphismus unitärer Ringe $f\colon R \to S$ haben wir nämlich das Diagramm

$$\begin{array}{ccc} \mathrm{Gl}_n(\mathbb{R}) & \xrightarrow{\mathrm{Gl}_n(f)} & \mathrm{Gl}_n(S) \\ {\scriptstyle\alpha_R}\downarrow & & \downarrow{\scriptstyle\alpha_S} \\ R^\times & \xrightarrow[f|_{R^\times}]{} & S^\times \end{array}$$

das wegen

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

kommutiert.

In der Kategorientherorie lässt man allgemeiner oft zu, dass für Objekte $A, B \in \mathrm{Obj}(\mathbf{C})$ auch $\mathrm{Mor}(A, B)$ Klassen sein dürfen. Falls die Klassen von Morphismen in Wahrheit immer Mengen sind, dann heißen die Kategorien *lokal klein*. Eine Kategorie $\mathbf{C}$, für die $\mathrm{Obj}(\mathbf{C})$ eine Menge ist, heißt *klein*.

Kategorientheorie ist den Vierzigerjahren des letzten Jahrhunderts aus der Topologie heraus erwachsen, die prägenden Namen sind hier MacLane und Eilenberg. Kategorientheorie wird gelegentlich auch von Mathematikern als „allgemeiner Unsinn“ bezeichnet, ist es aber nicht.

8. Universelle Objekte

Erinnerung: Wir starten eine kleine Reise durch den Inhalt der Linearen Algebra und wollen uns an Aussagen erinnern, in denen bereits von einer „universellen Eigenschaft“ die Rede war.

(i) „Der Fortsetzungssatz für lineare Abbildungen“, Satz 12, Lineare Algebra I: Seien n eine natürliche Zahl, $[n] := \{1, \dots, n\}$, K ein Körper und

$$b\colon [n] \longrightarrow K^n, \qquad i \longmapsto e_i = (\delta_{ij})_{1\le j\le n}.$$

Für jeden K-Vektorraum W und jede Abbildung $f\colon [n] \to W$ gibt es genau eine lineare Abbildung $\phi\colon K^n \to W$, sodass $\phi \circ b = f$.

(ii) „Der Homomorphiesatz“, Satz 15, Lineare Algebra I: Seien K ein Körper, V ein K-Vektorraum, $U \subseteq V$ ein Untervektorraum und $\pi\colon V \to V/U$ die kanonische Projektion. Für jede lineare Abbildung $\phi\colon V \to W$ mit $U \subseteq \mathrm{Kern}(\phi)$ gibt es genau eine lineare Abbildung $\psi\colon V/U \to W$, sodass $\psi \circ \pi = \phi$.

(iii) „Das Tensorprodukt“, Satz 29, Lineare Algebra II: Seien R ein kommutativer Ring, M_1, M_2 Moduln über R und $\tau\colon M_1 \times M_2 \to M_1 \otimes_R M_2$ die Tensorabbildung. Für jede bilineare Abbildung $\beta\colon M_1 \times M_2 \to N$ in einen weiteren R-Modul N gibt es genau eine R-lineare Abbildung $\phi_\beta\colon M_1 \otimes_R M_2 \to N$ mit $\phi_\beta \circ \tau = \beta$.

Für all diese Aussagen bietet die Kategorientheorie den richtigen Rahmen für eine befriedigende Beschreibung.

Definition IX.8.1 (Darstellbarkeit, universelles Objekt): Seien $\mathbf{C}$ eine Kategorie und $F\colon \mathbf{C} \to \mathsf{Set}$ ein kovarianter Funktor. Gibt es ein Objekt $C \in \mathrm{Obj}(\mathbf{C})$ und eine funktorielle Äquivalenz $\alpha\colon F \to \hom_C$, dann heißt F *darstellbar mit darstellendem Objekt* C. Das Objekt C heißt auch *universelles Objekt*.

Satz 42 (Tensorprodukt als universelles Objekt): *Seien R ein kommutativer Ring und M_1, M_2 Moduln über R. Der Funktor $F\colon R\text{-}\mathsf{Mod} \to \mathsf{Set}$, der auf Objekten durch $N \mapsto \mathrm{BL}(M_1 \times M_2, N)$ gegeben ist, und auf Morphismen durch $(f\colon M_1 \to M_2) \mapsto (f_*\colon \beta \mapsto f \circ \beta)$, ist darstellbar mit universellem Objekt $M_1 \otimes_R M_2$.*

Beweis: Wir schreiben $C := M_1 \otimes_R M_2 \in \mathrm{Obj}(R\text{-}\mathsf{Mod})$ und definieren eine natürliche Transformation $\alpha\colon F \to \hom_{\mathbf{C}}$ durch

$$\begin{aligned} \alpha_N\colon F(N) = \mathrm{BL}(M_1 \times M_2, N) &\longrightarrow \hom_C(N) = \mathrm{Hom}_R(M_1 \otimes_R M_2, N), \\ \beta &\longmapsto \phi_\beta. \end{aligned}$$

Wir zeigen zunächst, dass α eine natürliche Transformation ist. Für einen Morphismus $f\colon N_1 \to N_2$ haben wir das Diagramm

$$\begin{array}{ccc} \mathrm{BL}(M_1 \times M_2, N_1) & \xrightarrow{F(f)} & \mathrm{BL}(M_1 \times M_2, N_2) \\ \downarrow{\scriptstyle \alpha_{N_1}} & & \downarrow{\scriptstyle \alpha_{N_2}} \\ \mathrm{Hom}_R(M_1 \otimes_R M_2, N_1) & \xrightarrow[\hom_C(f)]{} & \mathrm{Hom}_R(M_1 \otimes M_2, N_2) \end{array}$$

das kommutiert, da $(f \circ \phi_\beta) \circ \tau = f \circ (\phi_\beta \circ \tau) = f \circ \beta \circ \tau = \phi_{f \circ \beta} \circ \tau$. Bleibt zu zeigen, dass α sogar eine funktorielle Äquivalenz ist. Die Abbildung α_W ist ein Isomorphismus Set, da wir die Umkehrabbildung einfach angeben kann:

$$\mathrm{Hom}_\beta(M_1 \otimes_R M_2, N) \longrightarrow \mathrm{BL}(M_1 \times M_2, N), \qquad \phi \mapsto \phi \circ \beta. \qquad \square$$

Kapitel X.

Multilineare Algebra - Teil 2

1. Tensorpotenz und Quotienten

Erinnerung X.1.1: Seien R ein kommutativer unitärer Ring, M und N Moduln über R, n eine nichtnegative ganze Zahl, und $h\colon M^n \to N$ eine multilineare Abbildung.

(i) Gilt für alle $\sigma \in S_n$ und $m_1, \ldots, m_n \in M$, dass

$$h(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) = h(m_1, \ldots, m_n),$$

dann heißt h *symmetrisch*.

(ii) Gilt für alle $m_1, \ldots, m_n \in M$, für die es $i \neq j$ mit $m_i = m_j$ gibt, dass $h(m_1, \ldots, m_n) = \mathbf{0}_N$, dann heißt h *alternierend*.

Wir erinnern uns, dass schiefsymmetrische multilineare Abbildungen spezielle alternierende Abbildungen sind. Wir folgen der Konvention $M^0 = R$, $M^1 = M$ und entsprechend $M^n = M^{n-1} \times M$ für $n \geq 2$.

Definition X.1.2: Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl. Dann schreiben wir:

(i) $\mathrm{Mult}_M^n(N) := \{h\colon M^n \to N \text{ multilinear}\}$,

(ii) $\mathrm{Sym}_M^n(N) := \{h\colon M^n \to N \text{ symmetrisch und multilinear}\}$,

(iii) $\mathrm{Alt}_M^n(N) := \{h\colon M^n \to N \text{ alternierend und multilinear}\}$.

Alle drei sind R-Moduln zusammen mit den punktweisen Verknüpfungen von Abbildungen.

