Vorlesungsskript
Lineare Algebra

Wintersemester 2025–Sommersemester 2026

Roland Herzog*

2026-04-13

*Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany
(roland.herzog@iwr.uni-heidelberg.de, https://scoop.iwr.uni-heidelberg.de)

🅭🅯🄏

Teile dieses Skripts orientieren sich an früheren Vorlesungen von Jan Johannes (Universität Heidelberg).

Änderungen gegenüber bereits veröffentlichten Versionen werden in dieser Farbe gekennzeichnet.

Material für 27 Vorlesungen (Lineare Algebra I).

Kommentare und Korrekturen bitte an roland.herzog@iwr.uni-heidelberg.de.

Inhaltsverzeichnis

Kapitel 1 Mathematische Grundlagen

Die **Algebra** (von arabisch الجبر, *al-ǧabr*, „das Zusammenfügen gebrochener Teile“, englisch: algebra) hat ihren Ursprung in der Beschreibung von Lösungsverfahren linearer und quadratischer Gleichungen. Heute verstehen wir den Begriff Algebra deutlich weiter, es geht jedoch immer um Strukturen, Abbildungen zwischen Strukturen und die in ihnen geltenden „Rechenregeln“. Speziell die **lineare Algebra** (englisch: linear algebra) befasst sich mit „linearen Strukturen“, das sind vor allem Vektorräume, Abbildungen zwischen Vektorräumen und lineare Gleichungssysteme.

Wie andere Wissenschaften auch hat die Mathematik eine eigene Sprache, die man erlernen muss, um die Gegenstände dieser Wissenschaft zu verstehen und sich sachgerecht ausdrücken und argumentieren zu können. Das Herz der Mathematik bilden Beweise. Jede Aussage, jeder Lehrsatz muss bewiesen werden, d. h., durch logische Verknüpfungen aus den verwendeten Grundaxiomen und bereits bewiesenen Aussagen hergeleitet werden.

Eine streng formale, axiomatische Einführung der Logik und logischer Schlussweisen ist im Rahmen dieser Lehrveranstaltung leider nicht möglich. Diese kann später bei Interesse in weiterführenden Veranstaltungen zur Logik nachgeholt werden. Wir beschränken uns hier auf eine „naive“ (nicht-axiomatische) Einführung in die Logik.

§ 1 Aussagenlogik

Literatur: Deiser, 2024b, Kapitel 1.1; Deiser, 2022, Anhang 1; Magnus u. a., 2023, Kapitel 1–14; Velleman, 2019

Definition 1.1 (Aussage, Wahrheitswert)**.**
Eine **Aussage** (englisch: statement, proposition) ist ein Satz (einer Sprache), dem eindeutig entweder der **Wahrheitswert wahr** (kurz: W oder $\top$, englisch: true, T) oder der **Wahrheitswert falsch** (kurz: F oder $\bot$, englisch: false, F) zugeordnet werden kann. △

Der Satz kann dabei der gewöhnlichen Sprache oder der mathematischen Sprache entstammen. Wir bezeichnen Aussagen in der Regel mit Großbuchstaben wie P, Q usw.

Beispiel 1.2 (Aussagen und Nicht-Aussagen)**.**

(i) P: 9 ist durch 3 teilbar.
Dieses ist eine wahre Aussage.

(ii) Q: Am 17.10.2025 ist London die Hauptstadt von Frankreich.
Dieses ist eine falsche Aussage.

(*iii*) R: München ist 781 km von Hamburg entfernt.
Dieses ist keine Aussage, da der Satz zu viel Interpretationsspielraum lässt. Was ist mit „München" und „Hamburg" gemeint? Mit welcher Toleranz ist die Entfernungsangabe zu verstehen?

(*iv*) S: Das Team des VfL Wolfsburg ist in der Saison 2025/26 deutscher Meister in der Frauen-Fußball-Bundesliga.
Dieses ist eine Aussage, deren Wahrheitswert wir im Moment aber nicht kennen.

(*v*) T: Es gibt unendlich viele Primzahlzwillinge.
Dieses ist ebenfalls eine Aussage, deren Wahrheitswert wir zur Zeit nicht kennen.[1] △

Ein grundlegendes Prinzip in der Mathematik ist es, aus bekannten Objekten durch Verknüpfung neue Objekte zu schaffen. In der Logik heißen diese Verknüpfungen **Junktoren** (englisch: logical operators, junction, lateinisch: iungere: verbinden, verknüpfen). Ein Junktor erschafft also aus einer oder aus mehreren Aussagen eine neue Aussage. Der Wahrheitswert der neuen Aussage ergibt sich aus den Wahrheitswerten der miteinander verknüpften Aussagen. Wir geben einen Junktor über seine **Wahrheitswerttabelle** (auch: **Wahrheitstafel**, englisch: truth table) an.

Definition 1.3 (Junktoren)**.**

Im Folgenden seien P und Q Aussagen. Wir definieren die folgenden ein- und zweistellige Junktoren:

(*i*) $\neg$ **Negation**[2] (**Verneinung**)

Die Operation $\neg P$ (sprich: „nicht P") heißt **Negation**. $\neg P$ ist wahr, wenn P falsch ist, und falsch, wenn P wahr ist.

P	$\neg P$
W	F
F	W

(*ii*) $\wedge$ **Konjunktion**[3] (**Und-Verknüpfung**)

Die Aussage $P \wedge Q$ (sprich: „P und Q") ist dann wahr, wenn P und Q beide wahr sind, ansonsten falsch.

P	Q	$P \wedge Q$
W	W	W
W	F	F
F	W	F
F	F	F

[1] siehe Primzahlzwillingsvermutung
[2] englisch: negation, lateinisch: negare: verneinen
[3] englisch: conjunction, lateinisch: coniungere: verbinden

(*iii*) $\vee$ **Disjunktion**[4] (**Oder-Verknüpfung**)

Die Aussage $P \vee Q$ (sprich: „P oder Q") ist wahr, wenn mindestens eine der Aussagen P und Q wahr ist, ansonsten falsch. Das „Oder" ist also in einem nicht-ausschließenden Sinne gemeint.

P	Q	$P \vee Q$
W	W	W
W	F	W
F	W	W
F	F	F

(*iv*) $\to$ **Implikation**[5] (**Konditional**[6], **Subjunktion**[7], **Wenn-Dann-Verknüpfung**)

Die Aussage $P \to Q$ ist über die nebenstehende Wahrheitswerttabelle definiert. Wir benennen diese Aussage auch als „P ist **hinreichend** für Q" (englisch: „P is sufficient for Q"), „Q ist **notwendig** für P" (englisch: „Q is necessary for P"), „P impliziert Q" (englisch: „P implies Q") oder „Wenn P, dann Q" (englisch: „If P, then Q"). In einer Implikation $P \to Q$ nennt man P auch das **Antezedens** (englisch: antecedent, lateinisch: antecedens: das Vorausgehende) und Q das **Konsequens** (englisch: consequent, lateinisch: consequentis: folgerichtig).

P	Q	$P \to Q$
W	W	W
W	F	F
F	W	W
F	F	W

Die Implikation behauptet keinerlei kausalen oder sonstigen inhaltlichen Zusammenhang zwischen den Aussagen P und Q. Man spricht auch von **materialer Implikation** (englisch: material implication, material conditional). Die häufig anzutreffende Sprechweise „Wenn P, dann Q" ist daher problematisch, weil wir diese intuitiv als Kausalität oder zeitliche Nähe interpretieren.

(*v*) $\leftrightarrow$ **Äquivalenz**[8] (**Bikonditional**[9], **Bisubjunktion**, **Genau-Dann-Wenn-Verknüpfung**)

Die Aussage $P \leftrightarrow Q$ ist wahr, wenn entweder P und Q beide wahr oder beide falsch sind, ansonsten falsch. Wir benennen diese Aussage auch als „P ist **notwendig und hinreichend** für Q" (englisch: „P is necessary and sufficient for Q"), „P ist äquivalent zu Q" (englisch: „P is equivalent to Q"), „P genau dann, wenn Q" oder „P dann und nur dann, wenn Q" (englisch: „P if and only if Q", „P iff Q").

P	Q	$P \leftrightarrow Q$
W	W	W
W	F	F
F	W	F
F	F	W

Auch hier gilt, dass die Äquivalenz nichts über einen eventuellen kausalen oder sonstigen inhaltlichen Zusammenhang zwischen den Aussagen P und Q aussagt. Man spricht auch von **materialer Äquivalenz** (englisch: material equivalence, material biconditional). △

Quizfrage 1.1: Wieviele verschiedene einstellige Junktoren gibt es? Wieviele zweistellige?

Quizfrage 1.2: Können Sie alle zweistelligen Junktoren aus den oben genannten Junktoren $\neg$

[4] englisch: disjunction, lateinisch: disiungere: trennen, unterscheiden
[5] englisch: implication, lateinisch: implicare: verwickeln
[6] englisch: conditional, lateinisch: conditio: Bedingung
[7] lateinisch: subiungere: unterordnen
[8] englisch: equivalence, lateinisch: aequivalens: gleichwertig
[9] englisch: biconditional

sowie $\wedge$, $\vee$, $\rightarrow$ und $\leftrightarrow$ zusammensetzen? Reicht evtl. sogar eine Teilmenge davon aus?

Beispiel 1.4 (Symbolisierung von Sätzen der Umgangssprache[10]).
Die Symbolisierung von Sätzen der Umgangssprache in logische Aussagen ist nicht immer ganz einfach. Es folgen einige Beispiele jeweils mit einer oder mehreren gleichwertigen Symbolisierungen.

(i) Zum Burger servieren wir Pommes oder Salat.
Das „oder“ ist hier im ausschließenden Sinne gemeint.

P: Zum Burger servieren wir Pommes.
S: Zum Burger servieren wir Salat.

- $(P \vee S) \wedge (\neg(P \wedge S))$
- $(P \wedge (\neg S)) \vee (S \wedge (\neg P))$

(ii) Obwohl Barbara energisch ist, ist sie nicht sportlich.

E: Barbara ist energisch.
S: Barbara ist sportlich.

- $E \wedge (\neg S)$

(iii) Du wirst keine Suppe bekommen, aber dafür den Salat.

S_1: Du wirst Suppe bekommen.
S_2: Du wirst Salat bekommen.

- $(\neg S_1) \wedge S_2$

(iv) Du wirst Dich erkälten, es sei denn, Du trägst eine Jacke.

J: Du trägst eine Jacke.
E: Du wirst Dich erkälten.

- $(\neg J) \rightarrow E$
- $J \vee E$

An den Beispielen sieht man, dass unter der formalen Symbolisierung Nuancen der Sprache zugunsten der Präzision verloren gehen. △

Lemma 1.5 (Umschreibung von $\rightarrow$ und $\leftrightarrow$).
Es seien P und Q Aussagen.

(i) Die Aussagen

- $P \rightarrow Q$
- $(\neg P) \vee Q$
- $(\neg Q) \rightarrow (\neg P)$

haben dieselben Wahrheitstafeln.

(ii) Die Aussagen

- $P \leftrightarrow Q$
- $(P \rightarrow Q) \wedge (Q \rightarrow P)$

[10] angelehnt an Beispiele aus Magnus u. a., 2023, Kapitel 5, genutzt unter der Lizenz CC-BY 4.0

haben dieselben Wahrheitstafeln.

Beweis. Wir stellen die Wahrheitstafeln für die drei Aussagen in Aussage (i) auf:

P	Q	$P \to Q$	$(\neg P) \vee Q$	$\neg Q$	$\neg P$	$(\neg Q) \to (\neg P)$
W	W	W	W	F	F	W
W	F	F	F	W	F	F
F	W	W	W	F	W	W
F	F	W	W	W	W	W

Der Beweis der Aussage (ii) ist Gegenstand der Übung. □

Da die Verknüpfung von Aussagen stets wieder auf Aussagen führt, können wir durch wiederholte Verknüpfung komplexe Aussagen aufbauen, wie etwa $(P \to D) \to ((Q \vee C) \to (D \wedge C))$. Zur Vereinfachung der Notation vereinbaren wir folgende Bindungsregeln:

$$\begin{aligned} \neg &\quad \text{bindet stärker als} \quad \wedge \\ \wedge &\quad \text{bindet stärker als} \quad \vee \\ \vee &\quad \text{bindet stärker als} \quad \to \\ \to &\quad \text{bindet stärker als} \quad \leftrightarrow . \end{aligned} \tag{1.1}$$

Diese Regeln erlauben uns, auf Klammern zu verzichten. Beispielsweise ist

$$\begin{aligned} (\neg P) \wedge Q &\quad \text{gleichwertig mit} \quad \neg P \wedge Q \\ \text{und} \quad (\neg(P \wedge Q)) \to (Q \vee (\neg Q)) &\quad \text{gleichwertig mit} \quad \neg(P \wedge Q) \to Q \vee \neg Q. \end{aligned}$$

Es gilt jedoch, dass Klammern zur Verdeutlichung nicht schaden können. Statt $(\cdot)$ können auch $[\cdot]$ oder $\{\cdot\}$ verwendet werden.

Wir bestimmen jetzt die Wahrheitstafeln einiger zusammengesetzter Aussagen.

Beispiel 1.6 (Wahrheitstafeln zusammengesetzter Aussagen)**.**

(i) $\neg(\neg P \wedge \neg Q)$

P	Q	$\neg P$	$\neg Q$	$\neg P \wedge \neg Q$	$\neg(\neg P \wedge \neg Q)$
W	W	F	F	F	W
W	F	F	W	F	W
F	W	W	F	F	W
F	F	W	W	W	F

Beachte: Diese Wahrheitstafel ist offenbar dieselbe wie die von $P \vee Q$.

(*ii*) $P \vee Q \rightarrow Q \wedge R$

P	Q	R	$P \vee Q$	$Q \wedge R$	$P \vee Q \rightarrow Q \wedge R$
W	W	W	W	W	W
W	W	F	W	F	F
W	F	W	W	F	F
W	F	F	W	F	F
F	W	W	W	W	W
F	W	F	W	F	F
F	F	W	F	F	W
F	F	F	F	F	W

(*iii*) $\neg(P \wedge Q) \leftrightarrow \neg P \vee \neg Q$

P	Q	$\neg(P \wedge Q)$	$\neg P \vee \neg Q$	$\neg(P \wedge Q) \leftrightarrow \neg P \vee \neg Q$
W	W	F	F	W
W	F	W	W	W
F	W	W	W	W
F	F	W	W	W

△

Die Aussage aus Punkt (*iii*) $\neg(P \wedge Q) \leftrightarrow \neg P \vee \neg Q$ hat also immer den Wahrheitswert W, unabhängig von den Wahrheitswerten der Aussagen P und Q. Eine solche Aussage nennt man eine **Tautologie**[11] (englisch: tautology) oder ein **logisches Gesetz**. Die Aussage aus Punkt (*i*) bedeutet, dass auch $\neg(\neg P \wedge \neg Q) \leftrightarrow P \vee Q$ eine Tautologie ist. Tautologien spielen eine entscheidende Rollen in mathematischen Beweisen, siehe § 3.

Definition 1.7 (logische Implikation, logische Äquivalenz)**.**
Es seien P und Q Aussagen.

(*i*) Die Aussage Q heißt eine **logische Implikation** (englisch: logical implication) der Aussage P, wenn $P \rightarrow Q$ eine Tautologie ist. P heißt dann **Prämisse** (englisch: premise), und Q heißt **Konklusion** (englisch: conclusion). Wir schreiben dann $P \Rightarrow Q$ und sagen „P impliziert Q" oder „Q folgt aus P".

(*ii*) Die Aussagen P und Q heißen **logisch äquivalent (zueinander)** (englisch: logically equivalent), wenn $P \leftrightarrow Q$ eine Tautologie ist. Wir schreiben $P \Leftrightarrow Q$ und sagen „P ist äquivalent zu Q" oder „P und Q sind (zueinander) äquivalent". △

Wir vereinbaren, dass $\Rightarrow$ und $\Leftrightarrow$ noch schwächer binden als die Junktoren in (1.1), also haben wir

$$\begin{array}{ccc} \neg & \text{bindet stärker als} & \wedge \\ \wedge & \text{bindet stärker als} & \vee \end{array}$$

[11] altgriechisch: *ταυτο*: dasselbe

$$
\begin{array}{rcl}
\vee & \text{bindet stärker als} & \rightarrow \\
\rightarrow & \text{bindet stärker als} & \leftrightarrow \\
\leftrightarrow & \text{bindet stärker als} & \Rightarrow \\
\Rightarrow & \text{bindet stärker als} & \Leftrightarrow .
\end{array} \tag{1.2}
$$

Beispiel 1.8 (logische Implikationen und Äquivalenzen).

(*i*) Die Aussage $(P \rightarrow Q) \wedge P$ impliziert die Aussage Q, kurz: $(P \rightarrow Q) \wedge P \Rightarrow Q$, denn $(P \rightarrow Q) \wedge P \rightarrow Q$ ist eine Tautologie:

P	Q	$P \rightarrow Q$	$(P \rightarrow Q) \wedge P$	$(P \rightarrow Q) \wedge P \rightarrow Q$
W	W	W	W	W
W	F	F	F	W
F	W	W	F	W
F	F	W	F	W

(*ii*) Die Aussage $(P \rightarrow Q) \wedge \neg Q$ impliziert die Aussage $\neg P$, kurz: $(P \rightarrow Q) \wedge \neg Q \Rightarrow \neg P$, denn $(P \rightarrow Q) \wedge \neg Q \rightarrow \neg P$ ist eine Tautologie:

P	Q	$P \rightarrow Q$	$(P \rightarrow Q) \wedge \neg Q$	$(P \rightarrow Q) \wedge \neg Q \rightarrow \neg P$
W	W	W	F	W
W	F	F	F	W
F	W	W	F	W
F	F	W	W	W

(*iii*) Die Aussagen $\neg(P \wedge Q)$ und $\neg P \vee \neg Q$ sind logisch äquivalent, kurz: $\neg(P \wedge Q) \Leftrightarrow \neg P \vee \neg Q$, denn $\neg(P \wedge Q) \leftrightarrow \neg P \vee \neg Q$ ist eine Tautologie, wie in Beispiel 1.6 gerade schon gezeigt wurde. △

Satz 1.9 (logische Implikationen und Äquivalenzen).
Es seien P, Q und R Aussagen. Es gelten die folgenden Implikationen und Äquivalenzen.

$$\neg(\neg P) \Leftrightarrow P \qquad \text{doppelte Verneinung}^{12} \tag{1.3}$$

$$P \Rightarrow \top \qquad \text{„Aus Beliebigem folgt Wahres.“}^{13} \tag{1.4a}$$

$$\bot \Rightarrow P \qquad \text{„Aus Falschem folgt Beliebiges.“}^{14} \tag{1.4b}$$

$$P \wedge P \Leftrightarrow P \qquad \textbf{Idempotenz von } \wedge^{15} \tag{1.5a}$$

$$P \vee P \Leftrightarrow P \qquad \textbf{Idempotenz von } \vee \tag{1.5b}$$

$$P \wedge \top \Leftrightarrow P \qquad \textbf{Neutralität von } \wedge \top^{16} \tag{1.6a}$$

$$P \vee \bot \Leftrightarrow P \qquad \textbf{Neutralität von } \vee \bot \tag{1.6b}$$

$$P \wedge \bot \Leftrightarrow \bot \qquad \textbf{Absorption bei } \wedge^{17} \tag{1.7a}$$

$$P \vee \top \Leftrightarrow \top \qquad \textbf{Absorption bei } \vee \tag{1.7b}$$

$$P \land \neg P \iff \bot \qquad \textbf{Komplementarität bei } \land^{18} \tag{1.8a}$$
$$P \lor \neg P \iff \top \qquad \textbf{Komplementarität bei } \lor^{19} \tag{1.8b}$$

$$P \land Q \iff Q \land P \qquad \textbf{Kommutativität von } \land^{20} \tag{1.9a}$$
$$P \lor Q \iff Q \lor P \qquad \textbf{Kommutativität von } \lor \tag{1.9b}$$

$$(P \lor Q) \lor R \iff P \lor (Q \lor R) \qquad \textbf{Assoziativität von } \lor^{21} \tag{1.10a}$$
$$(P \land Q) \land R \iff P \land (Q \land R) \qquad \textbf{Assoziativität von } \land \tag{1.10b}$$

$$\neg(P \land Q) \iff \neg P \lor \neg Q \qquad \textbf{De Morgansches Gesetz}^{22} \tag{1.11a}$$
$$\neg(P \lor Q) \iff \neg P \land \neg Q \qquad \textbf{De Morgansches Gesetz} \tag{1.11b}$$

$$P \land (Q \lor R) \iff (P \land Q) \lor (P \land R) \qquad \textbf{Distributivität}^{23} \tag{1.12a}$$
$$P \lor (Q \land R) \iff (P \lor Q) \land (P \lor R) \qquad \textbf{Distributivität} \tag{1.12b}$$

$$(P \to Q) \land P \implies Q \qquad \textbf{modus ponendo ponens} \tag{1.13a}$$
$$(P \to Q) \land \neg Q \implies \neg P \qquad \textbf{modus tollendo tollens} \tag{1.13b}$$
$$(P \to \neg Q) \land P \implies \neg Q \qquad \textbf{modus ponendo tollens}^{24} \tag{1.13c}$$
$$(\neg P \to Q) \land \neg P \implies Q \qquad \textbf{modus tollendo ponens}^{25} \tag{1.13d}$$

$$(P \to Q) \land (Q \to R) \implies P \to R \qquad \textbf{Kettenschluss}^{26}. \tag{1.14}$$

Beweis. Der Beweis erfolgt durch Aufstellen der Wahrheitstafeln und wird hier nicht ausgeführt. □

Quizfrage 1.3: Können Sie einfache Beispiele in Alltagssprache angeben, bei denen die vier logischen Implikationen (1.13a)–(1.13d) benutzt werden?

Ende der Vorlesung 1

---

[12] lateinisch: duplex negatio affirmat
[13] lateinisch: verum ex quolibet
[14] lateinisch: ex falso quodlibet
[15] englisch: idempotence, lateinisch: idem: dasselbe, lateinisch: potentia: Vermögen, Kraft
[16] englisch: neutrality, lateinisch: neutrum: keines von beiden
[17] englisch: absorption, lateinisch: absorbere: einsaugen, verschlingen
[18] englisch: complementarity, lateinisch: complementum: Ergänzung, Vollendung
[19] Gesetz vom ausgeschlossenen Dritten, lateinisch: tertium non datur
[20] englisch: commutativity, lateinisch: commutare: tauschen, vertauschen
[21] englisch: associativity, lateinisch: associare: verbinden, beigesellen
[22] englisch: De Morgan's law
[23] englisch: distributivity, lateinisch: distribuere: verteilen, aufteilen
[24] Der modus ponendo tollens wird häufig als $\neg(P \land Q) \land P \implies \neg Q$ geschrieben.
[25] Der modus tollendo ponens wird häufig als $(P \lor Q) \land \neg P \implies Q$ geschrieben.
[26] englisch: chain inference

§ 2 Prädikatenlogik

Literatur: Deiser, 2022, Anhang 1; Magnus u. a., 2023, Kapitel 22–39

Die Aussagenlogik reicht für die Bedürfnisse der Mathematik nicht aus. Beispielsweise lässt sich die Aussage „Wenn n eine gerade ganze Zahl ist, dann ist auch n^2 eine gerade ganze Zahl." innerhalb der Aussagenlogik aus § 1 nicht wie erforderlich symbolisieren. Die Schwierigkeit ist, dass wir in der Aussagenlogik keine Aussagen mit Variablen zur Verfügung haben. Wir benötigen dazu die **Prädikatenlogik**[27], eine Erweiterung der Aussagenlogik. In der Prädikatenlogik ist es möglich, eine Aussage von dem Gegenstand, über den sie gemacht wird, zu trennen. Neben den schon bekannten Junktoren verwendet die Prädikatenlogik zusätzlich

- **Aussageformen** (englisch: statement) oder **Prädikate** (englisch: predicate), das sind sprachliche Gebilde mit Variablen (Leerstellen), die nach Einsetzen der Variablen in Aussagen übergehen, z. B.

$$\begin{aligned} A(x) &: x \text{ wohnt in Aachen.} \\ Z(x) &: x \text{ ist eine gerade ganze Zahl.} \\ G(x, y) &: x \text{ ist mindestens so groß wie } y. \end{aligned}$$

  Die Anzahl der Variablen einer Aussageform heißt deren **Stelligkeit** (englisch: arity).
- **Quantoren** (englisch: quantifier), und zwar

$\forall$	„für alle"	**Allquantor**[28]
$\exists$	„es existiert (mindestens) ein"	**Existenzquantor**[29]
$\exists!$	„es existiert genau ein"	**Eindeutigkeitsquantor**[30].

Für die Symbole der Prädikatenlogik sollen folgende Bindungsregeln gelten:

$$\begin{aligned} \forall, \exists \text{ und } \exists! &\quad \text{binden stärker als} \quad \neg \\ \neg &\quad \text{bindet stärker als} \quad \wedge \\ \wedge &\quad \text{bindet stärker als} \quad \vee \\ \vee &\quad \text{bindet stärker als} \quad \rightarrow \\ \rightarrow &\quad \text{bindet stärker als} \quad \leftrightarrow \\ \leftrightarrow &\quad \text{bindet stärker als} \quad \Rightarrow \\ \Rightarrow &\quad \text{bindet stärker als} \quad \Leftrightarrow . \end{aligned} \tag{2.1}$$

Zu jedem Quantor geben wir den **Grundbereich** (auch: **Individuenbereich**, **Diskursuniversum**, **Domäne**, englisch: universe of discourse, domain of discourse) an. In der Regel nimmt

[27] genauer: Prädikatenlogik erster Stufe, englisch: predicate logic, first order logic
[28] englisch: universal quantifier
[29] englisch: existential quantifier
[30] englisch: uniqueness quantifier

man an, dass der Grundbereich nicht leer ist, um gewisse Komplikationen auszuschließen. Der Grundbereich ist wichtig und beeinflusst den Wahrheitswert einer quantorisierten Aussage:

$\forall x \in \mathbb{N}\ (x \geqslant 0)$	Alle natürlichen Zahlen sind nichtnegativ.	(wahre Aussage)
$\forall x \in \mathbb{R}\ (x \geqslant 0)$	Alle reellen Zahlen sind nichtnegativ.	(falsche Aussage)

Beispiel 2.1 (Symbolisierung von Sätzen der Umgangssprache mit Quantoren).
Wir betrachten die Aussageformen

$$E(x) : x \text{ hat } 100\,000 \text{ oder mehr Einwohner.}$$
$$S(x) : x \text{ ist eine Stadt.}$$

mit dem Grundbereich $O \coloneqq$ Menge aller Orte in Deutschland. Dann können wir die folgenden Aussagen wie angegeben symbolisieren:

(i) Es gibt mindestens eine Stadt in Deutschland, die 100 000 oder mehr Einwohner hat.
$\exists x \in O\ (E(x) \land S(x))$

(ii) Es gibt genau einen Ort in Deutschland, der 100 000 oder mehr Einwohner hat, der aber keine Stadt ist.
$\exists! x \in O\ (E(x) \land \neg S(x))$

(iii) Alle Städte in Deutschland haben 100 000 oder mehr Einwohner.
$\forall x \in O\ (S(x) \to E(x))$

(iv) Keine Stadt in Deutschland hat 100 000 oder mehr Einwohner.
$\neg \exists x \in O\ (E(x) \land S(x))$

Diese letzte Aussage (iv) könnten wir beispielsweise äquivalent auch auf folgende Weisen ausdrücken:

(v) Jeder Ort in Deutschland ist nicht gleichzeitig Stadt und hat mehr als 100 000 Einwohner.
$\forall x \in O\ (\neg(E(x) \land S(x)))$

(vi) Jeder Ort in Deutschland ist entweder keine Stadt oder hat weniger als 100 000 Einwohner oder beides.
$\forall x \in O\ (\neg E(x) \lor \neg S(x))$

(vii) Für alle Orte in Deutschland gilt: Wenn sie 100 000 oder mehr Einwohner haben, sind sie keine Stadt.
$\forall x \in O\ (E(x) \to \neg S(x))$

$(viii)$ Für alle Orte in Deutschland gilt: Wenn sie eine Stadt sind, haben sie weniger als 100 000 Einwohner.
$\forall x \in O\ (S(x) \to \neg E(x))$

Die Äquivalenz der Aussagen (iv) bis $(viii)$ können wir mit Hilfe von (2.2b), des De Morganschen Gesetzes (1.11a) und der Definition von $\to$ zeigen. △

Man sagt, dass die Variable einer Aussageform durch ihren Quantor **gebunden** (englisch: bound variable) wird. Auf den Namen der Variablen kommt es dabei übrigens nicht an, es sind also $\exists x\ (E(x) \land S(x))$ und $\exists y\ (E(y) \land S(y))$ äquivalente Aussagen.

Besonders mehrstellige Aussageformen spielen in vielen mathematischen Aussagen eine große Rolle. Die Reihenfolge verschiedener Quantoren ist dabei wichtig! Unterscheide zum Beispiel (siehe Lehrveranstaltung zur *Analysis* zu den Begriffen „Stetigkeit" und „gleichmäßige Stetigkeit")

Die Funktion $f\colon (a,b) \to \mathbb{R}$ ist stetig:

$$\forall x \in (a,b)\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in (a,b)\ \underbrace{\big(|x-y| < \delta \ \Rightarrow\ |f(x) - f(y)| < \varepsilon\big)}_{\text{vierstellige Aussageform}}.$$

Die Funktion $f\colon (a,b) \to \mathbb{R}$ ist gleichmäßig stetig:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in (a,b)\ \forall y \in (a,b)\ \underbrace{\big(|x-y| < \delta \ \Rightarrow\ |f(x) - f(y)| < \varepsilon\big)}_{\text{dieselbe vierstellige Aussageform}}.$$

Für Aussagen mit Quantoren gelten folgende Regeln (ohne Beweis).

Satz 2.2 (logische Implikationen und Äquivalenzen von Aussagen mit Quantoren)**.**
Es seien P, Q einstellige Aussageformen mit gemeinsamem Grundbereich und R eine zweistellige Aussageform. Es gelten die folgenden Implikationen und Äquivalenzen.[31]

$$\neg(\forall x\ P(x)) \quad\Leftrightarrow\quad \exists x\ (\neg P(x)) \qquad \text{Negation des Allquantors} \tag{2.2a}$$
$$\neg(\exists x\ P(x)) \quad\Leftrightarrow\quad \forall x\ (\neg P(x)) \qquad \text{Neg. des Existenzquantors} \tag{2.2b}$$

$$\forall x\ (P(x) \wedge Q(x)) \quad\Leftrightarrow\quad \forall x\ P(x) \wedge \forall x\ Q(x) \qquad \textbf{Distributivität} \tag{2.3a}$$
$$\exists x\ (P(x) \vee Q(x)) \quad\Leftrightarrow\quad \exists x\ P(x) \vee \exists x\ Q(x) \qquad \textbf{Distributivität} \tag{2.3b}$$

$$(\forall x\ P(x)) \vee (\forall x\ Q(x)) \quad\Rightarrow\quad \forall x\ (P(x) \vee Q(x)) \tag{2.4a}$$
$$\exists x\ (P(x) \wedge Q(x)) \quad\Rightarrow\quad (\exists x\ P(x)) \wedge (\exists x\ Q(x)) \tag{2.4b}$$

$$\forall x\ (P(x) \to Q(x)) \quad\Rightarrow\quad (\forall x\ P(x)) \to (\forall x\ Q(x)) \tag{2.5a}$$
$$\exists x\ (P(x) \to Q(x)) \quad\Leftrightarrow\quad (\forall x\ P(x)) \to (\exists x\ Q(x)) \tag{2.5b}$$

Auf https://de.wikipedia.org/wiki/Prädikatenlogik#Quantoren finden sich schöne Veranschaulichungen wahrer Aussagen mit zweistelligen Aussageformen und verschiedenen Quantoren.

§ 3 Beweismuster

Literatur: Deiser, 2024b, Kapitel 1.1; Magnus u. a., 2023, Kapitel 15–21

In einem Beweis versuchen wir in der Regel, für gegebene Aussagen P, Q die Implikation $P \Rightarrow Q$ nachzuweisen. Das heißt, wir müssen nachweisen, dass $P \to Q$ eine Tautologie ist. Meistens besteht die Prämisse P selbst aus einer Konjunktion (Und-Verknüpfung) und/oder

[31] Aus Gründen der Lesbarkeit lassen wir die Angabe des Grundbereichs bei den Quantoren hier weg.

Disjunktion (Oder-Verknüpfung) mehrerer einzelner Prämissen. Nicht alle Prämissen werden in der Formulierung eines mathematischen Satzes explizit genannt. Beispielsweise wird man die als wahr geltenden Grundannahmen (Axiome) z. B. über die reellen Zahlen nicht jedes Mal explizit erwähnen.

Ein Beweis wird oft in viele kleine Schritte zerlegt. Das Aufstellen einer Wahrheitswerttabelle ist nicht zielführend, vor allem dann nicht, wenn die Aussagen Variablen enthalten. Vielmehr werden wir Schlussregeln anwenden, die auf Tautologien beruhen. Solche Tautologien haben wir in Satz 1.9 und Satz 2.2 bereits aufgeführt.

Folgende Beweismuster für Implikationen $P \Rightarrow Q$ werden häufig verwendet:

(1) Beim **direkten Beweis** (englisch: direct proof) wird $P \Rightarrow Q$, typischerweise unter Verwendung von Axiomen und bereits bewiesenen Sätzen, direkt mit Hilfe von Schlussregeln hergeleitet. Ein Vergleich mit der Wahrheitswerttabelle der Implikation $P \rightarrow Q$ (Definition 1.3) zeigt, dass nur das Fall „P ist wahr, und Q ist falsch" ausgeschlossen werden muss. Dazu nehmen wir typischerweige die Aussage P als wahr an und zeigen, dass dann auch die Aussage Q wahr ist.

(2) Beim **indirekten Beweis** oder **Beweis durch Kontraposition** (englisch: indirect proof, proof by contrapositive, lateinisch: contra: gegen, lateinisch: positio: Position, Stellung) nutzen wir die Äquivalenz $(P \rightarrow Q) \Leftrightarrow (\neg Q \rightarrow \neg P)$ aus. Wir führen also einen direkten Beweis für $\neg Q \Rightarrow \neg P$. Dazu nehmen wir die Aussage Q als falsch an und zeigen, dass dann die Aussage P auch falsch ist.

(3) Beim **Widerspruchsbeweis** (englisch: proof by contradiction, lateinisch: reductio ad adsurdum: Zurückführung auf das Sinnlose) nutzen wir die Äquivalenz $(P \rightarrow Q) \Leftrightarrow (P \wedge \neg Q) \rightarrow \perp$ aus. Dazu nehmen wir die Aussage P als wahr und die Aussage Q als falsch an und zeigen, dass dann $\perp$ folgt.

(4) Beim **Beweis durch Fallunterscheidung** (englisch: proof by distinction of cases) nutzen wir die Äquivalenz $(P \wedge R \rightarrow Q) \wedge (P \wedge \neg R \rightarrow Q) \Leftrightarrow P \rightarrow Q$. Dabei ist R irgendeine weitere Aussage. Wir nehmen also zunächst die Aussagen P und R als wahr an und zeigen, dass dann auch die Aussage Q wahr ist. Anschließend nehmen wir die Aussage P weiterhin als wahr aber die Aussage R als falsch an und zeigen, dass dann wiederum die Aussage Q wahr ist. Manchmal wird die Aussage R in weitere Teilaussagen zerlegt, sodass mehr als die zwei Fälle „R ist wahr" und „R ist falsch" abgebildet werden können.

Beispiel 3.1 (verschiedene Beweismuster)**.**

(1) **direkter Beweis**

Behauptung: Für natürliche Zahlen m, n gelte $m^2 < n^2$, dann gilt auch $m < n$.

Wir symbolisieren die zugehörigen Aussagen über zweistellige Aussageformen:

$$P(m, n) : m^2 < n^2$$
$$Q(m, n) : m < n$$

und verwenden als Grundbereich für beide Variablen in beiden Aussageformen die Menge $\mathbb{N} \coloneqq \{1, 2, 3, \ldots\}$ der natürlichen Zahlen. Wir wollen zeigen:

$$\forall m \in \mathbb{N}\; \forall n \in \mathbb{N}\; (P(m, n) \Rightarrow Q(m, n)).$$

Beweis: Es seien dazu $m, n \in \mathbb{N}$.

$$\begin{aligned} & m^2 < n^2 && \text{nach Definition von } P \\ \Rightarrow\quad & 0 < n^2 - m^2 && \text{nach Subtraktion von } m^2 \\ \Rightarrow\quad & 0 < (n-m)(n+m) && \text{nach Rechenregeln in } \mathbb{N} \\ \Rightarrow\quad & 0 < n - m && \text{da } n+m > 0 \text{ und nach Regeln von } < \text{ in } \mathbb{N} \\ \Rightarrow\quad & m < n && \text{nach Rechenregeln in } \mathbb{N}. \end{aligned}$$

Ab sofort werden wir solche Beweise als Fließtext schreiben, etwa wie folgt: „Es seien $m, n \in \mathbb{N}$ und $m^2 < n^2$. Dann gilt auch $0 < n^2 - m^2 = (n-m)(n+m)$. Die Division durch die positive Zahl $n+m$ ergibt $0 < n - m$, also auch $m < n$, was zu zeigen war."

Die konkrete Benennung der verwendeten Aussageformen P und Q war für den Beweis auch nicht wesentlich, sodass wir im Folgenden darauf verzichten können.

(2) **Beweis durch Kontraposition**

Behauptung: Für natürliche Zahlen $n \in \mathbb{N}$ gilt: Wenn $4^n - 1$ eine Primzahl ist, dann ist notwendig n ungerade.

Kontraposition der Behauptung: Für natürliche Zahlen $n \in \mathbb{N}$ gilt: Wenn n gerade ist, dann ist $4^n - 1$ keine Primzahl.

Beweis: Es sei $n \in \mathbb{N}$ gerade, also gilt $n = 2k$ für eine Zahl $k \in \mathbb{N}$. Damit ist $4^n - 1 = 4^{2k} - 1 = (4^k - 1)(4^k + 1)$. Beide Faktoren sind > 1, d. h., $4^n - 1$ ist keine Primzahl.

(3) **Widerspruchsbeweis**[32]

Behauptung: Für alle reellen Zahlen $x \in \mathbb{R}$ gilt $\sin(x) + \cos(x) \neq \frac{3}{2}$.

Beweis: Wir nehmen an, es gäbe eine Zahl $x_0 \in \mathbb{R}$ mit der Eigenschaft $\sin(x) + \cos(x) = \frac{3}{2}$. Durch Quadrieren folgt dann $(\sin(x_0))^2 + (\cos(x_0))^2 + 2\,(\sin(x_0))(\cos(x_0)) = \frac{9}{4}$. Wegen $(\sin(x))^2 + (\cos(x))^2 = 1$ und $2\,(\sin(x))(\cos(x)) = \sin(2\,x)$ für alle $x \in \mathbb{R}$ (insbesondere auch für x_0) folgt also $\sin(2\,x_0) = \frac{5}{4} > 1$. Jedoch nimmt die sin-Funktion nur Werte zwischen -1 und 1 an, Widerspruch.

Weitere klassische Aussagen, die typischerweise mit Widerspruchsbeweisen gezeigt werden, sind „Es gibt unendlich viele Primzahlen" und „$\sqrt{2}$ ist keine rationale Zahl".

(4) **Beweis durch Fallunterscheidung**

Behauptung: Für jede ganze Zahl $n \in \mathbb{Z} = \{0, 1, -1, 2, -2, \ldots\}$ gilt: $n^2 + n$ ist gerade.

Beweis: Es sei $n \in \mathbb{Z}$. Wir unterscheiden zwei Fälle:

Fall 1: n ist ungerade.

In diesem Fall gilt also $n = 2\,k + 1$ für ein $k \in \mathbb{Z}$. Dann ist

$$n^2 + n = (2\,k+1)^2 + 2\,k + 1 = 4\,k^2 + 4\,k + 1 + 2\,k + 1 = 4\,k^2 + 6\,k + 2,$$

also eine gerade Zahl.

Fall 2: n ist gerade.

[32] Dieses Beispiel ist Thiele, 1979 entnommen.

In diesem Fall gilt also $n = 2\,k$ für ein $k \in \mathbb{Z}$. Dann ist

$$n^2 + n = (2\,k)^2 + 2\,k = 4\,k^2 + 2\,k,$$

also wiederum eine gerade Zahl. △

Das Ende eines Beweises wird oft mit der Abkürzung **q.e.d.** (lateinisch: quod erat demonstrandum: was zu zeigen war, englisch: what was to be proved) oder mit dem Symbol □ markiert.

Andere Sätze sind nicht als Implikation formuliert, sondern in Form mehrerer äquivalenter Aussagen $A_1 \Leftrightarrow A_2 \Leftrightarrow \cdots \Leftrightarrow A_n$. In diesem Fall verwenden wir häufig einen

(5) **Beweis durch Ringschluss** (englisch: closed chain inference). Bei diesem zeigen wir nacheinander die Implikationen $P_1 \Rightarrow P_2$, $P_2 \Rightarrow P_3$ usw. bis $P_{n-1} \Rightarrow P_n$ und $P_n \Rightarrow P_1$, was dann wiederum die gewünschten Äquivalenzen zur Folge hat. Das erfordert n Beweisschritte. Wir können sogar allgemeiner solange verschiedene Implikationen $P_i \Rightarrow P_j$ zeigen, bis wir mittels Kettenschluss von jeder der beteiligten Aussagen zu jeder anderen Aussage gelangen können. Die Anzahl der zu zeigenden Implikationen beträgt aber mindestens n.

    Quizfrage 3.1: Wieviele Implikationen wären zu zeigen, wenn man die Äquivalenz der Aussagen P_i und P_j für $i, j = 1, \ldots, n$ mit $i \neq j$ paarweise zeigen würde?

Schließlich betrachten wir noch den

(6) **Beweis durch vollständige Induktion** (englisch: proof by induction), der dann verwendet werden kann, wenn wir die Wahrheit einer Aussageform $P(n)$ für alle ganzen Zahlen $n \geqslant n_0$ ab einem gewissen Startindex $n_0 \in \mathbb{Z}$ zeigen wollen. In diesem Fall zeigen wir am **Induktionsanfang** (englisch: base case) die Wahrheit der Aussage $P(n_0)$. Oft wird der Induktionsanfang bei $n_0 = 0$ oder $n_0 = 1$ gesetzt.

    Im **Induktionsschritt** (englisch: induction step) wird für beliebiges $n \geqslant n_0$ die Aussage $P(n) \Rightarrow P(n+1)$ gezeigt. Dabei heißt $P(n)$ die **Induktionsannahme** oder **Induktionsvoraussetzung** (englisch: induction hypothesis). Bei Bedarf kann sogar auf alle vorgehenden Aussagen $P(n_0), \ldots, P(n)$ zurückgegriffen werden, also $P(n_0) \wedge \cdots \wedge P(n) \Rightarrow P(n+1)$ gezeigt werden.[33]

Expertenwissen: vollständige Induktion für allgemeinere Indexmengen

Man muss eine Indexmenge wie $\{n \in \mathbb{Z} \mid n \geqslant n_0\}$ nicht notwendig aufsteigend traversieren. Wenn man z. B. beweisen kann, dass $P(n) \Rightarrow P(2n)$ für alle $n \in \mathbb{N}$ mit $n \geqslant n_0$ gilt sowie $P(n) \Rightarrow P(n-1)$ für $n - 1 \geqslant n_0$, so hat man auch alles gezeigt.

Außerdem kann man die vollständige Induktion auch für Aussageformen über der gesamten Menge $\mathbb{Z}$ durchführen und benötigt dann zusätzlich zum Schritt von n auf $n+1$ den Schritt von n auf $n-1$.

Schließlich erfordert die Induktion nicht einmal totalgeordnete Indexmengen (Definition 5.28). Allgemein kann man fundierte Mengen nehmen, das sind halbgeordnete Mengen, deren nichtleere Teilmengen mindestens ein minimales Element besitzen.

[33] Ein schönes Beispiel für einen fehlerhaft ausgeführten Induktionsbeweis ist das Pferde-Paradoxon, bei dem „bewiesen“ wird, dass alle Pferde dieselbe Farbe haben.

Beispiel 3.2 (vollständige Induktion).
Behauptung: Die Summe der ersten n natürlichen Zahlen ist gleich $\frac{1}{2}n\,(n+1)$, also: Für alle $n \in \mathbb{N}$ gilt

$$P(n) : \sum_{j=1}^{n} j = \frac{1}{2} n\,(n+1).$$

Induktionsanfang bei $n_0 = 1$: $P(1)$ lautet: $\sum_{j=1}^{1} j = \frac{1}{2} \cdot 1 \cdot 2$, was eine wahre Aussage ist. Wir zeigen nun im Induktionsschritt, dass $P(n)$ auch $P(n+1)$ impliziert. Zu zeigen ist für beliebiges $n \in \mathbb{N}$ also $\sum_{j=1}^{n+1} j = \frac{1}{2}(n+1)\,(n+2)$, wobei die Aussagen $\sum_{j=1}^{n} j = \frac{1}{2}n\,(n+1)$ verwendet werden darf.

$$\begin{aligned}
\sum_{j=1}^{n+1} j &= n+1+\sum_{j=1}^{n} j && \text{wegen der Assoziativität der Addition}\\
&= n+1+\frac{1}{2}n\,(n+1) && \text{nach Induktionsannahme, dass } P(n) \text{ wahr ist}\\
&= (n+1)\left[1+\frac{1}{2}n\right] && \text{wegen des Distributivgesetzes für Addition und Multiplikation}\\
&= (n+1)\left[\frac{2+n}{2}\right] && \text{wegen } 1 = \frac{2}{2} \text{ und } \frac{1}{2}n = \frac{n}{2}\\
&= \frac{1}{2}(n+1)\,(n+2) && \text{wegen der Kommutativität der Multiplikation,}
\end{aligned}$$

was $P(n+1)$ entspricht. △

Ende der Vorlesung 2

Ende der Woche 1

§ 4 Mengenlehre

Literatur: Deiser, 2024b, Kapitel 1.2; Deiser, 2024a, Kapitel 1.1; Jänich, 2008, Kapitel 1.1; Jänich, 2008, Kapitel 6

Mengen sind das klassische Fundament der Mathematik. Georg Cantor, Begründer der Mengenlehre, hat 1895 folgenden Versuch der Definition einer Menge angegeben:

> „Unter einer **Menge** verstehen wir jede Zusammenfassung X von bestimmten wohlunterschiedenen Objekten x unserer Anschauung oder unseres Denkens (welche die **Elemente** von X genannt werden) zu einem Ganzen."

Diese ursprüngliche Definition hat allerdings Schwächen, wie wir gleich noch sehen werden.

Wir bezeichnen Mengen oft mit Großbuchstaben. Ist X eine Menge (englisch: set) und x ein Element (englisch: element) von X, so notieren wir diese Beziehung als $x \in X$ (seltener auch $X \ni x$) und lesen „x ist Element von X" oder kurz „x in X" oder auch „X enthält x". Das Symbol $x \notin X$ (oder $X \not\ni x$) drückt aus, dass x *kein* Element von X ist.

Mengen sind vollständig durch ihre Elemente bestimmt. Zwei Mengen X und Y sind also genau dann **gleich** (englisch: equality of sets), wenn sie dieselben Elemente enthalten. In Symbolen:

$$X = Y \quad \text{ist definiert als die Aussage} \quad \forall x \in X \ (x \in Y) \ \wedge \ \forall y \in Y \ (y \in X).$$

Mengen können beispielsweise durch Aufzählung ihrer Elemente in geschweiften Klammern $\{\}$ angegeben werden, etwa

$$X \coloneqq \{2, 3, 5\}.$$

Da Mengen nur aus „wohlunterschiedenen" Elementen bestehen und es auf die Reihenfolge nicht ankommt, könnten wir dieselbe Menge auch als

$$X \coloneqq \{5, 2, 3, 2\}$$

beschreiben. Bei der „Konstruktion" der Menge wird das doppelte Vorkommen des Elements 2 also ignoriert.

Wichtige Mengen sind die **Zahlbereiche** (englisch: number systems)

$\mathbb{N} \coloneqq \{1, 2, 3, \dots\}$	Menge der **natürlichen Zahlen**[34]	(4.1a)
$\mathbb{N}_0 \coloneqq \{0, 1, 2, 3, \dots\}$	Menge der **natürlichen Zahlen mit Null**	(4.1b)
$\mathbb{Z} \coloneqq \{0, 1, -1, 2, -2, \dots\}$	Menge der **ganzen Zahlen**[35]	(4.1c)
$\widetilde{\mathbb{Q}} \coloneqq \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\} \right\}$	vorläufige Menge der **rationalen Zahlen**[36]	(4.1d)
$\mathbb{R}$	Menge der **reellen Zahlen**[37]	(4.1e)
$\mathbb{C} \coloneqq \{a + b\,i \mid a, b \in \mathbb{R}\}$	Menge der **komplexen Zahlen**[38],	(4.1f)

die hier nur informell definiert werden. Für die tatsächliche Definition der rationalen Zahlen $\mathbb{Q}$ verweisen wir auf (5.18). Die mengentheoretische Konstruktion der Zahlbereiche aus (4.1) wird in Anhang A angedeutet.

Eine weitere Möglichkeit, Mengen anzugeben, besteht darin, Elemente anhand bestimmter Eigenschaften zu sammeln. Es sei dazu P eine Aussageform mit Grundbereich X, der eine Menge sein soll. Dann können wir

$$Y \coloneqq \{x \in X \mid P(x)\} \tag{4.2}$$

betrachten, bestehend aus den Elementen von X, für die $P(x)$ eine wahre Aussage ist. Diese Konstruktion heißt **Mengenkomprehension** (englisch: set comprehension).

Hier erkennt man ein Problem der sehr freien Definition einer Menge nach Cantor. Sie lässt es beispielsweise zu, X als die Menge aller Mengen zu definieren. Wählen wir dann $P(x)$ als die Aussageform „enthält sich nicht selbst als Element", so definiert

$$Z \coloneqq \{x \in X \mid x \notin x\}$$

also die „Menge aller Mengen, die sich nicht selbst als Element enthalten". Stellen wir jetzt die Frage, ob Z sich selbst enthält, so ergibt sich folgendes Problem:

[34] englisch: natural numbers
[35] englisch: integer numbers, lateinisch: integer: ganz, unversehrt
[36] englisch: rational numbers, lateinisch: ratio: Verhältnis
[37] englisch: real numbers
[38] englisch: complex numbers

- Falls Z sich selbst enthält ($Z \in Z$), dann liegt das daran, dass Z die Komprehensionsbedingung $Z \notin Z$ erfüllt.
- Falls Z sich nicht selbst enthält ($Z \notin Z$), dann erfüllt Z die Komprehensionsbedingung $Z \notin Z$ nicht, also gilt $Z \in Z$.

In Kurzform erhalten wir den Widerspruch $Z \in Z \Leftrightarrow Z \notin Z$. Dieser Widerspruch ist als **Russell-Paradoxon** (englisch: Russell's paradox) oder **Russell-Antinomie** der „naiven" Cantorschen Mengenlehre bekannt geworden, entdeckt 1901 von Russell und unabhängig etwa zeitgleich von Zermelo.[39]

Die Auflösung in der axiomatischen Mengenlehre nach Zermelo und Fraenkel (**ZF-Mengenlehre**, englisch: ZF set theory) besteht darin, den Mengenbegriff geeignet einzuschränken, sodass Konstruktionen wie die „Menge aller Mengen" nicht mehr möglich sind. In dieser Lehrveranstaltung können wir die zugehörigen Axiome[40] nicht behandeln und verweisen auf spätere Spezialveranstaltungen. Wir weisen aber darauf hin, dass die Mengenkomprehension (4.2) in Form des sogenannten **Aussonderungsaxioms** (englisch: axiom schema of separation) als Konstruktionsprinzip von Mengen weiterhin vorkommt. Wesentlich ist nur eben, dass der Grundbereich X der Aussageform P eine Menge im Sinne der ZF-Axiome sein muss.[41]

Intervalle in den reellen Zahlen lassen sich beispielsweise über Mengenkomprehension definieren:[42]

Beispiel 4.1 (Mengenkomprehension)**.**

Es seien $a, b \in \mathbb{R}$. Dann heißt

$[a, b] \coloneqq \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ **abgeschlossenes Intervall**[43]

$(a, b] \coloneqq \{x \in \mathbb{R} \mid a < x \leqslant b\}$ **links offenes, rechts abgeschlossenes Intervall**[44]

$[a, b) \coloneqq \{x \in \mathbb{R} \mid a \leqslant x < b\}$ **links abgeschlossenes, rechts offenes Intervall**[45]

$(a, b) \coloneqq \{x \in \mathbb{R} \mid a < x < b\}$ **offenes Intervall**[46]

$[a, \infty) \coloneqq \{x \in \mathbb{R} \mid a \leqslant x\}$ **rechtsseitig unendliches abgeschlossenes Intervall**[47]

$(a, \infty) \coloneqq \{x \in \mathbb{R} \mid a < x\}$ **rechtsseitig unendliches offenes Intervall**[48]

$(-\infty, b] \coloneqq \{x \in \mathbb{R} \mid x \leqslant b\}$ **linksseitig unendliches abgeschlossenes Intervall**[49]

$(-\infty, b) \coloneqq \{x \in \mathbb{R} \mid x < b\}$ **linksseitig unendliches offenes Intervall**[50]

$(-\infty, \infty) \coloneqq \{x \in \mathbb{R} \mid \top\} = \mathbb{R}$ **beidseitig unendliches Intervall**[51].

[39] Eine bekannte andere Formulierung des Russell-Paradoxons ist die folgende. In einem Dorf lebt ein (männlicher) Barbier, der alle Männer rasiert, die sich nicht selbst rasieren. Rasiert der Dorfbarbier sich selbst?

[40] Bei Interesse können Sie sich aber unter https://de.wikipedia.org/wiki/Zermelo-Fraenkel-Mengenlehre einen Eindruck verschaffen.

[41] Ist der Grundbereich keine Menge, so landet man beim Begriff der **Klasse** (englisch: class), siehe etwa Deiser, 2024a, Kapitel 3. Stark vereinfacht gesagt ist eine Klasse ein Objekt, das „zu groß" ist, um eine Menge zu sein. Ein wichtiges Beispiel ist die **Klasse aller Mengen** (englisch: class of all sets).

[42] Hierbei ist „$\leqslant$" die übliche Totalordnung auf $\mathbb{R}$. Mehr dazu in Definition 5.28 und Anhang A.

Dabei ist $\{x \in \mathbb{R} \,|\, a \leqslant x \leqslant b\}$ eine gebräuchliche Kurzschreibweise für $\{x \in \mathbb{R} \,|\, a \leqslant x \,\wedge\, x \leqslant b\}$.

Die Intervalle der Form $[a, b]$, $(a, b]$, $[a, b)$ und (a, b) heißen **endliche Intervalle** (englisch: finite intervals) oder **beschränkte Intervalle** (englisch: bounded intervals) mit **Endpunkten** (englisch: end points) $a, b \in \mathbb{R}$. Diese Intervalle sind leer, falls $b < a$ bzw. $b \leqslant a$ sein sollte. Die Bedeutung der Eigenschaften **offen** (englisch: open) und **abgeschlossen** (englisch: closed) wird typischerweise in Lehrveranstaltungen zur *Analysis* behandelt.

Wir definieren für $a, b \in \mathbb{Z}$ auch das **ganzzahlige Intervall** (englisch: integer interval)

$$[\![a, b]\!] \coloneqq [a, b] \cap \mathbb{Z}.$$

Schließlich werden wir selbsterklärende Symbole wie $\mathbb{R}_{\geqslant 0}$, $\mathbb{Q}_{\neq 0}$ usw. verwenden. △

Definition 4.2 (Teilmenge, Obermenge)**.**
Für Mengen X und Y definieren wir:

(*i*) X ist eine **Teilmenge** (englisch: subset) von Y, kurz: $X \subseteq Y$, wenn jedes Element von X auch ein Element von Y ist, kurz: $\forall x \in X \; (x \in Y)$. In diesem Fall sagen wir auch, Y sei eine **Obermenge** (englisch: superset) von X, und schreiben $Y \supseteq X$.

(*ii*) X ist eine **echte Teilmenge** (englisch: proper subset) von Y, kurz: $X \subsetneq Y$, falls $X \subseteq Y$ und $X \neq Y$ gilt. In diesem Fall sagen wir auch, Y sei eine **echte Obermenge** (englisch: proper superset) von X, und schreiben $Y \supsetneq X$.

Die Teilmengenbeziehung $\subseteq$ zwischen Mengen heißt auch **Inklusion** (englisch: inclusion, lateinisch: includere: einschließen) △

Beispielsweise erzeugt die Mengenkomprehension (4.2), also $Y = \{x \in X \,|\, P(x)\}$, immer eine Teilmenge $Y \subseteq X$. Außerdem gelten die echten Inklusionen

$$\mathbb{N} \subsetneq \mathbb{N}_0 \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}.$$

für die Zahlbereiche aus (4.1). **Quizfrage 4.1:** Wie kann man sich davon überzeugen, dass die Inklusionen echt sind?

In der axiomatischen Mengenlehre gibt es genau eine Menge, die keine Elemente enthält: die **leere Menge** (englisch: empty set) $\emptyset$. Eine Menge heißt **nichtleer** (englisch: non-empty set), wenn sie von $\emptyset$ verschieden ist, also mindestens ein Element enthält. Wir schreiben dann $X \neq \emptyset$.

[43] englisch: closed interval
[44] englisch: left-open, right-closed interval
[45] englisch: left-closed, right-open interval
[46] englisch: open interval. Bei der Notation (a, b) für offene Intervalle besteht eine Verwechslungsgefahr mit den Elementen (a, b) des kartesischen Produkts von zwei Mengen, siehe Definition 4.8.
[47] englisch: unbounded above, closed interval
[48] englisch: unbounded above, open interval
[49] englisch: unbounded below, closed interval
[50] englisch: unbounded below, open interval
[51] englisch: unbounded above and below interval

Definition 4.3 (Durchschnitt, disjunkte Mengen, Vereinigung, disjunkte Vereinigung).

(i) Es sei $\mathcal{A}$ eine nichtleere Menge von Mengen. Dann heißt die Menge

$$\bigcap \mathcal{A} := \{x \mid \forall A \in \mathcal{A}\ (x \in A)\} \tag{4.3a}$$

die **Schnittmenge**, der **Durchschnitt** oder kurz der **Schnitt** (englisch: intersection) von $\mathcal{A}$. Sind die Elemente von $\mathcal{A}$ über eine nichtleere Indexmenge I indiziert, gilt also $\mathcal{A} = \{A_i \mid i \in I\}$, so schreiben wir statt $\bigcap\{A_i \mid i \in I\}$ auch

$$\bigcap_{i \in I} A_i := \{x \mid \forall i \in I\ (x \in A_i)\}. \tag{4.3b}$$

Ist speziell $\mathcal{A} = \{A_1, A_2\}$, so schreiben wir statt $\bigcap\{A_1, A_2\}$ auch

$$A_1 \cap A_2 := \{x \mid x \in A_1 \ \wedge\ x \in A_2\}. \tag{4.3c}$$

Gilt $\bigcap \mathcal{A} = \emptyset$ bzw. $\bigcap_{i \in I} A_i = \emptyset$ bzw. $A_1 \cap A_2 = \emptyset$, so heißen die Elemente von $\mathcal{A}$ bzw. die Mengen A_i bzw. die Mengen A_1 und A_2 **disjunkt** (englisch: disjoint).

(ii) Es sei $\mathcal{A}$ eine (möglicherweise leere) Menge von Mengen. Dann heißt die Menge

$$\bigcup \mathcal{A} := \{x \mid \exists A \in \mathcal{A}\ (x \in A)\} \tag{4.4a}$$

die **Vereinigungmenge** oder die **Vereinigung** (englisch: union) von $\mathcal{A}$.[52] Sind die Elemente von $\mathcal{A}$ über eine Indexmenge I indiziert, gilt also $\mathcal{A} = \{A_i \mid i \in I\}$, so schreiben wir statt $\bigcup\{A_i \mid i \in I\}$ auch

$$\bigcup_{i \in I} A_i := \{x \mid \exists i \in I\ (x \in A_i)\}. \tag{4.4b}$$

Ist speziell $\mathcal{A} = \{A_1, A_2\}$, so schreiben wir statt $\bigcup\{A_1, A_2\}$ auch

$$A_1 \cup A_2 := \{x \mid x \in A_1 \ \vee\ x \in A_2\}. \tag{4.4c}$$

(iii) Sind die Elemente von $\mathcal{A}$ aus Aussage (ii) **paarweise disjunkt** (englisch: pairwise disjoint), gilt also $A \cap B = \emptyset$ für alle $A, B \in \mathcal{A}$ mit $A \neq B$, dann können wir das verdeutlichen, indem wir für die Vereinigungsmenge (4.4a)

$$\dot{\bigcup} \mathcal{A} := \{x \mid \exists A \in \mathcal{A}\ (x \in A)\} \tag{4.5a}$$

schreiben und von der **disjunkten Vereinigungmenge** oder von der **disjunkten Vereinigung** (englisch: disjoint union) von $\mathcal{A}$ sprechen. Sind die Elemente von $\mathcal{A}$ über eine Indexmenge I indiziert, gilt also $\mathcal{A} = \{A_i \mid i \in I\}$ und gilt $A_i \cap A_j = \emptyset$ für alle $i, j \in I$ mit $i \neq j$, so schreiben wir auch

$$\dot{\bigcup_{i \in I}} A_i := \{x \mid \exists i \in I\ (x \in A_i)\}. \tag{4.5b}$$

Ist speziell $\mathcal{A} = \{A_1, A_2\}$ mit $A_1 \cap A_2 = \emptyset$, so schreiben wir auch

$$A_1 \dot{\cup} A_2 := \{x \mid x \in A_1 \ \vee\ x \in A_2\}. \tag{4.5c}$$

△

[52] Die Existenz der Vereinigungsmenge (4.4a) folgt direkt aus einem der ZF-Axiome. In (4.3a) sollte dann explizit $\bigcap \mathcal{A} := \{x \in \bigcup \mathcal{A} \mid \forall A \in \mathcal{A}\ (x \in A)\}$ stehen.

Definition 4.4 (Differenz, symmetrische Differenz, Komplement)**.**
Für Mengen X und Y definieren wir

(i) die **Differenzmenge** (englisch: set difference) **von** Y **in** X

$$X \setminus Y := \{x \in X \mid x \notin Y\}, \tag{4.6}$$

kurz auch als „X **ohne** Y" bezeichnet.

(ii) die **symmetrische Differenz** (englisch: symmetric difference) **von** X **und** Y

$$X \mathbin{\triangle} Y := (X \setminus Y) \cup (Y \setminus X). \tag{4.7}$$

Ist weiter X eine Menge und $Y \subseteq X$ eine Teilmenge, so definieren wir

(iii) das **Komplement** (englisch: complement) **von** Y **in** X

$$Y^c := X \setminus Y = \{x \in X \mid x \notin Y\}. \tag{4.8}$$

Da die Menge X im Symbol Y^c nicht angegeben wird, muss sie aus dem Zusammenhang klar sein. △

Quizfrage 4.2: Was sind $X \mathbin{\triangle} X$ und $X \mathbin{\triangle} \emptyset$?

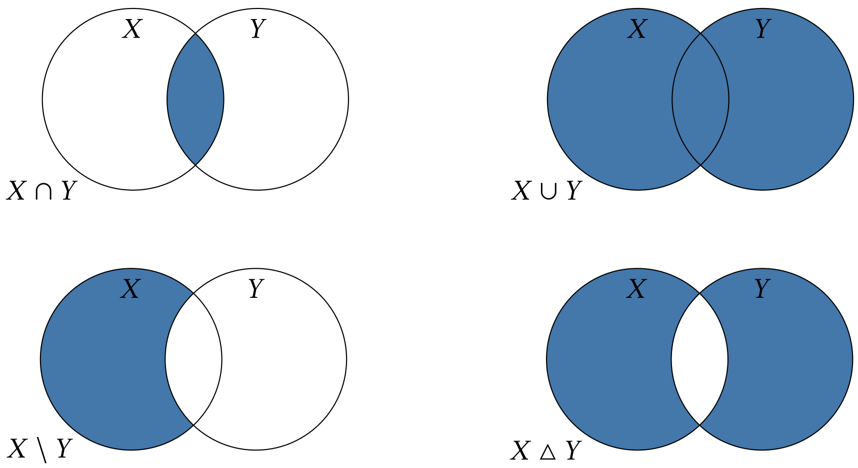


Abbildung 4.1.: Darstellung der Mengenoperationen $X \cap Y$, $X \cup Y$, $X \setminus Y$ und $X \mathbin{\triangle} Y$ aus Definition 4.4.

Lemma 4.5 (Eigenschaften von Schnitt und Vereinigung)**.**
Es seien X, Y und Z Mengen. Dann gilt:

$$X \cap Y = Y \cap X \qquad \textbf{Kommutativität von } \cap \tag{4.9a}$$
$$X \cup Y = Y \cup X \qquad \textbf{Kommutativität von } \cup \tag{4.9b}$$

$$(X \cap Y) \cap Z = X \cap (Y \cap Z) \qquad \textbf{Assoziativität von } \cap \tag{4.10a}$$
$$(X \cup Y) \cup Z = X \cup (Y \cup Z) \qquad \textbf{Assoziativität von } \cup \tag{4.10b}$$

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z) \qquad \textbf{Distributivität} \tag{4.11a}$$
$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z) \qquad \textbf{Distributivität} \tag{4.11b}$$

$$X \setminus Y = X \setminus (X \cap Y) \tag{4.12}$$

$$X \cap Y = X \quad \Leftrightarrow \quad X \subseteq Y \tag{4.13a}$$
$$X \cup Y = Y \quad \Leftrightarrow \quad X \subseteq Y. \tag{4.13b}$$

Sind Y und Z Teilmengen einer Menge X, bzgl. der wir das Komplement nehmen, so gilt weiter:

$$(Y \cap Z)^c = Y^c \cup Z^c \qquad \textbf{De Morgansches Gesetz} \tag{4.14a}$$
$$(Y \cup Z)^c = Y^c \cap Z^c \qquad \textbf{De Morgansches Gesetz} \tag{4.14b}$$

$$(Y^c)^c = Y \qquad \text{Komplementbildung ist } \textbf{involutorisch}^{53} \tag{4.15}$$

$$Y \subseteq Z \quad \Leftrightarrow \quad Z^c \subseteq Y^c. \tag{4.16}$$

Beweis. Der Beweis kann durch Ausnutzung von $X = Y \Leftrightarrow \forall x \; (x \in X \leftrightarrow x \in Y)$ und $X \subseteq Y \Leftrightarrow \forall x \; (x \in X \rightarrow x \in Y)$ auf Satz 1.9 zurückgeführt werden. Die Details werden hier nicht ausgeführt. □

Zur Vereinfachung der Notation vereinbaren wir auch hier wieder Bindungsregeln:

$$\begin{aligned} \cdot^c \quad & \text{bindet stärker als} \quad \setminus \\ \setminus \quad & \text{bindet stärker als} \quad \cap \\ \cap \quad & \text{bindet stärker als} \quad \cup\,. \end{aligned} \tag{4.17}$$

Dadurch könnten wir beispielsweise das erste Distributivgesetz (4.11a) auch als $X \cap (Y \cup Z) = X \cap Y \,\cup\, X \cap Z$ schreiben. Es gilt jedoch auch hier, dass Klammern zur Verdeutlichung nicht schaden können.

Definition 4.6 (Potenzmenge)**.**
Für jede Menge X heißt

$$\mathcal{P}(X) \coloneqq \{A \mid A \subseteq X\} \tag{4.18}$$

die **Potenzmenge** (englisch: power set) von X. △

Die Potenzmenge von X ist also die Menge aller Teilmengen von X. In der axiomatischen Mengenlehre nach Zermelo und Fraenkel gibt es das Potenzmengenaxiom, das garantiert, dass jede Menge eine Potenzmenge besitzt.

[53] auch: selbstinvers, englisch: involutory, involutive, self-inverse, lateinisch: involvere: einwickeln

Beispiel 4.7 (Potenzmenge).

(*i*) Für $X = \emptyset$ ist $\mathcal{P}(X) = \{\emptyset\}$.

(*ii*) Für $X = \{a\}$ ist $\mathcal{P}(X) = \big\{\emptyset, \{a\}\big\}$.

(*iii*) Für $X = \{a, b\}$ ist $\mathcal{P}(X) = \big\{\emptyset, \{a\}, \{b\}, \{a, b\}\big\}$. △

Definition 4.8 (kartesisches Produkt endlich vieler Mengen).

(*i*) Für Mengen X und Y definieren wir das **kartesische Produkt** (englisch: Cartesian product) oder **Kreuzprodukt** (englisch: cross product)

$$X \times Y \coloneqq \{(x, y) \,|\, x \in X \;\wedge\; y \in Y\}. \tag{4.19}$$

Die Elemente des kartesischen Produkts heißen **geordnete Paare** (englisch: ordered pairs) oder einfach **Paare** (englisch: pairs) (x, y).[54]

(*ii*) Analog können wir auch das kartesische Produkt von mehr als zwei Mengen definieren, etwa $X \times Y \times Z$, dessen Elemente **Tripel** (englisch: triplets) (x, y, z) sind.[55]

(*iii*) Allgemeiner heißen die Elemente $(x_1, x_2, \ldots, x_n)$ des Produkts $\bigtimes_{i=1}^n X_i = X_1 \times \cdots \times X_n$ von $n \in \mathbb{N}_0$ Mengen **n-Tupel** (englisch: n-tuples).[56] Dabei gilt $x_i \in X_i$ für $i = 1, \ldots, n$.

(*iv*) Die Einträge $x_1, \ldots, x_n$ eines n-Tupels $(x_1, \ldots, x_n)$ für $n \in \mathbb{N}_0$ heißen die **Komponenten** (englisch: components) des Tupels. Die Anzahl der Komponenten eines Tupels heißt dessen **Länge** (englisch: length).

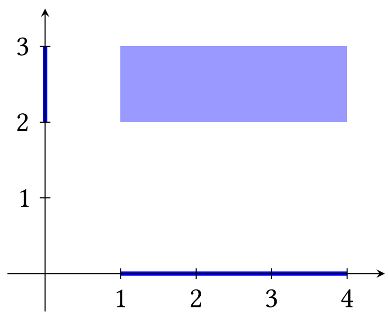
(*v*) Wir schreiben $X^2 = X \times X$ und allgemeiner $X^n = \bigtimes_{i=1}^n X$ für das kartesische Produkt einer Menge X mit sich selbst für $n \in \mathbb{N}_0$. Dabei verstehen wir unter X^0 die Menge $\{()\}$, die nur das **leere Tupel** (englisch: empty tuple) enthält. △

Beispiel 4.9 (kartesisches Produkt).

(*i*) Ist $X = \{\text{Kreuz}, \text{Pik}, \text{Herz}, \text{Karo}\}$ und $Y = \{7, 8, 9, 10, \text{Bube}, \text{Dame}, \text{König}, \text{As}\}$, so entsprechen die Elemente des kartesischen Produkts $X \times Y$ gerade den 32 Karten eines Skatspiels, also $(\text{Kreuz}, 7)$, $(\text{Kreuz}, 8)$ usw. bis (Karo, As).

(*ii*) Das kartesische Produkt von Intervallen in $\mathbb{R}$ heißt ein **mehrdimensionales Intervall** (englisch: multi-dimensional interval).

Für $X = [1, 4]$ und $Y = [2, 3]$ können wir das mehrdimensionale Intervall $X \times Y = \{(x_1, x_2) \,|\, 1 \leqslant x_1 \leqslant 4 \;\wedge\; 2 \leqslant x_2 \leqslant 3\} \subseteq \mathbb{R} \times \mathbb{R}$ wie folgt illustrieren:



△

Ende der Vorlesung 3

[54] Ein geordnetes Paar kann in der Sprache der Mengenlehre durch die **Kuratowski-Darstellung** $(x, y) \coloneqq \big\{\{x\}, \{x, y\}\big\}$ realisiert werden.

[55] Tripel können in der Sprache der Mengenlehre durch die Rekursion $(x, y, z) \coloneqq \big((x, y), z\big)$ realisiert werden.

[56] n-Tupel können in der Sprache der Mengenlehre rekursiv durch $\big(x_1, x_2, \ldots, x_n\big) \coloneqq \big((x_1, x_2, \ldots, x_{n-1}), x_n\big)$ für $n \geqslant 2$ und $(x_1) \coloneqq x_1$ für $n = 1$ realisiert werden. Ein Paar ist ein 2-Tupel, ein Tripel ist ein 3-Tupel.

§ 5 Relationen

Literatur: Deiser, 2024b, Kapitel 1.3

Relationen geben Beziehungen zwischen Objekten an wie beispielsweise „$1 \leqslant 3$“ oder „$5 \in \mathbb{N}$“ oder „$3 \mid 756$“ („3 teilt 756“) oder „Berlin ist die Hauptstadt von Deutschland“.

Definition 5.1 (Relation)**.**

Es seien X und Y Mengen.

(i) Ist $R \subseteq X \times Y$, so heißt (R, X, Y) eine (**zweistellige**) **Relation** (englisch: relation, lateinisch: relatio: Verhältnis, Beziehung) **zwischen** X und Y. Die Menge R heißt der **Graph der Relation** (englisch: graph of a relation).

(ii) Sind $R, S \subseteq X \times Y$ zwei Relationen zwischen X und Y und gilt $R \subseteq S$, dann heißt R eine **Teilrelation** (englisch: subrelation) **von** S, und S heißt eine **Oberrelation** (englisch: superrelation) **von** R.

(iii) Im Fall $Y = X$ sprechen wir von einer **homogenen Relation** (englisch: homogeneous relation) **auf** X. △

Wenn X und Y klar sind, sagt man auch oft, $R \subseteq X \times Y$ selbst sei die Relation. Statt $(x, y) \in R$ schreiben wir auch $x \mathrel{R} y$, um die Lesart „x steht in Relation zu y“ zu erleichtern.

Beachte: Zwei Relationen (R_1, X_1, Y_1) und (R_2, X_2, Y_2) sind genau dann gleich, wenn $X_1 = X_2$, $Y_1 = Y_2$ und $R_1 = R_2$ gilt.

Beispiel 5.2 (Relation)**.**

(i) Ist X die Menge der Teilnehmenden an der Lehrveranstaltung *Lineare Algebra I* und $Y = \{\text{Mathematik}, \text{Physik}, \text{Informatik}\}$ eine Menge von Studienfächern, so ergibt die Beziehung „Die teilnehmende Person x studiert das Fach y.“ eine Relation zwischen X und Y.

(ii) Wir sagen, die Zahl $x \in \mathbb{Z}$ **teilt** (englisch: divides) die Zahl $y \in \mathbb{Z}$, in Symbolen: $x \mid y$, wenn eine Zahl $n \in \mathbb{Z}$ existiert, sodass $y = n\,x$ gilt. Insbesondere teilt jede ganze Zahl die Zahl 0, und die Zahl 1 teilt jede ganze Zahl.

Die folgende Tabelle stellt die **Teilbarkeitsrelation** (englisch: divisibility relation) $x \mid y$ („Die Zahl x teilt die Zahl y.“) auf der Menge $X = Y = \{0, 1, 2, \ldots, 10\} = [\![0, 10]\!]$ dar:

$x \backslash y$	0	1	2	3	4	5	6	7	8	9	10
0	•										
1	•	•	•	•	•	•	•	•	•	•	•
2	•		•		•		•		•		•
3	•			•			•			•	
4	•				•				•		
5	•					•					•
6	•						•				
7	•							•			
8	•								•		
9	•									•	
10	•										•

(*iii*) Es sei $X = Y = \mathbb{R}$ und $R = \{(x, y) \in \mathbb{R}^2 \,|\, x \leqslant y\}$ die **gewöhnliche Kleiner-Gleich-Relation auf $\mathbb{R}$** (englisch: usual less-or-equal relation).

(*iv*) Es sei X eine Menge, $\mathcal{P}(X)$ die Potenzmenge und $R = \{(A, B) \in \mathcal{P}(X) \times \mathcal{P}(X) \,|\, A \subseteq B\}$ die **Teilmengenrelation** (englisch: subset relation) oder **Inklusionsrelation** (englisch: inclusion relation) auf $\mathcal{P}(X)$.

(*v*) Auf einer beliebigen Menge X heißt die Menge

$$\Delta_X \coloneqq \{(x, y) \in X \times X \,|\, x = y\} \tag{5.1}$$

die **Diagonale** (englisch: diagonal) in $X \times X$. Die Relation $\mathrm{id}_X \coloneqq (\Delta_X, X, X)$ heißt die **Gleichheitsrelation** (englisch: equality relation) oder **Identitätsrelation** (englisch: identity relation) auf der Menge X.

(*vi*) Auf einer beliebigen Menge X heißt die Relation $U_X \coloneqq (U, X, X)$ mit $U = X \times X$ die **universelle Relation** (englisch: universal relation). △

Quizfrage 5.1: Können Sie weitere Beispiele für Relationen benennen?

Definition 5.3 (Einschränkung von Relationen)**.**
Es seien X und Y Mengen und (R, X, Y) eine Relation zwischen X und Y. Sind $A \subseteq X$ und $B \subseteq Y$ Teilmengen, dann heißt $(R|_{A \times B}, A, B)$ mit

$$R|_{A \times B} \coloneqq \{(x, y) \in R \,|\, x \in A, \;\; y \in B\} \tag{5.2}$$

die **Einschränkung** oder **Restriktion** (englisch: restriction, lateinisch: restringere: zurückziehen) von f **auf** $A \times B$. △

Im Falle einer homogenen Relation auf X mit Teilmenge $A \subseteq X$ betrachtet man häufig die Einschränkung $R|_{A \times A}$, die dann wiederum eine homogene Relation, und zwar auf der Teilmenge $A \subseteq X$, ist.

Definition 5.4 (Komposition von Relationen)**.**

Es seien X, Y und Z Mengen sowie (R, X, Y) und (S, Y, Z) zwei Relationen. Dann heißt die Relation $(S \circ R, X, Z)$ mit

$$S \circ R \coloneqq \{(x, z) \in X \times Z \mid \exists y \in Y \text{ mit } (x, y) \in R \text{ und } (y, z) \in S\} \tag{5.3}$$

die **Komposition** (englisch: composition, lateinisch: componere: zusammenstellen), die **Hintereinanderausführung** oder die **Verkettung** von R und S. Um die Reihenfolge klar zu benennen, sagt man auch „S **nach** R“. △

Quizfrage 5.2: Durch die Komposition welcher Relationen kann man die Relation „ist Onkel von“ ausdrücken? Und die Relation „ist Urgroßmutter von“?

Lemma 5.5 (Rechenregeln für die Komposition von Relationen).
Es seien X, Y, Z und W Mengen sowie (R, X, Y), (R_1, X, Y), (R_2, X, Y) sowie (S, Y, Z), (S_1, Y, Z), (S_2, Y, Z) und (T, Z, W) Relationen.

(*i*) Die Komposition ist assoziativ, d. h., es gilt

$$(T \circ S) \circ R = T \circ (S \circ R). \tag{5.4}$$

(*ii*) Für $\circ$ und $\cup$ gelten die **Distributivgesetze**

$$(S_1 \cup S_2) \circ R = (S_1 \circ R) \cup (S_2 \circ R), \tag{5.5a}$$
$$S \circ (R_1 \cup R_2) = (S \circ R_1) \cup (S \circ R_2). \tag{5.5b}$$

Beweis. Der Beweis ist Gegenstand der Übung. □

Beachte: Für $\circ$ und $\cap$ gelten die Distributivgesetze nicht! (**Quizfrage 5.3:** Beispiel?)

Definition 5.6 (Umkehrrelation).
Es seien X und Y Mengen und (R, X, Y) eine Relation. Dann heißt (R^{-1}, Y, X) die **Umkehrrelation** (englisch: reverse relation) oder **inverse Relation** (englisch: inverse relation) von R, wobei

$$R^{-1} \coloneqq \big\{(y, x) \in Y \times X \,\big|\, (x, y) \in R\big\} \subseteq Y \times X$$

definiert ist. △

Quizfrage 5.4: Wie bezeichnet man die Umkehrrelationen von „kleiner oder gleich sein als“, „Teilmenge sein von“ bzw. „Teiler sein von“?

Quizfrage 5.5: Wie könnte man die Umkehrrelationen der Teilbarkeitsrelation auf $\mathbb{Z}$ bezeichnen? Welche Darstellung hat diese Umkehrrelation als Tabelle wie in Beispiel 5.2?

Lemma 5.7 (Die Umkehrrelation kommutiert mit der Vereinigung).

(*i*) Es seien R_1, R_2 Relationen zwischen den Mengen X und Y. Dann gilt

$$(R_1 \cup R_2)^{-1} = R_1^{-1} \cup R_2^{-1}. \tag{5.6a}$$

(*ii*) Es sei $\mathcal{R}$ eine Menge von Relationen zwischen den Mengen X und Y. Dann gilt

$$\Big(\bigcup \mathcal{R}\Big)^{-1} = \bigcup \{R^{-1} \,|\, R \in \mathcal{R}\} \tag{5.6b}$$

Beweis. Aussage (*i*):

$$\begin{aligned}(R_1 \cup R_2)^{-1} &= \{(y,x) \,|\, (x,y) \in R_1 \text{ oder } (x,y) \in R_2\} \\ &= \{(y,x) \,|\, (x,y) \in R_1\} \cup \{(y,x) \,|\, (x,y) \in R_2\} \\ &= R_1^{-1} \cup R_2^{-1}.\end{aligned}$$

Der Beweis von Aussage (*ii*) erfolgt analog. □

Quizfrage 5.6: Was ist die Umkehrrelation von „ist Onkel von“ und von „ist Urgroßmutter von“?

Definition 5.8 (Potenzen homogener Relationen)**.**
Es sei X eine Menge und (R, X, X) eine Relation auf X. Wir definieren die **Potenzen** (englisch: powers, lateinisch: potentia: Macht, Kraft, Vermögen) von R für $n \in \mathbb{Z}$ rekursiv durch

$$R^0 \coloneqq \Delta_X \tag{5.7a}$$
$$R^{n+1} \coloneqq R^n \circ R \quad \text{für } n \in \mathbb{N}_0 \tag{5.7b}$$
$$R^{-n} \coloneqq (R^n)^{-1} \quad \text{für } n \in \mathbb{N}. \tag{5.7c}$$

△

Quizfrage 5.7: Wenn R die Relation „ist Elternteil von“ ist, was ist dann die Relation R^2? Und was bedeutet R^{-2}?

Wir definieren nun einige wichtige Eigenschaften, die homogene Relationen besitzen können.

Definition 5.9 (Eigenschaften homogener Relationen)**.**
Es sei X eine Menge und R eine Relation auf X.

(*i*) R heißt **reflexiv** (englisch: reflexive), wenn gilt:

$$(x,x) \in R \quad \text{für alle } x \in X.$$

(*ii*) R heißt **irreflexiv** (englisch: irreflexive), wenn gilt:

$$(x,x) \notin R \quad \text{für alle } x \in X.$$

(*iii*) R heißt **symmetrisch** (englisch: symmetric), wenn gilt:

$$(x,y) \in R \quad \Rightarrow \quad (y,x) \in R.$$

(*iv*) R heißt **antisymmetrisch** (englisch: antisymmetric), wenn gilt:

$$(x,y) \in R \text{ und } (y,x) \in R \quad \Rightarrow \quad x = y.$$

(*v*) R heißt **transitiv** (englisch: transitive), wenn gilt:

$$(x, y) \in R \text{ und } (y, z) \in R \quad \Rightarrow \quad (x, z) \in R.$$

(*vi*) R heißt **total** (englisch: total), wenn gilt:

$$(x, y) \in R \text{ oder } (y, x) \in R \quad \text{für alle } x, y \in X.$$ △

Quizfrage 5.8: Die Reflexivität von R kann man auch als $\operatorname{id}_X \subseteq R$ ausdrücken. Wie sieht das für die anderen Eigenschaften aus?

Beispiel 5.10 (Eigenschaften homogener Relationen)**.**

(*i*) Die Teilbarkeitsrelation $|$ auf $\mathbb{Z}$ ist reflexiv und transitiv, aber nicht irreflexiv, symmetrisch, antisymmetrisch oder total.

(*ii*) Die Teilbarkeitsrelation $|$ auf $\mathbb{N}_0$ ist reflexiv, antisymmetrisch und transitiv, aber nicht irreflexiv, symmetrisch oder total.

(*iii*) Die Relation „x liebt y" auf einer Menge von Personen hat in der Regel keine der sechs genannten Eigenschaften. △

Da Relationen Mengen sind, können wir auch Durchschnitte von Relationen betrachten.

Definition 5.11 (Durchschnitt von Relationen)**.**
Es sei $\mathcal{R}$ eine nichtleere Menge von Relationen zwischen den Mengen X und Y. Dann heißt die Relation

$$\bigcap \mathcal{R} := \{(x, y) \,|\, \forall R \in \mathcal{R} \; ((x, y) \in R)\} \tag{5.8}$$

der **Durchschnitt** der Relationen in $\mathcal{R}$. △

Beispiel 5.12 (Durchschnitt von Relationen)**.**

(*i*) Der Durchschnitt der Relationen „$\leqslant$" und „$\geqslant$" auf der Menge $\mathbb{R}$ ist die Diagonale $\Delta_{\mathbb{R}}$.

(*ii*) Der Durchschnitt der Relation „teilbar durch 2" und „teilbar durch 3" auf $\mathbb{Z}$ ist die Relation „teilbar durch 6". △

Lemma 5.13 (Eigenschaften homogener Relationen unter Durchschnitten)**.**
Es sei X eine Menge und $\mathcal{R}$ eine nichtleere Menge von Relationen auf X.

(*i*) Sind alle $R \in \mathcal{R}$ reflexiv, dann auch $\bigcap \mathcal{R}$.

(*ii*) Sind alle $R \in \mathcal{R}$ symmetrisch, dann auch $\bigcap \mathcal{R}$.

(*iii*) Sind alle $R \in \mathcal{R}$ transitiv, dann auch $\bigcap \mathcal{R}$.

Quizfrage 5.9: Warum führen wir keine entsprechenden Aussagen für irreflexive, antisymmetrische und totale Relationen auf?

Beweis. Aussage (i): Es sei $x \in X$, dann gilt $(x,x) \in R$ für alle $R \in \mathcal{R}$, also auch $(x,x) \in \bigcap \mathcal{R}$.

Aussage (ii): Es sei $(x,y) \in \bigcap \mathcal{R}$, also $(x,y) \in R$ für alle $R \in \mathcal{R}$. Dann folgt $(y,x) \in R$ für alle $R \in \mathcal{R}$, also $(y,x) \in \bigcap \mathcal{R}$.

Aussage (iii): Es seien $(x,y) \in \bigcap \mathcal{R}$ und $(y,z) \in \bigcap \mathcal{R}$, also $(x,y) \in R$ und $(y,z) \in R$ für alle $R \in \mathcal{R}$. Dann folgt $(x,z) \in R$ für alle $R \in \mathcal{R}$, also $(x,z) \in \bigcap \mathcal{R}$. □

Die Beobachtung, dass Reflexivität, Symmetrie und Transitivität unter Durchschnittsbildung erhalten bleiben, ermöglicht es uns, entsprechende **Hüllen** homogener Relationen zu definieren, siehe auch Anhang C. Dabei geht es darum, um welche Paare eine Relation mindestens erweitert werden muss, damit sie die gewünschte Eigenschaft erhält.

Definition 5.14 (Hüllen homogener Relationen)**.**
Es sei X eine Menge und R eine Relation auf X.

(i) Die **reflexive Hülle** (englisch: reflexive hull) von R ist definiert als[57]

$$R^? \coloneqq \bigcap \{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\}. \tag{5.9a}$$

(ii) Die **symmetrische Hülle** (englisch: symmetrische hull) von R ist definiert als

$$R^{\text{sym}} \coloneqq \bigcap \{S \subseteq X \times X \mid S \text{ ist symmetrisch und } R \subseteq S\}. \tag{5.9b}$$

(iii) Die **transitive Hülle** (englisch: transitive hull) von R ist definiert als

$$R^+ \coloneqq \bigcap \{S \subseteq X \times X \mid S \text{ ist transitiv und } R \subseteq S\}. \tag{5.9c}$$

(iv) Die **reflexiv-transitive Hülle** (englisch: reflexive-transitive hull) von R ist definiert als

$$R^* \coloneqq \bigcap \{S \subseteq X \times X \mid S \text{ ist reflexiv und transitiv und } R \subseteq S\}. \tag{5.9d}$$

(v) Die **reflexiv-symmetrisch-transitive Hülle** (englisch: reflexive-symmetric-transitive hull) von R ist definiert als

$$R^\sim \coloneqq \bigcap \{S \subseteq X \times X \mid S \text{ ist reflexiv, symmetrisch und transitiv und } R \subseteq S\}. \tag{5.9e}$$

Anstelle der Bezeichnung **reflexive Hülle** wird auch der Begriff **reflexiver Abschluss** verwendet, analog für die anderen Begriffe.[58] △

Aufgrund der Konstruktion sollte klar sein, dass die Hüllenbildung eine Relation nur vergrößern kann, es gilt also

$$R \subseteq R^?, \quad R \subseteq R^{\text{sym}}, \quad R \subseteq R^+, \quad R \subseteq R^* \quad \text{und} \quad R \subseteq R^\sim. \tag{5.10}$$

Aufgrund von Lemma 5.13 gilt außerdem, dass die reflexive Hülle reflexiv ist, die symmetrische Hülle symmetrisch usw. Das folgende Resultat zeigt schließlich, dass die Hüllenbildung genau dann „nichts hinzufügt“, wenn die Relation die gewünschte Eigenschaft bereits besitzt.

[57] Andere Bezeichnungen für die reflexive Hülle sind auch $R^=$ oder R^r.
[58] Zusätzlich zu den Begriffen in Definition 5.14 könnten wir noch die **reflexiv-symmetrische Hülle** und die **symmetrisch-transitive Hülle** definieren, diese haben jedoch für uns geringere Bedeutung.

Lemma 5.15 (Hüllenbildung erkennt Eigenschaften)**.**
Es sei X eine Menge und R eine Relation auf X. Dann gilt:

(i) R ist reflexiv genau dann, wenn $R^? = R$ gilt.

(ii) R ist symmetrisch genau dann, wenn $R^{\text{sym}} = R$ gilt.

(iii) R ist transitiv genau dann, wenn $R^+ = R$ gilt.

(iv) R ist reflexiv und transitiv genau dann, wenn $R^* = R$ gilt.

(v) R ist reflexiv, symmetrisch und transitiv genau dann, wenn $R^\sim = R$ gilt.

Beweis. Aussage (i): Es sei zunächst R reflexiv, dann ist R ein Element der Menge $\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\}$, über die in (5.9a) der Durchschnitt gebildet wird. Damit folgt $R^? \subseteq R$. Andererseits gilt auch die umgekehrte Inklusion $R \subseteq R^?$ wegen (5.10).

Umgekehrt gelte $R^? = R$. Wäre R nicht reflexiv, dann gäbe es ein $x \in X$ mit $(x,x) \notin R$. Jedoch gehört (x,x) zu jeder reflexiven Relation S mit der Eigenschaft $R \subseteq S$ und damit auch zu

$$\bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\} = R^?.$$

Damit erhalten wir den Widerspruch $R \subsetneq R^?$.

Die weiteren Aussagen (ii) bis (v) zeigt man analog. □

Bemerkung 5.16 (Eigenschaften der verschiedenen Hüllen)**.**
Per Konstruktion ist die reflexive Hülle $R^?$ der Relation R auf X die kleinste reflexive Oberrelation von R auf X. Genauer: Es gilt $R^? \subseteq S$ für jede reflexive Oberrelation S von R. („Jede andere reflexive Relation auf X, die R enthält, ist größer als die reflexive Hülle.")

Mit Hilfe der Begriffe für Ordnungsrelationen (§ 5.2) können wir das auch wie folgt ausdrücken: $R^?$ ist das Minimum der Teilmenge $\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\}$ bzgl. der Inklusionshalbordnung in der Menge aller Relationen auf X (also in der Potenzmenge $\mathcal{P}(X \times X)$).

Analoge Aussagen gelten für die anderen Hüllen aus Definition 5.14. △

Die Definition 5.14 liefert eine recht abstrakte Definition der verschiedenen Hüllen homogener Relationen. Wir wollen diese Hüllen daher nun charakterisieren.

Satz 5.17 (Darstellung der Hüllen homogener Relationen)**.**
Es sei X eine Menge und R eine Relation auf X. Dann gilt:

$$R^? = \bigcup_{n \in \{0,1\}} R^n = R \cup \Delta_X \quad \text{reflexive Hülle} \tag{5.11a}$$

$$R^{\text{sym}} = \bigcup_{n \in \{-1,1\}} R^n = R \cup R^{-1} \quad \text{symmetrische Hülle} \tag{5.11b}$$

$$R^+ = \bigcup_{n \in \mathbb{N}} R^n \quad \text{transitive Hülle} \tag{5.11c}$$

$$R^* = \bigcup_{n \in \mathbb{N}_0} R^n \quad \text{reflexiv-transitive Hülle} \tag{5.11d}$$

$$R^{\sim} = \bigcup_{n \in \mathbb{N}_0} (R \cup R^{-1})^n \qquad \text{reflexiv-symmetrisch-transitive Hülle.} \tag{5.11e}$$

Beweis. Wir beweisen (5.11a). Die Gleichheit $\bigcup_{n \in \{0,1\}} R^n = R \cup \Delta_X$ ist klar wegen $R^1 = R$ und $R^0 = \Delta_X$. Wir führen den Beweis in zwei Schritten:

Schritt 1: $R^? \supseteq R \cup \Delta_X$: Es sei S eine beliebige reflexive Oberrelation von R, also eine der Mengen, die im Durchschnitt $\bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\}$ vorkommt, siehe (5.9a). Dann enthält S neben R wegen der Reflexivität notwendigerweise auch Δ_X als Teilmenge. Also gilt $R \cup \Delta_X \subseteq S$. Da S beliebig war, haben wir

$$R \cup \Delta_X \subseteq \bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\} = R^?.$$

Schritt 2: $R^? \subseteq R \cup \Delta_X$: Da die Relation $R \cup \Delta_X$ eine reflexive Oberrelation von R ist, kommt sie als eine der Mengen S im Durchschnitt in (5.9a) vor. Daher gilt:

$$R^? = \bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\} \subseteq R \cup \Delta_x.$$

Der Beweis der weiteren Aussagen läuft nach demselben Prinzip. Wir geben daher nur die wesentlichen Argumente an.

Wir müssen uns für den Beweis von (5.11b) davon überzeugen, dass jede symmetrische Oberrelation von R notwendigerweise auch R^{-1} enthält (für das Analogon von Schritt 1) und dass andererseits $R \cup R^{-1}$ selbst symmetrisch ist (für das Analogon von Schritt 2).

Für den Beweis von (5.11c) verwenden wir, dass jede transitive Oberrelation von R notwendigerweise $R \circ R$ und damit auch alle Potenzen R^n mit $n \in \mathbb{N}$ enthält und dass andererseits die Vereinigung aller Potenzen $\bigcup_{n \in \mathbb{N}} R^n$ selbst transitiv ist (**Quizfrage 5.10:** klar?).

Für den Beweis von (5.11d) verwenden wir, dass jede reflexiv-transitive Oberrelation von R notwendigerweise auch alle Potenzen R^n mit $n \in \mathbb{N}_0$ enthält und dass andererseits die Vereinigung aller Potenzen $\bigcup_{n \in \mathbb{N}_0} R^n$ selbst reflexiv und transitiv ist.

Schließlich verwenden wir für den Beweis von (5.11e), dass jede reflexiv-symmetrisch-transitive Oberrelation von R notwendigerweise auch alle Potenzen $(R \cup R^{-1})^n$ mit $n \in \mathbb{N}_0$ enthält und dass andererseits die Vereinigung aller Potenzen $\bigcup_{n \in \mathbb{N}_0} (R \cup R^{-1})^n$ selbst reflexiv, symmetrisch und transitiv ist. □

Das folgende Resultat zeigt, dass die reflexiv-transitive Hülle übereinstimmt mit der reflexiven Hülle der transitiven Hülle und auch mit der transitiven Hülle der reflexiven Hülle. Das ist keinesfalls selbstverständlich und gilt für allgemeine mehrfache Hüllen nicht!

Folgerung 5.18 (Sequenzielle Hüllenbildung)**.**
Es sei X eine Menge und R eine Relation auf X. Dann gilt für die reflexiv-transitive Hülle:

$$R^* = (R^+)^? = (R^?)^+. \tag{5.12}$$

Beweis. Unter Benutzung von (5.11) folgt

$$(R^+)^? = (R^+) \cup \Delta_X = \Big(\bigcup_{n\in\mathbb{N}} R^n\Big) \cup \Delta_X = \bigcup_{n\in\mathbb{N}_0} R^n = R^*.$$

Andererseits bekommen wir mit Hilfe des Distributivgesetzes (5.5) für $\cup$ und $\circ$ die Aussage

$$\begin{aligned}(R \cup \Delta_X)^2 &= (R \cup \Delta_X) \circ (R \cup \Delta_X)\\ &= R^2 \;\cup\; (\Delta_X \circ R) \;\cup\; (R \circ \Delta_X) \;\cup\; (\Delta_X \circ \Delta_X)\\ &= R^2 \cup R \cup \Delta_X\end{aligned}$$

und analog (per Induktion)

$$(R \cup \Delta_X)^n = \bigcup_{k=0}^{n} R^k$$

für alle $n \in \mathbb{N}$. Daher folgt

$$(R^?)^+ = (R \cup \Delta_X)^+ = \bigcup_{n\in\mathbb{N}} (R \cup \Delta_X)^n = \bigcup_{n\in\mathbb{N}} \bigcup_{k=0}^{n} R^k = \bigcup_{n\in\mathbb{N}_0} R^n = R^*. \qquad \square$$

Ende der Vorlesung 4

Ende der Woche 2

§ 5.1 Äquivalenzrelationen

Definition 5.19 (Äquivalenzrelation)**.**
Es sei X eine Menge. Eine Relation R auf X heißt eine **Äquivalenzrelation** (englisch: equivalence relation) auf X, wenn sie

reflexiv, symmetrisch und transitiv

ist. Elemente $x, y \in X$, die $x \mathrel{R} y$ erfüllen, heißen **(zueinander) äquivalent** (englisch: equivalent, lateinisch: aequivalens: gleichwertig). △

Äquivalenzrelationen werden oft mit Symbolen wie $=$, $\sim$ oder $\equiv$ notiert. Unter Verwendung der Notation $\sim$ können wir für eine Äquivalenzrelation auf X also festhalten, dass für alle $x, y, z \in X$ gilt:

$$x \sim x \quad \text{Reflexivität,} \tag{5.13a}$$

$$x \sim y \;\Rightarrow\; y \sim x \quad \text{Symmetrie,} \tag{5.13b}$$

$$x \sim y \;\text{ und }\; y \sim z \;\Rightarrow\; x \sim z \quad \text{Transitivität.} \tag{5.13c}$$

Die Idee von Äquivalenzrelationen ist es, die Elemente einer Menge, die eine bestimmte Eigenschaft gemeinsam haben, zusammenzugruppieren und als gleichwertig zu betrachten. Wir erhalten dadurch eine „gröbere Version“ der Menge.

Wir können jede Relation R durch Übergang zu ihrer reflexiv-symmetrisch-transitiven Hülle $R^\sim$, siehe (5.11e), zu einer Äquivalenzrelation „aufwerten“. Daher sprechen wir bei $R^\sim$ auch von der **Äquivalenzhülle** (englisch: equivalence hull).

Beispiel 5.20 (Äquivalenzrelationen)**.**

(i) Die Identitätsrelation id_X ist eine Äquivalenzrelation auf jeder Menge X.

(ii) Die universelle Relation U_X ist eine Äquivalenzrelation auf jeder Menge X.

(iii) Es sei $m \in \mathbb{N}$ fest gewählt. Auf der Menge $\mathbb{Z}$ ist durch

$$x \overset{m}{\equiv} y \quad \Leftrightarrow \quad \exists n \in \mathbb{Z}\, (x - y = n\, m) \tag{5.14}$$

eine Äquivalenzrelation erklärt (**Quizfrage 5.11:** Details?). Anders ausgedrückt, x und y unterscheiden sich nur um ein Vielfaches von m, also, $m \mid (x - y)$. Diese Relation heißt die **Kongruenzrelation modulo** m (englisch: congruence relation modulo m, lateinisch: congruere: zusammentreffen, übereinstimmen) auf $\mathbb{Z}$.[59] △

Definition 5.21 (Äquivalenzklasse, Repräsentant, Repräsentantensystem)**.**
Es sei X eine Menge mit der Äquivalenzrelation $\sim$.

(i) Für $x \in X$ heißt die Menge

$$[x] \coloneqq \{y \in X \mid y \sim x\} \tag{5.15}$$

die **Äquivalenzklasse** (englisch: equivalence class) von x bzgl. $\sim$.[60]

(ii) Statt $[x]$ schreibt man manchmal auch $[x]_\sim$ oder auch $x\,/\sim$, um die Äquivalenzrelation zu betonen.

(iii) Jedes Element einer Äquivalenzklasse heißt ein **Repräsentant** (englisch: representative, lateinisch: repraesentare: darstellen) dieser Äquivalenzklasse.

(iv) Eine Menge $S \subseteq X$, die aus jeder Äquivalenzklasse genau einen Repräsentanten enthält, heißt ein **Repräsentantensystem** (englisch: system of representatives) von $\sim$. △

Beispiel 5.22 (Äquivalenzklasse, Repräsentant)**.**

(i) Wir betrachten eine beliebige Menge X mit der Identitätsrelation. Dann gilt $[x] = \{x\}$ für alle $x \in X$. Jede Äquivalenzklasse hat also nur ein Element und damit einen eindeutigen Repräsentanten. Das einzig mögliche Repräsentantensystem ist X selbst.

(ii) Wir betrachten eine beliebige Menge X mit der universellen Relation. Dann gilt $[x] = X$ für alle $x \in X$. Falls $X \neq \emptyset$ ist, dann gibt es also nur eine einzige Äquivalenzklasse, und diese enthält alle Elemente von X. In diesem Fall ist jede einelementige Teilmenge von X ein Repräsentantensystem.

(iii) Die Äquivalenzklassen der Kongruenzrelation modulo m ($m \in \mathbb{N}$) heißen auch die **Restklassen modulo** m (englisch: residue classes).[61] Die Restklasse von $a \in \mathbb{Z}$ modulo m ist also[62]

$$[a] = \{y \in \mathbb{Z} \mid y \overset{m}{\equiv} a\}$$

[59] Oft wird diese Relation statt $x \overset{m}{\equiv} y$ als $x \equiv y \pmod m$ geschrieben.
[60] Obwohl im Begriff **Äquivalenzklasse** das Wort „Klasse“ vorkommt, handelt es sich um eine Menge.
[61] Der Name leitet sich aus der Tatsache ab, dass die Elemente einer Restklasse durch die Eigenschaft charakterisiert sind, dass sie bei ganzzahliger Division durch m denselben Rest lassen.
[62] Die Notationen $m\mathbb{Z} = \{m\,z \mid z \in \mathbb{Z}\}$ („Zahl mal Menge“) und $a + m\mathbb{Z} = \{a + m\,z \mid z \in \mathbb{Z}\}$ („Zahl plus Zahl mal Menge“) werden später in Bemerkung 7.20 nochmal in einem allgemeineren Kontext erklärt.

$$
\begin{aligned}
&= \{y \in \mathbb{Z} \mid \exists n \in \mathbb{Z}\ (y - a = n\,m)\} \\
&= \{a + n\,m \mid n \in \mathbb{Z}\} \\
&= a + m\mathbb{Z}.
\end{aligned}
$$

Das Repräsentantensystem $\{0, 1, \ldots, m-1\}$ heißt das **natürliche Repräsentantensystem** (englisch: natural system of representatives) der Kongruenzrelation modulo m.

(iv) Speziell im Fall $m = 2$ gibt es genau zwei Äquivalenzklassen (Restklassen):

$$
\begin{aligned}
[0] &= \{y \in \mathbb{Z} \mid y \overset{2}{\equiv} 0\} \\
&= \{y \in \mathbb{Z} \mid \exists n \in \mathbb{Z}\ (y - 0 = 2\,n)\} \\
&= \{y \in \mathbb{Z} \mid y \text{ ist gerade}\},
\end{aligned}
$$

$$
\text{und} \quad
\begin{aligned}
[1] &= \{y \in \mathbb{Z} \mid y \overset{2}{\equiv} 1\} \\
&= \{y \in \mathbb{Z} \mid \exists n \in \mathbb{Z}\ (y - 1 = 2\,n)\} \\
&= \{y \in \mathbb{Z} \mid y \text{ ist ungerade}\}.
\end{aligned}
$$

Das natürliche Repräsentantensystem ist $\{0, 1\}$, ein anderes ist $\{-2, 43\}$. △

Satz 5.23 (Äquivalenzklassen sind entweder gleich oder disjunkt)**.**
Es sei X eine Menge mit der Äquivalenzrelation $\sim$. Weiter seien $[x]$ und $[y]$ zwei Äquivalenzklassen. Dann sind diese entweder gleich oder disjunkt.

Beweis. Nehmen wir an, $[x]$ und $[y]$ seien nicht disjunkt. Das heißt, sie haben ein Element $z \in X$ gemeinsam. Es sei nun $\widetilde{x}$ ein beliebiges Element aus $[x]$. Dann gilt

$$
\widetilde{x} \sim x \sim z.
$$

Wegen der Transitivität von $\sim$ ist also $\widetilde{x}$ äquivalent zu z, das nach Voraussetzung zu $[y]$ gehört. Damit haben wir $[x] \subseteq [y]$ gezeigt. Die umgekehrte Inklusion folgt analog. □

Definition 5.24 (Partition)**.**
Es sei X eine nichtleere Menge und $\mathcal{A}$ eine Menge von Teilmengen von X, also $\mathcal{A} \subseteq \mathcal{P}(X)$. $\mathcal{A}$ heißt eine **Partition** (englisch: partition) oder **disjunkte Zerlegung** (englisch: disjoint decomposition) von X, wenn gilt:

(i) Für alle $x \in X$ gibt es eine Menge $A \in \mathcal{A}$, die x enthält.

(ii) Für alle $A, B \in \mathcal{A}$ gilt, dass A und B entweder gleich sind oder disjunkt.

(iii) $\emptyset \notin \mathcal{A}$. △

Zu Eigenschaft (i) sagen wir auch, dass die Mengen in $\mathcal{A}$ die Menge X **überdecken** (englisch: to cover) oder eine **Überdeckung** (englisch: cover, covering) von X darstellen. Eigenschaft (ii) besagt, dass die Mengen in $\mathcal{A}$ paarweise disjunkt sind. Kurz gesagt ist eine Partition von X also die Darstellung von X als disjunkte Vereinigung nichtleerer Teilmengen.

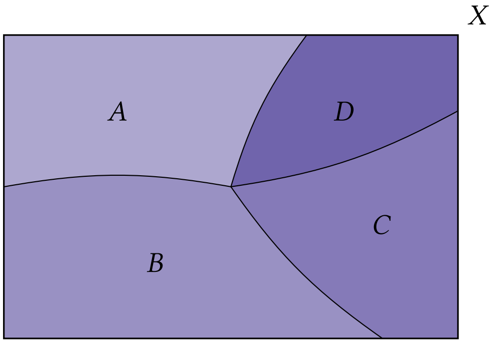


Abbildung 5.1.: Illustration der Partition einer Menge X in vier Teilmengen (Definition 5.24).

Satz 5.25 (Partitionen werden durch Äquivalenzrelationen erzeugt und umgekehrt)**.**
Es sei X eine nichtleere Menge.

(i) Ist $\sim$ eine Äquivalenzrelation, dann bildet die Menge der Äquivalenzklassen $\{[x] \mid x \in X\}$ eine Partition von X.

(ii) Ist $\mathcal{A}$ eine Partition von X, dann gibt es eine eindeutig bestimmte Äquivalenzrelation $\sim$, sodass $\mathcal{A}$ genau aus den Äquivalenzklassen von $\sim$ besteht.

Beweis. Aussage (i): Zur Abkürzung sei $\mathcal{A} \coloneqq \{[x] \mid x \in X\}$ die Menge der Äquivalenzklassen zur Äquivalenzrelation $\sim$. Wir weisen die drei Eigenschaften der Definition 5.24 nach. Zunächst ist jedes $x \in X$ Element seiner Äquivalenzklasse $[x]$, da ja $x \sim x$ gilt. Das zeigt Eigenschaft (i). Nach Satz 5.23 sind Äquivalenzklassen paarweise disunkt. Das zeigt Eigenschaft (ii). Schließlich sind Äquivalenzklassen nicht leer. Damit ist auch Eigenschaft (iii) gezeigt.

