Lineare Algebra I

gehalten von Prof. Weitze-Schmithüsen im Winter '20

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:

s9fhguen@stud.uni-saarland.de

Fehler können auch via Microsoft Teams per private Nachricht an Friedrich Günther gemeldet werden.

Inhaltsverzeichnis

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 Beispiel 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 $\mathrm{Abb}(\{1, \ldots, k\}, X)$. Dann sind die Abbildungen

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

und

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

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 $f \in \operatorname{Abb}(\{1,\dots,k\},X)$, 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. $F \circ G = \mathrm{id}_{\operatorname{Abb}(\{1,\dots,k\},X)}$. 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 $\operatorname{Abb}(\{1,\dots,k\},X)$. □

Definition I.5.14 (Leeres Produkt): Für eine Menge X heißt $X^0 := \operatorname{Abb}(\varnothing, X)$ das *leere Produkt*. X^0 besteht nach Bemerkung 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

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

bezeichnen wir die Menge der Permutationen von X.

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

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]

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

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	✓	✓	✓	✓	✓	—	✓

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

[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).

- *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 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 Definition 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 Definition 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 Definition 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 Definition 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 Definition 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* $I_p A = A = AI_p$.

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 Beispiel 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 Bemerkung 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 Definition 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 Definition 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 Definition 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

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

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 $1 \le i \le m$ setzen wir $a_i := \phi(e_i)$ und erklären eine Matrix $A \in \mathbb{R}^{n\times m}$ durch $A := (a_1|\dots|a_m)$. 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 $\mathrm{Hom}(G, H) := \{\varphi\colon (G, *) \to (H, \circ) \text{ ist Gruppenhomomorphismus}\}$.

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 Definition 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 Beispiel 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 Bemerkung 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 Definition 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 Beispiel 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 Bemerkung 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 Definition 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 Korollar 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 Korollar 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.

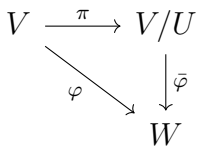
Beweis: Sobald wir uns davon überzeugt haben, dass die oben angegebenen Verknüpfungen wohldefiniert sind, sehen wir sofort dass V/U ein K-Vektorraum ist, da wir repräsentantenweise rechnen und wir wissen, dass V ein Vektorraum über K ist. Für die Wohldefiniertheit ist die Unabhängigkeit von der Wahl 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 Bemerkung IV.4.8 ist $\dim \operatorname{Kern} A = \dim \operatorname{Kern} \phi$, sodass $\operatorname{Rang} A = m - \dim \operatorname{Kern} \phi$. Wegen Bemerkung 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. $Z_k = A_{i,j}^{\alpha}$, oder $Z_k = V_{i,j}$ oder $Z_k = \operatorname{diag}(\alpha_1, \dots, \alpha_n)$, wobei α, $\alpha_1, \dots, \alpha_n \in K^\times$), 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

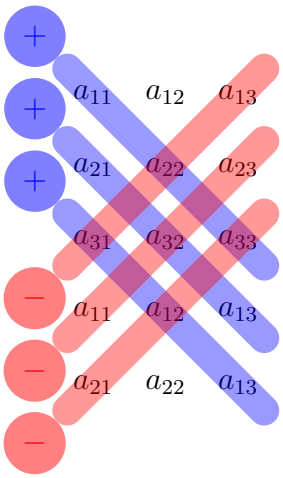
$$\det A = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = ad - bc.$$

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 Beispiel 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 Korollar 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 Bemerkung V.3.8 und Korollar 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 Korollar 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 Korollar 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 Korollar 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 Korollar 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 Korollar 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 Beispiel 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. □