Satz 43 (Tensorpotenzen): *Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl.*

(i) *Es gibt einen R-Modul $T^n(M)$ zusammen mit einer multilinearen Abbildung $t\colon M^n \to T^n(M)$ mit folgender universellen Abbildungseigenschaft: Für alle R-Moduln N und $h \in \mathrm{Mult}^n_M(N)$ gibt es genau eine lineare Abbildung $\phi\colon T^n(M) \to N$ mit $\phi \circ t = h$.*

(ii) *Es gibz einen R-Modul $S^n(M)$ zusammen mit einer symmetrischen multilineare Abbildung $s\colon M^n \to S^n(M)$ mit folgender universellen Abbildungseigenschaft: Für alle R-Moduln N und $h \in \mathrm{Sym}^n_M(N)$ gibt es genau eine lineare Abbildung $\phi\colon S^n \to N$ mit $\phi \circ s = h$.*

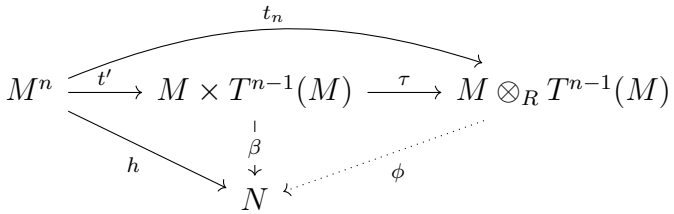
(iii) *Es gibt einen R-Modul $\bigwedge^n(M)$ zuammen mit einer alternierenden multilinearen Abbildung $a\colon M^n \to \bigwedge^n(M)$ mit folgender universellen Abbildungseigenschaft: Für alle R-Moduln N und alle $h \in \mathrm{Alt}^n_M(N)$ gibt es genau eine lineare Abbildung $\phi\colon \bigwedge^n(M) \to N$ mit $\phi \circ a = h$.*

Definition X.1.3: In der Situation von Satz 43 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 43): (i) Wir wollen per vollständiger Induktion vorgehen. Für $n = 0$ setzen wir $T^0(M) := R$, für $n = 1$ setzen wir $T^1(M) := M$ und für $n \geq 2$ setzen wir $T^n(M) := M \otimes_R T^{n-1}(M)$. Ferner setzen wir $t_0 = \mathrm{id}_R$, $t_1 := \mathrm{id}_M$ und für $n \geq 2$ definieren wir

$$t = t_n\colon M^n \longrightarrow T^n(M), \qquad (m_1, \ldots, m_n) \longmapsto m_1 \otimes t_{n-1}(m_2, \ldots, m_n).$$

Nun zeigen wir per vollständiger Induktion die Gültigkeit der universellen Abbildungseigenschaft. Für $n = 0$ und $n = 1$ ist alles klar. Die Aussage gelte nun für ein $n \geq 2$ und es sei $h\colon M^n \to N$ eine multilineare Abbildung. Dann sind wir in der Situation



und wir möchten die Abbildung ϕ konstruieren.

Zur Existenz der Abbildung ϕ: Wir definieren $t'\colon M^n \to M \times T^{n-1}(M)$ durch $(m_1, \ldots, m_n) \mapsto (m_1, t_{n-1}(m_2, \ldots, m_n))$. Für jedes $m_1 \in M$ ist dann $h_{m_1} := h(m_1, \cdot)\colon M^{n-1} \to N$, $(m_2, \ldots, m_n) \mapsto h(m_1, \ldots, m_n)$ multilinear,

d. h., per Induktionsvoraussetzung gibt es einen Vektorraumhomomorphismus $\phi_{m_1}\colon T^{n-1}(M) \to N$ mit $\phi_{m-1} \circ t_{n-1} = h_{m_1}$.

Nun definieren wir $\beta\colon M \times T^{n-1}(M) \to N$ per $(m_1, \omega) \mapsto \phi_{m_1}(\omega)$. Dann ist $\beta \circ t' = h$ und β ist bilinear.

Sei nun $\tau\colon M \times T^{n-1}(M) \to M \otimes_R T^{n-1}(M)$ die Tensorabbildung. Per Konstruktion ist dann auch $\tau \circ t' = t_n$. Wegen der universellen Abbildungseigenschaft des Tensorprodukts existiert eine lineare Abbildung $\phi\colon M \otimes_R T^{n-1}(M) \to N$ mit $\phi \circ \tau = \beta$, d. h. $\phi \circ t_n = \phi \circ \tau \circ t' = \beta \circ t' = h$.

Die Eindeutigkeit von ϕ folgt jetzt aus der Eindeutigkeit von β und ϕ.

(ii) Wir würden gerne als Kandidaten $T^n(M)$ zusammen mit t_n hernehmen. Leider ist t_n im Allgemeinen nicht symmetrisch, d. h. für $\sigma \in S_n - \{\mathrm{id}\}$ und $m_1, \ldots, m_n \in M$ ist im Allgemeinen

$$t(m_1, \ldots, m_n) = m_1 \otimes \cdots \otimes m_n \neq m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} = t(m_{\sigma(1)}, \ldots, m_{\sigma(n)}).$$

Das Problem können wir aber gut aus der Welt schaffen: Wir teilen den Untermodul

$$U_1 := \mathrm{Lin}(\{m_1 \otimes \cdots \otimes m_n - m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)} \mid \sigma \in S_n, m_1, \ldots, m_n \in M\})$$

aus $T^n(M)$ heraus und erhalten $S^n(M) := T^n(M)/U_1$ zusammen mit der kanonischen Projektion $\pi_1\colon T^n(M) \to T^n(M)/U_1 = S^n(M)$. Die Komposition $s := \pi_1 \circ t\colon M^n \to S^n(M)$ ist jetzt per Konstruktion sowohl multilinear, als auch symmetrisch.

Für ein multilineares symmetrisches $h\colon M^n \to N$ haben wir genau eine lineare Abbildung $\phi'\colon T^n(M) \to N$ mit $\phi' \circ t_n = h$. Wegen

$$\begin{aligned}\phi'(m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)}) &= \phi'(t(m_{\sigma(1)}, \ldots m_{\sigma(n)})) \\ &= h(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) \\ &= h(m_1, \ldots, m_n) = \phi'(m_1 \otimes \cdots \otimes m_n)\end{aligned}$$

ist ϕ' symmetrisch, d. h. $\phi'(U_1) = \{\mathbf{0}_N\}$. Der Homomorphiesatz liefert uns jetzt eine lineare Abbildung $\phi\colon S^n(M) \to N$ mit $\phi \circ \pi_1 = \phi'$. Dieses ϕ leistet wegen der universellen Abbildungseigenschaft von $T^n(M)$ und der Surjektivität von π_1 das Gewünschte und ist eindeutig.

(iii) Analog zu (ii) definieren wir

$$U_2 := \mathrm{Lin}(\{m_1 \otimes \cdots \otimes m_n \mid m_1, \ldots, m_n \in M, \exists\, i \neq j : m_i = m_j\})$$

und setzen $\bigwedge^n(M) := T^n(M)/U_2$. Es bezeichne $\pi_2\colon T^n(M) \to \bigwedge^n(M)$ die kanoische Projektion auf den Quotientenvektorraum. Dann erfüllen $\bigwedge^n(M)$ und $a := \pi_2 \circ t$ die behauptete universelle Abbildungseigenschaft. □

Bemerkung X.1.4 (Trivialitäten): Natürlich haben wir $T^0(M) = S^0(M) = \bigwedge^0(M) = R$ sowie $T^1(M) = S^1(M) = \bigwedge^1(M) = M$.

Bemerkung X.1.5: Ähnlich wie beim Tensorprodukt kann jeweils aus der universellen Abbildungseigenschaft zeigen, dass die Potenzen $T^n(M)$, $S^n(M)$ und $\bigwedge^n(M)$ eindeutig bis auf eindeutige Isomorphie sind.

Notation X.1.6: Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl. Dann schreiben wir:

(i) $m_1 \otimes \cdots \otimes m_n = t(m_1, \ldots, m_n)$,

(ii) $m_1 \odot \cdots \odot m_n = s(m_1, \ldots, m_n)$,

(iii) $m_1 \wedge \cdots \wedge m_n = a(m_1, \ldots, m_n)$.

Bemerkung X.1.7 (Erzeuger und Rechenregeln): Seien R ein kommutativer unitärer Ring, M und N Moduln über R und n eine nichtnegative ganze Zahl.