Der Beweis von Aussage (ii) ist Gegenstand der Übung. □

Beachte: Verschiedene Äquivalenzrelationen erzeugen verschiedene Partionen und umgekehrt. Etwas ungenau ausgedrückt sind also die Partitionen einer Menge X „dasselbe" wie die Äquivalenzrelationen auf X.[63]

Definition 5.26 (Faktormenge, Invarianz)**.**
Es sei X eine nichtleere Menge mit der Äquivalenzrelation $\sim$.

(i) Die Menge der Äquivalenzklassen

$$X/\sim \coloneqq \{[x] \mid x \in X\} \tag{5.16}$$

heißt auch die **Faktormenge** (englisch: factor set) oder die **Quotientenmenge** (englisch: quotient set) von X bzgl. $\sim$.

[63] Genauer werden wir mit Hilfe der Begriffe aus § 6 sagen können, dass die Menge aller Äquivalenzrelationen auf einer Menge X durch die Abbildung aus dem Beweis bijektiv auf die Menge aller Partitionen auf X abgebildet wird.

(*ii*) Eine Aussageform P auf X heißt **invariant** (englisch: invariant) oder **wohldefiniert** (englisch: well-defined) unter $\sim$, wenn $x \sim y$ impliziert, dass $P(x)$ und $P(y)$ denselben Wahrheitswert haben. △

Die Invarianz ist wichtig, wenn man eine Aussageform auf der Faktormenge dadurch definieren möchte, dass man sie auf den Elementen jeder Äquivalenzklasse definiert. Dabei ist sicherzustellen, dass sich tatsächlich für jeden Repräsentanten einer Äquivalenzklasse derselbe Wahrheitswert ergibt.

Beispiel 5.27 (wohldefinierte Aussageformen)**.**

(*i*) Die Kongruenzrelation $\overset{2}{\equiv}$ partitioniert die ganzen Zahlen in die beiden Äquivalenzklassen (auch Restklassen genannt) $[0]$ und $[1]$. Die Faktormenge ist also $\mathbb{Z} / \overset{2}{\equiv} = \{[0], [1]\}$.

Die Restklassen bestehen jeweils aus den geraden bzw. den ungeraden ganzen Zahlen. Die Aussageform „x ist eine gerade ganze Zahl" auf $\mathbb{Z}$ ist also unter der Kongruenzrelation $\overset{2}{\equiv}$ wohldefiniert, da sie für alle Repräsentanten der Äquivalenzklasse $[0]$ wahr und für alle Repräsentanten der Äquivalenzklasse $[1]$ falsch ist.

(*ii*) Dieselbe Aussageform auf $\mathbb{Z}$ ist jedoch unter der Kongruenzrelation $\overset{3}{\equiv}$ nicht wohldefiniert, da die Restklassen $[0]$, $[1]$ und $[2]$ jeweils sowohl gerade als auch ungerade ganze Zahlen enthalten. △

Die Menge der rationalen Zahlen wurde in (4.1) vorläufig als

$$\widetilde{\mathbb{Q}} \coloneqq \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\} \right\}$$

eingeführt. Darin sind aber beispielsweise $\frac{1}{2}$, $\frac{3}{6}$ und $\frac{-2}{-4}$ unterschiedliche Elemente, die wir jedoch miteinander identifizieren wollen. Zu diesem Zweck verwenden wir die Äquivalenzrelation

$$\frac{m_1}{n_1} \sim \frac{m_2}{n_2} \quad \Leftrightarrow \quad m_1 \cdot n_2 = m_2 \cdot n_1 \tag{5.17}$$

auf $\widetilde{\mathbb{Q}}$. Das führt uns zur Definition

$$\mathbb{Q} \coloneqq \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\} \right\} \Big/ \sim \tag{5.18}$$

für die **rationalen Zahlen**. Statt der Äquivalenzklasse $\left[\frac{m}{n}\right]$ schreiben wir üblicherweise weiter $\frac{m}{n}$, wir arbeiten also immer mit Repräsentanten. Das erklärt auch die übliche Notation $\frac{1}{2} = \frac{3}{6} = \frac{-2}{-4}$ anstelle von $\frac{1}{2} \sim \frac{3}{6} \sim \frac{-2}{-4}$.

§ 5.2 Ordnungsrelationen

Definition 5.28 (Ordnungsrelation)**.**
Es sei X eine Menge.

(*i*) Eine Relation R auf X heißt eine **Ordnungsrelation**, **Halbordnung** oder **partielle Ordnung** (englisch: order relation, partial order) auf X, wenn sie

reflexiv, antisymmetrisch und transitiv

ist. Das Paar (X, R) heißt dann eine **(halb-)geordnete Menge** oder **partiell geordnete Menge** (englisch: partially ordered set, poset).

(*ii*) Zwei Elemente $x, y \in X$ heißen **vergleichbar** (englisch: comparable), wenn $x \, R \, y$ oder $y \, R \, x$ gilt.

(*iii*) Ist eine Halbordnung R zusätzlich total, gilt also für beliebige $x, y \in X$ stets $x \, R \, y$ oder $y \, R \, x$, dann heißt sie eine **Totalordnung** (englisch: total order) oder eine **lineare Ordnung** (englisch: linear order). Das Paar (X, R) heißt dann eine **totalgeordnete Menge** (englisch: totally ordered set). △

Ordnungsrelationen werden oft mit Symbolen wie $\le$, $\preccurlyeq$ oder $\subseteq$ notiert. Unter Verwendung der Notation $\preccurlyeq$ können wir für eine Ordnungsrelation auf X also festhalten, dass für alle $x, y, z \in X$ gilt:

$$x \preccurlyeq x \quad \text{Reflexivität,} \tag{5.19a}$$

$$x \preccurlyeq y \quad \text{und} \quad y \preccurlyeq x \quad \Rightarrow \quad x = y \quad \text{Antisymmetrie,} \tag{5.19b}$$

$$x \preccurlyeq y \quad \text{und} \quad y \preccurlyeq z \quad \Rightarrow \quad x \preccurlyeq z \quad \text{Transitivität.} \tag{5.19c}$$

Die Idee von Ordnungsrelationen ist es, Elemente einer Menge bezüglich einer bestimmten Eigenschaft zu vergleichen. Bei einer Totalordnung ist dabei jedes Element mit jedem Element vergleichbar, bei einer Halbordnung nicht notwendigerweise.

Quizfrage 5.12: Lässt sich jede Relation auf einer Menge X durch eine geeignete Hüllenbildung zu einer Halbordnung machen?

Beispiel 5.29 (Halbordnungen und Totalordnungen)**.**

(*i*) Die Identitätsrelation id_X ist eine Halbordnung auf jeder Menge X.

(*ii*) Die universelle Relation U_X ist *keine* Halbordnung auf jeder Menge X, die mindestens zwei Elemente enthält, da sie dann nicht antisymmetrisch ist.

(*iii*) Die Kleiner-Gleich-Relation $\le$ ist eine Totalordnung auf jeder Teilmenge von $\mathbb{R}$.

(*iv*) Die Teilmengenrelation $\subseteq$ ist eine Halbordnung auf der Potenzmenge $\mathcal{P}(X)$ jeder beliebigen Menge X. Die Teilmengenrelation wird auch **Inklusionshalbordnung** (englisch: inclusion order) genannt. Sie ist eine totale Ordnung dann und nur dann, wenn X entweder kein oder genau ein Element enthält.

(*v*) Die Teilbarkeitsrelation $\mid$ ist eine Halbordnung auf $\mathbb{N}$, nicht aber auf $\mathbb{Z}$. (**Quizfrage 5.13:** Warum nicht?) △

Lemma 5.30 (Inverse von Ordnungsrelationen sind Ordnungsrelationen)**.**
Es sei $\preccurlyeq$ eine Halbordnung auf einer Menge X. Dann ist auch die inverse Relation $\succcurlyeq$ eine Halbordnung auf X. Ist $\preccurlyeq$ eine Totalordnung, dann auch $\succcurlyeq$.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

Wir klären jetzt, welche Art der Relation die **gewöhnliche Echt-Kleiner-Relation**, z. B. auf der Menge $\mathbb{R}$ der reellen Zahlen, ist.

Definition 5.31 (strenge Ordnungsrelation, vgl. Definition 5.28)**.**
Es sei X eine Menge.

(i) Eine Relation R auf X heißt eine **strenge Ordnungsrelation**, **strenge Halbordnung** oder **strenge partielle Ordnung** (englisch: strict order relation, strict partial order) auf X, wenn sie

irreflexiv und transitiv

ist. Das Paar (X, R) heißt dann eine **streng halbgeordnete Menge** (englisch: strictly partially ordered set).

(ii) In einer strengen Ordnungsrelation R auf X nennen wir zwei Elemente $x, y \in X$ **vergleichbar**, wenn $x\,R\,y$ oder $y\,R\,x$ oder $x = y$ gilt.

(iii) Sind in einer strengen Halbordnung R beliebige $x, y \in X$ stets vergleichbar, dann heißt sie eine **strenge Totalordnung** (englisch: strict total order). Das Paar (X, R) heißt dann eine **streng totalgeordnete Menge** (englisch: strictly totally ordered set). △

Strenge Ordnungsrelationen werden oft mit Symbolen wie $<$, $\lneq$, $\prec$, $\precneqq$ oder $\subsetneq$ notiert. Unter Verwendung der Notation $\prec$ können wir für eine strenge Ordnungsrelation auf X also festhalten, dass für alle $x, y, z \in X$ gilt:

$$x \not\prec x \quad \text{Irreflexivität,} \tag{5.20a}$$

$$x \prec y \quad \text{und} \quad y \prec z \quad \Rightarrow \quad x \prec z \quad \text{Transitivität.} \tag{5.20b}$$

Beachte: Strenge Ordnungsrelationen sind auch antisymmetrisch, denn die Annahme $x \prec y$ und $y \prec x$ impliziert $x \prec x$, was der Irreflexivität widerspricht. Es kann also nicht vorkommen, dass $x \prec y$ und $y \prec x$ gleichzeitig gelten, damit ist die Relation $\prec$ antisymmetrisch.

Das folgende Resultat besagt, dass es zu jeder Ordnungsrelation eine zugehörige strenge Ordnungsrelation gibt und umgekehrt.

Lemma 5.32 (Beziehungen zwischen Ordnungsrelationen und strengen Ordnungsrelationen)**.**
Es sei X eine Menge.

(i) Ist $\preccurlyeq$ eine Ordnungsrelation auf X, dann ist $\prec \coloneqq \preccurlyeq \setminus \Delta_X$ eine strenge Ordnungsrelation auf X, genannt die **zugehörige strenge Ordnungsrelation** (englisch: associated strict order relation). Wenn $\preccurlyeq$ eine Totalordnung ist, dann ist $\prec$ eine strenge Totalordnung.

(ii) Ist $\prec$ eine strenge Ordnungsrelation auf X, dann ist $\preccurlyeq \coloneqq \prec \cup \Delta_X$ eine Ordnungsrelation auf X, genannt die **zugehörige Ordnungsrelation** (englisch: associated order relation). Wenn $\prec$ eine strenge Totalordnung ist, dann ist $\preccurlyeq$ eine Totalordnung.

Beachte: Für die so definierten Relationen $\preccurlyeq$ und $\prec$ gilt also $x \prec y \Leftrightarrow (x \preccurlyeq y) \wedge (x \neq y)$.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

Bemerkung 5.33 (Überdeckungsrelation, Hasse-Diagramm)**.**

(*i*) Es sei $\preccurlyeq$ eine Halbordnung auf einer Menge X und $\prec$ die zugehörige strenge Halbordnung. Dann heißt die Teilrelation $\prec\!\!\cdot$ von $\prec$, definiert durch

$$x \prec\!\!\cdot\; y \quad \Leftrightarrow \quad \nexists z \in X \; (x \prec z \prec y) \tag{5.21}$$

die zu $\preccurlyeq$ gehörige **Überdeckungsrelation** (englisch: covering relation). Gilt $x \prec\!\!\cdot\; y$, dann sagen wir auch, x sei ein **direkter Vorgänger** (englisch: immediate predecessor) von y und y sei ein **direkter Nachfolger** (englisch: immediate successor) von x.

(*ii*) Das **Hasse-Diagramm** (englisch: Hasse diagram) oder (englisch: order diagram) der Halbordnung $\preccurlyeq$ ist ein gerichteter Graph, dessen Knoten die Elemente von X sind. Die Kante von x nach y wird genau dann eingeführt, wenn $x \prec\!\!\cdot\; y$ gilt. Das Hasse-Diagramm bildet also die Relation „direkter Nachfolger von“ ab.

(*iii*) Ist X eine endliche Menge[64], dann kann $\preccurlyeq$ als die reflexiv-transitive Hülle von $\prec\!\!\cdot$ wiedergewonnen werden: $\preccurlyeq \; = \; \prec\!\!\cdot^*$.[65] △

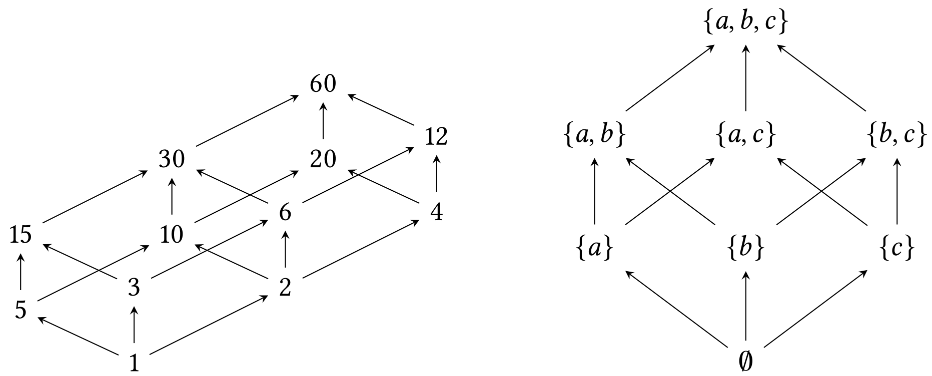


Abbildung 5.2.: Links: Hasse-Diagramm (Bemerkung 5.33) der Halbordnung | (Teilbarkeitsrelation) auf der Menge der Teiler von 60 in $\mathbb{N}$. Rechts: Hasse-Diagramm der Halbordnung $\subseteq$ (Inklusionshalbordnung) auf der Potenzmenge von $\{a, b, c\}$ mit paarweise verschiedenen Elementen. Die Pfeile der Kanten können hier auch weggelassen werden, weil sie sich aus der Anordnung der Knoten (von unten nach oben) ergeben.

Definition 5.34 (obere und untere Schranken, Supremum und Infimum, maximale und minimale Elemente, Maximum und Minimum)**.**

Es sei X mit der Relation $\preccurlyeq$ eine halbgeordnete Menge.

[64] Der Begriff der endlichen Menge wird in Definition 6.32 formal eingeführt.

[65] Für unendliche partiell oder totalgeordnete Mengen gilt zwar immer noch $\prec\!\!\cdot^* \; \subseteq \; \preccurlyeq$, aber möglicherweise gewinnen wir nicht die gesamte Ordnungsrelation $\preccurlyeq$ zurück. Die Überdeckungsrelation $\prec\!\!\cdot$ kann sogar leer sein, z. B. ist das für die gewöhnliche Kleiner-Gleich-Relation auf $\mathbb{R}$ der Fall.

(*i*) $b \in X$ heißt eine **obere Schranke** (englisch: upper bound) von $A \subseteq X$, wenn gilt:

$$x \preccurlyeq b \quad \text{für alle } x \in A. \qquad \text{(„Ganz } A \text{ ist } \preccurlyeq b.\text{")}$$

(*ii*) Wenn $A \subseteq X$ eine obere Schranke besitzt, so heißt A **nach oben beschränkt** (englisch: bounded above), ansonsten **nach oben unbeschränkt** (englisch: unbounded above).

(*iii*) $b \in X$ heißt das **Supremum** (englisch: supremum, lateinisch: supremum: das Größte) oder die **kleinste obere Schranke** (englisch: least upper bound) von $A \subseteq X$, wenn gilt:

$$\begin{aligned} &b \text{ ist eine obere Schranke von } A, \\ &\text{und für jede obere Schranke } \widehat{b} \text{ von } A \text{ gilt:} \quad b \preccurlyeq \widehat{b}. \end{aligned}$$

In diesem Fall schreiben wir $b = \sup A$.

(*iv*) $b \in X$ heißt das **Maximum** (englisch: maximum) von $A \subseteq X$, wenn b eine obere Schranke von A ist, die zu A gehört, wenn also gilt:

$$b \in A \quad \text{und} \quad x \preccurlyeq b \quad \text{für alle } x \in A.$$

In diesem Fall schreiben wir $b = \max A$.

(*v*) $b \in X$ heißt ein **maximales Element** (englisch: maximal element) von $A \subseteq X$, wenn gilt:

$$b \in A, \quad \text{und für alle } x \in A \text{ gilt:} \quad b \preccurlyeq x \;\Rightarrow\; x = b.$$

(„Kein Element von A ist echt größer als b.")

(*vi*) $a \in X$ heißt eine **untere Schranke** (englisch: lower bound) von $A \subseteq X$, wenn gilt:

$$a \preccurlyeq x \quad \text{für alle } x \in A. \qquad \text{(„Ganz } A \text{ ist } \succcurlyeq a.\text{")}$$

(*vii*) Wenn $A \subseteq X$ eine untere Schranke besitzt, so heißt A **nach unten beschränkt** (englisch: bounded below), ansonsten **nach unten unbeschränkt** (englisch: unbounded below).

(*viii*) $a \in X$ heißt das **Infimum** (englisch: infimum, lateinisch: infimum: das Kleinste) oder die **größte untere Schranke** (englisch: greatest lower bound) von $A \subseteq X$, wenn gilt:

$$\begin{aligned} &a \text{ ist eine untere Schranke von } A, \\ &\text{und für jede untere Schranke } \widehat{a} \text{ von } A \text{ gilt:} \quad \widehat{a} \preccurlyeq a. \end{aligned}$$

In diesem Fall schreiben wir $a = \inf A$.

(*ix*) $a \in X$ heißt das **Minimum** (englisch: minimum) von $A \subseteq X$, wenn a eine untere Schranke von A ist, die zu A gehört, wenn also gilt:

$$a \in A \quad \text{und} \quad a \preccurlyeq x \quad \text{für alle } x \in A.$$

In diesem Fall schreiben wir $a = \min A$.

(*x*) $a \in X$ heißt ein **minimales Element** (englisch: minimal element) von $A \subseteq X$, wenn gilt:

$$a \in A, \quad \text{und für alle } x \in A \text{ gilt:} \quad x \preccurlyeq a \;\Rightarrow\; x = a.$$

(„Kein Element von A ist echt kleiner als a.")

△

Beachte: Sind b_1, b_2 zwei maximale Elemente von A, dann sind sie entweder gleich oder nicht vergleichbar, denn: Aus der Vergleichbarkeit, etwa $b_1 \preccurlyeq b_2$, folgt nach Definition bereits $b_2 = b_1$. Eine analoge Aussage gilt für minimale Elemente.

Beachte: Die Bezeichnung **kleinste obere Schranke** für das **Supremum** ist gerechtfertigt, weil das Supremum einer Menge A (sofern es existiert) laut Definition 5.34 tatsächlich das Minimum der Menge der oberen Schranken von A ist. Eine analoge Aussage gilt für die größte untere Schranke (das Infimum).

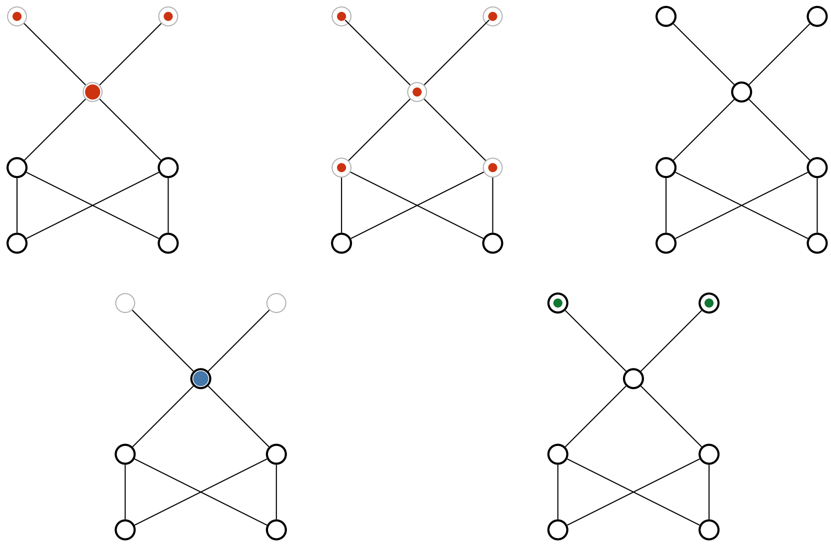


Abbildung 5.3.: Illustration der Begriffe aus Definition 5.34 für verschiedene Teilmengen $A \subseteq X$ einer halbgeordneten Menge X mit Hilfe ihres Hasse-Diagramms. Elemente von A sind als ○ gekennzeichnet, Elemente von $X \setminus A$ als ○. Oben links: Teilmenge A mit mehreren oberen Schranken (•), von denen eine (●) die kleinste (das Supremum von A) ist. Oben Mitte: Teilmenge A mit mehreren oberen Schranken ohne Supremum. Oben rechts: Die gesamte Menge X besitzt in diesem Beispiel keine oberen Schranken. Unten links: Teilmenge A mit einem Maximum (●), das gleichzeitig das einzige maximale Element von A ist. Unten rechts: Teilmenge A mit mehreren maximalen Elementen (•).

Wir zeigen nun einige ausgewählte Eigenschaften, die insbesondere die Sprechweisen „das" Supremum, „das" Maximum, „das" Infimum und „das" Minimum rechtfertigen.

Lemma 5.35 (Eigenschaften und Beziehungen zwischen Supremum und Maximum, Infimum und Minimum)**.**
Es sei X mit der Relation $\preccurlyeq$ eine halbgeordnete Menge und $A \subseteq X$.

(i) Existiert ein Supremum von A, so ist dieses eindeutig.

(ii) Existiert ein Maximum von A, so ist dieses eindeutig. Es ist dann auch das einzige maximale Element von A.

(*iii*) Ist b das Maximum von A, so ist b gleichzeitig das Supremum von A.

(*iv*) Hat A ein Supremum b, so gilt: Gehört b zu A, so ist b das Maximum von A. Gehört b nicht zu A, so besitzt A kein Maximum.

(*v*) Ist $\preccurlyeq$ eine Totalordnung auf A, dann gilt: Ist b ein maximales Element von A, so ist b auch das Maximum von A.

Analoge Aussagen gelten auch für das Infinum und Minimum von A.

Beweis. Aussage (*i*): Wir nehmen an, $b \in X$ und $\widehat{b} \in X$ seien beides Suprema von A. Dann sind b und $\widehat{b}$ beides obere Schranken. Da b ein Supremum von A ist, gilt $b \preccurlyeq \widehat{b}$. Da $\widehat{b}$ ein Supremum von A ist, gilt $\widehat{b} \preccurlyeq b$. Aufgrund der Antisymmetrie von $\preccurlyeq$ folgt nun $b = \widehat{b}$.

Aussage (*ii*): Wir nehmen an, $b_1 \in X$ und $b_2 \in X$ seien beides Maxima von A. Dann gehören b_1 und b_2 beide zu A. Da b_1 ein Maximum von A ist, gilt $b_2 \preccurlyeq b_1$, und da auch b_2 ein Maximum von A ist, gilt ebenfalls $b_1 \preccurlyeq b_2$. Aufgrund der Antisymmetrie von $\preccurlyeq$ folgt nun $b_1 = b_2$.

Das Maximum b_1 von A ist außerdem ein maximales Element von A, denn für beliebiges $x \in A$ gilt $x \preccurlyeq b_1$, sodass zusätzlich $b_1 \preccurlyeq x$ dann aufgrund der Antisymmetrie $x = b_1$ impliziert.

Wäre nun $b_2 \in A$ ein anderes maximales Element von A, dann wäre $b_2 \preccurlyeq b_1$ (aufgrund der Maximum-Eigenschaft von b_1), was der Nicht-Vergleichbarkeit verschiedener maximaler Elemente widerspricht.

Aussage (*iii*): Es sei b_1 das Maximum von A. Es gilt also $b_1 \in A$ und $x \preccurlyeq b_1$ für alle $x \in A$. Das heißt aber, dass b_1 eine obere Schranke von A ist. Ist nun b_2 eine weitere obere Schranke von A, dann gilt $x \preccurlyeq b_2$ für alle $x \in A$, insbesondere $b_1 \preccurlyeq b_2$. Das zeigt, dass b_1 das Supremum von A ist.

Aussage (*iv*): Es sei b das Supremum von A. Insbesondere ist b eine obere Schranke von A, es gilt also $x \preccurlyeq b$ für alle $x \in A$. Falls nun b zu A gehört, dann ist b per Definition das Maximum von A. Falls jedoch b nicht zu A gehört, so ist b per Definition kein Maximum von A. Ein Maximum von A kann auch nicht existieren, sonst wäre es nach Aussage (*iii*) gleichzeitig das Supremum, also gleich b.

Aussage (*v*): Es sei nun $\preccurlyeq$ eine Totalordnung auf X und b ein maximales Element von A. Für beliebiges $x \in A$ sind x und b vergleichbar. Die Annahme $b \prec x$ führt wegen der Eigenschaft, dass b maximales Element von A ist, zu dem Widerspruch $x = b$. Es muss also $x \preccurlyeq b$ gelten, was b als Maximum von A bestätigt. □

Beispiel 5.36 (Schranken, extremale Elemente, Maxima und Minima, Suprema und Infima)**.**

(*i*) In den natürlichen Zahlen $\mathbb{N}$ mit der gewöhnlichen Totalordnung $\leqslant$ ist die Zahl 1 das Minimum und damit das Infimum der Menge $\mathbb{N}$. Die Menge $\mathbb{N}$ besitzt keine obere Schranke, also auch kein Supremum und kein Maximum.

(*ii*) Es sei X eine beliebige nichtleere Menge. In der Potenzmenge $\mathcal{P}(X)$ mit der Halbordnung $\subseteq$ ist $\emptyset$ das Minimum von $\mathcal{P}(X)$ und X das Maximum von $\mathcal{P}(X)$.

Hat X mindestens zwei Elemente, dann besitzt die Teilmenge $P = \mathcal{P}(X) \setminus \{\emptyset\}$ das Infimum $\emptyset$, aber kein Minimum. Die minimalen Elemente von P sind genau die einelementigen Teilmengen von X. △

Quizfrage 5.14: Können Sie sich eine Menge mit einer Halbordnung oder einer totalen Ordnung vorstellen, die kein maximales Element besitzt?

Quizfrage 5.15: Können Sie sich eine Menge mit einer Halbordnung vorstellen, die ein eindeutiges maximales Element besitzt, das aber kein Maximum ist?

Expertenwissen: partielle Ordnung auf der Menge der Äquivalenzrelationen

Es seien X eine Menge und $\sim_1$ und $\sim_2$ zwei Äquivalenzrelationen auf X. Die Äquivalenzrelation $\sim_1$ heißt **feiner** (englisch: finer) als $\sim_2$, und $\sim_2$ heißt **gröber** (englisch: coarser) als $\sim_1$, wenn für alle $x, y \in X$ gilt:

$$x \sim_1 y \quad \Rightarrow \quad x \sim_2 y.$$

Mit anderen Worten: Jede Äquivalenzklasse von $\sim_1$ ist in einer Äquivalenzklasse von $\sim_2$ enthalten: $[x]_{\sim_1} \subseteq [x]_{\sim_2}$ für alle $x \in X$.

In dem Fall heißt $\sim_1$ auch eine **Verfeinerung** (englisch: refinement) von $\sim_2$, und $\sim_2$ heißt eine **Vergröberung** (englisch: coarsening) von $\sim_1$. Wir schreiben auch $\sim_1 \preccurlyeq \sim_2$.

Die Relation $\preccurlyeq$ ist eine partielle Ordnung auf der Menge aller Äquivalenzrelationen auf X. In dieser Ordnungsrelation ist die universelle Relation $X \times X$ das Maximum, und die Diagonale Δ_X ist das Minimum.

Ende der Vorlesung 5

§ 6 Abbildungen

Literatur: Deiser, 2024a, Kapitel 1.3; Deiser, 2024b, Kapitel 1.4; Jänich, 2008, Kapitel 1.2

In diesem Abschnitt geht es um den grundlegenden Begriff der Abbildung oder Funktion. Eine Abbildung ist dabei nichts anderes als eine spezielle Relation.

Definition 6.1 (weitere Eigenschaften von Relationen)**.**
Es seien X und Y Mengen. Eine Relation (R, X, Y) zwischen X und Y heißt

(i) **linkstotal** (englisch: left-total), falls für alle $x \in X$ ein $y \in Y$ existiert, sodass $x \mathrel{R} y$ gilt.

(ii) **rechtseindeutig** (englisch: right-unique), falls für alle $x \in X$ und alle $y_1, y_2 \in Y$ gilt: $(x \mathrel{R} y_1) \wedge (x \mathrel{R} y_2) \Rightarrow y_1 = y_2$. △

Definition 6.2 (Funktion)**.**
Es seien X und Y Mengen. Eine linkstotale und rechtseindeutige Relation (f, X, Y) zwischen X und Y heißt **Abbildung** (englisch: map) oder **Funktion** (englisch: function) **von** X **in** Y oder **auf** X **mit Werten in** Y. Die Menge X heißt der **Definitionsbereich** (englisch: domain) oder die **Definitionsmenge** und die Menge Y der **Zielmenge** (englisch: codomain) von f. Ist $Y = X$, so spricht man auch von einer Funktion von X **in sich**. △

Den Sachverhalt, dass f eine Funktion von X in Y ist, drücken wir kurz in der Form

$$f\colon X \to Y \quad \text{oder} \quad X \xrightarrow{f} Y \quad \text{oder} \quad Y \xleftarrow{f} X$$

aus. Zu gegebenem $x \in X$ heißt das eindeutige $y \in Y$ mit $x\,f\,y$ das **Bild** (englisch: image) oder der **Funktionswert** (englisch: function value) **von f an der Stelle x** oder **von x unter f**. Dieser wird auch mit $f(x)$ bezeichnet. Wir sagen auch, f **bilde** x auf $f(x)$ **ab** (englisch: f maps x to $f(x)$) und schreiben dafür $x \mapsto f(x)$. Für die Definition einer Funktion sind daher auch die kompakten Schreibweisen

$$X \ni x \mapsto f(x) \in Y \quad \text{oder} \quad f\colon \begin{cases} X \to Y \\ x \mapsto f(x) \end{cases}$$

üblich.

Beachte: Zwei Funktionen sind genau dann gleich, wenn sie in ihren Definitionsbereichen, Zielmengen und ihren Abbildungsvorschriften übereinstimmen.

Die Menge

$$\{(x, f(x)) \mid x \in X\} \subseteq X \times Y \tag{6.1}$$

heißt der **Graph** (englisch: graph) der Funktion $f\colon X \to Y$.[66]

Beispiel 6.3 (Abbildungen).

(*i*) Es seien X und Y Mengen und $y_0 \in Y$. Dann heißt die Abbildung f mit

$$X \ni x \mapsto f(x) \coloneqq y_0 \in Y$$

die **konstante Funktion** (englisch: constant function) auf X mit dem Wert y_0.

(*ii*) Die Funktion

$$X \ni x \mapsto \mathrm{id}_X(x) \coloneqq x \in X$$

heißt die **Identität** (englisch: identity) oder **identische Abbildung** (englisch: identity map) **von X in sich**. Der Graph von id_X ist die Diagonale $\Delta_X = \{(x, y) \in X \times X \mid x = y\}$, siehe (5.1).

(*iii*) Ist die Definitionsmenge $X = \emptyset$ die leere Menge und Y eine beliebige Menge, dann existiert genau eine Funktion $f\colon \emptyset \to Y$, die **leere Funktion** (englisch: empty function).

(*iv*) Ist die Zielmenge $Y = \emptyset$ die leere Menge, dann existiert eine Funktion $f\colon X \to \emptyset$ genau dann, wenn auch X die leere Menge ist. △

Definition 6.4 (Bild, Einschränkung, Fortsetzung).
Es sei $f\colon X \to Y$ eine Funktion.[67]

[66] Der Begriff des Graphen einer Funktion stimmt also überein mit dem Begriff des Graphen der Funktion als Relation, vgl. Definition 5.1.
[67] Wir sagen damit insbesondere, dass X und Y Mengen sind.

(*i*) Für $A \subseteq X$ heißt

$$f(A) \coloneqq \{f(x) \mid x \in A\} \subseteq Y \tag{6.2}$$

die **Bildmenge** oder kurz das **Bild** (englisch: image) von f **auf** A oder das **Bild** von A **unter** f.

(*ii*) Ist $A \subseteq X$, dann heißt die Funktion $f|_A$

$$A \ni x \mapsto f|_A(x) \coloneqq f(x) \in Y$$

die **Einschränkung** oder **Restriktion** von f **auf** A.[68]

(*iii*) Gilt zusätzlich $f(A) \subseteq B$, so bezeichnen wir mit $f|_A^B$ die **Einschränkung** von f auf A, wobei zusätzlich die Zielmenge durch B ersetzt wird, also die Funktion

$$A \ni x \mapsto f|_A^B(x) \coloneqq f(x) \in B.$$

Gilt insbesondere $f(X) \subseteq B$, dann bezeichnet $f|^B$ die Funktion

$$X \ni x \mapsto f|^B(x) \coloneqq f(x) \in B,$$

bei der gegenüber f nur die Zielmenge ersetzt wird.

Beachte: Es wird nicht gefordert, dass $B \subseteq Y$ gilt.

(*iv*) Ist $C \supseteq X$ und $D \supseteq Y$, dann heißt eine Funktion $g\colon C \to D$, die auf X mit f übereinstimmt, für die also $g|_X^Y = f$ gilt, eine **Fortsetzung** (englisch: extension) von f. △

Beispiel 6.5 (Bild, Einschränkung, Fortsetzung).
Wir betrachten die Funktionen[69]

$$\begin{aligned}
\mathbb{R} \ni x \mapsto f(x) &\coloneqq \sin(x) \in \mathbb{R} && \text{mit dem Bild } [-1,1],\\
\mathbb{R} \ni x \mapsto g(x) &\coloneqq \sin(x) \in [-1,1] && \text{mit dem Bild } [-1,1],\\
\frac{\pi}{2}\mathbb{Z} \ni x \mapsto h(x) &\coloneqq \sin(x) \in [-1,1] && \text{mit dem Bild } \{-1,0,1\},\\
\frac{\pi}{2}\mathbb{Z} \ni x \mapsto i(x) &\coloneqq \sin(x) \in \{-1,0,1\} && \text{mit dem Bild } \{-1,0,1\}.
\end{aligned}$$

Dann sind g, h und i Einschränkungen von f, und f ist eine Fortsetzung von g, h und i. △

Definition 6.6 (Urbild).
Es sei $f\colon X \to Y$ eine Funktion.

(*i*) Für $B \subseteq Y$ heißt die Menge

$$f^{-1}(B) \coloneqq \{x \in X \mid f(x) \in B\} \tag{6.3}$$

die **Urbildmenge** oder das **Urbild** (englisch: pre-image) von B **unter** f.

[68] vgl. Definition 5.3.
[69] Hierbei bedeutet $\frac{\pi}{2}\mathbb{Z} = \{\frac{\pi}{2}z \mid z \in \mathbb{Z}\}$ („Zahl mal Menge“) die Menge der ganzzahligen Vielfachen von $\frac{\pi}{2}$. Diese Notation wird später in Bemerkung 7.20 nochmal in einem allgemeineren Kontext erklärt.

(*ii*) Ist $B \subseteq Y$ einelementig, also $B = \{y\}$ für ein $y \in Y$, dann heißt das Urbild

$$f^{-1}(\{y\}) \coloneqq \{x \in X \mid f(x) = y\} \tag{6.4}$$

auch die **Faser** (englisch: fiber) von y **unter** f. △

Beispiel 6.7 (Urbild)**.**
Wir betrachten die Funktion

$$\mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}.$$

Dann ist

$$f^{-1}(\{y\}) = \begin{cases} \{\sqrt{y}, -\sqrt{y}\} & \text{falls } y > 0, \\ \{0\} & \text{falls } y = 0, \\ \emptyset & \text{falls } y < 0. \end{cases}$$

△

Satz 6.8 (Bilder und Urbilder von Vereinigungen und Durchschnitten)**.**
Es sei $f\colon X \to Y$ eine Funktion. Weiter seien I und J irgendwelche Indexmengen und $\{X_i \mid i \in I\}$ eine Menge von Teilmengen von X sowie $\{Y_j \mid j \in J\}$ eine Menge von Teilmengen von Y. Dann gilt:

$$f\Big(\bigcup_{i \in I} X_i\Big) = \bigcup_{i \in I} f(X_i) \tag{6.5a}$$

$$f\Big(\bigcap_{i \in I} X_i\Big) \subseteq \bigcap_{i \in I} f(X_i) \tag{6.5b}$$

$$f^{-1}\Big(\bigcup_{j \in J} Y_j\Big) = \bigcup_{j \in J} f^{-1}(Y_j) \tag{6.5c}$$

$$f^{-1}\Big(\bigcap_{j \in J} Y_j\Big) = \bigcap_{j \in J} f^{-1}(Y_j). \tag{6.5d}$$

Beweis. Wir beweisen hier nur (6.5a) und (6.5c). Die Aussagen (6.5b) und (6.5d) sind Gegenstand der Übung.

Zum Beweis von (6.5a):

$$\begin{aligned} & y \in f\Big(\bigcup_{i \in I} X_i\Big) \\ \Leftrightarrow\ & \exists i \in I\ \exists x \in X_i\ (y = f(x)) && \text{nach Definition (4.4b) der Vereinigungsmenge} \\ \Leftrightarrow\ & \exists i \in I\ (y \in f(X_i)) && \text{nach Definition (6.2) des Bildes } f(X_i) \\ \Leftrightarrow\ & y \in \bigcup_{i \in I} f(X_i) && \text{nach Definition (4.4b) der Vereinigungsmenge.} \end{aligned}$$

Zum Beweis von (6.5c):

$$x \in f^{-1}\Big(\bigcup_{j \in J} Y_j\Big)$$

$$
\begin{aligned}
&\Leftrightarrow \quad \exists y \in \bigcup_{j\in J} Y_j \; \big(y = f(x)\big) && \text{nach Definition (6.3) des Urbildes}\\
&\Leftrightarrow \quad \exists j \in J \; \exists y \in Y_j \; \big(y = f(x)\big) && \text{nach Definition (4.4b) der Vereinigungsmenge}\\
&\Leftrightarrow \quad \exists j \in J \; \big(x \in f^{-1}(Y_j)\big) && \text{nach Definition (6.3) des Urbildes}\\
&\Leftrightarrow \quad x \in \bigcup_{j\in J} f^{-1}(Y_j) && \text{nach Definition (4.4b) der Vereinigungsmenge.} \qquad \square
\end{aligned}
$$

Beispiel 6.9 (Bilder und Urbilder von Vereinigungen und Durchschnitten)**.**
In (6.5b) gilt i. A. nicht die Gleichheit, wie folgendes Beispiel zeigt: Es sei

$$\mathbb{R} \ni x \mapsto y_0 \in \mathbb{R}$$

eine konstante Funktion. Für die disjunkten Mengen $X_1 = \{0\}$ und $X_2 = \{1\}$ gilt

$$
\begin{aligned}
f(X_1 \cap X_2) &= f(\emptyset) = \emptyset,\\
\text{aber} \quad f(X_1) \cap f(X_2) &= \{y_0\} \cap \{y_0\} = \{y_0\}. \qquad \triangle
\end{aligned}
$$

§ 6.1 Injektivität und Surjektivität

Definition 6.10 (Injektivität, Surjektivität, Bijektivität)**.**
Eine Funktion $f\colon X \to Y$ heißt

(i) **surjektiv** (englisch: surjective, onto, französisch: sur: auf, lateinisch: iacere: werfen), eine **Surjektion** (englisch: surjection) oder **rechtstotal** (englisch: right-total), wenn $f(X) = Y$ gilt.[70] Man sagt auch, f bilde X **auf** Y ab.

Äquivalent dazu sind:

- Für alle $y \in Y$ gibt es mindestens ein $x \in X$ mit der Eigenschaft $f(x) = y$.
- Jedes Element der Zielmenge Y hat ein nichtleeres Urbild:

$$\forall y \in Y \; \big(f^{-1}(\{y\}) \neq \emptyset\big).$$

(ii) **injektiv** (englisch: injective, one-to-one, lateinisch: in: hinein), eine **Injektion** (englisch: injection), eine **Einbettung** (englisch: embedding) oder **linkseindeutig** (englisch: left-unique), wenn für alle $x_1, x_2 \in X$ gilt: $f(x_1) = f(x_2) \;\Rightarrow\; x_1 = x_2$.[71]

Äquivalent dazu sind:

- Für alle $y \in Y$ gibt es höchstens ein $x \in X$ mit der Eigenschaft $f(x) = y$.
- Das Urbild jedes Elements der Zielmenge Y ist maximal einelementig:

$$\forall y \in Y \; \big(f^{-1}(\{y\}) \text{ hat kein oder genau ein Element}\big).$$

[70] Die Surjektivität von f wird manchmal auch durch die Schreibweise $f\colon X \twoheadrightarrow Y$ ausgedrückt.
[71] Die Injektivität von f wird manchmal auch durch die Schreibweise $f\colon X \rightarrowtail Y$ ausgedrückt.

(iii) **bijektiv** (englisch: bijective, lateinisch: bi- (Vorsilbe): beide) oder eine **Bijektion** (englisch: bijection), wenn f surjektiv und injektiv ist.[72]

Äquivalent dazu sind:

- Für alle $y \in Y$ gibt es genau ein $x \in X$ mit der Eigenschaft $f(x) = y$.
- Das Urbild jedes Elements der Zielmenge Y ist genau einelementig ist:

$$\forall y \in Y \; \big(f^{-1}(\{y\}) \text{ hat genau ein Element}\big). \qquad \triangle$$

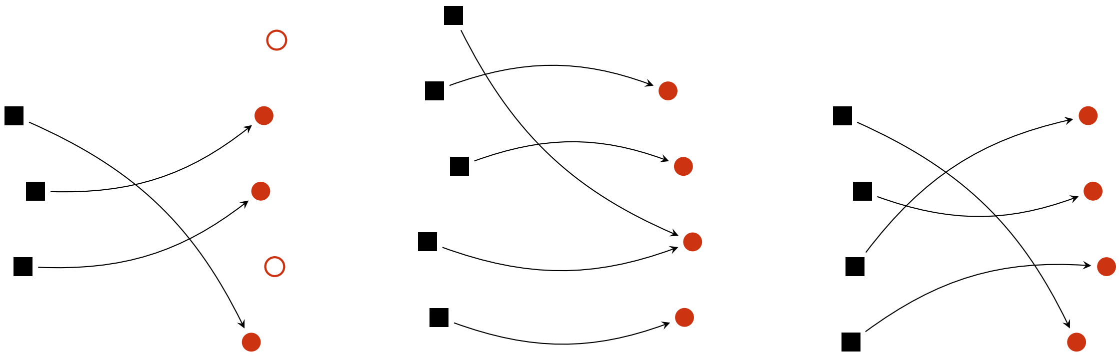


Abbildung 6.1.: Darstellung einer injektiven, aber nicht surjektiven Abbildung (links), einer surjektiven, aber nicht injektiven Abbildung (Mitte) sowie einer bijektiven Abbildung (rechts), siehe Definition 6.10.

Bemerkung 6.11 (Einbettung).

(i) Der Name **Einbettung** für eine injektive Abbildung $f: X \to Y$ rührt daher, dass durch f mit $f(X)$ ein Abbild der Menge X in der Menge Y entsteht, wobei jedes Element von $f(X)$ ein einelementiges Urbild in X hat: $f^{-1}(\{f(x)\}) = \{x\}$. Dadurch wird $f|^{f(X)}$ zu einer Bijektion (Lemma 6.12), und X kann mit seiner Bildmenge $f(X) \subseteq Y$ identifiziert werden. Durch diese Identifikation wird X in Y eingebettet.

(ii) Sind X und Y Mengen mit $X \subseteq Y$, dann heißt die injektive Abbildung $i_{Y \leftarrow X}$ mit

$$X \ni x \mapsto i_{Y \leftarrow X}(x) \coloneqq x \in Y$$

die **kanonische** oder **natürliche Injektion** (englisch: canonical injection, natural injection) oder die **kanonische** oder **natürliche Einbettung** (englisch: canonical embedding, natural embedding) von X in Y. △

Quizfrage 6.1: Können Sie (nicht-mathematische) Beispiele für injektive, surjektive bzw. bijektive Funktionen benennen?

Lemma 6.12 (Bijektiv-Machen einer injektiven Funktion).
Es sei $f: X \to Y$ eine injektive Funktion. Dann ist $f|^{f(X)}$ (erhalten durch die Einschränkung der Zielmenge auf die tatsächliche Bildmenge) bijektiv.

[72]Die Bijektivität von f wird manchmal auch durch die Schreibweise $f: X \rightarrowtail\!\!\!\!\rightarrow Y$ ausgedrückt.

Beweis. Der Beweis ist Gegenstand der Übung. □

Beispiel 6.13 (Injektivität, Surjektivität, Bijektivität).

(i) Die Funktion
$$\mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}$$
ist nicht surjektiv und nicht injektiv.

(ii) Die Funktion
$$\mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}_{\geqslant 0}$$
ist surjektiv, aber nicht injektiv. Hierbei ist $\mathbb{R}_{\geqslant 0} = \{x \in \mathbb{R} \,|\, x \geqslant 0\}$ die Menge der nichtnegativen reellen Zahlen.

(iii) Die Funktion
$$\mathbb{R}_{\geqslant 0} \ni x \mapsto x^2 \in \mathbb{R}$$
ist injektiv, aber nicht surjektiv.

(iv) Die Funktion
$$\mathbb{R}_{\geqslant 0} \ni x \mapsto x^2 \in \mathbb{R}_{\geqslant 0}$$
ist bijektiv.

(v) Ist X eine nichtleere Menge und $\sim$ eine Äquivalenzrelation auf X, dann heißt die Abbildung
$$\pi \colon X \ni x \mapsto [x] \in X/\sim, \tag{6.6}$$
die jedem Element seine Äquivalenzklasse zuordnet, die **kanonische Surjektion** (englisch: canonical surjection). Diese ist surjektiv.

(vi) Die leere Funktion $f \colon \emptyset \to Y$ ist für beliebige Mengen Y injektiv. Sie ist bijektiv genau dann, wenn $Y = \emptyset$ ist. △

Wenden wir das Konzept der Komposition von Relationen (Definition 5.4) speziell auf Funktionen an, so erhalten wir folgende Definition:

Definition 6.14 (Komposition von Funktionen).
Es seien $f \colon X \to Y$ und $g \colon Y \to Z$ Funktionen. Die Funktion $g \circ f$, definiert durch
$$X \ni x \mapsto (g \circ f)(x) \coloneqq g(f(x)) \in Z,$$
heißt die **Komposition** (englisch: composition, lateinisch: componere: zusammenstellen), die **Hintereinanderausführung** oder die **Verkettung** von f und g. Um die Reihenfolge klar zu benennen, sagt man auch „**g nach f**“. △

Quizfrage 6.2: Wie sieht man, dass $g \circ f$ wieder linkstotal und rechtseindeutig ist, also tatsächlich eine Funktion?

$$\begin{tikzcd} Z & Y \arrow[l, "g"'] & X \arrow[l, "f"'] \arrow[ll, bend left, "g\circ f"] \end{tikzcd}$$

Wir können den Sachverhalt aus Definition 6.14 durch das nebenstehende Bild illustrieren. Die Voraussetzung, dass die Zielmenge von f mit der Definitionsmenge von g übereinstimmt, kann relaxiert werden: Die Komposition $g \circ f$ von $f \colon X \to \widetilde{Y}$ und $g \colon Y \to Z$ ist definiert, sofern $f(X) \subseteq Y$ gilt.

Beispiel 6.15 (Komposition von Funktionen).
Es seien

$$\mathbb{R} \ni x \mapsto f(x) \coloneqq x^2 \in \mathbb{R},$$
$$\mathbb{R} \ni x \mapsto g(x) \coloneqq x + 1 \in \mathbb{R}.$$

Dann sind $f(\mathbb{R}) \subseteq \mathbb{R}$ und $g(\mathbb{R}) \subseteq \mathbb{R}$, also sind sowohl $g \circ f$ als auch $f \circ g$ definiert. Sie sind gegeben durch

$$\mathbb{R} \ni x \mapsto (g \circ f)(x) \coloneqq x^2 + 1 \in \mathbb{R},$$
$$\mathbb{R} \ni x \mapsto (f \circ g)(x) \coloneqq (x + 1)^2 \in \mathbb{R}. \qquad \triangle$$

Die Komposition von Funktionen f, g ist also, selbst wenn $g \circ f$ und $f \circ g$ beide definiert sind, i. A. nicht kommutativ.

Bemerkung 6.16 (Komposition mit der Identität und mit der kanonischen Einbettung).
Es sei $f \colon X \to Y$ eine Funktion. Dann gilt

$$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f \tag{6.7a}$$
$$f|_A = f \circ i_{X \leftarrow A} \qquad \text{für } A \subseteq X \tag{6.7b}$$
$$i_{Y \leftarrow B} \circ (f|^B) = f \qquad \text{für } Y \supseteq B \supseteq f(X). \tag{6.7c}$$

$\triangle$

Lemma 6.17 (Komposition von Funktionen ist assoziativ).
Es seien $f \colon X \to Y$, $g \colon Y \to Z$ und $h \colon Z \to W$ Funktionen. Dann gilt $(h \circ g) \circ f = h \circ (g \circ f)$, d. h., die Komposition von Funktionen ist assoziativ.

Beweis. Die Aussage ist ein Spezialfall von Lemma 5.5 (Komposition von Relationen ist assoziativ). Wir können aber auch einen eigenständigen Beweis führen: Für $x \in X$ gilt

$$\big((h \circ g) \circ f\big)(x) = (h \circ g)(f(x)) = h(g(f(x)))$$
$$\text{und} \quad \big(h \circ (g \circ f)\big)(x) = h\big((g \circ f)(x)\big) = h(g(f(x))).$$

Folglich stimmen $(h \circ g) \circ f \colon X \to W$ und $h \circ (g \circ f) \colon X \to W$ in Definitionsbereich, Zielmenge und Abbildungsvorschrift überein. $\square$

Definition 6.18 (Potenzen von Funktionen).
Es sei $f \colon X \to X$ eine Funktion.

(*i*) Wir definieren die **Potenzen** von f für $n \in \mathbb{N}_0$ rekursiv durch

$$f^0 \coloneqq \mathrm{id}_X \tag{6.8a}$$
$$f^{n+1} \coloneqq f^n \circ f \quad \text{für } n \in \mathbb{N}_0. \tag{6.8b}$$

(*ii*) f heißt eine **Involution** (englisch: involution), wenn $f^2 = \mathrm{id}_X$ gilt.

(*iii*) f heißt eine **Funktion endlicher Ordnung** (englisch: function of finite order), wenn es ein $n \in \mathbb{N}$ gibt mit der Eigenschaft $f^n = \mathrm{id}_X$. In diesem Fall heißt die kleinste Zahl $n \in \mathbb{N}$ mit dieser Eigenschaft die **Ordnung** von f. Falls kein $n \in \mathbb{N}$ mit der Eigenschaft $f^n = \mathrm{id}_X$ existiert, so heißt f eine **Funktion unendlicher Ordnung** (englisch: function of infinite order). △

Beachte: Jede Funktion $f: X \to X$ der Ordnung 2 ist eine Involution. Allerdings ist nicht jede Involution eine Funktion der Ordnung 2. (**Quizfrage 6.3:** Warum nicht?)

Quizfrage 6.4: Warum definieren wir (im Gegensatz zu Definition 5.8) keine negativen Potenzen von Funktionen?

Lemma 6.19 (Komposition injektiver und surjektiver Funktionen)**.**
Es seien $f: X \to Y$ und $g: Y \to Z$ Funktionen.

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

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

(*iii*) Ist $g \circ f$ injektiv, so ist f injektiv.

(*iv*) Ist $g \circ f$ injektiv und f surjektiv, dann ist g injektiv.

(*v*) Ist $g \circ f$ surjektiv, so ist g surjektiv.

(*vi*) Ist $g \circ f$ surjektiv und g injektiv, dann ist f surjektiv.

Beachte: Aus den Aussagen (*i*) und (*ii*) folgt sofort, dass die Komposition bijektiver Funktionen bijektiv ist.

Beweis. Aussage (*i*): Für $x_1, x_2 \in X$ gelte $(g \circ f)(x_1) = (g \circ f)(x_2)$, also $g(f(x_1)) = g(f(x_2))$. Aus der Injektivität von g folgt $f(x_1) = f(x_2)$, und aus der Injektivität von f folgt weiter $x_1 = x_2$. Also ist $g \circ f$ injektiv.

Aussage (*ii*): Es sei $z \in Z$. Aufgrund der Surjektität von g gibt es ein $y \in Y$, sodass $z = g(y)$ gilt. Wegen der Surjektivität von f gibt es ein $x \in X$, sodass $y = f(x)$ gilt. Es gilt also $z = g(y) = g(f(x)) = (g \circ f)(x)$, d. h., $z \in (g \circ f)(X)$.

Aussage (*iii*): Es seien $x_1, x_2 \in X$, sodass $f(x_1) = f(x_2)$ gilt. Dann gilt auch $g(f(x_1)) = g(f(x_2))$, und wegen der Injektivität von $g \circ f$ folgt $x_1 = x_2$, d. h., f ist injektiv.

Aussage (*iv*): Es seien $y_1, y_2 \in Y$, sodass $g(y_1) = g(y_2)$ gilt. Aufgrund der Surjektivität von f gibt es $x_1, x_2 \in X$, sodass $y_1 = f(x_1)$ und $y_2 = f(x_2)$ gilt. Es folgt $g(f(x_1)) = g(f(x_2))$, und aufgrund der Injektivität von $g \circ f$ folgt $x_1 = x_2$, also auch $y_1 = y_2$, d. h., g ist injektiv.

Aussage (*v*): Es sei $z \in Z$. Aufgrund der Surjektivität von $g \circ f$ gibt es ein $x \in X$, sodass $z = g(f(x))$ gilt. Das heißt aber $z = g(y)$ für $y = f(x)$, also ist g surjektiv.

Aussage (*vi*): Ist $g \circ f$ surjektiv und g injektiv, dann ist f surjektiv. Es sei $y \in Y$ und $z \coloneqq g(y)$. Aufgrund der Surjektivität von $g \circ f$ gibt es ein $x \in X$, sodass $g(f(x)) = z$ gilt. Nun gilt also $z = g(y) = g(f(x))$, und da g injektiv ist, folgt $y = f(x)$, d. h., f ist surjektiv. □

Folgerung 6.20 (Komposition zur Identität)**.**
Es seien $f: X \to Y$ und $g: Y \to X$ Funktionen. Wenn $g \circ f = \mathrm{id}_X$ ist, dann ist f injektiv und g surjektiv.

Beweis. Die Identitätsabbildung id_X ist bijektiv. Aus Lemma 6.19, Aussagen *(iii)* und *(v)* folgt daher, dass f injektiv und g surjektiv ist. □

Folgerung 6.21 (Funktionen endlicher Ordnung sind bijektiv)**.**
Es sei $f: X \to X$ eine Funktion endlicher Ordnung. Dann ist f bijektiv. Insbesondere ist jede Involution bijektiv.

Beweis. Nach Voraussetzung gibt es ein $n \in \mathbb{N}$, sodass $f^n = \mathrm{id}_X$ gilt. Ist $n = 1$, so ist nichts zu zeigen, denn id_X ist bijektiv. Andernfalls ergibt sich aus $f^n = f \circ f^{n-1} = \mathrm{id}_X$ mit Folgerung 6.20 die Surjektivität von f. Weiter ergibt sich aus $f^n = f^{n-1} \circ f = \mathrm{id}_X$ mit Folgerung 6.20 die Surjektivität von f. □

Ende der Vorlesung 6

Ende der Woche 3

§ 6.2 Umkehrfunktion

Definition 6.22 (Umkehrfunktion)**.**
Es sei $f: X \to Y$ eine Funktion. Wenn eine Funktion $g: Y \to X$ existiert, sodass

$$g \circ f = \mathrm{id}_X \quad \text{und} \quad f \circ g = \mathrm{id}_Y \tag{6.9}$$

gilt, dann heißt f **invertierbar** (englisch: invertible). Die (eindeutig bestimmte) Funktion $g: Y \to X$ heißt die **Umkehrfunktion** oder **Umkehrabbildung**, **inverse Funktion** (englisch: inverse function) oder **inverse Abbildung** (englisch: inverse map) von f. Sie wird auch mit $f^{-1}: Y \to X$ bezeichnet. △

Lemma 6.23 (Umkehrfunktionen sind eindeutig)**.**
Es sei $f: X \to Y$ eine invertierbare Funktion. Sind $g_1, g_2: Y \to X$ beides Umkehrfunktionen von f, dann ist $g_1 = g_2$.

Beweis. Es gilt

$$\begin{aligned} g_1 &= g_1 \circ \mathrm{id}_Y && \text{nach (6.7a)} \\ &= g_1 \circ (f \circ g_2) && \text{nach Voraussetzung} \\ &= (g_1 \circ f) \circ g_2 && \text{nach Lemma 6.17} \\ &= \mathrm{id}_X \circ g_2 && \text{nach Voraussetzung} \\ &= g_2 && \text{nach (6.7a).} \end{aligned}$$
□

Lemma 6.24 (Charakterisierung der Invertierbarkeit)**.**
Es sei $f: X \to Y$ eine Funktion. Dann sind äquivalent:

(i) f ist invertierbar.

(ii) f ist bijektiv.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $f: X \to Y$ invertierbar und $f^{-1}: Y \to X$ die Umkehrfunktion. Die Abbildung $f^{-1} \circ f = \mathrm{id}_X$ ist bijektiv, insbesondere injektiv. Aus Lemma 6.19 (iii) folgt also, dass f injektiv ist. Auch die Abbildung $f \circ f^{-1} = \mathrm{id}_Y$ ist bijektiv, insbesondere surjektiv. Aus Lemma 6.19 (v) folgt also, dass f auch surjektiv ist.

Aussage (ii) $\Rightarrow$ Aussage (i): Wir konstruieren eine Abbildung $g: Y \to X$ wie folgt: Wir definieren für beliebiges $y \in Y$ den Funktionswert $g(y)$ als das nach Voraussetzung eindeutig definierte $x \in X$, für das $y = f(x)$ gilt. Für diese Funktion haben wir also $g(y) = x \Leftrightarrow f(x) = y$ und daher

$$(g \circ f)(x) = g(f(x)) = x \quad \text{für alle } x \in X$$

sowie

$$(f \circ g)(y) = f(g(y)) = y \quad \text{für alle } y \in Y.$$

Damit ist $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$ gezeigt, d. h., f ist invertierbar, und g ist die zugehörige Umkehrfunktion. □

Bemerkung 6.25 (Umkehrfunktion)**.**

(i) Die Bedingung (6.9), also

$$f^{-1} \circ f = \mathrm{id}_X \quad \text{und} \quad f \circ f^{-1} = \mathrm{id}_Y$$

zeigt nicht nur, dass $f^{-1}: Y \to X$ die Umkehrfunktion von $f: X \to Y$ ist, sondern auch, dass f die Umkehrfunktion von f^{-1} ist. Insbesondere ist mit f auch f^{-1} bijektiv, und die Invertierung einer Funktion ist **involutorisch**, denn es gilt $(f^{-1})^{-1} = f$.

(ii) Für die Abbildungsvorschrift der Umkehrfunktion gilt $f^{-1}(y) = x \Leftrightarrow y = f(x)$.

(iii) Wenn wir wissen, dass eine Funktion $f: X \to Y$ bijektiv (invertierbar) ist, dann reicht es aus, für eine Funktion $g: Y \to X$ eine der beiden Bedingungen aus (6.9) zu prüfen, um zu bestätigen, dass g die Umkehrfunktion von f ist, denn:

$$\begin{aligned} f \circ g = \mathrm{id}_Y \quad &\Rightarrow \quad f^{-1} \circ f \circ g = f^{-1} \quad \Rightarrow \quad g = f^{-1}, \\ g \circ f = \mathrm{id}_X \quad &\Rightarrow \quad g \circ f \circ f^{-1} = f^{-1} \quad \Rightarrow \quad g = f^{-1}. \end{aligned}$$

Wenn die Bijektivität von f aber nicht bestätigt ist, dann kann aus der Erfüllung nur einer der beiden Bedingungen aus (6.9) nicht geschlossen werden, dass f invertierbar und g die Umkehrfunktion ist, vgl. Sätze 6.29 und 6.46. △

Beispiel 6.26 (Umkehrfunktion)**.**

(i) Die Funktion $f: \mathbb{R} \to \mathbb{R}$ mit $f(x) = 2\,x + 1$ ist invertierbar. Um ihre Umkehrfunktion zu bestimmen, lösen wir $y = 2\,x + 1$ nach x auf, erhalten also $x = \frac{1}{2}(y - 1)$. Die Umkehrfunktion ist also gegeben durch $f^{-1}(y) = \frac{1}{2}(y - 1)$.

(*ii*) Die Funktion $f\colon \mathbb{N} \to \mathbb{N}$ mit $f(x) = 2\,x - 1$ ist nicht invertierbar. Sie ist zar injektiv, aber nicht surjektiv. Durch Einschränkung der Zielmenge auf die ungeraden Zahlen $U \coloneqq \{n \in \mathbb{N} \mid n \text{ ist ungerade}\}$ wird $f|^U$ bijektiv (Lemma 6.12). Es ist dann $(f|^U)^{-1}(y) = \frac{1}{2}(y+1)$.

(*iii*) Die Funktion $f\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ mit $f(x) = x^2$ ist bijektiv (Beispiel 6.13). Die Umkehrfunktion f^{-1} heißt die **Wurzelfunktion** (englisch: square root function), geschrieben $f^{-1}(y) = \sqrt{y}$. △

Bemerkung 6.27 (Umkehrfunktion).

Das Symbol f^{-1} für die Umkehrfunktion muss vom Urbild der Funktion f unterschieden werden. Wenn die Umkehrfunktion von $f\colon X \to Y$ existiert, so gilt jedoch

$$\underbrace{f^{-1}(\{y\})}_{\text{Urbild von } \{y\}} = \{\underbrace{f^{-1}(y)}_{\text{Wert der Umkehrfunktion bei } y}\}.$$ △

Satz 6.28 (Umkehrfunktion der Komposition).

Es seien $f\colon X \to Y$ und $g\colon Y \to Z$ bijektive Funktionen. Dann ist auch $g \circ f$ bijektiv, und für die Umkehrfunktion gilt

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}. \tag{6.10}$$

Quizfrage 6.5: Wie erklärt man sich anschaulich, dass sich bei der Umkehrfunktion die Reihenfolge ändert?

Beweis. Die Bijektivität von $g \circ f$ folgt sofort aus Lemma 6.19, Aussagen (*i*) und (*ii*). Aufgrund der Assoziativität der Komposition von Funktionen (Lemma 6.17) und Bemerkung 6.16 gilt

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

Da die Bijektivität von $g \circ f$ bereits bekannt ist, reicht das nach Bemerkung 6.25 aus, um zu bestätigen, dass $f^{-1} \circ g^{-1}$ die Umkehrfunktion von $g \circ f$ ist. □

Satz 6.29 (Charakterisierung der Injektivität).

Es sei $f\colon X \to Y$ eine Funktion und $X \neq \emptyset$. Dann sind äquivalent:

(*i*) f ist injektiv.

(*ii*) Es existiert eine Abbildung $g\colon Y \to X$ mit der Eigenschaft $g \circ f = \mathrm{id}_X$. Eine solche Abbildung heißt eine **Linksinverse** (englisch: left inverse) von f. Sie ist notwendig surjektiv. Ihre Einschränkung $g|_{f(X)}$ auf das Bild von f ist eindeutig.

(*iii*) Für beliebige Mengen X_0 und beliebige Abbildungen $f_1, f_2\colon X_0 \to X$ gilt: Aus $f \circ f_1 = f \circ f_2$ folgt $f_1 = f_2$.

(**Quizfrage 6.6:** Welche Implikationen stimmen nicht mehr, wenn $X = \emptyset$ ist?)

Beweis. Aussage (*i*) $\Rightarrow$ Aussage (*ii*): Wir definieren zunächst eine Abbildung $\widetilde{g}\colon f(X) \to X$ wie folgt: Wir setzen für $y \in f(X)$ als $\widetilde{g}(y)$ das wegen der Injektivität eindeutig definierte $x \in X$, für das $y = f(x)$ gilt. Für diese Funktion haben wir also $\widetilde{g}(y) = x \Leftrightarrow f(x) = y$ und damit

$$(\widetilde{g} \circ f)(x) = \widetilde{g}(f(x)) = x \quad \text{für alle } x \in X.$$

Damit ist $\widetilde{g} \circ f = \mathrm{id}_X$ gezeigt. Aufgrund von Folgerung 6.20 ist $\widetilde{g}$ surjektiv. Wir setzen nun $\widetilde{g}\colon f(X) \to X$ zu $g\colon Y \to X$ fort. Dazu wählen wir irgendein $x_0 \in X$ und setzen $g(y) \coloneqq \widetilde{g}(y)$ für $y \in f(X)$ und $g(y) \coloneqq x_0$ für $y \in Y \setminus f(X)$. Die Funktion g erbt die Surjektivität von $\widetilde{g}$.

Angenommen, $h\colon Y \to X$ sei eine andere Linksinverse von f. Dann gilt für $y \in f(X)$ aufgrund der Injektivität von f: Es gibt genau ein $x \in X$ mit der Eigenschaft $y = f(x)$. Wegen $h(y) = h(f(x)) = x$ und ebenso $g(y) = g(f(x)) = x$ müssen g und h auf $f(X)$ übereinstimmen.

Die weiteren Implikationen sind Bestandteil einer Übung. □

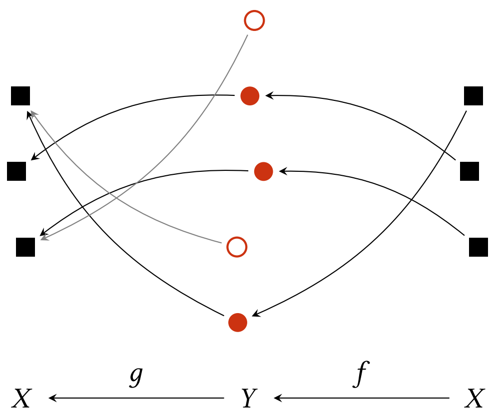


Abbildung 6.2.: Jede injektive Funktion $f\colon X \to Y$ mit $X \neq \emptyset$ besitzt eine Linksinverse $g\colon Y \to X$ mit $g \circ f = \mathrm{id}_X$ (Satz 6.29). Die Linksinverse ist surjektiv und i. A. nicht eindeutig, ihre Einschränkung $g|_{f(X)}$ jedoch schon.

Beispiel 6.30 (Linksinverse)**.**
Die Funktion $f\colon \mathbb{N} \to \mathbb{N}$ mit $f(n) = n + 3$ ist injektiv (aber nicht bijektiv). Jede Funktion $g\colon \mathbb{N} \to \mathbb{N}$ mit $g(n) = n - 3$ für $n \geqslant 4$ und $g(1), g(2), g(3) \in \mathbb{N}$ beliebig ist eine Linksinverse von f. △

Eine Charakterisierung der Surjektivität analog zu Satz 6.29 folgt erst in Satz 6.46, weil wir dafür interessanterweise das Auswahlaxiom benötigen.

§ 6.3 Mächtigkeit von Mengen

Mit Hilfe von Funktionen können wir Mengen in ihrer Mächtigkeit (vereinfacht gesagt, bzgl. der Anzahl ihrer Elemente) vergleichen.

Definition 6.31 (Gleichmächtigkeit von Mengen).
Es seien X und Y Mengen. Wir sagen, X sei **gleichmächtig** (englisch: equinumerous) zu Y, wenn es eine bijektive Abbildung $f : X \to Y$ gibt. Wir schreiben in diesem Fall $X \sim Y$. △

Die Gleichmächtigkeit von Mengen ist eine Äquivalenzrelation auf der Klasse aller Mengen, siehe Übung. Die Äquivalenzklassen heißen **Kardinalzahlen** (englisch: cardinal numbers).

Definition 6.32 (Endlichkeit, Abzählbarkeit, Überabzählbarkeit).
Es sei X eine Menge.

(i) X heißt **endlich** (englisch: finite), wenn $X \sim [\![1, n]\!]$ für ein $n \in \mathbb{N}_0$ gilt, ansonsten **unendlich** (englisch: infinite).

(ii) Wenn X endlich ist mit $X \sim [\![1, n]\!]$, dann heißt $n \in \mathbb{N}_0$ die **Mächtigkeit** oder **Kardinalität** (englisch: cardinality) von X. Wir schreiben dann: $\#X = n$.[73]

(iii) X heißt **abzählbar unendlich** (englisch: countably infinite), wenn $X \sim \mathbb{N}$ gilt.

(iv) X heißt **abzählbar** (englisch: countable), wenn X entweder endlich oder abzählbar unendlich ist, ansonsten **überabzählbar** (englisch: uncountable). △

Beachte: Die Mächtigkeit einer endlichen Menge ist eindeutig bestimmt. Die leere Menge $\emptyset$ ist nur zu sich selbst gleichmächtig. Sie ist die einzige Menge mit Mächtigkeit 0. Teilmengen endlicher Mengen sind endlich. Teilmengen abzählbarer Mengen sind abzählbarer.

Beispiel 6.33 (Gleichmächtigkeit von Mengen, Abzählbarkeit, Überabzählbarkeit).

(i) Die Menge der ganzen Zahlen $\mathbb{Z}$ ist gleichmächtig zur Menge der geraden ganzen Zahlen $\{2\,n \mid n \in \mathbb{Z}\}$. Sie ist abzählbar unendlich.

(ii) Die Menge der rationalen Zahlen $\mathbb{Q}$ ist abzählbar unendlich.[74]

(iii) Die Vereinigung abzählbar vieler abzählbarer Mengen ist wieder abzählbar.

(iv) Die Menge der reellen Zahlen $\mathbb{R}$ ist überabzählbar.[75] △

Lemma 6.34 (Veränderung der Kardinalität um 1).
Es sei X eine endliche Menge und $x \in X$. Dann gilt

$$\#X = \#(X \setminus \{x\}) + 1. \tag{6.11}$$

Beweis. Es sei $n = \#(X \setminus \{x\}) \in \mathbb{N}_0$. Es gibt also eine bijektive Abbildung $\widehat{f} : \{1, \ldots, n\} \to X \setminus \{x\}$. Wir definieren $f : \{1, \ldots, n+1\} \to X$ durch $f(i) \coloneqq \widehat{f}(i)$ für $i = 1, \ldots, n$ und $f(n+1) \coloneqq x$. Dann ist f ebenfalls bijektiv, d. h., $\#X = n + 1 = \#(X \setminus \{x\}) + 1$. □

Satz 6.35 (Funktionen auf endlichen Mengen).
Es seien X und Y endliche, gleichmächtige Mengen und $f : X \to Y$ eine Funktion. Dann sind äquivalent:

[73] In dieser Lehrveranstaltung verwenden wir das Symbol # nur für endliche Mengen.
[74] Ein Beweis erfolgt typischerweise in der Lehrveranstaltung *Analysis*.
[75] Ein Beweis erfolgt typischerweise in der Lehrveranstaltung *Analysis*.

(*i*) f ist injektiv.

(*ii*) f ist surjektiv.

(*iii*) f ist bijektiv.

Beweis. Aussage (*i*) $\Rightarrow$ Aussage (*ii*): Wir führen einen Induktionsbeweis nach der Mächtigkeit $n = \#X = \#Y \in \mathbb{N}_0$. Der Induktionsanfang ist der Fall $n = 0$, also $X = Y = \emptyset$. Dann ist die einzig mögliche Abbildung die leere Abbildung, diese ist bijektiv. Im Induktionsschritt schließen wir von n auf $n + 1$ für $n \in \mathbb{N}_0$. Es sei also nun $\#X = \#Y = n + 1$. Wir wählen ein $x \in X$ und setzen $y \coloneqq f(x)$. Dann gilt aufgrund von Lemma 6.34 $\#X \setminus \{x\} = \#Y \setminus \{y\} = n$.

Wir bezeichnen mit $\widetilde{f} \colon X \setminus \{x\} \to Y \setminus \{y\}$ die Einschränkung von f. Diese ist aufgrund der vorausgesetzten Injektivität definiert, denn x ist das einzige Element von X, das durch f auf y abgebildet wird. Außerdem erbt $\widetilde{f}$ die Injektivität von f. Nach Induktionsvoraussetzung ist $\widetilde{f}$ daher auch surjektiv, alle Elemente von $Y \setminus \{y\}$ liegen also im Bild von $\widetilde{f}$ und damit im Bild von f. Da auch y im Bild von f liegt, ist f tatsächlich surjektiv.

Aussage (*ii*) $\Rightarrow$ Aussage (*i*): Wir führen auch hier einen Induktionsbeweis nach der Mächtigkeit $n = \#X = \#Y$. Der Induktionsanfang beinhaltet die Fälle $n = 0$ und $n = 1$. Im Fall $n = 0$ ist $X = Y = \emptyset$, dann ist die einzig mögliche Abbildung die leere Abbildung, diese ist bijektiv. Im Fall $n = 1$ gibt es ebenfalls nur eine mögliche Abbildung, auch diese ist bijektiv.

Im Induktionsschritt schließen wir von n auf $n + 1$ für $n \in \mathbb{N}$. Es sei also nun $\#X = \#Y = n + 1$. Wir führen einen Widerspruchsbeweis, nehmen also an, dass f surjektiv, aber *nicht* injektiv ist. Dann gibt es ein $y_0 \in Y$, sodass das Urbild $f^{-1}(\{y_0\})$ aus (mindestens) zwei verschiedenen Elementen besteht, sagen wir $x_{01}, x_{02} \in f^{-1}(\{y_0\})$ und $x_{01} \neq x_{02}$. Wir wählen außerdem ein $y_1 \in Y \setminus \{y_0\}$ aus, was wegen $\#Y = n + 1 \geqslant 2$ möglich ist. Dazu existiert ein x_1 mit $f(x_1) = y_1$. Wegen $y_1 \neq y_0$ ist $x_1 \neq x_{01}$ und $x_1 \neq x_{02}$.

Wir konstruieren nun eine Funktion $\widetilde{f} \colon X \setminus \{x_{01}\} \to Y \setminus \{y_0\}$ durch

$$\widetilde{f}(x) \coloneqq \begin{cases} f(x) & \text{im Fall } f(x) \neq y_0, \\ y_1 & \text{im Fall } f(x) = y_0. \end{cases}$$

Dann ist $\widetilde{f}$ ebenfalls surjektiv, denn:

(1) Für jedes $y \in Y \setminus \{y_0, y_1\}$ existiert aufgrund der Surjektivität von f ein $x \in X$ mit $\widetilde{f}(x) = f(x) = y$, und wegen $f(x_{01}) = y_0$ ist $x \in X \setminus \{x_{01}\}$.

(2) Außerdem gilt $f(x_{02}) = y_0$, also $\widetilde{f}(x_{02}) = y_1$.

Aufgrund von Lemma 6.34 gilt wieder $\#X \setminus \{x_{01}\} = \#Y \setminus \{y_0\} = n$. Nach Induktionsvoraussetzung ist $\widetilde{f}$ daher auch injektiv. Jedoch enthält $\widetilde{f}^{-1}(\{y_1\})$ neben x_1 auch noch mindestens das weitere Element $x_{02} \in f^{-1}(\{y_0\})$. Das steht im Widerspruch zur Injektivität von $\widetilde{f}$.

Wir haben jetzt Aussage (*i*) $\Leftrightarrow$ Aussage (*ii*) bewiesen. Da die Bijektivität sich aus Surjektivität und Injektivität zusammensetzt, gilt auch Aussage (*i*) $\Leftrightarrow$ Aussage (*ii*) $\Leftrightarrow$ Aussage (*iii*). □

Beachte: Die Aussage von Satz 6.35 ist falsch, wenn X und Y zwar gleichmächtig, aber nicht endlich sind, siehe Übung.

Der Begriff der Gleichmächtigkeit erlaubt noch keinen Vergleich der Mächtigkeit von Mengen. Dazu dient folgende Definition.

Definition 6.36 (Vergleich der Mächtigkeit von Mengen)**.**
Es seien X und Y Mengen. Wir sagen, X sei **höchstens gleichmächtig** (englisch: at most equinumerous) zu Y, wenn es eine injektive Abbildung $f\colon X \to Y$ gibt. Wir schreiben in diesem Fall $X \preccurlyeq Y$ oder auch $Y \succcurlyeq X$. △

Bemerkung 6.37 (Mächtigkeit von Mengen)**.**

(i) Man kann zeigen, dass $X \preccurlyeq Y$ äquivalent ist zu: Es gilt $X = \emptyset$, oder es existiert eine surjektive Abbildung $g\colon Y \to X$.[76]

(ii) Die Relation $\preccurlyeq$ induziert eine Ordnungsrelation auf der Klasse aller Kardinalzahlen. Die Reflexivität und Transitivität sind dabei leicht einzusehen. (**Quizfrage 6.7:** Details?) Der Beweis der Antisymmetrie ist jedoch aufwändig und erfordert den Satz von Cantor-Bernstein-Schröder.

(iii) Der Satz von Cantor besagt, dass jede Menge echt weniger mächtig ist als ihre Potenzmenge, also $X \precnsim \mathcal{P}(X)$ mit der zu $\preccurlyeq$ gehörenden strengen Ordnungsrelation.[77]

(iv) Unter Zuhilfenahme des Auswahlaxioms (siehe § 6.5) kann man sogar zeigen, dass zwei Mengen bzgl. $\preccurlyeq$ stets vergleichbar sind. Es folgt, dass $\preccurlyeq$ sogar eine totale Ordnung auf der Klasse aller Kardinalzahlen induziert.

(v) Die natürlichen Zahlen sind ein Repräsentant der Äquivalenzklasse der kleinsten unendlichen Mengen. Es gilt also $\mathbb{N} \preccurlyeq X$ für alle unendlichen Mengen X. (Diese Aussage ist unabhängig vom Auswahlaxiom.) △

§ 6.4 Familien und Folgen

Definition 6.38 (Familie, Teilfamilie, Oberfamilie)**.**
Es seien I und Y Mengen.

(i) Eine Abbildung

$$y\colon I \ni i \mapsto y(i) \coloneqq y_i \in Y$$

heißt eine **Familie** (englisch: family) **in** Y oder **Familie mit Werten in** Y mit der **Indexmenge** (englisch: index set) I. Kurz wird diese auch mit $(y_i)_{i\in I}$ bezeichnet.

(ii) Für $i \in I$ heißt y_i das **Mitglied** (englisch: member) der Familie $(y_i)_{i\in I}$ zum Index i.

(iii) Ist $I_0 \subseteq I$, dann heißt $(y_i)_{i\in I_0}$ eine **Teilfamilie** (englisch: subfamily) von $(y_i)_{i\in I}$, und $(y_i)_{i\in I}$ heißt eine **Oberfamilie** (englisch: superfamily) von $(y_i)_{i\in I_0}$. Die Teil- bzw. Oberfamilie heißt **echt** (englisch: proper subfamiliy, proper superfamily) im Fall $I_0 \subsetneq I$.

[76] siehe etwa Deiser, 2024a, Kapitel 1.4

[77] Es gibt also für jede Menge X eine injektive Abbildung $X \to \mathcal{P}(X)$, aber niemals eine surjektive Abbildung $X \to \mathcal{P}(X)$.

(*iv*) Ist I endlich, gilt also $I \sim [\![1, n]\!]$ für ein $n \in \mathbb{N}_0$, so heißt $(y_i)_{i\in I}$ eine **endliche Familie** (englisch: finite family) mit n Mitgliedern.

(*v*) Ist I unendlich, so heißt $(y_i)_{i\in I}$ eine **unendliche Familie** (englisch: infinite family).

(*vi*) Ist I abzählbar unendlich, gilt also $I \sim \mathbb{N}$, so heißt $(y_i)_{i\in I}$ eine **abzählbar unendliche Familie** (englisch: countably infinite family).

(*vii*) Ist $I = \emptyset$, so heißt $(y_i)_{i\in I}$ die **leere Familie** (englisch: empty family) **in** Y, andernfalls eine **nichtleere Familie** (englisch: non-empty familiy) **in** Y. Wir schreiben die leere Familie kurz auch als $()$.

(*viii*) Zwei Familien $(y_i)_{i\in I}$ und $(z_j)_{j\in J}$ heißen **gleichmächtig**, wenn ihre Indexmengen gleichmächtig sind ($I \sim J$). △

In der Übung definieren wir noch die **Konkatenation** (englisch: concatenation) $F_1 \parallel F_2$ von zwei oder auch von beliebig vielen Familien $\big\|_{i\in I} F_i$.

Da die Indexmenge I einer Familie im Allgemeinen ungeordnet ist, also keine Ordnungsrelation auf I gegeben ist, haben auch die Mitglieder der Familie $(y_i)_{i\in I}$ keine natürliche Reihenfolge. Wenn I jedoch totalgeordnet ist, dann kann diese Ordnung auf die Familie übertragen werden.

Definition 6.39 (geordnete Familie, Folge, endliche Folge, Tupel)**.**
Es seien I und Y Mengen.

(*i*) Ist I totalgeordnet, dann heißt eine Familie $(y_i)_{i\in I}$ in Y auch eine **geordnete Familie**.

(*ii*) Ist speziell $I = \mathbb{N}$ oder allgemeiner $I = \{n \in \mathbb{Z} \mid n \geqslant n_0\}$ mit einem Startindex $n_0 \in \mathbb{Z}$, so heißt $(y_i)_{i\in I}$ eine **Folge** (englisch: sequence) **in** Y.[78] Wir nennen die **Mitglieder einer Folge** auch **Glieder** (englisch: terms) der Folge oder **Folgenglieder**.

(*iii*) Ist speziell $I = [\![1, n]\!]$, so heißt $(y_i)_{i\in I}$ eine **endliche Folge** (englisch: finite sequence) **in** Y mit n Mitgliedern oder der **Länge** n (englisch: length).[79]

(*iv*) Wir können eine endliche Folge mit der Indexmenge $I = [\![1, n]\!]$ auch als **n-Tupel** (englisch: n-tuple) $(y_1, y_2, \ldots, y_n)$ notieren.[80] Insbesondere kann die leere Folge als $()$ geschrieben werden. △

Bemerkung 6.40 (Mengen und Familien)**.**

(*i*) Mengen und Familien sind konzeptionell eng verwandt, aber keine identischen Konzepte.

Jeder Familie $(y_i)_{i\in I}$ in Y können wir die **Menge ihrer Mitglieder** (englisch: set of family members) $\{y_i \mid i \in I\} \subseteq Y$ zuordnen.

Umgekehrt können wir jeder Menge Y durch Indizierung über sich selbst eine Familie in Y zuordnen, nämlich die Abbildung $Y \ni y \mapsto y \in Y$.

(*ii*) Im Unterschied zu einer Menge, die jedes ihrer Elemente nur einmal enhält, können Mitglieder einer Familie $(y_i)_{i\in I}$ mehrfach mit verschiedenen Indizes i vorkommen. △

[78] Dabei wird die übliche Totalordnung „$\leqslant$“ auf $\mathbb{Z}$ genutzt, eingeschränkt auf I.
[79] Auch hier wird die übliche Totalordnung auf $\mathbb{N}$ genutzt, eingeschränkt auf I.
[80] Wir nutzen dabei also die übliche Totalordnung auf $\mathbb{N}$.

Beispiel 6.41 (Folge).

(i) Die Abbildung

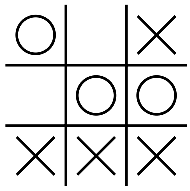
$$\mathbb{N} \ni n \mapsto y_n \coloneqq \frac{1}{n} \in \mathbb{R}$$

ist eine Folge in $\mathbb{R}$ mit der Standard-Indexmenge $\mathbb{N}$. Kurz wird diese Folge auch als $\left(\frac{1}{n}\right)_{n \in \mathbb{N}}$ notiert.

(ii) Die endliche Folge (notiert als Tupel) in $[\![1,3]\!] \times [\![1,3]\!]$

$$\big((1,3), (1,1), (3,1), (2,2), (3,3), (2,3), (3,2)\big)$$

könnte den Verlauf einer Partie Tic-Tac-Toe darstellen (bei der der erste Spieler (Kreuz) gewonnen hat).



△

§ 6.5 Das Auswahlaxiom

Das **Auswahlaxiom** (englisch: axiom of choice) der axiomatischen Mengenlehre nach Zermelo und Fraenkel besagt: Ist $\mathcal{A}$ eine Menge von nichtleeren Mengen, dann gibt es eine Funktion $F\colon \mathcal{A} \to \bigcup \mathcal{A}$, sodass gilt:

$$\forall A \in \mathcal{A} \ \big(F(A) \in A\big).$$

Eine solche Funktion F heißt eine **Auswahlfunktion** (englisch: choice function) für $\mathcal{A}$, weil sie aus jeder Menge $A \in \mathcal{A}$ irgendein Element auswählt und als Funktionswert $F(A)$ setzt. Das Auswahlaxiom besagt also, dass es möglich ist, aus jeder Menge $A \in \mathcal{A}$ ein Element auszuwählen, selbst wenn $\mathcal{A}$ aus überabzählbar vielen verschiedenen Mengen besteht und man daher nicht in der Lage ist, ein Verfahren anzugeben, nach dem die Auswahl geschehen soll.

Das Auswahlaxiom ist ein optionaler Bestandteil der axiomatischen Mengenlehre nach Zermelo und Fraenkel, es kann also dazugenommen werden oder auch nicht.[81] Es wird aber wohl von den meisten Mathematiker:innen akzeptiert. In Fällen, in denen $\mathcal{A}$ nur endlich viele Mengen enthält, wird das Auswahlaxiom nicht benötigt, weil seine Aussage bereits aus den anderen Axiomen folgt. Wir werden in der Vorlesung auf diese Weise[AoC] darauf hinweisen, wenn ein Resultat von der Hinzunahme des Auswahlaxioms abhängt. Einige Beispiele folgen bereits in diesem Abschnitt, siehe Satz 6.46.

Definition 6.42 (allgemeines kartesisches Produkt).

Es sei I eine Menge, und weiter sei A_i eine Menge für jedes $i \in I$. Dann ist das **kartesische Produkt** dieser Mengen gegeben durch

$$\bigtimes_{i \in I} A_i \coloneqq \Big\{ a\colon I \to \bigcup_{i \in I} A_i \;\Big|\; \underbrace{a(i)}_{\text{oder auch } a_i} \in A_i \text{ für alle } i \in I \Big\}. \tag{6.12}$$

△

[81]Man spricht von den ZF-Axiomen (ohne das Auswahlaxiom) und von den ZFC-Axiomen (mit Auswahlaxiom).

Das kartesische Produkt der Mengen A_i besteht also aus *Funktionen* auf der Indexmenge I, deren Funktionswerte jeweils im richtigen „Faktor“ A_i liegen.[82] Im Fall $I = \emptyset$ besteht das kartesische Produkt (6.12) aus dem einen einzigen Element $a\colon \emptyset \to \emptyset$, der leeren Funktion, besteht.

Bemerkung 6.43 (allgemeines kartesisches Produkt)**.**

(i) Wir hatten das **kartesische Produkt** bisher nur für endlich viele Mengen definiert, siehe Definition 4.8. Die allgemeine Definition 6.42 erfordert den Begriff der Funktion, der nun zur Verfügung steht.

(ii) Die Definition 6.42 lässt sich als Verallgemeinerung der Definition 4.8 verstehen: Ist nämlich die (totalgeordnete) Indexmenge $I = [\![1, n]\!]$ für $n \in \mathbb{N}_0$, so ist $\bigtimes_{i \in I} A_i$ nach (6.12) die Menge aller Funktionen $a\colon [\![1, n]\!] \to \bigcup_{i=1}^n A_i$ mit $a(i) =: a_i \in A_i$. Wenn wir die Funktionswerte als n-Tupel $(a_1, a_2, \dots, a_n)$ schreiben, so haben wir ein Element aus $\bigtimes_{i \in I} A_i$ gemäß Definition 4.8. Die Zuordnung $a \mapsto (a_1, a_2, \dots, a_n)$ (also Funktion auf das n-Tupel ihrer Funktionswerte) ist bijektiv.

(iii) Wenn alle Mengen $A_i = A$ sind, so schreiben wir statt $\bigtimes_{i \in I} A$ auch A^I. Es ist also beispielsweise

$\mathbb{R}^{\mathbb{N}}$ die Menge aller Folgen mit Werten in $\mathbb{R}$,

$\{0, 1\}^I$ die Menge aller $\{0, 1\}$-wertigen (binären) Funktionen auf einer Menge I.

Dabei wird $\{0, 1\}^A$ in manchen Büchern auch als Schreibweise für die Potenzmenge $\mathcal{P}(A)$ verwendet. (**Quizfrage 6.8:** Inwiefern ist diese Schreibweise gerechtfertigt?) △

Wir benötigen das Auswahlaxiom beispielsweise, um das Gegenstück zu Satz 6.29 für die Charakterisierung der Surjektivität zu beweisen:

Satz 6.44 (Charakterisierung der Surjektivität$^{\text{AoC}}$)**.**
Es sei $f\colon X \to Y$ eine Funktion. Dann sind äquivalent:

(i) f ist surjektiv.

(ii) Es existiert eine Abbildung $h\colon Y \to X$ mit der Eigenschaft $f \circ h = \mathrm{id}_Y$. Eine solche Abbildung heißt eine **Rechtsinverse** (englisch: right inverse) von f. Sie ist notwendig injektiv.

(iii) Für beliebige Mengen Z und beliebige Abbildungen $g_1, g_2\colon Y \to Z$ gilt: Aus $g_1 \circ f = g_2 \circ f$ folgt $g_1 = g_2$.

Wir zeigen wechselseitig alle Äquivalenzen untereinander, um zu erkennen, wo das Auswahlaxiom benötigt wird.

Beweis. Aussage (i) ⇒ Aussage (ii)$^{\text{AoC}}$: Es sei f surjektiv. Wir definieren die gesuchte Rechtsinverse $h\colon Y \to X$, indem wir zu jedem $y \in Y$ irgendein Element aus der nichtleeren Menge

[82] Im Unterschied dazu liegen bei einer Familie (Definition 6.38) alle Funktionswerte in derselben Menge und unterliegen keiner weiteren Einschränkung.

$f^{-1}(\{y\})$ auswählen.[AoC] Mit anderen Worten, h ist eine Auswahlfunktion zur Menge $\{f^{-1}(\{y\}) \mid y \in Y\}$. Für diese Funktion h gilt in der Tat $f(h(y)) = y$ für alle $y \in Y$, also $f \circ h = \mathrm{id}_Y$.

Aussage (ii) $\Rightarrow$ Aussage (i): Da id_Y bijektiv ist, impliziert $f \circ h = \mathrm{id}_Y$ insbesondere (Lemma 6.19), dass f surjektiv ist.

Aussage (i) $\Rightarrow$ Aussage (iii): Es seien f surjektiv, Z eine Menge und $g_1, g_2 \colon Y \to Z$ Funktionen mit der Eigenschaft $g_1 \circ f = g_2 \circ f$. Angenommen, $g_1 \neq g_2$. Dann gibt es ein $y \in Y$ mit $g_1(y) \neq g_2(y)$. Da f surjektiv ist, existiert ein $x \in X$ mit $f(x) = y$. Es gilt also $g_1(f(x)) = g_2(f(x))$, also auch $g_1(y) = g_2(y)$, ein Widerspruch.

Aussage (iii) $\Rightarrow$ Aussage (i): Angenommen, f sei nicht surjektiv, und es sei $y_0 \in Y \setminus f(X)$. Wähle $Z \coloneqq \{0, 1\}$ und definiere $g_1, g_2 \colon Y \to Z$ durch

$$g_1(y) \coloneqq 0 \quad \text{und} \quad g_2(y) \coloneqq \begin{cases} 1 & \text{für } y = y_0, \\ 0 & \text{sonst.} \end{cases}$$

Dann haben wir $g_1 \circ f = g_2 \circ f$, aber $g_1 \neq g_2$.

Aussage (ii) $\Rightarrow$ Aussage (iii): Es seien Z eine Menge und $g_1, g_2 \colon Y \to Z$ beliebige Funktionen. Aus $g_1 \circ f = g_2 \circ f$ folgt $(g_1 \circ f) \circ h = (g_2 \circ f) \circ h$, also aufgrund der Assoziativität auch $g_1 \circ (f \circ h) = g_2 \circ (f \circ h)$, d. h., $g_1 = g_2$.

Aussage (iii) $\Rightarrow$ Aussage (ii)[AoC]: Diese Aussage folgt aus den bereits bewiesenen Implikationen Aussage (iii) $\Rightarrow$ Aussage (i) und Aussage (i) $\Rightarrow$ Aussage (ii)[AoC].[83] □

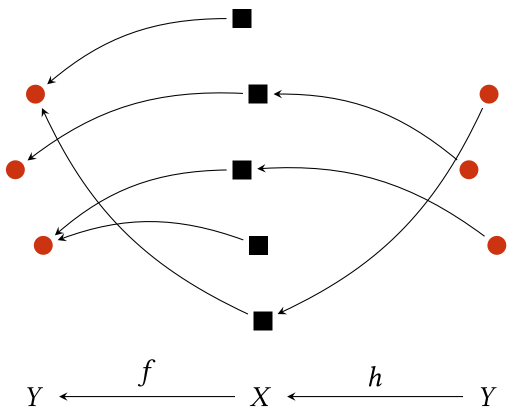


Abbildung 6.3.: Jede surjektive Funktion $f \colon X \to Y$ besitzt eine Rechtsinverse $h \colon Y \to X$ mit $f \circ h = \mathrm{id}_Y$ (Satz 6.44). Die Rechtsinverse ist injektiv und i. A. nicht eindeutig.

[83] Hier stellt sich die Frage, ob wir nicht vielleicht Aussage (iii) $\Rightarrow$ Aussage (ii) auch direkt beweisen könnten, ohne den Umweg über Aussage (i) zu gehen, um damit vielleicht die Abhängigkeit vom Auswahlaxiom loszuwerden. Das ist aber nicht möglich, denn könnten wir Aussage (iii) $\Rightarrow$ Aussage (ii) ohne das Auswahlaxiom beweisen, dann hätten wir zusammen mit Aussage (i) $\Rightarrow$ Aussage (iii) auch Aussage (i) $\Rightarrow$ Aussage (ii) ohne Auswahlaxiom bewiesen, was im Widerspruch zu Satz 6.46 steht.

Beispiel 6.45 (Rechtsinverse)**.**
Die Funktion $f\colon \mathbb{N} \to \mathbb{N}$ mit $f(n) = \frac{1}{2}n$ für $n \in \mathbb{N}$ gerade und $f(n) = \frac{1}{2}(n+1)$ für $n \in \mathbb{N}$ ungerade ist surjektiv (aber nicht bijektiv). Jede Funktion $h\colon \mathbb{N} \to \mathbb{N}$ mit $h(n) \in \{2\,n-1, 2\,n\}$ für $n \in \mathbb{N}$ ist eine Rechtsinverse von f. △

Das Auswahlaxiom hat eine ganze Menge äquivalenter, teilweise überraschender Charakterisierungen, von denen der nächste Satz (ohne Beweis) einige angibt.

Satz 6.46 (zum Auswahlaxiom äquivalente Aussagen)**.**
Folgende Aussagen sind in der Mengenlehre von Zermelo und Fraenkel äquivalent:

(i) Es gilt das Auswahlaxiom.

(ii) Es sei I eine Menge, und weiter sei A_i eine nichtleere Menge für jedes $i \in I$. Dann ist das kartesische Produkt $\bigtimes_{i \in I} A_i$ eine nichtleere Menge (Definition 6.42).

(iii) Jede Äquivalenzrelation besitzt ein Repräsentantensystem (Definition 5.21).

(iv) Jede surjektive Funktion besitzt eine Rechtsinverse (Satz 6.44).

(v) Es gilt der Wohlordnungssatz 6.47.

(vi) Es gilt das Lemma von Zorn 6.48.

Satz 6.47 (**Wohlordnungssatz**[84])**.**
Jede nichtleere Menge besitzt eine **Wohlordnung**, d. h., eine Totalordnung, bei der jede nichtleere Teilmenge ein kleinstes Element besitzt.[85]

Lemma 6.48 (**Lemma von Zorn**[86])**.**
Es sei X mit der Relation $\preccurlyeq$ eine halbgeordnete Menge. Weiter besitze jede totalgeordnete Teilmenge $A \subseteq X$ eine obere Schranke in X.[87] Dann existiert in X ein maximales Element.

Wir werden das Auswahlaxiom in Gestalt des Lemmas von Zorn 6.48 später noch verwenden (Satz 13.5) und aufgrund seiner Äquivalenz zum Auswahlaxiom auch dann durch die Markierung$^{\text{AoC}}$ darauf hinweisen.

Die Schwierigkeiten in der intuitiven Erfassung des Auswahlaxioms und des äquivalenten Wohlordnungssatzes 6.47 und des Lemmas von Zorn 6.48 werden in folgendem Zitat gut erfasst, das von dem Mathematiker Jerry Lloyd Bona stammt:

> „The Axiom of Choice is obviously true, the well-ordering theorem is obviously false; and who can tell about Zorn's Lemma?"

Ende der Vorlesung 7

[84] englisch: well-ordering theorem
[85] Beispielsweise ist die gewöhnliche Kleiner-Gleich-Relation eine Wohlordnung auf $\mathbb{N}$, aber nicht auf $\mathbb{Z}$, $\mathbb{Q}$ oder $\mathbb{R}$.
[86] englisch: Zorn's lemma
[87] X kann also nicht die leere Menge sein.

Kapitel 2 Algebraische Strukturen

In diesem Kapitel geht es um die grundlegenden algebraischen Strukturen, Abbildungen zwischen Strukturen und die in ihnen geltenden „Rechenregeln".

§ 7 Halbgruppen und Gruppen

Literatur: Beutelspacher, 2014, Kapitel 9; Deiser, 2024b, Kapitel 3.4; Fischer, Springborn, 2020, Kapitel 2.2

Definition 7.1 (Verknüpfung, Operation)**.**
Es sei M eine Menge. Eine Abbildung $\star\colon M \times M \to M$ heißt eine **(innere) Verknüpfung** oder **(innere) Operation** (englisch: (inner) operation) **auf** M. △

Wir schreiben $a \star b$ statt $\star(a, b)$ für $a, b \in M$.

Beispiel 7.2 (Verknüpfung)**.**

(i) Ist M endlich, so können wir eine Verknüpfung auf M mit Hilfe einer **Verknüpfungstafel** oder **Verknüpfungstabelle** (englisch: Cayley table) definieren. Beispielsweise sehen die Verknüpfungstafeln für die **Addition modulo** 2 (englisch: addition modulo 2) und die **Multiplikation modulo** 2 (englisch: multiplication modulo 2) wie folgt aus:

$$+_2\colon \{0,1\} \times \{0,1\} \to \{0,1\} \quad \text{mit der Verknüpfungstafel} \quad \begin{array}{c|cc} +_2 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

$$\cdot_2\colon \{0,1\} \times \{0,1\} \to \{0,1\} \quad \text{mit der Verknüpfungstafel} \quad \begin{array}{c|cc} \cdot_2 & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Beachte: Unsere Konvention ist, dass die Zeile das erste Argument (a) und die Spalte das zweite Argument (b) einer Verknüpfung $a \star b$ angibt.

Quizfrage 7.1: Wenn wir 0 mit *false* und 1 mit *true* identifizieren, welchen logischen Operationen (Junktoren) aus Definition 1.3 entsprechen dann $+_2$ und $\cdot_2$?

(ii) Die bekannten Rechenoperationen $+$ und $\cdot$ in $\mathbb{N}$

$$+\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N} \quad \text{mit } (a, b) \mapsto a + b$$
$$\cdot\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N} \quad \text{mit } (a, b) \mapsto a \cdot b$$

sind Verknüpfungen auf $\mathbb{N}$. Analoges gilt für die Mengen $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$, vgl. Anhang A.

(*iii*) Es seien X und Y Mengen, und auf Y sei die Verknüpfung $\star$ definiert. Dann können wir diese Verknüpfung auf die Menge der Funktionen $Y^X = \{f \mid f\colon X \to Y\}$ übertragen, indem wir sie punktweise anwenden:

$$\star\colon Y^X \times Y^X \to Y^X \quad \text{mit } f \star g \text{ definiert durch} \quad (f \star g)(x) \coloneqq f(x) \star g(x).$$

(*iv*) Beispielsweise übertragen sich die Verknüpfungen $+$ und $\cdot$ von der Menge $\mathbb{R}$ auf die Menge der Funktionen $\mathbb{R}^X = \{f \mid f\colon X \to \mathbb{R}\}$. Es sind also die punktweise Addition und die punktweise Multiplikation durch

$$\begin{aligned} +\colon \mathbb{R}^X \times \mathbb{R}^X \to \mathbb{R}^X \quad &\text{mit } f + g \text{ definiert durch} \quad (f + g)(x) \coloneqq f(x) + g(x), \\ \cdot\colon \mathbb{R}^X \times \mathbb{R}^X \to \mathbb{R}^X \quad &\text{mit } f \cdot g \text{ definiert durch} \quad (f \cdot g)(x) \coloneqq f(x) \cdot g(x), \end{aligned}$$

als Verknüpfungen auf der Menge der Funktionen $X \to \mathbb{R}$ definiert.

(*v*) Es sei X eine Menge und $X^X = \{f \mid f\colon X \to X\}$. Dann ist durch die Komposition

$$\circ\colon X^X \times X^X \to X^X \quad \text{mit } f \circ g \text{ definiert durch} \quad (f \circ g)(x) \coloneqq f(g(x))$$

eine Verknüpfung auf der Menge der Funktionen $X \to X$ definiert. △

§ 7.1 Halbgruppen

Definition 7.3 (Halbgruppe)**.**
Eine **Halbgruppe** (englisch: semigroup) $(H, \star)$ ist eine Menge H mit einer **assoziativen Verknüpfung** (englisch: associative operation) $\star$ auf H. Das heißt, es gilt $\star\colon H \times H \to H$ und

$$(a \star b) \star c = a \star (b \star c) \quad \text{für alle } a, b, c \in H. \tag{7.1}$$

△

Wegen der Assoziativität von $\star$ dürfen wir bei der Verknüpfung von drei oder mehr Elementen wie bei $a \star b \star c$ die Klammern weglassen.

Beispiel 7.4 (Halbgruppen, vgl. Beispiel 7.2 für die Verknüpfungen)**.**

(*i*) $(\mathbb{N}, +)$, $(\mathbb{N}_0, +)$, $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ sind Halbgruppen.

(*ii*) $(\mathbb{N}, \cdot)$, $(\mathbb{N}_0, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, \cdot)$ sind Halbgruppen.

(*iii*) $(\{0, 1\}, +_2)$ und $(\{0, 1\}, \cdot_2)$ aus Beispiel 7.2 sind Halbgruppen.

(*iv*) Es sei X eine Menge und $(H, \star)$ eine Halbgruppe. Dann ist $(H^X, \star)$ eine Halbgruppe. Die Assoziativität wird also auf die punktweise Verknüpfung vererbt.

(*v*) Insbesondere sind $(\mathbb{N}^X, +)$, $(\mathbb{N}_0^X, +)$, $(\mathbb{Z}^X, +)$, $(\mathbb{Q}^X, +)$, $(\mathbb{R}^X, +)$, $(\mathbb{C}^X, +)$ und $(\mathbb{N}^X, \cdot)$, $(\mathbb{N}_0^X, \cdot)$, $(\mathbb{Z}^X, \cdot)$, $(\mathbb{Q}^X, \cdot)$, $(\mathbb{R}^X, \cdot)$, $(\mathbb{C}^X, \cdot)$ Halbgruppen.

(*vi*) Es sei X eine Menge, dann ist $(X^X, \circ)$ eine Halbgruppe. Die Assoziativität der Komposition $\circ$ wurde in Lemma 6.17 gezeigt.

(*vii*) Ist X eine Menge, dann sind $(\mathcal{P}(X), \cap)$, $(\mathcal{P}(X), \cup)$ und $(\mathcal{P}(X), \triangle)$ Halbgruppen.

$(viii)$ Es sei Σ eine nichtleere Menge und $\Sigma^* \coloneqq \bigcup_{n \in \mathbb{N}_0} \Sigma^n$, also die Menge von Tupeln beliebiger Länge. Wir definieren eine Verknüpfung $\|$ auf Σ^* durch die **Konkatenation** von Tupeln:

$$(x_1, \ldots, x_n) \parallel (y_1, \ldots, y_m) \coloneqq (x_1, \ldots, x_n, y_1, \ldots, y_m).$$

Dann ist $(\Sigma^*, \circ)$ eine Halbgruppe.[1] △

Beispiel 7.5 (Gegenbeispiele)**.**
Keine Halbgruppen sind:

(i) $(\mathbb{N}, -)$, denn $-$ („Minus") ist keine Verknüpfung auf $\mathbb{N}$, da beispielsweise $1 - 1$ kein Wert in $\mathbb{N}$ zugeordnet ist.

(ii) $(\mathbb{Z}, -)$, denn $-$ („Minus") ist zwar eine Verknüpfung auf $\mathbb{Z}$, sie ist aber nicht assoziativ. Beispielsweise ist
$$3 - (2 - 1) = 2, \quad \text{aber} \quad (3 - 2) - 1 = 0.$$

(iii) $(\mathbb{N}, \wedge)$ mit $a \wedge b \coloneqq a^b$. Es ist zwar $\wedge\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ eine Verknüpfung, sie ist aber nicht assoziativ. Beispielsweise ist

$$2 \wedge (3 \wedge 2) = 2^9, \quad \text{aber} \quad (2 \wedge 3) \wedge 2 = 8^2.$$ △

Definition 7.6 (neutrales Element)**.**
Es sei $(H, \star)$ eine Halbgruppe. Ein Element $e \in H$ heißt **neutrales Element** (englisch: neutral element) von $(H, \star)$, wenn gilt:

$$e \star a = a \star e = a \quad \text{für alle } a \in H. \tag{7.2}$$

Falls in $(H, \star)$ ein neutrales Element existiert, dann heißt $(H, \star)$ auch ein **Monoid** (englisch: monoid). △

Beachte: Im Unterschied zu einer Halbgruppe ist ein Monoid immer nichtleer.

Lemma 7.7 (neutrale Elemente sind eindeutig)**.**
Es sei $(H, \star)$ eine Halbgruppe. Sind e_1 und e_2 beides neutrale Elemente von $(H, \star)$, dann gilt $e_1 = e_2$.

Beweis. Es gilt
$$e_1 = e_1 \star e_2 = e_2.$$ □

Beispiel 7.8 (Halbgruppen mit und ohne neutrale Elemente, vgl. Beispiel 7.4 zu Halbgruppen)**.**

(i) $(\mathbb{N}_0, +)$, $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ haben alle das neutrale Element 0.

(ii) $(\mathbb{N}, +)$ besitzt kein neutrales Element.

[1] Diese findet Anwendung bei der Definition formaler Sprachen in der Informatik. Dort ist Σ in der Regel endlich und heißt das **Alphabet** (englisch: alphabet) und Σ^* die **Kleenesche Hülle** (englisch: Kleene star) von Σ. Die Elemente von Σ^* heißen **Worte** über dem Alphabet Σ. Sie werden üblicherweise ohne die Klammern und Kommata notiert, also etwa $ab \circ ba = abba$.

(*iii*) $(\mathbb{N}, \cdot)$, $(\mathbb{N}_0, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, \cdot)$ haben alle das neutrale Element 1.

(*iv*) $(\{0, 1\}, +_2)$ aus Beispiel 7.2 besitzt das neutrale Element 0.

(*v*) $(\{0, 1\}, \cdot_2)$ aus Beispiel 7.2 besitzt das neutrale Element 1.

(*vi*) Es sei X eine Menge und $(H, \star)$ ein Monoid mit neutralem Element e. Dann ist $(H^X, \star)$ ein Monoid mit neutralem Element f, gegeben durch die konstante Funktion $f\colon X \to H$ mit $f(x) = e$ für alle $x \in X$.

(*vii*) $(\mathcal{P}(X), \cap)$ besitzt das neutrale Element X.

(*viii*) $(\mathcal{P}(X), \cup)$ besitzt das neutrale Element $\emptyset$.

(*ix*) $(\mathcal{P}(X), \triangle)$ besitzt das neutrale Element $\emptyset$.

(*x*) $(\Sigma^*, \circ)$ aus Beispiel 7.4 besitzt das neutrale Element $()$, genannt das **leere Tupel** (englisch: empty tuple) oder das **leere Wort** (englisch: empty word). △

Definition 7.9 (Links- und Rechtstranslation).
Es sei $(H, \star)$ eine Halbgruppe. Für festes $b \in H$ heißen die Abbildungen

$$\begin{aligned} {}_b\star\colon H \ni a \mapsto b \star a \in H \quad &\text{die } \textbf{Linkstranslation}^{2} \textbf{ um } b, && (7.3a)\\ \star_b\colon H \ni a \mapsto a \star b \in H \quad &\text{die } \textbf{Rechtstranslation}^{3} \textbf{ um } b. && (7.3b)\end{aligned}$$

△

In einer Verknüpfungstafel sehen wir die Bilder einer Linkstranslation in der zugehörigen Zeile (fest gewähltes erstes Element in der Verknüpfung) und die Bilder einer Rechtstranslation in der zugehörigen Spalte (fest gewähltes zweites Element in der Verknüpfung).