(i) Der R-Modul $T^n(M)$ wird erzeugt von der Menge der reinen Tensoren $\{m_1 \otimes \cdots \otimes m_n \mid m_1, \ldots, m_n \in M\}$. Entsprechend wird $S^n(M)$ erzeugt von $\{m_1 \odot \cdots \odot m_n \mid m_1, \ldots, m_n \in M\}$ und $\bigwedge^n(M)$ von $\{m_1 \wedge \cdots \wedge m_n \mid m_1, \ldots, m_n\}$.

(ii) Für $m_1, \ldots, m_n \in M$, $m_i' \in M$ und $r \in R$ gilt

$$\begin{aligned} &m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i + r m_i' \otimes m_{i+1} \otimes \cdots \otimes m_n \\ &\quad = m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i \otimes m_{i+1} \otimes \cdots \otimes m_n \\ &\qquad + r m_1 \otimes \cdots \otimes m_{i-1} \otimes m_i' \otimes m_{i+1} \otimes \cdots \otimes m_n \end{aligned}$$

und da auch „$\odot$“ sowie „$\wedge$“ multlinear sind, gilt die gleiche Aussage auch für diese Produkte.

(iii) Für alle $m_1, \ldots, m_n \in M$ und $\sigma \in S_n$ gilt

$$m_1 \odot \cdots \odot m_n = m_{\sigma(1)} \odot \cdots \odot m_{\sigma(n)}, \qquad m_1 \wedge \cdots \wedge m_n = \operatorname{sgn}(\sigma) m_{\sigma(1)} \wedge \cdots \wedge m_{\sigma(n)}.$$

Gibt es ferner $i \neq j$ mit $m_i = m_j$, dann ist $m_1 \wedge \cdots \wedge m_n = 0$.

Beispiel X.1.8: Seien $M = \mathbb{R}^3$ zusammen mit der Standardbasis (e_1, e_2, e_3) und Für $v = 3e_1 + 2e_2$, $w = e_1 + e_2 + 5e_3$.

(i) In $T^2(M)$ gilt

$$\begin{aligned} v \otimes w &= (3e_1 + 2e_2) \otimes (e_1 + e_2 + 5e_3) \\ &= 3e_1 \otimes e_1 + 3e_1 \otimes e_2 + 15e_1 \otimes e_3 + 2e_2 \otimes e_1 + 2e_2 \otimes e_2 + 10e_2 \otimes e_3. \end{aligned}$$

(ii) In $S^2(M)$ haben wir

$$\begin{aligned} v \odot w &= (3e_1 + 2e_2) \odot (e_1 + e_2 + 5e_3) \\ &= 3e_1 \odot e_1 + 5e_1 \odot e_2 + 15e_1 \odot e_3 + 2e_2 \odot e_2 + 10e_2 \odot e_3. \end{aligned}$$

(iii) In $\bigwedge^2(M)$ ist

$$v \wedge w = (3e_1 + 2e_2) \wedge (e_1 + e_2 + 5e_3) = e_1 \wedge e_2 + 15e_1 \wedge e_3 + 10e_2 \wedge e_3.$$

Proposition X.1.9: *Seien R ein kommutativer unitärer Ring und M ein freier R-Modul vom Rang r mit Basis $(b_1, \ldots, b_r)$. Dann gilt:*

(i) *$T^n(M)$ ist freier R-Modul mit Basis $\{b_{i_1} \otimes \cdots \otimes b_{i_n} \mid 1 \leq i_1, \ldots, i_n \leq r\}$.*

(ii) *$S^n(M)$ ist freier R-Modul mit Basis $\{b_1^{\nu_1} \odot \cdots \odot b_r^{\nu_r} \mid \sum_{i=1}^r \nu_i = n\}$.*[1]

(iii) *$\bigwedge^n(M)$ ist freier R-Modul mit Basis $\{b_{i_1} \wedge \cdots \wedge b_{i_n} \mid 1 \leq i_1 < \cdots < i_n \leq r\}$.*

Beweis: Dass die Mengen jeweils Erzeugendensysteme sind folgt aus Proposition X.1.7.

(i) Folgt aus Proposition III.3.11 (die dort verwendeten Argumente gehen auch für Moduln durch).

(ii) Fehlt.

(iii) Ist $n > r$, so ist $\bigwedge^n(M) = \{\mathbf{0}\}$, d. h. die Behauptung ist wahr. Ist $n = r$, so stimmt die Behauptung nach Proposition III.6.8.

Sei nun $n < r$ und seien $r_{(i_1,\ldots,i_n)} \in R$ für $1 \leq i_1 < \cdots < i_n \leq r$ mit

$$\sum_{1 \leq i_1 < \cdots < i_n \leq r} r_{(i_1,\ldots,i_n)} b_{i_1} \wedge \cdots \wedge b_{i_n} = \mathbf{0}.$$

Wir wollen für jeden n-Tupel $\boldsymbol{j} = (j_1, \ldots, j_n)$ mit $1 \leq j_1 < \cdots < j_n \leq r$ zeigen, dass $r_{\boldsymbol{j}} = \mathbf{0}$. Wähle dazu $\sigma_{\boldsymbol{j}} \in S_r$ sodass $\sigma_{\boldsymbol{j}}(1) = j_1, \ldots, \sigma_{\boldsymbol{j}}(n) = j_n$ und $\sigma_{\boldsymbol{j}}(n+1), \ldots, \sigma_{\boldsymbol{j}}(r)$ die Werte in $\{1, \ldots, r\} - \{j_1, \ldots, j_n\}$ sind.

Jetzt ist

$$b_{j_1} \wedge \cdots \wedge b_{j_n} \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} = (-1)^\ell b_1 \wedge \cdots \wedge b_r \qquad \text{(X.1)}$$

mit einem Exponent $\ell \in \mathbb{N}$ und es ist $b_{i_1} \wedge \cdots \wedge b_{i_n} \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} = \mathbf{0}$, falls $(i_1, \ldots, i_n) \neq (j_1, \ldots, j_n)$. Einsetzen in Gl. (X.1) gibt

$$\begin{aligned} \mathbf{0} &= \Big(\sum_{1 \leq i_1 < \cdots < i_n \leq r} r_{(i_1,\ldots,i_n)} b_{i_1} \wedge \cdots \wedge b_{i_n} \Big) \wedge b_{\sigma_{\boldsymbol{j}}(n+1)} \wedge \cdots \wedge b_{\sigma_{\boldsymbol{j}}(r)} \\ &= (-1)^\ell \cdot r_{(i_1,\ldots,i_n)} b_1 \wedge \cdots \wedge b_r, \end{aligned}$$

d. h. $r_{(i_1,\ldots,i_n)} = 0$, also die lineare Unabhängigkeit. □

[1] Natürlich meinen wir mit $b_i^{\nu_i}$ das symmetrische Produkt $b_i \odot \cdots \odot b_i$ mit ν_i-vielen Faktoren.

Korollar X.1.10 (Dimensionen der Potenzen): *Seien K ein Körper und V ein d-dimensionaler K-Vektorraum. Dann gilt:*

(i) *$T^n(V)$ ist ein K-Vektorraum der Dimension d^n.*

(ii) *$S^n(V)$ ist ein K-Vektorraum der Dimension $\binom{d+n-1}{n}$.*

(iii) *$\bigwedge^n(V)$ ist ein K-Vektorraum der Dimension $\binom{d}{n}$.*

Beweis: Wir können die in Proposition X.1.9 angegebenen Basen mithilfe der bekannten Urnenmodelle einfach abzählen: Die angegebenen Dimensionen sind die Möglichkeiten, aus einer Urne mit d Kugeln n Kugeln zu ziehen...