Beispiel 7.10 (Links- und Rechtstranslation).

(*i*) In der Halbgruppe $(\mathbb{R}, +)$ ist die Linkstranslation mit $b = \sqrt{2}$ gegeben durch die Abbildung $a \mapsto \sqrt{2} + a$. Sie ist wegen der Kommutativität von $+$ identisch zur Rechtstranslation mit b.

(*ii*) In der Halbgruppe $(\mathbb{R}^{\mathbb{R}}, \circ)$ ist die Linkstranslation mit g, definiert durch $g(x) = 2\,x$, gegeben durch

$$_g\circ\colon \mathbb{R}^{\mathbb{R}} \ni f \mapsto g \circ f \in \mathbb{R}^{\mathbb{R}}, \quad \text{wobei } (g \circ f)(x) = g(f(x)) = 2\,f(x) \text{ gilt,}$$

während die Rechtstranslation mit g gegeben ist durch

$$\circ_g\colon \mathbb{R}^{\mathbb{R}} \ni f \mapsto f \circ g \in \mathbb{R}^{\mathbb{R}}, \quad \text{wobei } (f \circ g)(x) = f(g(x)) = f(2\,x) \text{ ist.}$$ △

Quizfrage 7.2: Wie lässt sich der Begriff **neutrales Element** in einer Halbgruppe mit Hilfe der Begriffe **Linkststranslation** und **Rechtstranslation** ausdrücken?

Definition 7.11 (Unterhalbgruppe, Untermonoid).
Es sei $(H, \star)$ eine Halbgruppe.

(*i*) Eine Teilmenge $U \subseteq H$ heißt **abgeschlossen** (englisch: closed) bzgl. $\star$, wenn $\star\colon H \times H \to H$ eingeschränkt werden kann zu $\star_U\colon U \times U \to U$.[4] In diesem Fall heißt $\star_U$ die auf

[2] englisch: left translation, lateinisch: translatio: Übertragung, Verschiebung
[3] englisch: right translation
[4] Mit der Notation aus Bemerkung 7.20 können wir die Abgeschlossenheit von U auch als $U \star U \subseteq U$ schreiben.

U **induzierte (innere) Verknüpfung** (englisch: induced (inner) operation, lateinisch: inducere: hineinführen).

(*ii*) Eine bzgl. $\star$ abgeschlossene Teilmenge $U \subseteq H$ mit der Verknüpfung $\star_U$ heißt eine **Unterhalbgruppe** (englisch: subsemigroup) **von** $(H, \star)$.[5] Manchmal schreibt man dies als $(U, \star_U) \leqslant (H, \star)$.

(*iii*) Ist $(H, \star)$ ein Monoid mit neutralem Element e, dann heißt eine bzgl. $\star$ abgeschlossene Teilmenge $U \subseteq H$, die auch das neutrale Element e enthält, ein **Untermonoid** (englisch: submonoid) **von** $(H, \star)$.

(*iv*) Eine Unterhalbgruppe oder ein Untermonoid $(U, \star_U)$ von $(H, \star)$ heißt **echt** (englisch: proper subsemigroup, proper submonoid), wenn $U \subsetneq H$ gilt. △

Der Einfachheit halber werden wir in Zukunft statt $\star_U$ einfach $\star$ schreiben. Wir werden außerdem auch einfach die Menge U als Unterhalbgruppe bzw. Untermonoid von $(H, \star)$ bezeichnen, weil sich die Verknüpfung auf U ja durch Einschränkung der Verknüpfung $\star$ auf H ergibt.

Beispiel 7.12 (Unterhalbgruppe, Untermonoid)**.**

(*i*) Die geraden Zahlen bilden ein Untermonoid von $(\mathbb{Z}, +)$ mit neutralem Element 0.

(*ii*) Es seien a, b, c paarweise verschieden. $(\mathcal{P}(\{a, b\}), \cap)$ bildet zwar eine Unterhalbgruppe des Monoids $(\mathcal{P}(\{a, b, c\}), \cap)$, aber kein Untermonoid, denn das neutrale Element $\{a, b, c\}$ fehlt in $\mathcal{P}(\{a, b\})$. Vielmehr ist $(\mathcal{P}(\{a, b\}), \cap)$ ein Monoid mit einem anderen neutralen Element, nämlich $\{a, b\}$, siehe Beispiel 7.8. △

Bemerkung 7.13 („Unterhalbgruppe sein" und „Untermonoid sein" sind Ordnungsrelationen)**.**

(*i*) Die Relation „ist Unterhalbgruppe von" ist eine partielle Ordnung auf der Klasse aller Halbgruppen.

Insbesondere ist die Menge aller Unterhalbgruppen einer bestimmten Halbgruppe $(H, \star)$ durch die Unterhalbgruppenhalbordnung partiell geordnet. Diese Ordnung stimmt mit der Inklusionshalbordnung überein.[6]

(*ii*) Auch die Relation „ist Untermonoid von" ist eine partielle Ordnung auf der Klasse aller Monoide.

Insbesondere ist die Menge aller Untermonoide eines bestimmten Monoids $(H, \star)$ durch die Untermonoidhalbordnung partiell geordnet. Diese Ordnung stimmt mit der Inklusionshalbordnung überein. △

Definition 7.14 (invertierbare Elemente)**.**
Es sei $(H, \star)$ ein Monoid mit neutralem Element e. Ein Element $a \in H$ heißt **invertierbar** (englisch: invertible) oder eine **Einheit** (englisch: unit) von $(H, \star)$, wenn ein $a' \in H$ existiert mit

$$a \star a' = a' \star a = e. \tag{7.4}$$

[5] Da die Assoziativät durch die Einschränkung der Verknüpfung nicht verlorengeht, ist $(U, \star_U)$ wieder eine Halbgruppe mit der eingeschränkten Verknüpfung $\star_U$.

[6] Das heißt: Sind H_1, H_2 Unterhalbgruppen von H, und gilt $H_1 \subseteq H_2$, dann ist H_1 auch eine Unterhalbgruppe von H_2.

In diesem Fall heißt a' ein zu a **inverses Element** (englisch: inverse element) oder ein **Inverses** zu a. △

Lemma 7.15 (inverse Elemente sind eindeutig)**.**
Es sei $(H, \star)$ ein Monoid mit neutralem Element e. Ist $a \in H$ invertierbar und sind a_1' und a_2' beides Inverse zu a, dann gilt $a_1' = a_2'$.

Beweis. Es gilt

$$\begin{aligned} a_1' &= a_1' \star e && \text{denn } e \text{ ist das neutrale Element} \\ &= a_1' \star (a \star a_2') && \text{denn } a_2' \text{ ist invers zu } a \\ &= (a_1' \star a) \star a_2' && \text{wegen der Assoziativität von } \star \\ &= e \star a_2' && \text{denn } a_1' \text{ ist invers zu } a \\ &= a_2' && \text{denn } e \text{ ist das neutrale Element.} \end{aligned}$$ □

Beachte: Die Definition (7.4) besagt nicht nur, dass a' das Inverse zu a ist, sondern auch, dass a das Inverse zu a' ist. Die Invertierung ist also **involutorisch** (englisch: involutory), d. h., für alle invertierbaren $a \in H$ gilt

$$(a')' = a. \tag{7.5}$$

Quizfrage 7.3: Welches Element eines Monoids ist immer invertierbar? Was ist sein Inverses?

Lemma 7.16 (invertierbare Elemente bilden ein Untermonoid)**.**
Es sei $(H, \star)$ ein Monoid. Dann bildet die Teilmenge der invertierbaren Elemente

$$E \coloneqq \{a \in H \mid a \text{ ist invertierbar}\} \tag{7.6}$$

ein Untermonoid von $(H, \star)$. Sind $a, b \in E$ und a', b' die zugehörigen inversen Elemente in H, dann gilt für das zu $a \star b$ inverse Element

$$(a \star b)' = b' \star a'. \tag{7.7}$$

Beweis. Zunächst gilt für das neutrale Element $e \in E$ wegen $e \star e = e$, also $e' = e$. Es seien nun $a, b \in E$. Dann gilt

$$\begin{aligned} & (a \star b) \star (b' \star a') = a \star (b \star b') \star a' = a \star e \star a' = a \star a' = e \\ \text{und} \quad & (b' \star a') \star (a \star b) = b' \star (a' \star a) \star b = b' \star e \star b = b' \star b = e. \end{aligned}$$

Das zeigt, dass $a \star b$ wieder invertierbar und dass $b' \star a'$ das Inverse zu $a \star b$ ist, also (7.7). Das heißt aber auch, dass E bzgl. $\star$ abgeschlossen ist, also bildet E ein Untermonoid von $(H, \star)$. □

Beispiel 7.17 (invertierbare Elemente in Monoiden)**.**

(*i*) $(\mathbb{N}, +)$ besitzt kein neutrales Element, also können wir auch nicht von invertierbaren Elementen sprechen.

(*ii*) In $(\mathbb{N}_0, +)$ ist nur das neutrale Element 0 invertierbar.

(*iii*) In $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$ und $(\mathbb{C},+)$ sind alle Elemente invertierbar. Das Inverse von a wird mit $-a$ bezeichnet.

(*iv*) In $(\mathbb{N},\cdot)$ ist nur das neutrale Element 1 invertierbar.

(*v*) In $(\mathbb{Z},\cdot)$ sind nur 1 und -1 invertierbar. Beide sind zu sich selbst invers.

(*vi*) In $(\mathbb{Q},\cdot)$, $(\mathbb{R},\cdot)$ und $(\mathbb{C},\cdot)$ sind alle Elemente bis auf 0 invertierbar. Das Inverse von a wird mit a^{-1} oder $1/a$ oder $\frac{1}{a}$ bezeichnet. Allgemeiner steht $\frac{a}{b}$ für $a \cdot b^{-1} = b^{-1} \cdot a$, wobei $b \neq 0$ vorausgesetzt wird.[7]

(*vii*) In $(\{0,1\},+_2)$ aus Beispiel 7.2 sind beide Elemente invertierbar. Beide sind zu sich selbst invers.

(*viii*) In $(\{0,1\},\cdot_2)$ aus Beispiel 7.2 ist nur das neutrale Element 1 invertierbar.

(*ix*) In $(X^X,\circ)$ sind die invertierbaren Elemente genau die bijektiven Funktionen $X \to X$, also die invertierbaren Funktionen. △

Quizfrage 7.4: Was sind die invertierbaren Elemente in den Monoiden $(\mathcal{P}(X),\cap)$, $(\mathcal{P}(X),\cup)$ und $(\mathcal{P}(X),\triangle)$?

Lemma 7.18 (invertierbare Elemente können gekürzt werden)**.**
Es sei $(H,\star)$ ein Monoid. Für invertierbares $a \in H$ gelten die **Kürzungsregeln** (englisch: cancellation rules)

$$a \star b_1 = a \star b_2 \quad \Rightarrow \quad b_1 = b_2 \tag{7.8a}$$
$$b_1 \star a = b_2 \star a \quad \Rightarrow \quad b_1 = b_2 \tag{7.8b}$$

für beliebige $b_1, b_2 \in H$.

Beweis. Es gilt

$$\begin{aligned}
& a \star b_1 = a \star b_2 && \\
\Rightarrow\quad & a' \star (a \star b_1) = a' \star (a \star b_2) && \text{denn } a \text{ ist invertierbar} \\
\Rightarrow\quad & (a' \star a) \star b_1 = (a' \star a) \star b_2 && \text{wegen der Assoziativität von } \star \\
\Rightarrow\quad & e \star b_1 = e \star b_2 && \text{da } a' \text{ invers zu } a \text{ ist} \\
\Rightarrow\quad & b_1 = b_2 && \text{wegen der Eigenschaften von } e.
\end{aligned}$$

Die Aussage (7.8b) folgt analog. □

Lemma 7.19 (Nachweis von Inversen)**.**
Es sei $(H,\star)$ ein Monoid mit neutralem Element e und $a, b \in H$. Gilt $a \star b = e$ und ist a oder b bereits als invertierbar bekannt, dann sind a und b beide invertierbar und gegenseitig Inverse voneinander.

[7] Aus dem Symbol $\frac{a}{b}$ geht (anders als bei $a - b$) nicht hervor, ob $a \cdot b^{-1}$ oder $b^{-1} \cdot a$ gemeint ist. Wir verwenden daher das Symbol $\frac{a}{b}$ in allgemeinen, multiplikativ notierten Halbgruppen (Bemerkung 7.20) nur dann, wenn die Verknüpfung kommutativ ist (Definition 7.27), wenn also $a \cdot b^{-1} = b^{-1} \cdot a$ gilt.

Beachte: Wenn man also schon weiß, dass z. B. das Element a eines Monoids invertierbar ist, dann reicht es für den Nachweis, dass ein Element b das Inverse zu a ist, aus, a und b in einer der beiden Reihenfolgen miteinander zu verknüpfen

Beweis. Es sei a invertierbar und $a \star b = e$. Das inverse Element zu a sei a'. Dann gilt $a' = a' \star (a \star b) = (a' \star a) \star b = b$. Ist dagegen b invertierbar und $a \star b = e$ und ist b' das inverse Element zu b, dann gilt $b' = (a \star b) \star b' = a \star (b \star b') = a$. □

Bemerkung 7.20 (abkürzende Schreibweisen in Halbgruppen)**.**

(i) Wir haben als „neutrale" Notation der Verknüpfung einer Halbgruppe H das Symbol $\star$ gewählt, e als Bezeichnung für das neutrale Element und a' für das zu a inverse Element (sofern existent).

Für Teilmengen $A, B \subseteq H$ und $c \in H$ definieren wir die Mengenschreibweisen[8]

$$\begin{aligned} c \star A &\coloneqq \{c \star a \mid a \in A\}, \\ A \star c &\coloneqq \{a \star c \mid a \in A\}, \\ A \star B &\coloneqq \{a \star b \mid a \in A,\ b \in B\}. \end{aligned}$$

Ist $(H, \star)$ ein Monoid und sind alle Elemente in A invertierbar, so definieren wir auch

$$A' \coloneqq \{a' \mid a \in A\}.$$

(ii) Bezeichnen wir dagegen die Verknüpfung einer Halbgruppe H als „Addition" und notieren sie als „+" o. ä. (**additive Notation**, englisch: additive notation), so nennen wir ein eventuell existierendes neutrales Element auch **Nullelement** (englisch: additive identity, zero element) oder **Null** (englisch: zero), geschrieben als „0_H" oder einfach 0.

Für $n \in \mathbb{N}$ und $a \in H$ ist $n\, a$ eine Abkürzung für $a + \cdots + a$ (n-mal).[9] (**Quizfrage 7.5:** Warum ist $a + \cdots + a$ auch ohne Setzen von Klammern wohldefiniert?)

Besitzt $(H, +)$ das neutrale Element 0_H, so definieren wir auch $0\, a \coloneqq 0_H$.

Ist weiter $a \in H$ invertierbar, so notieren wir das Inverse als $-a$. Dann ist auch $n\, a$ invertierbar für $n \in \mathbb{N}_0$, und wir setzen $(-n)\, a \coloneqq -(n\, a) = n\, (-a)$. (**Quizfrage 7.6:** Warum gilt $-(n\, a) = n\, (-a)$?) Insbesondere ist $(-1)\, a \coloneqq -a$ und $(-0)\, a \coloneqq -(0\, a) = -0_H = 0_H$.

Es gilt

$$n\, (m\, a) = (n \cdot m)\, a \quad \text{und} \quad (n + m)\, a = n\, a + m\, a \tag{7.9}$$

für alle $n, m \in \mathbb{N}$ bzw. $n, m \in \mathbb{N}_0$ bzw. $n, m \in \mathbb{Z}$. (**Quizfrage 7.7:** Ist Ihnen in (7.9) bei jeder Operation klar, worum es sich handelt?)

Die Bezeichnung $a - b$ steht für $a + (-b)$, vorausgesetzt, b ist invertierbar.

Für Teilmengen $A, B \subseteq H$ und $c \in H$ definieren wir die Mengenschreibweisen

$$c + A \coloneqq \{c + a \mid a \in A\},$$

[8] Im Fall $A = \emptyset$ sind auch die Mengen $c \star A$, $A \star c$, $A \star B$ und A' leer. Auch im Fall $B = \emptyset$ ist $A \star B$ die leere Menge.

[9] Zur Verdeutlichung notieren wir vorübergehend alle Vorfaktoren in dieser Farbe. Diese sind keine Elemente der Halbgruppe, sondern Elemente aus $\mathbb{N}$ bzw. $\mathbb{Z}$.

$$A + c \coloneqq \{a + c \mid a \in A\},$$
$$A + B \coloneqq \{a + b \mid a \in A,\ b \in B\}.$$

Ist $(H, +)$ ein Monoid und sind alle Elemente in A invertierbar, so definieren wir auch

$$-A \coloneqq \{-a \mid a \in A\}.$$

(iii) Bezeichnen wir die Verknüpfung einer Halbgruppe H als „Multiplikation" und notieren sie als „$\cdot$" o. ä. (**multiplikative Notation**, englisch: multiplicative notation), so nennen wir ein eventuell existierendes neutrales Element auch **Einselement** (englisch: multiplicative identity, one element, unit element) oder **Eins** (englisch: one, unit), geschrieben als „1_H" oder einfach 1.

Für $n \in \mathbb{N}$ und $a \in H$ ist a^n eine Abkürzung für $a \cdot \cdots \cdot a$ (n-mal).

Besitzt $(H, \cdot)$ das neutrale Element 1_H, so definieren wir auch $a^0 \coloneqq 1_H$.

Ist weiter $a \in H$ invertierbar, so notieren wir das Inverse als a^{-1}. Dann ist auch a^n invertierbar für $n \in \mathbb{N}_0$, und wir setzen $a^{-n} \coloneqq (a^n)^{-1} = (a^{-1})^n$. Insbesondere ist $a^{-0} \coloneqq (a^0)^{-1} = 1_H^{-1} = 1_H$.

Es gilt

$$(a^n)^m = a^{n \cdot m} \quad \text{und} \quad a^{n+m} = a^n \cdot a^m \tag{7.10}$$

für alle $n, m \in \mathbb{N}$ bzw. $n, m \in \mathbb{N}_0$ bzw. $n, m \in \mathbb{Z}$.

Für Teilmengen $A, B \subseteq H$ und $c \in H$ definieren wir die Mengenschreibweisen

$$c \cdot A \coloneqq \{c \cdot a \mid a \in A\},$$
$$A \cdot c \coloneqq \{a \cdot c \mid a \in A\},$$
$$A \cdot B \coloneqq \{a \cdot b \mid a \in A,\ b \in B\}.$$

Ist $(H, \cdot)$ ein Monoid und sind alle Elemente in A invertierbar, so definieren wir auch

$$A^{-1} \coloneqq \{a^{-1} \mid a \in A\}.$$

(iv) Bezeichnen wir die Verknüpfung einer Halbgruppe H als „Komposition" und notieren sie als „$\circ$" o. ä. (**Kompositionsnotation**, englisch: compositional notation), so nennen wir ein eventuell existierendes neutrales Element auch **Identität** (englisch: identity), geschrieben als „id". In diesem Fall benutzen wir dieselbe Notation wie im Fall multiplikativer Notation:

Für $n \in \mathbb{N}$ und $a \in H$ ist a^n eine Abkürzung für $a \circ \cdots \circ a$ (n-mal).

Besitzt $(H, \circ)$ das neutrale Element id, so definieren wir auch $a^0 \coloneqq \mathrm{id}$.

Ist weiter $a \in H$ invertierbar, so notieren wir das Inverse als a^{-1}. Dann ist auch a^n invertierbar für $n \in \mathbb{N}_0$, und wir setzen $a^{-n} \coloneqq (a^n)^{-1} = (a^{-1})^n$. Insbesondere ist $a^{-0} \coloneqq (a^0)^{-1} = \mathrm{id}^{-1} = \mathrm{id}$.

Es gilt

$$(a^n)^m = a^{n \cdot m} \quad \text{und} \quad a^{n+m} = a^n \circ a^m \tag{7.11}$$

für alle $n, m \in \mathbb{N}$ bzw. $n, m \in \mathbb{N}_0$ bzw. $n, m \in \mathbb{Z}$.

Für Teilmengen $A, B \subseteq H$ und $c \in H$ definieren wir die Mengenschreibweisen

$$\begin{aligned} c \circ A &\coloneqq \{c \circ a \,|\, a \in A\}, \\ A \circ c &\coloneqq \{a \circ c \,|\, a \in A\}, \\ A \circ B &\coloneqq \{a \circ b \,|\, a \in A,\ b \in B\}. \end{aligned}$$

Ist $(H, \circ)$ ein Monoid und sind alle Elemente in A invertierbar, so definieren wir auch

$$A^{-1} \coloneqq \{a^{-1} \,|\, a \in A\}. \qquad \triangle$$

Ende der Vorlesung 8

Ende der Woche 4

§ 7.2 Gruppen

Definition 7.21 (Gruppe).
Ein Monoid $(H, \star)$ heißt eine **Gruppe** (englisch: group), wenn jedes Element aus H ein Inverses besitzt. △

Beachte: Es gilt also: $(G, \star)$ Gruppe $\Rightarrow$ $(G, \star)$ Monoid $\Rightarrow$ $(G, \star)$ Halbgruppe.

Beispiel 7.22 (Gruppen und Gegenbeispiele, vgl. Beispiel 7.17 zu Monoiden und ihren invertierbaren Elementen).

(i) $(\mathbb{N}_0, +)$ ist ein Monoid, aber keine Gruppe, da nur das neutrale Element 0 invertierbar ist.

(ii) $(\mathbb{Z}, +)$ ist eine Gruppe mit neutralem Element 0. Das Inverse zu $a \in \mathbb{Z}$ ist $-a \in \mathbb{Z}$, denn $a + (-a) = 0 = (-a) + a$. Dasselbe gilt für $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$.

(iii) $(\mathbb{Z}, \cdot)$ ist ein Monoid, aber keine Gruppe, da nur 1 und -1 invertierbar sind.

(iv) $(\mathbb{Q}_{\neq 0}, \cdot)$ ist eine Gruppe mit neutralem Element 1. Das Inverse zu $a \in \mathbb{Q}_{\neq 0}$ ist $a^{-1} = 1/a \in \mathbb{Q}_{\neq 0}$. Dasselbe gilt für $(\mathbb{R}_{\neq 0}, \cdot)$ und $(\mathbb{C}_{\neq 0}, \cdot)$.

(v) Für $m \in \mathbb{N}$ bildet die Menge $\mathbb{Z}_m \coloneqq \{0, 1, \ldots, m-1\}$ mit der Verknüpfung $+_m$ eine Gruppe. Dabei ist die **Addition modulo m** (englisch: addition modulo m) $+_m$ definiert als[10]

$$\begin{aligned} a +_m b \coloneqq{}& \begin{cases} a + b, & \text{falls } a + b \leqslant m - 1 \\ a + b - m, & \text{falls } a + b \geqslant m \end{cases} \\ ={}& \text{der natürliche Repräsentant von } a + b \text{ in der Restklasse } [a + b] \text{ modulo } m \\ ={}& \text{Rest von } a + b \text{ bei ganzzahliger Division durch } m. \end{aligned} \tag{7.12}$$

Diese Gruppe heißt die **additive Gruppe von $\mathbb{Z}$ modulo m** (englisch: additive group of $\mathbb{Z}$ modulo m), geschrieben $(\mathbb{Z}_m, +_m)$.

Den Fall $m = 2$ kennen wir bereits als $(\{0, 1\}, +_2)$ aus Beispiel 7.2.

[10] Beispielsweise ist $3 +_6 5 = 2$, weil $3 + 5 = 8$ ist und $8 \overset{6}{\equiv} 2$ gilt.

(*vi*) Für $m \in \mathbb{N}$ bildet die Menge $\mathbb{Z}_m$ mit der Verknüpfung $\cdot_m$ ein Monoid. Dabei ist $\cdot_m$ die **Multiplikation modulo** m (englisch: multiplication modulo m) definiert als[11]

$$\begin{aligned} a \cdot_m b &\coloneqq \text{der natürliche Repräsentant von } a \cdot b \text{ in der Restklasse } [a \cdot b] \text{ modulo } m \\ &= \text{Rest von } a \cdot b \text{ bei ganzzahliger Division durch } m. \end{aligned} \tag{7.13}$$

Dieses Monoid heißt das **multiplikative Monoid von $\mathbb{Z}$ modulo m** (englisch: multiplicative monoid of $\mathbb{Z}$ modulo m), geschrieben $(\mathbb{Z}_m, \cdot_m)$.

$(\mathbb{Z}_m, \cdot_m)$ ist genau dann eine Gruppe, wenn $m = 1$ ist, also wenn $\mathbb{Z}_m = \{0\}$ gilt.[12]

Den Fall $m = 2$ kennen wir bereits als $(\{0, 1\}, \cdot_2)$ aus Beispiel 7.2.

(*vii*) Es sei X eine Menge und $(G, \star)$ eine Gruppe. Dann ist $(G^X, \star)$ eine Gruppe.

(*viii*) Insbesondere ist $(\mathbb{R}^X, +)$ eine Gruppe.

(*ix*) $(\mathbb{R}^X, \cdot)$ ist keine Gruppe, wenn $X \neq \emptyset$ ist, da die Funktionen, die irgendwo den Wert 0 annehmen, keine invertierbaren Elemente sind. $((\mathbb{R}_{\neq 0})^X, \cdot)$ ist jedoch für jede Menge X eine Gruppe.

(*x*) $(X^X, \circ)$ ist keine Gruppe, sobald X zwei oder mehr Elemente enthält, denn dann gibt es Funktionen $X \to X$, die nicht bijektiv sind. Wenn X jedoch null- oder einelementig ist, dann ist $(X^X, \circ)$ eine Gruppe.

(*xi*) Die **Kleinsche Vierergruppe** (englisch: Klein four-group) K_4 ist eine kommutative Gruppe mit vier Elementen und der folgenden Verknüpfungstafel:

$\circ$	e	h	v	r
e	e	h	v	r
h	h	e	r	v
v	v	r	e	h
r	r	v	h	e

Das neutrale Element wird hier als e bezeichnet. Jedes Element ist selbstinvers. Die Gruppe kann verstanden werden als die **Symmetriegruppe** (englisch: symmetry group) eines Rechtecks $\begin{smallmatrix} D & & C \\ & \square & \\ A & & B \end{smallmatrix}$, also als die Menge der Abbildungen eines Rechtecks auf sich selbst. Dabei steht

e	für die identische Abbildung	$\begin{smallmatrix} D & & C \\ & \square & \\ A & & B \end{smallmatrix}$
h	für die horizontale Spiegelung	$\begin{smallmatrix} C & & D \\ & \square & \\ B & & A \end{smallmatrix}$
v	für die vertikale Spiegelung	$\begin{smallmatrix} A & & B \\ & \square & \\ D & & C \end{smallmatrix}$
r	für die Drehung um $180°$	$\begin{smallmatrix} B & & A \\ & \square & \\ C & & D \end{smallmatrix}$.

△

Lemma 7.23 (Einheitengruppe, vgl. Lemma 7.16).

[11] Beispielsweise ist $3 \cdot_6 5 = 3$, weil $3 \cdot 5 = 15$ ist und $15 \overset{6}{\equiv} 3$ gilt.

[12] In diesem Fall ist $(\mathbb{Z}_1, \cdot_1)$ isomorph (Definition 8.1) zu $(\mathbb{Z}_1, +_1)$, also abgesehen von der Notation dieselbe Gruppe.

Es sei $(H, \star)$ ein Monoid. Dann ist das Untermonoid der invertierbaren Elemente (Lemma 7.16)

$$E \coloneqq \{a \in H \mid a \text{ ist invertierbar}\} \tag{7.14}$$

eine Gruppe, genannt die **Einheitengruppe** (englisch: unit group, group of units) von $(H, \star)$.

Beweis. Wir wissen nach Lemma 7.16 bereits, dass E ein Monoid ist (nämlich ein Untermonoid von H). Per Definition sind alle Elemente von E invertierbar, also ist E eine Gruppe. □

Beispiel 7.24 (Einheitengruppe).

(*i*) $(\mathbb{Q}_{\neq 0}, \cdot)$, $(\mathbb{R}_{\neq 0}, \cdot)$ bzw. $(\mathbb{C}_{\neq 0}, \cdot)$ sind die Einheitengruppen der Monoide $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$ bzw. $(\mathbb{C}, \cdot)$.

(*ii*) Es sei X eine Menge und $(H, \star)$ ein Monoid. Die Einheitengruppe von $(H^X, \star)$ besteht genau aus denjenigen Funktionen $X \to H$, die nur Funktionswerte in der Einheitengruppe von H annehmen. (**Quizfrage 7.8:** Wie sieht das Inverse eines Elements der Einheitengruppe von H^X aus?)

(*iii*) Es sei X eine Menge. Die Einheitengruppe des Monoids $(X^X, \circ)$ (Beispiel 7.22) besteht genau aus den invertierbaren (bijektiven) Funktionen $X \to X$. △

Quizfrage 7.9: Können Sie die Multiplikationstabellen für $(\mathbb{Z}_m, \cdot_m)$ im Fall $m = 5$ und $m = 8$ aufstellen (Beispiel 7.22)? Haben Sie eine Vermutung, was die Einheitengruppe von $(\mathbb{Z}_m, \cdot_m)$ für $m \in \mathbb{N}$ ist, also welche Elemente in $(\mathbb{Z}_m, \cdot_m)$ invertierbar sind?

Das folgende Lemma gibt mit Hilfe von Links- und Rechtstranslationen eine notwendige und eine hinreichende Bedingung dafür an, wann eine Halbgruppe sogar eine Gruppe ist.

Lemma 7.25 (Gruppenkriterium mit Links- und Rechtstranslationen („**Sudoku-Kriterium**“)).

(*i*) Ist $(G, \star)$ eine Gruppe, so sind die Links- und Rechtstranslationen $\star_a$ und $_a\star$ für alle $a \in G$ bijektive Abbildungen $G \to G$.

(*ii*) Ist $(H, \star)$ eine nichtleere Halbgruppe und gilt für alle $a \in H$, dass die Links- und Rechtstranslationen $\star_a$ und $_a\star$ surjektive Abbildungen sind, dann ist $(H, \star)$ eine Gruppe.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

Beispiel 7.26 (Sudoku-Kriterium).
Gegeben ist die Menge $M \coloneqq$ {♡, ⊠, ◐, ⊗, ⬙, ⊳} mit zwei Verknüpfungen $\star$ und □ wie folgt:

$\star$	⊠	◐	⊗	⬙	♡	⊳
⬙	◐	⊳	♡	⊗	⬙	⊠
⊠	♡	⊗	◐	⊳	⊠	⬙
♡	⊠	◐	⊗	⬙	♡	⊳
◐	⬙	♡	⊳	⊠	◐	⊗
⊗	⊳	⊠	⬙	♡	⊗	◐
⊳	⊗	⬙	⊠	◐	⊳	♡

□	⊠	◐	⊗	⬙	♡	⊳
⬙	⬙	◐	♡	⬙	♡	◐
⊠	⊠	◐	⊗	⬙	♡	⊳
♡	♡	♡	♡	♡	♡	♡
◐	◐	⬙	♡	◐	♡	⬙
⊗	⊗	◐	⊠	♡	♡	⊗
⊳	⊳	⬙	⊗	◐	♡	⊠

Beide Verknüpfungen sind assoziativ, d. h., $(M, \star)$ und $(M, \square)$ sind beides Halbgruppen. Die Funktionswerte der Linkstranslationen finden wir in den Zeilen der jeweiligen Verknüpfungstafel, während sich die Funktionswerte der Rechtstranslationen in den Spalten wiederfinden.

Bei $(M, \star)$ handelt es sich tatsächlich um eine Gruppe, denn alle Linkstranslationen und alle Rechtstranslationen sind surjektiv. Das erkennen wir daran, dass in jeder Zeile und in jeder Spalte der Verknüpfungstafel alle Elemente der Menge M vorkommen.[13] Das neutrale Element ist $\heartsuit$, denn die entsprechende Zeile ist identisch mit der Zeile aus dem Tabellenkopf.

$(M, \square)$ ist dagegen keine Gruppe, da beispielsweise die Linkstranslation (Zeile) mit $\heartsuit$ nicht surjektiv ist. △

Definition 7.27 (kommutative Verknüpfung, Halbgruppe, Monoid, Gruppe)**.**

- (i) Es sei M eine Menge mit einer Verknüpfung $\star$. Die Elemente $a, b \in M$ **vertauschen** oder **kommutieren** (englisch: commute) bzgl. der Verknüpfung $\star$, wenn $a \star b = b \star a$ gilt.
- (ii) Die Verknüpfung $\star$ auf der Menge M heißt **kommutativ** (englisch: commutative) oder **abelsch** (englisch: Abelian), wenn $a \star b = b \star a$ für alle $a, b \in M$ gilt.
- (iii) Eine Halbgruppe bzw. ein Monoid bzw. eine Gruppe $(H, \star)$ heißt **kommutativ** oder **abelsch**, wenn die Verknüpfung $\star$ kommutativ ist. △

Bemerkung 7.28 (zur abkürzenden Notation in Halbgruppen (Bemerkung 7.20))**.**

- (i) Es ist üblich, die additive Notation (Bemerkung 7.20) nur für kommutative Verknüpfungen zu verwenden. Nur im kommutativen Fall verwenden wir auch das Summenzeichen, etwa in den Ausdrücken
$$\sum_{j=1}^{n} a_j \quad \text{oder} \quad \sum_{a \in A} a.$$
- (ii) Multiplikativ notierte Verknüpfungen können kommutativ oder nicht kommutativ sein. Nur im kommutativen Fall verwenden wir auch das Produktzeichen, etwa in den Ausdrücken
$$\prod_{j=1}^{n} a_j \quad \text{oder} \quad \prod_{a \in A} a.$$
- (iii) Als Komposition notierte Verknüpfungen sind i. d. R. nicht kommutativ. △

Beispiel 7.29 (kommutative Halbgruppen und Gruppen, vgl. Beispiele 7.4 und 7.22)**.**

Die Verknüpfungen „$+$“ und „$\cdot$“ auf $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sind kommutativ. Beispielsweise ist also $(\mathbb{N}, +)$ eine kommutative Halbgruppe, $(\mathbb{N}_0, +)$ ein kommutatives Monoid, und $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ sind kommutative Gruppen. △

[13]Da M endlich ist, kommen die Elemente von M in jeder Zeile und jeder Spalte jeweils genau einmal vor. Die Translationen sind also sogar alle bijektiv, vgl. Satz 6.35.

§ 7.3 Die symmetrische Gruppe

Definition 7.30 (symmetrische Gruppe)**.**
Es sei X eine Menge und $S(X) \coloneqq \{f : X \to X \mid f \text{ ist bijektiv}\}$. Dann heißt $(S(X), \circ)$ die **symmetrische Gruppe** (englisch: symmetric group) auf X. Jedes Element von $S(X)$ heißt eine **Permutation** (englisch: permutation, lateinisch: permutare: vertauschen) von X.

Ist $X = [\![1, n]\!]$ für $n \in \mathbb{N}_0$, so schreiben wir statt $S([\![1, n]\!])$ auch S_n und sprechen von der **symmetrischen Gruppe vom Grad** n (englisch: symmetric group of degree n). Jedes $\sigma \in S_n$ heißt eine **Permutation** (englisch: permutation) von $[\![1, n]\!]$. △

Beachte: Nach Beispiel 7.24 ist $S(X)$ tatsächlich eine Gruppe, nämlich die Einheitengruppe von $(X^X, \circ)$. Das neutrale Element ist id_X. Wenn X drei oder mehr Elemente enthält, dann ist $(S(X), \circ)$ nicht kommutativ, ansonsten kommutativ.

Im Folgenden werden wir nur noch symmetrische Gruppen S_n für $n \in \mathbb{N}_0$ betrachten. Eine Permutation $\sigma \in S_n$ können wir z. B. in der Form

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

notieren. Die Anzahl der Elemente von S_n für $n \in \mathbb{N}_0$ ist gleich $n!$ („n Fakultät"). Das stimmt auch für $n = 0$, denn es gilt $0! = 1$, und die einzige Permutation ist die leere Permutation.

Beispiel 7.31 (symmetrische Gruppe vom Grad 3)**.**
Die symmetrische Gruppe S_3 hat $3! = 6$ Elemente:

$$\sigma_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \text{(Drehungen)},$$
$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{(Spiegelungen)}.$$

Sie lassen sich identifizieren mit den Kongruenzabbildungen, die ein gleichseitiges Dreieck mit den Eckpunkten 1, 2 und 3 auf sich selbst überführen, vgl. Abbildung 7.1. Wegen

$$\begin{aligned} \sigma_4 \circ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \sigma_2 \\ \sigma_3 \circ \sigma_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \sigma_1 \end{aligned}$$

ist S_3 wie oben behauptet tatsächlich nicht kommutativ. △

Definition 7.32 (Transposition)**.**
Eine Permutation $\sigma \in S_n$, $n \in \mathbb{N}_0$, heißt eine **Transposition** (englisch: transposition, lateinisch: transponere: umstellen), wenn es Zahlen $i, j \in [\![1, n]\!]$ mit $i \neq j$ gibt, sodass gilt:

$$\sigma(k) = \begin{cases} j & \text{für } k = i, \\ i & \text{für } k = j, \\ k & \text{sonst.} \end{cases} \tag{7.15}$$

Wir notieren σ dann auch als $\tau(i, j)$. △

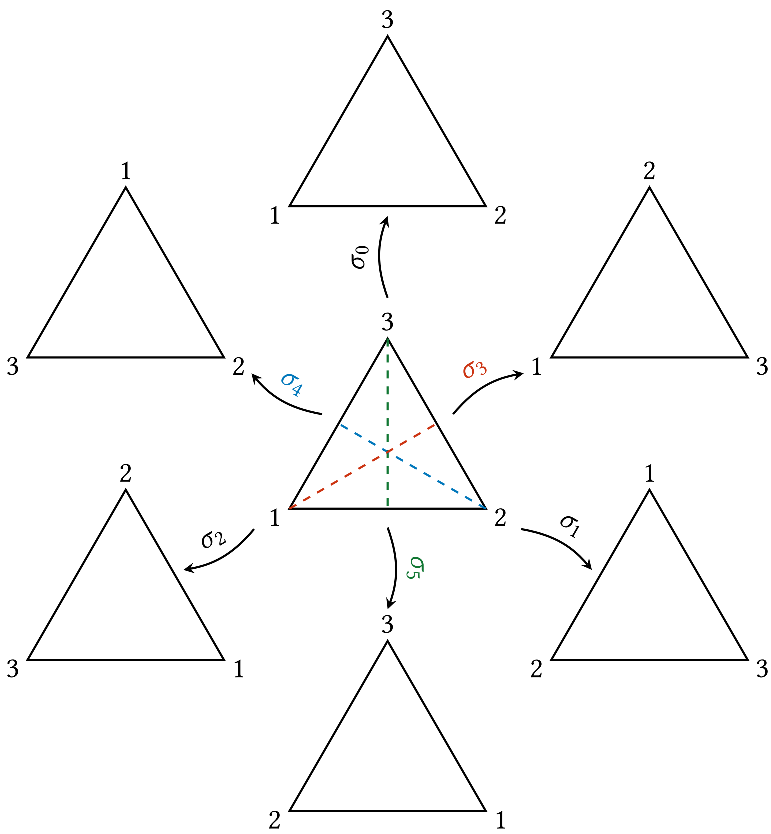


Abbildung 7.1.: Illustration der Elemente der symmetrischen Gruppe S_3, siehe Beispiel 7.31. Die Permutationen σ_0, σ_1 und σ_2 entsprechenden Drehungen. Die Permutationen σ_3, σ_4 und σ_5 sind genau die Transpositionen. Sie entsprechenden den Spiegelungen um die farbig eingezeichneten Achsen.

Eine Transposition von links vertauscht also genau zwei verschiedene Elemente von $[\![1, n]\!]$ und lässt den Rest unverändert. (Daher gibt es Transpositionen nur im Fall $n \geqslant 2$.) Offenbar gilt für jede Transposition

$$\tau^2 = \tau \circ \tau = \mathrm{id}, \quad \text{also } \tau^{-1} = \tau. \tag{7.16}$$

Transpositionen sind also Involutionen (Definition 6.18) und damit selbstinverse Elemente der Gruppe S_n.

(**Quizfrage 7.10:** Wieviele verschiedene Transpositionen gibt es in S_n?)

(**Quizfrage 7.11:** Können Sie eine Vermutung anstellen, welche Permutationen von S_n neben den Transpositionen noch selbstinvers sind?)

Satz 7.33 (Darstellung von Permutationen als Komposition von Transpositionen)**.**
Es sei $n \in \mathbb{N}_0$. Jede Permutation $\sigma \in S_n$ lässt sich als Komposition von $0 \leqslant r \leqslant n-1$ Transpositionen schreiben (bzw. $r = 0$ im Fall $n = 0$).[14]

[14]Insbesondere bildet die Menge der Transpositionen also ein Erzeugendensystem von S_n.

Beweis. Wir zeigen die Behauptung für $n \in \mathbb{N}_0$ durch vollständige Induktion. Induktionsanfang: Das einzige Element von

$$S_0 = \{\emptyset \to \emptyset\}$$

ist eine Komposition von $r = 0$ Transpositionen. Ebenso ist das einzige Element von

$$S_1 = \{\mathrm{id}_{\{1\}}\}$$

eine Komposition von $r = 0$ Transpositionen.

Induktionsannahme: Die Behauptung sei für $n \in \mathbb{N}$ bereits bewiesen. Induktionsschritt: Wir betrachten eine Permutation $\sigma \in S_{n+1}$.

Fall 1: Falls $\sigma(n+1) = n+1$ gilt, dann gilt für die Einschränkung $\widehat{\sigma} \colon [\![1,n]\!] \to [\![1,n]\!]$ von σ die Eigenschaft $\widehat{\sigma} \in S_n$. Aufgrund der Induktionsannahme besitzt $\widehat{\sigma}$ die Darstellung $\widehat{\sigma} = \tau_1 \circ \cdots \circ \tau_r$ mit $0 \leqslant r \leqslant n-1$ mit Transpositionen τ_i auf S_n. Setzen wir diese Transpositionen durch $n+1 \mapsto n+1$ zu Transpositionen auf S_{n+1} fort, die wir weiterhin mit τ_i bezeichnen, so ergibt sich die Darstellung $\sigma = \tau_1 \circ \cdots \circ \tau_r$.

Fall 2: Falls $\sigma(n+1) = m$ für ein $1 \leqslant m \leqslant n$ gilt, dann betrachte die Transposition $\tau(m, n+1) \in S_{n+1}$. Für $\widetilde{\sigma} \coloneqq \tau(m, n+1) \circ \sigma \in S_{n+1}$ gilt dann $\widetilde{\sigma}(n+1) = n+1$. Aufgrund von Fall 1 gilt $\widetilde{\sigma} = \tau_1 \circ \cdots \circ \tau_r$ mit $0 \leqslant r \leqslant n-1$. Die Behauptung folgt jetzt aus der Darstellung $\sigma = \tau(m, n+1) \circ \widetilde{\sigma} = \tau(m, n+1) \circ \tau_1 \circ \cdots \circ \tau_r$. □

Beispiel 7.34 (Darstellung von Permutationen als Komposition von Transpositionen)**.**
Wir können die Zerlegung einer Permutation $\sigma \in S_n$ bestimmen, indem wir die Bilder durch wiederholte Anwendung von Transpositionen z. B. von hinten nach vorne in die richtige Reihenfolge bringen, etwa:

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} &\xrightarrow{\tau(4,1)\circ} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} && \text{die 4 ist jetzt an der richtigen Stelle} \\
&\xrightarrow{\tau(3,1)\circ} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} && \text{auch die 3 ist jetzt an der richtigen Stelle} \\
&\xrightarrow{\tau(2,1)\circ} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} && \text{auch die 1 und die 2 sind jetzt an der richtigen Stelle.}
\end{aligned}$$

Das heißt also,

$$\tau(2,1) \circ \tau(3,1) \circ \tau(4,1) \circ \sigma = \mathrm{id} \quad \text{oder aber} \quad \sigma = \tau(4,1) \circ \tau(3,1) \circ \tau(2,1). \qquad \triangle$$

Man kann allgemein zeigen, dass genau solche zyklischen Vertauschungen σ wie in Beispiel 7.34 nicht mit weniger als $r = n-1$ (hier also $r = 3$) Permutationen dargestellt werden können. Die obere Schranke für die benötigte Anzahl an Transpositionen aus Satz 7.33 ist also scharf.

Die Darstellung einer Permutation als Komposition von Transpositionen ist nicht eindeutig. Jedoch ist die Anzahl der benötigten Transpositionen entweder immer gerade oder immer ungerade, wie wir gleich beweisen werden (Folgerung 7.41).

Definition 7.35 (Fehlstand, Signum einer Permutation)**.**
Es sei $n \in \mathbb{N}_0$ und σ eine Permutation in S_n.

(*i*) Ein Indexpaar $(i, j) \in [\![1, n]\!]^2$ heißt ein **Fehlstand** (englisch: inversion) von σ, wenn $i < j$ und $\sigma(i) > \sigma(j)$ gilt.

(*ii*) Die Zahl[15]

$$\operatorname{sgn} \sigma \coloneqq \prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} \tag{7.17}$$

heißt das **Signum** (englisch: sign, lateinisch: signum: Zeichen) von σ. △

Beispiel 7.36 (Fehlstand, Signum einer Permutation)**.**
Die Permutation

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

von S_3 hat genau zwei Fehlstände, nämlich $(1, 3)$ und $(2, 3)$. Es gilt

$$\begin{aligned} \operatorname{sgn} \sigma_1 &= \frac{\sigma(2) - \sigma(1)}{2 - 1} \frac{\sigma(3) - \sigma(1)}{3 - 1} \frac{\sigma(3) - \sigma(2)}{3 - 2} \\ &= \frac{3 - 2}{2 - 1} \frac{1 - 2}{3 - 1} \frac{1 - 3}{3 - 2} \\ &= 1. \end{aligned}$$

Die Permutation

$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

hat genau drei Fehlstände, nämlich $(1, 2)$, $(1, 3)$ und $(2, 3)$. Es gilt

$$\begin{aligned} \operatorname{sgn} \sigma_4 &= \frac{\sigma(2) - \sigma(1)}{2 - 1} \frac{\sigma(3) - \sigma(1)}{3 - 1} \frac{\sigma(3) - \sigma(2)}{3 - 2} \\ &= \frac{2 - 3}{2 - 1} \frac{1 - 3}{3 - 1} \frac{1 - 2}{3 - 2} \\ &= -1. \end{aligned}$$

△

Bemerkung 7.37 (zu Definition 7.35)**.**
Da in den Faktoren des Produkts in (7.17) dieselben ganzen Zahlen — abgesehen vom Vorzeichen — jeweils einmal im Zähler und einmal im Nenner vorkommen, ist das Signum einer Permutation immer entweder +1 oder −1. Das Signum einer Permutation gibt die **Parität** (englisch: parity) der Anzahl der Fehlstände an, also ob diese gerade oder ungerade ist, da wir für jedes Indexpaar (i, j) mit $i < j$ den Faktor −1 erhalten, wenn es sich um ein Fehlstand handelt, und ansonsten den Faktor +1. Es gilt also

$$\operatorname{sgn} \sigma = (-1)^{\text{Anzahl der Fehlstände von } \sigma}. \tag{7.18}$$

Dementsprechend nennen wir $\sigma \in S_n$ eine **gerade Permutation** (englisch: even permutation), wenn $\operatorname{sgn} \sigma = 1$ ist und eine **ungerade Permutation** (englisch: odd permutation), wenn $\operatorname{sgn} \sigma = -1$ gilt. △

[15]Definitionsgemäß wird das im Fall $n = 0$ und $n = 1$ leere Produkt als das neutrale Element der Multiplikation (hier in $\mathbb{Q}$), also als 1, interpretiert.

Lemma 7.38 (Transpositionen sind ungerade)**.**
Ist $\tau \in S_n$, $n \in \mathbb{N}_0$, eine Transposition, so gilt $\operatorname{sgn} \tau = -1$.

Beweis. Wir betrachten eine beliebige Transposition $\tau(i,j)$ in S_n, wobei notwendigerweise $n \geqslant 2$ gilt. O. B. d. A. können wir $i < j$ voraussetzen, also haben wir

$$\tau(i,j) = \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots n \\ 1 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots n \end{pmatrix}.$$

$\tau(i,j)$ hat also genau die Fehlstände

$$\begin{aligned} &(i, i+1), \dots, (i, j) && \text{Anzahl: } j - i \\ &(i+1, j), \dots, (j-1, j) && \text{Anzahl: } j - i - 1. \end{aligned}$$

Daher gilt $\operatorname{sgn} \tau(i,j) = (-1)^{2(j-i)-1} = -1$. □

Beispiel 7.39 (Transpositionen sind ungerade)**.**
Zur Veranschaulichung des Beweises von Lemma 7.38 betrachten wir die Permutation in S_7

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

mit $i = 3$ und $j = 6$. Diese hat genau die Fehlstände $(3,4)$, $(3,5)$, $(3,6)$ sowie $(4,6)$, $(5,6)$. △

Satz 7.40 (Signum ist verträglich mit der Komposition von Permutationen)**.**
Es sei $n \in \mathbb{N}_0$ und σ_1, σ_2 zwei Permutationen in S_n. Dann gilt

$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = (\operatorname{sgn} \sigma_1) \cdot (\operatorname{sgn} \sigma_2). \tag{7.19}$$

Beweis. Wir führen den Beweis in drei Schritten. Notwendigerweise gilt wieder $n \geqslant 2$, damit überhaupt Permutationen existieren.

Schritt 1: Wir beweisen den Satz zunächst für den Spezialfall, dass σ_1 eine Transposition benachbarter Elemente ist, sagen wir $\sigma_1 = \tau(k, k+1)$ für ein $k \in [\![1, n-1]\!]$.

Wenn $\sigma_2^{-1}(k) < \sigma_2^{-1}(k+1)$ gilt, dann ist $(\sigma_2^{-1}(k), \sigma_2^{-1}(k+1))$ kein Fehlstand von σ_2, jedoch ein Fehlstand von $\tau(k,k+1) \circ \sigma_2$. Wenn andererseits $\sigma_2^{-1}(k) > \sigma_2^{-1}(k+1)$ gilt, dann ist $(\sigma_2^{-1}(k), \sigma_2^{-1}(k+1))$ ein Fehlstand von σ_2, aber kein Fehlstand von $\tau(k,k+1) \circ \sigma_2$. Die anderen Fehlstände von σ_2 und $\tau(k,k+1) \circ \sigma_2$ sind dieselben. Daher sind die Anzahlen der Fehlstände von σ_2 und von $\tau(k,k+1) \circ \sigma_2$ um 1 verschieden. Damit ist

$$\operatorname{sgn}\big(\tau(k,k+1) \circ \sigma_2\big) = -\operatorname{sgn} \sigma_2 = \big(\operatorname{sgn} \tau(k,k+1)\big) \cdot (\operatorname{sgn} \sigma_2)$$

gezeigt.

Schritt 2: Wir beweisen den Satz für den Spezialfall, dass $\tau(k,\ell)$ eine beliebige Transposition ist.

Wir können o. B. d. A. $\ell > k$ annehmen, daher können wir $\tau(k,\ell)$ in der Form

$$\tau(k,\ell) = \underbrace{\tau(k,k+1) \circ \cdots \circ \tau(\ell-2,\ell-1)}_{} \circ \underbrace{\tau(\ell,\ell-1) \circ \cdots \circ \tau(k+1,k)}_{},$$

also als Komposition von $(2\,(\ell-k)-1)$ Transpositionen jeweils benachbarter Elemente schreiben.[16] Aufgrund von Schritt 1 und der Assoziativität der Komposition haben wir nun also

$$\begin{aligned}
&\operatorname{sgn}\big(\tau(k,\ell) \circ \sigma_2\big) \\
&\qquad = \operatorname{sgn}\big(\tau(k,k+1) \circ \cdots \circ \tau(\ell-2,\ell-1) \circ \tau(\ell,\ell-1) \circ \cdots \circ \tau(k+1,k) \circ \sigma_2\big) \\
&\qquad = \big(\operatorname{sgn}\tau(k,k+1)\big) \cdot \operatorname{sgn}\big(\cdots \circ \tau(\ell-2,\ell-1) \circ \tau(\ell,\ell-1) \circ \cdots \circ \tau(k+1,k) \circ \sigma_2\big) \\
&\qquad = \cdots \\
&\qquad = \big(\operatorname{sgn}\tau(k,k+1)\big) \cdots \big(\operatorname{sgn}\tau(\ell,\ell-1)\big) \cdots \big(\operatorname{sgn}\tau(k+1,k)\big)(\operatorname{sgn}\sigma_2) \\
&\qquad = \cdots \\
&\qquad = \big(\operatorname{sgn}\tau(k,\ell)\big) \cdot (\operatorname{sgn}\sigma_2).
\end{aligned}$$

Schritt 3: Schließlich können wir den allgemeinen Fall zeigen.

Ist $\sigma_1 \in S_n$ eine beliebige Permutation, so können wir sie nach Satz 7.33 als Komposition von Transpositionen $\sigma_1 = \tau_1 \circ \cdots \circ \tau_r$ schreiben. Unter Benutzung von Schritt 2 und der Assoziativität der Komposition folgt nun ähnlich wie im Beweis von Schritt 2:

$$\begin{aligned}
&\operatorname{sgn}(\sigma_1 \circ \sigma_2) \\
&\qquad = \operatorname{sgn}(\tau_1 \circ \cdots \circ \tau_r \circ \sigma_2) \\
&\qquad = (\operatorname{sgn}\tau_1) \cdot \operatorname{sgn}(\cdots \circ \tau_r \circ \sigma_2) \\
&\qquad = \cdots \\
&\qquad = (\operatorname{sgn}\tau_1) \cdots (\operatorname{sgn}\tau_r)(\operatorname{sgn}\sigma_2) \\
&\qquad = \cdots \\
&\qquad = \operatorname{sgn}(\tau_1 \circ \cdots \circ \tau_r) \cdot \operatorname{sgn}(\sigma_2).
\end{aligned}$$

□

Folgerung 7.41 (zu Satz 7.40)**.**
Es sei $n \in \mathbb{N}_0$ und σ eine Permutation in S_n.

(*i*) Ist $\sigma = \sigma_1 \circ \cdots \circ \sigma_s$ dargestellt als Komposition[17] von $s \in \mathbb{N}_0$ Permutationen $\sigma_i \in S_n$, so gilt $\operatorname{sgn}\sigma = (\operatorname{sgn}\sigma_1) \cdots (\operatorname{sgn}\sigma_s)$.

---

[16] Beispielsweise im Fall $k = 3$ und $\ell = 6$ haben wir

$$\begin{pmatrix} 1&2&3&4&5&6&7 \\ 1&2&3&4&5&6&7 \end{pmatrix} \xrightarrow{\tau(4,3)\circ} \begin{pmatrix} 1&2&3&4&5&6&7 \\ 1&2&4&3&5&6&7 \end{pmatrix} \xrightarrow{\tau(5,4)\circ} \begin{pmatrix} 1&2&3&4&5&6&7 \\ 1&2&5&3&4&6&7 \end{pmatrix} \xrightarrow{\tau(6,5)\circ} \begin{pmatrix} 1&2&3&4&5&6&7 \\ 1&2&6&3&4&5&7 \end{pmatrix}$$
$$\xrightarrow{\tau(4,5)\circ} \begin{pmatrix} 1&2&3&4&5&6&7 \\ 1&2&6&3&5&4&7 \end{pmatrix} \xrightarrow{\tau(3,4)\circ} \begin{pmatrix} 1&2&3&4&5&6&7 \\ 1&2&6&4&5&3&7 \end{pmatrix}.$$

wodurch in der Tat $3 \leftrightarrow 6$ getauscht sind.

[17] Vereinbarungsgemäß ist die Verknüpfung von null Permutationen das neutrale Element in S_n, also die identische Abbildung id.

(*ii*) Ist insbesondere $\sigma = \tau_1 \circ \cdots \circ \tau_r$ dargestellt als Komposition von $r \in \mathbb{N}_0$ Transpositionen in S_n, so gilt $\operatorname{sgn} \sigma = (-1)^r$.

(*iii*) Es gilt $\operatorname{sgn} \operatorname{id} = 1$.

(*iv*) Es gilt $\operatorname{sgn} \sigma = \operatorname{sgn} \sigma^{-1}$.

Beweis. Aussage (*i*): Nach Satz 7.40 ist das Signum einer Komposition von zwei Permutationen gleich dem Produkt der Signa der beiden Faktoren. Wie im Beweis von Satz 7.40 können wir die Aussage leicht auf mehr als zwei Faktoren ausdehnen. Die Fälle $s = 0$ und $s = 1$ sind trivial.

Aussage (*ii*): Das Signum einer Transposition ist nach Lemma 7.38 gleich -1. Aussage (*ii*) folgt damit aus Aussage (*i*).

Aussage (*iii*): Die identische Abbildung ist Produkt von null Transpositionen, also gilt $\operatorname{sgn} \operatorname{id} = (-1)^0 = 1$.

Aussage (*iv*): Schließlich gilt

$$1 = \operatorname{sgn} \operatorname{id} = \operatorname{sgn}(\sigma \circ \sigma^{-1}) = (\operatorname{sgn} \sigma) \cdot (\operatorname{sgn} \sigma^{-1}),$$

also $\operatorname{sgn} \sigma^{-1} = 1/\operatorname{sgn} \sigma = \operatorname{sgn} \sigma$, da $\operatorname{sgn} \sigma \in \{\pm 1\}$ ist. □

Ende der Vorlesung 9

§ 7.4 Untergruppen

Definition 7.42 (Untergruppe)**.**
Es sei $(G, \star)$ eine Gruppe.

(*i*) Eine Teilmenge $U \subseteq G$ heißt eine **Untergruppe** (englisch: subgroup) **von** $(G, \star)$, wenn U bzgl. $\star$ abgeschlossen und wenn $(U, \star)$ selbst wieder eine Gruppe ist. Manchmal schreibt man dies als $(U, \star) \leqslant (G, \star)$.

(*ii*) Eine Untergruppe $(U, \star)$ von $(G, \star)$ heißt **echt** (englisch: proper subgroup), wenn $U \subsetneq G$ gilt. △

Beachte: Die Assoziativität wird von $\star$ auf $\star$ vererbt. Ist $(G, \star)$ kommutativ, dann auch $(U, \star)$. Wie bereits bei Unterhalbgruppen werden wir in Zukunft auch einfach die Menge U als Untergruppe von $(G, \star)$ bezeichnen.

Bei Untermonoiden (Definition 7.11) hatten wir gefordert, dass ihr neutrales Element dasselbe ist wie im „Obermonoid“. Bei Untergruppen ergibt sich das von selbst:

Lemma 7.43 (neutrale und inverse Elemente in einer Untergruppe)**.**
Es sei U eine Untergruppe der Gruppe $(G, \star)$. Dann ist das neutrale Element e_U von $(U, \star)$ gleich dem neutralen Element e von $(G, \star)$. Außerdem gilt für alle $a \in U$, dass das Inverse von a in U übereinstimmt mit dem Inversen von a in G.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

Aufgrund dieser Erkenntnis benötigen wir also keine neue Notation für das neutrale Element und die Inversen in einer Untergruppe. Wir halten weiterhin fest, dass eine Untergruppe die Eigenschaft $U \star U = U$ erfüllt. (**Quizfrage 7.12:** Warum?)

Die Prüfung einer Teilmenge $U \subseteq G$ auf die Untergruppen-Eigenschaft lässt sich mit folgendem Kriterium erreichen:

Satz 7.44 (Untergruppenkriterium)**.**
Es sei $(G, \star)$ eine Gruppe und $U \subseteq G$. Dann sind äquivalent:

(i) $(U, \star)$ ist eine Untergruppe von $(G, \star)$.

(ii) $U \neq \emptyset$, und für alle $a, b \in U$ gilt $a \star b' \in U$.[18]

(iii) $U \neq \emptyset$, und für alle $a, b \in U$ gilt $a' \star b \in U$.[19]

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $(U, \star)$ eine Untergruppe von $(G, \star)$. Dann enthält U notwendigerweise das neutrale Element e von $(G, \star)$, da es nach Lemma 7.43 auch das neutrale Element in $(U, \star)$ ist. Für $a, b \in U$ gilt $b' \in U$ nach Lemma 7.43. Da U bzgl. $\star$ abgeschlossen ist, folgt $a \star b' \in U$.

Aussage (ii) $\Rightarrow$ Aussage (i):

Schritt 1: U enthält das neutrale Element e von $(G, \star)$:

Da U nichtleer ist, existiert ein $a \in U$. Mit dem dazu inversen Element a' gilt aufgrund der Voraussetzung $a \star a' \in U$, also $e \in U$ für das neutrale Element e von $(G, \star)$.

Schritt 2: U enthält die Inversen seiner Elemente:

Es sei $a \in U$, dann gilt $a' = e \star a'$, und aufgrund der Voraussetzung liegt $a' \in U$.

Schritt 3: U ist abgeschlossen bzgl. $\star$:

Für $a, b \in U$ liegt auch $b' \in U$, also ist $a \star b = a \star (b')'$ aufgrund der Voraussetzung ebenfalls ein Element von U.

Zusammenfassend haben wir also gezeigt, dass U bzgl. $\star$ abgeschlossen ist (Schritt 3), also bildet $(U, \star)$ eine Halbgruppe. Weiter zeigt Schritt 1, dass $(U, \star)$ ein Monoid mit dem neutralen Element e von $(G, \star)$ ist. Schließlich zeigt Schritt 2, dass alle Elemente von U ein Inverses in U besitzen, also ist $(U, \star)$ eine Gruppe und wegen $U \subseteq G$ eine Untergruppe von $(G, \star)$.

Der Beweis von Aussage (i) $\Rightarrow$ Aussage (iii) und Aussage (iii) $\Rightarrow$ Aussage (i) läuft analog. □

Beispiel 7.45 (Untergruppen)**.**

(i) Es sei $(G, \star)$ eine Gruppe mit neutralem Element e. Dann sind $\{e\}$ und G Untergruppen von $(G, \star)$. Diese heißen die **trivialen Untergruppen** (englisch: trivial subgroups).

(ii) $\mathbb{Q}_{>0}$ ist eine Untergruppe von $(\mathbb{Q}_{\neq 0}, \cdot)$, und $\mathbb{R}_{>0}$ ist eine Untergruppe von $(\mathbb{R}_{\neq 0}, \cdot)$.

[18] kurz: $U \star U' \subseteq U$
[19] kurz: $U' \star U \subseteq U$

(*iii*) Für jede Zahl $m \in \mathbb{N}_0$ ist $m\mathbb{Z} = \{m\,z \mid z \in \mathbb{Z}\}$ mit der Verknüpfung $+$ eine Untergruppe der Gruppe $(\mathbb{Z}, +)$. Das sind auch bereits alle möglichen Untergruppen von $(\mathbb{Z}, +)$.

(*iv*) Die Menge $\{\pm 1\}$ ist eine Untergruppe von $(\mathbb{Q}_{\neq 0}, \cdot)$, von $(\mathbb{R}_{\neq 0}, \cdot)$ und von $(\mathbb{C}_{\neq 0}, \cdot)$.

(*v*) Die Menge $\{\pm 1, \pm i\}$ ist eine Untergruppe von $(\mathbb{C}_{\neq 0}, \cdot)$.

(*vi*) Die Menge $\{z \in \mathbb{C} \mid |z| = 1\}$ ist eine Untergruppe von $(\mathbb{C}_{\neq 0}, \cdot)$.

(*vii*) Für $n \in \mathbb{N}_0$ ist

$$\begin{aligned} A_n &\coloneqq \{\sigma \in S_n \mid \sigma \text{ ist Komposition einer geraden Anzahl von Transpositionen}\} \\ &= \{\sigma \in S_n \mid \operatorname{sgn} \sigma = 1\} \end{aligned} \tag{7.20}$$

eine Untergruppe von S_n, genannt die **alternierende Gruppe** (englisch: alternating group) vom Grad n. Für $n = 0, 1$ stimmt sie mit S_n überein. Für $n \geqslant 2$ ist A_n eine echte Untergruppe von S_n mit $\frac{1}{2}n!$ Elementen.

(*viii*) In S_3 besteht die alternierende Untergruppe A_3 in der Notation von Beispiel 7.31 gerade aus $\{\sigma_0, \sigma_1, \sigma_2\}$. Diese entsprechen bei Interpretation als Kongruenzabbildungen eines gleichseitigen Dreiecks auf sich selbst (Abbildung 7.1) gerade den Drehungen. △

Quizfrage 7.13: Können Sie eine Gruppe finden, die außer den trivialen Untergruppen keine weiteren Untergruppen besitzt?

Bemerkung 7.46 („Untergruppe sein" ist eine Ordnungsrelation, vgl. Bemerkung 7.13 zu Unterhalbgruppen und Untermonoiden)**.**

(*i*) Die Relation „ist Untergruppe von" ist eine partielle Ordnung auf der Klasse aller Gruppen.

(*ii*) Insbesondere ist die Menge aller Untergruppen einer bestimmten Gruppe $(G, \star)$ durch die Untergruppenhalbordnung partiell geordnet. Diese Ordnung stimmt mit der Inklusionshalbordnung überein (Abbildung 7.2).

(*iii*) Ist $U \subseteq H$ ein Untermonoid des Monoids $(H, \star)$ und ist $(U, \star)$ zusätzlich eine Gruppe, dann sprechen wir auch kurz von einer **Untergruppe** $(U, \star)$ **des Monoids** $(H, \star)$. Das trifft genau dann zu, wenn $(U, \star)$ eine Gruppe ist und das neutrale Element $e \in H$ enthält. (**Quizfrage 7.14:** Klar?)

(*iv*) In diesem Sinne ist die Einheitengruppe E eines Monoids $(H, \star)$ mit neutralem Element e die größte Untergruppe von $(H, \star)$, also: E ist das Maximum der Menge

$$\{U \subseteq H \mid U \text{ ist Untergruppe von } (H, \star)\}$$

bzgl. der Inklusionshalbordnung (und auch bzgl. der Untergruppen-Halbordnung). Alle weiteren Untergruppen von $(H, \star)$ sind also Teilmengen (und sogar Untergruppen) von E. △

Lemma 7.47 (Durchschnitt von Untergruppen)**.**
Es sei $(G, \star)$ eine Gruppe.

(*i*) Ist $(U_i)_{i \in I}$ eine nichtleere Familie von Untergruppen von $(G, \star)$, dann ist auch $\bigcap_{i \in I} U_i$ eine Untergruppe von $(G, \star)$.

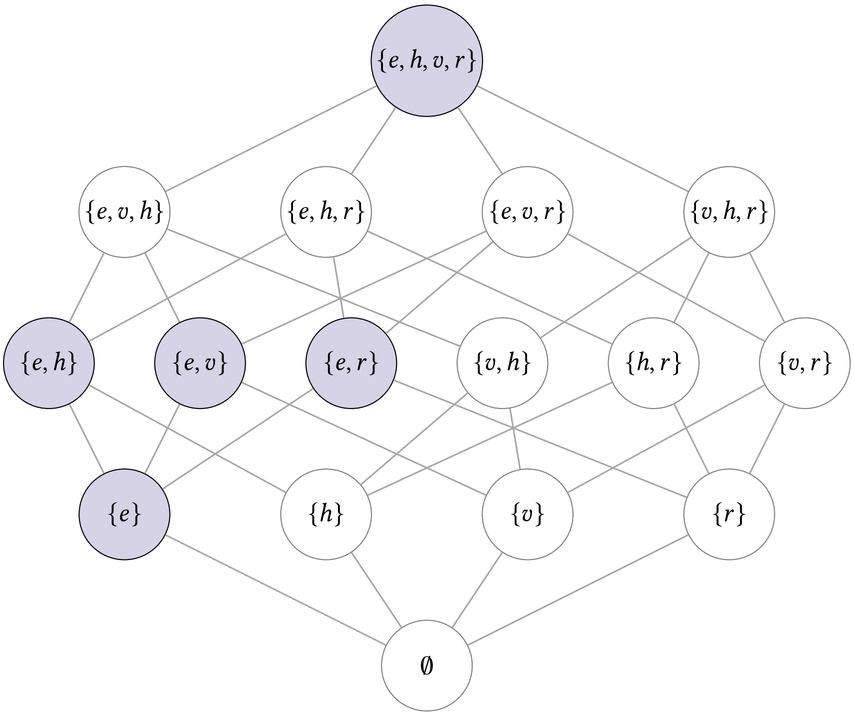


Abbildung 7.2.: Hasse-Diagramm der Inklusionshalbordnung auf der Kleinschen Vierergruppe $K_4 = \{e, h, v, r\}$ (Beispiel 7.22). Hervorgehoben sind die fünf Untergruppen von K_4. Diese Teilmenge ist durch die Halbordnung „ist Untergruppe von" partiell geordnet, welche mit der Inklusionshalbordnung übereinstimmt (Bemerkung 7.46). Für Untergruppen U_1, U_2 von K_4 gilt also: U_1 ist Untergruppe von U_2 genau dann, wenn U_1 Teilmenge von U_2 ist. Im Diagramm erkennen wir das durch einen aufsteigenden Pfad von U_1 nach U_2.
Anhand des Diagramms können wir außerdem die von einer Teilmenge $E \subseteq K_4$ erzeugte Untergruppe $\langle E \rangle$ (Definition 7.48) ablesen. Dazu suchen wir ausgehend von E einen kürzesten aufsteigenden Pfad zu einer der Untergruppen. Kürzeste Pfade sind i. A. nicht eindeutig, die dadurch erreichte Untergruppe jedoch schon. Beispielsweise gilt $\langle v, h \rangle = \{e, h, v, r\}$, und es gibt zwei mögliche Pfade.

(ii) Ist $\mathcal{U}$ eine nichtleere Menge von Untergruppen von $(G, \star)$, dann ist auch $\bigcap \mathcal{U}$ eine Untergruppe von $(G, \star)$.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

In Definition 5.14 hatten wir die Hüllenbildung einer Menge betrachtet, sodass die kleinstmögliche Oberrelation mit den gewünschten Eigenschaften entsteht. Analog dazu betrachten wir jetzt die Anreicherung der Teilmenge einer Gruppe zu einer Untergruppe.

Definition 7.48 (erzeugte Untergruppe, Erzeugendensystem, zyklische Gruppe)**.**
Es sei $(G, \star)$ eine Gruppe und $E \subseteq G$.

(*i*) Die Menge

$$\langle E \rangle \coloneqq \bigcap \{U \mid U \text{ ist Untergruppe von } (G,\star) \text{ und } E \subseteq U\} \tag{7.21}$$

heißt die **von E erzeugte Untergruppe** (englisch: subgroup generated by E) oder auch die **Untergruppenhülle** (englisch: subgroup hull) oder der **Untergruppenabschluss** (englisch: subgroup closure) **von E in** $(G,\star)$.

Ist speziell E die endliche Menge $E = \{a_1, \dots, a_n\}$ mit $a_i \in G$ und $n \in \mathbb{N}_0$, so schreiben wir auch $\langle a_1, \dots, a_n \rangle$ statt $\langle \{a_1, \dots a_n\} \rangle$.

(*ii*) Die von einem einzelnen Element $a \in G$ erzeugte Untergruppe $\langle a \rangle$ heißt die **von a erzeugte zyklische Untergruppe** (englisch: cyclic subgroup) von $(G,\star)$.

(*iii*) Gilt $\langle E \rangle = G$, dann heißt E ein **Erzeugendensystem** (englisch: generating set) von $(G,\star)$. Falls ein endliches Erzeugendensystem von G existiert, so heißt G **endlich erzeugt** (englisch: finitely generated).

(*iv*) Gilt $\langle a \rangle = G$ für ein $a \in G$, so heißt die Gruppe $(G,\star)$ **zyklisch** (englisch: cyclic) oder **zyklisch erzeugt** (englisch: cyclically generated). Ein solches Element $a \in G$ heißt ein **Erzeuger** (englisch: generator) **von** G.

(*v*) Ein Element $a \in G$ heißt ein **Gruppenelement endlicher Ordnung** (englisch: group element of finite order), wenn es ein $n \in \mathbb{N}$ gibt mit der Eigenschaft (in multiplikativer Notation) $a^n = 1$. In diesem Fall heißt die kleinste Zahl $n \in \mathbb{N}$ mit dieser Eigenschaft die **Ordnung** von a. Falls kein $n \in \mathbb{N}$ mit der Eigenschaft $a^n = 1$ existiert, so heißt a ein **Gruppenelement unendlicher Ordnung** (englisch: group element of infinite order).[20] △

Bemerkung 7.49 (Eigenschaften der Untergruppenhülle, vgl. Bemerkung 5.16 zu Eigenschaften von Hüllen von Relationen)**.**
Per Konstruktion ist $\langle E \rangle$ ist die kleinste Untergruppe von $(G,\star)$, die die Teilmenge $E \subseteq G$ enthält. Genauer: $\langle E \rangle$ ist das Minimum der Menge

$$\{U \mid U \text{ ist Untergruppe von } (G,\star) \text{ und } E \subseteq U\}$$

bzgl. der Untergruppen-Halbordnung (und auch bzgl. der Inklusionshalbordnung). △

Da Definition 7.48 nur eine abstrakte Definition der von einer Menge erzeugten Untergruppe liefert, wollen wir diese nun charakterisieren.

Satz 7.50 (Darstellung der erzeugten Untergruppe)**.**
Es sei $(G,\star)$ eine Gruppe und $E \subseteq G$. Dann gilt für die von E erzeugte Untergruppe:

$$\langle E \rangle = \left\{ a_1 \star \dots \star a_n \,\middle|\, \exists n \in \mathbb{N}_0 \ \forall i = 1, \dots, n \ (a_i \in E \cup E') \right\}, \tag{7.22}$$

wobei E' die Menge der Inversen von E bezeichnet, vgl. Bemerkung 7.20.[21]

[20] Die Ordnung einer bijektiven Funktion $f\colon X \to X$ (Definition 6.18) ist also nichts anderes als die Ordnung von f als Element der Gruppe der bijektiven Funktionen $X \to X$ mit der Komposition $\circ$ als Verknüpfung.

[21] Im Fall $n = 0$ interpretieren wir wie üblich die Verknüpfung von null Elementen als das neutrale Element e der Gruppe. Insbesondere im Fall $E = \emptyset$ ergibt sich also $\langle E \rangle = \{e\}$.

Beweis. Zur Abkürzung bezeichnen wir die Menge auf der rechten Seite von (7.22) mit M. Wir führen den Beweis in zwei Schritten.

Schritt 1: $\langle E\rangle \supseteq M$: Es sei U eine beliebige Untergruppe von G, die im Durchschnitt (7.21) vorkommt. U enthält also E als Teilmenge. Da U eine Untergruppe ist, enthält U auch E'. Da schließlich U abgeschlossen bzgl. $\star$ ist, enthält U auch alle Verknüpfungen endlich vieler Elemente aus $E \cup E'$. Also gilt $U \supseteq M$. Da dies für jede beliebige Untergruppe aus dem Durchschnitt in (7.21) gilt, gilt auch $\langle E\rangle \supseteq M$.

Schritt 2: $\langle E\rangle \subseteq M$: Wir zeigen zunächst, dass M selbst eine Untergruppe von G ist. Dazu überprüfen wir das Untergruppenkriterium (Satz 7.44). Offensichtlich ist $M \neq \emptyset$, denn M enthält mindestens e. Sind $a_1 \star \cdots \star a_n$ und $b_1 \star \cdots \star b_m$ zwei Elemente aus M, so ist auch $(a_1 \star \cdots \star a_n) \star (b_1 \star \cdots \star b_m)'$ ein Element aus M. Also ist M eine Untergruppe von G. Zusätzlich ist klar, dass $E \subseteq M$ gilt. (**Quizfrage 7.15:** Details?) Das heißt, M ist eine derjenigen Untergruppen von G, über die in der Definition von $\langle E\rangle$ der Durchschnitt gebildet wird. Folglich gilt $\langle E\rangle \subseteq M$. □

Beispiel 7.51 (erzeugte Untergruppe, Erzeugendensystem, zyklische Gruppe)**.**

(*i*) In der Gruppe $(\mathbb{Z}, +)$ erzeugt das Element $m \in \mathbb{Z}$ die zyklische Untergruppe $\langle m\rangle = m\mathbb{Z}$.

(*ii*) Die Gruppe $(\mathbb{Z}, +)$ ist zyklisch. Sie hat die Erzeuger 1 und -1, es gilt also $\langle 1\rangle = \langle -1\rangle = \mathbb{Z}$.

(*iii*) In S_3 gilt mit den Bezeichnungen aus Beispiel 7.31, also

$$\sigma_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \text{(Drehungen)}$$
$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{(Spiegelungen)}$$

die Beziehung

$$\sigma_1^2 = \sigma_2 \quad \text{und} \quad \sigma_1^3 = \sigma_0 = \mathrm{id}_{\{1,2,3\}}.$$

Folglich ist

$$\langle \sigma_1\rangle = \{\sigma_0, \sigma_1, \sigma_2\} = A_3$$

die alternierende Untergruppe, vgl. Beispiel 7.45. Wegen $\sigma_2^2 = \sigma_1$ und $\sigma_2^3 = \sigma_0$ gilt auch $\langle \sigma_2\rangle = A_3$. Wegen $\sigma_3^2 = \sigma_4^2 = \sigma_5^2 = \mathrm{id}_{\{1,2,3\}}$ gilt $\langle \sigma_3\rangle = \{\sigma_0, \sigma_3\}$, $\langle \sigma_4\rangle = \{\sigma_0, \sigma_4\}$ und $\langle \sigma_5\rangle = \{\sigma_0, \sigma_5\}$. Wollen wir ganz S_3 erzeugen, so müssen wir mindestens zwei Permutationen auswählen. Beispielsweise ist $\{\sigma_1, \sigma_3\}$ (eine Drehung, eine Spiegelung) ein Erzeugendensystem von S_3.

(*iv*) Die Bewegungen des Zauberwürfels (ausgehend von der gelösten Position in einer beliebig, aber fest gewählten Orientierung) können als eine Untergruppe der symmetrischen Gruppe S_{48} auf den 48 nicht-zentralen Facetten des Würfels modelliert werden. Diese Untergruppe wird z. B. erzeugt von der sechselementigen Menge der Drehungen der sechs Seiten des Würfels im Uhrzeigersinn. Sie hat 43 252 003 274 489 856 000 $\approx 4.3 \cdot 10^{19}$ Elemente, während S_{48} etwa $1.2 \cdot 10^{61}$ Elemente hat. △

Die Vereinigung $U_1 \cup U_2$ von zwei Untergruppen einer Gruppe G ist i. A. keine Untergruppe von G.[22] Wir betrachten daher stattdessen die kleinste Untergruppe von G, die $U_1 \cup U_2$ enthält, also die von $U_1 \cup U_2$ erzeugte Untergruppe $\langle U_1 \cup U_2 \rangle$ (englisch: join of two subgroups). Diese hat folgende Darstellung:

Folgerung 7.52 (zu Satz 7.50: Untergruppenhülle der Vereinigung zweier Untergruppen)**.**
Es seien $(G, \star)$ eine Gruppe und U_1, U_2 Untergruppen von G. Dann gilt für die von $U_1 \cup U_2$ erzeugte Untergruppe:

$$\langle U_1 \cup U_2 \rangle = \Big\{ a_1 \star b_1 \star a_2 \star b_2 \cdots \star a_n \star b_n \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; (a_i \in U_1, b_i \in U_2) \Big\}. \tag{7.23}$$

Beweis. □

Abschließend geben wir noch ein Resultat zur Untergruppenhülle von Vereinigung und Schnitt zweier beliebiger Teilmengen einer Gruppe an:

Folgerung 7.53 (zu Satz 7.50: Untergruppenhülle von Vereinigung und Schnitt)**.**
Es sei $(G, \star)$ eine Gruppe und $E_1, E_2 \subseteq G$. Dann gilt:

$$\langle E_1 \cup E_2 \rangle = \big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle, \tag{7.24a}$$

$$\langle E_1 \cap E_2 \rangle \subseteq \big\langle \langle E_1 \rangle \cap \langle E_2 \rangle \big\rangle = \langle E_1 \rangle \cap \langle E_2 \rangle. \tag{7.24b}$$

Beweis. Übung □

§ 7.5 Untergruppen induzieren Äquivalenzrelationen

Definition 7.54 (Links- und Rechtsnebenklassen einer Untergruppe)**.**
Es sei $(G, \star)$ eine Gruppe und U eine Untergruppe.

(i) Für $a \in G$ heißt $a \star U$ heißt eine **Linksnebenklasse** (englisch: left coset) **von** U.[23]

(ii) Für $a \in G$ heißt $U \star a$ heißt eine **Rechtsnebenklasse** (englisch: right coset) **von** U.[24] △

Satz 7.55 (Links- und Rechtsnebenklassen partitionieren eine Gruppe)**.**
Es sei $(G, \star)$ eine Gruppe, U eine Untergruppe und $a, b \in G$.

(i) Die Linksnebenklassen $a \star U$ und $b \star U$ sind entweder gleich oder disjunkt.

(ii) Die Rechtsnebenklassen $U \star a$ und $U \star b$ sind entweder gleich oder disjunkt.

[22] Tatsächlich gilt: $U_1 \cup U_2$ ist genau dann eine Untergruppe, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ gilt (Übung).

[23] Die Linksnebenklasse $a \star U$ ist also die Linkstranslation von U um das Element a. Bei den Linksnebenklassen $a \star U$ steht der Repräsentant a links vom U.

[24] Die Rechtsnebenklasse $U \star a$ ist also die Rechtstranslation von U um das Element a. Bei den Rechtsnebenklassen $a \star U$ steht der Repräsentant a rechts vom U.

Beweis. Aussage (i): Wir nehmen an, dass $a \star U$ und $b \star U$ nicht disjunkt sind. Es existiert also ein $c \in (a \star U) \cap (b \star U)$, d. h., es existieren $u_1, u_2 \in U$, sodass $a \star u_1 = b \star u_2$ gilt. Daraus folgt

$$a \star u = a \star u_1 \star u_1' \star u = b \star u_2 \star u_1' \star u \quad \text{für alle } u \in U.$$

Da U abgeschlossen ist unter Verknüpfung und Inversenbildung, gehört $u_2 \star u_1' \star u$ zu U. Das zeigt $a \star U \subseteq b \star U$. Analog folgt $b \star U \subseteq a \star U$ aus

$$b \star u = b \star u_2 \star u_2' \star u = a \star u_1 \star u_2' \star u \quad \text{für alle } u \in U.$$

Insgesamt erhalten wir also, dass zwei nicht-disjunkte Linksnebenklassen $a \star U$ und $b \star U$ bereits gleich sind.

Der Beweis von Aussage (ii) erfolgt analog. □

Jede Nebenklasse $a \star U$ enthält das Element a (ist also nichtleer), da das neutrale Element von G auch zu U gehört. Damit bilden die Linksnebenklassen eine Partition von G. Nach Satz 5.25 („Partitionen sind dasselbe wie Äquivalenzrelationen") existiert also eine eindeutig bestimmte Äquivalenzrelation $\sim^U$, deren Äquivalenzklassen genau die Linksnebenklassen sind. Analog existiert eine eindeutig bestimmte Äquivalenzrelation ${}^U\!\!\sim$, deren Äquivalenzklassen genau die Rechtsnebenklassen sind. Wir halten dieses Ergebnis fest als

Folgerung 7.56 (von einer Untergruppe induzierte Äquivalenzrelationen)**.**
Es sei $(G, \star)$ eine Gruppe und U eine Untergruppe.

(i) Dann sind durch

$$a \sim^U b \quad \Leftrightarrow \quad a \star U = b \star U \quad \Leftrightarrow \quad b \in a \star U \quad \Leftrightarrow \quad a \in b \star U \tag{7.25a}$$

$$a \;{}^U\!\!\sim b \quad \Leftrightarrow \quad U \star a = U \star b \quad \Leftrightarrow \quad b \in U \star a \quad \Leftrightarrow \quad a \in U \star b \tag{7.25b}$$

für $a, b \in G$ zwei Äquivalenzrelationen[25] auf G erklärt, deren Äquivalenzklassen gerade die Links- bzw. die Rechtsnebenklassen von U sind.

(ii) Jede der Äquivalenzklassen ist gleichmächtig zu U.

(iii) Ist L ein Repräsentantensystem der Linksnebenklassen, dann ist $R \coloneqq L'$ ein Repräsentantensystem der Rechtsnebenklassen.

Beweis. Aussage (i): Aus Satz 5.25 („Partitionen sind dasselbe wie Äquivalenzrelationen") folgt, dass die durch $a \sim^U b \Leftrightarrow a \star U = b \star U$ definierte Relation eine Äquivalenzrelation ist, deren Äquivalenzklassen genau die Linksnebenklassen von U sind. Es bleiben die weiteren Äquivalenzen in (7.25a) zu zeigen. Die Mengen $a \star U$ und $b \star U$ enthalten a bzw. b und sind beide Linksnebenklassen von U. Sie sind genau dann gleich, wenn $b \in a \star U$ gilt. Analog sind sie auch genau dann gleich, wenn $a \in b \star U$ gilt.

Die Aussagen in (7.25b) zeigt man analog.

[25] Für diese Relationen gibt es in der Literatur keine einheitliche Notation. Wir benutzen $\sim^U$ mit dem Symbol $\sim$ links von U für die Relation, deren Äquivalenzklassen die Linksnebenklassen sind. Entsprechend bezeichnet ${}^U\!\!\sim$ mit dem Symbol $\sim$ rechts von U die Äquivalenzrelation, deren Äquivalenzklassen die Rechtsnebenklassen sind.

Aussage (ii): Für jedes $a \in G$ ist die Abbildung bildet die Linkstranslation um a die Menge U bijektiv auf die Linksnebenklasse $a \star U$ ab. Das zeigt die Gleichmächtigkeit von U und $a \star U$. Analog zeigt die Rechtstranslation um a die Gleichmächtigkeit von U und $U \star a$.

Aussage (iii): Es sei L ein Repräsentantensystem der Linksnebenklassen und $U \star a$ eine der Rechtsnebenklassen. Zu zeigen ist, dass $U \star a$ durch genau ein Element aus $R := L'$ repräsentiert wird.

Schritt 1: Zur Existenz: Die Linksnebenklasse $a' \star U$ wird durch (genau) ein $\ell \in L$ repräsentiert, also gilt $\ell \in a' \star U$, d. h., $\ell = a' \star u$ für ein $u \in U$. Daher ist $\ell' = (a' \star u)' = u' \star a \in L' = R$ ein Element der Rechtsnebenklasse $U \star a$.

Schritt 2: Zur Eindeutigkeit: Nehmen wir an, $r_1, r_2 \in R$ repräsentieren die gleiche Rechtsnebenklasse $U \star a$. Es existieren also $u_1, u_2 \in U$ mit $r_1 = u_1 \star a$ und $r_2 = u_2 \star a$. Das zeigt $r_1' = a' \star u_1 \in a' \star U$ und $r_2' = a' \star u_2 \in a' \star U$. Damit repräsentieren $r_1', r_2' \in R' = L$ dieselbe Linksnebenklasse $a' \star U$. Es folgt $r_1' = r_2'$ und damit $r_1 = r_2$. □

Die durch die Links- bzw. Rechtsnebenklassen definierten Äquivalenzrelationen $\sim^U$ und ${}^U\!\sim$ heißen die von der Untergruppe U **induzierten Äquivalenzrelationen** (englisch: induced equivalence relations) **auf** G. Wir schreiben:

$$G / U \text{ für die Faktormenge } G / \sim^U \quad \text{(die Menge der Linksnebenklassen),} \tag{7.26a}$$

$$U \backslash G \text{ für die Faktormenge } G / {}^U\!\sim \quad \text{(die Menge der Rechtsnebenklassen).} \tag{7.26b}$$

Folgerung 7.57 (induzierte Äquivalenzrelationen in abelschen Gruppen)**.**

Es sei $(G, \star)$ eine abelsche Gruppe und U eine Untergruppe. Dann gilt $a \star u = u \star a$ für alle $u \in U$ und $a \in G$. Insbesondere gilt also $a \star U = U \star a$ für alle $a \in G$. Die Äquivalenzrelationen $a \sim^U b$ und $a \,{}^U\!\sim b$ sind also identisch.

Beweis. Der Beweis ist Gegenstand der Übung. □

Wann immer $a \sim^U b$ und $a \,{}^U\!\sim b$ identisch sind, schreiben wir auch einfach $a \overset{U}{\sim} b$ und sprechen von **Nebenklassen** (englisch: cosets) $a \star U = U \star a$ von U. Wir werden später in § 8.1 sehen, dass solche Untergruppen auch in nicht-kommutativen Gruppen vorkommen.

Bemerkung 7.58 (induzierte Äquivalenzrelationen verallgemeinern die Gleichheitsrelation)**.**

(i) Im Fall der trivialen Untergruppe $U = \{e\}$ gilt $a \star U = U \star a$ für alle $a \in G$ (Links- und Rechtsnebenklassen stimmen überein) und weiter

$$a \overset{U}{\sim} b \quad \Leftrightarrow \quad b \in a \star U = \{a\} \quad \Leftrightarrow \quad a = b.$$

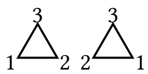
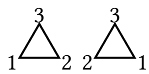
Jede Nebenklasse $a \star U = U \star a$ enthält also nur das Element a.

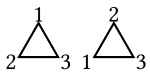
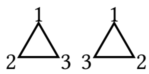
(ii) Wenn die Untergruppe U größer wird, werden auch die Äquivalenzklassen größer. Wir erhalten dadurch eine Verallgemeinerung, eine „gröbere Version“ der Gleichheit. Die Wahl von U bestimmt, welche Unterschiede „nicht gesehen“ oder „ausfaktorisiert“ werden sollen. Zwei Elemente a, b der Gruppe werden als „gleichwertig“ (äquivalent) betrachtet, wenn sie sich nur um ein Element in U unterscheiden (bei Multiplikation von links bzw. von rechts), also wenn $b \in a \star U$ gilt bei Verwendung von Linksnebenklassen bzw. $b \in U \star a$ bei Rechtsnebenklassen.

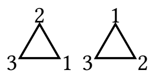
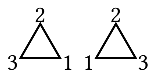
(iii) Im Extremfall $U = G$ gibt es in G/U und $U \backslash G$ jeweils nur eine einzige Äquivalenzklasse, nämlich die gesamte Gruppe G. Es werden also überhaupt keine Elemente mehr unterschieden. △

Beispiel 7.59 (Links- und Rechtsnebenklassen).

(i) In der symmetrischen Gruppe $(S_3, \circ)$ sind die Links- bzw. Rechtsnebenklassen der Untergruppe $U = \{\sigma_0, \sigma_5\}$ (mit den Bezeichnungen aus Beispiel 7.31 und Abbildung 7.1) gerade die Mengen

$\sigma_0 \circ U = \{\sigma_0, \sigma_5\}$  $\qquad U \circ \sigma_0 = \{\sigma_0, \sigma_5\}$ 

$\sigma_1 \circ U = \{\sigma_1, \sigma_3\}$  $\qquad U \circ \sigma_1 = \{\sigma_1, \sigma_4\}$ 

$\sigma_2 \circ U = \{\sigma_2, \sigma_4\}$  $\qquad U \circ \sigma_2 = \{\sigma_2, \sigma_3\}$  .

Die beiden durch U induzierten Äquivalenzrelationen sind hier also verschieden.

Man erkennt, dass innerhalb einer Rechtsnebenklasse (rechte Spalte) zwei Permutationen nicht unterschieden werden, wenn sie sich nur um die horizontale Spiegelung σ_5 unterscheiden. Eine eventuelle nachträgliche Spiegelung σ_5 wird also aus einer Permutation σ „ausfaktorisiert“. Im Unterschied dazu wird bei den Linksnebenklassen nicht unterschieden zwischen Permutation, die mit einer eventuellen Spiegelung σ_5 beginnen.

(ii) In der abelschen Gruppe $(\mathbb{Z}, +)$ erzeugt die Untergruppe $m\mathbb{Z}$ für $m \in \mathbb{N}$ gerade die Kongruenzrelation modulo m, d. h., die Äquivalenzrelationen $\overset{m\mathbb{Z}}{\sim}$ und $\overset{m}{\equiv}$ stimmen überein. Die Nebenklassen $[a] = a + m\mathbb{Z}$ werden auch **Restklassen modulo** m genannt, vgl. Beispiel 5.22), und sie partitionieren die ganzen Zahlen $\mathbb{Z}$ in m gleichmächtige Restklassen, $[0], [1], \dots, [m-1]$.

(iii) Die Standardkonstruktion einer nicht messbaren Teilmenge von $\mathbb{R}$ (Satz von Vitali) verwendet die Nebenklassen von $\mathbb{Q}$ in der abelschen Gruppe $(\mathbb{R}, +)$, zusammen mit dem Auswahlaxiom.[AoC] △

Aus der in Folgerung 7.57 festgestellten Gleichmächtigkeit der Äquivalenzklassen folgt der folgende wichtige **Satz von Lagrange** (englisch: Lagrange’s theorem) der Gruppentheorie:

Satz 7.60 (Satz von Lagrange).
Es sei $(G, \star)$ eine endliche Gruppe und U eine Untergruppe. Dann gilt $\#U \mid \#G$, d. h., die Kardinalität der Untergruppe ist ein Teiler der Kardinalität der Gruppe.

Beweis. Der Beweis ist Gegenstand der Übung. □

Folgerung 7.61 (Gruppen, deren Elementanzahl eine Primzahl ist)**.**
Es sei $(G, \star)$ eine endliche Gruppe, deren Kardinalität eine Primzahl ist. Dann gilt:

(i) G besitzt nur die trivialen Untergruppen $\{e\}$ und G.

(ii) G ist zyklisch, und jedes Element $a \in G \setminus \{e\}$ ist ein Erzeuger.

(iii) G ist abelsch.

Beweis. Aussage (i): Nach dem Satz von Lagrange 7.60 kommen als Kardinalität von U nur 1 oder $\#G$ in Frage.

Aussage (ii): Für $a \in G \setminus \{e\}$ ist $\langle a \rangle$ eine Untergruppe von G, die von $\{e\}$ verschieden ist. Es muss also $\langle a \rangle = G$ gelten.

Aussage (iii): Zyklisch erzeugte Gruppen sind immer abelsch. (**Quizfrage 7.16:** Warum ist das so?) □

Ende der Vorlesung 10

Ende der Woche 5

§ 8 Homomorphismen von Halbgruppen und Gruppen

Literatur: Beutelspacher, 2014, Kapitel 9.2.3; Fischer, Springborn, 2020, Kapitel 2.2

Homomorphismen (englisch: homomorphisms, altgriechisch: *ομος*: gemeinsam, altgriechisch: *μορφή*: Form) sind die **strukturverträglichen Abbildungen** (englisch: structurally compatible maps) zwischen algebraischen Strukturen. In diesem Abschnitt geht es speziell um Homomorphismen von Halbgruppen und Gruppen.

Definition 8.1 (Homomorphismus von Halbgruppen)**.**
Es seien $(H_1, \star)$ und $(H_2, \square)$ zwei Halbgruppen.

(i) Eine Abbildung $f: H_1 \to H_2$ heißt **strukturverträglich** oder ein **Homomorphismus von Halbgruppen** (englisch: semigroup homomorphism), wenn gilt:

$$f(a \star b) = f(a) \square f(b) \quad \text{für alle } a, b \in H_1. \tag{8.1}$$

(ii) Ist $f: H_1 \to H_2$ strukturverträglich und gilt $(H_1, \star) = (H_2, \square)$, so sprechen wir auch von einem **Endomorphismus einer Halbgruppe** (englisch: semigroup endomorphism, altgriechisch: *ένδον*: innen).

(*iii*) Ist $f\colon H_1 \to H_2$ strukturverträglich und bijektiv, so heißt f auch **strukturerhaltend** oder ein **Isomorphismus von Halbgruppen** (englisch: semigroup isomorphism, altgriechisch: *ίσος*: gleich). In diesem Fall nennen wir $(H_1, \star)$ und $(H_2, \square)$ auch zueinander **isomorphe Halbgruppen** (englisch: isomorphic semigroups) und schreiben

$$(H_1, \star) \cong (H_2, \square).$$

(*iv*) Ist $f\colon H_1 \to H_2$ strukturverträglich und bijektiv und gilt $(H_1, \star) = (H_2, \square)$, so sprechen wir auch von einem **Automorphismus einer Halbgruppe** (englisch: semigroup automorphism, altgriechisch: *αυτος*: selbst).[26] △

Expertenwissen: Halbgruppenhomomorphismus als kommutatives Diagramm

Wir können den Sachverhalt, dass $f\colon (H_1, \star) \to (H_2, \square)$ ein Homomorphismus von Halbgruppen ist, auch durch das folgende **kommutative Diagramm** (englisch: commutative diagram) ausdrücken:

$$\begin{tikzcd} H_1 \times H_1 \arrow[r, "f \times f"] \arrow[d, "\star"] & H_2 \times H_2 \arrow[d, "\square"] \\ H_1 \arrow[r, "f"] & H_2 \end{tikzcd}$$

Die Abbildung $f \times f$ ist dabei definiert durch $f \times f\colon H_1 \times H_1 \ni (a, b) \mapsto (f(a), f(b)) \in H_2 \times H_2$. Ein solches Diagramm heißt **kommutativ** (englisch: commutative diagram), wenn alle Pfade mit demselben Ausgangs- und demselben Endpunkt dasselbe Ergebnis produzieren.

Beachte: Einen surjektiven Homomorphismus von Halbgruppen nennt man manchmal auch einen **Epimorphismus von Halbgruppen** (englisch: semigroup epimorphism). Einen injektiven Homomorphismus von Halbgruppen nennt man manchmal auch einen **Monomorphismus von Halbgruppen** (englisch: semigroup monomorphism). Wir werden diese Bezeichnungen aber nicht verwenden.

Satz 8.2 (Komposition von Halbgruppenhomomorphismen, Inverse von Halbgruppenisomorphismen)**.**
Es seien $(H_1, \star)$, $(H_2, \square)$ und $(H_3, \bullet)$ drei Halbgruppen.

(*i*) Sind $f\colon H_1 \to H_2$ und $g\colon H_2 \to H_3$ Halbgruppenhomomorphismen, dann ist auch $g \circ f\colon H_1 \to H_3$ ein Halbgruppenhomomorphismus.

(*ii*) Ist $f\colon H_1 \to H_2$ ein Halbgruppenisomorphismus, dann ist auch $f^{-1}\colon H_2 \to H_1$ ein Halbgruppenisomorphismus.

Beweis. □

[26] Ein Automorphismus einer Halbgruppe ist somit ein bijektiver Endomorphismus oder auch ein Isomorphismus von einer Halbgruppe auf sich selbst.

Folgerung 8.3 (Isomorphie von Halbgruppen ist eine Äquivalenzrelation).
Isomorphie ist eine Äquivalenzrelation auf der Klasse aller Halbgruppen.

Beweis. Jede Halbgruppe ist zu sich selbst isomorph über die identische Abbildung. Die Transitivität und Symmetrie folgen aus Satz 8.2. □

Definition 8.4 (Homomorphismus von Monoiden).
Wenn $(H_1, \star)$ und $(H_2, \square)$ beides Monoide sind, so können wir ganz analog zu Definition 8.1 die Begriffe **Homomorphismus**, **Endomorphismus**, **Isomorphismus** und **Automorphismus** (englisch: monoid homomorphism, endomorphism, isomorphism, automorphism) definieren. Zusätzlich zu (8.1) fordern wir dabei aber noch, dass für die Einselemente gilt:

$$f(e_1) = e_2. \tag{8.2}$$

Die Monoide $(H_1, \star)$ und $(H_2, \square)$ heißen zueinander **isomorph**, wenn es zwischen ihnen einen Isomorphismus von Monoiden gibt. Wir schreiben dann $(H_1, \star) \cong (H_2, \square)$. △

Bemerkung 8.5 (zu Monoidhomomorphismen).

(*i*) Homomorphismen sind wie oben bemerkt die strukturverträglichen Abbildungen zwischen jeglicher Art algebraischer Strukturen. Für Monoide besteht diese Struktur aus der Verknüpfung und dem Einselement, daher kommt die Forderung (8.2) hinzu. Sie folgt nicht bereits aus (8.1).

(*ii*) Es reicht allerdings anstelle von (8.2) aus, lediglich zu fordern, dass $f(e_1)$ invertierbar ist, denn daraus folgt bereits (8.2):

$$\begin{aligned} f(e_1) \square f(e_1) &= f(e_1 \star e_1) && \text{da } f \text{ Halbgruppenhomomorphismus ist} \\ &= f(e_1) && \text{da } e_1 \text{ neutrales Element in } (H_1, \star) \text{ ist} \\ &= f(e_1) \square e_2 && \text{da } e_2 \text{ neutrales Element in } (H_2, \square) \text{ ist.} \end{aligned} \tag{8.3}$$

Da $f(e_1)$ als invertierbar vorausgesetzt wurde, zeigt die Kürzungsregel (7.8a) nun $f(e_1) = e_2$. △

Definition 8.6 (Homomorphismus von Gruppen).
Wenn $(G_1, \star)$ und $(G_2, \square)$ beides Gruppen sind, so können wir ganz analog zu Definition 8.1 die Begriffe **Homomorphismus**, **Endomorphismus**, **Isomorphismus** und **Automorphismus** (englisch: group homomorphism, endomorphism, isomorphism, automorphism) definieren. Die Bedingung (8.2) müssen wir für Gruppen nicht explizit fordern, denn sie folgt ja schon wie in (8.3), da $f(e_1)$ als Element der Gruppe $(G_2, \square)$ notwendig invertierbar ist.

Die Gruppen $(G_1, \star)$ und $(G_2, \square)$ heißen zueinander **isomorph**, wenn es zwischen ihnen einen Gruppenisomorphismus gibt. Wir schreiben dann $(G_1, \star) \cong (G_2, \square)$. △

Beachte: Satz 8.2 und Folgerung 8.3 gelten sinngemäß auch für Monoide und Gruppen.

Beispiel 8.7 (Homomorphismen von Halbgruppen und Gruppen).

(*i*) Es sei Σ eine nichtleere Menge und $(\Sigma^*, \circ)$ das Monoid der Tupel über Σ mit der Konkatenation $\circ$, siehe Beispiele 7.4 und 7.8. Die Abbildung $\#\colon (\Sigma^*, \|) \to (\mathbb{N}_0, +)$, die die Länge eines Tupels angibt, ist ein Monoidhomomorphismus, denn es gilt

$$\begin{aligned} &\#\big((x_1, \dots, x_n) \parallel (y_1, \dots, y_m)\big) = \#(x_1, \dots, x_n) + \#(y_1, \dots, y_m) = n + m \\ \text{und} \quad &\#() = 0. \end{aligned}$$

Genau dann, wenn Σ einelementig ist, ist # auch bijektiv, also ein Monoidisomorphismus.

(*ii*) Es seien X eine Menge und $(Y, \star)$ eine Halbgruppe oder ein Monoid oder eine Gruppe. Dann ist die Abbildung $\Phi\colon (Y^X, \star) \ni f \mapsto f(x_0) \in (Y, \star)$, die eine Funktion $f\colon X \to Y$ an einer festen Stelle $x_0 \in X$ auswertet, ein Homomorphismus von Halbgruppen bzw. Monoiden bzw. Gruppen, denn es gilt

$$\Phi(f \star g) = (f \star g)(x_0) = f(x_0) \star g(x_0) = \Phi(f) \star \Phi(g).$$

(*iii*) Für $a > 0$, $a \neq 1$, ist $\log_a\colon (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +)$ aufgrund des Logarithmusgesetzes

$$\log_a(x \cdot y) = \log_a(x) + \log_a(y)$$

ein Gruppenhomomorphismus. Weiter ist $\log_a$ bijektiv, also sogar ein Gruppenisomorphismus. Die Umkehrabbildung

$$(\mathbb{R}, +) \ni x \mapsto a^x \in (\mathbb{R}_{>0}, \cdot)$$

ist folglich ebenfalls ein Gruppenisomorphismus, erfüllt also insbesondere

$$a^{x+y} = a^x \cdot a^y.$$

(*iv*) Die Betrags-Abbildung

$$(\mathbb{C}, \cdot) \ni z \mapsto |z| \in (\mathbb{R}_{\geq 0}, \cdot)$$

ist ein Homomorphismus von Monoiden, denn es gilt $|w \cdot z| = |w| \cdot |z|$ für alle $w, z \in \mathbb{C}$. Ihre Einschränkung

$$(\mathbb{C}_{\neq 0}, \cdot) \ni z \mapsto |z| \in (\mathbb{R}_{>0}, \cdot)$$

ist ein Homomorphismus von Gruppen.

(*v*) Zwischen beliebigen Gruppen $(G_1, \star)$ und $(G_2, \square)$ gibt es immer den **trivialen Homomorphismus** (englisch: trivial homomorphism) $f\colon G_1 \ni a \mapsto f(a) \coloneqq e_2 \in G_2$. Für einige Paare von Gruppen ist das auch der einzig mögliche Homomorphismus. (**Quizfrage 8.1:** Können Sie ein Beispiel finden?)

(*vi*) Es sei $(H, \star)$ ein Monoid und $a \in H$ ein invertierbares Element. Dann ist die Abbildung

$$H \ni h \mapsto a \star h \star a' \in H, \tag{8.4}$$

genannt die **Konjugation mit a** (englisch: conjugation), ein Endomorphismus des Monoids $(H, \star)$. Ist $(G, \star)$ eine Gruppe, dann ist die Konjugation mit einem beliebigem $a \in G$

$$G \ni g \mapsto a \star g \star a' \in G$$

sogar ein Gruppenautomorphismus.

(*vii*) In einer Gruppe $(G, \cdot)$ ist für jedes feste $a \in G$ die Abbildung

$$(\mathbb{Z}, +) \ni n \mapsto a^n \in G$$

ein Gruppenhomomorphismus. Dieser ist surjektiv genau dann, wenn a ein Erzeuger (Definition 7.48) von G ist.

(*viii*) Für festes $n \in \mathbb{Z}$ ist die Abbildung (vgl. (7.10))

$$G \ni a \mapsto a^n \in G$$

in einer *abelschen* Gruppe $(G, \cdot)$ ein Gruppenendomorphismus. (**Quizfrage 8.2:** Wo geht die Kommutativität der Verknüpfung ein?)

(*ix*) Die sgn-Abbildung ist ein Gruppenhomomorphismus von der symmetrischen Gruppe S_n (für festes $n \in \mathbb{N}$) in die Gruppe $(\{\pm 1\}, \cdot)$, denn es gilt nach Satz 7.40

$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = (\operatorname{sgn} \sigma_1) \cdot (\operatorname{sgn} \sigma_2).$$

Genau für $n = 2$ ist sgn auch bijektiv, also ein Gruppenisomorphismus.

(*x*) Die in Beispiel 7.31 und Abbildung 7.1 vorgenommene „Identifikation" der symmetrischen Gruppe S_3 mit der Gruppe der Kongruenzabbildungen eines gleichseitigen Dreiecks stellt einen Gruppenisomorphismus dar.

(*xi*) Die Abbildung

$$\mathbb{R} \ni f(x) \coloneqq \exp(ix) \coloneqq \cos(x) + i \sin(x) \in \mathbb{C}$$

ist ein Gruppenhomomorphismus $(\mathbb{R}, +) \to (\mathbb{C}, \cdot)$, denn es gilt $f(x + y) = f(x) \cdot f(y)$, also

$$\begin{aligned} \exp(i(x+y)) &= \exp(ix) \cdot \exp(iy) \\ \Leftrightarrow \quad \cos(x+y) + i \sin(x+y) &= \big(\cos(x) + i \sin(x)\big) \cdot \big(\cos(y) + i \sin(y)\big). \end{aligned}$$

Nehmen wir den Real- bzw. Imaginärteil der linken und der rechten Seite, so ergeben sich daraus die **Additionstheoreme** (englisch: angle addition theorems) für die Winkelsumme

$$\cos(x+y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y), \tag{8.5a}$$

$$\sin(x+y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y). \tag{8.5b}$$

(*xii*) Die Abbildung

$$(\mathbb{Z}, \cdot) \ni n \mapsto 0 \in (\mathbb{Z}, \cdot)$$

ist ein Homomorphismus von Halbgruppen, aber kein Homomorphismus von Monoiden, denn $f(1) = 0 \neq 1$. △

Expertenwissen: Wann sind Halbgruppenhomomorphismen auch Monoidhomomorphismen?

Das Beispiel 8.7 (*xii*) zeigt, dass Homomorphismen von Halbgruppen tatsächlich nicht notwendigerweise das neutrale Element auf das neutrale Element abbilden, dass also (8.2) tatsächlich i. A. nicht aus der Strukturverträglichkeit (8.1) folgt. Es gibt aber ein Kriterium (für das Ziel-Monoid), das absichert, dass doch *jeder* Halbgruppenhomomorphismus auch ein Monoidhomomorphismus ist:

Es seien $(H_1, \star)$ und $(H_2, \square)$ Monoide mit den neutralen Elementen e_1 bzw. e_2. Dann sind äquivalent:

(i) Jeder Halbgruppenhomomorphismus $f\colon H_1 \longrightarrow H_2$ ist auch ein Monoidhomomorphismus.

(ii) Es gibt in $(H_2, \square)$ genau ein Element mit der Eigenschaft $y \square y = y$. (Dieses ist dann notwendigerweise gleich e_2.)

Quizfrage 8.3: Beweis?

Das folgende Resultat zeigt, dass Monoid- und Gruppenhomomorphismen inverse Elemente auf inverse Elemente abbilden:

Lemma 8.8 (Eigenschaften von Monoid- und Gruppenhomomorphismen)**.**

Es seien $(H_1, \star)$ und $(H_2, \square)$ Monoide. Weiter sei $f\colon H_1 \longrightarrow H_2$ ein Monoidhomomorphismus. Ist $a \in H_1$ invertierbar, dann ist auch $f(a) \in H_2$ invertierbar, und es gilt $f(a)' = f(a')$. Insbesondere gilt das auch für Gruppenhomomorphismen.

Beweis. Wir bezeichnen die neutralen Elemente von H_1 bzw. H_2 mit e_1 bzw. e_2. Es gilt

$$\begin{aligned} f(a') \square f(a) &= f(a' \star a) && \text{da } f \text{ Monoidhomomorphismus ist} \\ &= f(e_1) && \text{da } e_1 \text{ neutrales Element in } (H_1, \star) \text{ ist} \\ &= e_2 && \text{wegen (8.2).} \end{aligned}$$

Ganz analog folgt auch $f(a) \square f(a') = e_2$. Das zeigt, dass $f(a)$ invertierbar und dass $f(a')$ das zu $f(a)$ inverse Element ist. □

Quizfrage 8.4: Kann $f\colon (\mathbb{Z}, +) \ni n \mapsto n + 1 \in (\mathbb{Z}, +)$ ein Gruppenhomomorphismus sein?

Bemerkung 8.9 (zu Definitionen 8.1, 8.4 und 8.6, Lemma 8.8)**.**

Zwei zueinander isomorphe Halbgruppen bzw. Monoide bzw. Gruppen können und müssen, was ihre algebraischen Eigenschaften als Halbgruppen bzw. Monoide bzw. Gruppen angeht, nicht unterschieden werden. △

Expertenwissen: Transport von Struktur

Bisher haben wir gesehen, dass die Homomorphismus-Eigenschaft die Kompatibilität einer Abbildungen mit *bestehenden* Strukturen wie Halbgruppen oder Gruppen beschreibt. Wir können aber auch mittels einer Abbildung Struktur auf eine andere Menge übertragen, wobei die Abbildung dann automatisch zu einem Homomorphismus wird.

Es sei $(H_1, \star)$ eine Halbgruppe und $(H_2, \square)$ eine Menge mit einer Verknüpfung, von der keine weiteren Eigenschaften bekannt sind. Ist $f\colon H_1 \to H_2$ irgendeine surjektive strukturverträgliche Abbildung, die also (8.1) erfüllt, dann ist $\square$ automatisch assoziativ, also $(H_2, \square)$ eine Halbgruppe. Ist $(H_1, \star)$ ein Monoid mit neutralem Element e_1, so ist auch $(H_2, \square)$ ein Monoid, und zwar mit dem neutralen Element $e_2 \coloneqq f(e_1)$. Ist $(H_1, \star)$ eine Gruppe, so ist auch $(H_2, \square)$ eine Gruppe. Ist die Verknüpfung $\star$ kommutativ, dann auch $\square$. (**Quizfrage 8.5:** Beweis dieser Aussagen?)

Durch die Surjektivität und die Strukturverträglichkeit von f werden also Eigenschaften von $(H_1, \star)$ auf $(H_2, \square)$ transportiert.

Wir wollen nun Gruppenhomomorphismen genauer studieren.

Definition 8.10 (Bild und Kern eines Gruppenhomomorphismus)**.**
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen mit den neutralen Elementen e_1 bzw. e_2. Weiter sei $f\colon G_1 \to G_2$ ein Homomorphismus.

(i) Das **Bild** (englisch: image) von f ist definiert als

$$\operatorname{Bild}(f) \coloneqq \{f(a_1) \in G_2 \mid a_1 \in G_1\} = f(G_1). \tag{8.6}$$

(ii) Der **Kern** (englisch: kernel) von f ist definiert als

$$\operatorname{Kern}(f) \coloneqq \{a_1 \in G_1 \mid f(a_1) = e_2\} = f^{-1}(\{e_2\}). \tag{8.7}$$

△

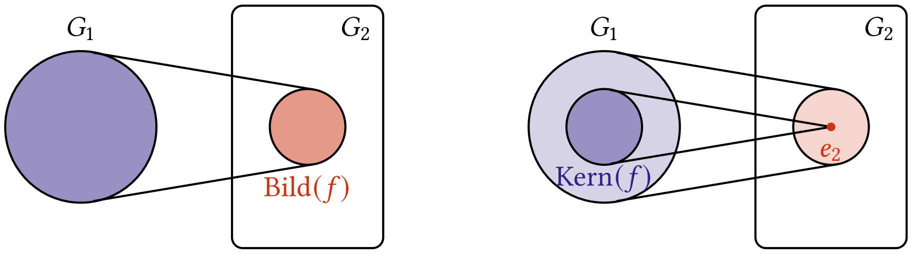


Abbildung 8.1.: Illustration des Bildes (links) und des Kerns (rechts) eines Gruppenhomomorphismus $f\colon G_1 \to G_2$, siehe Definition 8.10.

Lemma 8.11 (Bild und Kern sind Untergruppen)**.**
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen. Weiter sei $f\colon G_1 \to G_2$ ein Homomorphismus. Dann gilt:

(i) $\operatorname{Bild}(f)$ ist eine Untergruppe von $(G_2, \square)$.

(*ii*) $\operatorname{Kern}(f)$ ist eine Untergruppe von $(G_1, \star)$.

Beweis. Wir bezeichnen die neutralen Elemente von G_1 bzw. G_2 mit e_1 bzw. e_2.

Aussage (*i*): Wir überprüfen das Untergruppenkriterium (Satz 7.44). Nach Lemma 8.8 gilt $e_2 = f(e_1)$, also $e_2 \in \operatorname{Bild}(f)$ und $\operatorname{Bild}(f) \neq \emptyset$. Weiter seien a_2, b_2 irgendwelche Elemente in $\operatorname{Bild}(f)$. Wir müssen zeigen: $a_2 \mathbin{\square} b_2' \in \operatorname{Bild}(f)$.

Nach Definition von $\operatorname{Bild}(f)$ gibt es $a_1, b_1 \in G_1$ mit $f(a_1) = a_2$ und $f(b_1) = b_2$. Daher ist

$$
\begin{aligned}
a_2 \mathbin{\square} b_2' &= f(a_1) \mathbin{\square} f(b_1)' && \text{nach Voraussetzung} \\
&= f(a_1) \mathbin{\square} f(b_1') && \text{nach Lemma 8.8} \\
&= f(a_1 \star b_1') && \text{da } f \text{ Homomorphismus ist}
\end{aligned}
$$

und damit $a_2 \mathbin{\square} b_2' \in \operatorname{Bild}(f)$.

Aussage (*ii*): Wir überprüfen wiederum das Untergruppenkriterium. Es gilt $f(e_1) = e_2$, also $e_1 \in \operatorname{Kern}(f)$ und $\operatorname{Kern}(f) \neq \emptyset$. Weiter seien a_1, b_1 irgendwelche Elemente in $\operatorname{Kern}(f)$. Wir müssen zeigen: $a_1 \star b_1' \in \operatorname{Kern}(f)$.

$$
\begin{aligned}
f(a_1 \star b_1') &= f(a_1) \mathbin{\square} f(b_1') && \text{da } f \text{ Homomorphismus ist} \\
&= f(a_1) \mathbin{\square} f(b_1)' && \text{nach Lemma 8.8} \\
&= e_2 \mathbin{\square} e_2' && \text{da } a_1, b_1 \in \operatorname{Kern}(f) \text{ liegen} \\
&= e_2 && \text{da } e_2' = e_2 \text{ ist.}
\end{aligned}
$$

Damit ist $a_1 \star b_1' \in \operatorname{Kern}(f)$ gezeigt. □

Beispiel 8.12 (Bild und Kern sind Untergruppen).

(*i*) Für die Abbildung $\#\colon (\Sigma^*, \circ) \to (\mathbb{N}_0, +)$ aus Beispiel 8.7 gilt:

$$\operatorname{Bild}(\#) = \mathbb{N}_0 \quad \text{und} \quad \operatorname{Kern}(\#) = \{()\},$$

wobei () das leere Tupel kennzeichnet. Mit anderen Worten: Das leere Tupel ist das einzige Tupel der Länge 0.

(*ii*) Für die Abbildung $\operatorname{sgn}\colon (S_n, \circ) \to (\{\pm 1\}, \cdot)$ aus Beispiel 8.7 gilt im Fall $n \geqslant 2$:

$$\operatorname{Bild}(\operatorname{sgn}) = \{\pm 1\} \quad \text{und} \quad \operatorname{Kern}(\operatorname{sgn}) = A_n,$$

die alternierende Gruppe, vgl. (7.20). Mit anderen Worten: Die alternatierende Gruppe besteht genau aus den geraden Permutationen.

(*iii*) Für die Abbildung $f\colon (\mathbb{R}_{\neq 0}, \cdot) \ni x \mapsto x^2 \in (\mathbb{R}_{\neq 0}, \cdot)$ gilt

$$\operatorname{Bild}(f) = \mathbb{R}_{>0} \quad \text{und} \quad \operatorname{Kern}(f) = \{\pm 1\}. \qquad \triangle$$

Das folgende Resultat zeigt, dass Gruppenhomomorphismen genau dann injektiv sind, wenn ihr Kern die minimale Größe hat:

Lemma 8.13 (Charakterisierung der Injektivität von Gruppenhomomorphismen).
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen mit den neutralen Elementen e_1 bzw. e_2. Weiter sei $f : G_1 \to G_2$ ein Homomorphismus. Dann sind äquivalent:

(i) f ist injektiv.

(ii) $\operatorname{Kern}(f) = \{e_1\}$.

(iii) Die einzige Lösung der Gleichung $f(a) = e_2$ ist $a = e_1$.

Beachte: Um die Injektivität einer beliebigen Abbildung zwischen zwei Mengen zu zeigen, müssen wir sicherstellen, dass niemals zwei verschiedene Elemente der Definitionsmenge auf dasselbe Element in der Zielmenge abgebildet werden (Definition 6.10). Wenn wir aber wissen, dass diese Abbildung ein Gruppenhomomorphismus $G_1 \to G_2$ ist, vereinfacht sich dieser Nachweis erheblich. Wir müssen dann nur noch zeigen, dass neben dem neutralen Element $e_1 \in G_1$ kein weiteres Element auf das neutrale Element $e_2 \in G_2$ abgebildet wird.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Nach Lemma 8.8 gilt $f(e_1) = e_2$. Ist f injektiv, dann wird kein weiteres Element von G_1 auf e_2 abgebildet, also gilt $\operatorname{Kern}(f) = \{e_1\}$.

Aussage (ii) $\Rightarrow$ Aussage (i): Umgekehrt gelte $\operatorname{Kern}(f) = \{e_1\}$. Es seien weiter $a, b \in G_1$ mit $f(a) = f(b)$. Dann folgt

$$\begin{aligned} f(a \star b') &= f(a) \square f(b') \\ &= f(a) \square f(b)' \\ &= f(a) \square f(a)' \\ &= e_2, \end{aligned}$$

also $a \star b' \in \operatorname{Kern}(f) = \{e_1\}$. Daher muss $a \star b' = e_1$ gelten, also wegen der Eindeutigkeit inverser Elemente $a = b$. Das zeigt die Injektivität von f.

Die Äquivalenz von Aussage (ii) und Aussage (iii) ist einfach zu sehen, weil $\operatorname{Kern}(f)$ gerade aus den Lösungen der Gleichung $f(a) = e_2$ besteht und nach Lemma 8.8 $f(e_1) = e_2$ gilt. □

Wir zeigen nun, dass Gruppenhomomorphismen bereits durch ihre Bilder auf einem Erzeugendensystem eindeutig festgelegt sind.

Satz 8.14 (Eindeutigkeitssatz für Gruppenhomomorphismen).
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen. Weiter seien $f, g : G_1 \to G_2$ Homomorphismen, und für ein Erzeugendensystem $E \subseteq G_1$ gelte $f(e) = g(e)$ für alle $e \in E$. Dann ist $f = g$.

Beweis. Es sei $a \in G_1$ beliebig. Dann gibt es nach (7.22) ein $n \in \mathbb{N}_0$ und $a_1, \ldots, a_n \in E \cup E'$ mit

$$a = a_1 \star \cdots \star a_n.$$

Für diejenigen a_j mit $a_j \in E$ gilt $f(a_j) = g(a_j)$ nach Voraussetzung. Für diejenigen a_j mit $a_j \in E'$ gilt $a_j' \in E$ und daher nach Lemma 8.8 und nach Vorausetzung $f(a_j)' = f(a_j') = g(a_j') = g(a_j)'$, also wiederum $f(a_j) = g(a_j)$. Wir haben also

$$f(a) = f(a_1 \star \cdots \star a_n) \qquad \text{aufgrund der Darstellung von } a$$

$$
\begin{aligned}
&= f(a_1) \square \cdots \square f(a_n) && \text{da } f \text{ Homomorphismus ist} \\
&= g(a_1) \square \cdots \square g(a_n) && \text{nach Voraussetzung} \\
&= g(a_1 \star \cdots \star a_n) && \text{da } g \text{ Homomorphismus ist} \\
&= g(a) && \text{aufgrund der Darstellung von } a.
\end{aligned}
$$
□

Ende der Vorlesung 11

§ 8.1 Normalteiler und Faktorgruppen

Wir hatten in Satz 7.55 und Folgerung 7.57 gesehen, dass jede Untergruppe $(U, \star)$ einer Gruppe $(G, \star)$ zwei Äquivalenzrelationen auf G induziert, deren Äquivalenzklassen die Gestalt $a \star U$ bzw. $U \star a$ haben. Diese von U induzierten Äquivalenzrelationen sind (außer in Abelschen Gruppen, Folgerung 7.57) i. A. verschieden (Beispiel 7.59). Wir betrachten im Folgenden aber den Fall, dass sie übereinstimmen:

Definition 8.15 (Normalteiler)**.**
Es sei $(G, \star)$ eine Gruppe. Eine Untergruppe N heißt eine **normale Untergruppe** (englisch: normal subgroup) oder ein **Normalteiler** von $(G, \star)$, wenn gilt:

$$a \star N = N \star a \quad \text{für alle } a \in G. \tag{8.8}$$

Manchmal notiert man die Eigenschaft, dass $(N, \star)$ ein Normalteiler der Gruppe $(G, \star)$ ist, als $(N, \star) \trianglelefteq (G, \star)$. △

Beachte: In (8.8) steht die Gleichheit der beiden Mengen $a \star N$ und $N \star a$. Es wird nicht gefordert, dass $a \star n = n \star a$ für alle $n \in N$ gilt!

Beachte: Es ist hinreichend, (8.8) für alle a aus einem Erzeugendensystem von G zu überprüfen. (**Quizfrage 8.6:** Warum?)

Bemerkung 8.16 (Normalteiler)**.**

(*i*) Die definierende Gleichung (8.8) können wir auch so lesen, dass Normalteiler N genau diejenigen Untergruppen einer Gruppe sind, die gegenüber der Konjugation (8.4) mit beliebigen Gruppenelementen invariant sind:

$$a \star N \star a' \subseteq N \quad \text{für alle } a \in G. \tag{8.9}$$

(**Quizfrage 8.7:** Warum folgt aus der Inklusion (8.9) bereits die Gleichheit $a \star N \star a' = N$?)

(*ii*) Die Relation „ist Normalteiler von“ ist zwar reflexiv und antisymmetrisch, aber im Gegensatz zur Relation „ist Untergruppe von“ i. A. nicht transitiv! Im Gegensatz zur Untergruppenrelation ist die Normalteilerrelation also keine Ordnungsrelation.

Als Beispiel betrachten wir die alternierende Gruppe A_4 vom Grad 4, deren zur Kleinschen Vierergruppe (Beispiel 7.22) isomorphe Untergruppe

$$K_4 = \left\{ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \right\}$$

sowie deren Untergruppe

$$H_4 = \left\{ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \right\}.$$

Wir bestätigen zunächst, dass tatsächlich $H_4 \trianglelefteq K_4$ und $K_4 \trianglelefteq A_4$ gelten. Da K_4 kommutativ ist, ist die Untergruppe H_4 ein Normalteiler von K_4. Wir können außerdem überprüfen, dass $\sigma \circ K_4 = K_4 \circ \sigma$ für alle $\sigma \in A_4$ gilt, also $K_4 \trianglelefteq A_4$. Dafür ist es ausreichend, dies für ein Erzeugendensystem von A_4 zu zeigen, zum Beispiel für

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \quad \text{und} \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

Allerdings ist H_4 kein Normalteiler von A_4, denn beispielsweise ist

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}}_{\sigma \in A_4} \circ \underbrace{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}}_{\in H_4} \circ \underbrace{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}}_{\sigma^{-1} \in A_4} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \notin H_4.$$

Aus (8.9) folgt also tatsächlich $K_4 \ntrianglelefteq A_4$.

Zusammenfassend haben wir also $H_4 \trianglelefteq K_4 \trianglelefteq A_4$, aber $H_4 \ntrianglelefteq A_4$. △

Beispiel 8.17 (Normalteiler)**.**

(i) In jeder Gruppe $(G, \star)$ sind die trivialen Untergruppen $\{e\}$ und G Normalteiler.

(ii) In einer abelschen Gruppe $(G, \star)$ ist *jede* Untergruppe ein Normalteiler (Folgerung 7.57).

(iii) Das **Zentrum**[27] (englisch: center)

$$Z \coloneqq \{z \in G \mid a \star z = z \star a \text{ für alle } a \in G\} \tag{8.10}$$

einer Gruppe $(G, \star)$ ist ein Normalteiler.

(iv) Der **Kommutator** (englisch: commutator) der Elemente a, b einer Gruppe $(G, \star)$ ist definiert als

$$[a, b] \coloneqq a \star b \star a' \star b' = (a \star b) \star (b \star a)'. \tag{8.11}$$

Quizfrage 8.8: Was ist $[a, b]'$? **Beachte:** a und b kommutieren genau dann (bzgl. $\star$), wenn $[a, b]$ das neutrale Element e der Gruppe $(G, \star)$ ergibt. (**Quizfrage 8.9:** Klar?)

Die **Kommutator(unter)gruppe** (englisch: commutator subgroup) der Gruppe $(G, \star)$ ist die von den Kommutatoren von G erzeugte Untergruppe, also

$$\big\langle \big\{ [a, b] \,\big|\, a, b \in G \big\} \big\rangle. \tag{8.12}$$

Sie wird kurz auch in der Form $\big\langle [G, G] \big\rangle$ oder (ungenau) als $[G, G]$ notiert. Die Kommutatorgruppe ist ein Normalteiler von $(G, \star)$. △

Lemma 8.18 (Kerne von Gruppenhomomorphismen sind Normalteiler)**.**
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen und $f\colon G_1 \to G_2$ ein Homomorphismus. Dann gilt:

[27] Das Zentrum einer Gruppe besteht also aus denjenigen Elementen, die mit allen Gruppenelementen kommutieren.

(*i*) Die Elemente von G_1, die denselben Funktionswert wie $a \in G_1$ haben, sind genau die Elemente der Nebenklasse von $\mathrm{Kern}(f)$ zu a:

$$f^{-1}(\{f(a)\}) = a \star \mathrm{Kern}(f) = \mathrm{Kern}(f) \star a.$$

(*ii*) $\mathrm{Kern}(f)$ ist ein Normalteiler von G_1.

Beweis. Wir bezeichnen die neutralen Elemente von G_1 und G_2 mit e_1 bzw. e_2.

Wir zeigen zunächst die Aussage (*i*) in mehreren Schritten.

Schritt 1: $f^{-1}(\{f(a)\}) \subseteq \mathrm{Kern}(f) \star a$:
Es sei $b \in f^{-1}(\{f(a)\})$, also $f(b) = f(a)$. Dann gilt also

$$\begin{aligned} e_2 &= f(b) \square f(a)' \\ &= f(b) \square f(a') && \text{nach Lemma 8.8} \\ &= f(b \star a') && \text{da } f \text{ Homomorphismus ist.} \end{aligned}$$

Das heißt aber, dass $b \star a' \in f^{-1}(\{e_2\}) = \mathrm{Kern}(f)$ liegt. Mit anderen Worten: $b \in \mathrm{Kern}(f) \star a$.

Schritt 2: $\mathrm{Kern}(f) \star a \subseteq f^{-1}(\{f(a)\})$:
Es sei $b \in \mathrm{Kern}(f)$. Wir müssen $b \star a \in f^{-1}(\{f(a)\})$ zeigen, also $f(b \star a) = f(a)$. Das folgt aber sofort aus

$$\begin{aligned} f(b \star a) &= f(b) \square f(a) && \text{da } f \text{ Homomorphismus ist} \\ &= e_2 \square f(a) && \text{da } b \in \mathrm{Kern}(f) \text{ ist} \\ &= f(a). \end{aligned}$$

Schritt 3: $f^{-1}(\{f(a)\}) \subseteq a \star \mathrm{Kern}(f)$:
Ganz analog zu Schritt 1 gilt auch

$$\begin{aligned} e_2 &= f(a)' \square f(b) \\ &= f(a') \square f(b) && \text{nach Lemma 8.8} \\ &= f(a' \star b) && \text{da } f \text{ Homomorphismus ist.} \end{aligned}$$

Das heißt aber $a' \star b \in f^{-1}(\{e_2\}) = \mathrm{Kern}(f)$ und daher $b \in a \star \mathrm{Kern}(f)$.

Schritt 4: $a \star \mathrm{Kern}(f) \subseteq f^{-1}(\{f(a)\})$:
Es sei $b \in \mathrm{Kern}(f)$. Wir müssen $a \star b \in f^{-1}(\{f(a)\})$ zeigen, also $f(a \star b) = f(a)$. Das folgt aber sofort aus

$$\begin{aligned} f(a \star b) &= f(a) \square f(b) && \text{da } f \text{ Homomorphismus ist} \\ &= f(a) \square e_2 && \text{da } b \in \mathrm{Kern}(f) \text{ ist} \\ &= f(a). \end{aligned}$$

Aus Lemma 8.11 wissen wir, dass $\operatorname{Kern}(f)$ eine Untergruppe von G_1 ist. Aus Aussage (i) folgt $a \star \operatorname{Kern}(f) = \operatorname{Kern}(f) \star a$ für alle $a \in G_1$, also ist $\operatorname{Kern}(f)$ ein Normalteiler von G_1. Das zeigt Aussage (ii). □

Bemerkung 8.19 (Urbilder von Normalteilern sind Normalteiler)**.**

Es gilt sogar folgende Verallgemeinerung der Aussage (ii) aus Lemma 8.18: Urbilder von Normalteilern unter Gruppenhomomorphismen sind Normalteiler. △

Quizfrage 8.10: Ist auch das Bild eines Gruppenhomomorphismus immer ein Normalteiler?

Lemma 8.20 (Durchschnitt von Normalteilern, vgl. Lemma 7.47 zu Untergruppen)**.**

Es sei $(G, \star)$ eine Gruppe.

(i) Ist $(N_i)_{i \in I}$ eine nichtleere Familie von Normalteilern von $(G, \star)$, dann ist auch $\bigcap_{i \in I} N_i$ ein Normalteiler von G.

(ii) Ist $\mathcal{N}$ eine nichtleere Menge von Normalteilern von $(G, \star)$, dann ist auch $\bigcap \mathcal{N}$ ein Normalteiler von $(G, \star)$.

Beweis. Der Beweis ist Gegenstand der Übung. □

Der nun folgende Satz zeigt: Wenn $(N, \star)$ ein Normalteiler der Gruppe $(G, \star)$ ist, dann können wir die Faktormenge $G / N \coloneqq G / \overset{N}{\sim}$ mit einer Gruppenverknüpfung $\tilde{\star}$ ausstatten, die mit $\star$ kompatibel ist. Aus der Faktormenge wird damit die **Faktorgruppe** (englisch: factor group) oder **Quotientengruppe** (englisch: quotient group) **von G nach N**, die auch als **G modulo N** bezeichnet wird. Wir sagen auch: „Der Normalteiler N wird aus der Gruppe $(G, \star)$ ausfaktorisiert."

Satz 8.21 (Faktorgruppe)**.**

(i) Es sei $(G, \star)$ eine Gruppe und $(N, \star)$ einer ihrer Normalteiler. Dann gilt:

(a) Auf der Faktormenge[28]

$$G / N = \big\{[a] = a \star N \,\big|\, a \in G\big\}$$

ist $\tilde{\star}$, definiert als

$$[a] \mathbin{\tilde{\star}} [b] \coloneqq [a \star b] \quad \text{für } a, b \in G, \tag{8.13}$$

eine assoziative Verknüpfung, bzgl. der $(G / N, \tilde{\star})$ eine Gruppe bildet. Das neutrale Element ist $[e] = N$, und für die Inversen gilt $[a]' = [a']$.

[28] Aus Gründen der Lesbarkeit schreiben wir für die Äquivalenzklasse $a \star N = N \star a$ (Nebenklasse) auch $[a]$.

(b) Die **kanonische Surjektion** (englisch: canonical surjection) **von** G **auf** $G\,/\,N$[29]

$$\pi\colon \begin{cases} G \to G\,/\,N \\ a \mapsto [a], \end{cases} \tag{8.14}$$

die jedem Element $a \in G$ seine Nebenklasse $[a] = a \star N$ zuordnet, ist ein surjektiver Gruppenhomomorphismus. Es gilt $\operatorname{Kern}(\pi) = N$.

(c) Wenn $(G, \star)$ abelsch ist, dann auch $(G\,/\,N, \tilde{\star})$.

(ii) Es sei $(G, \star)$ eine Gruppe und U irgendeine Untergruppe. Ist die Verknüpfung (8.13) auf der Menge der Linksnebenklassen $G\,/\,U$ (oder auf der Menge der Rechtsnebenklassen $U \setminus G$) wohldefiniert, dann ist U notwendigerweise ein Normalteiler von G.

Beweis. Aussage (i): Wir zeigen zunächst Aussage (a) in mehreren Schritten.

Schritt 1: Wir müssen zunächst zeigen, dass $\tilde{\star}$ überhaupt eine Verknüpfung auf $G\,/\,N$ darstellt, also dass (8.13) wohldefiniert ist, da wir dort ja Bezug auf konkrete Repräsentanten $a, b \in G$ der Äquivalenzklassen $[a], [b]$ nehmen. Es seien also $a_1, a_2, b_1, b_2 \in G$ gegeben, wobei $[a_1] = [a_2]$ und $[b_1] = [b_2]$ angenommen wird, d. h., $a_1 \star N = a_2 \star N = N \star a_1 = N \star a_2$ und $b_1 \star N = b_2 \star N = N \star b_1 = N \star b_2$. Dann gilt

$$\begin{aligned} [a_1] \,\tilde{\star}\, [b_1] &= [a_1 \star b_1] && \text{per Definition von } \tilde{\star} \\ &= (a_1 \star b_1) \star N && \text{nach Definition der Äquivalenzklassen} \\ &= (a_1 \star b_1) \star (N \star N) && \text{da } N \text{ Untergruppe ist, also } N \star N = N \\ &= ((a_1 \star b_1) \star N) \star N && \text{da } \star \text{ assoziativ ist} \\ &= (N \star (a_1 \star b_1)) \star N && \text{da } N \text{ Normalteiler ist} \\ &= (N \star a_1) \star (b_1 \star N) && \text{da } \star \text{ assoziativ ist.} \end{aligned}$$

Ganz analog gilt auch

$$[a_2] \,\tilde{\star}\, [b_2] = (N \star a_2) \star (b_2 \star N),$$

und mit der Voraussetzung folgt $[a_1] \tilde{\star} [b_1] = [a_2] \tilde{\star} [b_2]$. Damit ist $\tilde{\star}$ als Verknüpfung auf $G\,/\,N$ wohldefiniert.

Schritt 2: Die Assoziativität von $\tilde{\star}$ folgt aus

$$\begin{aligned} &\big([a] \,\tilde{\star}\, [b]\big) \,\tilde{\star}\, [c] = [a \star b] \,\tilde{\star}\, [c] = [(a \star b) \star c] \\ \text{und}\quad &[a] \,\tilde{\star}\, \big([b] \,\tilde{\star}\, [c]\big) = [a] \,\tilde{\star}\, [b \star c] = [a \star (b \star c)]. \end{aligned}$$

Aufgrund der Assoziativität von $\star$ sind die Äquivalenzklassen gleich. Damit haben wir zunächst $(G\,/\,N, \tilde{\star})$ als Halbgruppe bestätigt.

Schritt 3: Wir notieren das neutrale Element von G als e und zeigen, dass $[e] = e \star N = N$ das neutrale Element von $(G\,/\,N, \tilde{\star})$ ist. Dazu sei $a \in G$ beliebig. Dann gilt gemäß Definition (8.13)

$$[e] \,\tilde{\star}\, [a] = [e \star a] = [a] \quad \text{sowie} \quad [a] \,\tilde{\star}\, [e] = [a \star e] = [a].$$

Also ist $(G\,/\,N, \tilde{\star})$ ein Monoid mit neutralem Element $[e]$.

[29] Die kanonische Surjektion als Abbildung auf eine Faktormenge einer Äquivalenzrelation wurde in (6.6) schon eingeführt.

Schritt 4: Nun zeigen wir, dass jedes $[a] \in G / N$ invertierbar ist mit Inverser $[a]' = [a']$:

$$[a] \mathbin{\tilde{\star}} [a'] = [a \star a'] = [e] \quad \text{sowie} \quad [a'] \mathbin{\tilde{\star}} [a] = [a' \star a] = [e].$$

Aussage (b): Die Eigenschaft, ein Gruppenhomomorphismus zu sein, bedeutet $\pi(a \star b) = \pi(a) \mathbin{\tilde{\star}} \pi(b)$. Nach Definition von π heißt das aber $[a \star b] = [a] \mathbin{\tilde{\star}} [b]$, was gerade die Definition von $\tilde{\star}$ war. Die Surjektivität von π ist klar, denn ein beliebiges Element $[a]$ von G / N ist gerade das Bild von a unter π. Es gilt $\operatorname{Kern}(\pi) = \pi^{-1}(\{[e]\}) = N$.

Aussage (c): Falls $(G, \star)$ abelsch ist, dann gilt

$$[a] \mathbin{\tilde{\star}} [b] = [a \star b] = [b \star a] = [b] \mathbin{\tilde{\star}} [a],$$

also ist auch $(G / N, \tilde{\star})$ abelsch.

Nun zur umgekehrten Aussage *(ii)*: Es sei dazu U irgendeine Untergruppe von G. Wir führen den Beweis nur für den Fall, dass (8.13) auf der Menge der Linksnebenklassen G / U wohldefiniert ist. Es seien dazu $a \in G$ und $u \in U$ beliebig. Dann gilt $[u] = U = [e]$, und die Wohldefiniertheit liefert

$$[a] = [e \star a] = [e] \mathbin{\tilde{\star}} [a] = [u] \mathbin{\tilde{\star}} [a] = [u \star a],$$

also $a \star U = u \star a \star U$ oder $U = a' \star u \star a \star U$. Das heißt aber, dass $a' \star u \star a \in U$ liegt. (**Quizfrage 8.11:** Warum?) Da $u \in U$ beliebig war, haben wir $a' \star U \star a \subseteq U$ für alle $a \in G$. Nach Bemerkung 8.16 ist U ein Normalteiler von G. □

Bemerkung 8.22 (Faktorgruppe)**.**

Praktisch können wir die Faktorgruppe $(G / N, \tilde{\star})$ benutzen, um wie in der Gruppe $(G, \star)$ zu „rechnen“, wobei jedoch Elemente a, b in derselben Äquivalenzklasse (für die also $b \in a \star N$ gilt) nicht mehr unterschieden, sondern miteinander identifiziert werden. Die Faktorgruppe $(G / N, \tilde{\star})$ ist also eine „gröbere Version“ der Gruppe $(G, \star)$. Wegen $[a] \mathbin{\tilde{\star}} [b] = [a \star b]$ rechnen wir mit Nebenklassen, indem wir stellvertretend mit Repräsentanten rechnen. △

Beispiel 8.23 (Faktorgruppe)**.**

(i) Es sei $(G, \star)$ eine beliebige Gruppe. Dann ist $N = \{e\}$, eine der beiden trivialen Untergruppen von G, nach Beispiel 8.17 ein Normalteiler. Die zugehörige Faktorgruppe $(G / \{e\}, \tilde{\star})$ ist isomorph zur Ausgangsgruppe $(G, \star)$ selbst.

(ii) Es sei $(G, \star)$ eine beliebige Gruppe und $N = G$ die andere triviale Untergruppe von G. G ist nach Beispiel 8.17 ein Normalteiler. Die zugehörige Faktorgruppe $(G / G, \tilde{\star})$ ist isomorph zu $(\{e\}, \star)$.

(iii) Es sei $(G, \star)$ eine beliebige Gruppe und K die Kommutatoruntergruppe von G (Beispiel 8.17). Dann ist die Faktorgruppe $(G / K, \tilde{\star})$ kommutativ.

Tatsächlich ist $(G / N, \tilde{\star})$ genau dann kommutativ, wenn der ausfaktorisierte Normalteiler N die Kommutatoruntergruppe von G enthält.

(*iv*) Für $m \in \mathbb{N}$ ist $m\mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$. Da $(\mathbb{Z}, +)$ abelsch ist, ist jede dieser Untergruppen ein Normalteiler. Die Elemente der Faktorgruppe $(\mathbb{Z}\,/\,m\mathbb{Z}, \tilde{+})$ sind die Nebenklassen von $m\mathbb{Z}$, also die Mengen der Form $[a] = a + m\mathbb{Z}$, vgl. Beispiel 7.59. Es gilt

$$[a] \,\tilde{+}\, [b] = [a + b].$$

Die Faktorgruppe $(\mathbb{Z}\,/\,m\mathbb{Z}, \tilde{+})$ ist isomorph zu einer uns bereits bekannten Gruppe, nämlich zur additiven Gruppe modulo m $(\mathbb{Z}_m, +_m)$ aus Beispiel 7.22 mittels des Isomorphismus $[a] \mapsto$ natürlicher Repräsentant von a in $\mathbb{Z}_m$. Beispielsweise können wir für $m = 5$ wie folgt rechnen:

$$\begin{array}{cccccc} \text{in } (\mathbb{Z}\,/\,5\mathbb{Z}, \tilde{+}) & [-21] & \tilde{+} & [9] & = & [-12] \\ & \downarrow & & \downarrow & & \downarrow \\ \text{in } (\mathbb{Z}_5, +_5) & 4 & +_5 & 4 & = & 3 \end{array}$$

(*v*) In der abelschen Gruppe $(\mathbb{Q}_{\neq 0}, \cdot)$ ist die Untergruppe $(\{\pm 1\}, \cdot)$ ein Normalteiler. Die Elemente der Faktorgruppe sind die Nebenklassen

$$[a] = a \cdot \{\pm 1\} = \{a, -a\}$$

für $a \in \mathbb{Q}_{\neq 0}$. Ein mögliches Repräsentantensystem sind die positiven rationalen Zahlen $\mathbb{Q}_{>0}$. Durch $\mathbb{Q}_{\neq 0}\,/\,\{\pm 1\}$ wird also „das Vorzeichen ausfaktorisiert". Dieselbe Konstruktion können wir in $\mathbb{R}_{\neq 0}$ und $\mathbb{C}_{\neq 0}$ durchführen. △

Bemerkung 8.24 (Normalteiler sind genau die Kerne von Gruppenhomomorphismen)**.**
Es sei $(G_1, \star)$ eine Gruppe.

(*i*) Nach Lemma 8.18 ist die Untergruppe $\operatorname{Kern}(f)$ für jeden beliebigen Gruppenhomomorphismus $f\colon G_1 \to G_2$ in irgendeine Gruppe $(G_2, \square)$ immer ein Normalteiler von $(G_1, \star)$.

(*ii*) Umgekehrt gilt auch, dass jeder Normalteiler N von $(G_1, \star)$ der Kern eines Gruppenhomomorphismus ist. Dazu wählen wir einfach $G_2 \coloneqq (G_1\,/\,N, \tilde{\star})$ als Zielgruppe und die kanonische Surjektion $\pi\colon G_1 \to G_1\,/\,N$ als Gruppenhomomorphismus. Dann gilt $\operatorname{Kern}(\pi) = N$. △

§ 8.2 Der Homomorphiesatz für Gruppen

Mit Hilfe des Wissens aus § 8.1 können wir nun die Struktur von Gruppenhomomorphismen analysieren. Der folgende Struktursatz besagt, dass ein Gruppenhomomorphismus $f\colon G_1 \to G_2$ „nebenklassenweise" wirkt. Er bildet also eine gesamte Nebenklasse von $\operatorname{Kern}(f)$ auf ein- und dasselbe Element von G_2 ab und verschiedene Nebenklassen auf verschiedene Elemente.[30] Das geschieht zudem strukturverträglich. Dadurch ist das $\operatorname{Bild}(f)$ eines solchen Gruppenhomomorphismus bereits im Wesentlichen (d. h. bis auf Isomorphie) festgelegt ist durch $(G_1, \star)$ und den Normalteiler $\operatorname{Kern}(f)$.

[30] Anders ausgedrückt: Die Fasern (Definition 6.6) von f sind gerade die Nebenklassen von $\operatorname{Kern}(f)$.

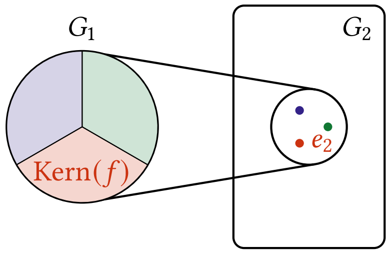


Abbildung 8.2.: Illustration des Homomorphiesatzes für Gruppenhomomorphismen 8.25. Alle Elemente einer Nebenklasse des Normalteilers Kern(f) werden auf ein- und dasselbe Element von G_2 abgebildet und verschiedene Nebenklassen auf verschiedene Elemente.

Satz 8.25 (Homomorphiesatz für Gruppen[31]).
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen. Weiter sei $f\colon G_1 \to G_2$ ein Homomorphismus. Dann gilt

$$G_1 \,/\, \mathrm{Kern}(f) \cong \mathrm{Bild}(f) \tag{8.15a}$$

mit dem Isomorphismus

$$I([a]) \coloneqq f(a) \quad \text{für } [a] = a \star \mathrm{Kern}(f) \in G_1 \,/\, \mathrm{Kern}(f). \tag{8.15b}$$

Beweis. Wir bezeichnen die neutralen Elemente von G_1 und G_2 mit e_1 bzw. e_2.

Wir definieren $I\colon G_1 \,/\, \mathrm{Kern}(f) \to \mathrm{Bild}(f)$ wie in (8.15).

Schritt 1: Wir müssen zunächst zeigen, dass I als Abbildung wohldefiniert ist, da wir in der Definition (8.15b) Bezug auf den konkreten Repräsentanten $a \in G_1$ nehmen.

Es seien dazu $a, b \in G_1$ gegeben mit $a \overset{U}{\sim} b$ für $U = \mathrm{Kern}(f)$, d. h., $a \star \mathrm{Kern}(f) = b \star \mathrm{Kern}(f)$. Dann folgt

$$\begin{aligned} f(a \star \mathrm{Kern}(f)) &= f(a) \square f(\mathrm{Kern}(f)) && \text{da } f \text{ Homomorphismus ist} \\ &= \{f(a)\} && \text{da } f(\mathrm{Kern}(f)) = \{e_2\} \text{ gilt} \end{aligned}$$

und analog $f(b \star \mathrm{Kern}(f)) = \{f(b)\}$. Aus $a \star \mathrm{Kern}(f) = b \star \mathrm{Kern}(f)$ folgt also $f(a) = f(b)$. Außerdem ist nach Definition von I klar, dass I in Bild(f) abbildet. Damit ist I wohldefiniert.

Schritt 2: Als nächstes zeigen wir, dass I ein Homomorphismus ist. In der Tat gilt

$$\begin{aligned} I([a] \mathbin{\tilde\star} [b]) &= I([a \star b]) && \text{nach Definition (8.13) von } \tilde\star \\ &= f(a \star b) && \text{nach Definition von } I \\ &= f(a) \square f(b) && \text{da } f \text{ Homomorphismus ist} \\ &= I([a]) \square I([b]) && \text{nach Definition von } I. \end{aligned}$$

[31]englisch: fundamental theorem on group homomorphisms

Schritt 3: Es bleibt zu zeigen, dass I surjektiv und injektiv ist. Wenn $a_2 \in \operatorname{Bild}(f)$ ist, dann existiert $a_1 \in G_1$ mit

$$a_2 = f(a_1) = I([a_1]).$$

Das zeigt die Surjektivität von I.

Für die Injektivität genügt es nach Lemma 8.13 zu zeigen, dass $\operatorname{Kern}(I)$ nur aus dem neutralen Element des Definitionsbereichs $G_1 / \operatorname{Kern}(f)$ besteht, d. h., aus $[e_1] = \operatorname{Kern}(f)$, vgl. Satz 8.21. Es gilt

$$\begin{aligned}
&\operatorname{Kern}(I) \\
&\quad = \big\{[a] \in G_1 / \operatorname{Kern}(f) \,\big|\, I([a]) = e_2\big\} && \text{nach Definition von } \operatorname{Kern}(I) \\
&\quad = \big\{[a] \in G_1 / \operatorname{Kern}(f) \,\big|\, f(a) = e_2\big\} && \text{nach Definition von } I \\
&\quad = \big\{[a] \in G_1 / \operatorname{Kern}(f) \,\big|\, a \in \operatorname{Kern}(f)\big\} && \text{nach Definition von } \operatorname{Kern}(f) \\
&\quad = \big\{a \star \operatorname{Kern}(f) \,\big|\, a \in \operatorname{Kern}(f)\big\} && \text{wegen } [a] = a \star \operatorname{Kern}(f) \\
&\quad = \big\{\operatorname{Kern}(f)\big\} && \text{da } \operatorname{Kern}(f) \text{ Untergruppe ist.} \quad \square
\end{aligned}$$

Wir stellen den Homomorphiesatz für Gruppen 8.25 auch noch einmal schematisch mit Hilfe eines kommutativen Diagrammes dar.[32] Dazu sei $i\colon \operatorname{Bild}(f) \ni a \mapsto a \in G_2$ der injektive Homomorphismus der kanonischen Einbettung.

$$\begin{tikzcd}
G_2 & G_1 \arrow[l, "f"'] \arrow[d, "\pi"] \\
\operatorname{Bild}(f) \arrow[u, "i"] & G_1 / \operatorname{Kern}(f) \arrow[l, "I"]
\end{tikzcd}
\qquad\qquad
f = \underbrace{i}_{\text{einbetten}} \circ \overbrace{I}^{\text{isomorph abbilden}} \circ \underbrace{\pi}_{\text{vergröbern}}$$

Expertenwissen: universelle Eigenschaft von Faktorgruppen

Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen und $f\colon G_1 \to G_2$ ein Homomorphismus. Weiter sei N eine normale Untergruppe von G_1. Dann sind äquivalent:

(i) $N \subseteq \operatorname{Kern}(f)$, also $f(N) = \{e_2\}$.

(ii) Der Homomorphismus f **faktorisiert durch** (englisch: factors through) die kanonische Surjektion $\pi\colon G_1 \to G_1 / N$, d. h., es gibt einen eindeutig bestimmten Homomorphismus $g\colon G_1 / N \to G_2$ mit $f = g \circ \pi$.

$$\begin{tikzcd}
G_1 \arrow[r, "\pi"] \arrow[dr, "f"'] & G_1 / N \arrow[d, "g"] \\
& G_2
\end{tikzcd}$$

Diese Eigenschaft nennt sich die **universelle Eigenschaft von Faktorgruppen** (englisch: universal property of quotient groups), denn sie charakterisiert Faktorgruppen bis

[32] Ein solches Diagramm heißt **kommutativ** (englisch: commutative diagram), wenn alle Pfade mit demselben Ausgangs- und demselben Endpunkt dasselbe Ergebnis produzieren.

auf Isomorphie eindeutig. Gilt also die obige Äquivalenzaussage mit irgendeiner Gruppe H anstelle von G_1 / N und einem surjektiven Gruppenhomomorphismus $\pi\colon G_1 \to H$, dann ist H isomorph zu G_1 / N. Die universelle Eigenschaft ermöglicht es, Faktorgruppen (bis auf Isomorphie) zu definieren, ohne die konkrete Konstruktion über Nebenklassen zu verwenden.

Bemerkung 8.26 (mögliche Gruppenhomomorphismen)**.**
Wegen $G_1 / \operatorname{Kern}(f) \cong \operatorname{Bild}(f)$ legen die Normalteiler einer Gruppe $(G_1, \star)$ im Wesentlichen die möglichen Gruppenhomomorphismen fest, die von G_1 aus möglich sind. Zu jedem Normalteiler N gibt es einen natürlichen, surjektiven Gruppenhomomorphismus $\pi\colon G_1 \to G_1 / N$. Dieser ist gewissermaßen der „Prototyp" eines Gruppenhomomorphismus auf G_1, der N als Kern hat. Jeder andere Gruppenhomomorphismus $f\colon G_1 \to G_2$ mit $\operatorname{Kern}(f) = N$ ist dann nur eine „eingebettete" Version dieses Prototyps, also $f = i \circ \pi$ mit dem injektiven Homomorphismus $i\colon G_1 / N \to G_2$ der kanonischen Einbettung und $\operatorname{Bild}(i) = \operatorname{Bild}(f)$. △

Beispiel 8.27 (Homomorphiesatz für Gruppen)**.**

(*i*) Wir betrachten für festes $n \in \mathbb{N}$ die Abbildung $\operatorname{sgn}\colon S_n \to (\{\pm 1\}, \cdot)$, vgl. Beispiele 8.7 und 8.12. Es gilt $\operatorname{Kern}(\operatorname{sgn}) = A_n$. Für $n \geqslant 2$ sind die Elemente der Faktorgruppe $S_n / \operatorname{Kern}(\operatorname{sgn}) = S_n / A_n$ die beiden gleichmächtigen Nebenklassen

$$[\mathrm{id}] = \mathrm{id} \circ A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\} \quad \text{(gerade Permutationen)},$$
$$[\tau] = \tau \circ A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = -1\} \quad \text{(ungerade Permutationen)},$$

wobei τ irgendeine Transposition in S_n ist. Gemäß Homomorphiesatz 8.25 ist

$$S_n / \operatorname{Kern}(\operatorname{sgn}) = S_n / A_n \cong \operatorname{Bild}(\operatorname{sgn}) = \{\pm 1\}.$$

Es werden alle geraden Permutationen $A_n = \operatorname{Kern}(\operatorname{sgn})$ ausfaktorisiert.

Im Fall $n = 1$ gilt $A_1 = S_1$, daher gibt es nur die eine Nebenklasse

$$[\mathrm{id}] = \mathrm{id} \circ S_1 = \{\mathrm{id}\}.$$

Der Homomorphiesatz 8.25 besagt daher in diesem Fall

$$S_1 / \operatorname{Kern}(\operatorname{sgn}) = S_1 / A_1 \cong \operatorname{Bild}(\operatorname{sgn}) = \{1\}.$$

Ähnliches gilt im Fall $n = 0$.

(*ii*) Für die Abbildung $f\colon (\mathbb{R}_{\neq 0}, \cdot) \ni x \mapsto x^2 \in (\mathbb{R}_{\neq 0}, \cdot)$ aus Beispiel 8.12 und Beispiel 8.23 gilt

$$\mathbb{R}_{\neq 0} / \operatorname{Kern}(f) = \mathbb{R}_{\neq 0} / \{\pm 1\} \cong \operatorname{Bild}(f) = \mathbb{R}_{>0}.$$

Durch $\operatorname{Kern}(f) = \{\pm 1\}$ wird das Vorzeichen ausfaktorisiert.

(*iii*) Für die Abbildung $f\colon (\mathbb{C}_{\neq 0}, \cdot) \ni z \mapsto |z| \in (\mathbb{R}_{\neq 0}, \cdot)$ gilt

$$\mathbb{C}_{\neq 0} / \operatorname{Kern}(f) = \mathbb{R}_{\neq 0} / K \cong \operatorname{Bild}(f) = \mathbb{R}_{>0},$$

wobei $K \coloneqq \{z \in \mathbb{C}_{\neq 0}\} |z| = 1$ der Einheitskreis in $\mathbb{C}$ ist. Durch $\operatorname{Kern}(f) = K$ wird die Lage von z auf der Kreislinie mit Radius $|z|$ ausfaktorisiert. △

Ende der Vorlesung 12

Ende der Woche 6

§ 9 Ringe

Literatur: Bosch, 2014, Kapitel 5.1; Fischer, Springborn, 2020, Kapitel 2.3

Ein Ring ist eine algebraische Struktur mit zwei Verknüpfungen, die gewissen Gesetzmäßigkeiten folgen. In Anlehnung an das Leit-Beispiel $\mathbb{Z}$ mit den Verknüpfungen „Addition" und „Multiplikation" bezeichnen wir diese Verknüpfungen häufig mit $+$ und $\cdot$.

Definition 9.1 (Ring).
Ein **Ring** (englisch: ring) $(R, +, \cdot)$ ist eine Menge R mit zwei (inneren) Verknüpfungen $+$ („Addition") und $\cdot$ („Multiplikation"), die die folgenden Bedingungen erfüllen:

(*i*) $(R, +)$ ist eine abelsche Gruppe.

(*ii*) $(R, \cdot)$ ist eine Halbgruppe.

(*iii*) Es gelten die **Distributivgesetze** (englisch: distributive laws)

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \tag{9.1a}$$

$$(a + b) \cdot c = (a \cdot c) + (b \cdot c) \tag{9.1b}$$

für alle $a, b, c \in R$.

Ein Ring $(R, +, \cdot)$ heißt **kommutativ** (englisch: commutative ring), wenn die Halbgruppe $(R, \cdot)$ kommutativ ist.[33] In diesem Fall fallen die beiden Distributivgesetze (9.1a) und (9.1b) zusammen, sind also untereinander äquivalent. △

Wie in Gruppen in additiver Notation üblich (Bemerkung 7.20), bezeichnen wir das neutrale Element eines Ringes $(R, +, \cdot)$ bzgl. $+$ als **Nullelement** (englisch: additive identity) und schreiben dafür zunächst „0_R". Außerdem benennen wir das bzgl. $+$ inverse Element zu $a \in R$ mit $-a$. Die Bezeichnung $a - b$ steht für $a + (-b)$.

Falls die Halbgruppe $(R, \cdot)$ sogar ein Monoid ist, so bezeichnen wir das neutrale Element bzgl. $\cdot$ als **Einselement** (englisch: multiplicative identity) und schreiben dafür zunächst „1_R". In diesem Fall heißt $(R, +, \cdot)$ auch ein **Ring mit Eins** (englisch: ring with unity) oder ein **unitärer Ring** (englisch: unitary ring, unital ring). Existiert dann zu $a \in R$ bzgl. $\cdot$ ein inverses Element, so bezeichnen wir dieses mit a^{-1}.

Wie üblich vereinbaren wir, dass $\cdot$ stärker bindet als $+$ („Punkt- vor Strichrechnung"), also könnten wir z. B. die rechte Seite in (9.1a) auch in der Form $a \cdot b + a \cdot c$ schreiben. Außerdem können wir $-a \cdot b$ schreiben statt $-(a \cdot b)$.

Beispiel 9.2 (Ring, vgl. Beispiel 7.22 zu Gruppen).

(*i*) $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind kommutative Ringe mit Eins.

[33] Die Bezeichnung „abelscher Ring" ist nicht üblich, sie wird manchmal sogar für eine andere Eigenschaft verwendet als für die Kommutativität der Multiplikation.

(ii) Ein Ring $(R, +, \cdot)$ heißt ein **Nullring** (englisch: zero ring), wenn R nur aus einem Element besteht, also wenn $R = \{0_R\}$ gilt. Dadurch sind die Verknüpfungen eindeutig bestimmt: $0_R + 0_R = 0_R$ und $0_R \cdot 0_R = 0_R$. Da 0_R notwendigerweise auch das neutrale Element bzgl. $\cdot$ ist, ist ein Nullring ein Ring mit Eins, und es gilt $1_R = 0_R$.

Nullringe sind die einzigen Ringe, in dem das Nullelement und das Einselement identisch sind, siehe Lemma 9.3. Ein Nullring ist bis auf Isomorphie (Definition 9.17) eindeutig bestimmt, daher wird oft auch die Bezeichnung „**der Nullring**“ verwendet.

(iii) Für $m \in \mathbb{N}$ ist $(m\mathbb{Z}, +, \cdot)$ ein kommutativer Ring. Im Fall $m \neq 1$ besitzt er kein Einselement. Im Fall $m = 1$ ist $1 \in \mathbb{Z}$ das Einselement.

(iv) Für $m \in \mathbb{N}$ ist $(\mathbb{Z}_m, +_m, \cdot_m)$ ein kommutativer Ring mit Einselement $1 \in \mathbb{Z}_m$, denn nach Beispiel 7.22 ist $(\mathbb{Z}_m, +_m)$ eine abelsche Gruppe und $(\mathbb{Z}_m, \cdot_m)$ ein kommutatives Monoid. Er wird der **Ring von $\mathbb{Z}$ modulo m** (englisch: ring of $\mathbb{Z}$ modulo m) genannt. Im Fall $m = 1$ ist $(\mathbb{Z}_m, +_m, \cdot_m)$ ein Nullring.

(v) Es sei X eine Menge und $(R, +, \cdot)$ ein Ring. Dann ist $R^X = \{f \mid f\colon X \to R\}$, ausgestattet mit der punktweisen Addition und der punktweisen Multiplikation, ein Ring. Das Nullelement in $(R^X, +, \cdot)$ ist die **Nullfunktion** (englisch: zero function, constant function zero) $x \mapsto 0_R$. Besitzt R das Einselement 1_R, dann ist die **Einsfunktion** (englisch: constant function one) $x \mapsto 1_R$ das Einselement von $(R^X, +, \cdot)$.

Quizfrage 9.1: Wann ist $(R^X, +, \cdot)$ ein Nullring? Wann ist er kommutativ?

(vi) Es sei $(G, +)$ eine abelsche Gruppe. Wir definieren

$$\operatorname{Endo}(G) \coloneqq \{f\colon G \to G \mid f \text{ ist Endomorphismus}\} \tag{9.2}$$

und statten $\operatorname{Endo}(G)$ mit den Verknüpfungen

$$+\colon \begin{cases} \operatorname{Endo}(G) \times \operatorname{Endo}(G) \to \operatorname{Endo}(G) \\ (f, g) \mapsto f + g, \end{cases}$$

definiert durch die punktweise Addition $(f + g)(x) \coloneqq f(x) + g(x)$ für $x \in G$, und

$$\circ\colon \begin{cases} \operatorname{Endo}(G) \times \operatorname{Endo}(G) \to \operatorname{Endo}(G) \\ (f, g) \mapsto f \circ g, \end{cases}$$

definiert durch die Komposition $(f \circ g)(x) \coloneqq f(g(x))$, aus. Dann ist $(\operatorname{Endo}(G), +, \circ)$ ein Ring mit Einselement id_G, genannt der **Endomorphismenring** (englisch: ring of endomorphisms) der abelschen Gruppe $(G, +)$. Er ist i. A. nicht kommutativ.

Quizfrage 9.2: Warum definieren wir den Endomorphismenring nur auf abelschen Gruppen und nicht allgemeiner auf beliebigen Gruppen?

(vii) Ist X eine Menge, dann ist $(\mathcal{P}(X), \triangle, \cap)$ ein kommutativer Ring mit Einselement X, gennant der **Potenzmengenring** (englisch: power set ring) über der Menge X.

($viii$) Ist X eine nichtleere Menge, dann ist $(\mathcal{P}(X), \triangle, \cup)$ ist kein Ring, da das Distributivgesetz nicht gilt. △

Lemma 9.3 (Rechenregeln in Ringen).
Es sei $(R, +, \cdot)$ ein Ring mit dem Nullelement 0_R. Für $a, b \in R$ gilt:

(i) $0_R \cdot a = 0_R = a \cdot 0_R$.

(ii) $a \cdot (-b) = -a \cdot b = (-a) \cdot b$.

(iii) $(-a) \cdot (-b) = a \cdot b$.

(iv) Ist $(R, +, \cdot)$ ein Ring mit Einselement 1_R, aber kein Nullring, dann gilt $1_R \neq 0_R$.

Beachte: Hat $(R, +, \cdot)$ das Einselement 1_R, dann folgt aus Aussage (ii) insbesondere $-b = (-1_R) \cdot b$. („Das additive Inverse ergibt sich auch durch Multiplikation mit dem additiven Inversen des Einselements.")

Beweis. Aussage (i): Es gilt

$$\begin{aligned} 0_R + 0_R \cdot a &= 0_R \cdot a && \text{da } 0_R \text{ das neutrale Element von } (R, +) \text{ ist} \\ &= (0_R + 0_R) \cdot a && \text{da } 0_R \text{ das neutrale Element von } (R, +) \text{ ist} \\ &= 0_R \cdot a + 0_R \cdot a && \text{wegen des Distributivgesetzes (9.1b).} \end{aligned}$$

Die Anwendung der Kürzungsregel (7.8b) in der Gruppe $(R, +)$, also die Addition von $-(0_R \cdot a)$ zu beiden Seiten der Gleichung, zeigt $0_R = 0_R \cdot a$. Das zweite Resultat, $a \cdot 0_R = 0_R$, folgt analog.

Aussage (ii): Wir zeigen zunächst, dass $a \cdot (-b) = -a \cdot b$ gilt, also dass $a \cdot (-b)$ das Inverse zu $a \cdot b$ in der Gruppe $(R, +)$ ist. Da $(R, +)$ eine abelsche Gruppe ist (oder auch wegen Lemma 7.19) reicht dafür der Nachweis von $a \cdot (-b) + a \cdot b = 0_R$ aus, also der einseitige Test. In der Tat haben wir

$$\begin{aligned} a \cdot (-b) + a \cdot b &= a \cdot (-b + b) && \text{wegen des Distributivgesetzes (9.1a)} \\ &= a \cdot 0_R \\ &= 0_R && \text{nach Aussage (i).} \end{aligned}$$

Die Aussage $(-a) \cdot b = -a \cdot b$ folgt analog.

Aussage (iii): Wir haben

$$\begin{aligned} (-a) \cdot (-b) &= -(a \cdot (-b)) && \text{nach Aussage (ii)} \\ &= -(-a \cdot b) && \text{nach Aussage (ii)} \\ &= a \cdot b && \text{denn Invertierung ist involutorisch, siehe (7.5).} \end{aligned}$$

Aussage (iv): Es sei R ein Ring mit Einselement 1_R. Wir führen den Beweis durch Kontraposition. Wir nehmen also $1_R = 0_R$ an. Nun sei $a \in R$ beliebig. Dann gilt

$$\begin{aligned} a &= a \cdot 1_R && \text{da } 1_R \text{ das neutrale Element von } (R, \cdot) \text{ ist} \\ &= a \cdot 0_R && \text{da } 1_R = 0_R \text{ angenommen wurde} \\ &= 0_R && \text{nach Aussage (i).} \end{aligned}$$

Der Ring R besteht also nur aus dem Nullelement 0_R, d. h., R ist ein Nullring. □

Wir verwenden auch in Ringen $(R, +, \cdot)$ und insbesondere in der Gruppe $(R, +)$ die Schreibweisen aus Bemerkung 7.20. Es gilt also für $n \in \mathbb{N}$

$$n\,a \coloneqq a + \cdots + a \quad \text{und} \quad a^n \coloneqq a \cdot \cdots \cdot a.$$

Weiter ist $(-n)\,a \coloneqq -(n\,a) = n\,(-a)$ und $0\,a \coloneqq 0_R$. Besitzt $(R, +, \cdot)$ das Einselement 1_R, dann gilt nach dem Distributivgesetz (9.1b) weiter

$$n\,a = a + \cdots + a = 1_R \cdot a + \cdots + 1_R \cdot a = (1_R + \cdots + 1_R) \cdot a = (n\,1_R) \cdot a.$$

und analog $n\,a = a \cdot (n\,1_R)$. Dann definieren wir auch $a^0 \coloneqq 1_H$. Ist $a \in R$ zudem multiplikativ invertierbar, so ist $a^{-n} \coloneqq (a^n)^{-1} = (a^{-1})^n$.

Definition 9.4 (Charakteristik eines Ringes)**.**
Es sei $(R, +, \cdot)$ ein Ring mit Einselement 1_R.

(i) Wenn es eine Zahl $n \in \mathbb{N}$ gibt, sodass $n\,1_R = 0_R$ gilt, so nennen wir die kleinste solche Zahl

$$\min\{n \in \mathbb{N} \mid n\,1_R = 0_R\}$$

die **Charakteristik** (englisch: characteristic) von R, kurz $\operatorname{char}(R)$.

(ii) Gilt hingegen $n\,1_R \neq 0_R$ für alle $n \in \mathbb{N}$, so sagen wir, R habe die **Charakteristik** 0 und schreiben $\operatorname{char}(R) = 0$. △

Beispiel 9.5 (Charakteristik eines Ringes)**.**

(i) $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ haben Charakteristik 0.

(ii) Nullringe sind die einzigen Ringe mit Charakteristik 1, also die einzigen Ringe, in denen $1_R = 0_R$ gilt, vgl. Lemma 9.3.

(iii) Der Ring von $\mathbb{Z}$ modulo m $(\mathbb{Z}_m, +_m, \cdot_m)$ hat Charakteristik $m \in \mathbb{N}$.

(iv) Der Restklassenring modulo m $(\mathbb{Z} / m\,\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ aus dem folgenden Beispiel 9.6 hat ebenfalls Charakteristik $m \in \mathbb{N}$. △

Beispiel 9.6 (Restklassenring modulo m)**.**
Es sei $m \in \mathbb{N}$. Wir erinnern an die Faktorgruppe $\mathbb{Z} / m\,\mathbb{Z}$ aus Beispiel 8.23 mit den Elementen $[a] = a + m\,\mathbb{Z}$ (für $a \in \mathbb{Z}$), der kommutativen Verknüpfung $[a] \mathbin{\tilde{+}} [b] = [a + b]$ und dem neutralen Element $[0]$.

Weiter bildet $(\mathbb{Z} / m\,\mathbb{Z}, \tilde{\cdot})$ mit der Verknüpfung $[a] \mathbin{\tilde{\cdot}} [b] = [a \cdot b]$ ein kommutatives Monoid mit dem neutralen Element $[1]$, siehe auch Übung.

Schließlich können wir zeigen, dass die Distributivgesetze (9.1a) und (9.1b) gelten, denn:

$$\begin{aligned} [a] \mathbin{\tilde{\cdot}} ([b] \mathbin{\tilde{+}} [c]) &= [a] \mathbin{\tilde{\cdot}} [b + c] && \text{nach Definition von } \tilde{+} \\ &= [a \cdot (b + c)] && \text{nach Definition von } \tilde{\cdot} \\ &= [a \cdot b + a \cdot c] && \text{nach Distributivgesetz in } \mathbb{Z} \\ &= [a \cdot b] \mathbin{\tilde{+}} [a \cdot c] && \text{nach Definition von } \tilde{+} \end{aligned}$$

$$= [a] \mathbin{\tilde{\cdot}} [b] \mathbin{\tilde{+}} [a] \mathbin{\tilde{\cdot}} [c] \quad \text{nach Definition von } \tilde{\cdot}.$$

Das zweite Distributivgesetz (9.1b) ist wegen der Kommutativität der Halbgruppe $(\mathbb{Z}/m\mathbb{Z}, \tilde{\cdot})$ ebenfalls erfüllt. Daher bildet $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ einen kommutativen Ring mit Eins, genannt der **Restklassenring modulo m**. Im Fall $m = 1$ ist $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ein Nullring.

Die Verknüpfungstafeln für $\mathbb{Z}/m\mathbb{Z}$ für $m \in \{1, 2, 3, 4\}$ lauten:

$\tilde{+}$	[0]
[0]	[0]

$\tilde{\cdot}$	[0]
[0]	[0]

$\tilde{+}$	[0]	[1]
[0]	[0]	[1]
[1]	[1]	[0]

$\tilde{\cdot}$	[0]	[1]
[0]	[0]	[0]
[1]	[0]	[1]

$\tilde{+}$	[0]	[1]	[2]
[0]	[0]	[1]	[2]
[1]	[1]	[2]	[0]
[2]	[2]	[0]	[1]

$\tilde{\cdot}$	[0]	[1]	[2]
[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]
[2]	[0]	[2]	[1]

$\tilde{+}$	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

$\tilde{\cdot}$	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

△

Die im Fall $m = 4$ in $\mathbb{Z}/m\mathbb{Z}$ erstmalig auftretende Situation $[2] \mathbin{\tilde{\cdot}} [2] = [0]$ wollen wir benennen:

Definition 9.7 (Nullteiler, Nullteilerfreiheit, Integritätsring)**.**
Es sei $(R, +, \cdot)$ ein Ring.

(*i*) Das Element $a \in R$ heißt ein **Linksnullteiler** (englisch: left zero divisor) **von** R, wenn es ein $b \in R \setminus \{0_R\}$ gibt, sodass $a \cdot b = 0_R$ gilt.

(*ii*) Das Element b heißt ein **Rechtsnullteiler** (englisch: right zero divisor) **von** R, wenn es ein $a \in R \setminus \{0_R\}$ gibt, sodass $a \cdot b = 0_R$ gilt.

(*iii*) Ein Element, das Links- oder Rechtsnullteiler ist, heißt auch einfach ein **Nullteiler** (englisch: zero divisor). Ein Element, das gleichzeitig Links- und Rechtsnullteiler ist, heißt auch einfach ein **zweiseitiger Nullteiler** (englisch: two-sided zero divisor).

(*iv*) Der Ring $(R, +, \cdot)$ heißt **nullteilerfrei** (englisch: ring with no zero divisors), wenn es außer dem trivialen (zweiseitigen) Nullteiler 0_R keine weiteren Links- oder Rechtsnullteiler gibt, wenn also gilt:

$$\forall a, b \in R \ (a \cdot b = 0_R \implies a = 0_R \text{ oder } b = 0_R). \tag{9.3}$$

(„Ein Produkt ist nur dann Null, wenn mindestens einer der Faktoren gleich Null ist.“)

(*v*) Der Ring $(R, +, \cdot)$ heißt ein **Integritätsring** oder **Integritätsbereich** (englisch: integral domain), wenn gilt: $(R, +, \cdot)$ ist ein kommutativer, nullteilerfreier Ring mit Eins, der kein Nullring ist.[34] △

Lemma 9.8 (Charakterisierung von Nullteilern)**.**

Es sei $(R, +, \cdot)$ ein Ring.

(*i*) Für $a \in R$ sind äquivalent:

(a) a ist kein Linksnullteiler von R.

(b) Der Gruppenhomomorphismus $(R, +) \ni b \mapsto a \cdot b \in (R, +)$ ist injektiv.

(c) Für alle $b, c \in R$ gilt: $a \cdot b = a \cdot c$ impliziert $b = c$, d. h., a ist **linkskürzbar** (englisch: left-cancellative).

(*ii*) Für $b \in R$ sind äquivalent:

(a) b ist kein Rechtsnullteiler von R.

(b) Der Gruppenhomomorphismus $(R, +) \ni a \mapsto a \cdot b \in (R, +)$ ist injektiv.

(c) Für alle $a, c \in R$ gilt: $a \cdot b = c \cdot b$ impliziert $a = c$, d. h., b ist **rechtskürzbar** (englisch: right-cancellative).

Beweis. Der Beweis ist Gegenstand der Übung. □

Quizfrage 9.3: Ist ein Nullring nullteilerfrei?

Beispiel 9.9 (Integritätsringe und Gegenbeispiele)**.**

(*i*) $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind Integritätsringe.

(*ii*) Der Restklassenring modulo m $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ist ein Integritätsring genau dann, wenn $m \in \mathbb{N}$ eine Primzahl ist, siehe Satz 9.11.

(*iii*) Es sei X eine Menge, $(R, +, \cdot)$ ein Ring. Wir betrachten den Ring $(R^X, +, \cdot)$, siehe Beispiel 9.2. Dieser ist genau dann nullteilerfrei, wenn R ein Nullring ist oder wenn $X = \emptyset$ gilt oder wenn (X genau ein Element hat und R nullteilerfrei ist). (**Quizfrage 9.4:** Nachweis?)

(*iv*) $(\mathcal{P}(X), \triangle, \cap)$ ist genau dann nullteilerfrei, wenn X höchstens ein Element hat. Sobald X zwei verschiedene Elemente a und b hat, sind $A = \{a\}$ und $B = \{b\}$ Mengen ungleich dem Nullelement (der leeren Menge), deren „Multiplikation" $A \cap B = \emptyset$ das Nullelement ergibt. $(\mathcal{P}(X), \triangle, \cap)$ ist also genau dann ein Integritätsring, wenn X genau ein Element hat. Im Fall $X = \emptyset$ ist $(\mathcal{P}(X), \triangle, \cap)$ ein Nullring. △

Lemma 9.10 (notwendige Bedingung für die Nullteilerfreiheit)**.**

Für jeden nullteilerfreien Ring R mit Eins ist $\operatorname{char}(R)$ entweder gleich 0 oder gleich 1 oder eine Primzahl.

[34]Als Merkhilfe für die definierenden Eigenschaften eines Integritätsringes kann man sich an $(\mathbb{Z}, +, \cdot)$ orientieren.

Beweis. Wir nehmen an, $\operatorname{char}(R)$ wäre eine natürliche Zahl $\geqslant 2$, die keine Primzahl ist. Wir hätten dann also $\operatorname{char}(R) = n_1\, n_2$ mit $n_1, n_2 \in \mathbb{N}_{\geqslant 2}$, und aufgrund der Distributivgesetze

$$0_R = (n_1\, n_2)\, 1_R = (n_1\, 1_R) \cdot (n_2\, 1_R).$$

Aus der Nullteilerfreiheit würde $n_1\, 1_R = 0_R$ oder $n_2\, 1_R = 0_R$ folgen. n_1 und n_2 sind aber beide echt kleiner als $n_1\, n_2$. Das steht im Widerspruch dazu, dass $\operatorname{char}(R) = n_1\, n_2$ nach Definition 9.4 die kleinste natürliche Zahl n ist, für die $n\, 1_R = 0_R$ gilt. □

Satz 9.11 (Nullteilerfreiheit des Restklassenringes)**.**
Der Restklassenring modulo m $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ist ein Integritätsring genau dann, wenn $m \in \mathbb{N}$ eine Primzahl ist.

Beweis. Für $m = 1$ ist $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ein Nullring und damit kein Integritätsring. Wir betrachten also im Weiteren nur den Fall $m \geqslant 2$. Das heißt, $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ist ein kommutativer Ring ungleich einem Nullring mit dem Einselement $[1]$ (Beispiel 9.6). Die Frage, ob dieser Ring ein Integritätsring ist, hängt also genau an der Nullteilerfreiheit. Das Nullelement von $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ist $[0]$.

Es sei zunächst $m \in \mathbb{N}$, $m \geqslant 2$, eine Primzahl. Wir nehmen an, $[a]$ und $[b]$ seien Elemente aus $\mathbb{Z}/m\mathbb{Z}$ mit $[0] = [a] \,\tilde{\cdot}\, [b] = [a \cdot b]$. Das heißt aber, da 0 und $a \cdot b$ in derselben Restklasse modulo m liegen, dass $a \cdot b = m\, z$ gilt für irgendein $z \in \mathbb{Z}$. Da m eine Primzahl ist, kommt m in der (vorzeichenbehafteten) Primfaktorzerlegung von $a \cdot b$ vor. Das heißt, dass a oder b den Primfaktor m enthält, also gilt $m \mid a$ oder $m \mid b$, woraus $[a] = [0]$ oder $[b] = [0]$ folgt. Damit ist $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ nullteilerfrei.

Es sei nun umgekehrt $m \in \mathbb{N}$, $m \geqslant 4$, *keine* Primzahl; sie lässt sich also schreiben als $m = a \cdot b$ mit Zahlen $a, b \in [\![2, m-1]\!]$. Die zugehörigen Restklassen $[a]$ und $[b]$ sind ungleich $[0]$ (**Quizfrage 9.5:** Warum?) Es gilt

$$\begin{aligned} [0] &= [m] && \text{da } 0 \overset{m}{\equiv} m \\ &= [a \cdot b] && \text{da } m = a \cdot b \\ &= [a] \,\tilde{\cdot}\, [b] && \text{nach Definition von } \tilde{\cdot}. \end{aligned}$$

Damit ist $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ *nicht* nullteilerfrei. □

Definition 9.12 (Unterring, vgl. Definition 7.42 einer Untergruppe)**.**
Es sei $(R, +, \cdot)$ ein Ring.

(*i*) Eine Teilmenge $U \subseteq R$ heißt ein **Unterring** (englisch: subring) von $(R, +, \cdot)$, wenn U bzgl. $+$ und bzgl. $\cdot$ abgeschlossen und wenn $(U, +, \cdot)$ selbst wieder ein Ring ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, +)$ eine Untergruppe von $(R, +)$ und wenn U eine Unterhalbgruppe von $(R, \cdot)$ ist.

(*ii*) Ist $(R, +, \cdot)$ ein Ring mit Einselement 1_R, dann heißt ein Unterring U, der auch das Einselement 1_R enthält, ein **Unterring mit Eins** (englisch: subring with unity).[35]

Beachte: Es reicht nicht aus, zu fordern, dass $(U, \cdot)$ *irgendein* neutrales Element besitzt.

[35] Dadurch ist der Unterring $(U, +, \cdot)$ dann natürlich selbst wieder ein Ring mit demselben Einselement 1_R.

(*iii*) Ein Unterring U von $(R, +, \cdot)$ heißt **echt** (englisch: proper subring), wenn $U \subsetneq R$ gilt. △

Die Prüfung einer Teilmenge $U \subseteq R$ auf die Unterring-Eigenschaft lässt sich mit folgendem Kriterium erreichen:

Satz 9.13 (Unterringkriterium, vgl. Satz 7.44 zum Untergruppenkriterium).
Es sei $(R, +, \cdot)$ ein Ring und $U \subseteq R$. Dann sind äquivalent:

(*i*) $(U, +, \cdot)$ ist ein Unterring von $(R, +, \cdot)$.

(*ii*) $U \neq \emptyset$, und für alle $a, b \in U$ gilt $a - b \in U$ und $a \cdot b \in U$.[36]

Beweis. Aussage (*i*) $\Rightarrow$ Aussage (*ii*): Es sei $(U, +, \cdot)$ ein Unterring von $(R, +, \cdot)$. Dann enthält U notwendigerweise das Nullelement 0_R von $(R, +, \cdot)$, also das neutrale Element der Untergruppe $(U, +)$ von $(R, +)$. Für $a, b \in U$ gilt $-b \in U$ nach Lemma 7.43. Da U bzgl. $+$ abgeschlossen ist, folgt $a - b = a + (-b) \in U$, und da U bzgl. $\cdot$ abgeschlossen ist, folgt $a \cdot b \in U$.

Aussage (*ii*) $\Rightarrow$ Aussage (*i*): Nach Voraussetzung erfüllt $(U, +)$ das Untergruppenkriterium (Satz 7.44), also ist $(U, +)$ eine Untergruppe von $(R, +)$. Außerdem ist U nach Voraussetzung abgeschlossen bzgl. $\cdot$, d. h., $(U, \cdot)$ ist eine Unterhalbgruppe von $(R, \cdot)$. Damit ist $(U, +, \cdot)$ ein Unterring von $(R, +, \cdot)$. □

Beispiel 9.14 (Unterring).

(*i*) $(\mathbb{Z}, +, \cdot)$ ist ein Unterring mit Eins von $(\mathbb{Q}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$ ist ein Unterring mit Eins von $(\mathbb{R}, +, \cdot)$, und $(\mathbb{R}, +, \cdot)$ ist ein Unterring mit Eins von $(\mathbb{C}, +, \cdot)$.

(*ii*) Der Nullring $\{0_R\}$ und R selbst sind Unterringe in jedem Ring $(R, +, \cdot)$.

(*iii*) Für $m \in \mathbb{N}$ ist $(m\,\mathbb{Z}, +, \cdot)$ ein Unterring von $(\mathbb{Z}, +, \cdot)$. Im Fall $m \neq 1$ handelt es sich nicht um einen Unterring mit Eins.

(*iv*) Für $k, m \in \mathbb{N}$ ist

$$\{n \,|\, n \in k\,\mathbb{Z} \cap \mathbb{Z}_m\} = \{0, k, 2\,k, \ldots, \} \cap \{0, 1, \ldots, m-1\}$$

mit den Operationen $+_m$ und $\cdot_m$ genau dann ein Unterring von $(\mathbb{Z}_m, +_m, \cdot_m)$, wenn $k \mid m$ gilt. Beispielsweise ist $(2\,\mathbb{Z} \cap \mathbb{Z}_4, +_4, \cdot_4)$ ein Unterring von $(\mathbb{Z}_4, +_4, \cdot_4)$, der nur aus den geraden Zahlen, also $\{0, 2\}$ besteht. Die Verknüpfungstafeln hatten wir in Beispiel 9.6 bereits angegeben (für den zu $\mathbb{Z}_4$ isomorphen Restklassenring $\mathbb{Z} / 4\,\mathbb{Z}$, siehe Beispiel 9.32):

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot_4$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Zu beachten sind in $2\,\mathbb{Z} \cap \mathbb{Z}_4$ nur diejenigen Zeilen und Spalten, die zu geraden Zahlen gehören, diese sind farbig hinterlegt.

[36]kurz: $U - U \subseteq U$ und $U \cdot U \subseteq U$

(*v*) Ist X eine Menge und $Y \subseteq X$, dann ist $(\mathcal{P}(Y), \triangle, \cap)$ ein Unterring von $(\mathcal{P}(X), \triangle, \cap)$. Im Fall $Y \subsetneq X$ handelt es sich nicht um einen Unterring mit Eins. $(\mathcal{P}(Y), \triangle, \cap)$ hat zwar das Einselement Y, dieses ist aber vom Einselement X in $(\mathcal{P}(X), \triangle, \cap)$ verschieden.

(*vi*) Das **Zentrum**[37]

$$Z \coloneqq \{z \in R \mid a \cdot z = z \cdot a \text{ für alle } a \in R\} \tag{9.4}$$

eines Ringes $(R, +, \cdot)$ ist ein kommutativer Unterring. Wenn R ein Ring mit Eins ist, dann ist Z ein Unterring mit Eins. △

Bemerkung 9.15 („Unterring sein" ist eine Ordnungsrelation, vgl. Bemerkung 7.46 zu Untergruppen)**.**

(*i*) Die Relation „ist Unterring von" ist eine partielle Ordnung auf der Klasse aller Ringe.

(*ii*) Insbesondere ist die Menge aller Unterringe eines bestimmten Ringes $(R, +, \cdot)$ durch die Unterringhalbordnung partiell geordnet. Diese Ordnung stimmt mit der Inklusionshalbordnung überein. △

Lemma 9.16 (Durchschnitt von Unterringen, vgl. Lemma 7.47 zu Untergruppen)**.**
Es sei $(R, +, \cdot)$ ein Ring.

(*i*) Ist $(U_i)_{i \in I}$ eine nichtleere Familie von Unterringen von $(R, +, \cdot)$, dann ist auch $\bigcap_{i \in I} U_i$ ein Unterring von R.

(*ii*) Ist $\mathcal{U}$ eine nichtleere Menge von Unterringen von $(R, +, \cdot)$, dann ist auch $\bigcap \mathcal{U}$ ein Unterring von $(R, +, \cdot)$.

Beweis. Der Beweis ist eine einfache Anwendung des Unterringkriteriums (Satz 9.13). □

Definition 9.17 (Ringhomomorphismus, vgl. Definitionen 8.1, 8.4 und 8.6 für Homomorphismen von Halbgruppen bzw. Monoiden bzw. Gruppen)**.**
Es seien $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ zwei Ringe.

(*i*) Eine Abbildung $f \colon R_1 \to R_2$ heißt **strukturverträglich** oder ein **Homomorphismus von Ringen** (englisch: ring homomorphism), wenn gilt:

$$f(a +_1 b) = f(a) +_2 f(b) \quad \text{für alle } a, b \in R_1, \tag{9.5a}$$

$$f(a \cdot_1 b) = f(a) \cdot_2 f(b) \quad \text{für alle } a, b \in R_1. \tag{9.5b}$$

Besitzen beide Ringe ein Einselement 1_{R_1} bzw. 1_{R_2} und gilt zusätzlich

$$f(1_{R_1}) = 1_{R_2}, \tag{9.5c}$$

dann nennen wir f genauer einen **Homomorphismus von Ringen mit Eins** (englisch: homomorphism of rings with unity).

[37]Das Zentrum eines Ringes besteht also aus denjenigen Elementen, die multiplikativ mit allen Gruppenelementen kommutieren, vgl. Definition (8.10) für das Zentrum einer Gruppe.

(*ii*) Ist $f: R_1 \to R_2$ strukturverträglich und gilt $(R_1, +_1, \cdot_1) = (R_2, +_2, \cdot_2)$, so sprechen wir auch von einem **Endomorphismus eines Ringes** (englisch: ring endomorphism) bzw. von einem **Endomorphismus eines Ringes mit Eins** (englisch: endomorphism of a ring with unity).

(*iii*) Ist $f: R_1 \to R_2$ strukturverträglich und bijektiv, so heißt f auch **strukturerhaltend** oder ein **Isomorphismus von Ringen** bzw. ein **Isomorphismus von Ringen mit Eins** (englisch: isomorphism of a ring with unity). In diesem Fall nennen wir $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ auch zueinander **isomorphe Ringe** (englisch: isomorphic rings) bzw. zueinander **isomorphe Ringe mit Eins** (englisch: isomorphic rings with unity) und schreiben

$$(R_1, +_1, \cdot_1) \cong (R_2, +_2, \cdot_2).$$

(*iv*) Ist $f: R_1 \to R_2$ strukturverträglich und bijektiv und gilt $(R_1, +_1, \cdot_1) = (R_2, +_2, \cdot_2)$, so sprechen wir auch von einem **Automorphismus** (englisch: ring automorphism) **eines Ringes** bzw. von einem **Automorphismus eines Ringes mit Eins** (englisch: automorphism of a ring with unity). △

Beachte: Die Beziehung (9.5a) besagt, dass $f: (R_1, +_1) \to (R_2, +_2)$ ein Gruppenhomomorphismus ist. Aus Lemma 8.8 folgt damit für die Nullelemente 0_{R_1} bzw. 0_{R_2} notwendigerweise

$$f(0_{R_1}) = 0_{R_2}. \tag{9.6}$$

Weiter bedeutet (9.5b), dass $f: (R_1, \cdot_1) \to (R_2, \cdot_2)$ ein Halbgruppenhomomorphismus ist. (9.5b) und (9.5c) zusammen bedeuten, dass $f: (R_1, \cdot_1) \to (R_2, \cdot_2)$ ein Monoidhomomorphismus ist.

Anstelle der Bedingung (9.5c) reicht es auch aus, zu fordern, dass $f(1_{R_1})$ ein multiplikativ invertierbares Element in R_2 ist, vgl. (8.3) in Bemerkung 8.5.

Analog zu Satz 8.2 und Folgerung 8.3 gilt:

Satz 9.18 (Komposition von Ringhomomorphismen, Inverse von Ringisomorphismen, vgl. Satz 8.2 zu Halbgruppenhomomorphismen)**.**
Es seien $(R_1, +_1, \cdot_1)$, $(R_2, +_2, \cdot_2)$ und $(R_3, +_3, \cdot_3)$ drei Ringe.

(*i*) Sind $f: R_1 \to R_2$ und $g: R_2 \to R_3$ Ringhomomorphismen, dann ist auch $g \circ f: R_1 \to R_3$ ein Ringhomomorphismus.

(*ii*) Ist $f: R_1 \to R_2$ ein Ringisomorphismus, dann ist auch $f^{-1}: R_2 \to R_1$ ein Ringisomorphismus.

Analoge Aussagen gelten für Homomorphismen von Ringen mit Eins.

Folgerung 9.19 (Isomorphie von Ringen ist eine Äquivalenzrelation, vgl. Folgerung 8.3 zur Isomorphie von Halbgruppen)**.**
Isomorphie ist eine Äquivalenzrelation auf der Klasse aller Ringe bzw. auf der Klasse aller Ringe mit Eins.

Lemma 9.20 (Ringe mit Eins und Charakteristik 0 enthalten $\mathbb{Z}$)**.**
Es sei $(R, +, \cdot)$ ein Ring mit Eins und $\operatorname{char}(R) = 0$. Dann enthält R einen Unterring, der isomorph zu $\mathbb{Z}$ ist.

Beweis. □

Definition 9.21 (Bild und Kern eines Ringhomomorphismus, vgl. Definition 8.10 von Bild und Kern eines Gruppenhomomorphismus)**.**
Es seien $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ Ringe mit den Nullelementen 0_{R_1} bzw. 0_{R_2}. Weiter sei $f\colon R_1 \to R_2$ ein Homomorphismus.

(i) Das **Bild** von f ist definiert als

$$\operatorname{Bild}(f) \coloneqq \{f(a_1) \in R_2 \mid a_1 \in R_1\} = f(R_1). \tag{9.7}$$

(ii) Der **Kern** von f ist definiert als

$$\operatorname{Kern}(f) \coloneqq \{a_1 \in R_1 \mid f(a_1) = 0_{R_2}\} = f^{-1}(\{0_{R_2}\}). \tag{9.8}$$

△

Lemma 9.22 (Bild und Kern sind Unterringe, vgl. Lemma 8.11 zur Untergruppeneigenschaft von Bild und Kern eines Gruppenhomomorphismus)**.**
Es seien $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ Ringe. Weiter sei $f\colon R_1 \to R_2$ ein Homomorphismus. Dann gilt:

(i) $\operatorname{Bild}(f)$ ist ein Unterring von $(R_2, +_2, \cdot_2)$.

(ii) $\operatorname{Kern}(f)$ ist ein Unterring von $(R_1, +_1, \cdot_1)$.

Beweis. Wir bezeichnen die Nullelemente von R_1 bzw. R_2 mit 0_{R_1} bzw. 0_{R_2}.

Aussage (i): Wir überprüfen das Unterringkriterium (Satz 9.13). Es gilt $f(0_{R_1}) = 0_{R_2}$ nach (9.6), also folgt $0_{R_2} \in \operatorname{Bild}(f)$ und $\operatorname{Bild}(f) \neq \emptyset$. Weiter seien a_2, b_2 irgendwelche Elemente in $\operatorname{Bild}(f)$. Wir müssen zeigen: $a_2 -_2 b_2 \in \operatorname{Bild}(f)$ sowie $a_2 \cdot_2 b_2 \in \operatorname{Bild}(f)$.

Nach Definition von $\operatorname{Bild}(f)$ gibt es $a_1, b_1 \in G_1$ mit $f(a_1) = a_2$ und $f(b_1) = b_2$. Daher ist

$$\begin{aligned} a_2 -_2 b_2 &= f(a_1) -_2 (f(b_1)) && \text{nach Voraussetzung} \\ &= f(a_1 -_1 b_1) && \text{nach Lemma 8.8} \end{aligned}$$

und damit $a_2 -_2 b_2 \in \operatorname{Bild}(f)$. Weiterhin gilt

$$\begin{aligned} a_2 \cdot_2 b_2 &= f(a_1) \cdot_2 (f(b_1)) && \text{nach Voraussetzung} \\ &= f(a_1 \cdot_1 b_1) && \text{da } f\colon (R_1, \cdot_1) \to (R_2, \cdot_2) \text{ Halbgruppenhomomorphismus ist} \end{aligned}$$

und damit $a_2 \cdot_2 b_2 \in \operatorname{Bild}(f)$.

Aussage (ii): Wir überprüfen wiederum das Unterringkriterium. Es gilt $f(0_{R_1}) = 0_{R_2}$ nach (9.6), also $0_{R_1} \in \operatorname{Kern}(f)$ und $\operatorname{Kern}(f) \neq \emptyset$. Weiter seien a_1, b_1 irgendwelche Elemente in $\operatorname{Kern}(f)$. Wir müssen zeigen: $a_1 -_1 b_1 \in \operatorname{Kern}(f)$ sowie $a_1 \cdot_1 b_1 \in \operatorname{Kern}(f)$. Es gilt

$$\begin{aligned} f(a_1 -_1 b_1) &= f(a_1) -_2 f(b_1) && \text{nach Lemma 8.8} \\ &= 0_{R_2} -_2 0_{R_2} && \text{da } a_1, b_1 \in \operatorname{Kern}(f) \text{ liegen} \\ &= 0_{R_2} && \text{da } -0_{R_2} = 0_{R_2} \text{ ist} \end{aligned}$$

und damit $a_1 -_1 b_1 \in \operatorname{Kern}(f)$. Weiterhin gilt

$$\begin{aligned} f(a_1 \cdot_1 b_1) &= f(a_1) \cdot_2 f(b_1) && \text{da } f\colon (R_1, \cdot_1) \longrightarrow (R_2, \cdot_2) \text{ Halbgruppenhomomorphismus ist} \\ &= 0_{R_2} \cdot_2 0_{R_2} && \text{da } a_1, b_1 \in \operatorname{Kern}(f) \text{ liegen} \\ &= 0_{R_2} && \text{nach Lemma 9.3} \end{aligned}$$

und damit $a_1 \cdot_1 b_1 \in \operatorname{Kern}(f)$. □

Beispiel 9.23 (Ringhomomorphismen).

(i) Die Abbildung

$$f\colon (\mathbb{Z}, +, \cdot) \ni a \mapsto [a] = a + m\mathbb{Z} \in (\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$$

ist ein surjektiver Ringhomomorphismus zwischen zwei kommutativen Ringen mit Eins, denn: f ist als kanonische Surjektion der Faktorgruppe nach Satz 8.21 und Beispiel 8.23 ein surjektiver Gruppenhomomorphismus $f\colon (\mathbb{Z}, +) \longrightarrow (\mathbb{Z}/m\mathbb{Z}, \tilde{+})$, und außerdem ist $f\colon (\mathbb{Z}, \cdot) \longrightarrow (\mathbb{Z}/m\mathbb{Z}, \tilde{\cdot})$ ein Monoidhomomorphismus, siehe Beispiel 9.6 und Übung.

Es gilt

$$\begin{aligned} \operatorname{Bild}(f) &= \mathbb{Z}/m\mathbb{Z}, \\ \operatorname{Kern}(f) &= f^{-1}([0]) = m\mathbb{Z}. \end{aligned}$$

(ii) Für $m \in \mathbb{N}$ ist die Abbildung

$$f\colon (\mathbb{Z}_m, +_m, \cdot_m) \ni a \mapsto [a] = a + m\mathbb{Z} \in (\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$$

ein Ringisomorphismus zwischen dem Ring von $\mathbb{Z}$ modulo m (Beispiel 9.2) und dem Restklassenring modulo m (Beispiel 9.6), beides kommutative Ringe mit Eins, denn: f ist nach Beispiel 8.23 ein Gruppenisomorphismus $f\colon (\mathbb{Z}_m, +_m) \longrightarrow (\mathbb{Z}/m\mathbb{Z}, \tilde{+})$, und außerdem ist $f\colon (\mathbb{Z}_m, \cdot_m) \longrightarrow (\mathbb{Z}/m\mathbb{Z}, \tilde{\cdot})$ nach Übung ein Monoidisomorphismus.

Es gilt

$$\begin{aligned} \operatorname{Bild}(f) &= \mathbb{Z}/m\mathbb{Z}, \\ \operatorname{Kern}(f) &= f^{-1}([0]) = \{0\}. \end{aligned}$$

△

Lemma 9.24 (Charakterisierung der Injektivität von Ringhomomorphismen, vgl. Lemma 8.13 für Gruppenhomomorphismen).
Es seien $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ Ringe mit den Nullelementen 0_{R_1} bzw. 0_{R_2}. Weiter sei $f\colon R_1 \longrightarrow R_2$ ein Homomorphismus. Dann sind äquivalent:

(i) f ist injektiv.

(ii) $\operatorname{Kern}(f) = \{0_{R_1}\}$.

(iii) Die einzige Lösung der Gleichung $f(a) = 0_{R_2}$ ist $a = 0_{R_1}$.

Beweis. □

Ende der Vorlesung 13

§ 9.1 Ideale und Faktorringe

Ideale in Ringen sind spezielle Unterringe, die dieselbe Funktion einnehmen wie Normalteiler in Gruppen (§ 8.1). Mit ihrer Hilfe können wir Faktorringe definieren (also „gröbere Versionen“ gegebener Ringe) und einen Homomorphiesatz für Ringe erhalten.

Definition 9.25 (Ideal, vgl. Definition 8.15 von Normalteilern)**.**
Es sei $(R, +, \cdot)$ ein Ring.

(i) Eine Teilmenge $J \subseteq R$ heißt ein **Ideal** (englisch: ideal) von $(R, +, \cdot)$, wenn J ein Unterring von R ist und zusätzlich
$$R \cdot J \subseteq J \quad \text{und} \quad J \cdot R \subseteq J \tag{9.9}$$
gilt.[38]

Beachte: Das ist genau dann erfüllt, wenn $(J, +)$ eine Untergruppe von $(R, +)$ ist und wenn (9.9) gilt, denn (9.9) impliziert ja bereits, dass J bzgl. $\cdot$ abgeschlossen ist.

Manchmal notiert man die Eigenschaft, dass $(J, +, \cdot)$ ein Ideal des Ringes $(R, +, \cdot)$ ist, als $(J, +, \cdot) \trianglelefteq (R, +, \cdot)$.

(ii) Ein Ideal $(J, +, \cdot)$ von $(R, +, \cdot)$ heißt **echt** (englisch: proper ideal), wenn $J \subsetneq R$ gilt. △

Bemerkung 9.26 (zum Begriff des Ideals)**.**

(i) In einer Gruppe $(G, \star)$ war die definierende Eigenschaft einer normalen Untergruppe ($a \star N = N \star a$ für alle $a \in G$) genau die Eigenschaft, die wir benötigt haben, um die Gruppenoperation $\star$ in natürlicher Weise auf die Faktormenge G / N zu vererben (Satz 8.21).

Dieselbe Eigenschaft wird man analog auch in Ringen benötigen. Da $(R, +)$ kommutativ ist, muss sie allerdings nicht explizit gefordert werden, da jede Untergruppe von $(R, +)$ bereits ein Normalteiler ist (Beispiel 8.17). Die Bedingung (9.9) bezieht sich daher ausschließlich auf die zweite Verknüpfung $\cdot$.

(ii) Es gibt den Begriff des **Ideals** auch bereits in Halbgruppen $(H, \square)$. Damit ist eine Unterhalbgruppe I gemeint, für die $H \square I \subseteq I$ und $I \square H \subseteq I$ gefordert wird. In diesem Sinne ist ein Ideal in einem Ring $(R, +, \cdot)$ also ein Normalteiler der abelschen Gruppe $(R, +)$ und ein Ideal der Halbgruppe $(R, \cdot)$. △

[38] Ausgeschrieben heißt das also: $a \in R$ und $j \in J$ impliziert $a\,j \in J$ und $j\,a \in J$.

Lemma 9.27 (Kerne von Ringhomomorphismen sind Ideale, vgl. Lemma 8.18 für Kerne von Gruppenhomomorphismen).
Es seien $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ Ringe und $f\colon R_1 \to R_2$ ein Homomorphismus. Dann gilt:

(i) Die Elemente von R_1, die denselben Funktionswert wie $a \in R_1$ haben, sind genau die Elemente der additiven Nebenklasse von $\operatorname{Kern}(f)$ zu a:

$$f^{-1}(\{f(a)\}) = a +_1 \operatorname{Kern}(f) = \operatorname{Kern}(f) +_1 a.$$

(ii) $\operatorname{Kern}(f)$ ist ein Ideal von R_1.

Beweis. Die Aussage (i) folgt sofort aus der Aussage (i) in Lemma 8.18, da f ja insbesondere ein Gruppenhomomorphismus $f\colon (R_1, +_1) \to (R_2, +_2)$ ist. Außerdem folgt aus Lemma 8.18, dass $\operatorname{Kern}(f) = \{a \in R_1 \mid f(a) = 0_{R_2}\}$ ein Normalteiler von $(R_1, +_1)$ ist, also insbesondere eine Untergruppe.

Zu zeigen bleibt nach Definition 9.25 eines Ideals $R_1 \cdot_1 \operatorname{Kern}(f) \subseteq \operatorname{Kern}(f)$ und $\operatorname{Kern}(f) \cdot_1 R_1 \subseteq \operatorname{Kern}(f)$. Dazu sei nun $a \in R_1$ und $j \in \operatorname{Kern}(f)$, dann gilt

$$f(a \cdot_1 j) = f(a) \cdot_2 f(j) = f(a) \cdot_2 0_{R_2} = 0_{R_2}$$

nach Lemma 9.3 und analog

$$f(j \cdot_1 a) = f(j) \cdot_2 f(a) = 0_{R_2} \cdot_2 f(a) = 0_{R_2}.$$

Also gehören $a \cdot_1 j$ und $j \cdot_1 a$ wieder zu J, kurz: $R_1 \cdot_1 J \subseteq J$ und $J \cdot_1 R_1 \subseteq J$. □

Bemerkung 9.28 (Urbilder von Idealen sind Ideale, vgl. Bemerkung 8.19 zu Urbildern von Normalteilern).
Es gilt sogar folgende Verallgemeinerung von Lemma 9.27: Urbilder von Idealen unter Ringhomomorphismen sind Ideale. △

Beispiel 9.29 (Ideal).

(i) In jedem Ring $(R, +, \cdot)$ sind $\{0\}$ (das **Nullideal**, englisch: zero ideal) und R (das **Einsideal**, englisch: unit ideal) Ideale. Diese heißen die **trivialen Ideale** (englisch: trivial ideals).

(ii) Die Mengen der Form $m\mathbb{Z}$ mit $m \in \mathbb{N}_0$ sind genau die Ideale des Ringes $(\mathbb{Z}, +, \cdot)$. Für $m = 0$ ergibt sich das Nullideal, für $m = 1$ das Einsideal.

(iii) Ist X eine Menge und $Y \subseteq X$, dann ist $(\mathcal{P}(Y), \triangle, \cap)$ ein Ideal von $(\mathcal{P}(X), \triangle, \cap)$. △

In einer Gruppe $(G, \star)$ konnten wir einen Normalteiler N ausfaktorisieren und dabei die Gruppenoperation $\star$ in natürlicher Weise auf die Faktormenge $G\,/\,N$ vererben. Dieselbe Konstruktion werden wir jetzt in Ringen durchführen. Aus der Faktormenge $R\,/\,J$ (bestehend aus den additiven Nebenklassen von J in $(R, +, \cdot)$) wird damit der **Faktorring** (englisch: factor ring) oder **Quotientenring** (englisch: quotient ring) **von R nach J**. Man sagt auch: „Aus dem Ring $(R, +, \cdot)$ wird das Ideal J ausfaktorisiert."

Satz 9.30 (Faktorring, vgl. Satz 8.21 über Faktorgruppen).

(*i*) Es sei $(R, +, \cdot)$ ein Ring und J eines seiner Ideale. Dann gilt:

(a) Auf der Faktormenge
$$R\,/\,J = \big\{[a] = a + J \,\big|\, a \in R\big\}$$
sind $\tilde{+}$ und $\tilde{\cdot}$, definiert als
$$[a] \,\tilde{+}\, [b] \coloneqq [a + b] \quad \text{für } a, b \in R, \tag{9.10a}$$
$$[a] \,\tilde{\cdot}\, [b] \coloneqq [a \cdot b] \quad \text{für } a, b \in R, \tag{9.10b}$$
assoziative Verknüpfungen, bzgl. der $(R\,/\,J, \tilde{+}, \tilde{\cdot})$ einen Ring bildet. Das Nullelement ist $[0_R] = J$, und für die additiven Inversen gilt $\tilde{-}[a] = [-a]$.

(b) Die **kanonische Surjektion von R auf $R\,/\,J$**
$$\pi\colon \begin{cases} R \to R\,/\,J \\ a \mapsto [a], \end{cases} \tag{9.11}$$
die jedem Element $a \in R$ seine additive Nebenklasse $[a]$ zuordnet, ist ein surjektiver Ringhomomorphismus. Es gilt $\operatorname{Kern}(\pi) = J$.

(c) Besitzt R das Einselement 1_R, dann besitzt $R\,/\,J$ das Einselement $[1_R]$. Ist dann $a \in R$ bzgl. $\cdot$ invertierbar, so ist auch $[a] \in R\,/\,J$ bzgl. $\tilde{\cdot}$ invertierbar, und es gilt $[a]^{-1} = [a^{-1}]$.

(d) Wenn $(R, +, \cdot)$ kommutativ ist, dann auch $(R\,/\,J, \tilde{+}, \tilde{\cdot})$.

(*ii*) Es sei $(R, +, \cdot)$ ein Ring und U irgendein Unterring. Ist die Verknüpfung (9.10b) auf der Menge der Nebenklassen[39] $R\,/\,U$ wohldefiniert, dann ist U notwendigerweise ein Ideal von R.

Beweis. Aussage (*i*): Wir zeigen zunächst Aussage (a). Da $(J, +)$ ein Normalteiler der abelschen Gruppe $(R, +)$ ist, folgt aus Satz 8.21 sofort, dass $(R\,/\,J, \tilde{+})$ ebenfalls eine abelsche Gruppe ist. Außerdem folgt $[0_R] = J$ und $\tilde{-}[a] = [-a]$. Es bleibt zu zeigen, dass die Multiplikation (9.10b) wohldefiniert und assoziativ ist und dass die Distributivgesetze gelten. Das erfolgt in mehreren Schritten:

Schritt 1: Wir müssen zunächst zeigen, dass $\tilde{\cdot}$ wohldefiniert ist. Dazu seien $a_1, a_2, b_1, b_2 \in R$ gegeben, wobei $[a_1] = [a_2]$ und $[b_1] = [b_2]$ angenommen wird, d. h., $a_1 + J = a_2 + J$ und $b_1 + J = b_2 + J$ oder äquivalent $a_1 - a_2 \in J$ sowie $b_1 - b_2 \in J$. Dann gilt
$$\begin{aligned} &a_1 \cdot b_1 - a_2 \cdot b_2 \\ &\quad = (a_1 - a_2) \cdot b_1 + a_2 \cdot (b_1 - b_2) && \text{nach den Distributivgesetzen (9.1)} \\ &\quad \in J \cdot b_1 + a_2 \cdot J && \text{wegen } a_1 - a_2 \in J \text{ und } b_1 - b_2 \in J \\ &\quad \subseteq J + J && \text{wegen der Idealeigenschaft von } J \\ &\quad = J && \text{da } (J, +) \text{ eine Untergruppe von } (R, +) \text{ ist.} \end{aligned}$$
Das bedeutet aber gerade $[a_1 \cdot b_1] = [a_2 \cdot b_2]$ und damit $[a_1] \,\tilde{\cdot}\, [b_1] = [a_2] \,\tilde{\cdot}\, [b_2]$.

[39] Im Unterschied zu Satz 8.21 müssen hier Links- und Rechtsnebenklassen nicht unterschieden werden, da $(R, +)$ ja kommutativ ist.

Schritt 2: Die Assoziativität von $\tilde{\cdot}$ folgt aus

$$\begin{aligned} &\big([a] \mathbin{\tilde{\cdot}} [b]\big) \mathbin{\tilde{\cdot}} [c] = [a \cdot b] \mathbin{\tilde{\cdot}} [c] = [(a \cdot b) \cdot c] \\ \text{und}\quad &[a] \mathbin{\tilde{\cdot}} \big([b] \mathbin{\tilde{\cdot}} [c]\big) = [a] \mathbin{\tilde{\cdot}} [b \cdot c] = [a \cdot (b \cdot c)]. \end{aligned}$$

Aufgrund der Assoziativität der Multiplikation im Ring R sind die Äquivalenzklassen gleich.

Schritt 3: Für die Distributivgesetze (9.1) argumentieren wir ähnlich:

$$\begin{aligned} &[a] \mathbin{\tilde{\cdot}} \big([b] \mathbin{\tilde{+}} [c]\big) = [a] \mathbin{\tilde{\cdot}} [b + c] = [a \cdot (b + c)] \\ \text{und}\quad &[a] \mathbin{\tilde{\cdot}} [b] \mathbin{\tilde{+}} [a] \mathbin{\tilde{\cdot}} [c] = [a \cdot b] \mathbin{\tilde{+}} [a \cdot c] = [(a \cdot b) + (a \cdot c)]. \end{aligned}$$

Aufgrund des Distributivgesetzes (9.1a) im Ring R sind die Äquivalenzklassen gleich. Analog können wir das zweite Distributivgesetz (9.1b) zeigen, also $\big([a] \mathbin{\tilde{+}} [b]\big) \mathbin{\tilde{\cdot}} [c] = [a] \mathbin{\tilde{\cdot}} [c] \mathbin{\tilde{+}} [b] \mathbin{\tilde{\cdot}} [c]$.

Aussage (b): Wir wissen bereits aus Satz 8.21, dass π ein surjektiver Gruppenhomomorphismus $(R, +) \longrightarrow (R\,/\,J, \tilde{+})$ mit $\operatorname{Kern}(\pi) = J$ ist. Die Strukturverträglichkeit mit der Multiplikation ist gerade die Aussage (9.10b).

Aussage (c): Wir zeigen, dass $[1_R]$ das Einselement von $(R\,/\,J, \tilde{\cdot})$ ist. Für beliebiges $a \in R$ enthält die Äquivalenzklasse $[1_R] \mathbin{\tilde{\cdot}} [a]$ das Element a, also gilt $[1_R] \mathbin{\tilde{\cdot}} [a] = [a]$. Auch die Äquivalenzklasse $[a] \mathbin{\tilde{\cdot}} [1_R]$ enthält das Element a, also gilt auch $[a] \mathbin{\tilde{\cdot}} [1_R] = [a]$.

Ist $a \in R$ invertierbar, so zeigen wir nun, dass $[a]^{-1} = [a^{-1}]$ gilt:

$$[a] \mathbin{\tilde{\cdot}} [a^{-1}] = [a \cdot a^{-1}] = [1_R]$$

und analog

$$[a^{-1}] \mathbin{\tilde{\cdot}} [a] = [a^{-1} \cdot a] = [1_R].$$

Aussage (d): Wenn $\cdot$ kommutativ ist, dann gilt

$$[a] \mathbin{\tilde{\cdot}} [b] = [a \cdot b] = [b \cdot a] = [b] \mathbin{\tilde{\cdot}} [a],$$

also ist auch $\tilde{\cdot}$ kommutativ.

Aussage (*ii*): Wir wollen zeigen, dass aus der Wohldefiniertheit (9.10b) folgt, dass U ein Ideal ist, dass also $R \cdot U \subseteq U$ und $U \cdot R \subseteq U$ gilt.[40] Dazu seien $a \in R$ und $u \in U$ beliebig. Dann gilt $[u] = U = [0_R]$, und die Wohldefiniertheit liefert

$$[a \cdot u] = [a] \mathbin{\tilde{\cdot}} [u] = [a] \mathbin{\tilde{\cdot}} [0] = [a \cdot 0] = [0] = U,$$

also $a \cdot u \in U$. Das zeigt $R \cdot U \subseteq U$. Die Aussage $U \cdot R \subseteq U$ folgt analog aus

$$[u \cdot a] = [u] \mathbin{\tilde{\cdot}} [a] = [0] \mathbin{\tilde{\cdot}} [a] = [0 \cdot a] = [0] = U. \qquad \square$$

[40] Die Wohldefiniertheit von (9.10a) würde wie im Beweis von Aussage (*ii*) von Satz 8.21 nur die Information liefern, dass $(U, +)$ ein Normalteiler der Gruppe $(R, +)$ ist, was aber wegen der Kommutativität von $(R, +)$ ohnehin klar ist.

Bemerkung 9.31 (Faktorring, vgl. Bemerkung 8.22 zu Faktorgruppen).
Praktisch können wir den Faktorring $(R / J, \tilde{+}, \tilde{\cdot})$ benutzen, um wie im Ring $(R, +, \cdot)$ zu „rechnen", wobei jedoch Elemente a, b in derselben Äquivalenzklasse (für die also $a - b \in J$ gilt) nicht mehr unterschieden, sondern miteinander identifiziert werden. Der Faktorring $(R / J, \tilde{+}, \tilde{\cdot})$ ist also eine „gröbere Version" des Ringes $(R, +, \cdot)$. △

Beispiel 9.32 (Faktorring, vgl. Beispiel 9.29 zu Idealen).