- ...mit Zurücklegen und mit Beachtung der Reihenfolge für $T^n(V)$,
- ...mit Zurücklegen und ohne Beachtung der Reihenfolge für $S^n(V)$,
- ...ohne Zurücklegen und ohne Beachtung der Reihenfolge für $\bigwedge^n(V)$. □

Proposition X.1.11 (Induzierte Morphismen auf Potenzen): *Seien R ein kommutativer unitärer Ring, M_1 und M_2 Moduln über R und $\varphi\colon M_1 \to M_2$ eine Homomorphismus von Moduln über R. 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 $v_1, \dots, v_n$ gilt:

(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: Wir zeigen exemplarisch die Behauptung aus (iii). Dazu definieren wir die alternierende multilineare Abbildung

$$h\colon M_1^n \longrightarrow \bigwedge^n(M_2), \qquad (m_1, \dots, m_n) \longmapsto \varphi(m_1) \wedge \cdots \wedge \varphi(m_n)$$

und erhalten aus der universellen Abbildungseigenschaft von $\bigwedge^n(M_1)$ eine lineare Abbildung $\bigwedge^n(\varphi)\colon \bigwedge^n(M_1) \to \bigwedge^n(M_2)$, die auf reinen Tensoren (und damit überall) das Gewünschte leistet. □

Proposition X.1.12 (Funktorialität): *Seien R ein kommutativer unitärer Ring, M_1, M_2 und M_3 Moduln über R und $\varphi_1\colon M_1 \to M_2$ sowie $\varphi_2\colon M_2 \to M_3$ Homomorphismen von Moduln über R. Dann gilt:*

(i) $T^n(\varphi_2 \circ \varphi_1) = T^n(\varphi_2) \circ T^n(\varphi_1)$, $S^n(\varphi_2 \circ \varphi_1) = S^n(\varphi_2) \circ S^n(\varphi_1)$ *sowie* $\bigwedge^n(\varphi_2 \circ \varphi_1) = \bigwedge^n(\varphi_2) \circ \bigwedge^n(\varphi_1)$.

(ii) *Für jeden R-Modul M gilt* $T^n(\mathrm{id}_M) = \mathrm{id}_{T^n(M)}$, $S^n(\mathrm{id}_M) = \mathrm{id}_{S^n(M)}$, $\bigwedge^n(\mathrm{id}_M) = \mathrm{id}_{\bigwedge^n(M)}$.

Beweis: Wir zeigen exemplarisch die Aussagen über die Tensorpotenz. Für alle $m_1, \ldots, m_n \in M$ haben wir

$$\begin{aligned}(T^n(\varphi_2) \circ T^n(\varphi_1))(m_1 \otimes \cdots \otimes m_n) &= T^n(\varphi_2)(\varphi_1(m_1) \otimes \cdots \otimes \varphi_1(m_1)) \\ &= \varphi_2(\varphi_1(m_1)) \otimes \cdots \otimes \varphi_2(\varphi_1(m_n)) \\ &= T^n(\varphi_2 \circ \varphi_1)(m_1 \otimes \cdots \otimes m_n)\end{aligned}$$

und da die reinen Tensoren die Tensorpotenz $T^n(M_1)$ erzeugen, wissen wir, dass die Abbildungen auch insgesamt übereinstimmen. □

2. Determinante und äußere Potenz

Erinnerung X.2.1: Seien R ein kommutativer unitärer Ring und n eine nichtnegative ganze Zahl.

(i) Wir hatten in der Linearen Algebra I die Determinante als die Abbildung

$$\det\colon \prod_{i=1}^{n} R^n \longrightarrow R, \qquad (m_1, \ldots, m_n) \longmapsto \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) m_{1,\sigma(1)} \cdots m_{n,\sigma(n)}$$

eingeführt und gesehen, dass die Determinante eine alternierende multilineare Abbildung mit $\det(e_1, \ldots, e_n) = 1$ ist.

(ii) Auf $R^{n\times n}$ haben wir die Determinante definiert als $\det\colon R^{n\times n} \to R$, $A \mapsto \det(Ae_1, \ldots, Ae_n)$.

Bemerkung X.2.2 („Es kann nur eine geben“): Seien R ein kommutativer unitärer Ring und n eine nichtnegative ganze Zahl.

(i) Wegen der universellen Abbildungseigenschaft von $\bigwedge^n(R^n)$ gibt es genau eine lineare Abbildung $\phi_{\det}\colon \bigwedge^n(R^n) \to R$ mit $\phi_{\det} = a \circ \det$.

(ii) Die äußere Potenz $\bigwedge^n(R^n)$ ist ein freier R-Modul vom Rang 1 mit Basis $\{e_1 \wedge \cdots \wedge e_n\}$, d. h. es gibt genau eine lineare Abbildung $\phi_0 \colon \bigwedge^n(R^n) \to R$ mit $\phi_0(e_1 \wedge \cdots \wedge e_n) = 1$; es ist also $\phi_0 = \phi_{\det}$ und wir erhalten unsere ursprüngliche Determinante zurück als $\det = \phi_0 \circ a$.

(iii) Aus (i) und (ii) folgt: $\phi_{\det}$ ist die eindeutige lineare Abbildung von $\bigwedge^n(R^n) \to R$ mit $\phi_{\det}(e_1 \wedge \cdots \wedge e_n) = 1$ und $\det = \phi_{\det} \circ a$.

Der dritte Punkt der vorangegangenen Bemerkung wird auch verwendet, um $\det \colon \prod_{i=1}^n R^n \to R$ und damit auch $\det \colon R^{n\times n} \to R$ zu definieren. Insbesondere folgt dann:

$$\det A = \det(Ae_1, \ldots, Ae_n) = \phi_{\det}(Ae_1 \wedge \cdots \wedge Ae_n).$$

Proposition X.2.3 (Rechenregel für det A): *Seien R ein kommutativer unitärer Ring und $A \in R^{n\times n}$ eine Matrix. Dann gilt:*

$$Ae_1 \wedge \cdots \wedge Ae_n = \det Ae_1 \wedge \cdots \wedge e_n.$$

Beweis: Unsere Abbildung $\phi_{\det} \colon \bigwedge^n(R^n) \to R$ ist ein Isomorphismus mit Umkehrabbildung $\phi' \colon R \to \bigwedge^n(R^n)$ mit $\phi'(r) = re_1 \wedge \cdots \wedge e_n$. Damit ist

$$\begin{aligned} Ae_1 \wedge \cdots \wedge Ae_n &= (\phi' \circ \phi_{\det})(Ae_1 \wedge \cdots \wedge Ae_n) \\ &= \phi'(\det(A)) = \det(A)e_1 \wedge \cdots \wedge e_n. \end{aligned}$$ □

Für die nächste Aussage stellen wir zu nächst eine Vorüberlegung an. SInd R ein kommutativer unitärer Ring, M ein freier R-Modul vom Rang 1 und φ ein Endomorphismus von M, dann gibt es irgendein $r \in R$, sodass $\varphi(m) = rm$ für alle $m \in M$.

Bemerkung X.2.4: Seien R ein kommutativer unitärer Ring, M ein freier R-Modul vom Rang n und φ ein Endomorphismus von M. Dann induziert φ eine R-lineare Abbildung $\bigwedge^n(\varphi) \colon \bigwedge^n(M) \to \bigwedge^n(M)$. Da $\bigwedge^n(M)$ frei und vom Rang 1 ist, gibt es $r \in R$, sodass $\bigwedge^n(\varphi)$ als Endomorphismus von $\bigwedge^n(M)$ von der Form $\bigwedge^n(\varphi)(m_1 \wedge \cdots \wedge m_n) = rm_1 \wedge \cdots \wedge m_n$ ist. Wir setzen $\det \varphi := r$. Insbesondere gilt dann:

(i) Für alle $m_1, \ldots, m_n \in M$ ist

$$\varphi(m_1) \wedge \cdots \wedge \varphi(m_n) = (\bigwedge^n \varphi)(m_1 \wedge \cdots \wedge m_n) = \det \varphi(m_1 \wedge \cdots \wedge m_n).$$

und $\phi_{\det}(\varphi(m_1) \wedge \cdots \wedge \varphi(m_n)) = \phi_{\det}(\det \varphi(m_1 \wedge \cdots \wedge m_n)) = \det \varphi \phi_{\det}(m_1 \wedge \cdots \wedge m_n)$.