(i) Für $m \in \mathbb{N}$ ist der Faktorring $\mathbb{Z} / m\mathbb{Z}$ von $(\mathbb{Z}, +, \cdot)$ der in Beispiel 9.6 bereits eingeführte Restklassenring modulo m. Nach Beispiel 9.23 ist dieser isomorph zu $(\mathbb{Z}_m, +_m, \cdot_m)$, dem Ring von $\mathbb{Z}$ modulo m.

(ii) Ist X eine Menge und $Y \subseteq X$, dann ist der Faktorring $\mathcal{P}(X) / \mathcal{P}(Y)$ von $(\mathcal{P}(X), \triangle, \cap)$ isomorph zu $(\mathcal{P}(X \setminus Y), \triangle, \cap)$. △

Bemerkung 9.33 (Ideale sind genau die Kerne von Ringhomomorphismen, vgl. Bemerkung 8.24 für Gruppenhomomorphismen).
Es sei $(R_1, +_1, \cdot_1)$ ein Ring.

(i) Nach Lemma 9.27 ist der Unterring $\operatorname{Kern}(f)$ für jeden beliebigen Ringhomomorphismus $f\colon R_1 \to R_2$ in irgendeinen Ring $(R_2, +_2, \cdot_2)$ immer ein Ideal von $(R_1, +_1, \cdot_1)$.

(ii) Umgekehrt gilt auch, dass jedes Ideal J von $(R_1, +_1, \cdot_1)$ der Kern eines Ringhomomorphismus ist: Dazu wählen wir einfach $R_2 \coloneqq (R_1 / J, \tilde{+}, \tilde{\cdot})$ als Zielring und die kanonische Surjektion $\pi\colon R_1 \to R_1 / J$ als Ringhomomorphismus. Dann gilt $\operatorname{Kern}(\pi) = J$. △

Lemma 9.34 (Durchschnitt von Idealen, vgl. Lemma 8.20 zu Normalteilern).
Es sei $(R, +, \cdot)$ ein Ring.

(i) Ist $(J_i, +, \cdot)_{i \in I}$ eine nichtleere Familie von Idealen von $(R, +, \cdot)$, dann ist auch $\bigcap_{i \in I} J_i$ ein Ideal von R.

(ii) Ist $\mathcal{J}$ eine nichtleere Menge von Idealen von $(R, +, \cdot)$, dann ist auch $\bigcap \mathcal{J}$ ein Ideal von $(R, +, \cdot)$.

Beweis. Der Beweis ist Gegenstand der Übung. □

Definition 9.35 (erzeugtes Ideal, Hauptideal).
Es sei $(R, +, \cdot)$ ein Ring und $E \subseteq R$.

(i) Dann heißt

$$(E) \coloneqq \bigcap \{J \mid (J, +, \cdot) \text{ ist Ideal von } (R, +, \cdot) \text{ und } E \subseteq J\} \tag{9.12}$$

das von E **erzeugte Ideal** (englisch: ideal generated by E) in $(R, +, \cdot)$.

(ii) Ist speziell $E = \{a\}$ für ein $a \in R$, so schreiben wir auch (a) statt $(\{a\})$ und nennen (a) das **von a erzeugte Hauptideal** (englisch: principal ideal).

(iii) Ein Ideal $(J, +, \cdot)$ heißt ein **Hauptideal**, wenn es ein $a \in R$ gibt, sodass $(a) = J$ gilt. △

Satz 9.36 (Darstellung des erzeugten Ideals, vgl. Satz 7.50).
Es sei $(R, +, \cdot)$ ein Ring, $E \subseteq R$ und $a \in R$.

(*i*) Dann gilt für das von E bzw. von a erzeugte Ideal:[41]

$$(E) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in E \,\cup\, -E \,\cup\, R\,E \,\cup\, E\,R \,\cup\, R\,E\,R\big) \Big\}, \tag{9.13a}$$

$$(a) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in \{\pm a\} \,\cup\, R\,a \,\cup\, a\,R \,\cup\, R\,a\,R\big) \Big\}. \tag{9.13b}$$

(*ii*) Ist $(R, +, \cdot)$ ein Ring mit Eins, dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in R\,E\,R\big) \Big\}, \tag{9.14a}$$

$$(a) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in R\,a\,R\big) \Big\}. \tag{9.14b}$$

Insbesondere ist $(1) = R$.

(*iii*) Ist $(R, +, \cdot)$ ein kommutativer Ring, dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in E \,\cup\, -E \,\cup\, R\,E\big) \Big\}, \tag{9.15a}$$

$$(a) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in \{\pm a\} \,\cup\, R\,a\big) \Big\}. \tag{9.15b}$$

(*iv*) Ist $(R, +, \cdot)$ ein kommutativer Ring mit Eins, dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^{n} a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in R\,E\big) \Big\}, \tag{9.16a}$$

$$(a) = R\,a. \tag{9.16b}$$

Insbesondere ist $(1) = R$.

In jedem Fall gilt $(\emptyset) = (0_R) = \{0_R\}$.

Beweis. Der Beweis ist Gegenstand der Übung. □

Beispiel 9.37 (erzeugtes Ideal, vgl. Beispiel 7.51 zu erzeugten Untergruppen)**.**

(*i*) Der **Kommutator** der Elemente a, b eines Ringes $(R, +, \cdot)$ ist definiert als

$$[a, b] \coloneqq a \cdot b - b \cdot a. \tag{9.17}$$

Beachte: Im Unterschied zum Kommutator zweier Elemente in einer Gruppe (8.11) werden hier beide Verknüpfungen zur Definition verwendet. Es gilt aber auch hier wieder, dass a

[41] Der leichteren Lesbarkeit wegen lassen wir in den folgenden Darstellungen das Symbol für die „Multiplikation“ $\cdot$ weg und schreiben beispielsweise $R\,E$ anstelle von $R \cdot E$.

und b genau dann (bzgl. $\cdot$) kommutieren, wenn der Kommutator das neutrale Element (bzgl. $+$) ergibt, wenn also $[a, b] = 0$ gilt. (**Quizfrage 9.6:** Klar?)

Das **Kommutatorideal** (englisch: commutator ideal) eines Ringes $(R, +, \cdot)$ ist das von Kommutatoren von R erzeugte Ideal[42], also

$$\big(\big\{[a, b] \,\big|\, a, b \in R\big\}\big). \tag{9.18}$$

Es wird kurz auch in der Form $\big([R, R]\big)$ oder (ungenau) als $[R, R]$ notiert.

Ist $(R, +, \cdot)$ ein beliebiger Ring und K sein Kommutatorideal, dann ist der Faktorring $(R / K, \tilde{+}, \tilde{\cdot})$ kommutativ.

Tatsächlich ist $(R / J, \tilde{+}, \tilde{\cdot})$ genau dann kommutativ, wenn das ausfaktorisierte Ideal J das Kommutatorideal von R enthält.

(ii) Das Zentrum eines Ringes (Beispiel 9.14) ist i. A. kein Ideal. △

§ 9.2 Der Homomorphiesatz für Ringe

Analog zum Homomorphiesatz für Gruppen Satz 8.25 gibt es einen Homomorphiesatz für Ringe. Er besagt wiederum, dass ein Ringhomomorphismus $f\colon R_1 \to R_2$ „nebenklassenweise“ wirkt. Er bildet eine gesamte Nebenklasse von $\operatorname{Kern}(f)$ auf ein- und dasselbe Element von R_2 ab und verschiedene Nebenklassen auf verschiedene Elemente. Das geschieht zudem strukturverträglich. Dadurch ist das $\operatorname{Bild}(f)$ eines solchen Ringhomomorphismus bereits im Wesentlichen (d. h. bis auf Isomorphie) festgelegt ist durch $(R_1, +, \cdot)$ und das Ideal $\operatorname{Kern}(f)$.

Satz 9.38 (**Homomorphiesatz für Ringe**[43], vgl. Satz 8.25 für Gruppen)**.**
Es seien $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ Ringe. Weiter sei $f\colon R_1 \to R_2$ ein Homomorphismus. Dann gilt

$$R_1 / \operatorname{Kern}(f) \cong \operatorname{Bild}(f) \tag{9.19a}$$

mit dem Isomorphismus

$$I([a]) \coloneqq f(a) \quad \text{für } [a] = a +_1 \operatorname{Kern}(f) \in R_1 / \operatorname{Kern}(f). \tag{9.19b}$$

Beweis. Der Ringhomomorphismus f ist insbesondere auch ein Gruppenhomomorphismus $f\colon (R_1, +_1) \to (R_2, +_2)$. Aus Satz 8.25 folgt daher, dass I ein Isomorphismus der Gruppen $(R_1 / \operatorname{Kern}(f), \tilde{+}_1)$ und $(\operatorname{Bild}(f), +_2)$ ist. Es bleibt zu zeigen, dass I auch ein Ringisomorphismus, also verträglich mit der Multiplikation ist. Dazu seien $a, b \in R_1$ beliebig, dann gilt:

$$\begin{aligned}
I([a] \,\tilde{\cdot}_1\, [b]) &= I([a] \cdot_1 [b]) && \text{per Definition von } \tilde{\cdot}_1 \\
&= f(a \cdot_1 b) && \text{nach Definition von } I \\
&= f(a) \cdot_2 f(b) && \text{da } f \text{ ein Ringhomomorphismus ist} \\
&= I([a]) \cdot_2 I([b]) && \text{nach Definition von } I.
\end{aligned}$$
□

[42]**Beachte:** Die Kommutatoruntergruppe einer Gruppe (Beispiel 8.17) wurde als die von den Kommutatoren erzeugte Untergruppe definiert, und die Normalteilereigenschaft konnte gezeigt werden. Im Unterschied dazu reicht es hier nicht aus, den von den Kommutatoren erzeugten Unterring zu betrachten, weil dieser i. A. kein Ideal ist.

[43]englisch: fundamental theorem on ring homomorphisms

§ 10 Körper

Literatur: Beutelspacher, 2014, Kapitel 2; Bosch, 2014, Kapitel 1.3; Fischer, Springborn, 2020, Kapitel 2.3; Deiser, 2024b, Kapitel 2.2

Ein Körper ist — wie ein Ring — eine algebraische Struktur mit zwei Verknüpfungen. In Anlehnung an die wichtigen Beispiele $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ mit den Verknüpfungen „Addition“ und „Multiplikation“ bezeichnen wir diese Verknüpfungen wieder mit $+$ und $\cdot$. Im Unterschied zu einem kommutativen Ring (wie etwa $\mathbb{Z}$) wird nun aber zusätzlich noch gefordert, dass alle Elemente (außer dem Nullelement) multiplikativ invertierbar sind:

Definition 10.1 (Körper)**.**
Ein **Körper** (englisch: field) $(K, +, \cdot)$ ist ein kommutativer Ring $(K, +, \cdot)$ mit der zusätzlichen Eigenschaft, dass $(K \setminus \{0_K\}, \cdot)$ eine abelsche Gruppe ist.[44] △

Da eine Gruppe nicht leer sein kann, besitzt ein Körper neben dem Nullelement 0_K mindestens noch ein weiteres Element, nämlich das multiplikativ neutrale Einselement $1_K \neq 0_K$. Anders gesagt ist ein Körper also ein kommutativer Ring mit Eins ungleich einem Nullring, in dem jedes Element außer dem Nullelement ein multiplikatives Inverses besitzt. (**Quizfrage 10.1:** Warum wird das Nullelement ausgenommen?)

Ausgeschrieben ist eine Menge mit zwei (inneren) Verknüpfungen $+$ („Addition“) und $\cdot$ („Multiplikation“) also genau dann ein Körper, wenn gilt:

(i) $(K, +)$ ist eine abelsche Gruppe.
(Das Nullelement bezeichnen wir mit 0_K.)

(ii) $(K, \cdot)$ ist ein abelsches Monoid mit Einselement $1_K \neq 0_K$, in dem alle Elemente außer 0_K invertierbar sind.

(iii) Es gelten die **Distributivgesetze** (englisch: distributive laws)

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \tag{10.1a}$$
$$(a + b) \cdot c = (a \cdot c) + (b \cdot c) \tag{10.1b}$$

für alle $a, b, c \in K$.[45]

Beispiel 10.2 (Körper und Gegenbeispiele)**.**

(i) $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind Körper mit dem Nullelement 0 und dem Einselement 1.

(ii) $(\mathbb{Z}_2, +_2, \cdot_2)$ aus Beispiel 7.22 mit den Verknüpfungstafeln aus Beispiel 7.2 ist ein Körper mit dem Nullelement 0 und dem Einselement 1. $\mathbb{Z}_2$ ist also der (bis auf Isomorphie eindeutige) kleinstmögliche Körper.

[44] Oft wird $K \setminus \{0_K\}$ in der Literatur als K^* oder als $K^\times$ abgekürzt. Wir verwenden diese Bezeichnungen jedoch hier nicht.

[45] Wie bereits in kommutativen Ringen fallen die beiden Distributivgesetze (10.1a) und (10.1b) zusammen. Es reicht also, eines von beiden zu prüfen.

(iii) Der Restklassenring $(\mathbb{Z}/4\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ mit dem Nullelement $[0]$ und dem Einselement $[1]$ aus Beispiel 9.6 ist *kein* Körper, da $[2]$ nicht das Nullelement, aber auch nicht multiplikativ invertierbar ist, denn es gilt $[2] \mathbin{\tilde{\cdot}} [a] \neq [1]$ für alle $a \in \mathbb{Z}$.

(iv) Es sei X eine Menge. Für die bisher besprochenen algebraischen Strukturen S (Halbgruppe, Monoid, Gruppe, Ring) galt, dass S^X, ausgestattet punktweise mit der oder den Verknüpfung(en) von S, die algebraische Struktur erbt, also ebenfalls Halbgruppe (Beispiel 7.4), Monoid (Beispiel 7.8), Gruppe (Beispiel 7.22) oder Ring (Beispiel 9.2) ist.

Wenn jedoch $(K, +, \cdot)$ ein Körper ist, dann ist $(K^X, +, \cdot)$ i. A. *kein* Körper, sondern nur ein kommutativer Ring mit Eins. (**Quizfrage 10.2:** Woran liegt das?) △

Satz 10.3 (Körper und Integritätsringe)**.**

(i) Jeder Körper $(K, +, \cdot)$ ist ein Integritätsring.[46]

(ii) Jeder endliche Integritätsring $(R, +, \cdot)$ ist ein Körper.

Beachte: Wie das Beispiel $\mathbb{Z}$ zeigt, sind unendliche Integritätsringe i. A. keine Körper.

Beweis. Aussage (i): Nach Definition 10.1 ist $(K, +, \cdot)$ ein kommutativer Ring mit dem Nullelement 0_K und dem Einselement $1_K \neq 0_K$. Insbesondere ist also K kein Nullring. Es bleibt zu zeigen, dass $(K, +, \cdot)$ nullteilerfrei ist. Es seien dazu $a, b \in K$ mit $a \cdot b = 0_K$. Wenn $a \neq 0_K$ ist, dann ist a als Element der Gruppe $(K \setminus \{0_K\})$ invertierbar, also gilt $b = a^{-1} \cdot 0_K = 0_K$. Ist dagegen $b \neq 0_K$, dann ist b invertierbar, und es folgt $a = 0_K \cdot b^{-1} = 0_K$. Das heißt, $(K, +, \cdot)$ ist nullteilerfrei.

Aussage (ii): Es sei $(R, +, \cdot)$ ein Integritätsring, also ein kommutativer, nullteilerfreier Ring mit dem Einselement 1_R ungleich einem Nullring. Wir wissen also bereits: $(R, +)$ ist eine abelsche Gruppe mit dem Nullelement 0_R, und $(R, \cdot)$ ist eine abelsche Halbgruppe mit dem Einselement $1_R \neq 0_R$ (Lemma 9.3). Aus der Nullteilerfreiheit (9.3) folgt, dass $R \setminus \{0_R\}$ bzgl. $\cdot$ abgeschlossen ist, also ist auch $(R \setminus \{0_R\}, \cdot)$ ein abelsches Monoid mit dem Einselement 1_R.

Es bleibt zu zeigen, dass $(R \setminus \{0_R\}, \cdot)$ sogar eine Gruppe ist. Dazu nutzen wir das Gruppenkriterium Lemma 7.25. Zu beliebigem $a \in R \setminus \{0\}$ betrachten wir die Rechtstranslation $\cdot_a$ auf dem Monoid $R \setminus \{0\}$. Diese ist injektiv, denn nach Distributivgesetz (9.1b) gilt

$$b \cdot a = c \cdot a \quad \Rightarrow \quad b \cdot a - c \cdot a = 0_R \quad \Rightarrow \quad (b - c) \cdot a = 0_R,$$

und da R nullteilerfrei und $a \neq 0_R$ ist, folgt $b = c$. Da nun R und damit $R \setminus \{0\}$ eine endliche Menge ist, gilt nach Satz 6.35, dass $\cdot_a$ auch surjektiv ist.

Ein analoges Argument zeigt, dass auch alle Linkstranslationen auf $R \setminus \{0\}$ surjektiv sind. Aus dem Gruppenkriterium Lemma 7.25 folgt nun, dass $(R \setminus \{0\}, \cdot)$ eine Gruppe ist. □

Folgerung 10.4 (Körpereigenschaft des Restklassenringes $\mathbb{Z}/m\mathbb{Z}$ und des Ringes $\mathbb{Z}_m$ von $\mathbb{Z}$ modulo m, vgl. Satz 9.11 zur Nullteilerfreiheit)**.**
Der Restklassenring modulo m $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ sowie der zu ihm isomorphe Ring (Beispiel 9.23) von $\mathbb{Z}$ modulo m $(\mathbb{Z}_m, +_m, \cdot_m)$ sind Körper genau dann, wenn $m \in \mathbb{N}$ eine Primzahl ist. In diesem

[46] also ein kommutativer, nullteilerfreier Ring mit Eins ungleich einem Nullring, siehe Definition 9.7

Fall nennen wir sie auch **Restklassenkörper modulo m** oder **Körper von $\mathbb{Z}$ modulo m** (englisch: field of $\mathbb{Z}$ modulo m).

Beweis. In Satz 9.11 haben wir gezeigt, dass $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ genau dann ein Integritätsring ist, wenn $m \in \mathbb{N}$ eine Primzahl ist. Da aber $\mathbb{Z}/m\mathbb{Z}$ nur endlich viele (nämlich m) Elemente hat, ist Integritätsring zu sein gleichbedeutend mit der Körpereigenschaft. Nach Beispiel 9.23 sind $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ und $(\mathbb{Z}_m, +_m, \cdot_m)$ als Ringe isomorph, also gelten dieselben Eigenschaften auch für $(\mathbb{Z}_m, +_m, \cdot_m)$. □

Bemerkung 10.5 (Charakteristik von Körpern)**.**

Die Definition 9.4 der Charakteristik eines Ringes wird auch auf Körper angewendet. Für Körper ist die Charakteristik nach Lemma 9.10 also entweder 0 oder eine Primzahl.

Für die Körper aus Beispiel 10.2 gilt $\operatorname{char}(\mathbb{Q}) = \operatorname{char}(\mathbb{R}) = \operatorname{char}(\mathbb{C}) = 0$ und $\operatorname{char}(\mathbb{Z}_2) = 2$. Für die Körper $\mathbb{Z}/m\mathbb{Z}$ und $\mathbb{Z}_m$ mit $m \in \mathbb{N}$ prim gilt $\operatorname{char}(\mathbb{Z}/m\mathbb{Z}) = \operatorname{char}(\mathbb{Z}_m) = m$, vgl. Beispiel 9.5. △

Definition 10.6 (Unterkörper, vgl. Definition 9.12 eines Unterringes)**.**

Es sei $(K, +, \cdot)$ ein Körper.

(i) Eine Teilmenge $U \subseteq K$ heißt ein **Teilkörper** oder **Unterkörper** (englisch: subfield) von $(K, +, \cdot)$, wenn U bzgl. $+$ und bzgl. $\cdot$ abgeschlossen und wenn $(U, +, \cdot)$ selbst wieder ein Körper ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, +)$ eine Untergruppe von $(K, +)$ und wenn $(U \setminus \{0\}, \cdot)$ eine Untergruppe von $(K \setminus \{0\}, \cdot)$ ist.

(ii) Ein Unterkörper $(U, +, \cdot)$ von $(K, +, \cdot)$ heißt **echt** (englisch: proper subfield), wenn $U \subsetneq K$ gilt. △

Beachte: Insbesondere folgt aus Lemma 7.43, dass das Nullelement 0_K von K notwendigerweise auch das Nullelement des Unterkörpers ist, und ebenso, dass das Einselement 1_K auch das Einselement des Unterkörpers ist.

Die Prüfung einer Teilmenge $U \subseteq K$ auf die Unterkörper-Eigenschaft lässt sich mit folgendem Kriterium erreichen:

Satz 10.7 (Unterkörperkriterium, vgl. Unterringkriterium Satz 9.13)**.**

Es sei $(K, +, \cdot)$ ein Körper und $U \subseteq K$. Dann sind äquivalent:

(i) $(U, +, \cdot)$ ist ein Unterkörper von $(K, +, \cdot)$.

(ii) U besitzt mindestens zwei Elemente, und für alle $a, b \in U$ gilt $a - b \in U$ sowie $a \cdot b^{-1} \in U$, sofern $b \neq 0_K$ ist.[47]

[47]kurz: $U - U \subseteq U$ und $U \cdot (U \setminus \{0_K\})^{-1} \subseteq U$

Beweis. Aussage $(i) \Rightarrow$ Aussage (ii): Es sei $(U, +, \cdot)$ ein Unterkörper von $(K, +, \cdot)$. Dann enthält U notwendigerweise das Nullelement 0_K von $(K, +, \cdot)$, also das neutrale Element der Untergruppe $(U, +)$ von $(K, +)$. Für $a, b \in U$ gilt $-b \in U$ nach Lemma 7.43. Da U bzgl. $+$ abgeschlossen ist, folgt $a - b = a + (-b) \in U$. Analog gilt nach Lemma 7.43 auch $b^{-1} \in U$ für $b \neq 0_K$, und da U bzgl. $\cdot$ abgeschlossen ist, folgt $a \cdot b^{-1} \in U$.

Aussage $(ii) \Rightarrow$ Aussage (i): Nach Voraussetzung erfüllt $(U, +)$ das Untergruppenkriterium (Satz 7.44), also ist $(U, +)$ eine Untergruppe von $(K, +)$. Zudem erfüllt $(U \setminus \{0_K\}, \cdot)$ nach Voraussetzung ebenfalls das Untergruppenkriterium, also ist $(U \setminus \{0_K\}, \cdot)$ eine Untergruppe von $(K \setminus \{0_K\}, \cdot)$. Damit ist $(U, +, \cdot)$ ein Unterkörper von $(K, +, \cdot)$. □

Beispiel 10.8 (Unterkörper)**.**

(i) $(\mathbb{Q}, +, \cdot)$ ist ein Unterkörper von $(\mathbb{R}, +, \cdot)$, und $(\mathbb{R}, +, \cdot)$ ist ein Unterkörper von $(\mathbb{C}, +, \cdot)$.

(ii) Der Restklassenkörper $\mathbb{Z}/m\mathbb{Z}$ (mit $m \in \mathbb{N}$ prim) und der zu ihm isomorphe Körper $\mathbb{Z}_m$ besitzen keine echten Unterkörper. (**Quizfrage 10.3:** Wie sieht man das?) △

Bemerkung 10.9 („Unterkörper sein" ist eine Ordnungsrelation, vgl. Bemerkung 9.15 zu Unterringen)**.**

(i) Die Relation „ist Unterkörper von" ist eine partielle Ordnung auf der Klasse aller Körper.

(ii) Insbesondere ist die Menge aller Unterkörper eines bestimmten Körpers $(K, +, \cdot)$ durch die Unterkörperhalbordnung partiell geordnet. Diese Ordnung stimmt mit der Inklusionshalbordnung überein. △

Lemma 10.10 (Durchschnitt von Unterkörpern, vgl. Lemma 9.16 zu Unterringen)**.**
Es sei $(K, +, \cdot)$ ein Körper.

(i) Ist $(U_i, +, \cdot)_{i \in I}$ eine nichtleere Familie von Unterkörpern von $(K, +, \cdot)$, dann ist auch $\bigcap_{i \in I} U_i$ ein Unterkörper von $(K, +, \cdot)$.

(ii) Ist $\mathcal{U}$ eine nichtleere Menge von Unterkörpern von $(K, +, \cdot)$, dann ist auch $\bigcap \mathcal{U}$ ein Unterkörper von $(K, +, \cdot)$.

Beweis. Der Beweis ist eine einfache Anwendung des Unterkörperkriteriums (Satz 10.7). □

Definition 10.11 (Körperhomomorphismus, vgl. Definition 9.17 eines Ringhomomorphismus)**.**
Es seien $(K_1, +_1, \cdot_1)$ und $(K_2, +_2, \cdot_2)$ zwei Körper.

(i) Eine Abbildung $f\colon K_1 \to K_2$ heißt **strukturverträglich** oder ein **(Homomorphismus von Körpern** (englisch: field homomorphism), wenn gilt:

$$f(a +_1 b) = f(a) +_2 f(b) \quad \text{für alle } a, b \in K_1, \tag{10.2a}$$
$$f(a \cdot_1 b) = f(a) \cdot_2 f(b) \quad \text{für alle } a, b \in K_1, \tag{10.2b}$$
$$f(1_{K_1}) = 1_{K_2}. \tag{10.2c}$$

(ii) Ist $f\colon K_1 \to K_2$ strukturverträglich und gilt $(K_1, +_1, \cdot_1) = (K_2, +_2, \cdot_2)$, so sprechen wir auch von einem **Endomorphismus eines Körpers** (englisch: field endomorphism).

(*iii*) Ist $f\colon K_1 \longrightarrow K_2$ strukturverträglich und bijektiv, so heißt f auch **strukturerhaltend** oder ein **Isomorphismus von Körpern** (englisch: field isomorphism). In diesem Fall nennen wir $(K_1, +_1, \cdot_1)$ und $(K_2, +_2, \cdot_2)$ auch zueinander **isomorphe Körper** (englisch: isomorphic fields) und schreiben

$$(K_1, +_1, \cdot_1) \cong (K_2, +_2, \cdot_2).$$

(*iv*) Ist $f\colon K_1 \longrightarrow K_2$ strukturverträglich und bijektiv und gilt $(K_1, +_1, \cdot_1) = (K_2, +_2, \cdot_2)$, so sprechen wir auch von einem **Automorphismus eines Körpers** (englisch: field automorphism). △

Da die Bedingungen (10.2) mit denen aus (9.5) übereinstimmen, ist ein Körperhomomorphismus nichts anderes als ein Homomorphismus von Ringen mit Eins, der speziell zwischen Körpern eingesetzt wird. Insbesondere haben wir auch hier wie in (9.6) wieder

$$f(0_{K_1}) = 0_{K_2}. \tag{10.3}$$

Anstelle der Bedingung (10.2c) reicht es auch aus, zu fordern, dass $f(1_{K_1}) \neq 0_{K_2}$ gilt, denn daraus ergibt sich notwendig $f(1_{K_1}) = 1_{K_2}$, vgl. (8.3). (**Quizfrage 10.4:** Wie nämlich?)

Auch für Körper gilt analog zu Satz 9.18 und Folgerung 9.19:

Satz 10.12 (Komposition von Körperhomomorphismen, Inverse von Körperisomorphismen, vgl. Satz 9.18 zu Ringhomomorphismen)**.**

Es seien $(K_1, +_1, \cdot_1)$, $(K_2, +_2, \cdot_2)$ und $(K_3, +_3, \cdot_3)$ drei Körper.

(*i*) Sind $f\colon K_1 \longrightarrow K_2$ und $g\colon K_2 \longrightarrow K_3$ Körperhomomorphismen, dann ist auch $g \circ f\colon K_1 \longrightarrow K_3$ ein Körperhomomorphismus.

(*ii*) Ist $f\colon K_1 \longrightarrow K_2$ ein Körperisomorphismus, dann ist auch $f^{-1}\colon K_2 \longrightarrow K_1$ ein Körperisomorphismus.

Folgerung 10.13 (Isomorphie von Körpern ist eine Äquivalenzrelation, vgl. Folgerung 9.19 zur Isomorphie von Ringen)**.**

Isomorphie ist eine Äquivalenzrelation auf der Klasse aller Körper.

Im Unterschied zu Homomorphismen anderer algebraischer Strukturen sind Körperhomomorphismen immer injektiv:

Lemma 10.14 (Körperhomomorphismen sind injektiv)**.**

Es seien $(K_1, +_1, \cdot_1)$ und $(K_2, +_2, \cdot_2)$ zwei Körper und $f\colon K_1 \longrightarrow \mathbb{K}_2$ ein Homomorphismus. Dann ist f injektiv.

Beweis. Wir nehmen $a \neq b$, aber $f(a) = f(b)$ an. Dann ergibt sich der Widerspruch

$$\begin{aligned}
1_{K_2} &= f(1_{K_1}) && \text{wegen (10.2c)} \\
&= f\big((a -_1 b)^{-1} \cdot_1 (a -_1 b)\big) && \text{da } a -_1 b \neq 0_{K_1} \text{ vorausgesetzt wurde} \\
&= f\big((a -_1 b)^{-1}\big) \cdot_2 f\big(a -_1 b\big) && \text{wegen (10.2a)} \\
&= f\big((a -_1 b)^{-1}\big) \cdot_2 \big(f(a) -_2 f(b)\big) && \text{wegen (10.2b)} \\
&= f\big((a -_1 b)^{-1}\big) \cdot_2 0_{K_2} && \text{da } f(a) = f(b) \text{ vorausgesetzt wurde} \\
&= 0_{K_2} && \text{nach Lemma 9.3.}
\end{aligned}$$

□

Beispiel 10.15 (Körperhomomorphismen)**.**

(*i*) Die Einbettungen $(\mathbb{Q}, +, \cdot) \ni x \mapsto x \in (\mathbb{R}, +, \cdot)$ und $(\mathbb{R}, +, \cdot) \ni x \mapsto x \in (\mathbb{C}, +, \cdot)$ sind Körperhomomorphismen.

(*ii*) Der Körper der rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ besitzt außer der Identität keine weiteren Körperautomorphismen.

(*iii*) Auch der Körper der rellen Zahlen $(\mathbb{R}, +, \cdot)$ besitzt außer der Identität keine weiteren Körperautomorphismen.

(*iv*) Die komplexe Konjugation $(\mathbb{C}, +, \cdot) \ni x \mapsto \overline{x} \in (\mathbb{C}, +, \cdot)$ ist ein Körperautomorphismus.

(*v*) Realteil $(\mathbb{C}, +, \cdot) \ni x \mapsto \operatorname{Re} x \in (\mathbb{R}, +, \cdot)$ und Imaginärteil $(\mathbb{C}, +, \cdot) \ni x \mapsto \operatorname{Im} x \in (\mathbb{R}, +, \cdot)$ sind keine Körperhomomorphismen, da sie nicht injektiv sind. △

Lemma 10.16 (Körper mit Charakteristik 0 enthalten $\mathbb{Q}$, vgl. Lemma 9.20 für Ringe mit Eins)**.**
Es sei $(K, +, \cdot)$ ein Körper mit $\operatorname{char}(K) = 0$. Dann enthält K einen Unterkörper, der isomorph zu $\mathbb{Q}$ ist.

Beweis. Der Beweis ist Gegenstand der Übung. □

Definition 10.17 (Bild und Kern eines Körperhomomorphismus, vgl. Definition 9.21 von Bild und Kern eines Ringhomomorphismus)**.**
Es seien $(K_1, +_1, \cdot_1)$ und $(K_2, +_2, \cdot_2)$ Körper mit den Nullelementen 0_{K_1} bzw. 0_{K_2}. Weiter sei $f\colon K_1 \to K_2$ ein Homomorphismus.

(*i*) Das **Bild** von f ist definiert als

$$\operatorname{Bild}(f) \coloneqq \{f(a_1) \in K_2 \mid a_1 \in K_1\} = f(K_1). \tag{10.4}$$

(*ii*) Der **Kern** von f ist definiert als

$$\operatorname{Kern}(f) \coloneqq \{a_1 \in K_1 \mid f(a_1) = 0_{K_2}\} = f^{-1}(\{0_{K_2}\}) \tag{10.5}$$

△

Analog zu Lemma 9.22 gilt auch bei Körperhomomorphismen wieder, dass das Bild eines Körperhomomorphismus $f\colon K_1 \longrightarrow K_2$ ein Unterkörper von K_2 ist. (**Quizfrage 10.5:** Beweis?) Der Kern eines Körperhomomorphismus eines Körperhomomorphismus $f\colon K_1 \longrightarrow K_2$ ist dagegen niemals ein Unterkörper von K_2, da aufgrund der Injektivität von f notwendigerweise $\operatorname{Kern}(f) = \{0_{K_1}\}$ gilt, ein Unterkörper aber mindestens zwei Elemente hat.

Bemerkung 10.18 (Ausfaktorisieren bei Körpern)**.**
Wir hatten in § 8.1 durch Ausfaktorisieren einer normalen Untergruppe eines Gruppe auf der Faktormenge wieder eine Gruppenstruktur erhalten. Analog konnten wir in § 9.1 ein Ideal aus einem Ring ausfaktorisieren und haben auf der Faktormenge eine Ringstruktur erhalten. Diese Idee funktioniert bei Körpern aber nicht. Würden wir für einen Körper K und einen Unterkörper U versuchen, wie in (9.10) die Faktorverknüpfungen $\widetilde{+}$ und $\widetilde{\cdot}$ einzuführen in der Hoffnung, auf der Faktormenge $K \,/\, U$ wieder eine Körperstruktur zu bekommen, dann könnten wir gleichzeitig K auch als Ring betrachten und würden insbesondere auf $K \,/\, U$ auch eine Ringstruktur bekommen. Nach Satz 9.30 wird das aber nur dann passieren, wenn U ein Ideal des Ringes K ist. Man kann jedoch zeigen, dass es in einem Körper nur die beiden trivialen Ideale $U_1 = \{0\}$ und $U_2 = K$ gibt. U_1 ist aber kein Unterkörper von K (da einelementig), und $K \,/\, U_2$ besteht nur aus einem einzigen Element (Nebenklasse), worauf sich keine Körperstruktur definieren lässt. △

Abschließend kombinieren wir noch die Begriffe Körper und Ordnungsrelation, was uns zum Begriff des **geordneten Körpers** führt. Dabei handelt es sich um einen Körper mit einer Totalordnung, die in gewissem Sinne mit den Verknüpfungen der Körpers verträglich ist. Das Nullelement spielt dabei eine besondere Rolle:

Definition 10.19 (geordneter Körper)**.**
Es seien $(K, +, \cdot)$ ein Körper mit dem Nullelement 0_K und $\leqslant$ eine Totalordnung auf K.

(i) Der Körper heißt **geordnet** (englisch: ordered field) bzgl. der Totalordnung $\leqslant$, wenn

$$\alpha \leqslant \beta \quad \Rightarrow \quad \alpha + \gamma \leqslant \beta + \gamma \tag{10.6a}$$
$$\alpha \geqslant 0_K \text{ und } \beta \geqslant 0_K \quad \Rightarrow \quad \alpha \cdot \beta \geqslant 0_K \tag{10.6b}$$

für alle $\alpha, \beta, \gamma \in K$ gilt.

(ii) $\alpha \in K$ heißt **nichtnegativ**, wenn $\alpha \geqslant 0_K$ ist.

(iii) $\alpha \in K$ heißt **positiv**, wenn $\alpha \geqslant 0_K$ und $\alpha \neq 0_K$ ist.

(iv) $\alpha \in K$ heißt **nichtpositiv**, wenn $\alpha \leqslant 0_K$ ist.

(v) $\alpha \in K$ heißt **negativ**, wenn $\alpha \leqslant 0_K$ und $\alpha \neq 0_K$ ist. △

Lemma 10.20 (Rechenregeln in geordneten Körpern)**.**
Es sei $(K, +, \cdot)$ mit der Totalordnung $\leqslant$ ein geordneter Körper. Dann gilt für $\alpha, \beta, \gamma, \delta \in K$:

(i) $\alpha \geqslant 0_K \quad \Leftrightarrow \quad -\alpha \leqslant 0_K$

(ii) $\alpha \leqslant \beta$ und $\gamma \leqslant \delta \quad \Rightarrow \quad \alpha + \gamma \leqslant \beta + \delta$

(iii) $\alpha \leqslant \beta$ und $\gamma \geqslant 0_K \quad \Rightarrow \quad \alpha \cdot \gamma \leqslant \beta \cdot \gamma$

(*iv*) $\alpha \leqslant \beta$ und $\gamma \leqslant 0_K \quad \Rightarrow \quad \alpha \cdot \gamma \geqslant \beta \cdot \gamma$

(*v*) $\alpha^2 \geqslant 0_K$

(*vi*) $\alpha \neq 0_K \quad \Rightarrow \quad \alpha^2 > 0_K$. Insbesondere gilt $1_K > 0_K$.

(*vii*) $\alpha > 0_K \quad \Rightarrow \quad \alpha^{-1} > 0_K$

(*viii*) $\beta > \alpha > 0_K \quad \Rightarrow \quad \alpha^{-1} > \beta^{-1} > 0_K$

(*ix*) $n\,1_K > 0_K$ für alle $n \in \mathbb{N}$. Insbesondere gilt notwendigerweise $\operatorname{char}(K) = 0 \in \mathbb{Z}$.

Beweis. Aussage (*i*):

$$\begin{aligned} & \alpha \geqslant 0_K && \Rightarrow && 0_K = \alpha + (-\alpha) \geqslant 0_K + (-\alpha) = -\alpha \\ \text{und} \quad & -\alpha \leqslant 0_K && \Rightarrow && 0_K = -\alpha + \alpha \leqslant 0_K + \alpha = \alpha. \end{aligned}$$

Aussage (*ii*):

$$\begin{aligned} & \alpha \leqslant \beta && \Rightarrow && \alpha + \gamma \leqslant \beta + \gamma \\ \text{und} \quad & \gamma \leqslant \delta && \Rightarrow && \beta + \gamma \leqslant \beta + \delta. \end{aligned}$$

Die Transitivität der Ordnung zeigt, dass dann auch $\alpha + \gamma \leqslant \beta + \delta$ gilt.

Aussage (*iii*):

$$\begin{aligned} & \alpha \leqslant \beta \\ \Rightarrow \quad & 0_K \leqslant \beta - \alpha \\ \Rightarrow \quad & 0_K \leqslant (\beta - \alpha) \cdot \gamma && \text{wegen } \gamma \geqslant 0_k \\ \Rightarrow \quad & 0_K \leqslant \beta \cdot \gamma - \alpha \cdot \gamma && \text{wegen des Distributivgesetzes (10.1b)} \\ \Rightarrow \quad & \alpha \cdot \gamma \leqslant \beta \cdot \gamma. \end{aligned}$$

Aussage (*iv*): Aus $\gamma \leqslant 0_K$ folgt $-\gamma \geqslant 0_K$ nach Aussage (*i*). Weiter folgt mit Aussage (*iii*) dann $\alpha \cdot (-\gamma) \leqslant \beta \cdot (-\gamma)$, also $\beta \cdot \gamma \leqslant \alpha \cdot \gamma$.

Aussage (*v*): Nehmen wir zunächst $\alpha \geqslant 0_K$ an. Es folgt $\alpha^2 = \alpha \cdot \alpha \geqslant 0_K$. Im Fall $\alpha \leqslant 0_K$ ist $-\alpha \geqslant 0_K$. Es folgt $\alpha^2 = (-\alpha) \cdot (-\alpha) \geqslant 0_K$.

Aussage (*vi*): Die Behauptung folgt aus der Nullteilerfreiheit des Körpers K (Satz 10.3). Die Wahl $\alpha = 1_K$ zeigt $\alpha^2 = 1_K \cdot 1_K = 1_K > 0_K$.

Aussage (*vii*): Wegen $\alpha > 0_K$ ist $\alpha \neq 0_K$, also multiplikativ invertierbar. Das Inverse α^{-1} ist ebenfalls $\neq 0_K$. Die Annahme $\alpha^{-1} < 0_K$ würde $1_K = \alpha \cdot \alpha^{-1} < 0_K$ ergeben. Andererseits ist aber $1_K > 0_K$ nach Aussage (*vi*), Widerspruch. Also muss $\alpha^{-1} > 0_K$ gelten.

Aussage (*viii*):

$$\begin{aligned} & \beta > \alpha > 0_K \\ \Rightarrow \quad & 1_K > \alpha \cdot \beta^{-1} > 0_K && \text{wegen } \beta^{-1} > 0_K \text{ nach Aussage (vii)} \\ \Rightarrow \quad & \alpha^{-1} > \beta^{-1} > 0_K && \text{wegen } \alpha^{-1} > 0_K \text{ nach Aussage (vii)} \end{aligned}$$

Aussage (*ix*): Aus Aussage (*vi*) folgt $1_K > 0_K$. Es folgt weiter $1_K + 1_K > 1_K + 0_K = 1_K > 0_K$. Induktiv zeigt man $n\,1_K > 0_K$ für alle $n \in \mathbb{N}$. Folglich ist $n\,1_K \neq 0_K$ für all $n \in \mathbb{N}$, d. h., $\operatorname{char}(K) = 0_K$. □

Beispiel 10.21 (geordneter Körper)**.**

(i) Die rationalen Zahlen $\mathbb{Q}$ mit der bekannten Totalordnung bilden einen geordneten Körper.

(ii) Die reellen Zahlen $\mathbb{R}$ mit der bekannten Totalordnung bilden einen geordneten Körper.

(iii) Die komplexen Zahlen $\mathbb{C}$ sind mit keiner Totalordnung ein geordneter Körper, da $i^2 = -1$ der Aussage (vi) aus Lemma 10.20 widerspricht. △

Quizfrage 10.6: Wie würden Sie den Begriff **geordneter Ring** definieren? Welche der Eigenschaften aus Lemma 10.20 gelten noch, welche nicht?

Quizfrage 10.7: Ist auch der Begriff der **geordneten Gruppe** sinnvoll? Wie sieht die Definition aus, und welche Eigenschaften gelten?

Ende der Vorlesung 14

Ende der Woche 7

Kapitel 3 Vektorräume

§ 11 Vektorräume

Literatur: Beutelspacher, 2014, Kapitel 3; Bosch, 2014, Kapitel 1; Fischer, Springborn, 2020, Kapitel 2.4–2.6; Jänich, 2008, Kapitel 2

Vektorräume sind die zentralen Strukturen in der *linearen* Algebra. Zu einem Vektorraum V gehört immer ein zugrundeliegender Körper $(K, +, \cdot)$. In Anlehnung an dessen Verknüpfungen bezeichnen wir die beiden Verknüpfungen im Vektorraum V vorübergehend mit $\oplus$ und $\odot$.

Definition 11.1 (Vektorraum)**.**
Es sei $(K, +, \cdot)$ ein Körper. Ein **Vektorraum** (englisch: vector space) oder **linearer Raum** (englisch: linear space) $(V, \oplus, \odot)$ **über dem Körper** K (kurz: ein K**-Vektorraum**) ist eine Menge V mit einer inneren Verknüpfung $\oplus\colon V \times V \to V$ und einer **äußeren Verknüpfung** (englisch: (outer) operation) $\odot\colon K \times V \to V$, die die folgenden Bedingungen erfüllen:

(*i*) $(V, \oplus)$ ist eine abelsche Gruppe. Das Nullelement bezeichnen wir mit 0_V, genannt der **Nullvektor** (englisch: zero vector).

(*ii*) Für die Verknüpfung $\odot$, genannt **S-Multiplikation**[1] (englisch: S-multiplication) oder **skalare Multiplikation** (englisch: scalar multiplication), gelten die folgenden Gesetze für alle $\alpha, \beta \in K$ und $u, v \in V$: das **„gemischte" Assoziativgesetz** (englisch: „mixed" associative law)

$$(\alpha \cdot \beta) \odot v = \alpha \odot (\beta \odot v) \tag{11.1a}$$

sowie die **„gemischten" Distributivgesetze**[2] (englisch: „mixed" distributive laws)

$$\alpha \odot (u \oplus v) = (\alpha \odot u) \oplus (\alpha \odot v) \tag{11.1b}$$

$$(\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v). \tag{11.1c}$$

Weiterhin ist das neutrale Element 1_K bzgl. der Multiplikation $\cdot$ im Körper K auch neutral bzgl. der S-Multiplikation $\odot$:

$$1_K \odot v = v. \tag{11.1d}$$

In einem K-Vektorraum V heißen die Elemente von V auch **Vektoren** (englisch: vectors). Die Elemente von K heißen **Skalare** (englisch: scalars), sie werden häufig mit griechischen Kleinbuchstaben $\alpha, \beta, \ldots$ bezeichnet. K selbst heißt der **Skalarkörper** (englisch: scalar field) von V. △

[1] **S-Multiplikation** ist der von uns bevorzugte Begriff, um Verwechslungen mit dem Begriff **Skalarprodukt** auszuschließen, siehe später.

Bemerkung 11.2 (abkürzende Schreibweisen in Vektorräumen)**.**

(*i*) Das Inverse zu $v \in V$ bzgl. $\oplus$ bezeichnen wir mit $\ominus v$. Die Bezeichnung $u \ominus v$ steht für $u \oplus (\ominus v)$.

(*ii*) Mit der in Bemerkung 7.20 eingeführten Notation haben wir

$$n\, v = \underbrace{v \oplus \cdots \oplus v}_{n\text{-mal}} = (1_K \odot v) \oplus \cdots \oplus (1_K \odot v) = (1_K + \cdots + 1_K) \odot v = (n\, 1_K) \odot v$$

für $n \in \mathbb{N}$. Bezeichnen wir das Element $n\, 1_K \in K$ einfach mit n, so erhalten wir also $n\, v = n \odot v$. Tatsächlich gilt das auch für $n \in \mathbb{Z}$. Das werden wir später noch nutzen, um unsere Notation zu vereinfachen (Bemerkung 11.13).

(*iii*) Wir vereinbaren, dass $\odot$ stärker bindet als $\oplus$, also könnten wir die rechte Seite in (11.1b) auch in der Form $\alpha \odot u \;\oplus\; \alpha \odot v$ schreiben.

(*iv*) Die Konvention, dass bei der skalaren Multiplikation $\odot\colon K \times V \to V$ die Skalare auf der linken Seite stehen, ist willkürlich. Wir können parallel auch die andere äußere Verknüpfung $\boxdot\colon V \times K \to V$ definieren, und zwar durch $v \boxdot \alpha \coloneqq \alpha \odot v$. Dann gelten auch dafür die Gesetze (11.1), wobei überall $\odot$ durch $\boxdot$ ersetzt wird und die beiden Argumente dieser Verknüpfung vertauscht werden. Aufgrund der Ähnlichkeit unterscheiden wir nicht zwischen der linken skalaren Multiplikation $\odot$ und der rechten skalaren Multiplikation $\boxdot$, sondern schreiben in Zukunft einfach $\odot$ für beide.

(*v*) Wir behalten die unterschiedliche Notation der Verknüpfungen „$+$" in K und „$\oplus$" in V wie auch von „$\cdot$" in K und „$\odot$" in V zur Verdeutlichung zunächst bei. Später werden wir diese jedoch nur noch als „$+$" bzw. „$\cdot$" notieren, siehe Bemerkung 11.13. △

Beispiel 11.3 (Vektorraum)**.**

(*i*) Über jedem Körper $(K, +, \cdot)$ gibt es den (bis auf Isomorphie eindeutigen) Vektorraum $(V, \oplus, \odot)$ mit nur einem einzigen Element, nämlich $V = \{0_V\}$. Die Verknüpfungen $\oplus$ und $\odot$ sind dann eindeutig festgelegt. (**Quizfrage 11.1:** Wie nämlich?) Dieser Raum heißt ein **Nullraum** (englisch: zero vector space) über K.

(*ii*) Jeder Körper $(K, +, \cdot)$, ausgestattet mit den Verknüpfungen $\oplus \coloneqq +\colon K \times K \to K$ und $\odot \coloneqq \cdot\colon K \times K \to K$, ist ein Vektorraum über sich selbst.

(*iii*) Allgemeiner sei $(K, +, \cdot)$ ein Körper und U ein Unterkörper. Dann ist $(K, +, \cdot)$ ein U-Vektorraum mit den Verknüpfungen $\oplus \coloneqq +\colon K \times K \to K$ und $\odot \coloneqq \cdot\colon U \times K \to K$. Beispielsweise ist $\mathbb{R}$ ein $\mathbb{Q}$-Vektorraum, und $\mathbb{C}$ ist sowohl ein $\mathbb{R}$-Vektorraum als auch ein $\mathbb{Q}$-Vektorraum.

(*iv*) Es sei $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Dann ist die Menge

$$K_n \coloneqq \Big\{ \big(x_1 \quad \cdots \quad x_n\big) \;\Big|\; x_i \in K \text{ für } i = 1, \ldots, n \Big\}, \tag{11.2}$$

ausgestattet mit der **komponentenweisen Addition** (englisch: componentwise addition) und der **komponentenweisen S-Multiplikation** (englisch: componentwise S-multiplication)

$$\big(x_1 \quad \cdots \quad x_n\big) \oplus \big(y_1 \quad \cdots \quad y_n\big) \coloneqq \big(x_1 + y_1 \quad \cdots \quad x_n + y_n\big)\,, \tag{11.3a}$$

$$\alpha \odot \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \coloneqq \begin{pmatrix} \alpha \cdot x_1 & \cdots & \alpha \cdot y_n \end{pmatrix}, \tag{11.3b}$$

ein K-Vektorraum, genannt der **Vektorraum der Zeilenvektoren** (englisch: row vector space) über K der **Dimension** n.[3] Der Nullvektor ist $\begin{pmatrix} 0_K & \cdots & 0_K \end{pmatrix} \in K_n$.

Es wird sich als praktisch erweisen, auch den Fall $n = 0$ zuzulassen. Der Raum K_0 besteht dann nur aus einem Element, dem leeren Zeilenvektor $()$. K_0 ist also ein Nullraum über K. Es gilt $\alpha \odot () = ()$ für alle $\alpha \in K$ und $() \oplus () = ()$.

(*v*) Es sei $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Dann ist die Menge[4]

$$K^n \coloneqq \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \, x_i \in K \text{ für } i = 1, \dots, n \right\}, \tag{11.4}$$

ausgestattet mit der **komponentenweisen Addition** (englisch: componentwise addition) und der **komponentenweisen S-Multiplikation** (englisch: componentwise S-multiplication)

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \coloneqq \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \tag{11.5a}$$

$$\alpha \odot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \coloneqq \begin{pmatrix} \alpha \cdot x_1 \\ \vdots \\ \alpha \cdot x_n \end{pmatrix}, \tag{11.5b}$$

ein K-Vektorraum, genannt der **Vektorraum der Spaltenvektoren** (englisch: column vector space) oder auch der **Standardvektorraum** (englisch: standard vector space) oder der **Koordinatenraum**[5] (englisch: coordinate space) über K der **Dimension** n. Der Nullvektor ist $\begin{pmatrix} 0_K \\ \vdots \\ 0_K \end{pmatrix}$ in K^n.

Es wird sich auch hier als praktisch erweisen, den Fall $n = 0$ zuzulassen. Der Raum K^0 besteht dann nur aus einem Element, dem leeren Spaltenvektor $()$. K^0 ist also ein Nullraum über K. Es gilt $\alpha \odot () = ()$ für alle $\alpha \in K$ und $() \oplus () = ()$.

(*vi*) Es sei $(K, +, \cdot)$ ein Körper und X eine Menge. Dann ist die Menge $K^X = \{f \mid f : X \to K\}$, ausgestattet mit den punktweisen Verknüpfungen

$$\begin{aligned} (f \oplus g)(x) &\coloneqq f(x) + g(x), \\ (\alpha \odot f)(x) &\coloneqq \alpha \cdot f(x), \end{aligned}$$

ein K-Vektorraum.[6] Der Nullvektor ist die **Nullfunktion** (englisch: zero function, constant function zero) $X \ni x \mapsto 0_K \in K$.

[3] Der Begriff der Dimension für beliebige Vektorräume wird in Definition 13.11 eingeführt. Hier dient er zunächst zur Angabe der Anzahl der Einträge in einem Zeilenvektor.

[4] K^n ist hier also nicht als die Menge von n-Tupeln über K zu lesen, siehe Definition 4.8.

[5] Der Begriff **Koordinatenraum** wird später in Satz 19.1 klar werden.

[6] In Beispiel 10.2 hatten wir gesehen, dass K^X i. A. kein Körper ist, sondern nur ein kommutativer Ring mit Eins. Hier zeigt sich nun, dass K^X außerdem ein K-Vektorraum ist.

Insbesondere ist die Menge der Folgen $K^{\mathbb{N}}$ mit Werten in K ein K-Vektorraum, genannt der **Folgenraum** (englisch: sequence space) **über** K.

(*vii*) Allgemeiner sei $(K, +, \cdot)$ ein Körper, $(V, \oplus, \odot)$ ein K-Vektorraum und X eine Menge. Dann ist die Menge $V^X = \{f \mid f\colon X \to V\}$, ausgestattet mit den punktweisen Verknüpfungen

$$(f \oplus g)(x) \coloneqq f(x) \oplus g(x),$$
$$(\alpha \odot f)(x) \coloneqq \alpha \odot f(x),$$

ein K-Vektorraum. Der Nullvektor ist die **Nullfunktion** (englisch: zero function, constant function zero) $X \ni x \mapsto 0_V \in V$.

Insbesondere ist die Menge der Folgen $V^{\mathbb{N}}$ mit Werten in V ein K-Vektorraum, genannt der **Folgenraum über** V. △

Definition 11.4 (kartesisches Produkt von Vektorräumen, vgl. Definition 6.42 des kartesischen Produkts von Mengen[AoC]).
Es seien $(K, +, \cdot)$ ein Körper, I eine Menge, und weiter sei $(V_i, \oplus_i, \odot_i)$ ein K-Vektorraum für jedes $i \in I$.

(*i*) Das **kartesische Produkt** dieser Vektorräume

$$W \coloneqq \bigtimes_{i \in I} V_i \coloneqq \Big\{F\colon I \to \bigcup_{i \in I} V_i \,\Big|\, F(i) \in V_i \text{ für alle } i \in I\Big\}, \tag{11.6}$$

ausgestattet mit der punktweisen Addition $\oplus$ und der punktweisen S-Multiplikation $\odot$

$$\oplus\colon W \times W \to W \quad \text{mit } F \oplus G \text{ definiert durch} \quad (F \oplus G)(i) \coloneqq F(i) \oplus_i G(i),$$
$$\odot\colon K \times W \to W \quad \text{mit } \alpha \odot F \text{ definiert durch} \quad (\alpha \cdot F)(i) \coloneqq \alpha \odot_i F(i),$$

ist ein K-Vektorraum. Dieser wird auch der **Produktraum** (englisch: product space) der Vektorräume V_i, $i \in I$, genannt.

(*ii*) Ist $I = [\![1, n]\!]$ für $n \in \mathbb{N}$, so schreiben wir statt $\bigtimes_{i \in I} V_i$ auch $\bigtimes_{i=1}^n V_i$ oder $V_1 \times \cdots \times V_n$ und identifizieren die Elemente $F\colon [\![1, n]\!] \to \bigcup_{i=1}^n V_i$ mit den n-Tupeln $(F(1), \ldots, F(n))$, vgl. Bemerkung 6.43.

(*iii*) Wenn alle Vektorräume $(V_i, \oplus_i, \odot_i) = (V, \oplus, \odot)$ sind, so schreiben wir statt $\bigtimes_{i \in I} V$ auch V^I. Ist zusätzlich $I = [\![1, n]\!]$, so schreiben wir auch $\bigtimes_{i=1}^n V$ oder V^n.

(*iv*) Im Fall $I = \emptyset$ besteht das kartesische Produkt (11.6) aus dem einzigen Element $F\colon \emptyset \to \emptyset$, also der leeren Funktion bzw. dem leeren Tupel $()$. In diesem Fall ist $W = \bigtimes_{i \in I} V_i$ also ein Nullraum über K. △

Lemma 11.5 (Rechenregeln in Vektorräumen, vgl. Rechenregeln in Ringen (Lemma 9.3)).
Es sei $(K, +, \cdot)$ ein Körper und $(V, \oplus, \odot)$ ein K-Vektorraum. Dann gilt für $\alpha \in K$ und $v \in V$:

(*i*) $0_K \odot v = 0_V$.

(*ii*) $\alpha \odot 0_V = 0_V$.

(*iii*) $\alpha \odot (\ominus v) = \ominus(\alpha \odot v) = (-\alpha) \odot v$ und insbesondere $\ominus v = (-1_K) \odot v$.

(*iv*) $(-\alpha) \odot (\ominus v) = \alpha \odot v$.

(*v*) $\alpha \odot v = 0_V \quad \Rightarrow \quad \alpha = 0_K$ oder $v = 0_V$.

Beweis. Aussage (*i*): Wir haben

$$\begin{aligned} 0_V \oplus 0_K \odot v &= 0_K \odot v && \text{da } 0_V \text{ das neutrale Element von } (V, \oplus) \text{ ist} \\ &= (0_K + 0_K) \odot v && \text{da } 0_K \text{ das neutrale Element von } (K, +) \text{ ist} \\ &= 0_K \odot v \oplus 0_K \odot v && \text{nach Distributivgesetz (11.1c).} \end{aligned}$$

Aus der Kürzungsregel (7.8b) in der Gruppe $(V, \oplus)$ folgt nun $0_V = 0_K \odot v$.

Aussage (*ii*): Für beliebiges $\alpha \in K$ gilt

$$\begin{aligned} \alpha \odot 0_V \oplus 0_V &= \alpha \odot 0_V && \text{da } 0_V \text{ das neutrale Element von } (V, \oplus) \text{ ist} \\ &= \alpha \odot (0_V \oplus 0_V) && \text{da } 0_V \text{ das neutrale Element von } (V, \oplus) \text{ ist} \\ &= \alpha \odot 0_V \oplus \alpha \odot 0_V && \text{nach Distributivgesetz (11.1b).} \end{aligned}$$

Aus der Kürzungsregel (7.8a) in der Gruppe $(V, \oplus)$ folgt nun $0_V = \alpha \odot 0_V$.

Aussage (*iii*): Es seien $\alpha \in K$ und $v \in V$. Wir zeigen zunächst, dass $\alpha \odot v$ das Inverse zu $(-\alpha) \odot v$ in der Gruppe $(V, \oplus)$ ist:

$$\begin{aligned} \big((-\alpha) \odot v\big) \oplus \big(\alpha \odot v\big) &= (-\alpha + \alpha) \odot v && \text{nach Distributivgesetz (11.1c)} \\ &= 0_K \odot v && \text{da } \alpha \text{ in der Gruppe } (K, +) \text{ das Inverse } -\alpha \text{ besitzt} \\ &= 0_V && \text{nach Aussage } (i). \end{aligned}$$

Das heißt, es gilt $\ominus(\alpha \odot v) = (-\alpha) \odot v$. Insbesondere für $\alpha = 1_K$ erhalten wir

$$\ominus v = \ominus(1_K \odot v) = (-1_K) \odot v. \tag{11.7}$$

Weiter haben wir

$$\begin{aligned} (-\alpha) \odot v &= \big((-1_K) \cdot \alpha\big) \odot v && \text{nach Lemma 9.3} \\ &= \big(\alpha \cdot (-1_K)\big) \odot v && \text{da } (K, \cdot) \text{ kommutativ ist} \\ &= \alpha \odot \big((-1_K) \odot v\big) && \text{nach Assoziativgesetz (11.1a)} \\ &= \alpha \odot (\ominus v) && \text{nach (11.7).} \end{aligned}$$

Aussage (*iv*): Aus Aussage (*iii*) mit $-\alpha$ statt α sowie $-(-\alpha) = \alpha$ folgt sofort

$$(-\alpha) \odot (\ominus v) = (-(-\alpha)) \odot v = \alpha \odot v.$$

Aussage (*v*): Es seien $\alpha \in K$ und $v \in V$ sowie $\alpha \odot v = 0_V$. Nehmen wir $\alpha \neq 0_K$ an, dann gilt

$$\begin{aligned} v &= 1_K \odot v && \text{nach (11.1d)} \\ &= (\alpha^{-1} \cdot \alpha) \odot v && \text{da } \alpha \neq 0_K \text{ in der Gruppe } (K \setminus \{0\}, \cdot) \text{ das Inverse } \alpha^{-1} \text{ besitzt} \\ &= \alpha^{-1} \odot (\alpha \odot v) && \text{nach Assoziativgesetz (11.1a)} \\ &= \alpha^{-1} \odot 0_V && \text{nach Voraussetzung} \\ &= 0_V && \text{nach Aussage } (ii). \end{aligned}$$

□

Bemerkung 11.6 (Mengen und Familien in Vektorräumen).
Die nun folgenden Definitionen und Resultate bis zum Ende von § 14 (Ende von Kapitel 3) geben wir jeweils in zwei Versionen an: für Mengen und für Familien. Beispielsweise geht es in der folgenden Definition 11.7 um Linearkombinationen einer Menge bzw. einer Familie von Vektoren. Die Aussagen sind jeweils konzeptionell gleich, unterscheiden sich jedoch im Detail, vgl. auch Bemerkung 6.40 zu den Unterschieden von Mengen und Familien.

Die Unterscheidung zwischen Mengen und Familien hätten wir beispielsweise auch bereits bei erzeugten Untergruppen (Definition 7.48) und bei erzeugten Idealen (Definition 9.35) machen können. Dort hatten wir nur Mengen als Erzeuger betrachtet. Der Grund, diese Unterscheidung nun bei Vektorräumen zu treffen, liegt darin, dass wir später vor allem bei der Darstellung linearer Abbildungen in Form von Matrizen von der Ordnung auf der Indexmenge einer (endlichen) Familie profitieren werden. △

Definition 11.7 (Linearkombination).
Es sei $(K, +, \cdot)$ ein Körper und $(V, \oplus, \odot)$ ein K-Vektorraum.

Es sei $E \subseteq V$ eine Menge von Vektoren in V. Ist $E_0 \subseteq E$ eine endliche Teilmenge und sind $\alpha_v \in K$, dann heißt der Vektor[7]

$$\sum_{v \in E_0} \alpha_v \odot v \tag{11.8}$$

eine **Linearkombination**[8] **der Menge** E oder eine **Linearkombination von Vektoren der Menge** E.

Es sei $F = (v_i)_{i \in I}$ eine Familie von Vektoren in V. Ist $F_0 = (v_i)_{i \in I_0}$ eine endliche Teilfamilie und sind $\alpha_i \in K$, dann heißt der Vektor

$$\sum_{i \in I_0} \alpha_i \odot v_i \tag{11.9}$$

eine **Linearkombination der Familie** F oder eine **Linearkombination von Vektoren der Familie** F.

Die Skalare α_v bzw. α_i heißen die **Koeffizienten** (englisch: coefficients) der Linearkombination. Eine Linearkombination heißt **trivial** (englisch: trivial linear combination), wenn alle Koeffizienten gleich 0_K sind.

Im Fall $E_0 = \emptyset$ bzw. $I_0 = \emptyset$ interpretieren wir wie üblich die Addition von null Elementen in (11.8) bzw. (11.9) als das neutrale Element, also als den Nullvektor 0_V. △

Beachte: Eine Linearkombination besteht immer aus endlich vielen Termen.

Bemerkung 11.8 (Linearkombination).
In der Literatur werden Linearkombinationen anstelle von (11.8) bzw. (11.9) oft in der Form

$$\alpha_1 \odot v_1 \oplus \cdots \oplus \alpha_n \odot v_n \qquad \text{oder kurz} \quad \sum_{j=1}^{n} \alpha_j \odot v_j \tag{11.10}$$

mit endlich vielen Skalaren $\alpha_j \in K$ und Vektoren $v_j \in E$

$$\alpha_1 \odot v_{i_1} \oplus \cdots \oplus \alpha_n \odot v_{i_n} \qquad \text{oder kurz} \quad \sum_{j=1}^{n} \alpha_j \odot v_{i_j} \tag{11.11}$$

mit endlich vielen Indizes $i_j \in I$, Skalaren $\alpha_j \in K$ und Vektoren v_{i_j} aus F

für $j = 1, \dots, n$ und $n \in \mathbb{N}_0$ geschrieben. Diese Notation hat aber einige Nachteile. Insbesondere ist es in (11.10) möglich, dass Vektoren aus E mehrfach gewählt werden, was die Formulierung einiger Resultate erschwert. In (11.11) können zudem Indizes aus I mehrfach vorkommen. (**Quizfrage 11.2:** Warum ist jede Linearkombination im Sinne von (11.8) eine Linearkombination im Sinne von (11.10) und umgekehrt? Und warum ist jede Linearkombination im Sinne von (11.9) auch eine Linearkombination im Sinne von (11.11) und umgekehrt?) △

Beispiel 11.9 (Linearkombination).

(i) Der Vektor $\binom{3}{-7}$ ist eine Linearkombination der Menge von Vektoren $\left\{\binom{1}{0}, \binom{0}{1}\right\}$ im Standardvektorraum $\mathbb{R}^2$, nämlich

$$\begin{pmatrix} 3 \\ -7 \end{pmatrix} = 3 \odot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus (-7) \odot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Die Koeffizienten 3 und -7 sind hier offensichtlich.

Derselbe Vektor $\binom{3}{-7}$ kann auch als

$$\begin{pmatrix} 3 \\ -7 \end{pmatrix} = 8 \odot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus (-7) \odot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \oplus (-5) \odot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

geschrieben werden, also als Linearkombination der Form (11.10).

(ii) Der Vektor $\binom{3}{-7}$ ist eine Linearkombination der Menge von Vektoren $\left\{\binom{2}{-1}, \binom{4}{1}\right\}$ im Standardvektorraum $\mathbb{R}^2$, nämlich

$$\begin{pmatrix} 3 \\ -7 \end{pmatrix} = \frac{31}{6} \odot \begin{pmatrix} 2 \\ -1 \end{pmatrix} \oplus \frac{-11}{6} \odot \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Die Koeffizienten sind hier nicht offensichtlich. Sie können aber durch Lösen eines linearen Gleichungssystems bestimmt werden, siehe § 16.

(iii) Die Funktion $[0, 2\pi] \ni x \mapsto \sin(x) \ominus \sqrt{2} \odot \cos(x) \in \mathbb{R}$ ist eine Linearkombination der Menge von Vektoren (Funktionen) $\{\sin, \cos\}$ im $\mathbb{R}$-Vektorraum der Funktionen $\mathbb{R}^{[0,2\pi]}$. △

Definition 11.10 (Unterraum, vgl. Definition 10.6 eines Unterkörpers).
Es sei $(K, +, \cdot)$ ein Körper und $(V, \oplus, \odot)$ ein K-Vektorraum.

(i) Eine Teilmenge $U \subseteq V$ heißt ein **Untervektorraum** oder kurz: ein **(linearer) Unterraum** (englisch: vector subspace, linear subspace) von $(V, \oplus, \odot)$, wenn U bzgl. $\oplus$ und bzgl. der S-Multiplikation $\odot$ mit Elementen in K abgeschlossen und wenn $(U, \oplus, \odot)$ selbst wieder ein K-Vektorraum ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, \oplus)$ eine Untergruppe von $(V, \oplus)$ und wenn U bzgl. der S-Multiplikation $\odot$ mit Elementen in K abgeschlossen ist.

(ii) Ein Unterraum $(U, \oplus, \odot)$ von $(V, \oplus, \odot)$ heißt **echt** (englisch: proper subspace), wenn $U \subsetneq V$ gilt. △

[7] Wir verwenden das Summenzeichen also auch für die Addition $\oplus$ im Vektorraum $(V, \oplus, \odot)$.
[8] englisch: linear combination

Beachte: Da $(U, \oplus)$ eine Untergruppe von $(V, \oplus)$ ist, enthält ein Unterraum U immer den Nullvektor 0_V.

Die Prüfung einer Teilmenge $U \subseteq V$ auf die Unterraum-Eigenschaft lässt sich mit folgendem Kriterium erreichen:

Satz 11.11 (Unterraumkriterium, vgl. Unterkörperkriterium Satz 10.7)**.**
Es sei $(K, +, \cdot)$ ein Körper, $(V, \oplus, \odot)$ ein K-Vektorraum und $U \subseteq V$. Dann sind äquivalent:

(i) $(U, \oplus, \odot)$ ist ein Unterraum von $(V, \oplus, \odot)$.

(ii) $U \neq \emptyset$, und für alle $u, v \in U$ und $\alpha \in K$ gilt $u \oplus v \in U$ und $\alpha \odot u \in U$.[9]

(iii) $U \neq \emptyset$, und für alle $u, v \in U$ und $\alpha, \beta \in K$ gilt $\alpha \odot u \oplus \beta \odot v \in U$.[10]

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $(U, \oplus, \odot)$ ein Unterraum von $(V, \oplus, \odot)$. Dann ist per Definition 11.10 und Definition 11.1 $(U, \oplus)$ eine Gruppe, also eine Untergruppe von $(V, \oplus)$. Insbesondere gilt $0_V \in U$, also $U \neq \emptyset$. Weiter ist U als Unterraum abgeschlossen bzgl. $\oplus$ und bzgl. der skalaren Multiplikation $\odot$ mit Elementen von K.

Aussage (ii) $\Rightarrow$ Aussage (i): Wir müssen zeigen, dass $(U, \oplus, \odot)$ unter der Annahme von Aussage (ii) wieder ein Vektorraum ist. Diese Annahme zeigt insbesondere, dass $\oplus \colon U \times U \to U$ und $\odot \colon K \times U \to U$ eingeschränkt werden können. Die Eigenschaften aus (11.1) bleiben bei dieser Einschränkung erhalten. Es bleibt also nur zu zeigen, dass $(U, \oplus)$ eine abelsche Gruppe ist, also eine Untergruppe von $(V, \oplus)$. Dazu wenden wir das Untergruppenkriterium (Satz 7.44) an. Es gilt nach Voraussetzung $U \neq \emptyset$. Wegen $\ominus u = (-1_K) \odot u$ für $u \in U$ und der Annahme $K \odot U \subseteq U$ gilt $\ominus U \subseteq U$. Aus der Annahme $U \oplus U \subseteq U$ folgt daher weiter $U \oplus (\ominus U) \subseteq U$. Nach dem Untergruppenkriterium ist $(U, \oplus)$ damit eine Untergruppe von $(V, \oplus)$.

Aussage (ii) $\Rightarrow$ Aussage (iii): Wir haben $K \odot U \subseteq U$ nach Voraussetzung, also auch $K \odot U \oplus K \odot U \subseteq U \oplus U \subseteq U$, wobei die letzte Inklusion wiederum nach Voraussetzung gilt.

Aussage (iii) $\Rightarrow$ Aussage (ii): Es gilt nach Voraussetzung $U \oplus U \subseteq K \odot U \oplus K \odot U \subseteq U$ und außerdem $K \odot U = K \odot U \oplus \{0_K\} \odot U \subseteq K \odot U \oplus K \odot U \subseteq U$. □

Beispiel 11.12 (Unterräume)**.**

(i) Es sei $(V, \oplus, \odot)$ ein Vektorraum. Dann sind $(\{0_V\}, \oplus, \odot)$ und $(V, \oplus, \odot)$ Unterräume von $(V, \oplus, \odot)$. Diese heißen die **trivialen Unterräume** (englisch: trivial subspaces). Speziell $(\{0_V\}, \oplus, \odot)$ heißt der **Nullunterraum** (englisch: zero vector subspace) von V.

(ii) Wir betrachten den Standardvektorraum $V = \mathbb{R}^2$ über dem Körper $\mathbb{R}$.

(a) Die Menge

$$U_1 \coloneqq \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1 - 2\,x_2 = 0 \right\}$$

[9] kurz: $U \oplus U \subseteq U$ und $K \odot U \subseteq U$
[10] kurz: $K \odot U \oplus K \odot U \subseteq U$

ist ein Unterraum von $V = \mathbb{R}^2$, denn die Aussage (ii) des Unterraumkriteriums Satz 11.11 ist erfüllt: Zunächst gilt $\binom{0}{0} \in U_1$, also ist $U_1 \neq \emptyset$. Weiter gilt für alle $\alpha \in \mathbb{R}$ und $u = \binom{x_1}{x_2} \in U_1$ sowie $v = \binom{y_1}{y_2} \in U_1$:

$$\begin{aligned} u \oplus v &= \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \in U_1, \quad \text{denn } (x_1 + y_1) - 2\,(x_2 + y_2) = (x_1 - 2\,x_2) + (y_1 - 2\,y_2) = 0 \\ \alpha \odot u &= \begin{pmatrix} \alpha\, x_1 \\ \alpha\, x_2 \end{pmatrix} \in U_1, \quad \text{denn } \alpha\, x_1 - 2\,(\alpha\, x_2) = \alpha\,(x_1 - 2\,x_2) = 0. \end{aligned}$$

(b) Die Menge

$$U_2 \coloneqq \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1 - 2\,x_2 = 1 \right\}$$

ist kein Unterraum von $V = \mathbb{R}^2$, denn sie enthält nicht den Nullvektor $\binom{0}{0}$. Der Nullvektor ist aber notwendigerweise in jedem Unterraum enthalten, wie wir im Anschluss an Definition 11.10 bereits gesehen haben.

(c) Die Menge

$$U_3 \coloneqq \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1 \geqslant 0,\ x_2 \geqslant 0 \right\}$$

ist kein Unterraum von $V = \mathbb{R}^2$, denn sie erfüllt das Unterraumkriterium nicht. Beispielsweise ist $\binom{1}{1} \in U_3$, aber $(-1) \odot \binom{1}{1} = \binom{-1}{-1}$ nicht.

(iii) Die Menge

$$U \coloneqq \left\{ a + b\,i \in \mathbb{C} \,\middle|\, a, b \in \mathbb{R},\ a - 2\,b = 0 \right\}$$

ist ein Unterraum des $\mathbb{R}$-Vektorraumes $\mathbb{C}$, jedoch kein Unterraum des $\mathbb{C}$-Vektorraumes $\mathbb{C}$.

(iv) Es sei $(K, +, \cdot)$ ein Körper und $(K^{\mathbb{N}}, \oplus, \odot)$ der Folgenraum über K. Der **Träger** (englisch: support) einer Folge $(x_n)_{n\in\mathbb{N}}$ ist die Menge derjenigen Indizes, für die $x_n \neq 0_K$ ist, also

$$\operatorname{supp}(x_n)_{n\in\mathbb{N}} = \{n \in \mathbb{N} \mid x_n \neq 0_K\} \subseteq \mathbb{N}. \tag{11.12}$$

Die Teilmenge der **Folgen mit endlichem Träger** (englisch: sequences with finite support) oder **endlich getragenen Folgen** (englisch: finitely supported sequences)

$$(K^{\mathbb{N}})_{00} \coloneqq \left\{ (x_n)_{n\in\mathbb{N}} \,\middle|\, \operatorname{supp}(x_n)_{n\in\mathbb{N}} \text{ ist endlich} \right\} \tag{11.13}$$

ist ein echter Unterraum von $(K^{\mathbb{N}}, \oplus, \odot)$. (**Quizfrage 11.3:** Kann man den Begriff der endlich getragenen Folgen auf Folgen über einem K-Vektorraum V ausdehnen?) △

Expertenwissen: noch mehr Unterräume des Folgenraumes über $\mathbb{R}$

In Lehrveranstaltungen zur *Analysis* werden die Folgenräume

$$\begin{aligned} (\mathbb{R}^{\mathbb{N}})_b &\coloneqq \left\{ (y_n)_{n\in\mathbb{N}} \,\middle|\, \exists C \geqslant 0 \text{ für alle } n \in \mathbb{N}\ (|y_n| \leqslant C) \right\} && \text{der beschränkten Folgen[11]} \\ (\mathbb{R}^{\mathbb{N}})_c &\coloneqq \left\{ (y_n)_{n\in\mathbb{N}} \,\middle|\, (y_n)_{n\in\mathbb{N}} \text{ ist konvergent} \right\} && \text{der konvergenten Folgen[12]} \\ (\mathbb{R}^{\mathbb{N}})_0 &\coloneqq \left\{ (y_n)_{n\in\mathbb{N}} \,\middle|\, (y_n)_{n\in\mathbb{N}} \text{ konvergiert gegen } 0 \right\} && \text{der Nullfolgen[13]} \end{aligned}$$

eingeführt. Es gilt

$$(\mathbb{R}^\mathbb{N})_{00} \subsetneq (\mathbb{R}^\mathbb{N})_0 \subsetneq (\mathbb{R}^\mathbb{N})_c \subsetneq (\mathbb{R}^\mathbb{N})_b \subsetneq \mathbb{R}^\mathbb{N}$$

im Sinne von Teilmengen und von Unterräumen. Diese Definitionen und Aussagen können auf normierte Vektorräume über $\mathbb{R}$ verallgemeinert werden.

Ende der Vorlesung 15

Bemerkung 11.13 (Vereinfachung der Notation).

(i) Zur Vereinfachung der Notation werden wir in Zukunft für die Verknüpfungen $\oplus$ und $\odot$ in K-Vektorräumen V einfach dieselbe Notation verwenden wie für die Verknüpfungen $+$ und $\cdot$ des zugrundeliegenden Körpers K. Aus den verknüpften Objekten ist entweder ersichtlich, ob die jeweilige Verknüpfung mit Skalaren oder mit Vektoren gemeint ist, oder aber (im Fall $K \subseteq V$) ist die Unterscheidung wegen (11.1d) unerheblich.

(ii) Das Zeichen für die Multiplikation $\cdot$ von Vektoren mit Elementen aus dem zugrundeliegenden Körper (oder von zwei Körperelementen) wird sogar oft ganz weggelassen, sofern sich dadurch keine Unklarheiten ergeben. Beispielsweise schreiben wir in Zukunft einfach

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{anstelle von} \quad \begin{pmatrix} 3 \\ 1 \end{pmatrix} \ominus 2 \odot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

im Vektorraum $\mathbb{R}^2$ oder

$$\sin - \sqrt{2} \cos \quad \text{anstelle von} \quad \sin \ominus \sqrt{2} \odot \cos$$

im Vektorraum $\mathbb{R}^{[0,2\pi]}$. Wie in Bemerkung 11.2 erklärt, müssen die Ausdrücke $n\,v$ und $n \odot v \coloneqq (n\,1_K) \odot v$ für $n \in \mathbb{Z}$ ohnehin nicht unterschieden werden.

(iii) Wir werden außerdem das Nullelement 0_K des Skalarkörpers K in Zukunft einfach als 0 notieren, das Einselement 1_K von K als 1 und den Nullvektor von V ebenfalls als 0.

(iv) Schließlich werden wir in Zukunft auch von Vektorräumen sprechen, ohne den zugrundeliegenden Körper explizit zu erwähnen, wenn dieser für die Formulierung des jeweiligen Resultats nicht relevant ist. Dennoch besitzt natürlich jeder Vektorraum immer einen konkreten zugrundeliegenden Körper. △

Wie bereits bei Untergruppen in § 7.4 sowie bei Idealen in § 9.1 beschäftigt uns nun die Frage, wie man Unterräume erzeugen kann.

Lemma 11.14 (Durchschnitt von Unterräumen, vgl. Lemma 10.10 zu Unterkörpern).
Es sei V ein Vektorraum.

(i) Ist $(U_i)_I$ eine nichtleere Familie von Unterräumen von $(V, +, \cdot)$, dann ist auch $\bigcap_{i \in I} U_i$ ein Unterraum von $(V, +, \cdot)$.

[11] englisch: bounded sequences
[12] englisch: convergent sequences
[13] englisch: null sequences

(*ii*) Ist $\mathcal{U}$ eine nichtleere Menge von Unterräumen von $(V,+,\cdot)$, dann ist auch $\bigcap \mathcal{U}$ ein Unterraum von $(V,+,\cdot)$.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

Definition 11.15 (erzeugter Unterraum, lineare Hülle, Spann, Erzeugendensystem, erzeugende Familie)**.**
Es sei V ein Vektorraum.

(*i*) Ist $E \subseteq V$ eine Teilmenge von Vektoren in V, dann heißt

$$\langle E \rangle \coloneqq \bigcap \left\{ U \,\middle|\, \begin{array}{l} U \text{ ist Unterraum von } V \\ \text{und } E \subseteq U \end{array} \right\} \tag{11.14}$$

der von E **erzeugte** oder **aufgespannte Unterraum**[14], die **lineare Hülle**[15] $\operatorname{Lin}(E)$ oder auch der **Spann**[16] $\operatorname{Span}(E)$ von E in V.

Ist speziell E die endliche Menge $E = \{v_1, \ldots, v_n\}$ für $v_i \in V$ und $n \in \mathbb{N}_0$, so schreiben wir auch $\langle v_1, \ldots, v_n \rangle$ oder $\operatorname{Lin}(v_1, \ldots, v_n)$ oder $\operatorname{Span}(v_1, \ldots, v_n)$ statt $\langle \{v_1, \ldots, v_n\} \rangle$ oder $\operatorname{Lin}(\{v_1, \ldots, v_n\})$ oder $\operatorname{Span}(\{v_1, \ldots, v_n\})$.

(*ii*) Gilt $\langle E \rangle = V$, dann heißt E eine **erzeugende Menge**[17] oder ein **Erzeugendensystem**[18] von V.

(*iii*) Falls eine endliche erzeugende Menge von V existiert, so heißt V **endlich erzeugt**[19].

(*i*) Ist $F = (v_i)_{i \in I}$ eine Familie von Vektoren in V ist, dann heißt

$$\langle F \rangle \coloneqq \bigcap \left\{ U \,\middle|\, \begin{array}{l} U \text{ ist Unterraum von } V \\ \text{und } \{v_i \mid i \in I\} \subseteq U \end{array} \right\} \tag{11.15}$$

der von F **erzeugte** oder **aufgespannte Unterraum**, die **lineare Hülle** $\operatorname{Lin}(F)$ oder auch der **Spann** $\operatorname{Span}(F)$ von F in V.

Ist speziell F die endliche Familie $F = (v_1, \ldots, v_n)$ für $v_i \in V$ und $n \in \mathbb{N}_0$, so schreiben wir auch $\langle v_1, \ldots, v_n \rangle$ oder $\operatorname{Lin}(v_1, \ldots, v_n)$ oder $\operatorname{Span}(v_1, \ldots, v_n)$ statt $\langle (v_1, \ldots, v_n) \rangle$ oder $\operatorname{Lin}((v_1, \ldots, v_n))$ oder $\operatorname{Span}((v_1, \ldots, v_n))$.

(*ii*) Gilt $\langle F \rangle = V$, dann heißt F eine **erzeugende Familie**[20] oder ein **Erzeugendensystem** von V.

(*iii*) Falls eine endliche erzeugende Familie von V existiert, so heißt V **endlich erzeugt**. △

Beachte: Bezeichnen wir mit $\mathcal{R}$ die Menge auf rechten Seite von (11.14), über die der Durchschnitt gebildet wird, dann ist $\langle E \rangle$ das Minimum der Menge $\mathcal{R}$ bzgl. der Inklusionshalbordnung und auch das Minimum der Menge $\mathcal{R}$ bzgl. der Halbordnung „ist Unterraum von" auf der Menge aller Unterräume von V.

Quizfrage 11.4: An der verkürzten Schreibweise $\langle v_1, \ldots, v_n \rangle$ kann man nicht mehr erkennen, ob $\{v_1, \ldots, v_n\}$ als Menge oder $(v_1, \ldots, v_n)$ als Familie gemeint war. Warum ist das kein Problem,

[14] englisch: subspace generated by E, subspace spanned by E
[15] englisch: linear hull
[16] englisch: span
[17] englisch: generating set
[18] englisch: generator
[19] englisch: finitely generated
[20] englisch: generating family

auch wenn Vektoren mehrfach vorkommen?

Quizfrage 11.5: Ist der Begriff eines endlich erzeugten Vektorraumes wohldefiniert, unabhängig davon, ob man mit erzeugenden Mengen oder mit erzeugenden Familien arbeitet?

Satz 11.16 (Darstellung des erzeugten Unterraumes, vgl. Satz 9.36)**.**
Es sei $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein K-Vektorraum.

Ist $E \subseteq V$ eine Teilmenge von Vektoren in V, dann gilt für den von E erzeugten Unterraum:

$$\langle E \rangle = \left\{ \sum_{v \in E_0} \alpha_v \, v \;\middle|\; \begin{matrix} E_0 \subseteq E \text{ endlich,} \\ \forall v \in E_0 \; (\alpha_v \in K) \end{matrix} \right\}. \tag{11.16}$$

Ist $F = (v_i)_{i \in I}$ eine Familie von Vektoren in V, dann gilt für den von F erzeugten Unterraum:

$$\langle F \rangle = \left\{ \sum_{i \in I_0} \alpha_i \, v_i \;\middle|\; \begin{matrix} I_0 \subseteq I \text{ endlich,} \\ \forall i \in I_0 \; (\alpha_i \in K) \end{matrix} \right\}. \tag{11.17}$$

Beachte: Der von einer Menge E (oder einer Familie F) erzeugte Unterraum ist also nicht anderes als die Menge der Linearkombinationen von E (bzw. F).[21] Zur Erinnerung: Im Fall $E_0 = \emptyset$ in (11.16) bzw. $I_0 = \emptyset$ in (11.17) interpretieren wir die Linearkombination von null Elementen als den Nullvektor 0. Insbesondere im Fall $E = \emptyset$ ist also $\langle E \rangle = \langle \emptyset \rangle = \{0\}$ der Nullunterraum von V, und auch im Fall der leeren Familie $F = ()$ gilt $\langle F \rangle = \langle () \rangle = \{0\}$.

Beweis. Wir führen den Beweis nur für Mengen E und die Darstellung (11.16). Der Beweis für den Fall (11.17) geht ähnlich. Zur Abkürzung bezeichnen wir die Menge auf der rechten Seite von (11.16) mit M. Wir führen den Beweis analog zu Satz 7.50 (Darstellung der erzeugten Untergruppe) in zwei Schritten.

Schritt 1: $\langle E \rangle \supseteq M$: Es sei U ein beliebiger Unterraum von V, der im Durchschnitt (11.14) vorkommt. U enthält also E als Teilmenge. Da U ein Unterraum ist, enthält U auch alle Linearkombinationen von E. Also gilt $U \supseteq M$. Da dies für jeden beliebigen Unterraum aus dem Durchschnitt in (11.14) gilt, gilt auch $\langle E \rangle \supseteq M$.

Schritt 2: $\langle E \rangle \subseteq M$: Wir zeigen zunächst, dass M selbst ein Unterraum von V ist. Dazu überprüfen wir das Unterraumkriterium (Satz 11.11). Offensichtlich ist $M \neq \emptyset$, denn M enthält mindestens den Nullvektor 0. Sind $\sum_{v \in E_0} \alpha_v \, v$ und $\sum_{v \in E_1} \beta_v \, v$ zwei Elemente aus M, so ist auch $\sum_{v \in E_0} \alpha_v \, v + \sum_{v \in E_1} \beta_v \, v$ ein Element aus M. (**Quizfrage 11.6:** Klar?) Zudem ist für jedes $\alpha \in K$ auch $\alpha \sum_{v \in E_0} \alpha_v \, v = \sum_{v \in E_0} \alpha \, \alpha_v \, v$ ein Element aus M. Also ist M ein Unterraum von V. Zusätzlich ist klar, dass $E \subseteq M$ gilt. (**Quizfrage 11.7:** Warum?) Das heißt, M ist einer derjenigen Unterräume von V, über die in der Definition von $\langle E \rangle$ der Durchschnitt gebildet wird. Folglich gilt $\langle E \rangle \subseteq M$. ◻

Beispiel 11.17 (lineare Hülle)**.**

(i) Es seien K ein Körper und $K^{\mathbb{N}}$ der Folgenraum über K, siehe Beispiel 11.12. Ein Element von $K^{\mathbb{N}}$, also eine Folge $(y_n)_{n \in \mathbb{N}}$ mit Werten in K, heißt die j-te **Standardfolge** (englisch: standard sequence) für $j \in \mathbb{N}$, wenn $y_j = 1$ und $y_n = 0$ für alle $n \neq j$ gilt. Wir bezeichnen die j-te Standardfolge auch mit dem Symbol e_j.[22]

[21] In manchen Büchern wird daher auch (11.16) als Definition des erzeugten Unterraumes verwendet.

[22] Sollten wir das n-te Glied der Folge e_j bezeichnen müssen, so würden wir einen Doppelindex verwenden: $e_{j,n}$.

Die Menge E aller Standardfolgen erzeugt den Unterraum der endlich getragenen Folgen, also $\langle E \rangle = (K^{\mathbb{N}})_{00}$, siehe Beispiel 11.12. (**Quizfrage 11.8:** Warum erzeugen die Standardfolgen nicht den ganzen Folgenraum $K^{\mathbb{N}}$?)

Die Menge $\{e_1, \ldots, e_n\}$, bestehend aus der ersten bis zur n-ten Standardfolge, $n \in \mathbb{N}_0$, erzeugt den Unterraum derjenigen Folgen, deren Träger in $[\![1, n]\!]$ liegt, die also nur an den ersten n Stellen von 0 verschieden sein dürfen.

(ii) In einem Vektorraum V heißt die lineare Hülle

$$\langle v \rangle = \{\alpha\, v \mid \alpha \in K\}$$

eines einzelnen Vektors $v \neq 0$ eine **Gerade** (englisch: line) durch 0 und v. Die lineare Hülle

$$\langle v, w \rangle = \{\alpha\, v + \beta\, w \mid \alpha, \beta \in K\}$$

von zwei Vektoren $v, w \neq 0$ mit $w \notin \langle v \rangle$ heißt eine **Ebene** (englisch: plane) durch 0, v und w. (**Quizfrage 11.9:** Was passiert im Fall $w \in \langle v \rangle$?) △

Folgerung 11.18 (zu Satz 11.16: lineare Hülle von Vereinigung und Schnitt, vgl. Folgerung 7.53)**.** Es sei V ein Vektorraum.

Sind $E_1, E_2 \subseteq V$ Teilmengen von Vektoren in V, dann gilt

$$\langle E_1 \cup E_2 \rangle = \big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle \tag{11.18a}$$

$$\begin{aligned} \langle E_1 \cap E_2 \rangle &\subseteq \big\langle \langle E_1 \rangle \cap \langle E_2 \rangle \big\rangle \\ &= \langle E_1 \rangle \cap \langle E_2 \rangle. \end{aligned} \tag{11.18b}$$

Sind $F_1, F_2 \subseteq V$ Familien von Vektoren in V, dann gilt

$$\langle F_1 \parallel F_2 \rangle = \big\langle \langle F_1 \rangle \cup \langle F_2 \rangle \big\rangle \tag{11.19a}$$

$$\begin{aligned} \langle F_{12} \rangle &\subseteq \big\langle \langle F_1 \rangle \cap \langle F_2 \rangle \big\rangle \\ &= \langle F_1 \rangle \cap \langle F_2 \rangle. \end{aligned} \tag{11.19b}$$

Dabei ist F_{12} diejenige Teilfamilie von F_1, deren Mitglieder auch Mitglieder von F_2 sind.

Beweis. Wir beweisen nur die Aussagen für Mengen, zunächst (11.18a): Die lineare Hülle ist offenbar ordnungserhaltend, d. h., es gilt $E_1 \subseteq \langle E_1 \rangle$ und $E_2 \subseteq \langle E_2 \rangle$. Daraus folgt $E_1 \cup E_2 \subseteq \langle E_1 \rangle \cup \langle E_2 \rangle$ und durch Bildung der linearen Hülle $\langle E_1 \cup E_2 \rangle \subseteq \big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle$.

Umgekehrt besteht $\langle E_1 \rangle$ nach Satz 11.16 gerade aus den Linearkombinationen von E_1, und $\langle E_2 \rangle$ besteht aus den Linearkombinationen von E_2. Das heißt, $\big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle$ besteht aus Linearkombinationen solcher Vektoren, die ihrerseits eine Linearkombination von E_1 oder eine Linearkombination von E_2 sind. Mit Hilfe des Assoziativgesetzes (11.1a) erhalten wir, dass wir jeden Vektor aus $\big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle$ als Linearkombination von $E_1 \cup E_2$ schreiben können, also folgt $\big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle \subseteq \langle E_1 \cup E_2 \rangle$.

Nun zum Beweis von (11.18b): Wegen $E_1 \subseteq \langle E_1 \rangle$ und $E_2 \subseteq \langle E_2 \rangle$ gilt $E_1 \cap E_2 \subseteq \langle E_1 \rangle \cap \langle E_2 \rangle$. Da $\langle E_1 \rangle \cap \langle E_2 \rangle$ als Durchschnitt zweier Unterräume bereits ein Unterraum ist (Lemma 11.14), folgt $\big\langle \langle E_1 \rangle \cap \langle E_2 \rangle \big\rangle = \langle E_1 \rangle \cap \langle E_2 \rangle$. □

§ 12 Lineare Unabhängigkeit

In diesem Abschnitt geht es um die Begriffe der linearen Unabhängigkeit bzw. linearen Abhängigkeit und ihre Konsequenzen.

Definition 12.1 (lineare (Un-)abhängigkeit)**.**
Es sei V ein Vektorraum.

(i) Eine Menge $E \subseteq V$ von Vektoren in V heißt **linear unabhängig**[23], wenn in jeder Linearkombination von Vektoren aus E, die den Nullvektor ergibt, notwendig alle Koeffizienten gleich 0 sind, wenn also für jede endliche Teilmenge $E_0 \subseteq E$ gilt:

$$\sum_{v \in E_0} \alpha_v \, v = 0 \quad \Rightarrow \quad \alpha_v = 0 \text{ für alle } v \in E_0. \tag{12.1}$$

(ii) Wenn dagegen eine Linearkombination von Vektoren aus E möglich ist, die den Nullvektor ergibt, wobei nicht alle Koeffizienten gleich 0 sind, dann heißt die Menge E **linear abhängig**[24].

(i) Eine Familie F von Vektoren in V heißt **linear unabhängig**, wenn in jeder Linearkombination von Vektoren aus F, die den Nullvektor ergibt, notwendig alle Koeffizienten gleich 0 sind, wenn also für jede endliche Teilmenge $I_0 \subseteq I$ der Indizes gilt:

$$\sum_{i \in I_0} \alpha_i \, v_i = 0 \quad \Rightarrow \quad \alpha_i = 0 \text{ für alle } i \in I_0. \tag{12.2}$$

(ii) Wenn dagegen eine Linearkombination von Vektoren aus F möglich ist, die den Nullvektor ergibt, wobei nicht alle Koeffizienten gleich 0 sind, dann heißt die Familie F **linear abhängig**. △

Erlauben wir Linearkombinationen der Form (11.10), so müssen wir die Definition 12.1 wie folgt anpassen: Eine Menge $E \subseteq V$ von Vektoren in V heißt **linear unabhängig**, wenn in jeder Linearkombination der Form (11.10) von $n \in \mathbb{N}$ paarweise verschiedenen Vektoren aus E, die den Nullvektor ergibt, notwendig alle Koeffizienten gleich 0 sind, wenn also für alle $n \in \mathbb{N}$ gilt:

$$v_1, \ldots, v_n \in E \text{ mit } v_i \neq v_j \text{ für } i \neq j \quad \text{und} \quad \sum_{\ell=1}^{n} \alpha_\ell \, v_\ell = 0$$
$$\Rightarrow \quad \alpha_\ell = 0 \text{ für alle } \ell = 1, \ldots, n.$$

(**Quizfrage 12.1:** Wie müsste die Definition für den Fall von Familien lauten, wenn Linearkombinationen der Form (11.11) zugelassen werden?)

Beispiel 12.2 (lineare (Un-)abhängigkeit)**.**

(i) Die Teilmenge $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} \right\}$ des Vektorraumes $\mathbb{R}^2$ über dem Körper $\mathbb{R}$ ist linear abhängig, denn es gilt

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

[23] englisch: linearly independent
[24] englisch: linearly dependent

(*ii*) Dieselbe Menge ist als Teilmenge des Vektorraumes $\mathbb{R}^2$ über dem Körper $\mathbb{Q}$ jedoch linear unabhängig, denn es gilt

$$\alpha_1 \begin{pmatrix}1\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}1\\\sqrt{2}\end{pmatrix} + \alpha_3 \begin{pmatrix}\sqrt{2}\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \alpha_3 = 0.$$

(*iii*) Wir betrachten den K-Vektorraum K^X der Funktionen $X \to K$, wobei $(K, +, \cdot)$ ein Körper und X eine Menge ist, siehe Beispiel 11.3. Für $y \in X$ definieren wir die **charakteristische Funktion** (englisch: characteristic function) $e_y \colon X \to K$ durch[25]

$$x \mapsto e_y(x) \coloneqq \delta_{xy} \coloneqq \begin{cases} 1, & \text{falls } x = y, \\ 0, & \text{falls } x \neq y. \end{cases} \tag{12.3}$$

Die Menge der charakteristischen Funktionen $\{e_y \mid y \in X\}$ ist linear unabhängig.

(*iv*) Es seien K ein Körper und $K^{\mathbb{N}}$ der Folgenraum über K, siehe Beispiel 11.12. Die Menge der Standardfolgen $\{e_j \mid j \in \mathbb{N}\} \subseteq K^{\mathbb{N}}$, siehe Beispiel 11.17, ist linear unabhängig. △

Bemerkung 12.3 (lineare (Un-)abhängigkeit)**.**

(*i*) Die lineare (Un-)abhängigkeit ist eine Eigenschaft, die sich auf eine Menge oder eine Familie von Vektoren bezieht. Sprechweisen wie „Der Vektor v ist linear unabhängig von $\{v_1, v_2\}$." sind nicht korrekt. Man kann aber beispielsweise sagen: „Die Menge $\{v\} \cup \{v_1, v_2\}$ ist linear unabhängig."

(*ii*) Die Menge $\{v\}$, bestehend aus einem einzigen Vektor, ist genau dann linear unabhängig, wenn $v \neq 0$ ist. Eine analoge Aussage gilt für eine Familie (v) mit nur einem Mitglied.

(*iii*) Eine Menge oder Familie von Vektoren, die den Nullvektor enthält, ist stets linear abhängig.

(*iv*) Die leere Menge und die leere Familie von Vektoren sind per Definition linear unabhängig. △

Lemma 12.4 (lineare (Un-)abhängigkeit von Teilmengen und Obermengen)**.**
Es sei V ein Vektorraum. Für eine Teilmenge $E \subseteq V$ bzw. für eine Familie $F = (v_i)_{i \in I}$ von Vektoren in V gilt:

(*i*) Ist E linear unabhängig, dann auch jede Teilmenge von E.

(*ii*) Ist jede endliche Teilmenge von E linear unabhängig, dann auch E.

(*iii*) Ist E linear abhängig, dann auch jede Obermenge von E.

(*i*) Ist F linear unabhängig, dann auch jede Teilfamilie von F.

(*ii*) Ist jede endliche Teilfamilie von F linear unabhängig, dann auch F.

(*iii*) Ist F linear abhängig, dann auch jede Oberfamilie von F.

Beweis. Der Beweis ergibt sich unmittelbar aus der Definition 12.1. □

[25] Das Symbol δ_{xy} ist das **Kronecker-Delta** (englisch: Kronecker delta). Allgemein ist die **charakteristische Funktion e_A einer Menge** $A \subseteq X$ definiert als $e_A(x) = 1$ für $x \in A$ und $e_A(x) = 0$ für $x \notin A$.

Im Folgenden führen wir einige zur linearen Abhängigkeit bzw. Unabhängigkeit äquivalente Eigenschaften an. Wir beginnen mit der Feststellung, dass die lineare Abhängigkeit einer Menge bzw. einer Familie von Vektoren bedeutet, dass man einen der Vektoren als Linearkombination der anderen darstellen kann:

Lemma 12.5 (lineare Abhängigkeit bedeutet Kombinierbarkeit)**.**
Es sei V ein Vektorraum. Für eine Teilmenge $E \subseteq V$ bzw. für eine Familie $F = (v_i)_{i \in I}$ von Vektoren in V sind äquivalent:

(i) E ist linear abhängig.

(ii) Es gibt einen Vektor $v \in E$, der als Linearkombination der Teilmenge $E \setminus \{v\}$ darstellbar ist.

(i) F ist linear abhängig.

(ii) Es gibt einen Index $i^* \in I$, sodass v_{i^*} als Linearkombination der Teilfamilie $(v_i)_{i \in I \setminus \{i^*\}}$ darstellbar ist.

Beweis. Dieser Beweis ist Gegenstand der Übung. □

Folgerung 12.6 (lineare Abhängigkeit bedeutet Redundanz)**.**
Es sei V ein Vektorraum. Für eine Teilmenge $E \subseteq V$ bzw. für eine Familie $F = (v_i)_{i \in I}$ von Vektoren in V sind äquivalent:

(i) E ist linear abhängig.

(ii) Es gibt ein $v \in E$, sodass gilt:

$$\langle E \setminus \{v\} \rangle = \langle E \rangle.$$

(i) F ist linear abhängig.

(ii) Es gibt einen Index $i^* \in I$, sodass gilt:

$$\langle (v_i)_{i \in I \setminus \{i^*\}} \rangle = \langle (v_i)_{i \in I} \rangle.$$

Beachte: Ist eine erzeugende Menge (bzw. eine erzeugende Familie) eines Vektorraumes linear abhängig, so kann man mindestens ein Element (bzw. ein Mitglied) aus der erzeugenden Menge (bzw. Familie) entfernen, ohne den erzeugten Unterraum zu verkleinern.

Beweis. Wir führen den Beweis nur für Familien durch, der Beweis für Mengen geht analog.

Aussage (i) $\Rightarrow$ Aussage (ii): Da F linear abhängig ist, gibt es nach Lemma 12.5 einen Index i^*, eine endliche Teilmenge $I_0 \subseteq I \setminus \{i^*\}$ und Koeffizienten α_i für $i \in I_0$, sodass

$$v_{i^*} = \sum_{i \in I_0} \alpha_i \, v_i \tag{12.4}$$

gilt. Es sei nun $w \in \langle (v_i)_{i \in I} \rangle$. Nach Satz 11.16 ist w dann eine Linearkombination von $(v_i)_{i \in I}$, d. h., es gibt eine endliche Teilmenge $I_1 \subseteq I$ und Koeffizienten β_i, $i \in I_1$, sodass

$$w = \sum_{i \in I_1} \beta_i \, v_i$$

gilt.

Wir wollen zeigen, dass wir w auch ohne Verwendung von v_{i^*} linearkombinieren können. Falls $i^* \notin I_1$ liegt, dann ist w bereits als Linearkombination von $(v_i)_{i \in I \setminus \{i^*\}}$ dargestellt und nichts zu zeigen. Liegt andererseits $i^* \in I_1$, so ersetzen wir das Vorkommen von v_{i^*} mit Hilfe von (12.4):

$$w = \sum_{i \in I_1} \beta_i \, v_i = \sum_{i \in I_1 \setminus \{i^*\}} \beta_i \, v_i + \beta_{i^*} \, v_{i^*} = \sum_{i \in I_1 \setminus \{i^*\}} \beta_i \, v_i + \sum_{i \in I_0} \beta_{i^*} \alpha_i \, v_i.$$

Also ist w auch bereits eine Linearkombination von $(v_i)_{i \in I \setminus \{i^*\}}$, also $w \in \langle (v_i)_{i \in I \setminus \{i^*\}} \rangle$.

Aussage (ii) $\Rightarrow$ Aussage (i): Der Vektor v_{i^*} ist Mitglied der Familie $(v_i)_{i \in I}$, also gilt erst recht $v_{i^*} \in \langle (v_i)_{i \in I} \rangle$ und nach Voraussetzung $v_{i^*} \in \langle (v_i)_{i \in I \setminus \{i^*\}} \rangle$. Nach Lemma 12.5 ist also $F = (v_i)_{i \in I}$ linear abhängig. □

Das folgende Resultat stellt klar, dass die lineare Unabhängigkeit einer Menge von Vektoren äquivalent dazu ist, dass die Vektoren aus ihrer linearen Hülle eine i. W. eindeutige Darstellung als Linearkombination besitzen:

Lemma 12.7 (lineare Unabhängigkeit und eindeutige Linearkombinationen)**.**
Es sei V ein Vektorraum. Für eine Teilmenge $E \subseteq V$ bzw. für eine Familie $F = (v_i)_{i \in I}$ von Vektoren in V sind äquivalent:

(i) E ist linear unabhängig.

(ii) Jeder Vektor $v \in \langle E \rangle$ lässt sich in eindeutiger Weise (bis auf Summanden mit Nullkoeffizienten) aus Vektoren der Menge E linearkombinieren.

Sind also

$$v = \sum_{v \in E_0} \alpha_v \, v = \sum_{v \in E_1} \beta_v \, v$$

zwei Darstellungen von v mit endlichen Teilmengen $E_0, E_1 \subseteq E$, dann gilt

$$\begin{aligned} \alpha_v &= \beta_v && \text{für } v \in E_0 \cap E_1, \\ \alpha_v &= 0 && \text{für } v \in E_1 \setminus E_0, \\ \beta_v &= 0 && \text{für } v \in E_0 \setminus E_1. \end{aligned}$$

(i) F ist linear unabhängig.

(ii) Jeder Vektor $v \in \langle (v_i)_{i \in I} \rangle$ lässt sich in eindeutiger Weise (bis auf Summanden mit Nullkoeffizienten) aus Vektoren der Familie F linearkombinieren.

Sind also

$$v = \sum_{i \in I_0} \alpha_i \, v_i = \sum_{i \in I_1} \beta_i \, v_i$$

zwei Darstellungen von v mit endlichen Teilmengen $I_0, I_1 \subseteq I$, dann gilt

$$\begin{aligned} \alpha_i &= \beta_i && \text{für } i \in I_0 \cap I_1, \\ \alpha_i &= 0 && \text{für } i \in I_1 \setminus I_0, \\ \beta_i &= 0 && \text{für } i \in I_0 \setminus I_1. \end{aligned}$$

Beweis. Wir beweisen nur die Aussagen für Familien.

Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $(v_i)_{i \in I}$ linear unabhängig und $v \in \langle (v_i)_{i \in I} \rangle$. Nach Satz 11.16 besteht $\langle (v_i)_{i \in I} \rangle$ gerade aus den Linearkombinationen von $(v_i)_{i \in I}$. Es gibt also eine endliche Teilmenge $I_0 \subseteq I$ und Skalare $\alpha_i \in K$, $i \in I_0$ mit der Eigenschaft $v = \sum_{i \in I_0} \alpha_i \, v_i$.

Zu zeigen ist noch, dass diese Darstellung i. W. eindeutig ist. Wir nehmen dazu an, dass

$$v = \sum_{i \in I_0} \alpha_i \, v_i = \sum_{j \in I_1} \beta_j \, v_j$$

zwei Darstellungen von v als Linearkombination von $(v_i)_{i\in I}$ ist mit endlichen Indexmengen I_0 und I_1 sind. Dann ist auch $I_{01} \coloneqq I_0 \cup I_1$ endlich. Wir ergänzen die Koeffizienten durch $\alpha_i \coloneqq 0$ für $i \in I_1 \setminus I_0$ und $\beta_i \coloneqq 0$ für $i \in I_0 \setminus I_1$. Dann haben wir

$$0 = v - v = \sum_{i\in I_{01}} \alpha_i \, v_i - \sum_{i\in I_{01}} \beta_i \, v_i = \sum_{i\in I_{01}} (\alpha_i - \beta_i) \, v_i.$$

Da nach Voraussetzung die Familie $(v_i)_{i\in I}$ linear unabhängig ist, muss $\alpha_i - \beta_i = 0$ für alle $i \in I_{01}$ gelten, also $\alpha_i = \beta_i$. Das zeigt die Behauptung.

Aussage (ii) $\Rightarrow$ Aussage (i): Es sei $I_0 \subseteq I$ eine beliebige endliche Teilmenge. Wir zeigen, dass $(v_i)_{i\in I_0}$ linear unabhängig ist. Wir untersuchen dazu die Linearkombination

$$\sum_{i\in I_0} \alpha_i \, v_i = 0.$$

Diese wird erreicht durch die Wahl von $\alpha_i = 0$ für alle $i \in I_0$. Nach Voraussetzung ist das auch die einzig mögliche Wahl der Koeffizienten. Das heißt aber, dass $(v_i)_{i\in I_0}$ linear unabhängig ist. Da I_0 eine beliebige endliche Teilmenge von I war, ist die gesamte Familie $(v_i)_{i\in I}$ linear unabhängig (Lemma 12.4). □

Quizfrage 12.2: Wie könnten wir die Aussage (ii) von Lemma 12.7 mit Hilfe von Linearkombinationen der Form (11.10) formulieren?

Ende der Vorlesung 16

Ende der Woche 8

§ 13 Basis und Dimension

Literatur: Beutelspacher, 2014, Kapitel 3; Bosch, 2014, Kapitel 1; Fischer, Springborn, 2020, Kapitel 2.5; Jänich, 2008, Kapitel 3

§ 13.1 Basis eines Vektorraumes

In diesem Abschnitt beantworten wir die Frage, wie wir einen Vektorraum durch eine möglichst kleine erzeugende Menge (oder erzeugende Familie) darstellen können und damit Redundanz in der Darstellung vermeiden.

Definition 13.1 (Basis)**.**
Es sei V ein Vektorraum.

Eine Menge $B \subseteq V$ von Vektoren in V heißt eine **Basis**[26] **von** V, wenn B linear unabhängig ist und $\langle B\rangle = V$ gilt.

Eine Familie B von Vektoren aus V heißt eine **Basis von** V, wenn B linear unabhängig ist und $\langle B\rangle = V$ gilt. △

[26] englisch: basis

Beachte: Eine Basis ist also eine linear unabhängige erzeugende Menge bzw. eine linear unabhängige erzeugende Familie. Zur Unterscheidung sprechen wir auch von einer **Basismenge** (englisch: basic set) bzw. einer **Basisfamilie** (englisch: basic family).

Beispiel 13.2 (Basis)**.**

(i) Die leere Menge $\emptyset$ ist die einzige Basis des Nullraumes $\{0\}$ über jedem Körper K.

(ii) Die Menge $E \coloneqq \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} \right\}$ ist eine erzeugende Familie des $\mathbb{R}$-Vektorraumes $\mathbb{R}^2$, jedoch keine Basis, da sie linear abhängig ist, siehe Beispiel 12.2. Wenn wir ein beliebiges Element aus E entfernen, so erhalten wir eine Basis von $\mathbb{R}^2$.

(iii) Im Standardvektorraum K^n über einem Körper K (Beispiel 11.3) ist

$$\{e_1, \dots, e_n\} \text{ mit } e_i \coloneqq \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-ter Eintrag} \tag{13.1}$$

eine Basis, genannt die **Standardbasis** (englisch: standard basis) von K^n.

(iv) $\{1, i\}$ ist eine Basis des $\mathbb{R}$-Vektorraumes $(\mathbb{C}, +, \cdot)$.

(v) Für den $\mathbb{Q}$-Vektorraum $\mathbb{R}$ ist keine Basis explizit bekannt.

(vi) Es seien K ein Körper und $K^{\mathbb{N}}$ der Folgenraum über K, siehe Beispiel 11.12. Die Menge $B = \{e_j \mid j \in \mathbb{N}\}$ aller Standardfolgen ist eine Basis des Unterraumes der endlich getragenen Folgen, also $\langle B \rangle = (K^{\mathbb{N}})_{00}$, siehe Beispiel 11.17. Diese Basis wird die **Standardbasis** von $(K^{\mathbb{N}})_{00}$ genannt. △

Satz 13.3 (Charakterisierung von Basen)**.**
Es sei V ein Vektorraum. Für eine Teilmenge $B \subseteq V$ bzw. für eine Familie $B = (v_i)_{i \in I}$ von Vektoren in V sind äquivalent:

(i) B ist eine Basis von V.

(ii) B ist eine maximale linear unabhängige Teilmenge von V.[27] Das heißt: B ist linear unabhängig, und jede echte Obermenge von B ist linear abhängig.

(iii) B ist eine minimale erzeugende Menge von V.[28] Das heißt: B ist eine erzeugen-

(i) B ist eine Basis von V.

(ii) B ist eine maximale linear unabhängige Teilfamilie von V.[29] Das heißt: B ist linear unabhängig, und jede echte Oberfamilie von B ist linear abhängig.

(iii) B ist eine minimale erzeugende Familie von V.[30] Das heißt: B ist eine

[27] Genauer: B ist ein maximales Element bzgl. der Mengeninklusion (Definition 5.34) in der Menge der linear unabhängigen Teilmengen von V.

[28] Genauer: B ist ein minimales Element bzgl. der Mengeninklusion in der Menge der erzeugenden Mengen von V.

de Menge, und jede echte Teilmenge von B ist keine erzeugende Menge.

(*iv*) Jeder Vektor $v \in V$ lässt sich auf eindeutige Weise (bis auf Summanden mit Nullkoeffizienten) aus Vektoren der Menge B linearkombinieren.

erzeugende Familie, und jede echte Teilfamilie von B ist keine erzeugende Familie.

(*iv*) Jeder Vektor $v \in V$ lässt sich auf eindeutige Weise (bis auf Summanden mit Nullkoeffizienten) aus Vektoren der Familie B linearkombinieren.

Beweis. Wir beweisen nur die Aussagen für Mengen.

Aussage (*i*) $\Rightarrow$ Aussage (*ii*): Es sei B eine Basis von V, d. h., B ist linear unabhängig und $\langle B \rangle = V$. Für einen beliebigen Vektor $v \in V \setminus B$ gilt $\langle B \cup \{v\} \rangle = V$. Aus Folgerung 12.6 („Redundanz bedeutet lineare Abhängigkeit") folgt nun, dass $B \cup \{v\}$ linear abhängig ist.

Aussage (*ii*) $\Rightarrow$ Aussage (*i*): Es sei B eine maximale linear unabhängige Teilmenge von V. Zu zeigen ist, dass B ganz V erzeugt. Es sei dazu $\widetilde{v} \in V$ beliebig. Nach Definition von $\langle B \rangle$ ist klar, dass $\langle B \rangle \supseteq B$ gilt. (**Quizfrage 13.1:** Klar?) Wenn also $\widetilde{v} \in B$ ist, dann auch $\widetilde{v} \in \langle B \rangle$, was zu zeigen war. Wir können also von $\widetilde{v} \in V \setminus B$ ausgehen. Nach Voraussetzung ist $B \cup \{\widetilde{v}\}$ als echte Obermenge von B linear abhängig. Es existieren also eine endliche Teilmenge $B_0 \subseteq B$ und Koeffizienten $\alpha_v \in K$ für $v \in B_0$, sodass gilt:

$$\sum_{v \in B_0} \alpha_v \, v + \widetilde{\alpha} \, \widetilde{v} = 0,$$

wobei nicht alle Koeffizienten gleich Null sind. Insbesondere ist $\widetilde{\alpha} \neq 0$, denn sonst wäre bereits B linear abhängig, was der Voraussetzung widerspricht. Wir erhalten also

$$\widetilde{v} = -\sum_{v \in B_0} \widetilde{\alpha}^{-1} \alpha_v \, v,$$

d. h., $\widetilde{v}$ lässt sich in der Tat aus Elementen von B linearkombinieren. Damit ist $\langle B \rangle = V$ gezeigt.

Aussage (*i*) $\Rightarrow$ Aussage (*iii*): Es sei B eine Basis von V, d. h., B ist linear unabhängig und $\langle B \rangle = V$. Aus Folgerung 12.6 („Redundanz bedeutet lineare Abhängigkeit") folgt nun, dass B keine redundanten Elemente enthält, dass also für alle $v \in B$ gilt: $\langle B \setminus \{v\} \rangle \subsetneq V$.

Aussage (*iii*) $\Rightarrow$ Aussage (*i*): Nach Voraussetzung haben wir $\langle B \rangle = V$, und für alle $v \in B$ gilt: $\langle B \setminus \{v\} \rangle \subsetneq V$. Aus Folgerung 12.6 („Redundanz bedeutet lineare Abhängigkeit") folgt daher, dass B linear unabhängig ist, also eine Basis.

Aussage (*i*) $\Rightarrow$ Aussage (*iv*): Es sei B eine Basis von V, d. h., B ist linear unabhängig und $\langle B \rangle = V$. Aus Lemma 12.7 folgt, dass sich jedes $v \in \langle B \rangle = V$ in i. W. eindeutiger Weise aus Elementen von B linearkombinieren lässt.

Aussage (*iv*) $\Rightarrow$ Aussage (*i*): Nach Voraussetzung lässt sich jeder Vektor $v \in V$ auf eindeutige Weise (bis auf Summanden mit Nullkoeffizienten) aus Elementen von B linearkombinieren. Also ist $\langle B \rangle = V$, und aus Lemma 12.7 folgt, dass B linear unabhängig ist, also eine Basis von V. □

[29] Genauer: Die Familie B ist ein maximales Element bzgl. der Ordnungsrelation „ist Teilfamilie von" in der Menge der linear unabhängigen Familien in V.

[30] Genauer: Die Familie B ist ein minimales Element bzgl. der Ordnungsrelation „ist Teilfamilie von" in der Menge der erzeugender Familien von V.

Folgerung 13.4 (verschiedene Basen sind bzgl. der Inklusionshalbordnung nicht vergleichbar).
Es sei V ein Vektorraum.

Sind die Mengen $B_1, B_2 \subseteq V$ zwei verschiedene Basen von V, dann sind B_1 und B_2 bzgl. der Halbordnung „$\subseteq$“ nicht vergleichbar, d. h., es gilt weder $B_1 \subseteq B_2$ noch $B_2 \subseteq B_1$.

Sind $B_1 = (v_i)_{i \in I_1}$ und $B_2 = (w_i)_{i \in I_2}$ Familien von Vektoren in V und sind die Mengen ihrer Mitglieder $E_1 \coloneqq \{v_i \mid i \in I_1\}$ und $E_2 \coloneqq \{w_i \mid i \in I_2\}$ verschieden, dann sind E_1 und E_2 bzgl. der Halbordnung „$\subseteq$“ nicht vergleichbar, d. h., es gilt weder $E_1 \subseteq E_2$ noch $E_2 \subseteq E_1$.

Beweis. Wir führen den Beweis für Mengen. Die Annahme von $B_1 \subsetneq B_2$ oder von $B_2 \subsetneq B_1$ für zwei Basen B_1, B_2 von V widerspräche Aussage (ii) von Satz 13.3. □

Wir geben nun einige wichtige Resultate zur Existenz von Basen an. Das Hauptresultat — der nachfolgende Basisergänzungssatz — besagt, dass zwischen einer linear unabhängigen, aber möglicherweise zu kleinen Menge (der erzeugte Unterraum ist nicht der ganze Vektorraum), und einer erzeugenden Menge, die aber möglicherweise zu groß (linear abhängig) ist, immer eine Basis liegt.

Satz 13.5 (**Basisergänzungssatz**[31] AoC).
Es sei V ein Vektorraum.

Es sei A eine linear unabhängige Menge von Vektoren aus V und E eine erzeugende Menge von V mit der Eigenschaft $A \subseteq E$. Dann existiert eine Basis B von V mit $A \subseteq B \subseteq E$.

Es sei $A = (v_i)_{i \in I_A}$ eine linear unabhängige Familie von Vektoren aus V und $E = (v_i)_{i \in I_E}$ eine erzeugende Familie von V mit der Eigenschaft $I_A \subseteq I_E$.[32] Dann existiert eine Basis $B = (v_i)_{i \in I_B}$ von V mit $I_A \subseteq I_B \subseteq I_E$.

Beweis. Wir führen den Beweis für Mengen.

Wir betrachten die Menge aller linear unabhängigen Teilmengen zwischen A und E, also

$$\mathcal{D} \coloneqq \{D \subseteq V \mid A \subseteq D \subseteq E, \text{ sodass } D \text{ linear unabhängig ist}\} \subseteq \mathcal{P}(E) \subseteq \mathcal{P}(V).$$

$\mathcal{D}$ ist bzgl. der Mengeninklusion eine halbgeordnete Menge. Wegen $A \in \mathcal{D}$ ist $\mathcal{D} \neq \emptyset$. Ziel ist die Anwendung des Lemmas von Zorn 6.48.

Schritt 1: Jede totalgeordnete Teilmenge $\mathcal{C} \subseteq \mathcal{D}$ besitzt eine obere Schranke S in $\mathcal{D}$:

Wir zeigen, dass $S \coloneqq \bigcup_{C \in \mathcal{C}} C$ eine obere Schranke (Definition 5.34) der Teilmenge $\mathcal{C}$ in $\mathcal{D}$ ist.

Dazu zeigen wir zunächst, dass S überhaupt Element von $\mathcal{D}$ ist:

(i) Für alle $C \in \mathcal{C} \subseteq \mathcal{D}$ gilt $A \subseteq C \subseteq E$, also auch $A \subseteq \bigcup_{C \in \mathcal{C}} C = S \subseteq E$.

[31] englisch: basis extension theorem
[32] A ist also eine Teilfamilie von E.

(ii) Weiter ist S linear unabhängig, denn: Es sei $\{c_1, \dots, c_n\} \subseteq S$ eine endliche Teilmenge. Für alle c_i existiert $C_i \in \mathcal{C}$ mit $c_i \in C_i$. $\mathcal{C}$ ist aber totalgeordnete Teilmenge, also existiert ein Maximum C_k der endlichen Teilmenge $\{C_1, \dots, C_n\}$ in $\mathcal{C}$. Folglich gilt $C_i \subseteq C_k$ für alle $i = 1, \dots, n$. Damit ist $\{c_1, \dots, c_n\} \subseteq \bigcup_{i=1}^n C_i = C_k$. Dabei ist $C_k \in \mathcal{C} \subseteq \mathcal{D}$ linear unabhängig. Also ist nach Lemma 12.4 auch die Teilmenge $\{c_1, \dots, c_n\}$ linear unabhängig.

Schließlich zeigt $C \subseteq \bigcup_{C \in \mathcal{C}} C = S$ für alle $C \in \mathcal{C}$, dass S tatsächlich eine obere Schranke von $\mathcal{C}$ in $\mathcal{D}$ ist.

Schritt 2: Das Lemma von Zorn 6.48 zeigt nun, dass ein maximales Element B von $\mathcal{D}$ existiert. Das heißt definitionsgemäß: B ist linear unabhängig und erfüllt $A \subseteq B \subseteq E$. Es bleibt zu zeigen, dass B tatsächlich ganz V erzeugt.

Falls $B = E$ gilt, so ist $V = \langle E \rangle = \langle B \rangle$ und der Beweis erbracht. Andernfalls gibt es ein $a \in E \setminus B$, also gilt $B \cup \{a\} \supsetneq B$. B ist aber ein maximales Element von $\mathcal{D}$, also kann $B \cup \{a\}$ nicht zu $\mathcal{D}$ gehören. Wegen $A \subseteq B \cup \{a\} \subseteq E$ muss das daran liegen, dass $B \cup \{a\}$ linear abhängig ist. Aus Folgerung 12.6 folgt also $\langle B \cup \{a\} \rangle = \langle B \rangle$ und insbesondere $a \in \langle B \rangle$. Da dieses Argument für jedes $a \in E \setminus B$ gilt, folgt $E \setminus B \subseteq \langle B \rangle$, und natürlich gilt auch $B \subseteq \langle B \rangle$. Es folgt $E \subseteq \langle B \rangle$. Der Übergang zur linearen Hülle zeigt $V = \langle E \rangle \subseteq \big\langle \langle B \rangle \big\rangle = \langle B \rangle$, aber natürlich gilt auch $\langle B \rangle \subseteq V$, also $\langle B \rangle = V$. □

Es folgen drei unmittelbare Folgerungen als Spezialfälle des Basisergänzungssatzes 13.5.

Der Fall $A = \emptyset$ und $E = V$ bzw. $A = ()$ und $E = (v)_{v \in V}$ ergibt:

Folgerung 13.6 (Basisexistenzsatz$^{\text{AoC}}$)**.**
Jeder Vektorraum V besitzt eine Basismenge.

Jeder Vektorraum V besitzt eine Basisfamilie.

Im Fall von $A = \emptyset$ bzw. $A = ()$ erhalten wir:

Folgerung 13.7 (Basisauswahlsatz$^{\text{AoC}}$)**.**
Aus jeder erzeugenden Menge E eines Vektorraumes V lässt sich eine Basis auswählen.

Aus jeder erzeugenden Familie F eines Vektorraumes V lässt sich eine Basis auswählen.

Und $E = V$ bzw. $E = (v)_{v \in V}$ führt schließlich zu:

Folgerung 13.8 (Basisergänzungssatz$^{\text{AoC}}$)**.**
Jede linear unabhängige Menge A eines Vektorraumes V kann zu einer Basismenge erweitert werden.

Jede linear unabhängige Familie A eines Vektorraumes V kann zu einer Basisfamilie erweitert werden.

Bemerkung 13.9 (zum Basisergänzungssatz 13.5)**.**

(i) Der Beweis des Basisergänzungssatzes 13.5 und seiner Folgerungen 13.6 bis 13.8 ist nicht konstruktiv, d. h., wir können ihn nicht zur Grundlage eines Verfahrens machen, um eine Basis zu konstruieren. Beispielsweise können wir mit Hilfe dieser Resultate keine konkrete

Basis von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum (Beispiel 11.3) konstruieren, obwohl eine solche existiert.[AoC] (**Quizfrage 13.2:** Ist eine Basis des $\mathbb{Q}$-Vektorraumes $\mathbb{R}$ abzählbar oder überabzählbar?)

(ii) Wenn der Vektorraum V jedoch endlich erzeugt ist, wenn es also eine endliche erzeugende Menge oder Familie gibt, dann lässt sich der Basisergänzungssatz 13.5 konstruktiv und ohne Verwendung des Zornschen Lemmas beweisen, indem man die linear unabhängige Menge A Schritt für Schritt durch einzelne Elemente von E ergänzt oder alternativ Schritt für Schritt einzelne Elemente von $E \setminus A$ entfernt. Das heißt, auch die Folgerungen 13.6 bis 13.8 benötigen in diesem Fall das Auswahlaxiom nicht. △

Lemma 13.10 (Basis endlicher kartesischer Produkte)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$. Weiter sei $V \coloneqq \bigtimes_{j=1}^n V_j$ der Produktraum der K-Vektorräume $V_1, \ldots, V_n$.

Ist B_j eine Basismenge von V_j für $j = 1, \ldots, n$, dann ist

$$\bigcup_{j=1}^n \Big\{ (0, \ldots, 0, \underbrace{v}_{\text{Position } j \text{ im Tupel}}, 0, \ldots, 0) \,\Big|\, v \in B_j \Big\}$$

eine Basismenge von V.

Ist $B_j = (v_i)_{i \in I_j}$ eine Basisfamilie von V_j für $j = 1, \ldots, n$, dann ist

$$\big\|_{j=1}^n (0, \ldots, 0, \underbrace{v_i}_{\text{Position } j \text{ im Tupel}}, 0, \ldots, 0)_{i \in I_j}$$

eine Basisfamilie von V.

(**Quizfrage 13.3:** Gilt ein ähnliches Resultat auch für unendliche Produkte von Vektorräumen?)

Beweis. Dieser Beweis ist Gegenstand der Übung. □

§ 13.2 Dimension eines Vektorraumes

Wir kommen nun zum wichtigen Begriff der Dimension, der in gewissem Sinne die „Größe" eines Vektorraumes beschreibt. Da sich dieser am Begriff der Basis orientiert, werden wir die Existenz von Basen voraussetzen müssen.

Definition 13.11 (Dimension eines Vektorraumes)**.**
Es sei V ein Vektorraum.

(i) Wenn V eine endliche Basismenge B der Kardinalität $n \in \mathbb{N}_0$ besitzt, so sagen wir, V habe **endliche Dimension** [33], genauer: die **Dimension**[34] n, in Symbolen: $\dim(V) = n$.

(ii) Wenn V keine endliche Basismenge

(i) Wenn V eine endliche Basisfamilie $B = (v_i)_{i \in I}$ mit $\#I = n \in \mathbb{N}_0$ Mitglieder besitzt, so sagen wir, V habe **endliche Dimension**, genauer: die **Dimension** n, in Symbolen: $\dim(V) = n$.

(ii) Wenn V keine endliche Basisfamilie

[33] englisch: finite dimension
[34] englisch: dimension

besitzt[35], so sagen wir, V habe **unendliche Dimension**[36], in Symbolen: $\dim(V) = \infty$.

besitzt[37], so sagen wir, V habe **unendliche Dimension**, in Symbolen: $\dim(V) = \infty$. △

Zur Verdeutlichung, welcher Körper K verwendet wird, schreiben wir manchmal auch $\dim_K(V)$. Beispielsweise gilt $\dim_{\mathbb{R}}(\mathbb{R}) = 1$, aber $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$.

Bevor wir mit dem Begriff der Dimension arbeiten, muss noch sichergestellt werden, dass dieser wohldefiniert ist, denn ein Vektorraum besitzt i. A. viele verschiedene Basen. Wir werden dazu beweisen, dass in dem Fall, dass ein Vektorraum eine endliche Basismenge besitzt, alle seine Basismengen endlich sind, und zwar alle mit dieselben Kardinalität (Bemerkung 13.15).

Lemma 13.12 (Austauschlemma)**.**
Es sei V ein Vektorraum über dem Körper K.

Die endliche Menge $B = \{v_1, \ldots, v_n\}$ sei eine Basis von V und $\#B = n \in \mathbb{N}_0$.[38] Ist $w = \sum_{i=1}^n \alpha_i\, v_i$ mit Koeffizienten $\alpha_i \in K$ und gilt $\alpha_j \neq 0$ für den Index $j \in [\![1, n]\!]$, dann ist auch $B_0 \coloneqq \{v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n\}$ eine Basis von V.

Die endliche Familie $B = (v_1, \ldots, v_n)$ sei eine Basis von V mit $n \in \mathbb{N}_0$. Ist $w = \sum_{i=1}^n \alpha_i\, v_i$ mit Koeffizienten $\alpha_i \in K$ und gilt $\alpha_j \neq 0$ für den Index $j \in [\![1, n]\!]$, dann ist auch $B_0 \coloneqq (v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n)$ eine Basis von V.

Beweis. Wir führen den Beweis für Mengen.

Da es bei einer Basismenge auf die Reihenfolge der Elemente nicht ankommt, nehmen wir aus Bequemlichkeit und o. B. d. A. an, dass $j = 1$ ist. Aus $w = \sum_{i=1}^n \alpha_i\, v_i$ mit $\alpha_1 \neq 0$ folgt

$$v_1 = \alpha_1^{-1}\Big(w - \sum_{i=2}^n \alpha_i\, v_i\Big). \tag{13.2}$$

Schritt 1: Wir zeigen: $\langle B_0 \rangle = V$.

Es sei dazu $v \in V$. Da B eine Basismenge ist, gibt es Koeffizienten $\beta_1, \ldots, \beta_n \in K$, sodass $v = \sum_{i=1}^n \beta_i\, v_i$ gilt. Durch Einsetzen von (13.2) folgt

$$\begin{aligned} v &= \beta_1\, \alpha_1^{-1}\Big(w - \sum_{i=2}^n \alpha_i\, v_i\Big) + \sum_{i=2}^n \beta_i\, v_i \\ &= \beta_1\, \alpha_1^{-1} w - \beta_1\, \alpha_1^{-1} \sum_{i=2}^n \alpha_i\, v_i + \sum_{i=2}^n \beta_i\, v_i \quad \text{nach Distributivgesetz (10.1a) in } K \\ &= \beta_1\, \alpha_1^{-1} w + \sum_{i=2}^n \big(\beta_i - \beta_1\, \alpha_1^{-1} \alpha_i\big)\, v_i \end{aligned}$$

[35] Das kann bedeuten, dass V überhaupt keine Basismenge besitzt (was ohne das Auswahlaxiom der Fall sein kann) oder dass jede Basismenge von V eine unendliche Menge ist.
[36] englisch: infinite dimension
[37] Das kann bedeuten, dass V überhaupt keine Basisfamilie besitzt (was ohne das Auswahlaxiom der Fall sein kann) oder dass jede Basisfamilie von V eine unendliche Familie ist.
[38] Das heißt also, dass die Vektoren $v_1, \ldots, v_n$ paarweise verschieden sind.

nach Kommutativgesetz und Distributivgesetz (10.1b). Damit ist gezeigt, dass in der Tat $\langle B_0 \rangle = \langle w, v_2, \dots, v_n \rangle = V$ gilt.

Schritt 2: Wir zeigen: B_0 ist linear unabhängig.

Wir betrachten dazu

$$\begin{aligned} & \beta_1 w + \sum_{i=2}^n \beta_i v_i = 0 \\ \Rightarrow \quad & \beta_1 \sum_{i=1}^n \alpha_i v_i + \sum_{i=2}^n \beta_i v_i = 0 \\ \Rightarrow \quad & \beta_1 \alpha_1 v_1 + \sum_{i=2}^n (\beta_1 \alpha_i + \beta_i) v_i = 0. \end{aligned}$$

Da $B = \{v_1, \dots, v_n\}$ linear unabhängig ist, ist das nur möglich, wenn alle Koeffizienten gleich Null sind, also

$$\beta_1 \alpha_1 = 0 \quad \text{und} \quad \beta_1 \alpha_i + \beta_i = 0 \text{ für } i = 2, \dots, n.$$

Wegen $\alpha_1 \neq 0$ muss $\beta = 0$ sein, woraus dann weiter $\beta_2 = \dots = \beta_n = 0$ folgt. Das heißt aber, dass $B_0 = \{w, v_2, \dots, v_n\}$ linear unabhängig ist. □

Satz 13.13 (**Austauschsatz von Steinitz**[39]).
Es sei V ein Vektorraum.

Die endliche Menge $B = \{v_1, \dots, v_n\}$ sei eine Basis von V mit $\#B = n \in \mathbb{N}_0$. Ist $A = \{a_1, \dots, a_m\}$ eine weitere linear unabhängige Menge in V mit $\#A = m \in \mathbb{N}_0$, dann gilt:

(*i*) $m \leqslant n$.

(*ii*) Es gibt eine $(n-m)$-elementige Teilmenge D von B, sodass $B_0 \coloneqq A \cup D$ ebenfalls eine Basis von V ist. Es gilt $\#B_0 = n$.

Die endliche Familie $B = (v_1, \dots, v_n)$ sei eine Basis von V mit $n \in \mathbb{N}_0$. Ist $A = (a_1, \dots, a_m)$ eine weitere linear unabhängige Familie in V mit $m \in \mathbb{N}_0$, dann gilt:

(*i*) $m \leqslant n$.

(*ii*) Es gibt eine Teilfamilie D von B mit $(n-m)$ Mitgliedern, sodass $B_0 \coloneqq A \parallel D$ ebenfalls eine Basis von V ist. B_0 hat n Mitglieder.

Beweis. Wir führen den Beweis für Mengen, und zwar mit Hilfe vollständiger Induktion nach der Mächtigkeit $m = \#A \in \mathbb{N}_0$. Den Induktionsanfang setzen wir bei $m = 0$. Dann ist $A = \emptyset$, und Aussage (*i*) gilt wegen $m = 0 \leqslant n$, und Aussage (*ii*) gilt mit $D = B$.

Es sei nun $m \geqslant 1$, und es gelten Aussagen (*i*) und (*ii*) bereits für $m-1$. Da $\{a_1, \dots, a_m\}$ linear unabhängig ist, ist auch $\{a_1, \dots, a_{m-1}\}$ linear unabhängig. Nach Induktionsannahme gilt $m - 1 \leqslant n$, und es existiert $D = \{v_{i_m}, \dots, v_{i_n}\} \subseteq B$, sodass $B_1 \coloneqq \{a_1, \dots, a_{m-1}, v_{i_m}, \dots, v_{i_n}\}$ eine Basis von V ist.

Falls nun $m - 1 = n$ wäre, also $D = \emptyset$, dann wäre bereits $\{a_1, \dots a_{m-1}\}$ eine Basis von V. Nach dem Satz 13.3 über die Charakterisierung von Basen hieße das aber, dass $\{a_1, \dots a_m\}$ linear

[39] englisch: Steinitz exchange theorem

abhängig wäre, im Widerspruch zur Voraussetzung. Es gilt also $m - 1 < n$, also $m \leqslant n$. Damit ist der Induktionsschritt für Aussage (i) gezeigt. Da B_1 eine Basis von V ist, können wir jeden Vektor, insbesondere a_m, durch die Basiselemente linearkombinieren. Es gibt also Skalare α_j, $j = 1, \dots n$, sodass gilt:

$$a_m = \sum_{j=1}^{m-1} \alpha_j \, a_j + \sum_{j=m}^{n} \alpha_j \, v_{i_j}.$$

Wären alle $\alpha_m = \alpha_{m+1} = \dots \alpha_n = 0$, so würde das die lineare Abhängigkeit von $A = \{a_1, \dots, a_m\}$ zeigen, im Widerspruch zur Voraussetzung. Demzufolge gibt es einen Index $j \in [\![m, n]\!]$ mit $\alpha_j \neq 0$. Nach dem Austauschlemma 13.12 ist $B_0 = B_1 \setminus \{v_{i_j}\} \cup \{a_m\}$ eine Basis von V. Die Kardinalität von B_0 ist $\#B_0 \leqslant \#B_1 - 1 + 1 \leqslant n$. Wäre $a_m \in B_1 \setminus \{v_{i_j}\}$, dann würde aus

$$0 = \sum_{j=1}^{m-1} \alpha_j \, a_j + \sum_{j=m}^{n} \alpha_j \, v_{i_j} - a_m$$

folgen, das B_1 linear abhängig ist, im Widerspruch zur Basiseigenschaft von B_1. Somit folgt $\#B_0 = n$, was den Induktionsschritt für Aussage (ii) zeigt. □

Folgerung 13.14 (endliche Basen sind gleichmächtig)**.**
Es sei V ein Vektorraum.

(i) Wenn V endlich erzeugt ist, dann besitzt V eine Basismenge. Jede Basismenge von V endlich, und alle Basismengen sind gleichmächtig.

(ii) Wenn V nicht endlich erzeugt ist, dann ist jede Basismenge von V unendlich. [40]

(i) Wenn V endlich erzeugt ist, dann besitzt V eine Basisfamilie. Jede Basisfamilie von V ist endlich, und alle Basisfamilien sind gleichmächtig.

(ii) Wenn V nicht endlich erzeugt ist, dann ist jede Basisfamilie von V unendlich. [40]

Beweis. Wir führen den Beweis für Mengen.

Aussage (i): Wenn V endlich erzeugt ist, dann gibt es eine endliche erzeugende Menge und nach Basisergänzungssatz 13.5 damit auch eine endliche Basis B von V, sagen wir mit Mächtigkeit $\#B = n \in \mathbb{N}_0$. Es sei B_0 eine weitere (möglicherweise unendliche) Basis von V. Insbesondere jede endliche Teilmenge $B_1 \subseteq B_0$ ist dann ebenfalls linear unabhängig, und nach Satz 13.13 ist $\#B_1 \leqslant n$. Damit muss B_0 selbst endlich sein mit $\#B_0 \leqslant n = \#B$. Durch Tausch der Rollen von B und B_0 folgt auch $\#B \leqslant \#B_0$, also zusammen $\#B = \#B_0$.

Aussage (ii): Wäre B eine endliche Basis, dann wäre insbesondere B eine endliche erzeugende Menge des Vektorraumes V. Der Vektorraum V wäre damit endlich erzeugt, im Widerspruch zur Voraussetzung. □

Bemerkung 13.15 (Dimensionsbegriff für Vektorräume)**.**

[40] Ohne das Auswahlaxiom heißt das nicht, dass eine Basis überhaupt existiert.

(i) Folgerung 13.14 zeigt, dass der Dimensionsbegriff aus Definition 13.11 wohldefiniert ist, wobei wir der Einfachheit halber nicht zwischen verschiedenen Unendlichkeiten unterschieden haben. Sogar für unendlich-dimensionale Vektorräume gilt aber, dass je zwei Basen gleichmächtig sind.[AoC] Daher könnten wir für die Dimension von Vektorräumen genauer auch Kardinalzahlen (Definition 6.31) verwenden und dadurch verschiedene Unendlichkeiten unterscheiden.

(ii) Der Beweis von Folgerung 13.14 verwendet den Basisergänzungssatz 13.5, jedoch nur die Version für endlich-dimensionale (endlich erzeugte) Vektorräume, die ohne das Zornsche Lemma und damit ohne das Auswahlaxiom auskommt. △

Beispiel 13.16 (Dimension eines Vektorraumes)**.**

(i) Der als „Standardvektorraum K^n der Dimension n" bezeichnete Vektorraum über einem Körper K (Beispiel 11.3) hat tatsächlich die Dimension $n \in \mathbb{N}_0$, da die Standardbasis $\{e_1, \ldots, e_n\}$ bzw. $(e_1, \ldots, e_n)$ die n paarweise verschiedenen Elemente bzw. Mitglieder hat (Beispiel 13.2).

(ii) Es gilt $\dim_{\mathbb{R}}(\mathbb{R}) = 1$ und $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$.

(iii) Es gilt $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ und $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. Eine Basismenge für den $\mathbb{R}$-Vektorraum $\mathbb{C}$ ist $\{1, i\}$.

(iv) Der Nullraum $\{0\}$ ist über jedem Körper der einzige Vektorraum der Dimension 0.

(v) Der Folgenraum $K^{\mathbb{N}}$ über einem Körper K hat unendliche Dimension. Auch der Unterraum $(K^{\mathbb{N}})_{00}$ der endlich getragenen Folgen hat unendliche Dimension, da die (abzählbar) unendliche Menge B aller Standardfolgen eine Basis von $(K^{\mathbb{N}})_{00}$ bildet (Beispiel 11.17).

Der Unterraum der Folgen, deren Träger in $[\![1, n]\!]$ liegt, hat Dimension n, da die Standardfolgen $\{e_1, \ldots, e_n\}$ eine Basis bilden. △

Folgerung 13.17 (Dimension, Unterräume und lineare Unabhängigkeit)**.**
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}_0$.

(i) Ist $A \subseteq V$ eine linear unabhängige Menge, dann gilt $\#A \leqslant n$.

(ii) $A \subseteq V$ ist genau dann eine Basis von V, wenn A linear unabhängig ist und $\#A = n$ gilt.

(i) Ist $A = (v_i)_{i \in I}$ eine linear unabhängige Familie in V, dann gilt $\#I \leqslant n$.

(ii) Eine Familie $A = (v_i)_{i \in I}$ in V ist genau dann eine Basis von V, wenn A linear unabhängig ist und $\#I = n$ gilt.

(iii) Für jeden Unterraum U von V gilt: $0 \leqslant \dim(U) \leqslant \dim(V)$.

(iv) Für jeden Unterraum U von V ist $U = V$ genau dann, wenn $\dim(U) = \dim(V)$ gilt.

Beweis. Wir führen den Beweis für Mengen.

Aussage (i): Nach Basisergänzungssatz 13.5 existiert eine Basis B von V mit $A \subseteq B$. Nach Folgerung 13.14 ist $\#B = n$ und daher $\#A \leqslant n$.

Aussage (ii): Ist A eine Basis von V, dann ist A definitionsgemäß linear unabhängig, und nach Folgerung 13.14 gilt $\#A = \dim(V) = n$. Ist umgekehrt A linear unabhängig und $\#A = n$, so gilt für jede Basis $B \supseteq A$ von V einerseits $\#B \geqslant \#A$, andererseits aber $\#B = \dim(V) = n$. Also muss $B = A$ sein, d. h., A ist bereits eine Basis.

Aussage (iii): Ist A eine Basis des Unterraumes U von V, dann ist A linear unabhängige Teilmenge von U und damit auch von V. Aus Aussage (i) folgt $\dim(U) = \#A \leqslant n = \dim(V)$.

Aussage (iv): Ist $U = V$, dann gilt $\dim(U) = \dim(V)$. Nun gelte andererseits $\dim(U) = \dim(V)$, und es sei A eine Basis von U. Dann ist A linear unabhängige Teilmenge von U und damit auch von V. Es gilt $\#A = \dim(U) = \dim(V) = n$. Aus Aussage (ii) folgt, dass A auch eine Basis von V ist, also gilt $U = \langle A \rangle = V$. □

Lemma 13.18 (Dimension des endlichen kartesischen Produkts endlich-dimensionaler Vektorräume)**.**

Es seien K ein Körper und $V_1, \ldots, V_n$ K-Vektorräume endlicher Dimension für $n \in \mathbb{N}_0$. Dann besitzt das kartesische Produkt (Definition 11.4) die Dimension

$$\dim\Big(\bigtimes_{i=1}^{n} V_i\Big) = \sum_{i=1}^{n} \dim(V_i). \tag{13.3}$$

Beweis. Das Resultat folgt sofort aus Lemma 13.10, weil sich die Kardinalität einer Basis von $\bigtimes_{i=1}^{n} V_i$ als Summe der Kardinalitäten der Basen von V_i, $i = 1, \ldots, n$, ergibt. □

Abschließend betrachten wir die Frage, wieviele Elemente (Vektoren) ein Vektorraum besitzt.

Bemerkung 13.19 (Mächtigkeit von Vektorräumen)**.**

Es sei V ein Vektorraum über einem Körper K.

(i) Hat der Vektorraum V unendliche Dimension, dann besitzt er auch unendliche viele Elemente. Besäße er nur endlich viele Elemente, dann gäbe es auch ein endliches Erzeugendensystem.

(ii) Gilt $\dim(V) = 0$, dann gilt $V = \{0\}$, V hat also genau ein Element, und zwar unabhängig vom Körper K.

(iii) Gilt $\dim(V) \in \mathbb{N}$, dann kommt es auf die Mächtigkeit des Körpers K an: Ist K endlich, dann besteht der Raum V aus genau $(\#K)^{\dim(V)}$ verschiedenen Vektoren. Ist K dagegen unendlich, dann ist auch V unendlich. △

Ende der Vorlesung 17

---

§ 14 Summen von Unterräumen

Literatur: Bosch, 2014, Kapitel 1; Fischer, Springborn, 2020, Kapitel 2.6

§ 14.1 Summen von zwei Unterräumen

Aus Lemma 11.14 wissen wir, dass der Durchschnitt $U \cap W$ zweier Unterräume U, W eines Vektorraumes V wieder ein Vektorraum ist. Die Vereinigung $U \cup W$ ist jedoch i. A. kein Unterraum von V.[42]

Statt $U \cup W$ können wir aber den kleinsten Unterraum von V betrachten, der $U \cup W$ enthält, also den von $U \cup W$ erzeugten Unterraum $\langle U \cup W \rangle$:

Lemma 14.1 (die lineare Hülle der Vereinigung zweier Unterräume ist die Summe, vgl. Folgerung 7.52 für Gruppen)**.**
Es seien V ein Vektorraum und U, W zwei Unterräume von V. Dann gilt

$$\langle U \cup W \rangle = U + W. \tag{14.1}$$

Die Menge $U + W$ ist zu verstehen im Sinne der Bemerkung 7.20, also als $U + W = \{u + w \mid u \in U, \ w \in W\}$. $U + W$ heißt die **Summe der Unterräume** U und W (englisch: sum of two subspaces).

Beweis. Aus Satz 11.16 wissen wir, dass $M \coloneqq \langle U \cup W \rangle$ übereinstimmt mit der Menge aller Linearkombinationen von $U \cup W$. Wir zeigen nun in zwei Schritten, dass $M = U + W$ gilt.

Schritt 1: In der Tat sind die Elemente $u + w$ von $U + W$ auch Linearkombinationen von $U \cup W$, nämlich $1\,u + 1\,w$. Also gilt $U + W \subseteq M$.

Schritt 2: Ist umgekehrt $v \in M$, dann hat v als Linearkombination von $U \cup W$ eine Darstellung der Form

$$v = \sum_{u \in U_0} \alpha_u \, u + \sum_{w \in W_0} \beta_w \, w$$

mit endlichen Teilmengen $U_0 \subseteq U$ und $W_0 \subseteq W$ und Koeffizienten $\alpha_u, \beta_w \in K$. Da U und W Unterräume sind, ergibt die erste Summe wieder ein Element von U, und die zweite Summe ergibt ein Element von W. Das zeigt $M \subseteq U + W$. □

Bemerkung 14.2 („Unterraum sein" ist eine Ordnungsrelation, vgl. Bemerkung 10.9 zu Unterkörpern)**.**

(*i*) Die Relation „ist Unterraum von" ist eine partielle Ordnung auf der Klasse aller Vektorräume (über beliebigen Körpern).

(*ii*) Insbesondere ist die Menge aller Unterräume eines bestimmten Vektorraumes V durch die Unterraumhalbordnung partiell geordnet. Diese Ordnung stimmt mit der Inklusionshalbordnung überein.

[42] Tatsächlich gilt: $U \cup W$ ist genau dann ein Unterraum, wenn $U \subseteq W$ oder $W \subseteq U$ gilt (Übung). Ein analoges Resultat gilt auch für Untergruppen (Übung), Unterringe und Unterkörper.

(*iii*) Für Unterräume U und W von V ist $U \cap W$ das Infimum von $\{U, W\}$ (Definition 5.34) und $U + W$ das Supremum von $\{U, W\}$. Genau dann, wenn $U \subseteq W$ oder $W \subseteq U$ gilt, ist das Infimum von $\{U, W\}$ ein Minimum und das Supremum ein Maximum von $\{U, W\}$. (**Quizfrage 14.1:** Wie lauten entsprechende Aussagen für Untergruppen, Unterringe und Unterkörper?) △

Satz 14.3 (Dimension der Summe von zwei Unterräumen).
Es seien V ein Vektorraum und U, W zwei endlich-dimensionale Unterräume von V. Dann gilt

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W). \tag{14.2}$$

Beweis. $U \cap W$ ist ebenfalls ein endlich-dimensionaler Unterraum von V, besitzt also eine endliche Basisfamilie $(v_1, \ldots, v_m)$ mit $m = \dim(U \cap W) \in \mathbb{N}_0$. Nach dem Basisergänzungssatz 13.5 kann diese Basis einerseits zu einer Basis von U und andererseits zu einer Basis von W ergänzt werden. Es gibt also eine Familie $(u_1, \ldots, u_k)$ von Vektoren in U und eine Familie $(w_1, \ldots, w_\ell)$ von Vektoren in W mit $k, \ell \in \mathbb{N}_0$, sodass

$$\begin{aligned} & B_U \coloneqq (v_1, \ldots, v_m,\ u_1, \ldots, u_k) && \text{Basis von } U \\ \text{und}\quad & B_W \coloneqq (v_1, \ldots, v_m,\ w_1, \ldots, w_\ell) && \text{Basis von } W \end{aligned}$$

ist. Wir zeigen nun, dass $B \coloneqq (v_1, \ldots, v_m,\ u_1, \ldots, u_k,\ w_1, \ldots, w_\ell)$ eine Basis von $U + W$ ist.

Offenbar ist B eine erzeugende Familie von $U + W$. Die lineare Unabhängigkeit von B zeigen wir wie folgt: Wir setzen die Linearkombination

$$\sum_{i=1}^m \alpha_i\, v_i + \sum_{i=1}^k \beta_i\, u_i + \sum_{i=1}^\ell \gamma_i\, w_i = 0$$

an mit Koeffizienten $\alpha_i, \beta_i, \gamma_i \in K$. Definieren wir $u \coloneqq \sum_{i=1}^m \alpha_i\, v_i + \sum_{i=1}^k \beta_i\, u_i \in U$, so gilt

$$u = \sum_{i=1}^\ell (-\gamma_i)\, w_i \in W,$$

also $u \in U \cap W$. Wir haben also

$$\begin{aligned} \text{einerseits}\quad & u \in U && = \langle v_1, \ldots, v_m,\ u_1, \ldots, u_k \rangle \\ \text{und andererseits}\quad & u \in U \cap W && = \langle v_1, \ldots, v_m \rangle. \end{aligned}$$

Aus der Eindeutigkeit der Darstellung (Lemma 12.7) folgt $\beta_1 = \ldots = \beta_k = 0$. Es gilt also

$$\sum_{i=1}^m \alpha_i\, v_i + \sum_{i=1}^\ell \gamma_i\, w_i = 0,$$

und da $B_W = (v_1, \ldots, v_m,\ w_1, \ldots, w_\ell)$ eine Basis ist, folgt $\alpha_1 = \ldots = \alpha_m = 0$ und $\gamma_1 = \ldots = \gamma_\ell = 0$. Das zeigt die lineare Unabhängigkeit von B, also ist B tatsächlich eine Basis von $U + W$.

Die Behauptung (14.2) folgt nun aus

$$\underbrace{m + k + \ell}_{\dim(U+W)} = \underbrace{m + k}_{\dim(U)} + \underbrace{m + \ell}_{\dim(W)} - \underbrace{m.}_{\dim(U \cap W)}$$

□

Beispiel 14.4 (Summe von zwei Unterräumen).

(*i*) Für die Unterräume

$$U = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \quad \text{und} \quad W = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle$$

von $\mathbb{R}^2$ gilt

$$U + W = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = \mathbb{R}^2 \quad \text{und} \quad U \cap W = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Die Dimensionsformel (14.2) ergibt

$$\underbrace{\dim(U+W)}_{2} = \underbrace{\dim(U)}_{1} + \underbrace{\dim(W)}_{1} - \underbrace{\dim(U \cap W)}_{0}.$$

(*ii*) Für die Unterräume

$$U = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \quad \text{und} \quad W = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

von $\mathbb{R}^3$ gilt

$$U + W = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^3 \quad \text{und} \quad U \cap W = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

Die Dimensionsformel (14.2) ergibt

$$\underbrace{\dim(U+W)}_{3} = \underbrace{\dim(U)}_{2} + \underbrace{\dim(W)}_{2} - \underbrace{\dim(U \cap W)}_{1}.$$

(*iii*) Für die Unterräume

$$U = \langle e_1, e_3, e_4, e_6 \rangle \quad \text{und} \quad W = \langle e_2, e_4, e_5 \rangle$$

des Vektorraumes V der Folgen mit Träger in $[\![1, 6]\!]$ über einem beliebigen Körper $(K, +, \cdot)$ gilt

$$U + W = \langle e_1, e_2, e_3, e_4, e_5, e_6 \rangle = V \quad \text{und} \quad U \cap W = \langle e_4 \rangle.$$

Die Dimensionsformel (14.2) ergibt

$$\underbrace{\dim(U+W)}_{6} = \underbrace{\dim(U)}_{4} + \underbrace{\dim(W)}_{3} - \underbrace{\dim(U \cap W)}_{1}. \qquad \triangle$$

Definition 14.5 (direkte Summe von zwei Unterräumen).
Es seien V ein Vektorraum und U, W zwei Unterräume von V. Die Summe $U + W$ heißt **direkt** (englisch: direct sum), wenn $U \cap W = \{0\}$ gilt. In dem Fall schreiben wir auch $U \oplus W$. △

Beispiel 14.6 (direkte Summe von zwei Unterräumen).
Von den Beispielen in Beispiel 14.4 ist nur

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = \mathbb{R}^2$$

eine direkte Summe. (**Quizfrage 14.2:** Woran erkennt man das?) △

Satz 14.7 (Charakterisierung direkter Summen von zwei Unterräumen)**.**
Es seien V ein Vektorraum und U, W zwei Unterräume von V. Dann sind äquivalent:

(i) $V = U \oplus W$.

(ii) Für alle $v \in V$ existieren eindeutige Vektoren $u \in U$ und $w \in W$, sodass $v = u + w$ gilt.

Sind U und W endlich-dimensional, dann sind diese Aussagen desweiteren äquivalent zu

(iii) $\dim(V) = \dim(U) + \dim(W)$ und $\dim(U \cap W) = 0$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): $V = U \oplus W$ heißt insbesondere $V = U + W$. Für gegebenes $v \in V$ gibt es also Vektoren $u_1 \in U$ und $w_1 \in W$, sodass $v = u_1 + w_1$ gilt. Gilt nun ebenfalls $v = u_2 + w_2$ für $u_2 \in U$ und $w_2 \in W$, dann gilt

$$u_1 - u_2 = w_2 - w_1 \in U \cap W = \{0\}$$

nach Voraussetzung. Daher muss $u_1 = u_2$ und $w_1 = w_2$ sein.

Aussage (ii) $\Rightarrow$ Aussage (i): Aus der Voraussetzung folgt sofort $V = U + W$. Zu zeigen ist $U \cap W = \{0\}$.

Für $v \in U \cap W$ folgt aus

$$\begin{aligned} v &= v + 0 && \text{mit } v \in U,\ 0 \in W \\ v &= 0 + v && \text{mit } 0 \in U,\ v \in W \end{aligned}$$

und der Eindeutigkeit der Zerlegung, dass $v = 0$ sein muss, also gilt tatsächlich $U \cap W = \{0\}$.

Aussage (i) $\Rightarrow$ Aussage (iii): Es gilt

$$\begin{aligned} \dim(V) &= \dim(U + W) && \text{da } V = U + W \text{ nach Voraussetzung} \\ &= \dim(U) + \dim(W) - \dim(U \cap W) && \text{nach Dimensionsformel (14.2)} \\ &= \dim(U) + \dim(W) - 0 && \text{da } U \cap W = \{0\} \text{ nach Voraussetzung.} \end{aligned}$$

Aussage (iii) $\Rightarrow$ Aussage (i): Es gilt

$$\begin{aligned} \dim(V) &= \dim(U) + \dim(W) && \text{nach Voraussetzung} \\ &= \dim(U + W) + \dim(U \cap W) && \text{nach Dimensionsformel (14.2)} \\ &= \dim(U + W) + 0 && \text{nach Voraussetzung.} \end{aligned}$$

$U + W$ ist also ein Unterraum von V maximaler Dimension und damit identisch zu V. Außerdem zeigt $\dim(U \cap W) = 0$, dass $U \cap W = \{0\}$ gilt, siehe Beispiel 13.16. □

Satz 14.8 (direkte Summe von zwei Unterräumen und disjunkte Zerlegung einer Basis)**.**
Es sei V ein Vektorraum. Dann gilt:

(i) Ist B eine Basismenge von V und bilden die Mengen B_1, B_2 eine disjunkte

(i) Ist $B = (v_i)_{i \in I}$ eine Basisfamilie von V und bilden die Mengen I_1, I_2 eine disjunkte Zerlegung von I, dann gilt $V = \langle B_1 \rangle \oplus \langle B_2 \rangle$ für die Teilfamilien

Zerlegung[43] von B in Teilmengen, dann gilt $V = \langle B_1 \rangle \oplus \langle B_2 \rangle$.

(ii) Sind U_1, U_2 Unterräume von V mit zugehörigen Basismengen B_1, B_2 und gilt $V = U_1 \oplus U_2$, so ist $B_1 \cup B_2$ eine Basismenge von V.

$B_1 \coloneqq (v_i)_{i \in I_1}$ und $B_2 \coloneqq (v_i)_{i \in I_2}$.

(ii) Sind U_1, U_2 Unterräume von V mit zugehörigen Basisfamilien B_1, B_2 und gilt $V = U_1 \oplus U_2$, so ist $B_1 \parallel B_2$ eine Basisfamilie on V.

Beweis. Wir führen den Beweis für Mengen.

Aussage (i): Wir zeigen zunächst $V = \langle B_1 \rangle + \langle B_2 \rangle$. Es gilt

$$\begin{aligned} V &= \langle B \rangle && B \text{ ist Basis von } V \\ &= \langle B_1 \cup B_2 \rangle && \text{nach Voraussetzung} \\ &= \big\langle \langle B_1 \rangle \cup \langle B_2 \rangle \big\rangle && \text{nach Folgerung 11.18} \\ &= \langle B_1 \rangle + \langle B_2 \rangle && \text{nach Lemma 14.1} \\ &\subseteq V. \end{aligned}$$

Damit gilt überall Gleichheit und insbesondere $\langle B_1 \rangle + \langle B_2 \rangle = V$.

Es bleibt $\langle B_1 \rangle \cap \langle B_2 \rangle = \{0\}$ zu zeigen. Nehmen wir also $\widetilde{v} \in \langle B_1 \rangle \cap \langle B_2 \rangle$ an, d. h. es gilt

$$\widetilde{v} = \sum_{v \in B_{1,0}} \alpha_v \, v = \sum_{v \in B_{2,0}} \beta_v \, v$$

für geeignete endliche Teilmengen $B_{1,0} \subseteq B_1$ und $B_{2,0} \subseteq B_2$ und Koeffizienten α_v, β_v. Es folgt

$$0 = \sum_{v \in B_{1,0}} \alpha_v \, v + \sum_{v \in B_{2,0}} (-\beta_v) \, v.$$

Da $B = B_1 \cup B_2$ eine Basis ist, ist auch die Teilmenge $B_{1,0} \cup B_{2,0}$ linear unabhängig. Wegen $B_1 \cap B_2 = \emptyset$ gilt auch $B_{1,0} \cap B_{2,0} = \emptyset$, also sind die beteiligten Vektoren paarweise verschieden. Daraus folgt $\alpha_v = 0$ für alle $v \in B_{1,0}$ und $\beta_v = 0$ für alle $v \in B_{2,0}$. Das heißt aber $\widetilde{v} = 0$ und damit $\langle B_1 \rangle \cap \langle B_2 \rangle = \{0\}$.

Aussage (ii): Es seien nun U_1, U_2 Unterräume von V mit Basismengen B_1, B_2. Wir nehmen $V = U_1 \oplus U_2$ an. Wir müssen zeigen, dass $B_1 \cup B_2$ linear unabhängig ist und ganz V erzeugt. Letzteres folgt aus

$$\begin{aligned} \langle B_1 \cup B_2 \rangle &= \big\langle \langle B_1 \rangle \cup \langle B_2 \rangle \big\rangle && \text{nach Folgerung 11.18} \\ &= \langle U_1 \cup U_2 \rangle && \text{da } B_1 \text{ Basis von } U_1 \text{ und } B_2 \text{ Basis von } U_2 \text{ ist} \\ &= U_1 + U_2 && \text{nach Lemma 14.1} \\ &= V && \text{nach Voraussetzung.} \end{aligned}$$

Um die lineare Unabhängigkeit von $B_1 \cup B_2$ zu zeigen, sei nun $B_0 \subseteq B_1 \cup B_2$ eine endliche Teilmenge. Wir können B_0 als disjunkte Zerlegung $B_0 = B_{1,0} \mathbin{\dot\cup} B_{2,0}$ schreiben, wobei $B_{1,0} \subseteq B_1$ und $B_{2,0} \subseteq B_2$ endliche Teilmengen sind. Wir betrachten nun die Linearkombination

$$\sum_{v \in B_{1,0}} \alpha_v \, v + \sum_{v \in B_{2,0}} \beta_v \, v = 0 \quad \text{bzw.} \quad u_1 \coloneqq \sum_{v \in B_{1,0}} \alpha_v \, v = \sum_{v \in B_{2,0}} (-\beta_v) \, v \eqqcolon u_2.$$

[43] Wir benutzen nicht das Wort „Partition“ (Definition 5.24), da $B_1 = \emptyset$ oder $B_2 = \emptyset$ erlaubt ist.

Dabei gehört u_1 zu U_1 und u_2 zu U_2. Wegen $U_1 \cap U_2 = \{0\}$ folgt $u_1 = u_2 = 0$, beide Linearkombinationen ergeben also den Nullvektor. Da aber B_1 und B_2 als Basen linear unbhängig sind, folgt weiter $\alpha_v = 0$ für alle $v \in B_{1,0}$ und $\beta_v = 0$ für alle $v \in B_{2,0}$. Damit ist gezeigt, dass $B_1 \cup B_2$ in der Tat linear unabhängig ist. □

Wir halten die Situation aus Satz 14.8, bei der zwei Unterräume in direkter Summe den ganzen Raum ergeben, in folgender Definition fest:

Definition 14.9 (komplementärer Unterraum, Kodimension)**.**
Es seien V ein Vektorraum und U ein Unterraum von V.

(i) Ein Unterraum W von V heißt ein **zu U komplementärer Unterraum** (englisch: complementary subspace) oder ein **Komplement** (englisch: complement) von U in V, wenn $V = U \oplus W$ gilt.

(ii) Die Dimension $\dim(W)$ eines zu U komplementären Unterraumes W heißt die **Kodimension** (englisch: codimension) von U in V, kurz: $\operatorname{codim}(U)$. △

Beachte: Komplementäre Unterräume eines Vektorraumes sind i. A. nicht eindeutig. Im Fall $U = V$ jedoch ist der einzige zu U komplementäre Unterraum der Nullraum $\{0\}$, und im Fall $U = \{0\}$ ist der einzige zu U komplementäre Unterraum der Raum V selbst.

Beachte: Komplementarität ist eine symmetrische Relation: Ein Unterraum W von V ist genau dann zu U komplementär, wenn U zu W komplementär ist.

Beachte: Die Kodimension von U in V ist wohldefiniert$^{\text{AoC}}$, denn sind W_1 und W_2 zwei zu U komplementäre Unterräume von V, dann gibt es nach dem Basisexistenzsatz 13.6$^{\text{AoC}}$ zugehörige Basisfamilien B_0 von U bzw. B_1 von W_1 und B_2 von W_2, sodass die Konkatenation $B_0 \parallel B_1$ bzw. $B_0 \parallel B_2$ jeweils eine Basisfamilie von V ist (Satz 14.8). Da alle Basen von V gleichmächtig sind (Bemerkung 13.15$^{\text{AoC}}$), folgt, dass B_1 und B_2 gleichmächtig sind.

Folgerung 14.10 (Existenz eines komplementären Unterraumes$^{\text{AoC}}$)**.**
Es seien V ein Vektorraum und U ein Unterraum von V. Dann existiert ein weiterer Unterraum W von V, sodass $V = U \oplus W$ gilt. Es gilt[44]

$$\dim(V) = \dim(U) + \dim(W) = \dim(U) + \operatorname{codim}(U). \tag{14.3}$$

Beweis. Es sei $B_U := (v_i)_{i \in I_U}$ eine Basisfamilie von U. Aus dem Basisergänzungssatz 13.5 folgt die Existenz einer Oberfamilie $B := (v_i)_{i \in I_B}$, die eine Basis von V ist. Dann ist die Familie $B_W := (v_i)_{i \in I_B \setminus I_U}$ linear unabhängig, und $W := \langle B_W \rangle$ ist nach Satz 14.8 ein Unterraum mit der gesuchten Eigenschaft.

Die Dimensionsformel (14.3) folgt direkt aus der Addition der Kardinalitäten der Indexmengen dieser Basen. □

Folgende Konstellationen sind für komplementäre Unterräume $U \oplus W = V$ möglich:

[44]Hierbei gilt $\infty + k = k + \infty = \infty + \infty = \infty$ für $k \in \mathbb{N}_0$.

$\dim(U)$	$\dim(W)$	$\dim(V)$
endlich	endlich	endlich
∞	endlich	∞
endlich	∞	∞
∞	∞	∞

In endlich-dimensionalen Räumen können wir eine Version von Folgerung 14.10 formulieren, deren Beweis ohne das Auswahlaxiom auskommt:

Folgerung 14.11 (Kodimension in endlich-dimensionalen Vektorräumen).
Es seien V ein endlich-dimensionaler Vektorraum und U ein Unterraum von V. Dann existiert ein weiterer Unterraum W von V, sodass $V = U \oplus W$ gilt. Es gilt

$$\operatorname{codim}(U) = \dim(V) - \dim(U). \tag{14.4}$$

Beweis. Einen zu U komplementären Unterraum W können wir wie in Bemerkung 13.9 ohne Verwendung des Auswahlaxioms durch Ergänzung einer Basis von U zu einer Basis von V konstruieren. Nach Satz 14.7 gilt

$$\dim(V) = \dim(U) + \dim(W) = \dim(U) + \operatorname{codim}(U). \qquad \square$$

Beispiel 14.12 (komplementärer Unterraum).

(*i*) In $\mathbb{R}^2$ ist jeder eindimensionale Unterraum W, der nicht identisch mit $U := \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$ ist, ein Komplement von U in $\mathbb{R}^2$.

(*ii*) In $\mathbb{R}^3$ ist jeder zweidimensionale Unterraum W, der den Unterraum $U = \left\langle \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\rangle$ nicht enthält, ein Komplement von U in $\mathbb{R}^3$.

(*iii*) Wir betrachten den Folgenraum $K^{\mathbb{N}}$ über einem Körper K und den Unterraum $U = (K^{\mathbb{N}})_{00}$ der endlich getragenen Folgen. Die zu U komplementären Unterräume haben keine einfache Beschreibung. (**Quizfrage 14.3:** Warum bilden — was man annehmen könnte — die nicht endlich getragenen Folgen plus die Nullfolge keinen solchen komplementären Unterraum?) △

§ 14.2 Summen von Familien von Unterräumen

Wir beschäftigen uns abschließend in Erweiterung von § 14.1 noch mit Summen von einer beliebigen Anzahl von Unterräumen eines Vektorraumes.

Definition 14.13 (Summe einer Familie von Unterräumen).
Es seien V ein Vektorraum und $(U_i)_{i \in I}$ eine nichtleere Familie von Unterräumen von V.

(*i*) Der Unterraum
$$\sum_{i \in I} U_i \coloneqq \Big\langle \bigcup_{i \in I} U_i \Big\rangle \tag{14.5}$$
heißt die **Summe der Familie von Unterräumen** $(U_i)_{i \in I}$ (englisch: sum of a family of subspaces). Im Fall $I = [\![1, n]\!]$ schreiben wir auch $U_1 + \cdots + U_n$ oder $\sum_{i=1}^n U_i$.

(*ii*) Die Summe (14.5) heißt **direkt** (englisch: direct sum), wenn gilt:
$$U_j \cap \sum_{i \in I \setminus \{j\}} U_i = \{0\} \quad \text{für alle } j \in I. \tag{14.6}$$
In dem Fall schreiben wir auch $\bigoplus_{i \in I} U_i$ und speziell im Fall $I = [\![1, n]\!]$ auch $U_1 \oplus \cdots \oplus U_n$ oder $\bigoplus_{i=1}^n U_i$. △

Beispiel 14.14 (Summe einer Familie von Unterräumen)**.**

(*i*) Für den Standardvektorraum K^n über einem Körper K mit den Basisvektoren e_i, $i = 1, \ldots, n \in \mathbb{N}_0$, siehe Beispiel 13.2, gilt
$$K^n = \bigoplus_{i=1}^n \langle e_i \rangle.$$

(*ii*) Für den Raum der endlich getragenen Folgen $(K^{\mathbb{N}})_{00}$ über einem Körper K gilt
$$(K^{\mathbb{N}})_{00} = \bigoplus_{j \in \mathbb{N}} \langle e_j \rangle. \qquad △$$

Bemerkung 14.15 (Summe einer Familie von Unterräumen)**.**

(*i*) Die Summe einer nichtleeren Familie $(U_i)_{i \in I}$ von Unterräumen hat die folgenden Darstellungen:
$$\sum_{i \in I} U_i = \bigcup \Big\{ \sum_{i \in I_0} U_i \,\Big|\, I_0 \subseteq I \text{ ist eine endliche Teilmenge} \Big\} \tag{14.7a}$$
$$= \Big\{ \sum_{i \in I_0} u_i \,\Big|\, I_0 \subseteq I \text{ ist eine endliche Teilmenge und } u_i \in U_i \text{ für alle } i \in I_0 \Big\}. \tag{14.7b}$$
Die Elemente von $\sum_{i \in I} U_i$ haben also jeweils eine Darstellung der Form $u_{i_1} + u_{i_2} + \cdots + u_{i_n}$ für $i_j \in I$, $j = 1, \ldots, n \in \mathbb{N}_0$. (**Quizfrage 14.4:** Ist klar, dass die rechte Seite von (14.7a) überhaupt ein Unterraum ist?)

(*ii*) Im Fall $\#I = 2$ stimmt die Bedingung (14.6) für die Direktheit der Summe einer Familie von Unterräumen überein mit Definition 14.5. Im Fall $\#I > 2$ reicht es jedoch für die Direktheit der Summe nicht aus, dass anstelle von (14.6) nur paarweise $U_i \cap U_j = \{0\}$ für $i \neq j$ gefordert wird.

Betrachte zum Beispiel die Unterräume
$$U_1 = \Big\langle \big(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\big) \Big\rangle, \quad U_2 = \Big\langle \big(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\big) \Big\rangle \quad \text{und} \quad U_3 = \Big\langle \big(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\big) \Big\rangle.$$
Dann gilt $U_i \cap U_j = \{0\}$ für alle $i \neq j$, aber $U_1 \cap (U_2 + U_3) = U_1 \cap \mathbb{R}^2 \supsetneq \{0\}$. Das heißt, die Summe der Unterräume U_1, U_2, U_3 ist nicht direkt. △

Satz 14.16 (Charakterisierung direkter Summen von Familien von Unterräumen, vgl. Satz 14.7)**.** Es seien V ein Vektorraum und $(U_i)_{i \in I}$ eine nichtleere Familie von Unterräumen von V. Dann sind äquivalent:

(i) $V = \bigoplus_{i \in I} U_i$.

(ii) Für alle $v \in V$ existiert eine endliche Teilmenge $I_0 \subseteq I$ und Vektoren $u_i \in U_i$, sodass $v = \sum_{i \in I_0} u_i$ gilt, und diese Darstellung ist (bis auf die Summation von Nullvektoren) eindeutig.

Beweis. Der Beweis ist Gegenstand der Übung. □

Satz 14.17 (direkte Summe von Unterräumen und disjunkte Zerlegung einer Basis, vgl. Satz 14.8)**.**

Es sei V ein Vektorraum. Dann gilt:

(i) Ist B eine Basismenge von V und bilden die Mengen B_i, $i \in I$, eine disjunkte Zerlegung von B in Teilmengen mit nichtleerer Indexmenge I, dann gilt $V = \bigoplus_{i \in I} \langle B_i \rangle$.

(ii) Ist $(U_i)_{i \in I}$ eine nichtleere Familie von Unterräumen von V mit Basismengen B_i, $i \in I$, und gilt $V = \bigoplus_{i \in I} U_i$, so ist $\bigcup_{i \in I} B_i$ eine Basismenge von V.

(i) Ist $B = (v_i)_{i \in I}$ eine Basisfamilie von V und bilden die Mengen I_j, $j \in J$, eine disjunkte Zerlegung von I mit nichtleerer Indexmenge J, dann gilt $V = \bigoplus_{j \in J} \langle B_j \rangle$ für die Teilfamilien $B_j \coloneqq (v_i)_{i \in I_j}$.

(ii) Ist $(U_i)_{i \in I}$ eine nichtleere Familie von Unterräumen von V mit Basisfamilien B_i, $i \in I$, und gilt $V = \bigoplus_{i \in I} U_i$, so ist die Konkatenation $\|_{i \in I} B_i$ der Familien B_i, $i \in I$, eine Basisfamilie von V.

Beweis. Der Beweis ist Gegenstand der Übung. □

Ende der Vorlesung 18

Ende der Woche 9

Kapitel 4 Matrizen und lineare Abbildungen

§ 15 Matrizen

Literatur: Fischer, Springborn, 2020, Kapitel 3.7; Bosch, 2014, Kapitel 3; Deiser, 2024b, Kapitel 3.3; Jänich, 2008, Kapitel 4.2 und 5.1–5.5

Matrizen sind ein universelles Mittel zur Darstellung verschiedener Sachverhalte, beispielsweise zur Beschreibung von Graphen, zur Erfassung von Rohstoffbedarfen in der Produktionsplanung, zur Darstellung chemischer Reaktionen und zur ersten Formulierung der Quantenmechanik in der Physik. Wir werden Matrizen in erster Linie zur Beschreibung linearer Abbildungen verwenden, das sind die Homomorphismen zwischen Vektorräumen.[1] Diese Bedeutung stellen wir aber bis § 19 zurück und betrachten Matrizen zunächst als eigenständiges Thema.

Definition 15.1 (Matrix)**.**
Es seien $(K, +, \cdot)$ ein Körper und $m, n \in \mathbb{N}_0$.

(*i*) Eine **Matrix** (englisch: matrix) **der Dimension** $n \times m$ (sprich: „n mal m" oder „n Kreuz m") oder eine **$n \times m$-Matrix A über dem Körper K** ist eine endliche, doppelt indizierte Familie in K mit der Indexmenge $[\![1, n]\!] \times [\![1, m]\!]$.[2] Wir schreiben sie in der Form[3]

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}. \tag{15.1}$$

Die Menge aller $n \times m$-Matrizen über K wird mit $K^{n \times m}$ bezeichnet.[4]

(*ii*) Die Indizes $(i, j) \in [\![1, n]\!] \times [\![1, m]\!]$ mit der Eigenschaft $j - i = k \in \mathbb{Z}$ bilden die **k-te Diagonale** (englisch: k-th diagonal) der Matrix. Die 0-te Diagonale heißt auch die **Hauptdiagonale** (englisch: main diagonal), die anderen Diagonalen heißen die **Nebendiagonalen** (englisch: off-diagonals) der Matrix.

(*iii*) Eine $n \times m$-Matrix heißt eine **Diagonalmatrix** (englisch: diagonal matrix), wenn alle Einträge außerhalb der Hauptdiagonale gleich 0 sind.[5]

[1] Obwohl nicht sofort offensichtlich, dienen Matrizen auch bei den genannten Sachverhalten zur Darstellung linearer Abbildungen.
[2] Wir können daher eine Matrix auch als eine Abbildung $[\![1, n]\!] \times [\![1, m]\!] \to K$ auffassen.
[3] Alternativ können statt der eckigen auch runde Klammern verwendet werden.
[4] Alternative Bezeichnungen sind $K^{n,m}$ oder $M_{n,m}(K)$.
[5] Man sagt auch, dass bei einer Diagonalmatrix die Nebendiagonalen „nicht besetzt" sind, d. h., dass dort nur Nullen stehen.

(*iv*) Eine Matrix heißt **quadratisch** (englisch: square matrix, quadratic matrix), wenn $m = n$ gilt.

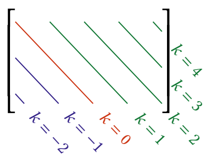
(*v*) Die quadratische $n \times n$-Diagonalmatrix

$$I_{n\times n} \coloneqq \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{15.2}$$

heißt die $n \times n$-**Einheitsmatrix** (englisch: identity matrix). Wir bezeichnen sie auch mit I_n oder einfach mit I, wenn die Dimension klar oder unerheblich ist. △

Beachte: Wir lassen explizit zu, dass eine oder beide Dimensionen einer Matrix gleich 0 sind. Das ist aus vielerlei Gründen praktisch. Eine Matrix der Dimension $n \times 0$ oder $0 \times m$ hat keine Elemente, besitzt aber dennoch ihre spezifische Form. Es gibt nur eine einzige Matrix der Dimension $n \times 0$ bzw. der Dimension $0 \times m$.

Wir illustrieren die Lage der Hauptdiagonale, der oberen Nebendiagonalen ($k > 0$) sowie der unteren Nebendiagonalen ($k < 0$) am Beispiel einer 3×5-Matrix:



Matrizen werden häufig mit lateinischen Großbuchstaben bezeichnet und ihre Elemente mit den zugehörigen Kleinbuchstaben, zum Beispiel

$$A = (a_{ij})_{i=1,\ldots,n,\; j=1,\ldots,m} \quad \text{oder} \quad A = (a_{ij})_{1\leqslant i\leqslant n,\; 1\leqslant j\leqslant m} \tag{15.3}$$

oder einfach $A = (a_{ij})$, wenn die Dimension klar ist.[6]

Der erste Index (häufig i) heißt der **Zeilenindex** (englisch: row index). Die i-te **Zeile** (englisch: row) von $A = (a_{ij}) \in K^{n\times m}$ ist die (einfach indizierte) Familie bzw. der Zeilenvektor

$$a_{i\bullet} \coloneqq \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{im} \end{pmatrix} \in K_m, \tag{15.4}$$

vgl. Beispiel 11.3.

Der zweite Index (häufig j) heißt der **Spaltenindex** (englisch: column index). Die j-te **Spalte** (englisch: column) von $A = (a_{ij}) \in K^{n\times m}$ ist die (einfach indizierte) Familie bzw. der Spaltenvektor

$$a_{\bullet j} \coloneqq \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \in K^n, \tag{15.5}$$

[6] Aus Gründen der Verdeutlichung schreiben wir manchmal auch $a_{i,j}$ anstelle von a_{ij}.

vgl. nochmals Beispiel 11.3. Wir werden $1 \times m$-Matrizen mit Zeilenvektoren und $n \times 1$-Matrizen mit Spaltenvektoren identifizieren.

Auf der Menge $K^{n \times m}$ der $n \times m$-Matrizen definieren wir komponentenweise die Addition und die S-Multiplikation wie folgt:

Definition 15.2 (Addition, S-Multiplikation von Matrizen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $m, n \in \mathbb{N}_0$. Auf der Menge der Matrizen $K^{n \times m}$ sind die innere Verknüpfung $+\colon K^{n \times m} \times K^{n \times m} \to K^{n \times m}$ (**Addition**[7]) und die äußere Verknüpfung $\cdot\colon K \times K^{n \times m} \to K^{n \times m}$ (**S-Multiplikation**[8]) durch

$$(A + B)_{ij} \coloneqq a_{ij} + b_{ij} \tag{15.6a}$$

$$(\alpha \cdot A)_{ij} \coloneqq \alpha \cdot a_{ij} \tag{15.6b}$$

für $A, B \in K^{n \times m}$ und $\alpha \in K$ erklärt. △

Wir werden das Zeichen $\cdot$ für die skalare Multiplikation in der Regel weglassen, vgl. Bemerkung 11.13.

Satz 15.3 ($(K^{n \times m}, +, \cdot)$ ist ein Vektorraum)**.**
Es seien $(K, +, \cdot)$ ein Körper und $m, n \in \mathbb{N}_0$. Dann bilden die $n \times m$-Matrizen mit den Verknüpfungen (15.6) einen K-Vektorraum. Dieser besitzt die Dimension $n\,m$, und die Matrizen

$$\{E_{11}, \dots, E_{nm}\} \text{ mit } E_{ij} \coloneqq \begin{array}{@{}cl@{}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{array}{l} \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \end{array} \\ \qquad\quad \uparrow j\text{-te Spalte} & \end{array} = (\delta_{ik}\,\delta_{j\ell})_{k=1,\dots,n,\ \ell=1,\dots,m} \tag{15.7}$$

bilden eine Basis, genannt die **Standardbasis** (englisch: standard basis) von $K^{n \times m}$.

Beweis. Die Eigenschaften eines Vektorraumes (Definition 11.1) sind leicht nachzurechnen:[9] $(K^{n \times m}, +)$ ist eine abelsche Gruppe, da $(K, +)$ eine abelsche Gruppe ist. Die Distributivgesetze der S-Multiplikation $\alpha\,(A + B) = \alpha\,A + \alpha\,B$ und $(\alpha + \beta)\,A = \alpha\,A + \beta\,A$ gelten wegen des Distributivgesetzes im Körper $(K, +, \cdot)$. Das gemischte Assoziativgesetz $(\alpha\,\beta)\,A = \alpha\,(\beta\,A)$ gilt wegen der Assoziativität der Multiplikation im Körper $(K, +, \cdot)$. Schließlich ist $1\,A = A$.

Die genannten Matrizen E_{ij} bilden nach Satz 13.3 eine Basis, da jede beliebige Matrix $A \in K^{n \times m}$ auf eindeutige Weise als Linearkombination darstellbar ist, nämlich als

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} \underbrace{a_{ij}}_{\text{Koeffizient}} E_{ij}. \tag{15.8}$$

Die Dimension von $K^{n \times m}$ ergibt sich aus der Anzahl $n\,m$ der Basiselemente E_{ij}. □

[7] Die Bezeichnung ist dieselbe wie die der „Addition“ im Körper K.

[8] Auch hier ist die Bezeichnung dieselbe wie die der „Multiplikation“ im Körper K.

[9] Wir könnten auch argumentieren, dass die Verknüpfungen (15.6) nichts anderes sind als die punktweise Addition und S-Multiplikation im Vektorraum der Abbildungen $[\![1, n]\!] \times [\![1, m]\!] \to K$, siehe Beispiel 11.3.

Bemerkung 15.4 (zum Vektorraum $K^{n\times m}$).

(i) Das neutrale Element bzgl. der Addition ist die **Nullmatrix** (englisch: zero matrix)

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in K^{n\times m}.$$

(ii) Der Vektorraum der 1×1-Matrizen über K ist ein eindimensionaler K-Vektorraum. Dieser kann identifiziert werden mit K selbst.[10]

(iii) Die Vektorräume der $0\times m$- und der $n\times 0$-Matrizen sind nulldimensionale K-Vektorräume (Nullräume), da sie nur aus einem Element bestehen. △

§ 15.1 Matrix-Matrix-Multiplikation

Für Matrizen passender Dimensionen können wir eine Matrix-Matrix-Multiplikation einführen. Diese ist von zentraler Bedeutung im Umgang mit Matrizen.

Definition 15.5 (Matrix-Matrix-Multiplikation).

Es seien K ein Körper und $m, n, \ell \in \mathbb{N}_0$. Für Matrizen passender Dimensionen ist die **Matrix-Matrix-Multiplikation** (englisch: matrix-matrix multiplication) oder kurz **Matrix-Multiplikation** (englisch: matrix multiplication) wie folgt definiert:

$$\cdot\colon K^{n\times m} \times K^{m\times \ell} \to K^{n\times \ell}. \tag{15.9a}$$

Dabei gilt für die Matrizen $A \in K^{n\times m}$, $B \in K^{m\times \ell}$ und ihr Matrix-Matrix-Produkt $C \coloneqq A \cdot B \in K^{n\times \ell}$:

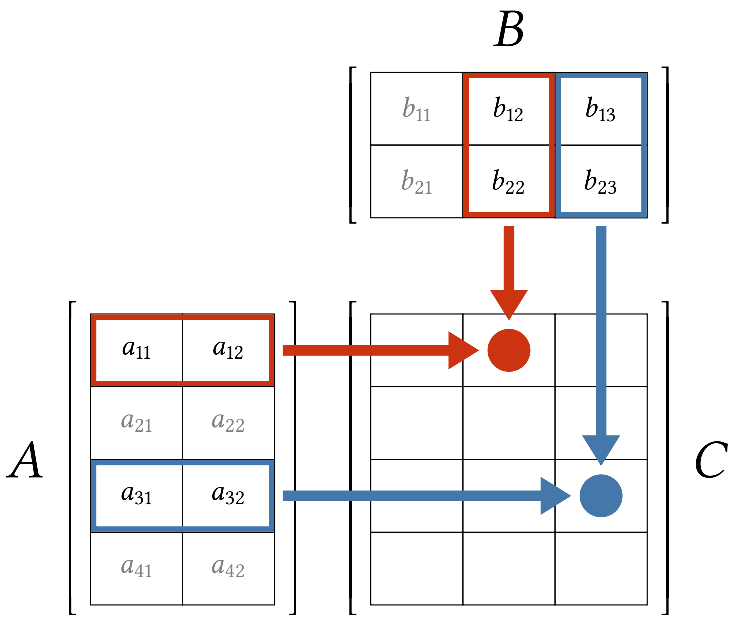
$$c_{ik} \coloneqq \sum_{j=1}^{m} a_{ij} \cdot b_{jk} \quad \text{für } 1 \leqslant i \leqslant n \text{ und } 1 \leqslant k \leqslant \ell. \tag{15.9b}$$

△

Beachte: Die Summation verwendet die „Addition“ + aus dem Körper K, und jeder Summand ergibt sich aus der „Multiplikation“ $\cdot$ zweier Elemente von K. Im Fall $m = 0$ sind die Summen in (15.9b) alle leer, ergeben also das Nullelement des Körpers K. Daher ist das Produkt einer $n \times 0$- und einer $0 \times m$-Matrix die $n \times m$-Nullmatrix.

Den beispielhaften Fall der Matrix-Multiplikation einer 4×2-Matrix A mit einer 2×3-Matrix B können wir wie folgt grafisch darstellen:

[10] Mit Hilfe des Begriffs der Isomorphie von Vektorräumen (Definition 17.1) können wir das später präzisieren.



$$c_{12} \coloneqq a_{11} \cdot b_{12} + a_{12} \cdot b_{22}$$

$$c_{33} \coloneqq a_{31} \cdot b_{13} + a_{32} \cdot b_{23}$$

Beispiel 15.6 (Matrix-Multiplikation).
Wir multiplizieren eine 4×2-Matrix mit einer 2×3-Matrix über dem Körper $\mathbb{Q}$:[11]

$$\begin{aligned}
&\begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 & -4 \\ 0 & -2 & 8 \end{bmatrix} \\
&= \begin{bmatrix} 1 \cdot 5 + (-3) \cdot 0 & 1 \cdot 2 + (-3) \cdot (-2) & 1 \cdot (-4) + (-3) \cdot 8 \\ 2 \cdot 5 + 4 \cdot 0 & 2 \cdot 2 + 4 \cdot (-2) & 2 \cdot (-4) + 4 \cdot 8 \\ 5 \cdot 5 + 0 \cdot 0 & 5 \cdot 2 + 0 \cdot (-2) & 5 \cdot (-4) + 0 \cdot 8 \\ (-3) \cdot 5 + (-6) \cdot 0 & (-3) \cdot 2 + (-6) \cdot (-2) & (-3) \cdot (-4) + (-6) \cdot 8 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 8 & -28 \\ 10 & -4 & 24 \\ 25 & 10 & -20 \\ -15 & 6 & -36 \end{bmatrix}.
\end{aligned}$$

△

Bemerkung 15.7 (Matrix-Multiplikation).

(*i*) Die Matrix-Matrix-Multiplikation (15.9) ist (außer im Fall $n = \ell = m$) keine innere Verknüpfung eines Vektorraumes von Matrizen, da die Dimensionen der beiden Faktoren und des Produkts verschieden sind.

[11] Wir würden dasselbe Ergebnis erhalten, wenn wir die Matrizen als Matrizen über $\mathbb{R}$ oder über $\mathbb{C}$ auffassen.

(*ii*) $A \cdot B$ ist genau dann definiert, wenn die Anzahl der Spalten des linken Faktors A übereinstimmt mit der Anzahl der Zeilen des rechten Faktors B.

(*iii*) Das Produkt $A \cdot B$ hat soviele Zeilen wie der linke Faktor A und soviele Spalten wie der rechte Faktor B. Die gemeinsame „mittlere Dimension“ ist nach der Bildung des Produkts $A \cdot B$ nicht mehr sichtbar.

(*iv*) Der Eintrag

$$c_{ik} = \sum_{j=1}^{m} a_{ij} \cdot b_{jk}$$

für Zeile i und Spalte k im Produkt $C = A \cdot B$ verwendet nur Informationen aus der i-ten Zeile $a_{i\bullet}$ des linken Faktors A und aus der k-ten Spalte $b_{\bullet k}$ des rechten Faktors B.

(*v*) Verantwortlich für die Position (i, k) im Ergebnis ist der Index i der Zeile des linken Faktors und der Index k der Spalte des rechten Faktors. △

Bei näherer Betrachtung ergeben sich zwei weitere Sichtweisen auf das Produkt $A \cdot B$:

(1) Die Spalten von $C = A \cdot B$ sind Linearkombinationen der Spalten des linken Faktors A. Beispielsweise ist die k-te Spalte von C gerade

$$c_{\bullet k} = \sum_{j=1}^{m} \underbrace{a_{\bullet j}}_{j\text{-te Spalte}} \cdot \underbrace{b_{jk}}_{\text{Koeffizient}}$$

Die Koeffizienten der Linearkombination stehen dabei in der k-ten Spalte von B.

In Beispiel 15.6 ist die erste Spalte des Ergebnisses gerade die Linearkombination „5 mal die erste Spalte von A plus 0 mal die zweite Spalte von A“:

$$\begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 & -4 \\ 0 & -2 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 8 & -28 \\ 10 & -4 & 24 \\ 25 & 10 & -20 \\ -15 & 6 & -36 \end{bmatrix}.$$

(2) Die Zeilen von $C = A \cdot B$ sind Linearkombinationen der Zeilen des rechten Faktors B. Beispielsweise ist die i-te Zeile von C gerade

$$c_{i\bullet} = \sum_{j=1}^{m} \underbrace{a_{ij}}_{\text{Koeffizient}} \cdot \underbrace{b_{j\bullet}}_{j\text{-te Zeile}}$$

Die Koeffizienten der Linearkombination stehen dabei in der i-ten Zeile von A.

In Beispiel 15.6 ist die zweite Zeile des Ergebnisses gerade die Linearkombination „2 mal die erste Zeile von B plus 4 mal die zweite Zeile von B“:

$$\begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 & -4 \\ 0 & -2 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 8 & -28 \\ 10 & -4 & 24 \\ 25 & 10 & -20 \\ -15 & 6 & -36 \end{bmatrix}.$$

Lemma 15.8 (Eigenschaften der Matrix-Multiplikation)**.**
Es seien K ein Körper und $m, n, p, q \in \mathbb{N}_0$. Für Matrizen $A, A_1, A_2 \in K^{n\times m}$ und $B, B_1, B_2 \in K^{m\times p}$ sowie $C \in K^{p\times q}$ und Skalare $\alpha \in K$ gelten die folgenden Eigenschaften:

$$A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2 \quad \textbf{Distributivgesetz}^{12} \tag{15.10a}$$

$$(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B \quad \textbf{Distributivgesetz} \tag{15.10b}$$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \quad \textbf{Assoziativgesetz}^{13} \tag{15.11}$$

$$A \cdot (\alpha \cdot B) = (\alpha \cdot A) \cdot B = \alpha \cdot (A \cdot B) \quad \text{„Skalare können überall stehen“} \tag{15.12}$$

$$I_n \cdot A = A \cdot I_m = A \quad \text{Neutralität der Einheitsmatrix.} \tag{15.13}$$

Beweis. Wir führen den Nachweis durch Vergleich der Einträge der Matrizen:

$$\begin{aligned}\big[A \cdot (B_1 + B_2)\big]_{ik} &= \sum_{j=1}^m a_{ij}\,(b_1 + b_2)_{jk} = \sum_{j=1}^m a_{ij}\,(b_1)_{jk} + \sum_{j=1}^m a_{ij}\,(b_2)_{jk} \\ &= \big[A \cdot B_1\big]_{ik} + \big[A \cdot B_2\big]_{ik} = \big[A \cdot B_1 + A \cdot B_2\big]_{ik}\end{aligned}$$

zeigt (15.10a). Analog folgt (15.10b). Die Rechnung

$$\begin{aligned}\big[(A \cdot B) \cdot C\big]_{i\ell} &= \sum_{j=1}^p (A \cdot B)_{ij}\, c_{j\ell} = \sum_{j=1}^p \sum_{k=1}^m (a_{ik}\, b_{kj})\, c_{j\ell} \\ &= \sum_{k=1}^m \sum_{j=1}^p a_{ik}\,(b_{kj}\, c_{j\ell}) = \sum_{k=1}^m a_{ik}\,\big[B \cdot C\big]_{k\ell} = \big[A \cdot (B \cdot C)\big]_{i\ell}\end{aligned}$$

zeigt (15.11). Weiter gilt

$$\begin{aligned}\big[A \cdot (\alpha \cdot B)\big]_{ik} &= \sum_{j=1}^m a_{ij}\,(\alpha \cdot B)_{jk} = \sum_{j=1}^m a_{ij}\,(\alpha\, b_{jk}) = \sum_{j=1}^m (\alpha\, a_{ij})\, b_{jk} = \big[(\alpha \cdot A) \cdot B\big]_{ik} \\ &= \alpha \sum_{j=1}^m a_{ij}\, b_{jk} = \alpha\,\big[A \cdot B\big]_{ik},\end{aligned}$$

also (15.12). Schließlich haben wir

$$\big[I_n \cdot A\big]_{ik} = \sum_{j=1}^m (I_n)_{ij}\, a_{jk} = \sum_{j=1}^m \delta_{ij}\, a_{jk} = a_{ik}$$

$$\text{und} \quad \big[A \cdot I_m\big]_{ik} = \sum_{j=1}^m a_{ij} (I_m)_{jk} = \sum_{j=1}^m a_{ij}\, \delta_{jk} = a_{ik},$$

also (15.13). □

[12] englisch: distributive law
[13] englisch: associative law

Auch bei der Matrix-Multiplikation werden wir in Zukunft das Multiplikationszeichen $\cdot$ in der Regel weglassen.

Bemerkung 15.9 (Matrix-Vektor-Multiplikation)**.**
Ein wichtiger Spezialfall der Matrix-Matrix-Multiplikation ist die **Matrix-Vektor-Multiplikation** (englisch: matrix-vector multiplication) $A\,x$, wobei $A \in K^{n \times m}$ ist und der Spaltenvektor $x \in K^m$ als $m \times 1$-Matrix aufgefasst wird. Beispielsweise gilt über dem Körper $K = \mathbb{Q}$

$$\begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 8 \\ -4 \\ 10 \\ 6 \end{pmatrix}.$$

Ein zweiter Spezialfall ist die **Vektor-Matrix-Multiplikation** (englisch: vector-matrix multiplication) $y\,A$, wobei $A \in K^{n \times m}$ ist und der Zeilenvektor $y \in K_n$ als $1 \times n$-Matrix aufgefasst wird. Beispielsweise gilt

$$\begin{pmatrix} 2 & 0 & 1 & -1 \end{pmatrix} \begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} = \begin{pmatrix} 10 & 0 \end{pmatrix}.$$ △

Ende der Vorlesung 19

§ 15.2 Zeilen- und Spaltenraum

Definition 15.10 (Zeilen- und Spaltenraum, Zeilen- und Spaltenrang)**.**
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in \mathbb{R}^{n \times m}$.

(i) Die lineare Hülle der Zeilenvektoren $a_{1\bullet}, \ldots, a_{n\bullet} \in K_m$ heißt der **Zeilenraum** (englisch: row space) von A:

$$\operatorname{ZR}(A) \coloneqq \langle a_{1\bullet}, \ldots, a_{n\bullet} \rangle \subseteq K_m. \tag{15.14a}$$

Die Dimension von $\operatorname{ZR}(A)$ heißt der **Zeilenrang** von A (englisch: row rank), also

$$\operatorname{ZRang}(A) \coloneqq \dim(\operatorname{ZR}(A)). \tag{15.14b}$$

(ii) Die lineare Hülle der Spaltenvektoren $a_{\bullet 1}, \ldots, a_{\bullet m} \in K^n$ heißt der **Spaltenraum** (englisch: column space) von A:

$$\operatorname{SR}(A) \coloneqq \langle a_{\bullet 1}, \ldots, a_{\bullet m} \rangle \subseteq K^n. \tag{15.15a}$$

Die Dimension von $\operatorname{SR}(A)$ heißt der **Spaltenrang** von A (englisch: column rank), also

$$\operatorname{SRang}(A) \coloneqq \dim(\operatorname{SR}(A)). \tag{15.15b}$$

△

Lemma 15.11 (Zeilenrang und Spaltenrang).
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n \times m}$. Dann gilt:

(i) Der Zeilenrang $\operatorname{ZRang}(A)$ ist gleich der maximalen Anzahl linearer unabhängiger Zeilenvektoren von A.

(ii) Der Spaltenrang $\operatorname{SRang}(A)$ ist gleich der maximalen Anzahl linearer unabhängiger Spaltenvektoren von A.

Beweis. Aussage (i): Die Zeilenvektoren $a_{1\bullet}, \dots, a_{n\bullet}$ von A spannen den Zeilenraum $\operatorname{ZR}(A)$ auf. Nach dem Basisauswahlsatz (Folgerung 13.7) können wir aus den n Zeilenvektoren eine Basis von $\operatorname{ZR}(A)$ auswählen. Deren Anzahl entspricht dem Zeilenrang $\operatorname{ZRang}(A)$. Jede echte Obermenge ist nach Satz 13.3 linear abhängig. Daher ist $\operatorname{ZRang}(A)$ gleich der maximalen Anzahl linear unabhängiger Zeilenvektoren von A.

Ein analoges Argument für die Spaltenvektoren zeigt Aussage (ii). □

Beispiel 15.12 (Zeilen- und Spaltenraum, Zeilen- und Spaltenrang).
Es sei

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 7 & 4 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 3}.$$

Dann gilt

$$\begin{aligned} &\operatorname{ZR}(A) = \langle \overbrace{\begin{pmatrix} 2 & 3 & -1 \end{pmatrix}, \begin{pmatrix} 7 & 4 & 0 \end{pmatrix}}^{\text{linear unabhängig}} \rangle && \text{mit } \operatorname{ZRang}(A) = \dim(\operatorname{ZR}(A)) = 2 \\ \text{und } &\operatorname{SR}(A) = \langle \underbrace{\begin{pmatrix} 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}}_{\text{linear abhängig}} \rangle = \langle \underbrace{\begin{pmatrix} 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}}_{\text{linear unabhängig}} \rangle && \text{mit } \operatorname{SRang}(A) = \dim(\operatorname{SR}(A)) = 2. \end{aligned}$$ △

Die in Beispiel 15.12 beobachtete Übereinstimmung von Zeilen- und Spaltenrang ist kein Zufall:

Satz 15.13 (Zeilenrang gleich Spaltenrang).
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n \times m}$. Dann gilt

$$0 \leqslant \operatorname{ZRang}(A) = \operatorname{SRang}(A) \leqslant \min\{m, n\}. \tag{15.16}$$

Beweis. Wir führen den Beweis in zwei Schritten.

Schritt 1: Wir zeigen $\operatorname{SRang}(A) \leqslant \operatorname{ZRang}(A)$.

Es sei $r \coloneqq \operatorname{ZRang}(A) \in \mathbb{N}_0$ und $C \in K^{r \times m}$ eine Matrix, deren linear unabhängige Zeilen $c_{1\bullet}, \dots, c_{r\bullet} \in K_m$ eine Basis des Zeilenraumes $\operatorname{ZR}(A)$ bilden, also $\langle c_{1\bullet}, \dots, c_{r\bullet} \rangle = \operatorname{ZR}(A)$. Die i-te Zeile $a_{i\bullet}$ von A, die ja zu $\operatorname{ZR}(A)$ gehört, ist also eine Linearkombination der Zeilen von C, sagen wir

$$a_{i\bullet} = b_{i1}\, c_{1\bullet} + \dots + b_{ir}\, c_{r\bullet}.$$

Da das Gesagte für jede Zeile $i = 1, \ldots, n$ von A gilt, erhalten wir die Darstellung

$$A = B\,C = \begin{bmatrix} b_{11} \cdots\cdots b_{1r} \\ \vdots \qquad \vdots \\ b_{i1} \cdots\cdots b_{ir} \\ \vdots \qquad \vdots \\ b_{n1} \cdots\cdots b_{nr} \end{bmatrix} \begin{bmatrix} \text{——} & c_{1\bullet} & \text{——} \\ & \vdots & \\ \text{——} & c_{r\bullet} & \text{——} \end{bmatrix}$$

mit einer Koeffizientenmatrix $B \in K^{n \times r}$.

Wegen $Ax = B\,(Cx)$ für alle $x \in K^m$ ist jede Linearkombination der Spalten von A auch eine Linearkombination der Spalten von B, es gilt also

$$\begin{aligned} & \operatorname{SR}(A) \subseteq \operatorname{SR}(B) \\ \Rightarrow\quad & \operatorname{SRang}(A) \leqslant \operatorname{SRang}(B) \leqslant r = \operatorname{ZRang}(A) \quad \text{nach Folgerung 13.17.} \end{aligned}$$

Schritt 2: Wir zeigen $\operatorname{ZRang}(A) \leqslant \operatorname{SRang}(A)$.

Es sei nun $s \coloneqq \operatorname{SRang}(A) \in \mathbb{N}_0$ und $B \in K^{n \times s}$ eine Matrix, deren lineare unabhängige Spalten $b_{\bullet 1}, \ldots b_{\bullet s} \in K^n$ eine Basis des Spaltenraumes $\operatorname{SR}(A)$ bilden, also $\langle b_{\bullet 1}, \ldots b_{\bullet s} \rangle = \operatorname{SR}(A)$. Die j-te Spalte $a_{\bullet j}$ von A, die ja zu $\operatorname{SR}(A)$ gehört, ist also eine Linearkombination der Spalten von B, sagen wir

$$a_{\bullet j} = b_{\bullet 1}\, c_{1j} + \cdots + b_{\bullet s}\, c_{sj}.$$

Da das Gesagte für jede Spalte $j = 1, \ldots, m$ von A gilt, erhalten wir die Darstellung

$$A = B\,C = \begin{bmatrix} | & & | \\ b_{\bullet 1} & \cdots & b_{\bullet s} \\ | & & | \end{bmatrix} \begin{bmatrix} c_{11} \cdots\cdots c_{1j} \cdots\cdots c_{1m} \\ \vdots \qquad\quad \vdots \\ c_{s1} \cdots\cdots c_{sj} \cdots\cdots c_{sm} \end{bmatrix}$$

mit einer Koeffizientenmatrix $C \in K^{s \times m}$.

Wegen $y\,A = (y\,B)\,C$ für alle $y \in K_n$ ist jede Linearkombination der Zeilen von A auch eine Linearkombination der Zeilen von C, es gilt also

$$\begin{aligned} & \operatorname{ZR}(A) \subseteq \operatorname{ZR}(C) \\ \Rightarrow\quad & \operatorname{ZRang}(A) \leqslant \operatorname{ZRang}(C) \leqslant s = \operatorname{SRang}(A) \quad \text{nach Folgerung 13.17.} \end{aligned}$$

Der Beweis bis hierher zeigt $0 \leqslant \operatorname{ZRang}(A) = \operatorname{SRang}(A)$ für beliebige Matrizen $A \in K^{n \times m}$. Wegen $\operatorname{ZRang}(A) \leqslant n$ und $\operatorname{SRang}(A) \leqslant m$ (Lemma 15.11) folgt die Behauptung (15.16). □

Da Zeilenrang und Spaltenrang übereinstimmen, sprechen wir ab sofort nur noch vom **Rang** einer Matrix:

Definition 15.14 (Rang einer Matrix)**.**
Es sei K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n \times m}$.

(*i*) Dann ist der **Rang** (englisch: rank) von A definiert als

$$\operatorname{Rang}(A) \coloneqq \operatorname{ZRang}(A) = \operatorname{SRang}(A).$$

(*ii*) Wir sagen, A habe **vollen Zeilenrang** (englisch: full row rank), wenn $\operatorname{Rang}(A) = n$ gilt, wenn also die Menge der Zeilenvektoren von A linear unabhängig ist.

(*iii*) Wir sagen, A habe **vollen Spaltenrang** (englisch: full column rank), wenn $\operatorname{Rang}(A) = m$ gilt, wenn also die Menge der Spaltenvektoren von A linear unabhängig ist.

(*iv*) Wir sagen, A habe **vollen Rang** (englisch: full rank), wenn $\operatorname{Rang}(A) = \min\{m, n\}$ gilt. Andernfalls heißt die Matrix A **rang-defizitär** (englisch: rank-deficient). △

Beispiel 15.15 (Rang einer Matrix)**.**

(*i*) Die Matrix

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 7 & 4 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

aus Beispiel 15.12 hat $\operatorname{Rang}(A) = 2$. A hat also vollen Rang, was hier vollen Zeilenrang bedeutet.

(*ii*) Die Matrix

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 7 & \frac{21}{2} & -\frac{7}{2} \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

hat dagegen nur $\operatorname{Rang}(A) = 1$. Sie ist rang-defizitär.

(*iii*) Der Rang einer Diagonalmatrix $D \in K^{n\times m}$ ist gleich der Anzahl der von Null verschiedenen Diagonaleinträge von D. △

Als direktes Resultat aus dem Beweis von Satz 15.13 halten wir fest:

Folgerung 15.16 (Rangfaktorisierung)**.**
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n\times m}$ mit $r = \operatorname{Rang}(A) \in \mathbb{N}_0$. Dann existieren Matrizen $B_{\text{Rang}} \in K^{n\times r}$ und $C_{\text{Rang}} \in K^{r\times m}$, sodass gilt:

$$A = B_{\text{Rang}}\, C_{\text{Rang}} = \begin{bmatrix} | & & | \\ b_{\bullet 1} & \cdots & b_{\bullet r} \\ | & & | \end{bmatrix} \begin{bmatrix} \text{---} & c_{1\bullet} & \text{---} \\ & \vdots & \\ \text{---} & c_{r\bullet} & \text{---} \end{bmatrix}. \tag{15.17}$$

Die Spalten von B_{Rang} bilden eine Basis von $\operatorname{SR}(A)$. Die Zeilen von C_{Rang} bilden eine Basis von $\operatorname{ZR}(A)$.

Eine Faktorisierung der Matrix A wie in (15.17), bei der die inneren Dimensionen der Faktoren mit $\operatorname{Rang}(A)$ übereinstimmt, heißt eine **Rangfaktorisierung** (englisch: rank factorization) von A. Der Rang einer Matrix A ist ein Maß für deren „Informationsgehalt“.

Satz 15.17 (Rang des Produkts von Matrizen)**.**
Es seien K ein Körper und $m, n, \ell \in \mathbb{N}_0$. Für Matrizen $A \in K^{n \times m}$ und $B \in K^{m \times \ell}$ gilt:

$$0 \leqslant \operatorname{Rang}(A\,B) \leqslant \min\{\operatorname{Rang}(A),\ \operatorname{Rang}(B)\} \leqslant \min\{\ell, m, n\}. \tag{15.18}$$

Beweis. Wir verwenden dieselbe Technik wie im Beweis von Satz 15.13: Jede Linearkombination der Spalten von $A\,B$, also $A\,B\,x = A\,(B\,x)$ mit Koeffizientenvektor $x \in K^\ell$, ist eine Linearkombination der Spalten von A. Also gilt $\operatorname{SR}(A\,B) \subseteq \operatorname{SR}(A)$ und somit $\operatorname{Rang}(A\,B) = \operatorname{SRang}(A\,B) \leqslant \operatorname{SRang}(A) = \operatorname{Rang}(A)$.

Außerdem ist jede Linearkombination der Zeilen von $A\,B$, also $y\,A\,B = (y\,A)\,B$ mit Koeffizientenvektor $y \in K_n$, eine Linearkombination der Zeilen von B. Also gilt $\operatorname{ZR}(A\,B) \subseteq \operatorname{ZR}(B)$ und somit $\operatorname{Rang}(A\,B) = \operatorname{ZRang}(A\,B) \leqslant \operatorname{ZRang}(B) = \operatorname{Rang}(B)$. Das zeigt die zweite Ungleichung in (15.18).

Die dritte Ungleichung folgt sofort aus $\operatorname{Rang}(A) \leqslant \min\{m, n\}$ und $\operatorname{Rang}(B) \leqslant \min\{\ell, m\}$, siehe Satz 15.13. □

§ 15.3 Zeilenstufenform

Wir geben in diesem Abschnitt eine konstruktive Möglichkeit an, eine Rangfaktorisierung (Folgerung 15.16) einer Matrix A und damit auch den $\operatorname{Rang}(A)$ zu bestimmen. Dazu bringen wir die Matrix A durch geschickte Umformungen auf eine Gestalt, aus der wir eine Basis von $\operatorname{ZR}(A)$ und damit die Dimension von $\operatorname{ZR}(A)$, also den Zeilenrang von A, ablesen können.

Definition 15.18 (elementare Zeilenumformungen, Elementarmatrizen)**.**
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und

$$A = \begin{bmatrix} \text{——} & a_{1\bullet} & \text{——} \\ & \vdots & \\ \text{——} & a_{i\bullet} & \text{——} \\ & \vdots & \\ \text{——} & a_{n\bullet} & \text{——} \end{bmatrix} \in \mathbb{K}^{n \times m}.$$

Unter **elementaren Zeilenumformungen** (englisch: elementary row operations) versteht man die folgenden Operationen, angewendet auf die Matrix A:[14]

Typ I: Multiplikation der i-ten Zeile mit einem Skalar $\alpha \in K \setminus \{0\}$, d. h., Multiplikation der Matrix A von links mit der $n \times n$-Diagonalmatrix

$$D_i(\alpha) \coloneqq \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & & \\ & & & \alpha & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} = I + (\alpha - 1)\,E_{ii}. \tag{15.19}$$

[14] Matrizen vom Typ I existieren für $n \geqslant 1$. Matrizen vom Typ II und III existieren für $n \geqslant 2$.

Es gilt (nur die gegenüber A geänderten Zeilen werden hervorgehoben)

$$D_i(\alpha)\,A = \begin{bmatrix} & \vdots & \\ \text{———} & \alpha\, a_{i\bullet} & \text{———} \\ & \vdots & \end{bmatrix}.$$

Typ II: Addition des α-Fachen der j-ten Zeile zur i-ten Zeile ($i \neq j$) mit einem Skalar $\alpha \in K$, d. h., Multiplikation der Matrix A von links mit der $n \times n$-Matrix[15]

$$S_{i,j}(\alpha) \coloneqq \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & \alpha & 1 & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = I + \alpha\, E_{ij}. \tag{15.20}$$

Es gilt

$$S_{i,j}(\alpha)\,A = \begin{bmatrix} & \vdots & \\ \text{———} & a_{j\bullet} & \text{———} \\ & \vdots & \\ \text{———} & a_{i\bullet} + \alpha\, a_{j\bullet} & \text{———} \\ & \vdots & \end{bmatrix}.$$

Typ III: Vertauschen der i-ten mit der j-ten Zeile ($i \neq j$), d. h. Multiplikation der Matrix A von links mit der $n \times n$-Matrix

$$T_{i,j} \coloneqq \begin{bmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & & 1 & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & \\ & & & 1 & & & & 0 & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{bmatrix} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}. \tag{15.21}$$

Es gilt

$$T_{i,j}\,A = \begin{bmatrix} & \vdots & \\ \text{———} & a_{j\bullet} & \text{———} \\ & \vdots & \\ \text{———} & a_{i\bullet} & \text{———} \\ & \vdots & \end{bmatrix}.$$

[15] Die Illustration zeigt den Fall $i > j$.

Die Matrizen $D_i(\cdot)$, $S_{i,j}(\cdot)$ und $T_{i,j}$ heißen **Elementarmatrizen** (englisch: elementary matrices) vom Typ I, Typ II bzw. Typ III. Die Matrizen $T_{i,j}$ (Typ III) nennen wir auch **Transpositionsmatrizen** (englisch: transposition matrices).[16] △

Transpositionsmatrizen (Typ III) verwenden wir nur aus Bequemlichkeit. Sie werden eigentlich nicht benötigt, da wir jede Elementarmatrix vom Typ III als ein Produkt von Elementarmatrizen vom Typ I und Typ II schreiben können. (**Quizfrage 15.1:** Wie nämlich?)

Lemma 15.19 (elementare Zeilenumformungen ändern den Zeilenraum und den Zeilenrang nicht)**.**
Es seien K ein Körper und $m, n \in \mathbb{N}_0$. Entsteht die Matrix $C \in K^{n\times m}$ aus $A \in K^{n\times m}$ durch elementare Zeilenumformungen, dann gilt $\operatorname{ZR}(C) = \operatorname{ZR}(A)$, also auch $\operatorname{Rang}(C) = \operatorname{Rang}(A)$.

Beweis. Es reicht zu zeigen, dass sich durch eine einzelne elementare Zeilenumformung der Zeilenraum der Matrix nicht ändert.

Typ I: Für $\alpha \in K \setminus \{0\}$ gilt offensichtlich

$$\langle a_{1\bullet}, \; \ldots, \; a_{n\bullet} \rangle = \langle a_{1\bullet}, \; \ldots, \; a_{i-1\bullet}, \; \alpha \, a_{i\bullet}, \; a_{i+1\bullet}, \; \ldots, \; a_{n\bullet} \rangle.$$

Typ II: Auch in diesem Fall haben wir

$$\langle a_{1\bullet}, \; \ldots, \; a_{n\bullet} \rangle = \langle a_{1\bullet}, \; \ldots, \; a_{i-1\bullet}, \; a_{i\bullet} + \alpha \, a_{j\bullet}, \; a_{i+1\bullet}, \; \ldots, \; a_{n\bullet} \rangle.$$

Quizfrage 15.2: Wie rechnen sich die Koeffizienten einer Linearkombination der Vektoren um?

Typ III: Offenbar ändert die Reihenfolge der Vektoren nicht das Resultat von

$$\langle a_{1\bullet}, \; \ldots, \; a_{n\bullet} \rangle.$$ □

Die gewünschte Gestalt, aus der man den Zeilenrang und eine Basis des Zeilenraumes einer Matrix gut ablesen kann, ist die folgende:

Definition 15.20 (Zeilenstufenform)**.**
Es seien K ein Körper und $m, n \in \mathbb{N}_0$. Eine Matrix $A \in K^{n\times m}$ heißt in **Zeilenstufenform** (englisch: row echelon form), wenn folgende Bedingungen erfüllt sind:

(i) Es gibt eine Zahl $r \in [\![0, n]\!]$, sodass die Zeilen $a_{1\bullet}, \ldots, a_{r\bullet}$ keine Nullzeilen sind und die Zeilen $a_{r+1\bullet}, \ldots, a_{n\bullet}$ sämtlich Nullzeilen sind.

(ii) Bezeichnet $j_i \coloneqq \min\{j \in [\![1, m]\!] \mid a_{ij} \neq 0\}$ den niedrigsten Spaltenindex in Zeile $i \in [\![1, r]\!]$, in der ein Eintrag ungleich 0 steht, dann gilt die **Stufenbedingung** (englisch: echelon condition) $j_1 < j_2 < \cdots < j_r$.

Die Elemente $a_{1j_1}, \ldots, a_{rj_r}$ heißen die **Pivot-Elemente** (englisch: pivot elements) der Zeilenstufenform. △

[16] Die Bezeichnung der Typen I, II und III ist in der Literatur nicht einheitlich.

Bemerkung 15.21 (zur Zeilenstufenform)**.**

(i) Die Stufenbedingung bedeutet, dass sowohl links als auch unterhalb von Pivot-Elementen nur Nullen stehen können. Die Pivot-Elemente rücken von Zeile zu Zeile mindestens eine Spalte weiter nach rechts.

(ii) Die Lage der Pivot-Elemente in einer Zeilenstufenform einer Matrix ist durch die Ausgangsmatrix eindeutig festgelegt. Die gesamte Zeilenstufenform an sich ist aber i. A. nicht eindeutig, da wir Nicht-Nullzeilen mit Skalaren aus $K \setminus \{0\}$ multiplizieren können und wiederum eine Zeilenstufenform erhalten. Außerdem können wir zu einer Zeile ein Vielfaches einer weiter unten stehenden Zeile addieren, ohne die Zeilenstufenform zu stören.