(ii) Für $M = R^n$, $A \in R^{n\times n}$ und $\varphi = \varphi_A \colon x \mapsto Ax$ ist $\det(A) = \det(\varphi_A)$, denn

$$\begin{aligned}\det(A) &= \phi_{\det}(Ae_1 \wedge \cdots \wedge Ae_n)\\ &= \phi_{\det}(\varphi_A(e_1) \wedge \cdots \wedge \varphi_A(e_n)) = \det(\varphi_A)\phi_{\det}(e_1 \wedge \cdots \wedge e_n) = \det(\varphi_A).\end{aligned}$$

Proposition X.2.5 (Regenregeln für Determinante): *Seien n eine nichtnegative ganze Zahl, R ein kommutativer unitärer Ring, M ein freier Modul vom Rang n, $A \in R^{n\times n}$ eine Matrix und $\varphi \in \mathrm{End}(M)$ ein Endomorphismus. Dann gilt:*

(i) *Für alle $x_1, \ldots, x_n \in R$ ist $Ax_1 \wedge \cdots \wedge Ax_n = \det(A)(x_1 \wedge \cdots \wedge x_n)$.*

(ii) *Für alle $m_1, \ldots, m_n \in M$ ist $\varphi(m_1)\wedge\cdots\wedge\varphi(m_n) = (\det\varphi)(m_1\wedge\cdots\wedge m_n)$.*

(iii) *Für $A_1, A_2 \in R^{n\times n}$ ist $\det(A_1A_2) = \det(A_1)\det(A_2)$.*

(iv) *Für $\varphi_1, \varphi_2 \in \mathrm{End}(M)$ ist $\det(\varphi_1 \circ \varphi_2) = \det(\varphi_1)\det(\varphi_2)$.*

(v) *Ist $B = (b_1, \ldots, b_n)$ eine Basis von M und $A = D_{B,B}(\varphi)$, dann ist $\det(A) = \det(\varphi)$.*

Beweis: (i) Folgt aus Bemerkung V.2.4 (ii) und Aussage (ii).

(ii) Folgt aus Bemerkung V.2.4 (i).

(iii) Folgt aus (iv) und Bemerkung V.2.4(ii).

(iv) Folgt aus Proposition V.1.12.

(v) Sei $D_B \colon M \to R^n$ die Koordinatenabbildung. Wir haben das kommutative Diagramm

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & M \\ {\scriptstyle D_B}\downarrow & & \downarrow{\scriptstyle D_B} \\ R^n & \xrightarrow[\varphi_A]{} & R^n \end{array}$$

d. h. $\det(A) = \det(\varphi_A) = \det(D_B \circ \varphi \circ D_B^{-1}) = \det(D_B)\det\varphi\det D_B^{-1} = \det\varphi$ was wir zeigen wollten. □

3. Tensoralgebra, symmetrische Algebra und äußere Algebra

Proposition X.3.1: *Seien R ein kommutativer unitärer Ring, M ein Modul über R und k und ℓ nichtnegative ganze Zahlen. Dann gibt es eine eindeutige*

bilineare Abbildung

$$h_{k,\ell}\colon T^k(M) \times T^\ell(M) \longrightarrow T^{k+\ell}(M),$$
$$h_{k,\ell}(m_1 \otimes \cdots \otimes m_k, m'_1 \otimes \cdots \otimes m'_\ell) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell.$$

Beweis: Wir zeigen die Behauptung mithilfe der universellen Abbildungseigenschaften der Tensorpotenzen. Dafür halten wir zunächst den zweiten Teil fest. Für jedes $m' = (m'_1, \ldots, m'_\ell) \in M^\ell$ sei $h_{m'}\colon M^k \to T^{k+\ell}(M)$, $(m_1, \ldots, m_k) \mapsto m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell$. Da $h_{m'}$ multilinear ist, erhalten wir genau eine lineare Abbildung $\varphi_{m'}\colon T^k(M) \to T^{k+\ell}(M)$ mit $\varphi_{m'}(m_1 \otimes \cdots \otimes m_k) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell$.

Nun erhalten wir eine multilineare Abbildung

$$h\colon M^\ell \longrightarrow \mathrm{Hom}_R(T^k(M), T^{k+\ell}(M)), \qquad m' \longmapsto \varphi(m')(m),$$

für die es wegen der universellen Abbildungseigenschaft eine eindeutige lineare Abbildung $\varphi\colon T^\ell(M) \to \mathrm{Hom}_R(T^k(M), T^{k+\ell}(M))$ mit $\varphi(m'_1 \otimes \cdots \otimes m'_\ell) = \varphi_{m'}$.

Definieren wir nun die Abbildung $h_{k,\ell}\colon T^k(M) \times T^\ell(M) \to T^{k+\ell}(M)$ über $(m, m') \mapsto \varphi_{m'}(m)$, dann gilt einerseits

$$\begin{aligned} h_{k,\ell}(m_1 \otimes \cdots \otimes m_k, m'_1 \otimes \cdots \otimes m'_\ell) &= \varphi_{m'}(m_1 \otimes \cdots \otimes m_k) \\ &= m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell \end{aligned}$$

und andererseits ist $h_{k,\ell}$ per Definition bilinear. □

Definition X.3.2 (Unendliche SUmme von *R*-Moduln): Seien R ein kommutativer unitärer Ring, I eine nichtleere Menge und für jedes $i \in I$ sei M_i ein Modul über R. Dann heißt

$$\begin{aligned} \bigoplus_{i\in I} M_i &:= \{(m_i)_{i\in I} \mid \text{Für jedes } i \in I \text{ ist } m_i \in M_i, \text{ nur endlich viele } m_i \neq 0\} \\ &:= \{f \in \mathrm{Abb}_0(I, \bigcup_{i\in I} M_i) \mid \text{Für jedes } i \in I \text{ ist } f(i) \in M_i\} \end{aligned}$$

direkte Summe der M_i. Mit komponentenweiser Addition und skalarer Multiplikation wird $\bigoplus_{i\in I} M_i$ zu einem R-Modul. Für die Familie $(m_i)_{i\in I}$ schreiben wir $\sum_{i\in I} m_i$ und halten fest, dass es sich per Definition immer um eine endliche Summe handelt.

Gelegentlich nennt man diese direkte Summe auch äußere direkte Summe, da sie – anders als die bisher aufgetretene direkte Summe – a priori nicht innerhalb eines größeren Vektorraums stattfindet. Tatsächlich lebt $\bigoplus_{i\in I} M_i$ im direkten Produkt $\prod_{i\in I} M_i$, dem im Gegensatz zur direkten Summe die Endlichkeitsforderung für die enthaltenen Familien fehlt. In diesem größeren Vektorraum fallen der alte und der neue Begriff für direkte Summe zusammen.

Definition X.3.3 (Die Hauptakteure): Seien R ein kommutativer unitärer Ring und M ein Modul über R. Dann setzen wir:

(i) $T(M) := \bigoplus_{n\in\mathbb{N}_0} T^n(M)$,

(ii) $S(M) := \bigoplus_{n\in\mathbb{N}_0} S^n(M)$,

(iii) $\bigwedge(M) := \bigoplus_{n\in\mathbb{N}_0} \bigwedge^n(M)$.

Wir nennen $T(M)$ die *Tensoralgebra von* M, $S(M)$ die *symmetrische Algebra von* M und $\bigwedge(M)$ die *Grassmann-Algebra von* M.

Wir werden im Laufe des Abschnittes sehen, dass die Bezeichnung der obigen R-Moduln als Algebren gerechtfertigt ist, d. h. wir werden die oben definierten R-Moduln mit entsprechenden Multiplikationen ausstatten.

Bemerkung X.3.4: Seien R ein kommutativer unitärer Ring und M ein Modul über R.

(i) Für $j \in I$ ist $\iota_j\colon M_i \hookrightarrow \bigoplus_{i\in I} M_i$, $m_j \mapsto (\delta_{ij}m_j)_{i\in I}$ ein injektiver R-Modulhomomorphismus. Vermöge ι_j fassen wir M_j als R-Untermodul von $\bigoplus_{i\in I} M_i$ auf.

(ii) Als R-Modul wird $T(M)$ von $\{m_1 \otimes \cdots \otimes m_n \mid n \in \mathbb{N}_0, m_i \in M\}$ erzeugt. Genau so werden $S(M)$ respektive $\bigwedge(M)$ von den reinen Tensoren endlicher Länge erzeugt.