(iii) Der Rang einer Matrix A in Zeilenstufenform ist wegen $r = \operatorname{Rang}(A)$ leicht abzulesen: Der Rang ist gleich der Anzahl der Nicht-Nullzeilen, also gleich der Anzahl der Pivot-Elemente. (**Quizfrage 15.3:** Warum?) △

Beispiel 15.22 (Zeilenstufenform)**.**
Nachfolgend sind die Besetzungsmuster einiger 3×4-Matrizen in Zeilenstufenform zu sehen, wobei ? jeweils für einen Eintrag aus dem Körper K steht und $\star$ für einen Eintrag ungleich 0 (die Pivot-Elemente).

$$\begin{array}{r}j_1 = 1 \longrightarrow\\ j_2 = 2 \longrightarrow\\ j_3 = 3 \longrightarrow\end{array}\begin{bmatrix}\star & ? & ? & ?\\ 0 & \star & ? & ?\\ 0 & 0 & \star & ?\end{bmatrix} \qquad \begin{array}{r}j_1 = 1 \longrightarrow\\ j_2 = 3 \longrightarrow\\ \end{array}\begin{bmatrix}\star & ? & ? & ?\\ 0 & 0 & \star & ?\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad \begin{array}{r}j_1 = 3 \longrightarrow\\ j_2 = 4 \longrightarrow\\ \end{array}\begin{bmatrix}0 & 0 & \star & ?\\ 0 & 0 & 0 & \star\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$r = 3 \qquad\qquad r = 2 \qquad\qquad r = 2$$

△

Wir geben nun einen Algorithmus an, der eine gegebene Matrix $A \in K^{n \times m}$ durch elementare Zeilenumformungen in eine Matrix $C \in K^{n \times m}$ in Zeilenstufenform überführt. Die Idee des Verfahrens ist folgende:

(1) Finde eines der am weitesten links stehenden Elemente in der Matrix. Dessen Spaltenindex sei j_1. Bringe das Element, sofern erforderlich, durch Zeilentausch (elementare Zeilenumformung vom Typ III) an die oberste Position $(1, j_1)$.

(2) Erzeuge in der Spalte j_1 unterhalb der Position $(1, j_1)$ Nullen durch Addition geeigneter Vielfacher der Zeile 1 zur entsprechenden Zeile 2 bis n (elementare Zeilenumformungen vom Typ II).

(3) Wiederhole die obigen Schritte mit der unteren rechten Teilmatrix ab Zeile 2 und ab Spalte $j_1 + 1$.

Ein vollständiges Verfahren wird in Algorithmus D.1 angegeben. Mit Hilfe von Algorithmus D.1 können wir folgenden Satz konstruktiv beweisen:

Satz 15.23 (Erreichbarkeit der Zeilenstufenform)**.**
Es seien K ein Körper und $n, m \in \mathbb{N}_0$. Jede Matrix $A \in K^{n \times m}$ lässt sich durch elementare Zeilenumformungen in eine Matrix $C \in K^{n \times m}$ in Zeilenstufenform überführen. Ist $r \in [\![0, n]\!]$ die Anzahl der Nicht-Nullzeilen in C, dann bilden die Zeilenvektoren $c_{1\bullet}, \ldots, c_{r\bullet}$ eine Basis von $\operatorname{ZR}(A) = \operatorname{ZR}(C)$, und es gilt $\operatorname{Rang}(A) = \operatorname{Rang}(C) = r$.

Beispiel 15.24 (Erreichbarkeit der Zeilenstufenform).
Wir betrachten ein Beispiel in $\mathbb{R}^{3\times 4}$:

$$\begin{bmatrix} 0 & 0 & 3 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 3 & 0 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 3 & 0 & 2 \end{bmatrix}$$

Tauschen der Zeilen 1 und 2

$$-3\quad\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 3 & 0 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -6 & 2 \end{bmatrix}$$

Erzeugen von Nullen unterhalb des Pivot-Elements
Addition des (-3)-Fachen der Zeile 1 zur Zeile 3

$$2\quad\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -6 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Erzeugen von Nullen unterhalb des Pivot-Elements
Addition des 2-Fachen der Zeile 2 zur Zeile 3.

Die ursprüngliche Matrix A wird also von links zunächst mit $T_{1,2}$ von links multipliziert, anschließend mit $S_{3,1}(-3)$ und schließlich mit $S_{3,2}(2)$. Da nun eine Zeilenstufenform erreicht ist, können wir aus dieser den Zeilenraum von A ablesen:

$$\operatorname{ZR}(A) = \Big\langle \begin{pmatrix} 0 & 1 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 3 & -1 \end{pmatrix} \Big\rangle,$$

und es gilt $\dim(\operatorname{ZR}(A)) = \operatorname{Rang}(A) = 2$. △

Bemerkung 15.25 (Berechnung einer Rangfaktorisierung).
Wir hatten eingangs des Abschnitts angekündigt, eine konstruktive Möglichkeit anzugeben, eine Rangfaktorisierung $A = B_{\text{Rang}}\, C_{\text{Rang}}$ einer Matrix $A \in K^{n\times m}$ zu bestimmen, sodass also $B_{\text{Rang}} \in K^{n\times r}$, $C_{\text{Rang}} \in K^{r\times m}$ und $r = \operatorname{Rang}(A)$ gilt. Tatsächlich erhalten wir aus der besprochenen Zeilenstufenform C den gesuchten rechten Faktor. Dazu müssen wir lediglich in C eventuell vorhandene Nullzeilen streichen, daraus ergibt sich die Matrix sei $C_{\text{Rang}} \in K^{r\times m}$. Wie aber kommen wir an den linken Faktor B_{Rang}?

Die Zeilenstufenform entstand aus A durch Zeilenmanipulationen, also durch Multiplikation von A mit Elementarmatrizen $E_j \in K^{n\times n}$ von links:

$$E_k \cdots E_2\, E_1\, A = C. \tag{15.22}$$

Wenn es gelänge, die Multiplikationen durch geeignete Matrizen $E'_j \in K^{n\times n}$ rückgängig zu machen, wobei also $E'_j\, E_j = I_n$ gelten soll[17], dann bekämen wir die Darstellung

$$E_k\, E_{k-1} \cdots E_2\, E_1\, A = C \quad\Rightarrow\quad E_{k-1} \cdots E_2\, E_1\, A = E'_k\, C \quad\Rightarrow\quad \cdots \quad\Rightarrow\quad A = E'_1\, E'_2 \cdots E'_k\, C.$$

In der Tat ist das möglich, da Elementarmatrizen invertierbar sind (Übung). Setzen wir nun zur Abkürzung $B \coloneqq E'_1\, E'_2 \cdots E'_k$, so erhalten wir eine Faktorisierung der Gestalt

$$A = {\scriptstyle n}\underbrace{\left[\begin{array}{c|c} \overset{r}{B_{\text{Rang}}} & \overset{n-r}{*} \end{array}\right]}_{B} \underbrace{\left[\begin{array}{c} \overset{m}{C_{\text{Rang}}} \\ \hline 0 \end{array}\right]}_{C} \begin{array}{l} {\scriptstyle r} \\ {\scriptstyle n-r} \end{array}$$

[17] Später (Definition 15.39, Lemma 15.43) werden wir solche Matrizen **invertierbar** nennen.

Da bei der Bildung des Matrix-Produkts $A = B\,C$ die letzten $n - r$ Spalten von B immer nur mit den Null-Koeffizienten aus den unteren Zeilen von C multipliziert werden, können wir, ohne das Ergebnis der Matrix-Multiplikation zu ändern, den rechten Faktor C durch seine oberen r Zeilen C_{Rang} und den linken Faktor B durch seine linken r Spalten B_{Rang} ersetzen. So erhalten wir die gewünschte Rangfaktorisierung

$$A = B_{\text{Rang}}\,C_{\text{Rang}}$$

mit $B_{\text{Rang}} \in K^{n\times r}$ und $C_{\text{Rang}} \in K^{r\times m}$. (**Quizfrage 15.4:** Wie müsste Algorithmus D.1 ergänzt werden, damit wir als Ergebnis auch den linken Faktor B der Faktorisierung $A = B\,C$ erhalten, aus der dann durch Streichen von Spalten bzw. Zeilen die Rangfaktorisierung $A = B_{\text{Rang}}\,C_{\text{Rang}}$ folgt? (Die Auflösung ergibt sich aus Algorithmus D.2.)) △

§ 15.4 Transposition von Matrizen

Definition 15.26 (Transposition).
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n\times m}$. Die Matrix $A^\intercal \in K^{m\times n}$, definiert durch $(A^\intercal)_{ij} = (A)_{ji}$ für $i \in [\![1, m]\!]$ und $j \in [\![1, n]\!]$, heißt die zu A **transponierte Matrix** (englisch: transposed matrix, lateinisch: transponere: umstellen) oder kurz die **Transponierte** (englisch: transpose) zu A. △

Die transponierte Matrix $A^\intercal$ entsteht aus A durch Spiegelung an der Hauptdiagonalen. Dadurch werden die Zeilen zu Spalten und die Spalten zu Zeilen. Die zu A transponierte Matrix hat die Darstellung

$$A^\intercal = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}. \tag{15.23}$$

Beispiel 15.27 (transponierte Matrix).

$$A = \begin{bmatrix} 8 & -5 & 2 & 0 \\ 0 & 2 & 3 & 0 \\ -1 & 7 & 4 & 4 \end{bmatrix} \quad\Rightarrow\quad A^\intercal = \begin{bmatrix} 8 & 0 & -1 \\ -5 & 2 & 7 \\ 2 & 3 & 4 \\ 0 & 0 & 4 \end{bmatrix}.$$ △

Lemma 15.28 (Rechenregeln für die Transposition).
Es seien K ein Körper und $m, n, \ell \in \mathbb{N}_0$. Für Matrizen $A, B \in K^{n\times m}$ und $C \in K^{m\times \ell}$ und Skalare $\alpha \in K$ gelten die folgenden Eigenschaften:

$$(A^\intercal)^\intercal = A \quad \text{Transposition ist **involutorisch**} \tag{15.24}$$

$$(A + B)^\intercal = A^\intercal + B^\intercal \quad \text{Transposition ist additiv} \tag{15.25a}$$

$$(\alpha\,A)^\intercal = \alpha\,A^\intercal \quad \text{Transposition ist homogen[18]} \tag{15.25b}$$

$$(A\,C)^\intercal = C^\intercal A^\intercal. \tag{15.26}$$

Beweis. (15.24), (15.25a) und (15.25b) sind offensichtlich. Für (15.26) betrachten wir

$$\begin{aligned}[(A\,C)^\intercal]_{ik} = [A\,C]_{ki} = \sum_{j=1}^m a_{kj}\,c_{ji} = \sum_{j=1}^m c_{ji}\,a_{kj} \\ = \sum_{j=1}^m [C^\intercal]_{ij}\,[A^\intercal]_{jk} = [C^\intercal A^\intercal]_{ik}.\end{aligned}$$ □

Lemma 15.29 (Rang der transponierten Matrix)**.**
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n\times m}$. Dann gilt $\operatorname{Rang}(A) = \operatorname{Rang}(A^\intercal)$.

Beweis. Es sei $A = B\,C$ eine Rangfaktorisierung (Folgerung 15.16) mit $B \in K^{n\times r}$, $C \in K^{r\times m}$ und $r = \operatorname{Rang}(A)$. Aufgrund von (15.26) gilt $A^\intercal = C^\intercal B^\intercal$. Aus Satz 15.17 und Satz 15.13 folgt $\operatorname{Rang}(A^\intercal) \leqslant \min\{\operatorname{Rang}(B^\intercal), \operatorname{Rang}(C^\intercal)\} \leqslant \min\{m, n, r\} \leqslant r = \operatorname{Rang}(A)$.

Indem wir A durch $A^\intercal$ ersetzen, erhalten wir aus der bereits bewiesenen Aussage $\operatorname{Rang}(A^\intercal) \leqslant \operatorname{Rang}(A)$ auch die umgekehrte Ungleichung $\operatorname{Rang}(A) = \operatorname{Rang}((A^\intercal)^\intercal) \leqslant \operatorname{Rang}(A^\intercal)$. □

Definition 15.30 (symmetrische und antisymmetrische Matrizen)**.**
Es seien K ein Körper, $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$.

(i) A heißt **symmetrisch** (englisch: symmetric), wenn $A = A^\intercal$ gilt.

(ii) A heißt **antisymmetrisch** (englisch: antisymmetric) oder **schiefsymmetrisch** (englisch: skew-symmetric), wenn $A = -A^\intercal$ gilt.

Die Menge der symmetrischen bzw. schiefsymmetrischen $n \times n$-Matrizen bezeichnen wir mit $K^{n\times n}_{\text{sym}}$ bzw. $K^{n\times n}_{\text{skew}}$. △

Beispiel 15.31 (symmetrische und antisymmetrische Matrizen)**.**

(i) Die Matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

ist symmetrisch.

(ii) Die Matrix

$$B = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

ist antisymmetrisch.

[18] Additivität und Homogenität ergibt zusammen **Linearität**. Die Sprechweise, dass die Transposition eine lineare Abbildung sei, wird in § 17 klar werden.

(*iii*) Die Matrix

$$C = \begin{bmatrix} 0 & 1 & 2 \\ 4 & 0 & 3 \\ 3 & 2 & 0 \end{bmatrix} \in \mathbb{Z}_5^{3\times 3}$$

ist antisymmetrisch.

(*iv*) Die Matrix

$$D = \begin{bmatrix} 0 & -i & -2+i \\ i & 0 & 3 \\ 2-i & -3 & 0 \end{bmatrix} \in \mathbb{C}^{3\times 3}$$

ist antisymmetrisch. △

Lemma 15.32 (symmetrische und antisymmetrische Anteile quadratischer Matrizen)**.**
Es seien K ein Körper der Charakteristik $\operatorname{char}(K) \neq 2$ und $n \in \mathbb{N}_0$. Dann sind $K^{n\times n}_{\text{sym}}$ und $K^{n\times n}_{\text{skew}}$ Unterräume von $K^{n\times n}$ der Dimensionen

$$\dim(K^{n\times n}_{\text{sym}}) = \frac{1}{2} n\,(n+1), \tag{15.27a}$$

$$\dim(K^{n\times n}_{\text{skew}}) = \frac{1}{2} n\,(n-1), \tag{15.27b}$$

und gilt

$$K^{n\times n} = K^{n\times n}_{\text{sym}} \oplus K^{n\times n}_{\text{skew}}. \tag{15.28}$$

Quizfrage 15.5: Was gilt stattdessen im Körper $\mathbb{Z}_2$ der Charakteristik $\operatorname{char}(K) = 2$?

Beweis. Der Beweis ist Gegenstand der Übung. □

Ende der Vorlesung 20

§ 15.5 Der Ring quadratischer Matrizen

Die Menge der quadratischen Matrizen $K^{n\times n}$ über einem Körper $(K, +, \cdot)$ ist bzgl. der Matrix-Multiplikation $\cdot$ abgeschlossen.[19] Zusammen mit der Matrix-Addition ergibt sich eine Ringstruktur $(K^{n\times n}, +, \cdot)$:

Lemma 15.33 (quadratische Matrizen bilden einen nicht-kommutativen Ring mit Eins)**.**
Für $n \in \mathbb{N}_0$ bilden die quadratischen $n \times n$-Matrizen mit der Matrixaddition (15.6a) und der Matrix-Multiplikation (15.9) einen Ring mit dem Einselement I_n. Dieser Ring $(K^{n\times n}, +, \cdot)$ heißt **Matrixring** oder **Matrizenring** (englisch: matrix ring) der Größe $n \times n$ über dem Körper K. Im Fall $n \geqslant 2$ ist dieser Ring nicht kommutativ.

[19] Vorübergehend schreiben wir das Multiplikationszeichen $\cdot$ wieder.

Beweis. Wir prüfen die Definition 9.1 nach. $(K^{n \times n}, +)$ ist eine abelsche Gruppe, da $(K^{n \times n}, +, \cdot)$ (mit der S-Multiplikation $\cdot$) ein Vektorraum ist (Satz 15.3).[20] Mit der Matrix-Multiplikation $\cdot$ bildet $(K^{n \times n}, \cdot)$ eine Halbgruppe, da $\cdot$ nach Lemma 15.8 assoziativ ist. Die Distributivgesetze

$$\begin{aligned} A \cdot (B_1 + B_2) &= A \cdot B_1 + A \cdot B_2 \\ (A_1 + A_2) \cdot B &= A_1 \cdot B + A_2 \cdot B \end{aligned}$$

und die Neutralität der Einheitsmatrix I_n wurden ebenfalls in Lemma 15.8 gezeigt.

Die Nicht-Kommutativität für $n \geqslant 2$ sehen wir am Beispiel

$$E_{11} \cdot E_{12} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 \end{bmatrix} = E_{12}$$

$$E_{12} \cdot E_{11} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = 0,$$

das für beliebige Körper K gültig ist. □

Der Ring der quadratischen Matrizen $K^{n \times n}$ enthält neben den Diagonalmatrizen noch einige weitere erwähnenswerte Teilmengen:

Definition 15.34 (obere und untere Dreiecksmatrix[21])**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$. Eine Matrix $A \in K^{n \times n}$ heißt

(i) eine **obere Dreiecksmatrix** (englisch: upper triangular matrix), wenn $a_{ij} = 0$ für alle $1 \leqslant j < i \leqslant n$ gilt.

(ii) eine **strikte obere Dreiecksmatrix** (englisch: strictly upper triangular matrix), wenn $n \geqslant 1$ und $a_{ij} = 0$ für alle $1 \leqslant j \leqslant i \leqslant n$ gilt.

(iii) eine **untere Dreiecksmatrix** (englisch: lower triangular matrix), wenn $a_{ij} = 0$ für alle $1 \leqslant i < j \leqslant n$ gilt.

(iv) eine **strikte untere Dreiecksmatrix** (englisch: strictly lower triangular matrix), $n \geqslant 1$ und wenn $a_{ij} = 0$ für alle $1 \leqslant i \leqslant j \leqslant n$ gilt.

(v) **nilpotent** (englisch: nilpotent), wenn es ein $k \in \mathbb{N}_0$ gibt mit der Eigenschaft $A^k = 0$. △

[20] Dass wir sowohl die Matrix-Multiplikation als auch die S-Multiplikation einer Matrix mit einem Skalar mit demselben Symbol $\cdot$ bezeichnen, sollte nicht zu Verwirrung führen, da sich aus den verknüpften Objekten ergibt, welche Multiplikation gemeint ist. (**Quizfrage 15.6:** Wieso stimmt das auch im Fall $n = 1$?)

[21] Um Dreiecksmatrizen von strikten Dreiecksmatrizen unterscheiden zu können, schließen wir für die folgenden Resultate die 0×0-Matrizen aus.

Wir bezeichnen die Menge der Diagonalmatrizen der Dimension $n \times n$ auch mit dem Symbol $K^{n\times n}_{\diagdown}$ und die Menge der oberen bzw. unteren Dreiecksmatrizen mit $K^{n\times n}_{◹}$ bzw. $K^{n\times n}_{◺}$. Es gilt

$$K^{n\times n}_{\diagdown} = K^{n\times n}_{◹} \cap K^{n\times n}_{◺}.$$

Beispiel 15.35 (obere und untere Dreiecksmatrix)**.**

(i)

$$\begin{bmatrix} 1 & 3 & -3 \\ 0 & 7 & 4 \\ 0 & 0 & 5 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

ist eine obere Dreiecksmatrix, aber keine strikte obere Dreiecksmatrix.

(ii)

$$\begin{bmatrix} 0 & 0 & 0 \\ 7 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

ist eine strikte untere Dreicksmatrix.

(iii)

$$\begin{bmatrix} 2 & 2 & -2 \\ 5 & 1 & -3 \\ 1 & 5 & -3 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

ist eine nilpotente Matrix, denn es gilt

$$\begin{bmatrix} 2 & 2 & -2 \\ 5 & 1 & -3 \\ 1 & 5 & -3 \end{bmatrix}^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad \triangle$$

Lemma 15.36 (strikte Dreiecksmatrizen sind nilpotent)**.**
Es seien K ein Körper und $n \in \mathbb{N}$. Für jede strikte obere und jede strikte untere Dreiecksmatrix $A \in K^{n\times n}$ gilt $A^n = 0$.

Beweis. Der Beweis ist Gegenstand der Übung. □

Beispiel 15.37 (strikte Dreiecksmatrizen sind nilpotent)**.**
Für

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

gilt

$$A^2 = \begin{bmatrix} 0 & 0 & 4 & 17 \\ 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \qquad \triangle$$

Lemma 15.38 (obere und untere Dreiecksmatrizen bilden Unterringe)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$. Dann gilt:

(i) $K^{n \times n}_{\diagdown}$, $K^{n \times n}_{◹}$ und $K^{n \times n}_{◺}$ bilden jeweils einen Unterring mit Eins (Definition 9.12) von $K^{n \times n}$. $K^{n \times n}_{\diagdown}$ ist sogar ein kommutativer Ring mit Eins.

(ii) Für $n \geqslant 1$ bilden die strikten oberen und strikten unteren Dreiecksmatrizen einen Unterring (aber keinen Unterring mit Eins) von $K^{n \times n}$. Im Fall $n = 1$ ist das ein Nullring.

(**Quizfrage 15.7:** Sind $K^{n \times n}_{◹}$, $K^{n \times n}_{◺}$ oder $K^{n \times n}_{\diagdown}$ sogar Ideale?)

Beweis. Aussage (i): Nach Definition 9.12 ist zu zeigen, dass die jeweilige Teilmenge mit der Addition eine Untergruppe von $(K^{n \times n}, +)$ bildet, dass sie bzgl. der Matrix-Multiplikation abgeschlossen ist und das Einselement (die Einheitsmatrix I_n) enthält.

Wir führen den Beweis nur für den Fall $K^{n \times n}_{◹}$ aus. Sind $A, B \in K^{n \times n}_{◹}$, dann ist auch $-B \in K^{n \times n}_{◹}$ und damit $A - B \in K^{n \times n}_{◹}$. Aus dem Untergruppenkriterium (Satz 7.44) folgt, dass $(K^{n \times n}_{◹}, +)$ eine Untergruppe von $(K^{n \times n}, +)$ ist. Da $I_n \in K^{n \times n}_{◹}$ klar ist, bleibt nur zu zeigen, dass das Matrix-Produkt $A \cdot B$ von zwei Matrizen $A, B \in K^{n \times n}_{◹}$ wieder in $K^{n \times n}_{◹}$ liegt. Für Indizes $1 \leqslant k < i \leqslant n$ gilt

$$[A \cdot B]_{ik} = \sum_{j=1}^{n} \underbrace{a_{ij}}_{= 0 \text{ für } j < i} \overbrace{b_{jk}}^{= 0 \text{ für } j > k} = 0.$$

Da alle Summanden gleich Null sind, ist $A \cdot B$ tatsächlich wieder eine obere Dreiecksmatrix.

Aussage (ii): Der Beweis der Unterring-Eigenschaft geht genauso wie in Aussage (i). Da das Einselement I_n von $K^{n \times n}$ keine strikte Dreiecksmatrix ist, handelt es sich nicht um einen Unterring mit Eins[22] von $K^{n \times n}$. □

§ 15.6 Invertierbare Matrizen

Wir interessieren uns nun für die bzgl. der Matrix-Multiplikation invertierbaren Elemente im Ring der quadratischen Matrizen, also für die Einheitengruppe (Beispiel 7.22) des Monoids $(K^{n \times n}, \cdot)$.

Definition 15.39 (invertierbare Matrix)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$.

(i) Eine Matrix $A \in K^{n \times n}$ heißt **invertierbar** (englisch: invertible) oder **regulär** (englisch: non-singular), wenn sie ein invertierbares Element (Definition 7.14) des Monoids $(K^{n \times n}, \cdot)$ ist. Das heißt, es gibt eine Matrix $B \in K^{n \times n}$ mit der Eigenschaft

$$A\,B = I \quad \text{und} \quad B\,A = I. \tag{15.29}$$

[22] Im Fall $n = 1$ besteht der Unterring nur aus der Nullmatrix, ist also ein Nullring. Der Nullring hat zwar 0 als Einselement (Beispiel 9.2), dieses ist aber verschieden von der 1×1-Einheitsmatrix, daher handelt es sich auch in diesem Fall nicht um einen Unterring mit Eins im Sinne von Definition 9.12.

In diesem Fall heißt B die zu A **inverse Matrix** (englisch: inverse) oder kurz die **Inverse** (englisch: inverse) von A, in Symbolen: $B = A^{-1}$.

(ii) Andernfalls heißt A **nicht invertierbar** (englisch: non-invertible) oder **singulär** (englisch: singular).

(iii) Die Menge der invertierbaren Matrizen in $K^{n\times n}$, also die Einheitengruppe des Monoids $(K^{n\times n}, \cdot)$, heißt die **Gruppe der invertierbaren $n \times n$-Matrizen über K** (englisch: group of invertible matrices) oder die **allgemeine lineare Gruppe** (englisch: general linear group) **vom Grad n über dem Körper K**, in Symbolen

$$\mathrm{GL}(n, K) \coloneqq \{A \in K^{n\times n} \mid A \text{ ist invertierbar}\}. \tag{15.30}$$

△

Bemerkung 15.40 (zu invertierbaren Matrizen).

(i) B ist die Inverse von A genau dann, wenn A die Inverse von B ist. Wie in jedem Monoid ist das neutrale Element, also die Einheitsmatrix I, immer invertierbar und selbstinvers.

(ii) Die Bezeichnung **allgemeine lineare Gruppe** für die Einheitengruppe des Monoids $(K^{n\times n}, \cdot)$ leitet sich wie folgt ab: Jede Untergruppe des Monoids $(K^{n\times n}, \cdot)$, ist notwendigerweise eine Untergruppe der Einheitengruppe $\mathrm{GL}(n, K)$. Mit anderen Worten: $\mathrm{GL}(n, K)$ ist das Maximum der Menge aller Untergruppen von $(K^{n\times n}, \cdot)$ bzgl. der Untergruppen-Halbordnung, daher das Attribut **allgemeine Gruppe**. Das Attribut **linear** kommt daher, dass wir Matrizen über die Matrix-Vektor-Multiplikation $\mathbb{K}^n \ni x \mapsto A\,x \in K^n$ als lineare Abbildungen (Endomorphismen) von K^n auffassen können (Beispiel 17.5). △

Beispiel 15.41 (invertierbare Matrix).
Die Matrix

$$A = \begin{bmatrix} 1 & 0 & -7 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

ist invertierbar, denn es gilt

$$\begin{bmatrix} 1 & 0 & -7 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} \frac{1}{17} \begin{bmatrix} 3 & 7 & 0 \\ 1 & -9 & 17 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{und} \quad \frac{1}{17} \begin{bmatrix} 3 & 7 & 0 \\ 1 & -9 & 17 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -7 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Die inverse Matrix ist also

$$A^{-1} = \frac{1}{17} \begin{bmatrix} 3 & 7 & 0 \\ 1 & -9 & 17 \\ -2 & 1 & 0 \end{bmatrix}.$$

△

Beachte: Wie wir die Invertierbarkeit einer Matrix überprüfen und ggf. die Inverse berechnen, wird gleich noch Thema sein.

Satz 15.42 (Rechenregeln für Inverse).
Es seien K ein Körper und $n \in \mathbb{N}_0$ sowie $A, B, B_1, B_2 \in K^{n\times n}$.

(*i*) Ist A invertierbar, dann gelten die **Kürzungsregeln**

$$A\,B_1 = A\,B_2 \quad \Rightarrow \quad B_1 = B_2 \tag{15.31a}$$
$$B_1\,A = B_2\,A \quad \Rightarrow \quad B_1 = B_2. \tag{15.31b}$$

(*ii*) Die Invertierung ist **involutorisch**, d. h., für invertierbares A gilt

$$(A^{-1})^{-1} = A. \tag{15.32}$$

(*iii*) Sind A und B invertierbar, dann auch $A\,B$, und es gilt

$$(A\,B)^{-1} = B^{-1}A^{-1}. \tag{15.33}$$

(*iv*) Ist A invertierbar, dann auch $A^\intercal$, und es gilt

$$(A^\intercal)^{-1} = (A^{-1})^\intercal. \tag{15.34}$$

Beachte: Wegen (15.34) können wir statt $(A^\intercal)^{-1}$ in Zukunft auch einfach $A^{-\intercal}$ schreiben.

Beweis. Die Aussagen (*i*) bis (*iii*) sind uns bereits aus allgemeinen Monoiden bekannt: Die Kürzungsregeln (15.31) wurden in Lemma 7.18 gezeigt. Die involutorische Eigenschaft der Invertierung (15.32) haben wir bereits in (7.5) festgestellt. Aussage (*iii*) und (15.33) für die Inverse eines Produkts invertierbarer Elemente eines Monoids finden sich in Lemma 7.16.

Es bleibt Aussage (*iv*) zu zeigen. Dazu sei A invertierbar, dann gilt nach (15.26) für die Transponierte eines Produkts

$$\begin{aligned} A^\intercal\,(A^{-1})^\intercal &= (A^{-1}A)^\intercal = I^\intercal = I \\ \text{und} \quad (A^{-1})^\intercal\,A^\intercal &= (A\,A^{-1})^\intercal = I^\intercal = I, \end{aligned}$$

also ist $(A^{-1})^\intercal$ in der Tat die Inverse zu $A^\intercal$, was (15.34) zeigt. □

Anhand der Definition 15.39 kann man die Invertierbarkeit oder Nicht-Invertierbarkeit einer Matrix schlecht erkennen. Stattdessen geben wir nun Kriterien für die Invertierbarkeit an. Dabei beginnen wir mit „einfachen“ Matrizen und verallgemeinern die Aussagen dann auf allgemeine Matrizen.

Das folgende Resultat haben wir in Bemerkung 15.25 bereits verwendet:

Lemma 15.43 (Elementarmatrizen sind invertierbar)**.**
Es sei K ein Körper. Die Elementarmatrizen aus Definition 15.18 sind invertierbar. Die Inverse einer Elementarmatrix ist wiederum eine Elementarmatrix.

Beweis. Der Beweis ist Gegenstand der Übung. □

Lemma 15.44 (Invertierbarkeit von Diagonal- und Dreiecksmatrizen)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$.

(*i*) Es sei $A \in K^{n \times n}_{\diagdown}$ eine Diagonalmatrix. A ist genau dann invertierbar, wenn alle Diagonaleinträge von A ungleich Null sind. In dem Fall ist die Inverse von A wiederum eine Diagonalmatrix, bestehend aus den Inversen der Diagonaleinträge von A. Die invertierbaren Diagonalmatrizen bilden also eine Untergruppe der Gruppe der invertierbaren Matrizen $\mathrm{GL}(n, K)$.

(*ii*) Es sei $A \in K^{n \times n}_{\triangledown}$ eine obere Dreiecksmatrix. A ist genau dann invertierbar, wenn alle Diagonaleinträge von A ungleich Null sind. In dem Fall ist die Inverse von A wiederum eine obere Dreiecksmatrix. Die invertierbaren oberen Dreiecksmatrizen bilden also eine Untergruppe der Gruppe der invertierbaren Matrizen $\mathrm{GL}(n, K)$.

(*iii*) Es sei $A \in K^{n \times n}_{\triangle}$ eine untere Dreiecksmatrix. A ist genau dann invertierbar, wenn alle Diagonaleinträge von A ungleich Null sind. In dem Fall ist die Inverse von A wiederum eine untere Dreiecksmatrix. Die invertierbaren unteren Dreiecksmatrizen bilden also eine Untergruppe der Gruppe der invertierbaren Matrizen $\mathrm{GL}(n, K)$.

Beweis. Aussage (*i*): Es sei $A \in K^{n \times n}_{\diagdown}$ eine Diagonalmatrix. Wäre der Diagonaleintrag $a_{ii} = 0$ für ein $1 \leqslant i \leqslant n$, dann ist die i-te Zeile des Produkts $A\,B$ für jede Matrix $B \in K^{n \times n}$ der Nullvektor. Damit kann $A\,B$ nicht die Einheitsmatrix ergeben, d. h., A ist in diesem Fall nicht invertierbar. Es seien nun umgekehrt alle Diagonaleinträge von A ungleich Null. Dann ist

$$\begin{bmatrix} a_{11}^{-1} & & \\ & \ddots & \\ & & a_{nn}^{-1} \end{bmatrix} \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \begin{bmatrix} a_{11}^{-1} & & \\ & \ddots & \\ & & a_{nn}^{-1} \end{bmatrix} = I,$$

also ist wie behauptet die Diagonalmatrix bestehend aus den Inversen der Diagonaleinträge von A die Inverse von A.

Da das Produkt von Diagonalmatrizen wieder eine Diagonalmatrix ist (Lemma 15.38) und insbesondere die Einheitsmatrix eine invertierbare Diagonalmatrix ist, erfüllt die Menge der Diagonalmatrizen das Untergruppenkriterium (Satz 7.44) und ist damit eine Untergruppe von $\mathrm{GL}(n, K)$.

Aussage (*ii*): Es sei $A \in K^{n \times n}_{\triangledown}$ eine obere Dreiecksmatrix. Im Fall $n = 0$ ist A auch eine Diagonalmatrix, und die Aussage folgt aus Aussage (*i*). Es sei nun also $n \in \mathbb{N}$.

Schritt 1: Alle Diagonaleinträge von A seien ungleich Null. Wir zeigen, dass A dann invertierbar und die Inverse von A wiederum eine obere Dreiecksmatrix ist.

Es sei D die (nach Aussage (*i*) invertierbare) Diagonalmatrix mit den Diagonaleinträgen von A. Wir könnnen A in der Form

$$A = D\,(I + N)$$

schreiben, wobei N eine strikte obere Dreiecksmatrix ist. (**Quizfrage 15.8:** klar?)

Wir rechnen nun nach, dass die Inverse von $I + N$ gegeben ist durch[23]

$$(I + N)^{-1} = \sum_{k=0}^{n} (-N)^k. \tag{15.35}$$

Es gilt nämlich nach dem Distributiv- und dem Kommutativgesetz der Matrix-Addition und Matrix-Multiplikation

$$\begin{aligned}(I + N) \sum_{k=0}^{n} (-N)^k &= \sum_{k=0}^{n} (I + N)\,(-N)^k = \sum_{k=0}^{n} (-N)^k - \sum_{k=1}^{n+1} (-N)^k \\ &= (-N)^0 - (-N)^{n+1} = I - 0 = I.\end{aligned}$$

Die letzte Gleichheit gilt aufgrund von Lemma 15.36 (strikte Dreiecksmatrizen sind nilpotent). Ganz ähnlich können wir auch zeigen:

$$\begin{aligned}\Big(\sum_{k=0}^{n} (-N)^k\Big)\,(I + N) &= \sum_{k=0}^{n} (-N)^k\,(I + N) = \sum_{k=0}^{n} (-N)^k - \sum_{k=1}^{n+1} (-N)^k \\ &= (-N)^0 - (-N)^{n+1} = I - 0 = I.\end{aligned}$$

Tatsächlich ist also $I + N$ invertierbar und besitzt die in (15.35) angegebene Inverse. Diese ist aufgrund von Lemma 15.38 ebenfalls wieder eine obere Dreiecksmatrix.

Sind nun alle Diagonaleinträge von A ungleich Null, dann ist die Digonalmatrix D von A nach Aussage *(i)* invertierbar. Damit ist $A = D\,(I + N)$ invertierbar, und die Inverse

$$A^{-1} = (I + N)^{-1}\,D^{-1} = \sum_{k=0}^{n} (-N)^k\,D^{-1}$$

ist eine Summe oberer Dreiecksmatrizen, also in der Tat selbst eine obere Dreiecksmatrix.

Schritt 2: Es seien nun nicht alle Diagonaleinträge von A ungleich Null. Wir zeigen, dass A in diesem Fall nicht invertierbar ist.

Es sei dazu $1 \leqslant k \leqslant n$ der kleinste Index mit $a_{kk} = 0$. Im Fall $k = 1$ ist $a_{\bullet 1}$ eine Nullspalte, damit ist die Familie der Spaltenvektoren von A linear abhängig (vgl. Bemerkung 12.3). Im Fall $k \geqslant 2$ kann die Spalte $a_{\bullet k}$ als Linearkombination der vorhergehenden Spalten geschrieben werden, sodass (Lemma 12.5) auch in diesem Fall die Familie der Spaltenvektoren von A linear abhängig ist. Der Rang von A (die maximale Anzahl linear unabhängiger Spalten, siehe Lemma 15.11) erfüllt daher $\operatorname{Rang}(A) < n$.

Wäre nun A invertierbar, so gäbe es eine Matrix $B \in K^{n \times n}$ mit $A\,B = I_n$. Daraus folgt aber mit Satz 15.17 der Widerspruch

$$n = \operatorname{Rang}(I_n) = \operatorname{Rang}(A\,B) \leqslant \min\{\operatorname{Rang}(A),\ \operatorname{Rang}(B)\} < n.$$

[23] Vergleiche hierzu die geometrische Reihe aus Lehrveranstaltungen zur *Analysis*: $(1 - q)^{-1} = \sum_{k=0}^{\infty} q^k$ für $|q| < 1$. Im vorliegenden Fall ist $q = -N$ eine Matrix, und die Reihe wird zu einer endlichen Summe, weil N und damit auch $-N$ nilpotent ist (Lemma 15.36).

Da das Produkt von oberen Dreiecksmatrizen wieder eine obere Dreiecksmatrix ist (Lemma 15.38) und insbesondere die Einheitsmatrix eine invertierbare obere Dreiecksmatrix ist, erfüllt die Menge der oberen Dreiecksmatrizen das Untergruppenkriterium (Satz 7.44) und ist damit eine Untergruppe von $\mathrm{GL}(n, K)$.

Aussage (iii): Ist $A \in K^{n \times n}$ eine untere Dreiecksmatrix, dann ist $A^\intercal$ eine obere Dreiecksmatrix mit denselben Diagonaleinträgen. Nach Satz 15.42 ist A genau dann invertierbar, wenn $A^\intercal$ invertierbar ist, und in dem Fall gilt für die Inverse nach (15.34)

$$A^{-1} = \big((A^{-1})^\intercal\big)^\intercal = \big(\underbrace{(A^\intercal)^{-1}}_{\text{obere Dreiecksmatrix nach Aussage } (ii)}\big)^\intercal,$$

A^{-1} ist also eine untere Dreiecksmatrix. □

Wir können nun ein Invertierbarkeitskriterium für allgemeine Matrizen angeben. Dieses basiert auf der Beobachtung, dass die Umformung einer Matrix in Zeilenstufenform ihre Invertierbarkeit nicht ändert. Für Matrizen in Zeilenstufenform ist aber die Invertierbarkeit leicht zu erkennen:

Satz 15.45 (Invertierbarkeit von Matrizen erkennen durch Zeilenstufenform)**.**
Es seien K ein Körper, $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$. Weiter sei C eine zu A gehörige Matrix in Zeilenstufenform (Satz 15.23). Dann sind äquivalent:

(i) A ist invertierbar.

(ii) Es gilt $\operatorname{Rang}(A) = n$.

(iii) C ist invertierbar.

(iv) Es gilt $\operatorname{Rang}(C) = n$.

(v) C hat keine Nullzeilen und keine Nullspalten.

Beachte: Eine quadratische Matrix ist also genau dann invertierbar, wenn sie vollen Rang (Definition 15.14) besitzt. Das ist genau dann der Fall, wenn die Menge der Zeilenvektoren linear unabhängig ist, und auch genau dann, wenn die Menge der Spaltenvektoren linear unabhängig ist.

Beweis. Es sei C eine zu A gehörige Matrix in Zeilenstufenform. C ist also aus A durch Multiplikation von links mit Elementarmatrizen E_i hervorgegangen:

$$C = E_k \cdots E_2 \, E_1 \, A.$$

Aussage (i) ⇔ Aussage (iii): Da die Elementarmatrizen nach Lemma 15.43 invertierbar sind, ist A genau dann invertierbar, wenn C invertierbar ist. (**Quizfrage 15.9:** Klar?)

Aussage (ii) ⇔ Aussage (iv): Wie in Lemma 15.19 gezeigt wurde, ändern elementare Zeilenumformungen den Zeilenraum und insbesondere den Rang nicht. Es gilt also $\operatorname{Rang}(A) = \operatorname{Rang}(C)$.

[Aussage (*iv*)](#) $\Rightarrow$ Aussage (*iii*): Im Fall $n = 0$ ist die einzig mögliche Matrix C die leere Matrix, diese hat $\operatorname{Rang}(C) = 0$ und ist selbstinvers. Für den Rest des Beweisschrittes betrachten wir nun $n \in \mathbb{N}$. $\operatorname{Rang}(C) = n$ bedeutet, dass die Zeilenstufenform von C die Gestalt

$$\begin{bmatrix} \star & ? & ? \\ 0 & \ddots & ? \\ 0 & 0 & \star \end{bmatrix}$$

besitzt, vgl. Beispiel 15.22. Dabei sind die Pivot-Elemente $\star \in K \setminus \{0\}$. Nach Lemma 15.44 ist C also invertierbar.

Aussage (*iv*) $\Leftrightarrow$ Aussage (*v*): Wir wissen: $\operatorname{Rang}(C)$ ist die Anzahl der Pivot-Elemente von C. Da C eine $n \times n$-Matrix ist, gilt:

$$\begin{aligned} & \operatorname{Rang}(C) = n \\ \Leftrightarrow \quad & C \text{ hat } n \text{ Pivot-Elemente} \\ \Leftrightarrow \quad & \text{die Hauptdiagonale von } C \text{ hat keine Nullen} \\ \Leftrightarrow \quad & C \text{ hat keine Nullzeilen und keine Nullspalten.} \end{aligned}$$

Aussage (*iii*) $\Rightarrow$ Aussage (*v*): Wir argumentieren mit einem Beweis durch Kontraposition. Angenommen, C habe eine Nullzeile. Dann hat $C\,X$ für jede Matrix $X \in K^{n \times n}$ ebenfalls eine Nullzeile, kann also nicht die Einheitsmatrix I_n sein. Also ist C nicht invertierbar.

Angenommen, C habe eine Nullspalte. Dann hat $X\,C$ für jede Matrix $X \in K^{n \times n}$ ebenfalls eine Nullspalte, kann also nicht die Einheitsmatrix I_n sein. Also ist C nicht invertierbar. □

Folgerung 15.46 (Multiplikation mit invertierbaren Matrizen ändert den Rang nicht)**.**
Es seien K ein Körper und $m, n \in \mathbb{N}_0$. Für beliebige Matrizen $A \in K^{n \times m}$ und invertierbare Matrizen $B \in K^{n \times n}$, $C \in K^{m \times m}$ gilt:

$$\operatorname{Rang}(B\,A\,C) = \operatorname{Rang}(A). \tag{15.36}$$

Beweis. Der Beweis ist Gegenstand der Übung. □

Abschließend zeigen wir, dass es für den Nachweis, dass zwei $n \times n$-Matrizen A und B Inverse voneinander sind, ausreichend ist, $A\,B = I$ *oder* $B\,A = I$ nachzuweisen.[24] Mit anderen Worten, jede **Rechtsinverse** einer quadratischen Matrix ist auch eine **Linksinverse** (und damit die eindeutige Inverse) und umgekehrt. Das ist eine durchaus bemerkenswerte Eigenschaft des Matrix-Monoids $(K^{n \times n}, \cdot)$, die in allgemeinen Monoiden nicht gilt (Übung).

Satz 15.47 (Rechtsinverse quadratischer Matrizen sind Linksinverse und umgekehrt)**.**
Es seien K ein Körper, $n \in \mathbb{N}_0$ und $A, B \in K^{n \times n}$. Dann sind äquivalent:

(*i*) B ist eine Rechtsinverse von A, d. h., es gilt $A\,B = I_n$.

(*ii*) B ist eine Linksinverse von A, d. h., es gilt $B\,A = I_n$.

[24] Wenn wir bereits wüssten, dass A oder B invertierbar ist, dann wäre das wegen Lemma 7.19 klar.

(iii) B ist die Inverse von A, d. h., es gilt $A\,B = B\,A = I_n$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (iii): Es sei $A\,B = I_n$. Dann gilt

$$\begin{aligned} n &= \operatorname{Rang}(I_n) && \text{nach Satz 15.45 } (I_n \text{ ist invertierbar und hat daher vollen Rang}) \\ &= \operatorname{Rang}(A\,B) && I_n = A\,B \text{ nach Voraussetzung} \\ &\le \min\{\operatorname{Rang}(A),\ \operatorname{Rang}(B)\} && \text{nach Satz 15.17 (Rang des Produkts von Matrizen)} \\ &\le n && \text{nach Satz 15.13 (Rang ist beschränkt durch die Dimensionen).} \end{aligned}$$

Es muss also überall Gleichheit gelten, insbesondere ist $\operatorname{Rang}(A) = \operatorname{Rang}(B) = n$, und nach Satz 15.45 sind A und B invertierbar. Wir müssen noch nachweisen, dass B tatsächlich die Inverse von A ist. Es gilt nämlich $A\,A^{-1} = I$, aber nach Voraussetzung auch $A\,B = I_n$. Nach Kürzungsregel (15.31a) muss $A^{-1} = B$ sein.

Der Beweis von Aussage (ii) $\Rightarrow$ Aussage (iii) geht analog.

Aussage (iii) $\Rightarrow$ Aussage (i) und Aussage (iii) $\Rightarrow$ Aussage (ii) sind klar. □

Ende der Vorlesung 21

Ende der Woche 10

§ 16 Lineare Gleichungssysteme

Literatur: Fischer, Springborn, 2020, Kapitel 1.4 und 3.3; Bosch, 2014, Kapitel 3.5; Beutelspacher, 2014, Kapitel 4.1; Deiser, 2024b, Kapitel 3.3; Jänich, 2008, Kapitel 7

Sehr viele Aufgabenstellungen in den quantitativen Wissenschaften führen früher oder später auf lineare Gleichungssysteme.

Definition 16.1 (lineares Gleichungssystem)**.**
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n \times m}$.

(i) Eine Gleichung der Form $A\,x = b$ mit dem unbekannten Koeffizientenvektor $x \in K^m$ heißt ein **lineares Gleichungssystem** (englisch: linear system of equations) oder kurz **LGS**. A heißt die **Koeffizientenmatrix** (englisch: coefficient matrix) und $b \in K^n$ der **Vektor der rechten Seite** (englisch: right-hand side vector) oder kurz die **rechte Seite** (englisch: right-hand side) des LGS. Der unbekannte Koeffizientenvektor x heißt auch die **Variable** (englisch: variable) des LGS.

(ii) Die Matrix $[A, b] \in K^{n \times (m+1)}$ heißt die **erweiterte Koeffizientenmatrix** (englisch: augmented coefficient matrix) zum LGS $A\,x = b$.

(iii) Das LGS $A\,x = b$ heißt **homogen** (englisch: homogeneous), wenn $b = 0 \in K^m$ ist, andernfalls **nichthomogen** (englisch: non-homogeneous) oder **inhomogen** (englisch: inhomogeneous).

(iv) Das LGS $A\,x = b$ heißt **lösbar** (englisch: solvable), wenn es ein $x_0 \in K^m$ gibt, das $A\,x_0 = b$ erfüllt, andernfalls **unlösbar** oder **nicht lösbar** (englisch: unsolvable). △

Beispiel 16.2 (lineares Gleichungssystem)**.**

(*i*) Wir betrachten folgendes Beispiel mit $m = 3$ (Anzahl der Gleichungen) und $n = 3$ (Anzahl der unbekannten Koeffizienten) über dem Körper $\mathbb{R}$:

$$\left.\begin{array}{rcrcrcr} 6\,x_1 & + & 1\,x_2 & + & 1\,x_3 & = & 11 \\ 3\,x_1 & + & 3\,x_2 & + & 1\,x_3 & = & 5 \\ -6\,x_1 & - & 6\,x_2 & - & 3\,x_3 & = & -9 \end{array}\right\} \quad\Leftrightarrow\quad \underbrace{\begin{bmatrix} 6 & 1 & 1 \\ 3 & 3 & 1 \\ -6 & -6 & -3 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 11 \\ 5 \\ -9 \end{pmatrix}}_{b}.$$

(*ii*) Wie in Beispiel 11.9 bemerkt, führt die Bestimmung der unbekannten Koeffizienten in einer Linearkombination ebenfalls auf ein lineares Gleichungssystem. Wollen wir etwa den Vektor $\left(\begin{smallmatrix} 3 \\ -7 \end{smallmatrix}\right)$ im $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ als Linearkombination der Vektoren $\left\{\left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 4 \\ 1 \end{smallmatrix}\right)\right\}$ darstellen, so müssen die gesuchten Koeffizienten $x_1, x_2 \in \mathbb{R}$ das LGS

$$\underbrace{\begin{bmatrix} 2 & 4 \\ -1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 3 \\ -7 \end{pmatrix}}_{b}$$

erfüllen. △

Beachte: Jede Spalte der Matrix gehört zu einer der Variablen $x_1, \ldots, x_m$. Jede Zeile der Matrix gehört zu einer Gleichung. Die Matrix $A \in \mathbb{R}^{n \times m}$, die rechte Seite $b \in K^n$ und die Variable $x \in K^m$ sind alles Objekte über demselben Körper K.

Unser Ziel ist es, *alle* Lösungen von $A\,x = b$ zu bestimmen, also die gesamte **Lösungsmenge** (englisch: solution set)

$$\mathcal{L}(A, b) \coloneqq \{x \in K^m \mid A\,x = b\}. \tag{16.1}$$

Dabei spielt die Zeilenstufenform der Koeffizientenmatrix A bzw. der erweiterten Koeffizientenmatrix $[A, b]$ eine entscheidende Rolle, wie wir am Beweis des folgenden Satzes und auch anschließend beim Lösungsverfahren sehen werden.

Satz 16.3 (Struktur der Lösungsmenge, Lösbarkeit)**.**
Es seien K ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n \times m}$ und $b \in K^n$.

(*i*) $\mathcal{L}(A, 0)$ ist ein Unterraum von K^m der Dimension $m - \operatorname{Rang}(A)$.

(*ii*) Ist $x_0 \in K^m$ irgendeine („**partikuläre**") **Lösung** (englisch: particular solution) von $A\,x = b$, dann gilt

$$\mathcal{L}(A, b) = x_0 + \mathcal{L}(A, 0). \tag{16.2}$$

Beachte: Die Lösungsmenge eines allgemeinen Systems $A\,x = b$ ergibt sich aus einer partikulären Lösung plus der Lösungsmenge des zugehörigen homogenen Systems.

(*iii*) Die folgenden Aussagen sind äquivalent:

(a) $A\,x = b$ ist lösbar.
(b) $b \in \operatorname{SR}(A)$.
(c) $\operatorname{Rang}(A) = \operatorname{Rang}([A, b])$.

(*iv*) Die folgenden Aussagen sind äquivalent:

(a) $A\,x = b$ ist eindeutig lösbar.
(b) $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = m$.

(*v*) Ist A quadratisch, gilt also $m = n$, dann sind die folgenden Aussagen äquivalent:

(a) $A\,x = b$ ist eindeutig lösbar.
(b) $A\,x = c$ ist für jedes $c \in K^n$ eindeutig lösbar.
(c) A ist invertierbar.

In diesem Fall ist die eindeutige Lösung von $A\,x = b$ gegeben durch $x = A^{-1}b$.

Beweis. Wir zerlegen den Beweis von Aussage (*i*) in mehrere Schritte.

Schritt 1: $\mathcal{L}(A, 0)$ ist ein Unterraum von K^m:

Das folgt aus dem Unterraumkriterium (Satz 11.11), denn: Gehören die Vektoren $x, y \in K^m$ zu $\mathcal{L}(A, 0)$, erfüllen also die Bedingungen $A\,x = 0$ und $A\,y = 0$, und sind $\alpha, \beta \in K$, dann gehört auch $\alpha\,x + \beta\,y$ zu $\mathcal{L}(A, 0)$, denn es gilt

$$A\,(\alpha\,x + \beta\,y) = A\,(\alpha\,x) + A\,(\beta\,y) = \alpha\,A\,x + \beta\,A\,y = 0.$$

Schritt 2: Wir bringen A in Zeilenstufenform C ohne Nullzeilen und zeigen $\mathcal{L}(A, 0) = \mathcal{L}(C, 0)$:

Es sei dazu $\widehat{C} = E_k \cdots E_2\,E_1\,A$ eine zu A gehörige Zeilenstufenform. Da die Elementarmatrizen E_i invertierbar sind, gilt $A\,x = 0 \Leftrightarrow \widehat{C}\,x = E_k \cdots E_2\,E_1\,x = 0$, also $\mathcal{L}(A, 0) = \mathcal{L}(\widehat{C}, 0)$. Zur Abkürzung setzen wir $r \coloneqq \operatorname{Rang}(A) = \operatorname{Rang}(\widehat{C})$, $0 \leqslant r \leqslant \min\{m, n\}$. Eventuell in $\widehat{C}$ auftretende Nullzeilen können wir streichen und erhalten die Matrix $C \in K^{r \times m}$ in Zeilenstufenform ohne Nullzeilen. Es gilt weiterhin $\mathcal{L}(A, 0) = \mathcal{L}(\widehat{C}, 0) = \mathcal{L}(C, 0)$.

Schritt 3: Wir bestätigen $\dim(\mathcal{L}(C, 0)) = m - r$, indem wir eine Basis dieses Unterraumes angeben:

Wir bezeichnen mit $\mathcal{A} \coloneqq (j_1, \ldots, j_r)$ die Familie der Pivot-Spalten (in aufsteigender Reihenfolge) und mit $\mathcal{I}$ die komplementäre Familie der Nicht-Pivot-Spalten $\mathcal{I} \coloneqq (k_1, \ldots, k_{m-r})$. Wir können jeden Vektor $x \in K^m$ in zwei Teilvektoren $x_{\mathcal{A}} \in K^r$ und $x_{\mathcal{I}} \in K^{m-r}$ zerlegen. Dies geschieht durch Einführung der beiden Matrizen

$$\Pi_{\mathcal{A}} \coloneqq [e_{j_i}^\intercal]_{i \in [\![1,r]\!]} \in K^{r \times m} \quad \text{und} \quad \Pi_{\mathcal{I}} \coloneqq\in K^{(m-r) \times m} [e_{k_i}^\intercal]_{i \in [\![1,m-r]\!]},$$

die aus komplementären Zeilen der $m \times m$-Einheitsmatrix bestehen. (**Quizfrage 16.1:** Können Sie sich davon überzeugen, dass die nachfolgenden Ausführungen auch in den Grenzfällen $r = 0$ und $r = m$ Sinn ergeben?) Dann gilt

$$x_{\mathcal{A}} = \Pi_{\mathcal{A}}\,x \quad \text{und} \quad x_{\mathcal{I}} = \Pi_{\mathcal{I}}\,x \quad \text{sowie} \quad x = \Pi_{\mathcal{A}}^\intercal x_{\mathcal{A}} + \Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}}.$$

Die wesentliche Erkenntnis hierbei ist, dass wir

$$C\,x = C\,\Pi_{\mathcal{A}}^\intercal x_{\mathcal{A}} + C\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}}$$

schreiben können, wobei die Matrix $C\,\Pi_{\mathcal{A}}^\intercal = [c_{\bullet j_1}, \ldots, c_{\bullet j_r}]$ aus den Pivot-Spalten von C besteht und damit eine invertierbare obere Dreiecksmatrix (Satz 15.45) der

Dimension $r \times r$ darstellt. Wir können also $C\,x = 0$ nach der Teilvariablen $x_{\mathcal{A}}$ auflösen:

$$\begin{aligned} & C\,x = 0 \\ \Leftrightarrow\quad & C\,\Pi_{\mathcal{A}}^\intercal x_{\mathcal{A}} + C\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}} = 0 \\ \Leftrightarrow\quad & C\,\Pi_{\mathcal{A}}^\intercal x_{\mathcal{A}} = -C\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}} \\ \Leftrightarrow\quad & x_{\mathcal{A}} = -\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}}, \end{aligned}$$

wobei $x_{\mathcal{I}} \in K^{m-r}$ frei gewählt werden kann. Setzen wir diese Darstellung in $x = \Pi_{\mathcal{A}}^\intercal x_{\mathcal{A}} + \Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}}$ ein, so erhalten wir durch

$$x = -\Pi_{\mathcal{A}}^\intercal\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}} + \Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}} = \Big[I_m - \Pi_{\mathcal{A}}^\intercal\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\Big]\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}} \tag{16.3}$$

mit beliebigem $x_{\mathcal{I}} \in K^{m-r}$ genau die Lösungen von $C\,x = 0$.

Wir zeigen nun, dass die Spalten der Matrix

$$\Big[I_m - \Pi_{\mathcal{A}}^\intercal\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\Big]\,\Pi_{\mathcal{I}}^\intercal \in K^{n\times(m-r)} \tag{16.4}$$

linear unabhängig sind und damit eine Basis von $\mathcal{L}(C,0) = \mathcal{L}(A,0)$ bilden. Eine solche Spalte ergibt sich durch Multiplikation der Matrix aus (16.4) mit einem der Standardbasisvektoren e_i von K^{m-r}. Die lineare Unabhängigkeit der Spalten folgt aus

$$\begin{aligned} & \sum_{i=1}^{m-r} \alpha_i\,\Big[I_m - \Pi_{\mathcal{A}}^\intercal\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\Big]\,\Pi_{\mathcal{I}}^\intercal e_i = 0 \\ \Rightarrow\quad & \Pi_{\mathcal{I}} \sum_{i=1}^{m-r} \alpha_i\,\Big[I_m - \Pi_{\mathcal{A}}^\intercal\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\Big]\,\Pi_{\mathcal{I}}^\intercal e_i = 0 \\ \Rightarrow\quad & \sum_{i=1}^{m-r} \alpha_i\,\Pi_{\mathcal{I}}\,\Big[I_m - \Pi_{\mathcal{A}}^\intercal\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\Big]\,\Pi_{\mathcal{I}}^\intercal e_i = 0 \\ \Rightarrow\quad & \sum_{i=1}^{m-r} \alpha_i\,\Big[\underbrace{\Pi_{\mathcal{I}}\,\Pi_{\mathcal{I}}^\intercal}_{=I_{m-r}} - \underbrace{\Pi_{\mathcal{I}}\Pi_{\mathcal{A}}^\intercal}_{=0}\big(C\,\Pi_{\mathcal{A}}^\intercal\big)^{-1} C\,\Pi_{\mathcal{I}}^\intercal\Big]\,e_i = 0 \\ \Rightarrow\quad & \sum_{i=1}^{m-r} \alpha_i\,e_i = 0. \end{aligned}$$

Da die Menge der Standardbasisvektoren e_i von K^{m-r} linear unabhängig ist, folgt $\alpha_i = 0$ für alle $i \in [\![1, m-r]\!]$.

Damit haben wir gezeigt, dass die Spalten der Matrix (16.4) eine Basis der Kardinalität $m - r$ von $\mathcal{L}(C,0) = \mathcal{L}(A,0)$ bilden.

Aussage *(ii)*: Es sei $x_0 \in K^m$ mit $A\,x_0 = b$ gegeben. Dann gilt

$$\begin{aligned} & x \in \mathcal{L}(A,b) \\ \Leftrightarrow\quad & A\,x = b \\ \Leftrightarrow\quad & A\,(x - x_0) = 0 \\ \Leftrightarrow\quad & x - x_0 \in \mathcal{L}(A,0) \\ \Leftrightarrow\quad & x \in x_0 + \mathcal{L}(A,0). \end{aligned}$$

Nun zu Aussage (iii):

Aussage (a) $\Leftrightarrow$ Aussage (b): $A\,x = b$ ist lösbar genau dann, wenn b als Linearkombination der Spalten von A darstellbar ist, also genau dann, wenn $b \in \mathrm{SR}(A)$ liegt.

Aussage (b) $\Leftrightarrow$ Aussage (c): Die Hinzunahme des Spaltenvektors b zu einer Matrix A erhöht genau dann den $\mathrm{Rang}(A) = \mathrm{SRang}(A) = \dim(\mathrm{SR}(A))$ nicht, wenn b bereits in $\mathrm{SR}(A)$ enthalten ist.

Wir beweisen Aussage (iv):

Aussage (a) $\Rightarrow$ Aussage (b): Es besitze $A\,x = b$ eine eindeutige Lösung x_0. Aufgrund der Lösbarkeit folgt aus Aussage (iii) $\mathrm{Rang}(A) = \mathrm{Rang}([A, b])$. Aufgrund der Eindeutigkeit der Lösung und der Darstellung (16.2) $\mathcal{L}(A, b) = x_0 + \mathcal{L}(A, 0)$ muss $\mathcal{L}(A, 0) = \{0\}$, also $\dim(\mathcal{L}(A, 0)) = 0$ sein. Das bedeutet nach Aussage (i) aber gerade $m = \mathrm{Rang}(A)$.

Aussage (b) $\Rightarrow$ Aussage (a): Es gelte $\mathrm{Rang}(A) = \mathrm{Rang}([A, b]) = m$. Dann ist aufgrund von Aussage (iii) $A\,x = b$ lösbar. Aufgrund der Darstellung (16.5) und $\dim(\mathcal{L}(A, 0)) = m - \mathrm{Rang}(A) = 0$ folgt, dass die Lösung eindeutig ist.

Schließlich Aussage (v): Es sei nun A quadratisch, also $m = n$.

Aussage (a) $\Rightarrow$ Aussage (c): Es sei $A\,x = b$ eindeutig lösbar. Nach Aussage (iv) folgt, dass $\mathrm{Rang}(A) = \mathrm{Rang}([A, b]) = n$ gilt. $\mathrm{Rang}(A) = n$ impliziert nach Satz 15.45 aber, dass A invertierbar ist.

Aussage (c) $\Rightarrow$ Aussage (b): Wenn A invertierbar ist, dann ist $A\,x = c$ äquivalent zu $A^{-1}A\,x = x = A^{-1}c$. Also ist $A\,x = c$ für jedes $c \in K^n$ eindeutig lösbar.

Aussage (b) $\Rightarrow$ Aussage (a): Das ist offensichtlich. □

Bemerkung 16.4 (zu Satz 16.3 über die Struktur der Lösungsmenge eines linearen Gleichungssystems)**.**

(i) Wir illustrieren den Beweis von Aussage (i) aus Satz 16.3 anhand eines Beispiels (vgl. Beispiel 15.22). Die Zeilenstufenform von A habe die Gestalt

$$\widehat{C} = \begin{bmatrix} \star & ? & ? & ? \\ 0 & 0 & \star & ? \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und nach Streichen von Nullenzeilen} \quad C = \begin{bmatrix} \star & ? & ? & ? \\ 0 & 0 & \star & ? \end{bmatrix}.$$

Hieraus lesen wir ab:

- Anzahl der Variablen $m = 4$
- Rang der Matrix $\mathrm{Rang}(A) = r = 2$
- Familie der Pivot-Spalten $\mathcal{A} = (1, 3)$
- Familie der Nicht-Pivot-Spalten $\mathcal{I} = (2, 4)$

Die Auswahlmatrizen $\Pi_{\mathcal{A}}$ und $\Pi_{\mathcal{I}}$ haben also die Form

$$\Pi_{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{und} \quad \Pi_{\mathcal{I}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Es gilt

$$x_{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} \quad \text{und} \quad x_{\mathcal{I}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix}$$

sowie

$$\begin{aligned} x &= \Pi_{\mathcal{A}}^{\mathsf{T}} \, x_{\mathcal{A}} + \Pi_{\mathcal{I}}^{\mathsf{T}} \, x_{\mathcal{I}} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} x_1 \\ 0 \\ x_3 \\ 0 \end{pmatrix}}_{\text{abhängige Variablen}} + \underbrace{\begin{pmatrix} 0 \\ x_2 \\ 0 \\ x_4 \end{pmatrix}}_{\text{unabhängige Variablen}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \end{aligned}$$

Wir nennen die Variablen x_2 und x_4 auch die **unabhängigen Variablen** (englisch: independent variables), während x_1 und x_3 als die **abhängigen Variablen** (englisch: dependent variables) bezeichnet werden.

Die Teilmatrix $C \, \Pi_{\mathcal{A}}^{\mathsf{T}}$ (Auswahl der $\mathcal{A}$-Spalten von C) hat die Gestalt

$$\begin{bmatrix} \star & ? \\ 0 & \star \end{bmatrix}$$

einer invertierbaren oberen $r \times r$-Dreiecksmatrix.

Wir halten fest, dass die kompliziert aussehende Darstellung

$$\left[I_m - \Pi_{\mathcal{A}}^{\mathsf{T}} \big(C \, \Pi_{\mathcal{A}}^{\mathsf{T}} \big)^{-1} C \right] \Pi_{\mathcal{I}}^{\mathsf{T}} \in K^{n \times (m-r)} \tag{16.5}$$

einer spaltenweisen Basis von $\mathcal{L}(A, 0)$ praktisch Folgendes bedeutet: Wir setzen genau eine der unabhängigen Variablen x_i auf den Wert 1 und die anderen unabhängigen Variablen auf den Wert 0. (Das entspricht einer Spalte von $\Pi_{\mathcal{I}}^{\mathsf{T}}$.) Dann rechnen wir die Werte der abhängigen Variablen $x_{\mathcal{A}}$ mit Hilfe der Gleichung

$$\big(C \, \Pi_{\mathcal{A}}^{\mathsf{T}} \big) \, x_{\mathcal{A}} = - \underbrace{C \, \Pi_{\mathcal{I}}^{\mathsf{T}} x_{\mathcal{I}}}_{\text{genau eine der } \mathcal{I}\text{-Spalten von } C} \tag{16.6}$$

aus. Weil $C \, \Pi_{\mathcal{A}}^{\mathsf{T}}$ eine obere Dreiecksmatrix ist, können wir die Werte der abhängigen Variablen $x_{\mathcal{A}}$ von hinten nach vorne bestimmen und an der richtigen Stelle in den Lösungsvektor einsortieren ($\Pi_{\mathcal{A}}^{\mathsf{T}} x_{\mathcal{A}}$). Ein Beispiel folgt in Beispiel 16.10.

(ii) Die Dimension des Lösungsraumes $\dim(\mathcal{L}(A, 0)) = m - \operatorname{Rang}(A)$ sollten wir uns merken in der Form „Anzahl der Variablen (m) minus Anzahl der wesentlichen (sprich: linear unabhängigen) Gleichungen ($\operatorname{Rang}(A)$) im System $A \, x = b$“.

(iii) Bei einem linearen Gleichungssystem $A\,x = b$ können drei mögliche Fälle auftreten:

(1) Das System ist eindeutig lösbar.
(2) Das System ist lösbar, aber nicht eindeutig lösbar. In diesem Fall hat die Lösungsmenge die Struktur (16.2), also irgendeine beliebige, aber feste Lösung x_0 von $A\,x = b$, plus den Unterraum $\mathcal{L}(A, 0)$ der Dimension $m - \operatorname{Rang}(A) \geqslant 1$.[25]
(3) Das System ist nicht lösbar. △

Wie berechnet man nun praktisch die Lösungsmenge eines linearen Gleichungssystems $A\,x = b$ bzw. stellt fest, dass das System nicht lösbar ist? Dazu gehen wir wie folgt vor:

(1) Wir bringen die erweiterte Koeffizientenmatrix $[A, b]$ zunächst in Zeilenstufenform. An ihr können wir bereits die Lösbarkeit und ggf. die Dimension des Lösungsraumes ablesen.

(2) Wenn das System lösbar ist, so überführen wir die erweiterte Koeffizientenmatrix in die sogenannte **reduzierte Zeilenstufenform**, aus der wir die Lösungsmenge ablesen können.[26]

Dieses Verfahren heißt **Gaußsches Eliminationsverfahren** (englisch: Gaussian elimination method).[27]

Definition 16.5 (reduzierte Zeilenstufenform, vgl. Definition 15.20)**.**
Es seien K ein Körper und $m, n \in \mathbb{N}_0$. Eine Matrix $A \in K^{n\times m}$ heißt in **reduzierter Zeilenstufenform** (englisch: reduced row echelon form), wenn sie in Zeilenstufenform ist (Definition 15.20) und zusätzlich gilt:

(i) Alle Pivot-Elemente sind gleich 1.

(ii) Elemente, die in einer Spalte oberhalb eines Pivot-Elements stehen, sind gleich 0. △

Beachte: Die reduzierte Zeilenstufenform einer Matrix ist eindeutig. Eine Matrix ist genau dann invertierbar, wenn ihre reduzierte Zeilenstufenform die Einheitsmatrix ist (Beweis in Folgerung 16.8).

Beispiel 16.6 (reduzierte Zeilenstufenform, vgl. Beispiel 15.22 zur Zeilenstufenform)**.**
Nachfolgend sind die Besetzungsmuster einiger 3×4-Matrizen in reduzierter Zeilenstufenform zu sehen, wobei ? jeweils für einen Eintrag aus dem Körper K steht. Die Zahl r gibt wieder den Rang der Matrix an.

$$\begin{array}{r}j_1 = 1 \longrightarrow \\ j_2 = 2 \longrightarrow \\ j_3 = 3 \longrightarrow\end{array}\begin{bmatrix} 1 & 0 & 0 & ? \\ 0 & 1 & 0 & ? \\ 0 & 0 & 1 & ? \end{bmatrix} \qquad \begin{array}{r}j_1 = 1 \longrightarrow \\ j_2 = 3 \longrightarrow \\ \end{array}\begin{bmatrix} 1 & ? & 0 & ? \\ 0 & 0 & 1 & ? \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{array}{r}j_1 = 3 \longrightarrow \\ j_2 = 4 \longrightarrow \\ \end{array}\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$r = 3 \qquad\qquad r = 2 \qquad\qquad r = 2$$
△

[25]Die Lösungsmenge $\mathcal{L}(A, b)$ ist dann also eine unendliche Menge, falls der Körper K eine unendliche Menge ist. Ist dagegen der Körper K endlich, dann gilt $\#\mathcal{L}(A, b) = (\#K)^{m-\operatorname{Rang}(A)}$.

[26]Der zusätzliche Aufwand für die Überführung in die reduzierte Zeilenstufenform ist gerechtfertigt, weil dann die $\mathcal{A}$-Spalten $C\,\Pi_{\mathcal{A}}^{\mathsf{T}}$, aus denen sich — siehe (16.6) — die Werte der abhängigen Variablen ergeben, nicht nur eine obere Dreiecksmatrix, sondern die Einheitsmatrix ist, sodass wir die Werte direkt ablesen können.

[27]Das Wort „Elimination“ bezieht sich darauf, dass durch die Überführung in Zeilenstufenform Variablen aus den Gleichungen derart eliminiert wurden (erkennbar an Null-Koeffizienten), dass zwischen benachbarten Zeilen von unten nach oben immer nur eine Variable (Pivot-Spalte) hinzukommt, nach der diese Zeile (Gleichung) aufgelöst werden kann.

Wir können Algorithmus D.1 erweitern, um die reduzierte Zeilenstufenform einer Matrix zu berechnen, siehe Algorithmus D.3. Mit desen Hilfe können wir analog zu Satz 15.23 beweisen:

Satz 16.7 (Erreichbarkeit der reduzierten Zeilenstufenform, vgl. Satz 15.23)**.**
Es seien K ein Körper und $n, m \in \mathbb{N}_0$. Jede Matrix $A \in K^{n\times m}$ lässt sich durch elementare Zeilenumformungen in eine Matrix $C \in K^{n\times m}$ in reduzierter Zeilenstufenform überführen. Ist $r \in [\![0, n]\!]$ die Anzahl der Nicht-Nullzeilen in C, dann bilden die Zeilenvektoren $c_{1\bullet}, \ldots, c_{r\bullet}$ eine Basis von $\operatorname{ZR}(A) = \operatorname{ZR}(C)$, und es gilt $\operatorname{Rang}(A) = \operatorname{Rang}(C) = r$.

Folgerung 16.8 (invertierbare Matrizen sind Produkte von Elementarmatrizen)**.**
Es seien K ein Körper, $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Dann sind äquivalent:

(i) Die Matrix A ist invertierbar.

(ii) Die reduzierte Zeilenstufenform von A ist die Einheitsmatrix I_n.

(iii) Es gibt Elementarmatrizen $E_1, \ldots, E_k \in K^{n\times n}$, $k \in \mathbb{N}_0$, so dass gilt:

$$A = E_1\, E_2 \cdots E_k.$$

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Ist A invertierbar, dann gilt nach Satz 15.45 $\operatorname{Rang}(A) = n$. Nach Satz 16.7 besitzt auch die reduzierte Zeilenstufenform von A den Rang n, ist also die Einheitsmatrix I_n.

Aussage (ii) $\Rightarrow$ Aussage (iii): Es sei $C \in K^{n\times n}$ die reduzierte Zeilenstufenform von A. Diese ist aus A durch elementare Zeilenumformungen entstanden, also gibt es Elementarmatrizen $G_1, \ldots, G_k \in K^{n\times n}$, $k \in \mathbb{N}_0$, so dass gilt:

$$C = I_n = G_k \cdots G_2\, G_1\, A.$$

Da Elementarmatrizen invertierbar sind (Lemma 15.43), folgt

$$A = G_1^{-1}\, G_2^{-1} \cdots G_k^{-1}.$$

Da die Inversen $E_i \coloneqq G_i^{-1}$, $i = 1, \ldots, k$, ebenfalls Elementarmatrizen sind (Lemma 15.43), folgt die gewünschte Darstellung.

Aussage (iii) $\Rightarrow$ Aussage (i): A ist als Produkt invertierbarer Elementarmatrizen wieder invertierbar (Satz 15.42). □

Bemerkung 16.9 (Erzeugendensystem der allgemeinen linearen Gruppe)**.**

(i) Die Folgerung 16.8 besagt, dass die Elementarmatrizen ein Erzeugendensystem der allgemeinen linearen Gruppe $\operatorname{GL}(n, K)$ der invertierbaren Matrizen bilden.

(ii) Da Elementarmatrizen vom Typ III ihrerseits als Produkte von Elementarmatrizen vom Typ I und Typ II dargestellt werden können, folgt, dass bereits die Elementarmatrizen vom Typ I und Typ II ein Erzeugendensystem von $\operatorname{GL}(n, K)$ bilden. △

(**Quizfrage 16.2:** Welche Untergruppe wird alleine von den Elementarmatrizen vom Typ I erzeugt? Welche von den Elementarmatrizen vom Typ II?)

Wir geben nun anhand von Beispielen die Bestimmung der reduzierten Zeilenstufenform für eine erweiterte Koeffizientenmatrizen $[A, b]$ an und wie wir daraus die Lösungsmenge des linearen Gleichungssystems $A\,x = b$ ablesen können. Dabei kommen alle drei Fälle (eindeutige Lösbarkeit, nicht-eindeutige Lösbarkeit und Nicht-Lösbarkeit) vor.

Beispiel 16.10 (Gaußsches Eliminationsverfahren)**.**
Wir betrachten drei lineare Gleichungssysteme über dem endlichen Körper $(\mathbb{Z}_5, +_5, \cdot_5)$, vgl. Folgerung 10.4. Der Einfachheit halber schreiben wir die Verknüpfungen als $+$ und $\cdot$ (statt $+_5$ und $\cdot_5$) und wiederholen sie hier:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

In allen Fällen werden wir $m = 3$ Variablen und $n = 3$ Gleichungen verwenden.

(i) Wir betrachten das lineare Gleichungssystem

$$\left[\begin{array}{ccc|c} 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 2 \\ 3 & 2 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 2 & 2 & 2 \end{array}\right] \quad \text{Tauschen der Zeilen 1 und 2}$$

$$2 \; \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 2 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 3 & 2 & 1 \end{array}\right] \quad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements}\end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 3 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \quad \text{Tauschen der Zeilen 2 und 3}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \quad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements}\end{array}$$

An dieser Stelle ist eine Zeilenstufenform hergestellt. Wir erkennen $\operatorname{Rang}(A) = 3 = m$, also ist die Matrix invertierb und das System für jede rechte Seite eindeutig lösbar. Wir gehen weiter zur reduzierten Zeilenstufenform:

$$3 \; \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \quad \text{Normierung des Pivot-Elements}$$

$$\begin{array}{c} \\ 3 \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Erzeugen von Nullen oberhalb des Pivot-Elements

$$\begin{array}{c} \\ \\ 2 \end{array}\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Normierung des Pivot-Elements

$$\begin{array}{c} 2 \\ \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Erzeugen von Nullen oberhalb des Pivot-Elements

Wir haben in diesem Fall also die abhängigen Indizes $\mathcal{A} = (1, 2, 3)$ und keinen unabhängigen Index, also die leere Familie $\mathcal{I} = ()$. Hier können wir nun die eindeutige Lösung ablesen, nämlich $x = (2, 0, 3)^\intercal$. Wir führen die Probe durch:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 2 & 2 \end{bmatrix} \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

(*ii*) Wir betrachten das lineare Gleichungssystem

$$\left[\begin{array}{ccc|c} 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 2 \\ 3 & 4 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 2 \end{array}\right]$$

Tauschen der Zeilen 1 und 2

$$\begin{array}{c} \\ \\ 2 \end{array}\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right]$$

Erzeugen von Nullen unterhalb des Pivot-Elements

$$\begin{array}{c} \\ \\ 4 \end{array}\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Erzeugen von Nullen unterhalb des Pivot-Elements

An dieser Stelle ist eine Zeilenstufenform hergestellt. Wir erkennen daran $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = 2 = m - 1$, also ist das System lösbar, und die Lösungsmenge des homogenen Systems hat $\dim(\mathcal{L}(A, 0)) = 1$. Wir gehen weiter zur reduzierten Zeilenstufenform:

$$\begin{array}{c} \\ 3 \end{array}\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Normierung des Pivot-Elements

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Erzeugen von Nullen oberhalb des Pivot-Elements

Wir haben in diesem Fall also die abhängigen Indizes $\mathcal{A} = (1, 3)$ und einen einzelnen unabhängigen Index $\mathcal{I} = (2)$.

Eine partikuläre Lösung erhalten wir, indem wir die unabhängige Variable $x_2 \coloneqq 0$ setzen und die abhängigen Variablen mit Hilfe des Systems berechnen. Da wir die Pivot-Elemente auf 1 normiert haben, müssen wir dazu lediglich die rechte Seite ablesen und die Elemente an der richtigen Stelle (wie durch die Pivot-Spalten vorgegeben) in den Lösungsvektor eintragen. Wir erhalten so die partikuläre Lösung $x_0 = (2, 0, 3)^\intercal$. Wir führen die Probe durch:

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

Eine Basis der Lösungsmenge des homogenen Systems

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

erhalten wir nach Bemerkung 16.4 dadurch, dass wir für die unabhängigen Variablen (hier nur x_2) einen der Standardbasisvektoren von K^{m-r} einsetzen (hier also nur die Zahl $x_2 \coloneqq 1$) und die abhängigen Variablen von hinten nach vorne mit Hilfe der Gleichungen ausrechnen. Im vorliegenden Beispiel lesen wir aus der zweiten Gleichung

$$x_3 = 0$$

ab und anschließend aus der ersten Gleichung

$$x_1 + 3\,x_2 = 0 \quad\Leftrightarrow\quad x_1 = -3\,x_2 = 2\,x_2 = 2 \cdot 1 = 2.$$

Die Lösungsmenge des homogenen Systems besteht also gerade aus allen Vielfachen des Vektors $(2, 1, 0)^\intercal$. Wir führen auch hier die Probe durch:

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

Da wir es in unserem Beispiel mit dem endlichen Körper $\mathbb{Z}_5$ zu tun haben, können wir die Elemente des eindimensionalen Lösungsraumes sogar aufzählen. Es gilt

$$\begin{aligned} \mathcal{L}(A, b) &= \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} + \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 3 \end{pmatrix} \right\}. \end{aligned}$$

(*iii*) Wir betrachten das lineare Gleichungssystem

$$\left[\begin{array}{ccc|c} 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 2 \\ 3 & 4 & 2 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 3 \end{array}\right] \quad \text{Tauschen der Zeilen 1 und 2}$$

$$\begin{array}{c} \\ \\ 2 \end{array} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements} \end{array}$$

$$\begin{array}{c} \\ \\ 4 \end{array} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements} \end{array}$$

An dieser Stelle ist eine Zeilenstufenform hergestellt. Wir erkennen nun $\operatorname{Rang}(A) = 2$, aber $\operatorname{Rang}([A, b]) = 3$. Das heißt, das System ist nicht lösbar, vgl. Satz 16.3. Die reduzierte Zeilenstufenform muss daher nicht bestimmt werden. △

Abschließend bemerken wir, dass wir mit Hilfe der Transformation in reduzierte Zeilenstufenform nicht nur ein einzelnes lineares Gleichungssystem $A\,x = b$ lösen können, sondern mehrere Systeme gleichzeitig, sofern diese sich nur in der rechten Seite unterscheiden. Haben wir $k \in \mathbb{N}_0$ rechte Seiten, so schreiben wir diese spaltenweise als $n \times k$-Matrix B. Das System für die k unbekannten Spaltenvektoren in der $n \times k$-Matrix X lautet dann $A\,X = B$. Die Transformation auf Zeilenstufenform geschieht einfach für alle Spalten der erweiterten Koeffizientenmatrix $[A, B]$ gleichzeitig. Wie gewohnt können wir an der Zeilenstufenform ablesen, für welche rechten Seiten das System lösbar ist. Die rechten Seiten, die zu unlösbaren Systemen führen, können wir vor der Herstellung der reduzierten Zeilenstufenform einfach herausstreichen (oder auch stehenlassen). Wir erhalten dann für jede rechte Seite, für die das System lösbar ist, eine partikuläre Lösung. Die Lösungsmenge des homogenen System ist für alle rechten Seite dieselbe, da das homogene System ja dasselbe ist.

Ein wichtiger Anwendungsfall ist die Bestimmung der inversen Matrix einer quadratischen $n \times n$-Matrix A. Dies kann man durch die rechte Seite $B = I_n$ und Lösen des Systems $A\,X = I_n$ erreichen. An der Zeilenstufenform erkennt man, ob A invertierbar ist. Wenn ja, hat man am Ende rechts die inverse Matrix stehen.

Beispiel 16.11 (Berechnung der inversen Matrix, vgl. Beispiel 16.10)**.**
Wir führen die Berechnung der inversen Matrix für die erste Matrix aus Beispiel 16.10 vor. Wir erinnern daran, dass wir es hier mit Matrizen über dem Körper $\mathbb{Z}_5$ zu tun haben.

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 \\ 3 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \quad \text{Tauschen der Zeilen 1 und 2}$$

$$2 \; \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{array}\right] \quad \begin{array}{l}\text{Erzeugen von Nullen}\\ \text{unterhalb des Pivot-Elements}\end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \quad \text{Tauschen der Zeilen 2 und 3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \quad \begin{array}{l}\text{Erzeugen von Nullen}\\ \text{unterhalb des Pivot-Elements}\end{array}$$

An dieser Stelle ist eine Zeilenstufenform hergestellt. Wir erkennen $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = 3 = m$ für jede Spalte b der rechten Seite, also ist das System für jede der rechten Seiten eindeutig lösbar. Wir gehen weiter zur reduzierten Zeilenstufenform:

$$3 \; \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \quad \text{Normierung des Pivot-Elements}$$

$$\begin{array}{c} \\ 3 \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 4 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right]$$ Erzeugen von Nullen oberhalb des Pivot-Elements

$$\begin{array}{c} \\ 2 \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 4 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right]$$ Normierung des Pivot-Elements

$$\begin{array}{c} 2 \\ \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 4 & 4 \\ 0 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right]$$ Erzeugen von Nullen oberhalb des Pivot-Elements

Die Matrix auf der rechten Seite ist die Inverse der Ausgangsmatrix.

Wir führen die Probe durch:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 4 & 4 \\ 3 & 4 & 2 \\ 3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \triangle$$

Ende der Vorlesung 22

§ 17 Homomorphismen von Vektorräumen

Literatur: Fischer, Springborn, 2020, Kapitel 3.1–3.2; Bosch, 2014, Kapitel 2; Beutelspacher, 2014, Kapitel 5.1 und 5.3; Deiser, 2024b, Kapitel 3.3; Jänich, 2008, Kapitel 4

In diesem Abschnitt geht es um die Homomorphismen, also die strukturverträglichen Abbildungen zwischen Vektorräumen über demselben Körper.

Definition 17.1 (Vektorraumhomomorphismus, vgl. Definition 10.11 eines Körperhomomorphismus)**.**
Es seien $(V, +_1, \cdot_1)$, $(W, +_2, \cdot_2)$ Vektorräume über demselben Körper K.

(i) Eine Abbildung $f\colon V \to W$ heißt **strukturverträglich** oder ein **Homomorphismus von Vektorräumen** (englisch: vector space homomorphism) oder eine **lineare Abbildung** (englisch: linear map) von $(V, +_1, \cdot_1)$ in $(W, +_2, \cdot_2)$, wenn gilt:

$$f(u +_1 v) = f(u) +_2 f(v) \quad \text{für alle } u, v \in V, \tag{17.1a}$$
$$f(\alpha \cdot_1 u) = \alpha \cdot_2 f(u) \quad \text{für alle } u \in V \text{ und alle } \alpha \in K. \tag{17.1b}$$

Man bezeichnet die Eigenschaft (17.1a) auch als die **Additivität** (englisch: additivity) und die Eigenschaft (17.1b) als die **Homogenität** (englisch: homogeneity) der Abbildung f.

(ii) Ist $f\colon V \to W$ strukturverträglich und gilt $(V, +_1, \cdot_1) = (W, +_2, \cdot_2)$, so sprechen wir auch von einem **Endomorphismus eines Vektorraumes** (englisch: vector space endomorphism) oder einem **linearen Endomorphismus** (englisch: linear endomorphism)..

(*iii*) Ist $f: V \to W$ strukturverträglich und bijektiv, so heißt f auch **strukturerhaltend** oder ein **Isomorphismus von Vektorräumen** (englisch: vector space isomorphism) oder ein **linearer Isomorphismus** (englisch: linear isomorphism). In diesem Fall nennen wir $(V, +_1, \cdot_1)$ und $(W, +_2, \cdot_2)$ auch zueinander **isomorphe Vektorräume** (englisch: isomorphic vector spaces) und schreiben

$$(V, +_1, \cdot_1) \cong (W, +_2, \cdot_2).$$

(*iv*) Ist $f: V \to W$ strukturverträglich und bijektiv und gilt $(V, +_1, \cdot_1) = (W, +_2, \cdot_2)$, so sprechen wir auch von einem **Automorphismus eines Vektorraumes** (englisch: vector space automorphism) oder einem **linearen Automorphismus** (englisch: linear automorphism). △

Bemerkung 17.2 (zu Definition 17.1)**.**

(*i*) Die Vektorräume $(V, +_1, \cdot_1)$ und $(W, +_2, \cdot_2)$ sind insbesondere (abelsche) Gruppen $(V, +_1)$ und $(W, +_2)$. Aufgrund der Bedingung (17.1a) ist jeder Vektorraumhomomorphismus insbesondere ein Gruppenhomomorphismus. Wir können daher Ergebnisse aus § 8 verwenden.

(*ii*) Im Folgenden werden wir zulassen, die Vektorraumoperationen $+$ und $\cdot$ in beiden Vektorräumen mit demselben Symbol (und demselben Symbol wie im zugrundeliegenden Körper) zu notieren. △

Satz 17.3 (Komposition von Vektorraumhomomorphismen, Inverse von Vektorraumisomorphismen, vgl. Satz 10.12 zu Körperhomomorphismen)**.**
Es seien U, V, W Vektorräume über demselben Körper K.

(*i*) Sind $f: U \to V$ und $g: V \to W$ lineare Abbildungen, dann ist auch $g \circ f: U \to W$ eine lineare Abbildung.

(*ii*) Ist $f: U \to V$ eine bijektive lineare Abbildung, dann ist auch $f^{-1}: V \to U$ eine bijektive lineare Abbildung.

Beweis. Der Beweis ist Gegenstand der Übung. □

Folgerung 17.4 (Isomorphie von Vektorräumen ist eine Äquivalenzrelation, vgl. Folgerung 10.13 zur Isomorphie von Körpern)**.**
Isomorphie ist eine Äquivalenzrelation auf der Klasse aller Vektorräume über demselben Körper K.

Beispiel 17.5 (Vektorraumhomomorphismus)**.**

(*i*) Die Vektorräume K_n und K^n über einem Körper K sind zueinander isomorph. Die Abbildung

$$\cdot^\intercal : K_n \ni x \mapsto x^\intercal \in K^n$$

ist ein linearer Isomorphismus, vgl. (15.25a) und (15.25b). (**Quizfrage 17.1:** Wie sieht der inverse Vektorraumisomorphismus aus? Gibt es weitere mögliche Isomorphismen zwischen K_n und K^n?)

Allgemeiner ist die Transposition ein Isomorphismus zwischen $K^{n \times m}$ und $K^{m \times n}$ für beliebige $n, m \in \mathbb{N}_0$.

(ii) Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei die Familie $B = (v_i)_{i \in I}$ eine Basis von V. Dann ist V isomorph zu $(K^I)_{00}$, dem Unterraum der endlich getragenen Funktionen $I \longrightarrow K$ (Satz 13.3).

(iii) Der Vektorraum der $n \times m$-Matrizen $K^{n \times m}$ über einem Körper K ist isomorph zum Vektorraum $K^{n\,m}$. Die **Vektorisierung** (englisch: vectorization)

$$\operatorname{vec}\colon K^{n \times m} \ni A \mapsto \operatorname{vec}(A) \in K^{n\,m},$$

definiert durch „Übereinanderstapeln" der Spalten, also

$$\operatorname{vec}\left(\begin{bmatrix} | & & | \\ a_{\bullet 1} & \cdots & a_{\bullet m} \\ | & & | \end{bmatrix}\right) \coloneqq \begin{pmatrix} a_{\bullet 1} \\ \vdots \\ a_{\bullet m} \end{pmatrix},$$

ist ein linearer Isomorphismus. (**Quizfrage 17.2:** Wie sieht der inverse Vektorraumisomorphismus aus?)

(iv) Im Standardvektorraum K^n über einem Körper K ist die **Projektion auf die i-te Koordinate** (englisch: projection onto the i-th coordinate), gegeben durch

$$\pi_i\colon K^n \ni x \mapsto x_i \in K \tag{17.2}$$

für $1 \leqslant i \leqslant n \in \mathbb{N}$, eine surjektive lineare Abbildung. (**Quizfrage 17.3:** Ist sie auch injektiv?)

(v) Es seien K ein Körper und $n, m \in \mathbb{N}_0$. Für eine Matrix $A \in K^{n \times m}$ ist die Matrix-Vektor-Multiplikation mit A

$$f_A\colon K^m \ni x \mapsto f_A(x) \coloneqq A\,x \in K^n \tag{17.3}$$

eine lineare Abbildung. Sie wird die **von A induzierte lineare Abbildung** genannt (englisch: linear map induced by A). Ihre Eigenschaften untersuchen wir in § 17.2.

Wie wir in § 19 sehen werden, ist die Matrix-Vektor-Multiplikation der Prototyp einer linearen Abbildung zwischen endlich-dimensionalen Vektorräumen. Jede lineare Abbildung kann in Form von Matrix-Vektor-Multiplikation geschrieben werden. △

Expertenwissen: weitere lineare Abbildungen

Mit Kenntnissen aus Lehrveranstaltungen zur *Analysis* können wir zum Beispiel die Linearität der folgenden Abbildungen bestätigen:

(i) Die **Grenzwert-Abbildung**, definiert beispielsweise auf dem Vektorraum der konvergenten $\mathbb{R}$-wertigen Folgen $(\mathbb{R}^{\mathbb{N}})_c$ nach $\mathbb{R}$, also

$$\lim\colon (\mathbb{R}^{\mathbb{N}})_c \ni (y_n)_{n \in \mathbb{N}_0} \mapsto \lim_{n \to \infty} (y_n) \in \mathbb{R}$$

ist eine surjektive lineare Abbildung. (**Quizfrage 17.4:** Ist sie auch injektiv?)

(*ii*) Die **Ableitungsabbildung** (englisch: differentiation map), definiert beispielsweise auf dem Vektorraum der differenzierbaren Funktionen $(a, b) \to \mathbb{R}$ in den Vektorraum der Funktionen $(a, b) \to \mathbb{R}$, also

$$\cdot' \colon \big\{f \in \mathbb{R}^{(a,b)} \,\big|\, f \text{ ist differenzierbar}\big\} \ni f \mapsto f' \in \mathbb{R}^{(a,b)}$$

ist eine lineare Abbildung, die nicht surjektiv und nicht injektiv ist. (**Quizfrage 17.5:** Warum?)

Lemma 17.6 (Eigenschaften linearer Abbildungen)**.**

Es seien V, W Vektorräume über demselben Körper K. Weiter sei $f \colon V \to W$ eine lineare Abbildung. Dann gilt:

(*i*) $f(0) = 0$.

(*ii*) $f(-v) = -f(v)$ für alle $v \in V$.

(*iii*) $f\left(\sum_{i=1}^n \alpha_i\, v_i\right) = \sum_{i=1}^n \alpha_i\, f(v_i)$ für alle $n \in \mathbb{N}_0$, $\alpha_i \in K$ und $v_i \in V$.
(Das Bild einer Linearkombination von Vektoren ist also die Linearkombination der Bilder dieser Vektoren.)

(*iv*) Ist $E \subseteq V$, dann gilt $f(\langle E \rangle) = \langle f(E) \rangle$.
(Das Bild der linearen Hülle einer Menge ist die lineare Hülle des Bildes der Menge.)

(*v*) Ist $F = (v_i)_{i \in I}$ eine Familie von Vektoren in V, dann gilt $f(\langle F \rangle) = \langle f(F) \rangle$.
(Das Bild der linearen Hülle einer Familie ist die lineare Hülle des Bildes der Familie.)

(*vi*) Ist $U \subseteq V$ ein Unterraum, dann ist $f(U) \subseteq W$ ein Unterraum.
(Das Bild eines Unterraumes ist wieder ein Unterraum.)

(*vii*) Ist $Z \subseteq W$ ein Unterraum, dann ist $f^{-1}(Z) \subseteq V$ ein Unterraum.
(Das Urbild eines Unterraumes ist wieder ein Unterraum.)

(*viii*) Ist $M \subseteq V$ eine linear abhängige Menge von Vektoren, dann ist auch $f(M) \subseteq W$ eine linear abhängige Menge von Vektoren.
(Lineare Abhängigkeit kann durch eine lineare Abbildung nicht geheilt werden.)

(*ix*) Ist $F = (v_i)_{i \in I}$ eine linear abhängige Familie von Vektoren in V, dann ist auch $(f(v_i))_{i \in I}$ eine linear abhängige Familie von Vektoren.
(Lineare Abhängigkeit kann durch eine lineare Abbildung nicht geheilt werden.)

Beweis. Aussage (*i*) und Aussage (*ii*) folgen sofort aus Lemma 8.8.

Aussage (*iii*):

$$\begin{aligned} f\Big(\sum_{i=1}^n \alpha_i\, v_i\Big) &= \sum_{i=1}^n f(\alpha_i\, v_i) && \text{durch wiederholte Anwendung von (17.1a)} \\ &= \sum_{i=1}^n \alpha_i\, f(v_i) && \text{durch Anwendung von (17.1b) auf jeden Summanden} \end{aligned}$$

Aussage (iv): Nach Satz 11.16 besteht $\langle E\rangle$ gerade aus den Linearkombinationen von E, während $\big\langle f(E)\big\rangle$ aus den Linearkombinationen von $f(E)$ besteht. Nach Aussage (iii) sind das aber dieselben Mengen.

Aussage (v): Nach Satz 11.16 besteht $\langle F\rangle$ gerade aus den Linearkombinationen von F, während $\big\langle f(F)\big\rangle$ aus den Linearkombinationen von $(f(v_i))_{i\in I}$ besteht. Nach Aussage (iii) sind das aber dieselben Mengen.

Aussage (vi): Wir verwenden das Unterraumkriterium (Satz 11.11). Wegen $0 \in U$ und Aussage (i) ist $0 \in f(U)$, also ist $f(U) \neq \emptyset$. Sind weiter $w_1, w_2 \in f(U)$, dann gibt es $u_1, u_2 \in U$ mit $w_1 = f(u_1)$ und $w_2 = f(u_2)$. Für $\alpha_1, \alpha_2 \in K$ gilt

$$\alpha_1\, w_1 + \alpha_2\, w_2 = \alpha_1\, f(u_1) + \alpha_2\, f(u_2) = f(\alpha_1\, u_1 + \alpha_2\, u_2).$$

Wegen der Unterraumeigenschaft gehört $\alpha_1\, u_1 + \alpha_2\, u_2$ zu U, also gehört $\alpha_1\, w_1 + \alpha_2\, w_2$ zu $f(U)$.

Aussage (vii): Wir verwenden nochmal das Unterraumkriterium (Satz 11.11). Wegen $0 \in Z$ und Aussage (i) ist $0 \in f^{-1}(Z)$, also ist $f^{-1}(Z) \neq \emptyset$. Sind weiter $u_1, u_2 \in f^{-1}(Z)$, dann gibt es $w_1, w_2 \in Z$ mit $w_1 = f(u_1)$ und $w_2 = f(u_2)$. Für $\alpha_1, \alpha_2 \in K$ gilt

$$\alpha_1\, w_1 + \alpha_2\, w_2 = \alpha_1\, f(u_1) + \alpha_2\, f(u_2) = f(\alpha_1\, u_1 + \alpha_2\, u_2).$$

Wegen der Unterraumeigenschaft gehört $\alpha_1\, w_1 + \alpha_2\, w_2$ zu Z, also gehört $\alpha_1\, u_1 + \alpha_2\, u_2$ zu $f^{-1}(Z)$.

Aussage $(viii)$: Es sei $M \subseteq V$ eine linear abhängige Menge, d. h., es gibt ein $n \in \mathbb{N}_0$ und paarweise verschiedene Vektoren $v_1, \ldots, v_n \in M$ sowie Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$, die nicht alle gleich 0 sind, sodass gilt: $\sum_{i=1}^n \alpha_i\, v_i = 0$. Es folgt mit Aussage (iii) und Aussage (i)

$$\sum_{i=1}^n \alpha_i\, f(v_i) = f\Big(\sum_{i=1}^n \alpha_i\, v_i\Big) = f(0) = 0.$$

Das bedeutet aber gerade die lineare Abhängigkeit der Menge $f(M)$.

Aussage (ix): Es sei $F = (v_i)_{i\in I}$ eine linear abhängige Familie von Vektoren in V, d. h., es gibt ein $n \in \mathbb{N}_0$ und paarweise verschiedene Indizes $i_1, \ldots, i_n \in I$ sowie Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$, die nicht alle gleich 0 sind, sodass gilt: $\sum_{\ell=1}^n \alpha_\ell\, v_{i_\ell} = 0$. Es folgt mit Aussage (iii) und Aussage (i)

$$\sum_{\ell=1}^n \alpha_\ell\, f(v_{i_\ell}) = f\Big(\sum_{\ell=1}^n \alpha_\ell\, v_{i_\ell}\Big) = f(0) = 0.$$

Das bedeutet aber gerade die lineare Abhängigkeit der Familie $(f(v_i))_{i\in I}$. □

Definition 17.7 (Bild und Kern einer linearen Abbildung, vgl. Definition 10.17 von Bild und Kern eines Körperhomomorphismus)**.**
Es seien V, W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ eine lineare Abbildung.

(i) Das **Bild** (englisch: image) von f ist definiert als

$$\operatorname{Bild}(f) \coloneqq \{f(u) \in W \mid u \in V\} = f(V). \tag{17.4}$$

(*ii*) Der **Kern** (englisch: kernel, null space[28]) von f ist definiert als

$$\operatorname{Kern}(f) \coloneqq \{u \in V \mid f(u) = 0_W\} = f^{-1}(\{0_W\}) \tag{17.5}$$

△

Lemma 17.8 (Kern und Bild linearer Abbildungen)**.**
Es seien V, W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ eine lineare Abbildung.

(*i*) $\operatorname{Bild}(f)$ ist ein Unterraum von W.

(*ii*) $\operatorname{Kern}(f)$ ist ein Unterraum von V.

Beweis. Aussage (*i*): V ist selbst ein Unterraum von V, daher ist $\operatorname{Bild}(f) = f(V)$ nach Lemma 17.6 (*vi*) ein Unterraum von W.

Aussage (*ii*): $\{0\}$ ist ein Unterraum von W, daher ist $\operatorname{Kern}(f) = f^{-1}(\{0\})$ nach Lemma 17.6 (*vii*) ein Unterraum von V. □

Lemma 17.9 (Charakterisierung der Injektivität linearer Abbildungen, vgl. Lemma 9.24 für Ringhomomorphismen)**.**
Es seien V, W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ eine lineare Abbildung. Dann sind äquivalent:

(*i*) f ist injektiv.

(*ii*) $\operatorname{Kern}(f) = \{0_V\}$.

(*iii*) Die einzige Lösung der Gleichung $f(v) = 0_W$ ist $v = 0_V$.

Beweis. Weder der Begriff der Injektivität noch die Definition von $\operatorname{Kern}(f) \coloneqq \{u \in V \mid f(u) = 0\}$ ändern sich, wenn wir die lineare Abbildung $f\colon (V, +, \cdot) \to (W, +, \cdot)$ als Gruppenhomomorphismus $f\colon (V, +) \to (W, +)$ auffassen. Das Resultat folgt daher aus Lemma 8.13 (Charakterisierung der Injektivität für Gruppenhomomorphismen). □

§ 17.1 Konstruktion linearer Abbildungen

Wir zeigen nun, dass eine lineare Abbildung $V \to W$ durch die Bilder auf einer Basis von V bereits eindeutig festgelegt ist. Das gilt gleichermaßen für Basismengen wie für Basisfamilien (Definition 13.1). In Vorbereitung auf die spätere Darstellung linearer Abbildungen durch Matrizen (§ 19) bietet es sich aber bereits jetzt an, nur noch mit Basisfamilien zu arbeiten. Der Einfachheit halber werden wir diese ab sofort auch einfach als Basis bezeichnen.

[28] nicht zu verwechseln mit dem Begriff *zero vector space* (Nullraum) $\{0\}$, siehe Beispiel 11.3

Satz 17.10 (Existenz- und Eindeutigkeitssatz für Vektorraumhomomorphismen$^{\text{AoC}}$, vgl. Satz 8.14 für Gruppenhomomorphismen)**.**

Es seien V, W Vektorräume über demselben Körper K. Weiter seien $(v_i)_{i \in I}$ eine Familie von Vektoren in V und $(w_i)_{i \in I}$ eine Familie von Vektoren in W mit gleicher Indexmenge I.

(i) Ist $B \coloneqq (v_i)_{i \in I}$ eine Basis von V, dann gibt es genau eine lineare Abbildung $f : V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i \in I$.

Diese Abbildung hat außerdem folgende Eigenschaften:

(a) $\operatorname{Bild}(f) = \langle (w_i)_{i \in I} \rangle$.
(Die Bilder der Basisvektoren von V erzeugen den Bildraum $\operatorname{Bild}(f)$.)

(b) f ist surjektiv genau dann, wenn $\langle (w_i)_{i \in I} \rangle = W$ gilt.

(c) f ist injektiv genau dann, wenn $(w_i)_{i \in I}$ linear unabhängig ist.

(d) f ist bijektiv genau dann, wenn $(w_i)_{i \in I}$ eine Basis von W ist.

(ii) Ist $(v_i)_{i \in I}$ linear unabhängig, dann gibt es eine lineare Abbildung $f : V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i \in I$.[29]

Beweis. Wir beginnen mit Aussage (i).

Schritt 1: Wir konstruieren eine Abbildung $f : V \to W$ mit den gesuchten Eigenschaften.

Da $B = (v_i)_{i \in I}$ eine Basis von V ist, kann jedes $v \in V$ nach Satz 13.3 auf (bis auf Nullkoeffizienten) eindeutige Art und Weise als Linearkombination $v = \sum_{i \in I_0} \alpha_i \, v_i$ geschrieben werden. Daher ergibt die die Setzung

$$f(v) \coloneqq \sum_{i \in I_0} \alpha_i \, w_i$$

eine wohldefinierte Funktion. (**Quizfrage 17.6:** Warum ist die Eindeutigkeit der Darstellung von v bis auf Nullkoeffizienten für die Wohldefiniertheit schon ausreichend?)

Wegen $v_i = 1 \, v_i$ gilt

$$f(v_i) = f(1 \, v_i) = 1 \, w_i = w_i$$

erfüllt f wie gefordert die Bedingung $f(v_i) = w_i$.

Schritt 2: Wir zeigen: Die so definierte Abbildung $f : V \to W$ ist linear.

Sind $u = \sum_{i \in I_0} \alpha_i \, v_i$ und $v = \sum_{i \in I_1} \beta \, v_i$ zwei beliebige Vektoren aus V, dann können wir durch Hinzufügen von Nullkoeffizienten erreichen, dass beide Darstellungen dieselben endlich vielen Indizes aus $I_{01} \coloneqq I_0 \cup I_1$ verwenden, also $u = \sum_{i \in I_{01}} \alpha_i \, v_i$ und

[29] Diese Aussage hängt im Fall, dass V unendlich-dimensional ist, über den Basisergänzungssatz 13.5 vom Auswahlaxiom ab.

$v = \sum_{i \in I_{01}} \beta\, v_i$. Es gilt

$$
\begin{aligned}
f(u+v) &= f\Big(\sum_{i \in I_{01}} \alpha_i\, v_i + \sum_{i \in I_{01}} \beta\, v_i\Big) && \text{nach Darstellung von } u \text{ und } v \\
&= f\Big(\sum_{i \in I_{01}} (\alpha_i + \beta_i)\, v_i\Big) && \text{nach Distributivgesetz (11.1c) und Kommutativität} \\
&= \sum_{i \in I_{01}} (\alpha_i + \beta_i)\, w_i && \text{nach Definition von } f \\
&= \sum_{i \in I_{01}} \alpha_i\, w_i + \sum_{i \in I_{01}} \beta_i\, w_i && \text{nach Distributivgesetz (11.1c)} \\
&= f(u) + f(v) && \text{nach Definition von } f
\end{aligned}
$$

und außerdem für $\alpha \in K$

$$
\begin{aligned}
f(\alpha\, u) &= f\Big(\alpha \sum_{i \in I_0} \alpha_i\, v_i\Big) && \text{nach Darstellung von } u \\
&= f\Big(\sum_{i \in I_0} (\alpha\, \alpha_i)\, v_i\Big) && \text{nach Distributivgesetz (11.1b)} \\
&= \sum_{i \in I_0} (\alpha\, \alpha_i)\, w_i && \text{nach Definition von } f \\
&= \alpha \sum_{i \in I_0} \alpha_i\, w_i && \text{nach Distributivgesetz (11.1b)} \\
&= \alpha\, f(u) && \text{nach Definition von } f.
\end{aligned}
$$

Schritt 3: Wir zeigen die Eindeutigkeit von f.

Dazu sei $g\colon V \longrightarrow W$ eine weitere lineare Abbildung mit der Eigenschaft $g(v_i) = w_i$. Ist $v \in V$ ein beliebiger Vektor mit der i. W. eindeutigen Darstellung $v = \sum_{i \in I_0} \alpha_i\, v_i$, dann gilt

$$
\begin{aligned}
g(v) &= g\Big(\sum_{i \in I_0} \alpha_i\, v_i\Big) && \text{wegen der Darstellung von } v \\
&= \sum_{i \in I_0} \alpha_i\, g(v_i) && \text{wegen der Linearität von } g \\
&= \sum_{i \in I_0} \alpha_i\, w_i && \text{wegen der Eigenschaft } g(v_i) = w_i \\
&= f(v) && \text{nach Definition von } f.
\end{aligned}
$$

Also muss g notwendig mit f übereinstimmen.

Wir kommen zu den weiteren Eigenschaften der Abbildung f.

Aussage (a): Es gilt

$$
\begin{aligned}
\operatorname{Bild}(f) &= f(V) && \text{nach Definition von } \operatorname{Bild}(f) \\
&= f\big(\langle (v_i)_{i \in I} \rangle\big) && \text{denn } (v_i)_{i \in I} \text{ ist eine Basis von } V \\
&= \big\langle (f(v_i))_{i \in I} \big\rangle && \text{nach Lemma 17.6 (v)} \\
&= \langle (w_i)_{i \in I} \rangle && \text{nach Definition von } f.
\end{aligned}
$$

Aussage (b): $f: V \to W$ ist nach Definition surjektiv genau dann, wenn $\operatorname{Bild}(f) = W$ ist, also nach Aussage (a) genau dann, wenn $\langle (w_i)_{i\in I}\rangle = W$ gilt.

Aussage (c): Es sei zunächst die Familie $(w_i)_{i\in I}$ linear abhängig, d. h., es gibt eine endliche Teilmenge $I_0 \subseteq I$ und Koeffizienten $\alpha_i \in K$ mit $i \in I_0$, die nicht alle gleich 0 sind, sodass gilt: $\sum_{i\in I_0} \alpha_i\, w_i = 0$. Dann ist $v \coloneqq \sum_{i\in I_0} \alpha_i\, v_i$ nicht der Nullvektor in V, da $(v_i)_{i\in I}$ eine Basis von V ist, jedoch gilt

$$f(v) = f\Big(\sum_{i\in I_0} \alpha_i\, v_i\Big) = \sum_{i\in I_0} \alpha_i\, w_i = 0.$$

Das bedeutet $f(v) = f(0) = 0$, also ist f nicht injektiv.