Definition X.3.5 (Multiplikation auf $T(M)$): Seien R ein kommutativer unitärer Ring und M ein Modul über R. Auf $T(M)$ erklären wir eine Multiplikation wie folgt: Für $m = \sum_{i\in\mathbb{N}_0} m_i$, $m' = \sum_{j\in\mathbb{N}_0} m'_j$ setzen wir

$$m \otimes m' := \sum_{i\in\mathbb{N}_0} \sum_{j\in\mathbb{N}_0} h_{i,j}(m_i, m'_j).$$

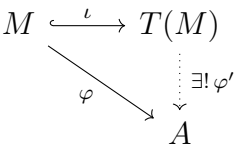
Insbesondere gilt für die reinen Tensoren $m_1 \otimes \cdots \otimes m_k$ und $m'_1 \otimes \cdots \otimes m'_\ell$, dass $(m_1 \otimes \cdots \otimes m_k) \otimes (m'_1 \otimes \cdots \otimes m'_\ell) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell$.

Bemerkung X.3.6: Seien R ein kommutativer unitärer Ring und M ein Modul über R.

(i) Mit der Multiplikation aus Proposition X.3.5 wird $T(M)$ zu einer unitären R-Algebra. Dabei kommt die Eins aus dem Ring R.

(ii) Auf analoge Weise erhalten wir Multiplikationen „$\odot$“ und „$\wedge$“ auf $S(M)$ und $\bigwedge(M)$, die $S(M)$ respektive $\bigwedge(M)$ zu R-Algebren machen.

Satz 44 (Universelle Abbildungseigenschaft): *Seien R ein kommutativer unitärer Ring und M ein Modul über R. Ist A eine unitäre R-Algebra und ist $\varphi\colon M \to A$ eine R-lineare Abbildung, dann gibt es genau einen Homomorpismus von unitären R-Algebren $\varphi'\colon T(M) \to A$, der das Diagramm*



kommutativ macht. Hierbei bezeichnet $\iota = \iota_1$ die Einbettung aus Proposition X.3.4.

Ist $\varphi\colon M \to A$ eine R-lineare Abbildung, die zusätzlich für alle $x, y \in M$ erfüllt, dass $\varphi(x)\varphi(y) = \varphi(y)\varphi(x)$, dann gibt es genau einen Homomorphismus von unitären R-Algebren $\varphi'\colon S(M) \to A$, sodass $\varphi' \circ \iota = \varphi$.

Ist $\varphi\colon M \to A$ eine R-lineare Abbildung, die zusätzlich für alle $x \in M$ erfüllt, dass $\varphi(x)\varphi(x) = 0$, dann gibt es genau einen Homomorphismus von unitären R-Algebren $\varphi'\colon \bigwedge(M) \to A$, sodass $\varphi' \circ \iota = \varphi$.

Beweis: Für diesen Beweis schreiben wir $t_n\colon M^n \to T^n(M)$, $(m_1, \ldots, m_n) \mapsto m_1 \otimes \cdots \otimes m_n$. Wir zeigen exemplarisch die erste universelle Abbildungseigenschaft. Die Strategie dafür wird sein, induktiv Abbildungen auf den Potenzen zu erklären, und dann zu einer großen Abbildung zusammenzusetzen.

Für $n = 0$ definieren wir $\varphi'^{(0)}\colon R \to A$ durch $r \mapsto r1_A$ und für $n = 1$ definieren wir $\varphi'^{(1)}\colon M \to A$ durch $m \mapsto \varphi(m)$. Für $n \geq 2$ wird $\varphi'^{(n)}\colon T^n(M) \to A$ von der multilinearen Abbildung

$$M^n \longrightarrow A, \qquad (m_1, \ldots, m_n) \longmapsto \varphi(m_1) \cdots \varphi(m_n)$$

induziert; insbesondere gilt $\varphi'^{(n)}(m_1 \otimes \cdots \otimes m_n) = \varphi(m_1) \cdots \varphi(m_n)$. Nun erhalten wir eine R-lineare Abbildung

$$\varphi'\colon T(M) \longrightarrow A, \quad \text{durch} \quad \sum_{i \in \mathbb{N}_0} \longmapsto \sum_{i \in \mathbb{N}_0} \varphi'^{(i)}(m_i).$$

Es bleibt zu zeigen, dass φ' auch ein Homomorphismus von R-Algebren ist, d. h. dass für $m, m' \in T(M)$ gilt, dass $\varphi'(m \otimes m') = \varphi'(m)\varphi'(m')$. Seien also

$m = \sum_{i\in\mathbb{N}_0} m_i$ und $m' = \sum_{j\in\mathbb{N}_0} m'_j$ Elemente von $T(M)$. Dann haben wir

$$\begin{aligned}\varphi'(m\otimes m') &= \varphi'\Big(\sum_{i\in\mathbb{N}_0}\sum_{j\in\mathbb{N}_0} m_i\otimes m'_j\Big)\\ &= \sum_{i\in\mathbb{N}_0}\sum_{j\in\mathbb{N}_0} \varphi'^{(i+j)}(m_i\otimes m'_j)\\ &= \sum_{i\in\mathbb{N}_0}\sum_{j\in\mathbb{N}_0} \varphi'^{(i)}(m_i)\varphi'^{(j)}(m'_j)\\ &= \Big(\sum_{i\in\mathbb{N}_0}\varphi'^{(i)}(m_i)\Big)\Big(\sum_{j\in\mathbb{N}_0}\varphi'^{(j)}(m'_j)\Big) = \varphi'(m)\varphi'(m').\end{aligned}$$

Der entscheidende Schritt war hierbei die folgende Gleichheit bilinearer Abbildungen von $T^i(M)\times T^j(M)$ nach A:

$$\varphi'^{(i+j)}\circ h_{i,j} = ((m_i, m'_j)) \mapsto \varphi'^{(i)}(m_i)\varphi'^{(j)}(m'_j).$$

Da beide Abbildungen bilinear sind, genügt es die Behauptung auf Erzeugern der Tensorpotenzen nachzuweisen. Seien also $m_1,\dots,m_n$ und $m'_1,\dots,m'_j$ Elemente von M. Dann ist

$$\begin{aligned}&\varphi'^{(i+j)}(h_{i,j}(m_1\otimes\cdots\otimes m_i, m'_1\otimes\cdots\otimes m'_j))\\ &\quad= \varphi'^{(i+j)}(m_1\otimes\cdots\otimes m_i\otimes m'_1\otimes\cdots\otimes m'_j)\\ &\quad= \varphi(m_1)\cdots\varphi(m_i)\cdot\varphi(m'_1)\cdots\varphi(m'_j)\\ &\quad= \varphi'^{(i)}(m_1\otimes\cdots\otimes m_i)\varphi'^{(j)}(m'_1\otimes\cdots\otimes m'_j)\end{aligned}$$

und wir sind fertig. Die Beweise für $S(M)$ und $\bigwedge(M)$ funktionieren analog. □

Proposition X.3.7 (Funktorialität): *Seien R ein kommutativer unitärer Ring und M_1, M_2 und M_3 Moduln über R. Ist $f\colon M_1\to M_2$ eine lineare Abbildung, dann gibt es eindeutige R-lineare Abbildungen $T(f)\colon T(M_1)\to T(M_2)$ bzw. $S(f)\colon S(M_1)\to S(M_2)$ bzw. $\bigwedge(f)\colon \bigwedge(S_1)\to\bigwedge(S_2)$, sodass*

$$\begin{aligned}T(f)(m_1\otimes\cdots\otimes m_k) &= f(m_1)\otimes\cdots\otimes f(m_k),\\ S(f)(m_1\odot\cdots\odot m_k) &= f(m_1)\odot\cdots\odot f(m_k),\\ \textstyle\bigwedge(f)(m_1\wedge\cdots\wedge m_k) &= f(m_1)\wedge\cdots\wedge f(m_k).\end{aligned}$$

Sind $f_1\colon M_1\to M_2$ und $f_2\colon M_2\to M_3$ Homomorphismen von R-Moduln, dann gilt $T(f_2\circ f_1) = T(f_2)\circ T(f_1)$ bzw. $S(f_2\circ f_1) = S(f_2)\circ S(f_1)$ bzw. $\bigwedge(f_2\circ f_1) = \bigwedge(f_2)\circ\bigwedge(f_1)$.

Für die identische Abbildung $\mathrm{id}_{M_1}\colon M_1\to M_1$ gilt $T(\mathrm{id}_{M_1}) = \mathrm{id}_{M_1}$ bzw. $S(\mathrm{id}_{M_1}) = \mathrm{id}_{M_1}$ bzw. $\bigwedge(\mathrm{id}_{M_1}) = \mathrm{id}_{M_1}$.

Beweis: Die erste Eigenschaft folgt aus Satz 44 für $A = T(M)$, $S(M)$ bzw. $\bigwedge(M)$. Die anderen beiden Eigenschaften folgen aus der Eindeutigkeit in der universellen Abbildungseigenschaft. □

4. Polynomringe in mehreren Variablen

Definition X.4.1 (Polynomring in n Variablen): Seien R ein kommutativer unitärer Ring, n eine natürliche Zahl und $I := \mathbb{N}_0^n$. Die Menge

$$\begin{aligned} R[X_1, \ldots, X_n] &:= \{(a_i)_{i \in I} \in R^I \mid \text{Nur für endlich viele } i \in I \text{ ist } a_i \neq 0\} \\ &:= \bigoplus_{i \in I} R = \mathrm{Abb}_0(I, R) \end{aligned}$$

wird mit den Verknüpfungen $(a_i)_{i \in I} + (b_i)_{i \in I} := (a_i + b_i)_{i \in I}$, $r(a_i)_{i \in I} := (ra_i)_{i \in I}$ und $(a_i)_{i \in I} \cdot (b_i)_{i \in I} := (\sum_{j+k=i} a_j b_k)_{i \in I}$, wobei $(a_i)_{i \in I}$, $(b_i)_{i \in I} \in R[X_1, \ldots, X_n]$ und $r \in R$, zu einer kommutativen R-Algebra mit Eins.

Proposition X.4.2 (Einsetzungshomomorphismus): *Seien R ein kommutativer unitärer Ring, A eine unitäre R-Algebra und $a_1, \ldots, a_n \in A$. Dann gibt es genau einen Homomorphismus von R-Algebren $\varphi\colon R[X_1, \ldots, X_n] \to A$, sodass $\varphi(X_1) = a_1, \ldots, \varphi(X_n) = a_n$. Der Homomorphismus φ heißt* Einsetzungshomomorphismus zu $a_1, \ldots, a_n$.

Beweis: Man zeigt die Aussage analog zum Einsetzungshomomorphismus für den Polynomring in einer Variable, d. h. man definiert φ durch

$$p = \sum_{i \in I} r_i X_1^{i_1} \cdots X_n^{i_n} \longmapsto \sum_{i \in I} r_i a_1^{i_1} \cdots a_n^{i_n}.$$

Alternativ kann man die Aussage per Induktion beweisen. □

Satz 45: *Seien n eine natürliche Zahl und R ein kommutativer unitärer Ring. Dann sind $S(R^n)$ und $R[X_1, \ldots, X_n]$ als R-Algebren isomorph.*

Beweis: Es bezeichne $(e_1, \ldots, e_n)$ die Standardbasis des R^n. Per Fortsetzungssatz erhalten wir eine eindeutige R-lineare Abbildung $\psi\colon R^n \to R[X_1, \ldots, X_n]$ mit $\psi(e_i) = X_i$. Wegen der universellen Abbildungseigenschaft der symmetrischen Algebra gibt es genau einen Homomorphismus unitärer R-Algebren $\psi'\colon S(R^n) \to R[X_1, \ldots, X_n]$ mit $\psi' \circ \iota = \psi$.

Weil $S(R^n)$ eine kommutative unitäre R-Algebra ist, gibt es wegen des Einsetzungshomomorphismus für den Polynomring in n Variablen genau einen Homomorphismus unitärer R-Algebren $\varphi'\colon R[X_1, \ldots, X_n] \to S(R^n)$ mit $\varphi'(X_1) = e_1$, $\ldots$, $\varphi'(X_n) = e_n$ und $\varphi'(1) = 1$.

Wegen der Eindeutigkeit in der universellen Abbildungseigenschaft für $S(R^n)$ und der Eindeutigkeit im Einsetzungshomomorphismus gelten sowohl $\varphi' \circ \psi' = \mathrm{id}$ und $\psi' \circ \varphi' = \mathrm{id}$. $\square$

Kapitel XI.

Unendlichdimensionale Vektorräume und das Zornsche Lemma

1. Motivation

Seien K ein Körper, V ein Vektorraum über K und B eine Basis von V. In der Linearen Algebra I haben wir gesehen, dass jedes Element von V eine eindeutige Linearkombination von Vektoren aus B hat, was gleichbedeutend dazu war, dass B sowohl linear unabhängig als auch erzeugend war.

Erinnerung XI.1.1: Seien K ein Körper, V ein K-Vektorraum und $B \subseteq V$ eine Teilmenge. Ist B linear unabhängig und maximal bezüglich Inklusion, dann ist B eine Basis.

Für den Beweis haben wir nirgends benötigt, dass V endlich erzeugt wäre. Für nicht endlich erzeugte Vektorräume müssen wir uns allerdings noch Gedanken machen, warum es überhaupt maximal linear unabhängige Teilmengen geben sollte.

Für einen nicht endlich erzeugten K-Vektorraum V betrachten wir also die partiell geordnete Menge $(\mathfrak{P}(V), \subseteq)$, und versuchen maximale Elemente ihrer Teilmenge $\mathfrak{S} := \{S \subseteq V \mid S \text{ ist linear unabhängig}\}$ zu finden.

Als ersten naheliegenden Ansatz setzen wir $B' := \bigcup_{S \in \mathfrak{S}} S$. Dann ist jedes S aus $\mathfrak{S}$ enthalten in B'. Dieses B' ist eine kleinste obere Schranke von $\mathfrak{S}$. Aber ist dieses B' weiterhin linear unabhängig?

Als zweiten Ansatz betrachten wir eine Teilmenge $K \subseteq \mathfrak{S}$, sodass K mit der Einschränkung der Ordnung von S total geordnet ist – solche K nennt man auch *Kette*. Setzen wir $B'_K := \bigcup_{S \in K} S$, dann sind alle S aus K in B'_K enthalten und außerdem können wir zeigen, dass B'_K linear unabhängig ist.

Proposition XI.1.2: *Sei K eine Kette in $\mathfrak{S}$. Dann hat K eine obere Schranke in $\mathfrak{S}$.*

Beweis: Seien $v_1, \dots, v_k$ Elemente von B'_K. Dann gibt es Mengen $S_1, \dots, S_k$ aus K, sodass $v_i \in S_i$. Weil K total geordnet ist, können wir je zwei Mengen vergleichen; genauer: Für $i, j \in \{, \dots, k\}$ gilt $S_i \subseteq S_j$ oder $S_j \subseteq S_i$. Wir finden deshalb einen Index i_0, sodass für $1 \leq i \leq k$ gilt: $S_i \subseteq S_{i_0}$. Insbesondere sind $v_1, \dots, v_k$ Elemente von S_{i_0}. Da S_{i_0} per Konstruktion linear unabhängig ist, ist auch $\{v_1, \dots, v_k\}$ linear unabhängig. □

Wir werden sehen, dass wir mithilfe von Proposition XI.1.2 zeigen können, dass $\mathfrak{S}$ maximale Elemente hat. Dies wird aus dem Lemma von Zorn folgen.

2. Das Zornsche Lemma

Definition XI.2.1: Seien X eine Menge und $\leqslant$ eine Relation auf X.

(i) Ist „$\leqslant$“ reflexiv, antisymmetrisch und transitiv, dann heißt die Relation eine *Ordungsrelation*.

(ii) Sind zusätzlich je zwei Elemente vergleichbar, d. h. gilt für $x, y \in X$ dass $x \leqslant y$ oder $y \leqslant x$, dann heißt „$\leqslant$“ *Totalordnung*.

(iii) Ist K eine nichtleere Teilmenge von X, sodass die Einschränkung von „$\leqslant$“ eine Totalordnung auf K ist, dann heißt K eine *Kette*.

Definition XI.2.2 (Maximale, minimale, größte und kleinste Elemente): Seien $(X, \leqslant)$ eine georndete Menge und x_0 ein Element von X.

(i) Gilt für alle $x \in X$ mit $x_0 \leqslant x$, dass $x = x_0$, dann heißt x *maximal*.

(ii) Gilt für alle $x \in X$ mit $x_0 \geqslant x$, dass $x_0 = x$, dann heißt x *minimal*.

(iii) Gilt für alle $x \in X$, dass $x \leqslant x_0$, dann heißt x_0 ein *größtes Element*.

(iv) Gilt für alle $x \in X$, dass $x \geqslant x_0$, dann heißt x_0 ein *kleinstes Element*.

Bemerkung XI.2.3: Eine geordnete Menge X kann viele maximale oder minimale Elemente haben, aber nur ein größtes beziehungsweise kleinstes Element.

Ist x_0 ein größtest Element, dann ist x_0 das einzige maximale Element; genau so für das kleinste Element.

Definition XI.2.4 (Obere Schranke, induktiv geordnet): Sei $(X, \leqslant)$ eine geordnete Menge.

(i) Seien S eine Teilmenge von X und x_0 ein Element von X. Gilt für alle $s \in S$, dass $s \leqslant x_0$, dann heißt x_0 eine *obere Schranke*. Das kleinste Element in $\{x \in X \mid x \text{ ist obere Schranke von } S\}$ heißt *kleinste obere Schranke*.

(ii) Hat jede Kette in X eine obere Schranke, dann heißt X *induktiv geordnet*. Hat jede Kette in X eine kleinste obere Schranke, dann heißt X eine *strikt induktiv geordnet*.

Beispiel XI.2.5: (i) Die Menge der natürlichen Zahlen mit dem gewöhnlichen „kleiner-gleich“ ist nicht induktiv geordnet, denn $\mathbb{N}$ selbst ist eine Kette, die keine obere Schranke hat.

(ii) Seien M eine Menge und $X = \mathfrak{P}(M)$ zusammen mit der Inklusion. Für jede Teilmenge $\mathfrak{S}$ von $\mathfrak{P}(M)$ gibt es eine kleinste obere Schranke S_0, nämlich $S_0 := \bigcup_{S \in S_0} S$. Insbesondere ist $(\mathfrak{P}(M), \subseteq)$ strikt induktiv geordnet.

Definition XI.2.6: Sei $(X, \leqslant)$ eine geordnete Menge. Hat jede nichtleere Teilmenge von X ein kleinstes Element, dann heißt „$\leqslant$“ eine *Wohlordnung*.

Beispiel XI.2.7: (i) Die totalgeordente Menge $(\mathbb{N}, \leq)$ ist wohlgeordnet („Prinzip des kleinsten Täters“).

(ii) Die partiell georndete Menge $(\mathfrak{P}(M), \subseteq)$ ist nicht wohlgeordnet, falls M mindestens zwei Elemente hat. Sind nämlich m_1 und m_2 zwei verschiedene Elemente von M, dann gehört $\{\{m_1\}, \{m_2\}\}$ zu $\mathfrak{P}(M)$, und dieses hat kein kleinstes Element.

Wir betrachten folgende Aussagen:

(i) **Auswahlaxiom:** Seien M eine nichtleere Menge, I eine nichtleere Indexmenge und $(M_i)_{i \in I}$ eine Familie nichtleerer Teilmengen von M. Dann gibt es eine Abbildung $f\colon I \to M$, sodass $f(i) \in M_i$. Dieses f nennt man *Auswahlfunktion*.

(ii) **Zorn’sches Lemma:** Sei $(M, \leqslant)$ eine nichtleere geordnete Menge. Ist $(M, \leqslant)$ induktiv geordnet, dann gibt es ein maximales Element in M.

(iii) **Wohlordnungssatz:** Jede Menge M hat eine Totalordnung bezüglich der sie wohlgeordnet ist.

Satz 46: *Das Auswahlaxiom, das Zorn’sche Lemma und der Wohlordnungssatz sind logisch äquivalent.*

Das Auswahlaxiom ist das Axiom in den ZFC-Axiomen, das nicht zu den ZF-Axiomen gehört. Man kann zeigen, dass das Zorn’sche Lemma und damit die Existenz von Basen in beliebigen Vektorräumen nicht aus den gewöhnlichen ZF-Axiomen folgen.

3. Existenz von Basen

Satz 47 (Existenz von Basen): *Seien K ein Körper und V ein Vektorraum über K. Dann hat V eine Basis.*

Der Satz über die Existenz von Basen folgt direkt aus dem allgemeinen Basisergänzungssatz, den wir als nächstes formulieren und beweisen möchten.

Satz 48 (Basisergänzungssatz): *Seien K ein Körper und V ein K-Vektorraum. Ist M eine linear unabhängige Teilmenge von V, dann gibt es eine Basis B von V mit $M \subseteq B$.*

Beweis: Wir setzen $X := \{S \subseteq V \mid M \subseteq S \text{ und } S \text{ ist linear unabhängig}\}$. Wegen $M \in X$ ist X nicht leer. Ferner ist X eine per Inklusion geordnete Menge. Wie in Proposition XI.1.2 zeigt man, dass $(X, \subseteq)$ induktiv geordnet ist. Nun liefert das Zorn'sche Lemma die Existenz eines maximalen Elements B in X, welches wegen Proposition XI.1.1 eine Basis von V ist. □

4. Beweis des Zorn'schen Lemmas

Proposition XI.4.1: *Seien $(X, \subseteq)$ eine nichtleere strikt induktiv georndete Menge mit einem kleinsten Element $x_{\min}$. Ist $F\colon X \to X$ eine Abbildung, sodass für alle $x \in X$ gilt, dass $F(x) \geq x$, dann hat F einen Fixpunkt.*

Mithilfe der Proposition XI.4.1 sind wir jetzt in der Lage, das Lemma von Kuratowski-Zorn zu beweisen.

Beweis: Seien $(X, \leqslant)$ eine nichtleere induktiv geordnete Menge X. Zunächst setzen wir voraus, dass $(X, \leqslant)$ sogar strikt induktiv geordnet ist. Ohne Einschränkung hat X ein kleinstes Element (wähle sonst irgendein x_0 aus X und ersetze X durch $X_{\geq x_0} := \{x \in X \mid x \geq x_0\}$).

Angenommen, X hätte kein maximales Element. Für jedes $x \in X$ könnten wir definieren $M_x := \{x' \in X \mid x' > x\}$. Per Annahme wäre für jedes $x \in X$ die Menge M_x nicht leer. Mithilfe des Auswahlaxioms erhielten wir eine Abbildung $F\colon X \to X$ mit $F(x) \in M_x$, d. h. für jedes $x \in X$ wäre $F(x) > x$. Nach Proposition XI.4.1 gäbe es dann einen Fixpunkt von F, d. h. es gäbe $x' \in X$ mit $F(x') = x'$. Das kann aber nach Konstruktion von F nicht sein, X muss also maximale Elemente enthalten.

Im zweiten Schritt setzen wir für $(X, \leqslant)$ nur noch induktive Ordnung voraus. Wir definieren $H := \{K \subseteq X \mid K \text{ ist eine Kette}\}$ und erhalten eine nichtleere strikt induktiv geordnete Menge $(H, \subseteq)$. Wegen des ersten Schritts wissen wir,

dass H ein maximale Element K_0 besitzt. Weil X induktiv geordnet ist, hat K_0 eine obere Schranke k_0 in X. Aber dieses k_0 ist sogar ein maximales Element, denn k_0 muss zu K_0 gehören ($K_0 \cup \{k_0\}$ wäre sonst eine größere Kette) und für $x \in X$ mit $x \geq k_0$ wäre x auch eine obere Schranke von K_0, d. h. es wäre wieder $x \in K_0$ und damit $x \leq k_0$, da k_0 eine obere Schranke von K_0 ist. □

Beweis (von Proposition XI.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 $x \leqslant F(x) \leqslant y$. □