Nun sei umgekehrt $(w_i)_{i\in I}$ linear unabhängig und $v \in V$ ein Vektor mit $f(v) = 0$. Der Vektor v hat eine Darstellung $v = \sum_{i\in I_0} \alpha_i\, v_i$ mit einer endlichen Indexmenge $I_0 \subseteq I$, und es gilt

$$0 = f(v) = f\Big(\sum_{i\in I_0} \alpha_i\, v_i\Big) = \sum_{i\in I_0} \alpha_i\, w_i.$$

Aufgrund der linearen Unabhängigkeit der Familie $(w_i)_{i\in I}$ ist das nur möglich, wenn alle $\alpha_i = 0$ sind, also nur dann, wenn $v = 0$ ist. Mit anderen Worten, $\operatorname{Kern}(f) = \{0\}$, und nach Lemma 17.9 ist f injektiv.

Aussage (d) folgt sofort aus Aussage (b) und Aussage (c).

Nun kommen wir zur Aussage *(ii)*. Es sei also $(v_i)_{i\in I}$ linear unabhängig. Wir nutzen den Basisergänzungssatz 13.5 (der im Fall, dass V unendlich-dimensional ist, das Zornsche Lemma und damit das Auswahlaxiom verwendet[AoC]) und ergänzen die Menge zu einer Basis $(v_i)_{i\in \widehat{I}}$ von V. Wir wählen die fehlenden w_i für $\widehat{I} \setminus I$ als beliebige Vektoren in W (beispielsweise alle als den Nullvektor). Nach Aussage *(i)* gibt es dann eine (eindeutige) lineare Abbildung $f: V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i \in \widehat{I}$, insbesondere für alle $i \in I$. □

Beispiel 17.11 (Konstruktion linearer Abbildungen)**.**

(i) Es sei V ein Vektorraum über dem Körper K. Dann ist eine lineare Abbildung $f: K \to V$ durch einen einzigen Funktionswert, etwa $f(1) \in V$, festgelegt.

(ii) Es seien K ein Körper und $(e_1, \ldots, e_m)$ die Standardbasis im Vektorraum K^m, $m \in \mathbb{N}_0$. Eine lineare Abbildung $f: K^m \to K^n$ ist dadurch festgelegt, dass wir die Bilder $f(e_1), \ldots, f(e_m)$ der m Basisvektoren angeben, also m Elemente von K^n. Tragen wir diese Bilder spaltenweise in eine Matrix

$$A \coloneqq \begin{bmatrix} | & & | \\ f(e_1) & \cdots & f(e_m) \\ | & & | \end{bmatrix}$$

ein, so gilt

$$f(x) = f\Big(\sum_{j=1}^m x_j\, e_j\Big) = \sum_{j=1}^m x_j\, f(e_j) = A\,x.$$

Jede lineare Abbildung $f: K^m \to K^n$ kann also durch Matrix-Vektor-Produkte $x \mapsto A\,x$ realisiert werden, wobei A spaltenweise die Bilder $f(e_j)$ der Standardbasisvektoren enthält.

(*iii*) Im Vektorraum $\mathbb{R}^2$ mit der Standardbasis $(e_1, e_2) = \left(\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right)$ legt

$$f(e_1) = \begin{pmatrix}\frac{1}{2}\sqrt{3}\\ \frac{1}{2}\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}-\frac{1}{2}\\ \frac{1}{2}\sqrt{3}\end{pmatrix}$$

eine lineare Abbildung (einen Endomorphismus) fest, und zwar eine Drehung um den Winkel 30° im mathematisch positiven Sinn (gegen den Uhrzeigersinn) um den Ursprung. Allgemeiner beschreibt

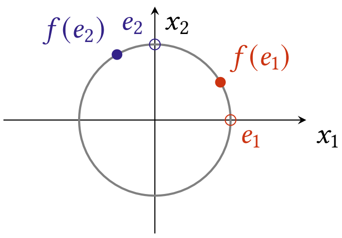
$$f(e_1) = \begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}-\sin(\alpha)\\ \cos(\alpha)\end{pmatrix}$$

eine Drehung um den Winkel α. Da $(f(e_1), f(e_2))$ eine Basis von $\mathbb{R}^2$ bildet, ist die Drehabbildung nach Satz 17.10 (*ii*) bijektiv, also ein Isomorphismus $\mathbb{R}^2 \to \mathbb{R}^2$.

Die Drehabbildung kann also durch Matrix-Vektor-Produkte mit der Matrix

$$\begin{bmatrix}\cos(\alpha) & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)\end{bmatrix} \tag{17.6}$$

realisiert werden.



(*iv*) Im Vektorraum $\mathbb{R}^2$ mit der Standardbasis $(e_1, e_2) = \left(\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right)$ legt

$$f(e_1) = \begin{pmatrix}1\\ 0\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}0\\ -1\end{pmatrix}$$

eine lineare Abbildung (einen Endomorphismus) fest, und zwar eine Spiegelung an der x_1-Achse. Allgemeiner beschreibt

$$f(e_1) = \begin{pmatrix}\cos^2(\alpha) - \sin^2(\alpha)\\ 2\cos(\alpha)\,\sin(\alpha)\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}2\cos(\alpha)\,\sin(\alpha)\\ \sin^2(\alpha) - \cos^2(\alpha)\end{pmatrix}$$

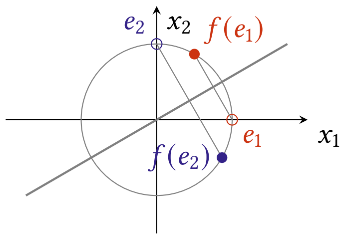
eine Spiegelung an derjenigen Achse durch den Ursprung, die den Winkel α gegen die x_1-Achse bildet. Da $(f(e_1), f(e_2))$ eine Basis von $\mathbb{R}^2$ bildet, ist die Spiegelungsabbildung nach Satz 17.10 bijektiv, also ein Isomorphismus $\mathbb{R}^2 \to \mathbb{R}^2$.

Die Spiegelungsabbildung kann also durch Matrix-Vektor-Produkte mit der Matrix

$$\begin{bmatrix}\cos^2(\alpha) - \sin^2(\alpha) & 2\cos(\alpha)\,\sin(\alpha)\\ 2\cos(\alpha)\,\sin(\alpha) & \sin^2(\alpha) - \cos^2(\alpha)\end{bmatrix} = \begin{bmatrix}\cos(2\alpha) & \sin(2\alpha)\\ \sin(2\alpha) & -\cos(2\alpha)\end{bmatrix} \tag{17.7}$$

realisiert werden.[30] Das bedeutet, dass beide Basisvektoren e_1 und e_2 um den Winkel 2α rotiert werden, während der Basisvektor e_2 anschließend noch am Ursprung punktgespiegelt wird.

[30] Die Gleichheit in (17.7) folgt aus den Additionstheoremen (8.5).



(*v*) Im Vektorraum $\mathbb{R}^2$ mit der Standardbasis $(e_1, e_2) = \left(\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right)$ legt

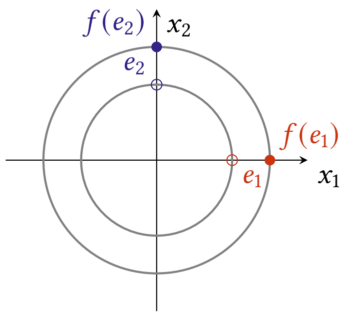
$$f(e_1) = \begin{pmatrix}\beta\\0\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}0\\\beta\end{pmatrix}$$

mit $\beta \in \mathbb{R}$ eine lineare Abbildung (einen Endomorphismus) fest, und zwar eine Streckung (im Fall $\beta > 0$) bzw. (im Fall $\beta < 0$) eine Streckung und Punktspiegelung am Ursprung. Da $(f(e_1), f(e_2))$ für $\beta \neq 0$ eine Basis von $\mathbb{R}^2$ bildet, ist die Streckungsabbildung nach Satz 17.10 (*ii*) in diesem Fall bijektiv, also ein Isomorphismus $\mathbb{R}^2 \to \mathbb{R}^2$.

Die Streckungsabbildung kann also durch Matrix-Vektor-Produkte mit der Matrix

$$\begin{bmatrix}\beta & 0\\0 & \beta\end{bmatrix} \tag{17.8}$$

realisiert werden.



(*vi*) Im Vektorraum $(K^{\mathbb{N}})_{00}$ der endlich getragenen Folgen über einem Körper K mit der Standardbasis $(e_j)_{j \in \mathbb{N}}$, siehe Beispiel 13.2, legt

$$f(e_j) = e_{j+1} \quad \text{für } j \in \mathbb{N}$$

eine lineare Abbildung (einen Endomorphismus) fest, eine sogenannte **Shift-Abbildung** (englisch: shift map). △

Ende der Vorlesung 23

Ende der Woche 11

§ 17.2 Die Matrix-Vektor-Multiplikation als lineare Abbildung

Die Matrix-Vektor-Multiplikation als Prototyp einer linearen Abbildung (Beispiele 17.5 und 17.11) ist von so zentraler Bedeutung, dass wir hier ihre Eigenschaften zusammenstellen. Kurz gesagt werden wir Folgendes sehen: Matrix-Vektor-Produkte sind die einzigen linearen Abbildungen $K^n \to K^m$, die es gibt; die Komposition linearer Abbildungen entspricht dem Produkt von Matrizen; und die inverse Abbildung entspricht der inversen Matrix.

Lemma 17.12 (Matrix-Vektor-Multiplikation als lineare Abbildung)**.**
Es seien K ein Körper und $n, m, k \in \mathbb{N}_0$.

(i) Ist $A \in K^{n \times m}$, dann ist

$$f_A \colon K^m \ni x \mapsto f_A(x) \coloneqq A\,x \in K^n \tag{17.9}$$

eine lineare Abbildung $K^m \to K^n$, genannt die **von A induzierte lineare Abbildung**.

(ii) Ist umgekehrt $f \colon K^m \to K^n$ eine lineare Abbildung, dann gibt es eine Matrix $A \in K^{n \times m}$, sodass $f = f_A$ gilt.

(iii) Sind $A \in K^{n \times m}$ und $B \in K^{m \times k}$, dann gilt

$$f_A \circ f_B = f_{A\,B}. \tag{17.10}$$

(„Die Komposition zweier linearer Abbildungen, die durch die Matrizen A und B induziert werden, entspricht derjenigen linearen Abbildung, die durch das Produkt der Matrizen A und B induziert wird.“)

(iv) $A \in K^{n \times n}$ ist genau dann invertierbar, wenn $f_A \colon K^n \to K^n$ invertierbar ist. In diesem Fall gilt

$$(f_A)^{-1} = f_{A^{-1}}. \tag{17.11}$$

(„Die Inverse einer linearen Abbildungen, die durch die Matrix A induziert ist, entspricht derjenigen linearen Abbildung, die durch die inverse Matrix A^{-1} induziert wird.“)

Beweis. Der Beweis ist Gegenstand der Übung. □

§ 17.3 Der Vektorraum der Vektorraumhomomorphismen

Definition 17.13 (Menge der Homomorphismen, Endomorphismen, Isomorphismen, Automorphismen von Vektorräumen)**.**
Es seien $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über demselben Körper K.

(i) Wir führen folgende Bezeichnungen ein:

$$\operatorname{Homo}(V, W) \coloneqq \{f \colon V \to W \mid f \text{ ist Vektorraumhomomorphismus}\} \tag{17.12a}$$
$$\operatorname{Endo}(V) \coloneqq \operatorname{Homo}(V, V) = \{f \colon V \to V \mid f \text{ ist Vektorraumendomorphismus}\} \tag{17.12b}$$
$$\operatorname{Iso}(V, W) \coloneqq \{f \colon V \to W \mid f \text{ ist Vektorraumisomorphismus}\} \tag{17.12c}$$
$$\operatorname{Auto}(V) \coloneqq \operatorname{Iso}(V, V) = \{f \colon V \to V \mid f \text{ ist Vektorraumautomorphismus}\}. \tag{17.12d}$$

(ii) Wollen wir den Skalarkörper betonen, so schreiben wir jeweils auch $\operatorname{Homo}_K(V, W)$, $\operatorname{Endo}_K(V)$, $\operatorname{Iso}_K(V, W)$ bzw. $\operatorname{Auto}_K(V)$. △

Satz 17.14 (die Homomorphismen zwischen Vektorräumen bilden einen Vektorraum)**.**
Es seien $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über demselben Körper K. Die Menge $(\operatorname{Homo}(V, W), +, \cdot)$ mit der punktweisen Addition $+$ und der punktweisen S-Multiplikation $\cdot$

$$
\begin{aligned}
+ \colon \operatorname{Homo}(V, W) \times \operatorname{Homo}(V, W) &\to \operatorname{Homo}(V, W) && \text{mit } f + g \text{ def. durch } (f + g)(u) \coloneqq f(u) + g(u) \\
\cdot \colon K \times \operatorname{Homo}(V, W) &\to \operatorname{Homo}(V, W) && \text{mit } \alpha \cdot f \text{ def. durch } (\alpha \cdot f)(u) \coloneqq \alpha\, f(u)
\end{aligned}
$$

bildet einen Vektorraum über K. Der Nullvektor in $(\operatorname{Homo}(V, W), +, \cdot)$ ist die **Nullabbildung** (englisch: zero map) $0 \colon V \to W$, die jeden Vektor in V auf den Nullvektor $0 \in W$ abbildet.

Beweis. Wir prüfen die Bedingungen aus der Definition 11.1 nach. Wir wissen bereits, dass die Menge der Funktionen, die auf irgendeiner Menge definiert sind und die Werte in einer abelschen Gruppe haben, selbst eine abelsche Gruppe mit der punktweisen Gruppenoperation bildet (Beispiel 7.22). Insbesondere ist also $(\operatorname{Homo}(V, W), +)$ eine abelsche Gruppe, da $(W, +)$ eine abelsche Gruppe ist.

Weiter gelten die Distributivgesetze

$$
\begin{aligned}
\alpha\,(f + g) &= \alpha\, f + \alpha\, g \\
\text{und} \quad (\alpha + \beta)\, f &= \alpha\, f + \beta\, f
\end{aligned}
$$

für alle $\alpha, \beta \in K$ und $f, g \in \operatorname{Homo}(V, W)$, denn wir haben

$$
\begin{aligned}
\big[\alpha\,(f + g)\big](u) &= \alpha\,\big[(f + g)(u)\big] = \alpha\,\big[f(u) + g(u)\big] = \alpha\, f(u) + \alpha\, g(u) \\
&= (\alpha\, f)(u) + (\alpha\, g)(u) = \big[\alpha\, f + \alpha\, g\big](u) \\
\text{und} \quad \big[(\alpha + \beta)\, f\big](u) &= (\alpha + \beta)\, f(u) = \alpha\, f(u) + \beta\, f(u) \\
&= \big[\alpha\, f + \beta\, f\big](u)
\end{aligned}
$$

für alle $u \in V$.

Das Assoziativgesetz

$$
(\alpha\,\beta)\, f = \alpha\,(\beta\, f)
$$

gilt, denn wir haben

$$
\big[(\alpha\,\beta)\, f\big](u) = (\alpha\,\beta)\, f(u) = \alpha\,(\beta\, f(u)) = \alpha\,([\beta\, f](u)) = [\alpha\,(\beta\, f)](u)
$$

für alle $u \in V$.

Schließlich gilt $(1\, f)(u) = 1\, f(u) = f(u)$ für alle $u \in V$, d. h., $1 \in K$ ist auch neutrales Element der S-Multiplikation. □

Als Spezialfall von Satz 17.14 ergibt sich, dass insbesondere die Endomorphismen $(\operatorname{Endo}(V), +, \cdot)$ einen Vektorraum bilden. Betrachten wir anstelle der S-Multiplikation als zweite Verknüpfung die Komposition $\circ$, so erhalten wir einen Ring wie schon beim Endomorphismenring einer abelschen Gruppe (vgl. Beispiel 9.2):

Satz 17.15 (die Endomorphismen eines Vektorraumes bilden einen Ring mit Eins, vgl. Beispiel 9.2)**.**

Es sei $(V, +, \cdot)$ ein Vektorraum. $(\operatorname{Endo}(V), +, \circ)$ mit der punktweisen Addition $+$ und der Komposition $\circ$

$$\begin{aligned} +\colon \operatorname{Endo}(V) \times \operatorname{Endo}(V) &\to \operatorname{Endo}(V) \quad \text{mit } f+g \text{ definiert durch} \quad (f+g)(u) \coloneqq f(u)+g(u), \\ \circ\colon \operatorname{Endo}(V) \times \operatorname{Endo}(V) &\to \operatorname{Endo}(V) \quad \text{mit } f\circ g \text{ definiert durch} \quad (f\circ g)(u) \coloneqq f(g(u)), \end{aligned}$$

bildet einen Ring mit dem Einselement id_V, genannt der **Endomorphismenring** (englisch: ring of endomorphisms) des Vektorraumes $(V, +, \cdot)$. Er ist i. A. nicht kommutativ.

Beweis. Der Beweis ist derselbe wie für den Endomorphismenring der abelschen Gruppe $(V, +)$ aus Beispiel 9.2, siehe Übung. Wir wissen bereits aus Satz 17.14 und Satz 17.3, dass Summen und Kompositionen von Vektorraumendomorphismen wieder Vektorraumendomorphismen sind. □

Definition 17.16 (allgemeine lineare Gruppe)**.**
Die Menge der bijektiven Endomorphismen $\operatorname{Auto}(V)$ eines K-Vektorraumes V, also die Einheitengruppe des Monoids $(\operatorname{Endo}(V), \circ)$, heißt die **Automorphismengruppe** (englisch: automorphism group) oder die **allgemeine lineare Gruppe** (englisch: general linear group) von V. Sie wird mit $\mathrm{GL}(V, K)$ bezeichnet. △

Der Name „allgemeine lineare Gruppe" kann dadurch motiviert werden, dass $\mathrm{GL}(V, K)$ die größte Untergruppe des Monoids $(\operatorname{Endo}(V), \circ)$ ist, die das neutrale Element id_V enthält. $\mathrm{GL}(V, K)$ ist also das Maximum der Menge

$$\{U \subseteq \operatorname{Endo}(V) \mid (U, \circ) \text{ ist Gruppe mit } \mathrm{id}_V \in U\}$$

bzgl. der Inklusionshalbordnung (und auch bzgl. der Untergruppen-Halbordnung). Alle weiteren Untergruppen von $(\operatorname{Endo}(V), \circ)$, die id_V enthalten, sind also Teilmengen (und sogar Untergruppen) von $\mathrm{GL}(V, K)$, vgl. Bemerkung 7.46. In diesem Sinne könnten wir also auch die Einheitengruppe eines Monoids als die „allgemeine Gruppe" des Monoids bezeichnen, vgl. Bemerkung 7.46. Das Attribut „linear" bezieht sich darauf, dass die Elemente von $\mathrm{GL}(V, K)$ lineare Abbildungen sind.

§ 17.4 Faktorräume

In § 8.1 hatten wir gesehen, dass bestimmte Untergruppen namens Normalteiler N aus einer Gruppe G ausfaktorisiert werden können, sodass die Faktormenge $G / N = \big\{[a] = a \star N \,\big|\, a \in G\big\}$ bestehend aus den Nebenklassen[31] von N wieder eine Gruppenstruktur trägt, die mit der Struktur der Gruppe verträglich ist. Die letzte Aussage bedeutete, dass die kanonische Surjektion $\pi\colon a \mapsto [a]$, also der Übergang von einem Gruppenelement zu seiner Nebenklasse, ein (surjektiver) Gruppenhomomorphismus ist, sodass also $[a \star b] = [a] \mathbin{\tilde{\star}} [b]$ gilt: „Erst verknüpfen, dann

[31] **Nebenklasse** war nur ein anderer Name für eine Äquivalenzklasse der durch einen Normalteiler N induzierten Äquivalenzrelation $a \overset{N}{\sim} b \Leftrightarrow a \star N = b \star N \Leftrightarrow a' \star b \in N$, siehe § 7.5.

zur Nebenklasse übergehen ergibt dasselbe wie erst zur Nebenklasse übergehen und dann verknüpfen."

Die Einführung der Faktorgruppe $(G / N, \tilde{\star})$ — einer gröberen Version der Gruppe G — war wichtig für das Verständnis der Wirkungsweise von Gruppenhomomorphismen (Homomorphiesatz für Gruppen 8.25). Dieser Satz besagte, dass ein Gruppenhomomorphismus $f\colon G_1 \to G_2$ „nebenklassenweise" in der Faktorgruppe $G_1 / \operatorname{Kern}(f)$ wirkt, also eine ganze Nebenklasse $[a] = a \star \operatorname{Kern}(f)$ auf ein Element in $\operatorname{Bild}(f)$ abbildet, und zwar verschiedene Nebenklassen auf verschiedene Bildelemente. Kurz gesagt: $G_1 / \operatorname{Kern}(f) \cong \operatorname{Bild}(f)$.

Die gleiche Konstruktion könnten wir in einem Vektorraum $(V, +, \cdot)$ einsetzen, da ja $(V, +)$ eine abelsche Gruppe ist und daher sogar jede Untergruppe einen Normalteiler bildet. Allerdings zielen wir darauf ab, dass die Faktormenge nicht nur eine Gruppenstruktur trägt, sondern wieder zu einem Vektorraum wird. Diese zusätzliche Kompatibilität der Nebenklassenbildung mit der S-Multiplikation erhalten wir genau dann, wenn wir anstelle beliebiger Normalteiler von $(V, +)$ nur Unterräume von $(V, +, \cdot)$ verwenden.

Aus der Faktormenge wird damit der **Faktorraum** (englisch: factor space) oder **Quotientenraum** (englisch: quotient space) **von** V **nach** U. Wir sagen auch: „Aus dem Vektorraum $(V, +, \cdot)$ wird der Unterraum U ausfaktorisiert." Jede additive Nebenklasse[32] $[v] = v + U$ wird ein **affiner Unterraum parallel zu** U (englisch: affine subspace parallel to U) genannt. Zwei Vektoren $v_1, v_2 \in V$ gehören zur selben Nebenklasse genau dann, wenn $v_1 + U = v_2 + U$ gilt, also genau dann, wenn $v_1 - v_2 \in U$ gilt. Man ordnet einem affinen Unterraum $v + U$ die **Dimension** (englisch: dimension) $\dim(v + U) \coloneqq \dim(U)$ zu.

Satz 17.17 (Faktorraum, vgl. Satz 9.30 über Faktorringe)**.**

(*i*) Es seien $(K, +, \cdot)$ ein Körper, $(V, +, \cdot)$ ein Vektorraum über K und U ein Unterraum von V. Dann gilt:

(a) Auf der Faktormenge
$$V / U = \left\{ [v] = v + U \,\middle|\, v \in V \right\}$$
sind $\tilde{+}$ und $\tilde{\cdot}$, definiert als
$$[v] \mathbin{\tilde{+}} [w] \coloneqq [v + w] \quad \text{für } v, w \in V, \tag{17.13a}$$
$$\alpha \mathbin{\tilde{\cdot}} [v] \coloneqq [\alpha \cdot v] \quad \text{für } \alpha \in K \text{ und } v \in V, \tag{17.13b}$$
eine innere bzw. äußere Verknüpfung, bzgl. denen $(V / U, \tilde{+}, \tilde{\cdot})$ einen Vektorraum über K bildet. Das neutrale Element bzgl. $\tilde{+}$ ist $[0] = U$, und für die Inversen gilt $\tilde{-}[v] = [-v]$.

(b) Die **kanonische Surjektion** (englisch: canonical surjection) von V auf V / U
$$\pi\colon \begin{cases} V \to V / U \\ v \mapsto [v], \end{cases} \tag{17.14}$$
die jedem Vektor $v \in V$ seine Nebenklasse $[v] = v + U$ zuordnet, ist ein surjektiver Vektorraumhomomorphismus. Es gilt $\operatorname{Kern}(\pi) = U$.

[32] Das Attribut „additiv" sagen wir gelegentlich dazu, obwohl wir in Vektorräumen i. A. keine anderen Nebenklassen definieren können.

(*ii*) Es seien $(K, +, \cdot)$ ein Körper, $(V, +, \cdot)$ ein Vektorraum über K und $(U, +)$ irgendeine Untergruppe von $(V, +)$. Sind die Verknüpfungen (17.13) auf der Menge der Nebenklassen $V\,/\,U$ wohldefiniert, dann ist U notwendigerweise ein Unterraum von V.

Beweis. Aussage (i): Wir zeigen zunächst die Aussage (a) und weisen die Bedingungen aus der Definition 11.1 eines Vektorraumes nach.

Nach Satz 8.21 ist $(V\,/\,U, \tilde{+})$ eine abelsche Gruppe, da $(U, +)$ eine Untergruppe und damit ein Normalteiler der abelschen Gruppe $(V, +)$ ist.

Schritt 1: Nun müssen wir uns zunächst davon überzeugen, dass die S-Multiplikation $\tilde{\cdot}$ überhaupt wohldefiniert ist. Es sei dazu $\alpha \in K$ und $v_1, v_2 \in V$ mit $[v_1] = [v_2]$, also $v_1 + U = v_2 + U$. Im Fall $\alpha \neq 0$ gilt $\alpha\, U = U$. (**Quizfrage 17.7:** Warum?) Damit folgt

$$\begin{aligned}
\alpha \,\tilde{\cdot}\, [v_1] &= [\alpha\, v_1] && \text{nach Definition von } \tilde{\cdot} \\
&= \alpha\, v_1 + U && \text{nach Definition von } [\cdot] \\
&= \alpha\,(v_1 + U) && \text{da } \alpha\, U = U \text{ ist} \\
&= \alpha\,(v_2 + U) && \text{da } [v_1] = [v_2] \text{ vorausgesetzt wurde} \\
&= \alpha\, v_2 + U && \text{da } \alpha\, U = U \text{ ist} \\
&= [\alpha\, v_2] && \text{nach Definition von } [\cdot] \\
&= \alpha \,\tilde{\cdot}\, [v_2] && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Im Fall $\alpha = 0$ gilt

$$\begin{aligned}
\alpha \,\tilde{\cdot}\, [v_1] &= [0\, v_1] && \text{nach Definition von } \tilde{\cdot} \\
&= [0\, v_2] && \text{da } 0\, v_1 = 0 = 0\, v_2 \text{ gilt} \\
&= \alpha \,\tilde{\cdot}\, [v_2] && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Schritt 2: Nun weisen wir das Assoziativgesetz (11.1a) nach:

$$\begin{aligned}
(\alpha\, \beta) \,\tilde{\cdot}\, [v] &= [(\alpha\, \beta)\, v] && \text{nach Definition von } \tilde{\cdot} \\
&= [\alpha\,(\beta\, v)] && \text{aufgrund des Assoziativgesetzes (11.1a) in } (V, +, \cdot) \\
&= \alpha \,\tilde{\cdot}\, [\beta\, v] && \text{nach Definition von } \tilde{\cdot} \\
&= \alpha \,\tilde{\cdot}\, (\beta \,\tilde{\cdot}\, [v]) && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Schritt 3: Es folgen die Distributivgesetze (11.1b) und (11.1c) in $(V\,/\,U, \tilde{+}, \tilde{\cdot})$:

$$\begin{aligned}
&\alpha \,\tilde{\cdot}\, ([v_1] \,\tilde{+}\, [v_2]) \\
&\quad = \alpha \,\tilde{\cdot}\, [v_1 + v_2] && \text{nach Definition von } \tilde{+} \\
&\quad = [\alpha\,(v_1 + v_2)] && \text{nach Definition von } \tilde{\cdot} \\
&\quad = [\alpha\, v_1 + \alpha\, v_2] && \text{aufgrund des Distributivgesetzes (11.1b) in } (V, +, \cdot) \\
&\quad = [\alpha\, v_1] \,\tilde{+}\, [\alpha\, v_2] && \text{nach Definition von } \tilde{+} \\
&\quad = \alpha \,\tilde{\cdot}\, [v_1] \,\tilde{+}\, \alpha \,\tilde{\cdot}\, [v_2] && \text{nach Definition von } \tilde{\cdot}
\end{aligned}$$

und

$$\begin{aligned}(\alpha+\beta) \,\tilde{\cdot}\, [v] &= [(\alpha+\beta)\,v] && \text{nach Definition von } \tilde{\cdot} \\ &= [\alpha\, v + \beta\, v] && \text{aufgrund des Distributivgesetzes (11.1c) in } (V,+,\cdot) \\ &= [\alpha\, v] \,\tilde{+}\, [\beta\, v] && \text{nach Definition von } \tilde{+} \\ &= \alpha \,\tilde{\cdot}\, [v] \,\tilde{+}\, \beta \,\tilde{\cdot}\, [v] && \text{nach Definition von } \tilde{\cdot}.\end{aligned}$$

Schritt 4: Und schließlich zeigen wir, dass $1 \in K$ neutrales Element bzgl. $\tilde{\cdot}$ ist:

$$\begin{aligned}1 \,\tilde{\cdot}\, [v] &= [1\, v] && \text{nach Definition von } \tilde{\cdot} \\ &= [v] && \text{da 1 neutrales Element bzgl. } \cdot \text{ in } (V,+,\cdot) \text{ ist.}\end{aligned}$$

Damit ist $(V/U, \tilde{+}, \tilde{\cdot})$ als Vektorraum bestätigt. Die Aussagen über das neutrale Element $[0] = U$ und über die Inversen $\tilde{-}[v] = [-v]$ folgen bereits aus dem Satz 8.21 über die Eigenschaften der abelschen Gruppe $(V/U, \tilde{+})$.

Aussage (b): Wir wissen bereits aus Satz 8.21, dass π ein surjektiver Gruppenhomomorphismus $(V,+) \to (V/U, \tilde{+})$ mit $\operatorname{Kern}(\pi) = U$ ist. Die Strukturverträglichkeit mit der Multiplikation ist gerade die Aussage (17.13b).

Aussage (*ii*): Die Eigenschaft, ein Vektorraumhomomorphismus zu sein, bedeutet

$$\begin{aligned}\pi(v+w) &= \pi(v) \,\tilde{+}\, \pi(v) \\ \text{und} \quad \pi(\alpha\, v) &= \alpha \,\tilde{\cdot}\, \pi(v)\end{aligned}$$

für alle $v, w \in V$ und $\alpha \in K$. Nach Definition von π heißt das aber

$$\begin{aligned}[v+w] &= [v] \,\tilde{+}\, [v] \\ \text{und} \quad [\alpha\, v] &= \alpha \,\tilde{\cdot}\, [v],\end{aligned}$$

was gerade die Definition von $\tilde{+}$ und $\tilde{\cdot}$ war. Die Surjektivität von π ist klar, denn ein beliebiges Element $[v]$ von V/U ist gerade das Bild von v unter π. Es gilt $\operatorname{Kern}(\pi) = \pi^{-1}(\{[0]\}) = U$.

Aussage (*ii*): Wir bemerken zunächst, dass die Addition (17.13a) unter den Voraussetzungen wohldefiniert ist, da $(U,+)$ ein Normalteiler der kommutativen Gruppe $(V,+)$ ist (Satz 8.21). Aus dem Unterraumkriterium (Satz 11.11) bleibt noch $\alpha\, U \subseteq U$ für beliebiges $\alpha \in K$ zu zeigen. In der Tat gilt für $v \in U$ die Beziehung $[v] = [0] = U$, da die additiven Nebenklassen von U eine Partition von V bilden. Aufgrund der Wohldefiniertheit von (17.13b) gilt für $\alpha \in K$ weiter

$$[\alpha\, v] = [\alpha\, 0] = [0],$$

also $\alpha\, v \in [0] = U$, was zu zeigen war. □

Bemerkung 17.18 (Faktorraum, vgl. Bemerkung 9.31 zu Faktorringen)**.**
Praktisch können wir den Faktorraum $(V/U, \tilde{+}, \tilde{\cdot})$ benutzen, um wie im Vektorraum $(V,+,\cdot)$ zu „rechnen", wobei jedoch Vektoren v, w in derselben Äquivalenzklasse (für die also $v - w \in U$ gilt) nicht mehr unterschieden werden. Der Faktorraum $(V/U, \tilde{+}, \tilde{\cdot})$ ist also eine „gröbere Version" des Vektorraumes $(V,+,\cdot)$.

Die Dimension von V/U werden wir später noch charakterisieren, siehe Satz 18.6. △

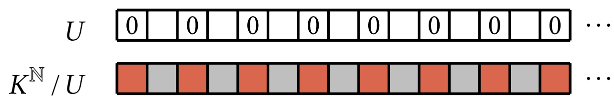
Beispiel 17.19 (Faktorraum, vgl. Beispiel 9.32 zu Faktorringen)**.**

(i) Es sei V ein beliebiger Vektorraum und $U = \{0\}$ der Nullraum, einer der beiden trivialen Unterräume von V. Der zugehörige Faktorraum $V\,/\,\{0\}$ ist isomorph zum Ausgangsraum V selbst.

(ii) Es sei V ein beliebiger Vektorraum und $U = V$ der andere triviale Unterraum von V. Der zugehörige Faktorraum $V\,/\,V$ ist isomorph zum Nullraum $\{0\}$.

(iii) Es sei $(K, +, \cdot)$ ein Körper und $(K^{\mathbb{N}}, +, \cdot)$ der Folgenraum über K. Wählen wir U als den Unterraum derjenigen Folgen $(x_n)_{n\in\mathbb{N}}$, deren Einträge an allen ungeraden Indizes $n \in \mathbb{N}$ gleich 0 sind, dann besteht der Faktorraum $K^{\mathbb{N}}\,/\,U$ aus Äquivalenzklassen von Folgen, wobei zwei Folgen genau dann in derselben Äquivalenzklasse (Nebenklasse) liegen, wenn sie in allen ungeraden Einträgen übereinstimmen. Die geraden Einträge der Folgen spielen keine Rolle mehr.

Wir illustrieren die Struktur des Unterraumes U und der Äquivalenzklassen in $K^{\mathbb{N}}\,/\,U$ wie folgt:



Die mit gekennzeichneten Einträge sind dabei für alle Elemente einer Äquivalenzklasse identisch. Diese Einträge legen die Äquivalenzklasse fest, bestimmen also das Element von $K^{\mathbb{N}}\,/\,U$. Verschiedene Repräsentanten einer Äquivalenzklasse unterscheiden sich nur in den mit markierten Einträgen. △

Bemerkung 17.20 (Unterräume sind genau die Kerne von Vektorraumhomomorphismen, vgl. Bemerkung 9.33 zu Kernen von Ringhomomorphismen)**.**
Es sei $(V, +, \cdot)$ über dem Körper K.

(i) Nach Lemma 17.8 ist der Unterraum $\operatorname{Kern}(f)$ für jeden beliebigen Vektorraumhomomorphismus $f\colon V \to W$ in irgendeinen Vektorraum $(W, +, \cdot)$ über demselben Körper K immer ein Unterraum von $(V, +, \cdot)$.

(ii) Umgekehrt gilt auch, dass jeder Unterraum U von $(V, +, \cdot)$ der Kern eines Vektorraumhomomorphismus ist. Dazu wählen wir einfach $W \coloneqq (V\,/\,U, \tilde{+}, \tilde{\cdot})$ als Zielraum und die kanonische Surjektion $\pi\colon V \to V\,/\,U$ als Vektorraumhomomorphismus. Dann gilt $\operatorname{Kern}(\pi) = U$. △

§ 17.5 Der Homomorphiesatz für Vektorräume

Mit Hilfe des Wissens über Faktorräume können wir nun die Struktur von Vektorraumhomomorphismen genauer analysieren. Der folgende Struktursatz besagt, dass ein Vektorraumhomomorphismus $f\colon V \to W$ „nebenklassenweise" wirkt. Er bildet also eine gesamte additive Nebenklasse von $\operatorname{Kern}(f)$ — das ist ein affiner Unterraum parallel zu $\operatorname{Kern}(f)$ — auf ein- und dasselbe Element von W ab und verschiedene Nebenklassen auf verschiedene Elemente. Dadurch ist das $\operatorname{Bild}(f)$ eines solchen Vektorraumhomomorphismus bereits im Wesentlichen (d. h. bis auf Isomorphie) festgelegt durch V und den Unterraum $\operatorname{Kern}(f)$.

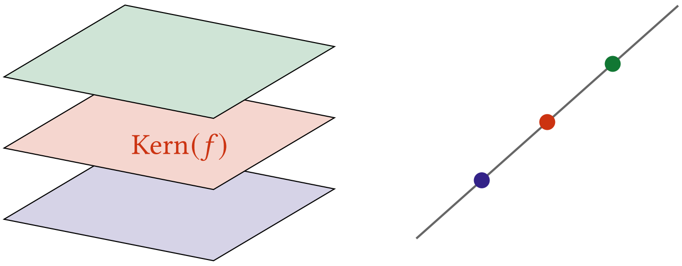


Abbildung 17.1.: Illustration des Homomorphiesatzes für Vektorraumhomomorphismen 17.21, vgl. Abbildung 8.2. Alle Elemente einer Nebenklasse des Unterraumes Kern(f) werden auf ein- und dasselbe Element von W abgebildet und verschiedene Nebenklassen auf verschiedene Elemente.

Satz 17.21 (**Homomorphiesatz für Vektorräume**[33], vgl. Homomorphiesatz für Ringe 9.38)**.** Es seien V, W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ ein Homomorphismus. Dann gilt

$$V \,/\, \operatorname{Kern}(f) \cong \operatorname{Bild}(f) \tag{17.15a}$$

mit dem Isomorphismus

$$I([v]) \coloneqq f(v) \quad \text{für } [v] = v + \operatorname{Kern}(f) \in V \,/\, \operatorname{Kern}(f). \tag{17.15b}$$

Beweis. Der Vektorraumhomomorphismus ist insbesondere ein Gruppenhomomorphismus $f\colon (V,+) \to (W,+)$, und $\operatorname{Kern}(f) = \{v \in V \mid f(v) = 0\}$ hängt nicht davon ab, ob wir f als Homomorphismus von Gruppen oder von Vektorräumen betrachten. Aus dem Homomorphiesatz für Gruppen 8.25 folgt also sofort, dass $V \,/\, \operatorname{Kern}(f)$ und $\operatorname{Bild}(f)$ im Sinne von Gruppen isomorph sind, und zwar vermöge des Isomorphismus (17.15b).

Es bleibt nur zu zeigen, dass I tatsächlich auch ein Isomorphismus im Sinne von Vektorräumen ist. Dazu fehlt nur der Nachweis der Homogenität (17.1b), der aber einfach zu erbringen ist:

$$\begin{aligned} I(\alpha \mathbin{\tilde{\cdot}} [v]) &= I([\alpha\, v]) && \text{nach Definition von } \tilde{\cdot} \\ &= f(\alpha\, v) && \text{nach Definition von } I \\ &= \alpha\, f(v) && \text{aufgrund der Homogenität von } f \\ &= \alpha\, I([v]) && \text{nach Definition von } I. \end{aligned}$$

□

Beispiel 17.22 (Homomorphiesatz für Vektorräume, vgl. Beispiel 8.27)**.**

(*i*) Es seien V, W Vektorräume über demselben Körper K. Wir betrachten den Nullhomomorphismus $f\colon V \to W$ mit $f(v) = 0$ für alle $v \in V$. Gemäß Homomorphiesatz für Vektorräume 17.21 ist

$$V \,/\, \operatorname{Kern}(f) = V \,/\, V \;\cong\; \operatorname{Bild}(f) = \{0\}.$$

[33]englisch: fundamental theorem on vector space homomorphisms

(*ii*) Es sei K^n der Standardvektorraum der Dimension $n \in \mathbb{N}$ über einem Körper K. Wir betrachten die Projektion auf die i-te Koordinate (Beispiel 17.5)

$$\pi_i \colon K^n \ni x \mapsto x_i \in K$$

für $1 \leqslant i \leqslant n \in \mathbb{N}$. Dann ist $\operatorname{Kern}(\pi_i) = \{x \in K^n \mid x_i = 0\}$, also besteht $K^n / \operatorname{Kern}(\pi_i)$ aus Äquivalenzklassen von Vektoren, die jeweils in der i-ten Koordinate übereinstimmen. Eine gesamte Äquivalenzklasse wird durch π_i auf ein- und dasselbe Element von K abgebildet.

Gemäß Homomorphiesatz für Vektorräume 17.21 gilt

$$K^n / \operatorname{Kern}(\pi_i) = K^n / \{x \in K^n \mid x_i = 0\} \;\cong\; \operatorname{Bild}(\pi_i) = K.$$

(*iii*) Es sei $K^{\mathbb{N}}$ der Folgenraum über einem Körper K. Wir betrachten auch hier die Projektion auf die i-Koordinate

$$\pi_i \colon K^{\mathbb{N}} \ni x \mapsto x_i \in K$$

für $i \in \mathbb{N}$, die eine Folge auf das i-te Folgenglied abbildet. Dann ist $\operatorname{Kern}(\pi_i) = \{(x_n)_{n \in \mathbb{N}} \mid x_i = 0\}$, also besteht $K^{\mathbb{N}} / \operatorname{Kern}(\pi_i)$ aus Äquivalenzklassen von Folgen, die jeweils im i-ten Folgenglied übereinstimmen. Eine gesamte Äquivalenzklasse wird auf ein- und dasselbe Element von K abgebildet. Gemäß Homomorphiesatz für Vektorräume 17.21 ist

$$K^{\mathbb{N}} / \operatorname{Kern}(\pi_i) = K^{\mathbb{N}} / \{x \in K^{\mathbb{N}} \mid x_i = 0\} \;\cong\; \operatorname{Bild}(\pi_i) = K. \qquad \triangle$$

Abschließend stellen wir den Homomorphiesatz für Vektorräume 17.21 nochmal schematisch mit Hilfe eines kommutativen Diagrammes dar. Dazu sei $i \colon \operatorname{Bild}(f) \ni w \mapsto w \in W$ der injektive Homomorphismus der kanonischen Einbettung.

$$\begin{tikzcd} W & V \arrow[l, "f"'] \arrow[d, "\pi"] \\ \operatorname{Bild}(f) \arrow[u, "i"] & V / \operatorname{Kern}(f) \arrow[l, "I"] \end{tikzcd} \qquad f = \underbrace{i}_{\text{einbetten}} \circ \overbrace{I}^{\text{isomorph abbilden}} \circ \underbrace{\pi}_{\text{vergröbern}}.$$

Expertenwissen: universelle Eigenschaft von Faktorräumen

Ähnlich wie Faktorgruppen (siehe Material im Anschluss an Satz 8.25) können auch Faktorräume durch eine universelle Eigenschaft charakterisiert werden:

Es seien V, W Vektorräume über demselben Körper K und $f \colon V \to W$ ein Homomorphismus. Weiter sei U ein Unterraum von V. Dann sind äquivalent:

(*i*) $U \subseteq \operatorname{Kern}(f)$, also $f(U) = \{0\}$.

(*ii*) Der Homomorphismus f **faktorisiert durch** die kanonische Surjektion $\pi \colon V \to V / U$, d. h., es gibt einen eindeutig bestimmten Homomorphismus $g \colon V / U \to W$ mit $f = g \circ \pi$.

$$\begin{tikzcd} V \arrow[r, "\pi"] \arrow[dr, "f"'] & V / U \arrow[d, "g"] \\ & W \end{tikzcd}$$

Diese Eigenschaft nennt sich die **universelle Eigenschaft von Faktorräumen** (englisch: universal property of quotient spaces), denn sie charakterisiert Faktorräume bis auf Isomorphie eindeutig. Gilt also die obige Äquivalenzaussage mit irgendeinem K-Vektorraum Z anstelle von V / U und einem surjektiven Vektorraumhomomorphismus $\pi \colon V \to Z$, dann ist Z isomorph zu V / U. Die universelle Eigenschaft ermöglicht es, Faktorräume (bis auf Isomorphie) zu definieren, ohne die konkrete Konstruktion über Nebenklassen zu verwenden.

Ende der Vorlesung 24

§ 18 Dimensionssätze

Literatur: Fischer, Springborn, 2020, Kapitel 3.2; Bosch, 2014, Kapitel 2; Beutelspacher, 2014, Kapitel 5.2

Wir wollen in diesem Abschnitt den Zusammenhang der Dimensionen der am Homomorphiesatz für Vektorräume 17.21 beteiligten Räume V, $\operatorname{Kern}(f)$ und $\operatorname{Bild}(f)$ untersuchen. Wir hatten am Beispiel 17.22 schon sehen können: Je höher die Dimension des ausfaktorisierten Unterraumes, desto geringer die verbleibende Dimension des Faktorraumes.

§ 18.1 Zusammenhang von Dimension und Isomorphie

Als Folgerungen aus dem Existenz- und Eindeutigkeitssatz für Vektorraumhomomorphismen 17.10 erhalten wir folgende bemerkenswerte Resultate:

Satz 18.1 (Vektorräume gleicher endlicher Dimension sind isomorph)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K. Dann sind äquivalent:

(i) V und W sind isomorphe Vektorräume.

(ii) $\dim(V) = \dim(W)$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $f \colon V \to W$ ein Isomorphismus. Nach Folgerung 13.6 existiert eine Basis $(v_1, \ldots, v_n)$ von V für ein $n \in \mathbb{N}_0$. Setze $w_i \coloneqq f(v_i)$ für $i \in [\![1, n]\!]$. Dann ist nach Satz 17.10 $(w_1, \ldots, w_n)$ eine Basis von W. Beide Basen sind gleichmächtig, also gilt $\dim(V) = \dim(W)$.

Aussage (ii) $\Rightarrow$ Aussage (i): Es gelte $\dim(V) = \dim(W) = n \in \mathbb{N}_0$. Es seien $(v_1, \ldots, v_n)$ eine beliebige Basis von V und $(w_1, \ldots, w_n)$ eine beliebige Basis von W. Nach Satz 17.10 gibt es genau eine lineare Abbildung $f : V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i = 1, \ldots, n$. Nach Satz 17.10 ist dieses f bijektiv, also sind V und W zueinander isomorph. □

Satz 18.1 besagt, dass es bis auf Isomorphie nur einen einzigen K-Vektorraum der Dimension $n \in \mathbb{N}_0$ gibt! Anders gesagt: Alle K-Vektorräume V mit $\dim(V) = n \in \mathbb{N}_0$ sind zueinander isomorph. Die Dimension beschreibt also einen endlich-dimensionalen K-Vektorraum bereits vollständig. Das werden wir später in § 19 noch ausnutzen.

Es gilt sogar folgende Verallgemeinerung von Satz 18.1:

Satz 18.2 (Vektorräume mit gleichmächtigen Basen sind isomorph$^{\text{AoC}}$)**.**
Es seien V, W Vektorräume über demselben Körper K. Dann sind äquivalent:

(i) V und W sind isomorphe Vektorräume.

(ii) Es gibt eine Basis $(v_i)_{i \in I}$ von V und eine Basis $(w_j)_{j \in J}$ von W, die gleichmächtig sind.

Es gibt also nicht nur zu jeder endlichen Zahl, sondern sogar zu jeder Kardinalzahl eine Äquivalenzklasse isomorpher Vektorräume.$^{\text{AoC}}$ Wir können die Vektorräume K^n als natürliche Repräsentanten der Äquivalenzklasse der Vektorräume mit einer Basis der endlichen Mächtigkeit $n \in \mathbb{N}_0$ ansehen. Weiter ist der Folgenraum $K^{\mathbb{N}}$ der natürliche Repräsentant der Vektorräume mit abzählbar unendlicher Basis.

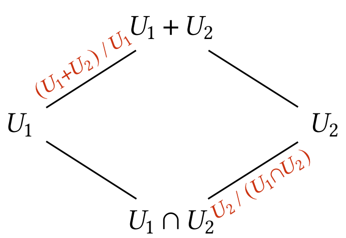
§ 18.2 Dimension von Faktorräumen

In diesem Abschnitt bestimmen wir die Dimension eines Faktorraumes V / U. Zur Vorbereitung benötigen wir folgendes Resultat.

Satz 18.3 (Faktorraum einer Summe von Unterräumen)**.**
Es seien V ein Vektorraum und U_1, U_2 zwei Unterräume von V. Dann gilt:

$$(U_1 + U_2) / U_1 \;\cong\; U_2 / (U_1 \cap U_2). \tag{18.1}$$

Beachte: Beide Seiten der Gleichung messen auf unterschiedliche Weise, „wieviel vom Unterraum U_2 nicht bereits im Unterraum U_1 liegt", und kommen dabei zum selben Ergebnis (bis auf Isomorphie). Grafisch wird das Ergebnis manchmal durch das folgende Diagramm dargestellt, weshalb Satz 18.3 im Englischen auch als diamond theorem bezeichnet wird:[34]



[34]Ein anderer Name für Satz 18.3 ist **zweiter Isomorphiesatz** (englisch: second isomorphism theorem). Der **erste Isomorphiesatz** (englisch: first isomorphism theorem) ist der Homomorphiesatz für Vektorräume 17.21.

Beweis. Der Beweis basiert auf der Idee, einen surjektiven Homomorphismus $f\colon U_2 \to (U_1 + U_2)\,/\,U_1$ anzugeben mit $\operatorname{Kern}(f) = U_1 \cap U_2$. Aus dem Homomorphiesatz für Vektorräume 17.21 folgt dann $U_2\,/\operatorname{Kern}(f) = U_2\,/\,(U_1 \cap U_2) \cong \operatorname{Bild}(f) = (U_1 + U_2)\,/\,U_1$, also die Behauptung (18.1).

$$\begin{tikzcd} U_2 \arrow[r, "\pi"] \arrow[dr, "f"'] & U_2\,/\,(U_1 \cap U_2) \arrow[d, "\cong"] \\ & (U_1 + U_2)\,/\,U_1 \end{tikzcd}$$

Wir definieren $f(u_2) \coloneqq [u_2] = u_2 + U_1$ für $u_2 \in U_2$. (Im gesamten Beweis bedeutet $[\cdot]$ immer eine Nebenklasse von U_1.)

Schritt 1: $f\colon U_2 \to (U_1 + U_2)\,/\,U_1$ ist ein Homomorphismus:

Zunächst stellen wir fest, dass $f(u_2) = u_2 + U_1$ in $U_2\,/\,U_1$ liegt, also erst recht in $(U_1 + U_2)\,/\,U_1$. Wir weisen die Linearität nach:

$$\begin{aligned} f(u_2 + v_2) &= [u_2 + v_2] && \text{nach Definition von } f \\ &= [u_2] \mathbin{\tilde{+}} [v_2] && \text{mit } \tilde{+} \text{ in } (U_1 + U_2)\,/\,U_1 \\ &= f(u_2) \mathbin{\tilde{+}} f(v_2) && \text{nach Definition von } f \end{aligned}$$

und

$$\begin{aligned} f(\alpha\, u_2) &= [\alpha\, u_2] && \text{nach Definition von } f \\ &= \alpha \mathbin{\tilde{\cdot}} [u_2] && \text{mit } \tilde{\cdot} \text{ in } (U_1 + U_2)\,/\,U_1 \\ &= \alpha \mathbin{\tilde{\cdot}} f(u_2) && \text{nach Definition von } f. \end{aligned}$$

Schritt 2: $\operatorname{Kern}(f) = U_1 \cap U_2$:

Es ist

$$\begin{aligned} \operatorname{Kern}(f) &= \{u_2 \in U_2 \mid f(u_2) = [0]\} && \text{nach Definition des Kerns und des} \\ & && \text{neutralen Elements } [0] \text{ in } (U_1 + U_2)\,/\,U_1 \\ &= \{u_2 \in U_2 \mid [u_2] = [0]\} && \text{nach Definition von } f \\ &= \{u_2 \in U_2 \mid u_2 - 0 \in U_1\} && \text{nach Definition von } [\cdot] \\ &= U_1 \cap U_2. \end{aligned}$$

Schritt 3: f ist surjektiv, d. h., $\operatorname{Bild}(f) = (U_1 + U_2)\,/\,U_1$.

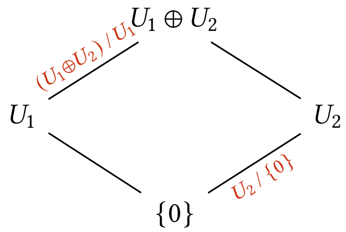
Es sei $[w] \in (U_1 + U_2)\,/\,U_1$. Wegen $w \in U_1 + U_2$ existieren $u_1 \in U_1$ und $u_2 \in U_2$ mit $w = u_1 + u_2$. Das heißt aber $[w] = u_1 + u_2 + U_1 = u_2 + U_1 = [u_2] = f(u_2)$. □

Für den Fall $U_1 \cap U_2 = \{0\}$ erhalten wir folgendes Korollar:

Folgerung 18.4 (Faktorraum einer direkten Summe von Unterräumen).
Es seien V ein Vektorraum und U_1, U_2 zwei Unterräume von V. Wenn $U_1 \cap U_2 = \{0\}$ gilt, dann ist

$$(U_1 \oplus U_2)\,/\,U_1 \cong U_2. \tag{18.2}$$

Beachte: Folgerung 18.4 besagt insbesondere, dass jeder zu einem Unterraum U_1 eines Vektorraumes V komplementäre Unterraum U_2 isomorph zum Faktorraum V / U_1 ist. Analog zu oben können wir das Resultat von Folgerung 18.4 wie folgt illustrieren:



Beweis. Nach Satz 18.3 gilt

$$(U_1 + U_2) / U_1 \cong U_2 / (U_1 \cap U_2).$$

Die Voraussetzung $U_1 \cap U_2 = \{0\}$ bedeutet, dass die Summe $U_1 \oplus U_2$ direkt ist, also folgt

$$(U_1 \oplus U_2) / U_1 \cong U_2 / \{0\} = \big\{[u_2] = u_2 + \{0\} \,\big|\, u_2 \in U_2\big\} \cong U_2,$$

vgl. auch Beispiel 17.19. □

Beispiel 18.5 (Faktorraum einer Summe von Unterräumen)**.**
Wir betrachten ein Beispiel im Folgenraum $V = K^{\mathbb{N}}$ mit den Unterräumen

$$U_1 = \big\{(x_n)_{n\in\mathbb{N}} \,\big|\, x_{2k} = 0 \text{ für alle } k \in \mathbb{N}\big\} \qquad \text{jedes zweite Folgenglied ist Null,}$$
$$U_2 = \big\{(x_n)_{n\in\mathbb{N}} \,\big|\, x_{3k} = 0 \text{ für alle } k \in \mathbb{N}\big\} \qquad \text{jedes dritte Folgenglied ist Null.}$$

Dann gilt

$$U_1 + U_2 = \big\{(x_n)_{n\in\mathbb{N}} \,\big|\, x_{6k} = 0 \text{ für alle } k \in \mathbb{N}\big\} \qquad \text{jedes sechste Folgenglied ist Null,}$$
$$U_1 \cap U_2 = \big\{(x_n)_{n\in\mathbb{N}} \,\big|\, x_{2k} = x_{3k} = 0 \text{ für alle } k \in \mathbb{N}\big\} \qquad \text{jedes zweite und dritte Folgenglied ist Null.}$$

Wir illustrieren die Struktur der Folgen in diesen vier Unterräumen wie folgt:

U_1		0		0		0		0		0		0		0		$\cdots$
U_2			0			0			0			0			0	$\cdots$
$U_1 + U_2$						0						0				$\cdots$
$U_1 \cap U_2$		0	0	0		0		0	0	0		0		0	0	$\cdots$

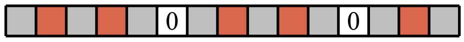
Wir betrachten nun die Faktorräume $(U_1 + U_2) / U_1$ und $U_2 / (U_1 \cap U_2)$:

- Jedes Element von $(U_1 + U_2) / U_1$ ist eine Äquivalenzklasse von Folgen. Zwei Folgen in $U_1 + U_2$ sind äquivalent, wenn ihre Differenz in U_1 liegt. Die Äquivalenzklasse einer Folge $(y_n)_{n\in\mathbb{N}} \in U_1 + U_2$ hat also die Gestalt

$$[(y_n)_{n\in\mathbb{N}}] = \big\{(x_n)_{n\in\mathbb{N}} \,\big|\, x_{6k} = 0 \text{ und } x_{2k} = y_{2k} \text{ für alle } k \in \mathbb{N}\big\}.$$

Eine solche Äquivalenzklasse ist also bereits durch die Werte y_{6k+2} und y_{6k+4} für alle $k \in \mathbb{N}_0$ eines Repräsentanten eindeutig bestimmt.

Wir illustrieren die Struktur der Äquivalenzklassen in $(U_1 + U_2) \,/\, U_1$ wie folgt:

$(U_1 + U_2) \,/\, U_1$  $\cdots$

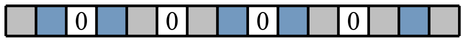
Die mit gekennzeichneten Einträge sind dabei für alle Elemente einer Äquivalenzklasse identisch. Diese Einträge legen die Äquivalenzklasse fest, bestimmen also das Element von $(U_1 + U_2) \,/\, U_1$. Verschiedene Repräsentanten einer Äquivalenzklasse unterscheiden sich nur in den mit markierten Einträgen.

- Jedes Element von $U_2 \,/\, (U_1 \cap U_2)$ ist ebenfalls eine Äquivalenzklasse von Folgen. Zwei Folgen in U_2 sind äquivalent, wenn ihre Differenz in $U_1 \cap U_2$ liegt. Die Äquivalenzklasse einer Folge $(z_n)_{n \in \mathbb{N}} \in U_2$ hat also die Gestalt

$$[(z_n)_{n \in \mathbb{N}}] = \left\{ (x_n)_{n \in \mathbb{N}} \,\middle|\, x_{3k} = 0 \text{ und } x_{2k} = z_{2k} \text{ und } x_{3k} = z_{3k} \text{ für alle } k \in \mathbb{N} \right\}.$$

Eine solche Äquivalenzklasse ist also bereits durch die Werte z_{6k+2} und z_{6k+4} für alle $k \in \mathbb{N}_0$ eines Repräsentanten eindeutig bestimmt.

Wir illustrieren die Struktur der Äquivalenzklassen in $U_2 \,/\, (U_1 \cap U_2)$ wie folgt:

$U_2 \,/\, (U_1 \cap U_2)$  $\cdots$

Die mit gekennzeichneten Einträge sind dabei für alle Elemente einer Äquivalenzklasse identisch. Diese Einträge legen die Äquivalenzklasse fest, bestimmen also das Element von $U_2 \,/\, (U_1 \cap U_2)$. Verschiedene Repräsentanten einer Äquivalenzklasse unterscheiden sich nur in den mit markierten Einträgen.

Offenbar ist die für $k \in \mathbb{N}_0$ durch

$$\begin{aligned} z_{6k+1} &\coloneqq 0 & z_{6k+3} &\coloneqq 0 & z_{6k+5} &\coloneqq 0 \\ z_{6k+2} &\coloneqq y_{6k+2} & z_{6k+4} &\coloneqq y_{6k+4} & z_{6k+6} &\coloneqq 0 \end{aligned}$$

definierte Zuordnung $[(y_n)_{n \in \mathbb{N}}] \mapsto [(z_n)_{n \in \mathbb{N}}]$ ein möglicher Isomorphismus zwischen den Faktorräumen $(U_1 + U_2) \,/\, U_1$ und $U_2 \,/\, (U_1 \cap U_2)$. △

Nun können wir die Dimension von Faktorräumen bestimmen:

Satz 18.6 (Dimension des Faktorraumes[35] AoC).
Es seien V ein Vektorraum und U ein Unterraum von V. Dann gilt

(i) Ist W ein zu U komplementärer Unterraum von V, dann gilt

$$V \,/\, U \cong W \tag{18.3a}$$
$$\dim(V \,/\, U) = \operatorname{codim}(U) \tag{18.3b}$$
$$\dim(V) = \dim(V \,/\, U) + \dim(U). \tag{18.3c}$$

[35] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

(*ii*) Ist V endlich-dimensional, dann gilt insbesondere

$$\dim(V\,/\,U) = \dim(V) - \dim(U). \tag{18.4}$$

Beweis. Aussage (*i*): Es sei W ein zu U komplementärer Unterraum von V. Dann gilt $V = U \oplus W$, und aus Folgerung 18.4 folgt

$$V\,/\,U = (U \oplus W)\,/\,U \cong W.$$

Aus Satz 18.2 folgt $\dim(V\,/\,U) = \dim(W) = \operatorname{codim}(U)$. Nach Folgerung 14.10 erhalten wir schließlich $\dim(V) = \dim(U) + \dim(W) = \dim(U) + \dim(V\,/\,U)$.

Aussage (*ii*): Ist V endlich-dimensional, dann gilt nach Folgerung 14.11 $\operatorname{codim}(U) = \dim(V) - \dim(U)$. □

§ 18.3 Dimensionen im Homomorphiesatz

Mit Hilfe von Satz 18.6 können wir nun die Dimensionen der Räume V, $\operatorname{Kern}(f)$ und $\operatorname{Bild}(f)$ im Homomorphiesatz für Vektorräume 17.21, also in der Aussage

$$V\,/\operatorname{Kern}(f) \;\cong\; \operatorname{Bild}(f),$$

untersuchen:

Satz 18.7 (Dimensionen der Räume im Homomorphiesatz für Vektorräume 17.21[36] AoC)**.**
Es seien V und W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ ein Homomorphismus. Dann gilt

$$\dim(\operatorname{Kern}(f)) + \dim(\operatorname{Bild}(f)) = \dim(V). \tag{18.5}$$

Beweis. Mit $U \coloneqq \operatorname{Kern}(f)$ folgt

$$\dim(V) = \dim(\operatorname{Kern}(f)) + \dim(V\,/\operatorname{Kern} f)$$

aus (18.3). Da nach dem Homomorphiesatz für Vektorräume 17.21 $\operatorname{Bild}(f)$ und $V\,/\operatorname{Kern} f$ isomorph sind und nach Satz 18.2 isomorphe Vektorräume dieselbe Dimension besitzen, ist (18.5) gezeigt. □

Definition 18.8 (Rang und Defekt eines Homomorphismus)**.**
Es seien V und W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ ein Homomorphismus. Dann heißt

$$\operatorname{Rang}(f) \coloneqq \dim(\operatorname{Bild}(f)) \tag{18.6}$$

der **Rang** (englisch: rank) der linearen Abbildung f, und

$$\operatorname{Defekt}(f) \coloneqq \dim(\operatorname{Kern}(f)) \tag{18.7}$$

heißt der **Defekt** (englisch: defect) von f. △

[36] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Wir können also die Dimensionsformel (18.5) auch in der Form

$$\mathrm{Defekt}(f) + \mathrm{Rang}(f) = \dim(V) \tag{18.8}$$

schreiben.

Das folgende Resultat erleichtert der Nachweis der Bijektivität (also der Isomorphismus-Eigenschaft) einer linearen Abbildung erheblich.

Folgerung 18.9 (Charakterisierung der Bijektivität von Homomorphismen endlich-dimensionaler Vektorräume)**.**
Es seien V und W Vektorräume über demselben Körper K. Weiter sei $f\colon V \to W$ ein Homomorphismus.

(i) Haben V und W dieselbe endliche Dimension $\dim(V) = \dim(W)$, dann sind äquivalent:
   (a) f ist injektiv.
   (b) $\mathrm{Defekt}(f) = 0$.
   (c) f ist surjektiv.
   (d) $\mathrm{Rang}(f) = \dim(V)$.
   (e) f ist bijektiv.

(ii) Ist V endlich-dimensional und gilt $\dim(V) < \dim(W) \in \mathbb{N} \cup \{\infty\}$, dann kann f nicht surjektiv sein.

(iii) Ist W endlich-dimensional und gilt $\dim(W) < \dim(V) \in \mathbb{N} \cup \{\infty\}$, dann kann f nicht injektiv sein.

(iv) Es seien V oder W endlich-dimensional. Ein Isomorphismus $V \to W$ existiert genau dann, wenn der andere Vektorraum auch endlich-dimensional ist und $\dim(V) = \dim(W)$ gilt.

Beweis. Der Beweis ist Gegenstand der Übung. □

Beispiel 18.10 (Charakterisierung der Bijektivität von Homomorphismen endlich-dimensionaler Vektorräume)**.**
In diesem Beispiel illustrieren wir verschiedene Fälle aus Folgerung 18.9.

(i) Ein injektiver Homomorphismus $f\colon \mathbb{R} \to \mathbb{R}^2$, der nicht surjektiv ist:

$$f(x) \coloneqq \begin{bmatrix} 1 \\ 0 \end{bmatrix} x.$$

(ii) Ein surjektiver Homomorphismus $f\colon \mathbb{R}^2 \to \mathbb{R}$, der nicht injektiv ist:

$$f(x) \coloneqq \begin{bmatrix} 1 & 0 \end{bmatrix} x.$$

(iii) Sind V und W beide unendlich-dimensional, so können alle Kombinationen von Injektivität und Surjektivität auftreten. Wir betrachten dazu Endomorphismen des Vektorraumes $(K^{\mathbb{N}})_{00}$ der endlich getragenen Folgen über einem Körper K. Zur Definition dieser Endomorphismen geben wir (siehe Satz 17.10) jeweils die Bilder der Standardfolgen e_j an (Beispiel 11.17).

- Die Shift-Abbildung nach links, gegeben durch $f(e_j) = e_{j-1}$ für $j \geqslant 2$ und $f(e_1) = 0$ (Nullfolge), ist surjektiv, aber nicht injektiv.
- Die Shift-Abbildung nach rechts, gegeben durch $f(e_j) = e_{j+1}$ für $j \in \mathbb{N}$, ist injektiv, aber nicht surjektiv.
- Die identische Abbildung, gegeben durch $f(e_j) = e_j$ für $j \in \mathbb{N}$, ist bijektiv.
- Die Nullabbildung, gegeben durch $f(e_j) = 0$ für $j \in \mathbb{N}$, ist weder injektiv noch surjektiv. △

Ende der Vorlesung 25

Ende der Woche 12

§ 19 Darstellungsmatrizen von Homomorphismen

Literatur: Fischer, Springborn, 2020, Kapitel 3.4–3.6 und 5.1, 5.3; Bosch, 2014, Kapitel 3.4 und 6.1; Beutelspacher, 2014, Kapitel 5.2 und 8.1–8.2; Jänich, 2008, Kapitel 9.1 und 11.1–11.2

Im gesamten § 19 sind alle Vektorräume endlich-dimensional. Wir erinnern an Satz 18.1, der besagt, dass alle K-Vektorräume der gemeinsamen Dimension $n \in \mathbb{N}_0$ zueinander isomorph sind. Es ist daher möglich und praktisch, für jeden n-dimensionalen K-Vektorraum V eine Art gemeinsame Standarddarstellung zu finden. Dafür bietet sich der Standardvektorraum K^n an.

§ 19.1 Koordinatendarstellung in endlich-dimensionalen Vektorräumen

Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Der folgende Satz gibt an, wie wir mit Hilfe einer Basis von V jeden beliebigen Vektor in V mit Hilfe seines Koordinatenvektors in K^n bzgl. dieser Basis darstellen können.

Satz 19.1 (Koordinatenvektor, Syntheseabbildung, Analyseabbildung)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei die Familie $B = (v_1, \ldots, v_n)$ eine Basis von V.

(i) Die Abbildung

$$\Phi_B \colon K^n \ni \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i \, v_i \in V \tag{19.1}$$

ist ein linearer Isomorphismus $K^n \to V$. Diese erzeugt aus einem **Koordinatenvektor**[37] (englisch: coordinate vector) $x \in K^n$ die zugehörige Linearkombination der Basisvektoren v_i. Wir bezeichnen die Abbildung Φ_B daher auch als die **Syntheseabbildung** (englisch: synthesis map) bzgl. der Basis B.

(ii) Der zu Φ_B inverse Isomorphismus ist die Abbildung

$$\Phi_B^{-1} \colon V \ni v \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n, \tag{19.2}$$

die jedem Vektor $v \in V$ seinen eindeutigen **Koordinatenvektor** $x \in K^n$ bzgl. der Basis B zuordnet. Wir nennen daher Φ_B^{-1} auch die **Koordinatenabbildung** (englisch: coordinate map) oder die **Analyseabbildung** (englisch: analysis map) bzgl. der Basis B. Der Eintrag $x_i \in K$ heißt die i-te **Koordinate** (englisch: coordinate) des Vektors $v \in V$ bzgl. der Basis B.

Beweis. Φ_B ist linear und bildet die Basis $(e_1, \ldots, e_n)$ in K^n auf die Basis $(v_1, \ldots, v_n)$ ab. Damit ist Φ_B nach Satz 17.10 bijektiv, also ein linearer Isomorphismus, und die Inverse Φ_B^{-1} ebenfalls (Satz 17.3). □

Wir werden Koordinatenvektoren typischerweise mit x bezeichnen.

Beispiel 19.2 (Koordinatendarstellung)**.**

(i) Im $\mathbb{R}$-Vektorraum $\mathbb{R}^{(-\frac{\pi}{2}, \frac{\pi}{2})}$ der Funktionen $(-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ hat der Vektor

$$5 \sin - 3 \cos + 7 \tan$$

bzgl. der Basis $B = (\sin, \cos, \tan)$ den Koeffizientenvektor $\left(\begin{smallmatrix} 5 \\ -3 \\ 7 \end{smallmatrix}\right)$, denn es gilt

$$5 \sin - 3 \cos + 7 \tan = 5 \sin + (-3) \cos + 7 \tan.$$

(ii) Um dieselbe Funktion $5 \sin - 3 \cos + 7 \tan$ in der Basis $B = (\tan - \cos + \sin,\ \tan + 3 \sin,\ \cos + \sin)$ darzustellen, schreiben wir sie als Linearkombination der Basisvektoren mit unbekanntem Koeffizientenvektor $\left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right)$ auf:

$$5 \sin + (-3) \cos + 7 \tan = x_1 (\tan - \cos + \sin) + x_2 (\tan + 3 \sin) + x_3 (\cos + \sin).$$

Ein Koeffizientenvergleich ergibt das lineare Gleichungssystem

$$\begin{array}{r} \text{Vergleich der sin-Terme}\to \\ \text{Vergleich der cos-Terme}\to \\ \text{Vergleich der tan-Terme}\to \end{array} \begin{bmatrix} 1 & 3 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ 7 \end{pmatrix}$$

[37] Hier zeigt sich der Grund, warum wir den Standardvektorraum K^n, der in Beispiel 11.3 eingeführt wurde, auch als **Koordinatenraum** bezeichnen. Statt **Koordinaten** kann man auch von den **Komponenten** (englisch: components) bzgl. der Basis B sprechen.

mit der eindeutigen Lösung $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 13 \\ -6 \\ 10 \end{pmatrix}$. (**Quizfrage 19.1:** Warum muss sich hier zwingend eine eindeutige Lösung ergeben?)

Wir führen die Probe durch:

$$\begin{aligned} &13\,(\tan - \cos + \sin) + (-6)\,(\tan + 3\sin) + 10\,(\cos + \sin) \\ &\quad = (13-6)\tan + (-13+10)\cos + (13-18+10)\sin \\ &\quad = 7\tan - 3\cos + 5\sin. \end{aligned}$$ △

Beachte: Zur Bestimmung des Koordinatenvektors $\Phi_B^{-1}(v)$ eines Vektors $v \in V$ bzgl. einer Basis B muss i. A. ein lineares Gleichungssystem gelöst werden!

§ 19.2 Darstellung linearer Abbildungen durch Matrizen

Aus Satz 17.10 wissen wir, dass eine lineare Abbildung bereits durch die Bilder der Vektoren einer Basis eindeutig festgelegt ist. Die wesentliche Idee bei der Darstellung einer linearen Abbildung $V \to W$ mit Hilfe einer Matrix ist nun,

- Vektoren in V durch ihre Koordinatenvektoren bzgl. einer Basis B_V darzustellen,
- ebenso Vektoren in W durch ihre Koordinatenvektoren bzgl. einer Basis B_W darzustellen
- und nur noch mit den Koordinatenvektoren zu rechnen, für die die lineare Abbildung notwendigerweise die Form von Matrix-Vektor-Produkten hat (Lemma 17.12).

Als Konsequenz können wir jede lineare Abbildung zwischen beliebigen endlich-dimensionalen Vektorräumen immer in der gleichen, einfachen und konkreten Form von Matrix-Vektor-Produkten auf der Ebene von Koordinatenvektoren darstellen.

Satz 19.3 (Darstellungssatz für lineare Abbildungen)**.**

Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W. Dann gibt es zu jeder linearen Abbildung $f\colon V \to W$ eine eindeutig definierte Matrix $A \in K^{n\times m}$ mit der Eigenschaft

$$f(v_j) = \sum_{i=1}^{n} a_{ij}\, w_i \quad \text{für alle } j = 1, \ldots, m. \tag{19.3}$$

Diese Matrix A heißt die **Darstellungsmatrix der linearen Abbildung f bzgl. der Basen B_V und B_W** (englisch: representation matrix), in Symbolen: $A = \mathcal{M}_{B_W \leftarrow B_V}(f)$.

Beweis. f ist durch die Bilder $f(v_j)$ der Basisvektoren v_j, $j = 1, \ldots, m$, nach Satz 17.10 eindeutig festgelegt. Für jeden Vektor $f(v_j) \in W$ ist wiederum sein Koordinatenvektor bzgl. der Basis $B_W = (w_1, \ldots, w_n)$ eindeutig festgelegt. Dieser Koordinatenvektor $\Phi_{B_W}^{-1}(f(v_j))$ bildet aber gerade die j-Spalte von A, die damit eindeutig festgelegt ist. □

Die Darstellung (19.3) definiert das Bild des j-ten Basisvektors v_j im Definitionsraum V als Linearkombination in der Basis des Zielraumes W. Die j-te Spalte $a_{\bullet j}$ der Darstellungsmatrix A enthält die Koordinaten von $f(v_j) \in W$ bzgl. der Basis B_W. Unsere Konvention beim Symbol der Darstellungsmatrix ist $\mathcal{M}_{\text{nach}\leftarrow\text{von}}$.

Beispiel 19.4 (Darstellungsmatrizen von Homomorphismen).

(i) Es seien K ein Körper, $V = K^m$ und $W = K^n$ und die lineare Abbildung $f_A\colon K^m \to K^n$ durch Matrix-Vektor-Multiplikation mit einer Matrix A (Lemma 17.12) gegeben, also $f_A(x) = A\,x$. Wählen wir als Standardbasen $B_V = (e_1, \dots, e_m)$ und $B_W = (e_1, \dots, e_n)$ in K^m und K^n, dann ist A selbst die Darstellungsmatrix $\mathcal{M}_{B_W \leftarrow B_V}(f_A)$.

Inbesondere hat beispielsweise die Drehabbildung aus Beispiel 17.11 als Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ bzgl. der Standardbasis (e_1, e_2) (in beiden Räumen) die Darstellungsmatrix

$$\begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix},$$

siehe (17.6).

(ii) Im Vektorraum $V = K^{[\![1,5]\!]}$ der endlichen Folgen (hier der Länge 5) über einem Körper K ist die **Shift-Abbildung** (englisch: shift map)

$$f\colon V \ni (y_1, y_2, y_3, y_4, y_5) \mapsto (0, y_1, y_2, y_3, y_4) \in W = V$$

definiert durch Einfügen einer Null am Anfang der Folge. Die **zyklische Shift-Abbildung** (englisch: cyclic shift map) ist definiert durch

$$g\colon V \ni (y_1, y_2, y_3, y_4, y_5) \mapsto (y_5, y_1, y_2, y_3, y_4) \in W = V.$$

Bzgl. der Standardbasen $B_V = B_W = \big(1, 0, 0, 0, 0, \dots, (0, 0, 0, 0, 1)\big)$ haben diese Endomorphismen die folgende Darstellung:

$$\mathcal{M}_{B_W \leftarrow B_V}(f) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{bzw.} \quad \mathcal{M}_{B_W \leftarrow B_V}(g) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

(**Quizfrage 19.2:** Können Sie die Form der Darstellungsmatrizen erklären?)

(iii) Auf dem Vektorraum $\mathbb{R}^{\mathbb{R}}$ der Funktionen $\mathbb{R} \to \mathbb{R}$ betrachten wir den vierdimensionalen Unterraum V der Funktionen, der von der Basis $B_V = (t \mapsto 1,\ t \mapsto t,\ t \mapsto t^2,\ t \mapsto t^3)$ aufgespannt wird. (**Quizfrage 19.3:** Warum ist das eine linear unabängige Familie?) Jede Funktionsauswertung eines Elements von V an einem festen Punkt ist eine lineare Abbildung $V \to \mathbb{R}$. (**Quizfrage 19.4:** Klar?) Wir betrachten diejenige lineare Abbildung $f\colon V \to W = \mathbb{R}^3$, die durch die drei Punktauswertungen an den Stellen -2, 0 und 2 gegeben ist, und wählen im Zielraum $W = \mathbb{R}^3$ die Standardbasis $B_W = (e_1, e_2, e_3)$. Dann hat f die Darstellungsmatrix

$$\mathcal{M}_{B_W \leftarrow B_V}(f) = \begin{array}{cl} \begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{bmatrix} & \begin{array}{l} \leftarrow\text{Auswertung bei } t = -2 \\ \leftarrow\text{Auswertung bei } t = 0 \\ \leftarrow\text{Auswertung bei } t = 2. \end{array} \end{array}$$

(*iv*) Wir betrachten den zweidimensionalen Vektorraum $V = \mathbb{C}$ über dem Körper $\mathbb{R}$ mit der Basis $B_V = (1, i)$ sowie den eindimensionalen Vektorraum $W = \mathbb{R}$ über dem Körper $\mathbb{R}$ mit der Standardbasis $B_W = (1)$. Die lineare Abbildung $\operatorname{Re}\colon \mathbb{C} \to \mathbb{R}$ (Realteil) hat dann die Darstellungsmatrix

$$\mathcal{M}_{B_W \leftarrow B_V}(\operatorname{Re}) = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

während die lineare Abbildung $\operatorname{Im}\colon \mathbb{C} \to \mathbb{R}$ (Imaginärteil) die Darstellungsmatrix

$$\mathcal{M}_{B_W \leftarrow B_V}(\operatorname{Im}) = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{19.4}$$

besitzt.

(*v*) Im Fall $V = W$ und für eine beliebige Wahl $B_V = B_W$ der Basis hat die Identitätsabbildung $\operatorname{id}_V\colon V \to V$ die Darstellungsmatrix

$$\mathcal{M}_{B_V \leftarrow B_V}(\operatorname{id}_V) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

also die Einheitsmatrix der passenden Dimension. △

Das folgende Resultat zeigt, dass durch eine geschickte Wahl der Basen B_V und B_W eine möglichst einfache Form der Darstellungsmatrix einer linearen Abbildung $V \to W$ erreicht werden kann.

Lemma 19.5 (möglichst einfache Darstellungsmatrizen für lineare Abbildungen)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter sei $f\colon V \to W$ ein Homomorphismus mit $r = \operatorname{Rang}(f)$ und $\operatorname{Defekt}(f) = \dim(\operatorname{Kern}(f)) = m - r$ nach Dimensionsformel (18.5). Wir können Basen B_V, B_W so wählen, dass gilt

$$B_W = (\underbrace{w_1, \ldots, w_r}_{\text{Basis von } \operatorname{Bild}(f)}, \underbrace{w_{r+1}, \ldots, w_n}_{\text{Ergänzung zu einer Basis von } W})$$

und

$$B_V = (\underbrace{v_1, \ldots, v_r}_{v_j \in f^{-1}(\{w_j\})}, \underbrace{v_{r+1}, \ldots, v_m}_{\text{Basis von } \operatorname{Kern}(f)}).$$

Die Darstellungsmatrix $\mathcal{M}_{B_W \leftarrow B_V}(f)$ von f bzgl. dieser Basen hat die Gestalt

$$\mathcal{M}_{B_W \leftarrow B_V}(f) = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \in K^{n \times m}. \tag{19.5}$$

Beweis. Der Beweis ist Gegenstand der Übung. □

Wir gehen nun der Frage nach, welche Eigenschaften die Zuordnung eines Homomorphismus zu seiner Darstellungsmatrix bzgl. fest gewählter Basen hat:

Satz 19.6 (die Zuordnung zur Darstellungsmatrix ist ein Vektorraumisomorphismus).
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \dots, v_m)$ und $B_W = (w_1, \dots, w_n)$ Basen von V bzw. von W. Die Zuordnung

$$\mathcal{M}_{B_W \leftarrow B_V} \colon \operatorname{Homo}(V, W) \ni f \mapsto \mathcal{M}_{B_W \leftarrow B_V}(f) \in K^{n \times m} \tag{19.6}$$

eines Homomorphismus zu seiner Darstellungsmatrix ist ein Isomorphismus von Vektorräumen.

Beachte: Dieses Resultat besagt, dass — wenn es um Vektorraumeigenschaften geht — unerheblich ist, ob wir Homomorphismen oder deren Darstellungsmatrizen betrachten.

Beweis. Wir zerlegen den Beweis in mehrere Schritte:

Schritt 1: Wir zeigen zunächst die Linearität von $\mathcal{M}_{B_W \leftarrow B_V}$.

Es seien dazu $f, g \in \operatorname{Homo}(V, W)$, $A \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f)$ und $B \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(g)$. Dann gilt

$$\begin{aligned}(f + g)(v_j) &= f(v_j) + g(v_j) \\ &= \sum_{i=1}^n a_{ij}\, w_i + \sum_{i=1}^n b_{ij}\, w_i \\ &= \sum_{i=1}^n (a_{ij} + b_{ij})\, w_i\end{aligned}$$

für alle $j = 1, \dots, m$. Das heißt $\mathcal{M}_{B_W \leftarrow B_V}(f + g) = \mathcal{M}_{B_W \leftarrow B_V}(f) + \mathcal{M}_{B_W \leftarrow B_V}(g)$.
Weiter gilt für $\alpha \in K$

$$\begin{aligned}(\alpha f)(v_j) &= \alpha f(v_j) \\ &= \alpha \sum_{i=1}^n a_{ij}\, w_i \\ &= \sum_{i=1}^n (\alpha\, a_{ij})\, w_i\end{aligned}$$

für alle $j = 1, \dots, m$. Das heißt $\mathcal{M}_{B_W \leftarrow B_V}(\alpha f) = \alpha\, \mathcal{M}_{B_W \leftarrow B_V}(f)$.

Schritt 2: Wir zeigen: $\mathcal{M}_{B_W \leftarrow B_V}$ ist injektiv.

Es sei dazu $f \in \operatorname{Homo}(V, W)$ so, dass $\mathcal{M}_{B_W \leftarrow B_V}(f) = 0 \in K^{n \times m}$ (die Nullmatrix) ergibt. Das heißt,

$$f(v_j) = \sum_{i=1}^n 0\, w_i = 0$$

für alle $j = 1, \dots, m$. Damit ist $f \colon V \to W$ der Nullhomomorphismus, also der Nullvektor von $\operatorname{Homo}(V, W)$. Daher gilt $\operatorname{Kern}(\mathcal{M}_{B_W \leftarrow B_V}) = \{0\}$, und nach Lemma 17.9 ist $\mathcal{M}_{B_W \leftarrow B_V}$ injektiv.

Schritt 3: Wir zeigen: $\mathcal{M}_{B_W \leftarrow B_V}$ ist surjektiv.

Es sei dazu $A \in K^{n \times m}$. Nach Satz 17.10 gibt es (genau) einen Homomorphismus $f \colon V \to W$, der $f(v_j) = \sum_{i=1}^n a_{ij}\, w_i$ als Bilder und damit A als Darstellungsmatrix hat. □

Quizfrage 19.5: Warum haben wir, um die Bijektivität von $\mathcal{M}_{B_W \leftarrow B_V}$ zu zeigen, nicht die Charakterisierung der Bijektivität von Homomorphismen (bijektiv $\Leftrightarrow$ injektiv $\Leftrightarrow$ surjektiv) nach Folgerung 18.9 genutzt?

Folgerung 19.7 (Dimension des Vektorraumes der Homomorphismen)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Dann gilt

$$\dim(\operatorname{Homo}(V, W)) = n\, m. \tag{19.7}$$

Beweis. Das Resultat folgt aus $\dim(K^{n\times m}) = n\, m$ (Satz 15.3) und der Isomorphie $\operatorname{Homo}(V, W) \cong K^{n\times m}$, aufgrund derer beide Räume $K^{n\times m}$ und $\operatorname{Homo}(V, W)$ dieselbe Dimension haben (Satz 18.1). □

Die Darstellungsmatrix $A \in K^{n\times m}$ einer linearen Abbildung $f\colon V \to W$ bzgl. fest gewählter Basen B_V und B_W induziert eine lineare Abbildung $f_A\colon K^m \to K^n$ (Lemma 17.12) auf Koeffizientenvektoren. Das folgende Resultat stellt den Zusammenhang zwischen diesen beiden linearen Abbildungen her.

Satz 19.8 (Zusammenhang zwischen einem Homomorphismus und dem durch seine Darstellungsmatrix induzierten Homomorphismus)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W und $f\colon V \to W$ ein Homomorphismus. Dann gilt für die Darstellungsmatrix $A \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f)$ und die durch A induzierte lineare Abbildung f_A der Zusammenhang

$$f_A = \underbrace{\Phi_{B_W}^{-1}}_{\text{Koordinaten} \mapsfrom \text{Vektor}} \circ f \circ \underbrace{\Phi_{B_V}}_{\text{Vektor} \mapsfrom \text{Koordinaten}} \colon K^m \to K^n \tag{19.8a}$$

$$f = \underbrace{\Phi_{B_W}}_{\text{Vektor} \mapsfrom \text{Koordinaten}} \circ f_A \circ \underbrace{\Phi_{B_V}^{-1}}_{\text{Koordinaten} \mapsfrom \text{Vektor}} \colon V \to W. \tag{19.8b}$$

Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} W & V \arrow[l, "f"'] \arrow[d, "\Phi_{B_V}^{-1}"] \\ K^n \arrow[u, "\Phi_{B_W}"] & K^m \arrow[l, "f_A"] \end{tikzcd}$$

Beachte: (19.8a) besagt, dass Matrix-Vektor-Produkte $f_A(x) = A\, x$ wie folgt wirken: Der Koordinatenvektor $x \in K^m$ wird durch Linearkombinationen der Basisvektoren in B_V zum Vektor $\Phi_{B_V}(x) \in V$ synthetisiert, auf diesen wirkt dann die Abbildung f, und schließlich wird das Ergebnis durch $\Phi_{B_W}^{-1}$ als Koordinatenvektor bzgl. der Basis B_W angegeben. (**Quizfrage 19.6:** Wie lässt sich (19.8b) interpretieren?)

Beweis. Wir zeigen, dass die Abbildungen $\Phi_{B_W} \circ f_A \in \operatorname{Homo}(K^m, W)$ und $f \circ \Phi_{B_V} \in \operatorname{Homo}(K^m, W)$ übereinstimmen. Dazu reicht es nach Satz 17.10 aus, zu zeigen, dass ihre Bilder auf einer Basis gleich sind. Wir wählen dazu die Standardbasis $(e_1, \ldots, e_m)$ von K^m. Es gilt einerseits

$$\begin{aligned}
(\Phi_{B_W} \circ f_A)(e_j) &= \Phi_{B_W}(f_A(e_j)) && \text{nach Definition 6.14 der Komposition } \circ \\
&= \Phi_{B_W}(A\, e_j) && \text{nach Definition der von } A \text{ induzierten Abbildung } f_A \\
&= \Phi_{B_W}(a_{\bullet j}) && \text{nach Definition des Matrix-Vektor-Produkts} \\
&= \sum_{i=1}^{n} a_{ij}\, w_i && \text{nach Definition (19.1) von } \Phi_{B_W}
\end{aligned}$$

und andererseits

$$\begin{aligned}
(f \circ \Phi_{B_V})(e_j) &= f(\Phi_{B_V}(e_j)) && \text{nach Definition 6.14 der Komposition } \circ \\
&= f(v_j) && \text{nach Definition (19.1) von } \Phi_{B_V} \\
&= \sum_{i=1}^{n} a_{ij}\, w_i && \text{nach Definition (19.3) der Darstellungsmatrix } A.
\end{aligned}$$

Aus $\Phi_{B_W} \circ f_A \in \operatorname{Homo}(K^m, W) = f \circ \Phi_{B_V}$ können wir nun (19.8a) und (19.8b) leicht durch Auflösen herleiten, weil Φ_{B_V} und Φ_{B_W} bijektiv sind. □

Wir zeigen abschließend in diesem Abschnitt nun noch, dass die Komposition linearer Abbildungen durch das Matrix-Matrix-Produkt ihrer Darstellungsmatrizen dargestellt wird.

Satz 19.9 (Darstellungsmatrix der Komposition von Homomorphismen)**.**
Es seien U, V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(U) = k \in \mathbb{N}_0$, $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_U = (u_1, \ldots, u_k)$, $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von U bzw. von V bzw. von W und $g\colon U \to V$ sowie $f\colon V \to W$ Homomorphismen. Dann gilt für die Darstellungsmatrizen der Zusammenhang

$$\mathcal{M}_{B_W \leftarrow B_U}(f \circ g) = \mathcal{M}_{B_W \leftarrow B_V}(f)\, \mathcal{M}_{B_V \leftarrow B_U}(g). \tag{19.9}$$

Beachte: Im mittleren Raum V muss für die Darstellung der ankommenden Abbildung g und für die Darstellung der ausgehenden Abbildung f dieselbe Basis B_V verwendet werden.

Beweis. Wir führen den Beweis mit Hilfe eines Diagrammes:

$$\begin{tikzcd}
W \arrow[ddd, "\Phi_{B_W}^{-1}"'] & & U \arrow[ll, "f \circ g"'] \arrow[dl, "g"] \\
& V \arrow[ul, "f"] \arrow[d, shift left, "\Phi_{B_V}^{-1}"] & \\
& K^m \arrow[u, shift left, "\Phi_{B_V}"] \arrow[dl, "f_A"'] & \\
K^n & & K^k \arrow[ll, "f_{AB}"] \arrow[ul, "f_B"'] \arrow[uuu, "\Phi_{B_U}"']
\end{tikzcd}$$

Zur Abkürzung setzen wir $A \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f)$ und $B \coloneqq \mathcal{M}_{B_V \leftarrow B_U}(g)$. Das rechte und das linke Trapez sind kommutativ nach Satz 19.8, d. h., es gilt

$$f_A = \Phi_{B_W}^{-1} \circ f \circ \Phi_{B_V} \quad \text{und} \quad f_B = \Phi_{B_V}^{-1} \circ g \circ \Phi_{B_U}.$$

Das obere Dreieck ist kommutativ aufgrund der Definition von $f \circ g$. Das untere Dreieck ist kommutativ wegen $f_A \circ f_B = f_{A\,B}$, siehe Lemma 17.12. Damit folgt

$$\begin{aligned} f_{A\,B} &= f_A \circ f_B \\ &= \Phi_{B_W}^{-1} \circ f \circ \Phi_{B_V} \circ \Phi_{B_V}^{-1} \circ g \circ \Phi_{B_U} \\ &= \Phi_{B_W}^{-1} \circ f \circ g \circ \Phi_{B_U} \\ &= \Phi_{B_W}^{-1} \circ (f \circ g) \circ \Phi_{B_U}. \end{aligned}$$

Das heißt aber, $A\,B$ ist die Darstellungsmatrix $\mathcal{M}_{B_W \leftarrow B_U}(f \circ g)$, was zu beweisen war. □

Ende der Vorlesung 26

§ 19.3 Eigenschaften linearer Abbildungen und ihrer Darstellungsmatrizen

Der isomorphe Zusammenhang zwischen linearen Abbildungen und ihren Darstellungsmatrizen (bzgl. beliebiger, aber fester Basen) erlaubt es, Eigenschaften linearer Abbildungen an ihren Darstellungsmatrizen abzulesen und umgekehrt. Um die bisher für Matrizen bzw. lineare Abbildungen schon bekannten Begriffe einmal zu rekapitulieren und fehlendes Vokabular zu ergänzen, geben wir folgende Tabelle an. Dabei ist K ein Körper, und V und W sind endlich-dimensionale Vektorräume über K.

Matrix $A \in K^{n \times m}$		lineare Abbildung $f\colon V \to W$	
Begriff	siehe	Begriff	siehe
$\text{Bild}(A) \coloneqq \text{SR}(A)$ $= \{A\,x \in K^n \mid x \in K^m\}$	(15.15a)	$\text{Bild}(f) = \{f(u) \in W \mid u \in V\}$ $= f(V)$	(17.4)
$\text{SRang}(A) = \dim(\text{SR}(A))$	(15.15b)	$\text{Rang}(f) = \dim(\text{Bild}(f))$	(18.6)
A ist surjektiv: $\text{Bild}(A) = K^n$		f ist surjektiv: $\text{Bild}(f) = W$	
$\text{Kern}(A) \coloneqq \{x \in K^m \mid A\,x = 0\}$ $= f_A^{-1}(\{0\})$		$\text{Kern}(f) = \{u \in V \mid f(u) = 0\}$ $= f^{-1}(\{0\})$	(17.5)
$\text{Defekt}(A) \coloneqq \dim(\text{Kern}(A))$		$\text{Defekt}(f) = \dim(\text{Kern}(f))$	(18.7)
A ist injektiv: $\text{Kern}(A) = \{0\}$		f ist injektiv: $\text{Kern}(f) = \{0\}$	Lemma 17.9

Beachte: Die Eigenschaften, die man der Matrix A zuspricht, sind gleichzeitig auch die Eigenschaften der von der Matrix induzierten linearen Abbildung f_A, also z. B. $\text{Bild}(A) = \text{Bild}(f_A)$ und $\text{Kern}(A) = \text{Kern}(f_A)$.

Wir können nun bestätigen, wie die oben genannten Begriffe für lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen und ihren Darstellungsmatrizen zusammenhängen:

Satz 19.10 (Eigenschaften linearer Abbildungen und ihrer Darstellungsmatrizen)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W und $f \colon V \to W$ ein Homomorphismus. Schließlich sei $A \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f) \in K^{n \times m}$ die Darstellungsmatrix von f bzgl. dieser Basen. Dann gilt:

(i) $\operatorname{Bild}(f) = \Phi_{B_W}\big(\operatorname{Bild}(A)\big)$.

(ii) $\operatorname{Rang}(f) = \operatorname{SRang}(A) = \operatorname{Rang}(A)$.
Insbesondere ist f surjektiv genau dann, wenn die Zeilen von A linear unabhängig sind.

(iii) $\operatorname{Kern}(f) = \Phi_{B_V}\big(\operatorname{Kern}(A)\big)$.

(iv) $\operatorname{Defekt}(f) = \operatorname{Defekt}(A)$.
Insbesondere ist f injektiv genau dann, wenn die Spalten von A linear unabhängig sind.

Beweis. Aussage (i): Es gilt

$$\begin{aligned}
& w \in \operatorname{Bild}(f) \\
\Leftrightarrow\quad & w \in \operatorname{Bild}(\Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}) && \text{wegen } f = \Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}\text{, siehe (19.8b)} \\
\Leftrightarrow\quad & w \in \Phi_{B_W}\big(\operatorname{Bild}(f_A \circ \Phi_{B_V}^{-1})\big) && \text{nach Definition von Bild} \\
\Leftrightarrow\quad & w \in \Phi_{B_W}\big(\operatorname{Bild}(f_A)\big) && \text{wegen der Bijektivität von } \Phi_{B_V}^{-1} \\
\Leftrightarrow\quad & w \in \Phi_{B_W}\big(\operatorname{Bild}(A)\big) && \text{wegen } f_A(x) = A\,x.
\end{aligned}$$

Aussage (ii): $\operatorname{Bild}(f)$ und $\operatorname{Bild}(A)$ sind nach Aussage (i) zueinander isomorphe Unterräume von W bzw. von K^n. Nach Satz 18.1 haben sie also dieselbe Dimension.

Weiter gilt

$$\begin{aligned}
& f \text{ ist surjektiv} \\
\Leftrightarrow\quad & \operatorname{Bild}(f) = W && \text{nach Definition der Surjektivität} \\
\Leftrightarrow\quad & \dim(\operatorname{Bild}(f)) = n && \text{nach Folgerung 13.17} \\
\Leftrightarrow\quad & \operatorname{Rang}(f) = n && \text{nach Definition 18.8 von } \operatorname{Rang}(f) \\
\Leftrightarrow\quad & \operatorname{Rang}(A) = n && \text{wie gerade gezeigt} \\
\Leftrightarrow\quad & \text{die Zeilen von } A \text{ sind linear unabhängig} && \text{nach Lemma 15.11 und Satz 15.13.}
\end{aligned}$$

Aussage (iii): Es gilt

$$\begin{aligned}
& v \in \operatorname{Kern}(f) \\
\Leftrightarrow\quad & v \in \operatorname{Kern}(\Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}) && \text{wegen } f = \Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}\text{, siehe (19.8b)} \\
\Leftrightarrow\quad & v \in \Phi_{B_V}\big(\operatorname{Kern}(\Phi_{B_W} \circ f_A)\big) && \text{wegen der Bijektivität von } \Phi_{B_V}^{-1} \\
\Leftrightarrow\quad & v \in \Phi_{B_V}\big(\operatorname{Kern}(f_A)\big) && \text{wegen der Bijektivität von } \Phi_{B_W} \\
\Leftrightarrow\quad & v \in \Phi_{B_V}\big(\operatorname{Kern}(A)\big) && \text{wegen } f_A(x) = A\,x.
\end{aligned}$$

Aussage (iv): $\operatorname{Kern}(f)$ und $\operatorname{Kern}(A)$ sind nach Aussage (iii) zueinander isomorphe Unterräume von V bzw. von K^m. Nach Satz 18.1 haben sie also dieselbe Dimension.

Weiter gilt

$$\begin{aligned}
& f \text{ ist injektiv} && \\
\Leftrightarrow\quad & \operatorname{Kern}(f) = \{0\} && \text{nach Lemma 17.9} \\
\Leftrightarrow\quad & \dim(\operatorname{Kern}(f)) = 0 && \text{nach Beispiel 13.16} \\
\Leftrightarrow\quad & \operatorname{Defekt}(f) = 0 && \text{nach Definition 18.8 von } \operatorname{Defekt}(f) \\
\Leftrightarrow\quad & \operatorname{Rang}(f) = m && \text{nach Dimensionsformel (18.8)} \\
\Leftrightarrow\quad & \operatorname{SRang}(A) = \operatorname{Rang}(A) = m && \text{nach Aussage } (ii) \\
\Leftrightarrow\quad & \text{die Spalten von } A \text{ sind linear unabhängig} && \text{nach Lemma 15.11.}
\end{aligned}$$

□

Satz 19.10 erlaubt es uns also, das Bild und den Kern einer beliebigen linearen Abbildung f zwischen beliebigen endlich-dimensionalen Vektorräumen auf eine standardisierte Weise zu berechnen. Dazu wählen wir irgendwelche Basen B_V und B_W und bestimmen die Darstellungsmatrix $A = \mathcal{M}_{B_W \leftarrow B_V}(f)$. Anschließend berechnen wir eine Basis des Unterraumes $\operatorname{Bild}(A) \subseteq \mathbb{R}^n$ bzw. des Unterraumes $\operatorname{Kern}(A) \subseteq \mathbb{R}^m$ (dazu gleich mehr). Um daraus dann eine Basis von $\operatorname{Bild}(f) \subseteq W$ zu erhalten, müssen wir laut Satz 19.10 lediglich die Basisvektoren von $\operatorname{Bild}(A)$ als Koordinatenvektoren bzgl. der Basis B_W interpretieren. Analog: Um eine Basis von $\operatorname{Kern}(f) \subseteq V$ zu erhalten, interpretieren wir die Basisvektoren von $\operatorname{Kern}(A)$ als Koordinatenvektoren bzgl. der Basis B_V. (**Quizfrage 19.7:** Warum können wir sicher sein, wieder eine Basis zu erhalten?)

Wie können wir nun für eine gegebene Matrix $A \in K^{n \times m}$ eine Basis von $\operatorname{Bild}(A) \subseteq K^n$ bzw. eine Basis von $\operatorname{Kern}(A) \subseteq K^m$ berechnen?

Bemerkung 19.11 (zur Bestimmung von Bild und Kern einer Matrix)**.**

(i) Für $\operatorname{Bild}(A)$ haben wir folgende Möglichkeiten:

(1) Wir bestimmen mit Hilfe von Algorithmus D.2, wie in Bemerkung 15.25 beschrieben, eine Rangfaktorisierung $A = B_{\text{Rang}}\, C_{\text{Rang}}$ mit $B_{\text{Rang}} \in K^{n \times r}$ und $C_{\text{Rang}} \in K^{r \times m}$ (in Zeilenstufenform) und $r = \operatorname{Rang}(A)$. Dann bilden die Spalten von B_{Rang} eine Basis von $\operatorname{Bild}(A) = \operatorname{SR}(A)$.

(2) Wir nutzen die Beziehung $\operatorname{Bild}(A) = \operatorname{SR}(A) = [\operatorname{ZR}(A^\intercal)]^\intercal$, bringen $A^\intercal$ in Zeilenstufenform (Algorithmus D.1) und erhalten somit eine Basis von $\operatorname{ZR}(A^\intercal)$. Durch Transposition der Basisvektoren von $\operatorname{ZR}(A^\intercal)$ bekommen wir die gesuchte Basis von $\operatorname{Bild}(A)$.

Für die Rechnung per Hand erscheint die zweite Möglichkeit günstiger, weil wir den linken Faktor „B" nicht mitführen müssen.

(ii) Die Bestimmung von $\operatorname{Kern}(A) = \mathcal{L}(A, 0)$ haben wir in § 16 bei der Lösung linearer Gleichungssysteme bereits durchgeführt. Zur Erinnerung: Wir bringen die erweiterte Koeffizientenmatrix $[A, 0]$ zunächst in Zeilenstufenform. Daran können wir bereits $\operatorname{Rang}(A)$ und damit auch $\dim(\operatorname{Kern}(A)) = m - \operatorname{Rang}(A)$ ablesen. Letzteres ist die Anzahl der Nicht-Pivot-Spalten (unabhängige Variablen) in der Zeilenstufenform.

Wenn $\dim(\operatorname{Kern}(A)) = 0$ gilt, dann ist $\operatorname{Kern}(A) = \{0\}$ und die einzige Basis von $\operatorname{Kern}(A)$ ist die leere Menge. Wenn $\dim(\operatorname{Kern}(A)) > 0$ gilt, dann überführen wir das System weiter in die reduzierte Zeilenstufenform. Anschließend können wir nacheinander Basisvektoren von $\operatorname{Kern}(A)$ bestimmen, indem wir eine der unabhängigen Variablen x_i auf den Wert 1 und die anderen unabhängigen Variablen auf den Wert 0 setzen und die Werte der abhängigen Variablen von hinten nach vorne aus den Gleichungen ausrechnen (Bemerkung 16.4 und Beispiel 16.10). △

Beispiel 19.12 (zur Bestimmung von Bild und Kern einer Matrix)**.**
Wir betrachten wie in Beispiel 19.4 den vierdimensionalen Unterraum V der Funktionen $\mathbb{R} \to \mathbb{R}$, der von der Basis $B_V = (t \mapsto 1,\ t \mapsto t,\ t \mapsto t^2,\ t \mapsto t^3)$ aufgespannt wird, und darauf die lineare Abbildung $f \colon V \to W = \mathbb{R}^3$, die durch die drei Punktauswertungen an den Stellen -2, 0 und 2 gegeben ist. Bezüglich der Standardbasis $B_W = (e_1, e_2, e_3)$ hat f die Darstellungsmatrix

$$A = \begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{bmatrix}.$$

Eine Zeilenstufenform von $A^\intercal$ ist gegeben durch

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{bmatrix}.$$

Es gilt also $\operatorname{Rang}(A^\intercal) = \operatorname{Rang}(A) = 3$. Die ersten drei Zeilen bilden eine Basis von $\operatorname{ZR}(A^\intercal)$. Ihre Transponierten bilden also eine Basis von $\operatorname{SR}(A) = \operatorname{Bild}(A)$, d. h., wir haben

$$\operatorname{Bild}(A) = \left\langle \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\2\\4\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\0\\8\end{smallmatrix}\right) \right\rangle = \left\langle \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\\2\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right) \right\rangle.$$

Da $\operatorname{Bild}(A)$ aber maximale Dimension in $\mathbb{R}^3$ hat, also $\operatorname{Bild}(A) = \mathbb{R}^3$ gilt, können wir auch jede andere Basis von $\mathbb{R}^3$ (z. B. die Standardbasis) als Basis von $\operatorname{Bild}(A)$ verwenden.

Um $\operatorname{Kern}(A)$ zu bestimmen, bringen wir $[A, 0]$ zunächst in Zeilenstufenform

$$\left[\begin{array}{cccc|c} 1 & -2 & 4 & -8 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 & 0 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{cccc|c} 1 & -2 & 4 & -8 & 0 \\ 0 & 2 & -4 & 8 & 0 \\ 0 & 0 & 8 & 0 & 0 \end{array}\right],$$

wo wir wiederum $\operatorname{Rang}(A) = 3$ und damit $\operatorname{Defekt}(A) = \dim(\operatorname{Kern}(A)) = 4 - 3 = 1$ ablesen können. Wir gehen weiter zu reduzierten Zeilenstufenform

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right],$$

an der wir ablesen:[38]

$$\operatorname{Kern}(A) = \left\langle \left(\begin{smallmatrix}0\\-4\\0\\1\end{smallmatrix}\right) \right\rangle$$

[38]Noch geschickter wäre folgendes Vorgehen gewesen: Wir bestimmen eine (reduzierte) Zeilenstufenform von A, um damit $\operatorname{Rang}(A)$, eine Basis von $\operatorname{Kern}(A)$ und $\dim(\operatorname{Kern}(A))$ zu bestimmen. Falls wie hier im Beispiel $\operatorname{Rang}(A) = n$ gilt, können wir uns die Berechnung einer Basis von $\operatorname{Bild}(A)$ sparen, da $\operatorname{Bild}(A)$ den ganzen K^n ausfüllt.

Übersetzen wir $\operatorname{Kern}(A)$ zurück in den Vektorraum V, so bedeutet das nach Satz 19.10 Aussage (iii), dass wir $\begin{pmatrix} 0 \\ -4 \\ 0 \\ 1 \end{pmatrix}$ als Koordinatenvektor bzgl. der gewählten Basis $B_V = (t \mapsto 1,\, t \mapsto t,\, t \mapsto t^2,\, t \mapsto t^3)$ interpretieren müssen. Es sind also genau die Vielfachen der Funktion

$$t \mapsto p(t) \coloneqq 0\,1 - 4\,t + 0\,t^2 + 1\,t^3 = -4\,t + t^3,$$

die in $\operatorname{Kern}(f)$ liegen, die also die Eigenschaft besitzen, dass alle drei Punktauswertungen Null ergeben. In der Tat gilt

$$\begin{aligned} p(-2) &= -4 \cdot (-2) + (-2)^3 = 0 \\ p(0) &= -4 \cdot 0 + 0^3 = 0 \\ p(2) &= -4 \cdot 2 + 2^3 = 0. \end{aligned}$$

Wir haben also

$$\operatorname{Kern}(f) = \langle t \mapsto -4\,t + t^3 \rangle \subseteq V.$$

Für $\operatorname{Bild}(A)$ ist in unserem Beispiel keine solche Übersetzung zurück erforderlich, denn wegen der Wahl der Standardbasis in $\mathbb{R}^3$ gilt $\operatorname{Bild}(A) = \operatorname{Bild}(f)$. (**Quizfrage 19.8:** Wie äußert sich das in Aussage (i) von Satz 19.10?) △

Neben den Zusammenhängen in Satz 19.10 können wir auch die Invertierbarkeit einer linearen Abbildung in Verbindung bringen mit der Invertierbarkeit ihrer Darstellungsmatrix. Wie wir aus Folgerung 18.9 wissen, müssen dazu notwendigerweise beide endlich-dimensionalen Vektorräume dieselbe Dimension besitzen (und gleichbedeutend damit die Darstellungsmatrix quadratisch sein).

Satz 19.13 (Invertierbarkeit linearer Abbildungen und ihrer Darstellungsmatrizen).
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K gleicher Dimension $\dim(V) = \dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_n)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W und $f\colon V \to W$ ein Homomorphismus. Schließlich sei $A \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f) \in K^{n \times n}$ die Darstellungsmatrix von f bzgl. dieser Basen. Dann sind äquivalent:

(i) f ist bijektiv.

(ii) $\operatorname{Rang}(f) = n$.

(iii) $\operatorname{Defekt}(f) = 0$.

(iv) A ist invertierbar.

(v) $\operatorname{Rang}(A) = n$.

(vi) $\operatorname{Defekt}(A) = 0$.

Ist f bijektiv, dann gilt für die Darstellungsmatrix der Inversen $f^{-1}\colon W \to V$

$$\mathcal{M}_{B_V \leftarrow B_W}(f^{-1}) = A^{-1}. \tag{19.10}$$

Beweis. Die Äquivalenz der Aussagen (i) bis (iii) wurde in Folgerung 18.9 gezeigt. Die Äquivalenz der Aussagen (iv) und (v) wurde in Satz 15.45 gezeigt. Nach Satz 19.10 gilt $\operatorname{Rang}(f) =$

Rang(A), also sind auch Aussagen (ii) und (v) äquivalent. Ebenfalls nach Satz 19.10 gilt Defekt(f) = Defekt(A), also sind auch Aussagen (iii) und (vi) äquivalent.

Ist f bijektiv, dann gilt

$$\begin{aligned} & f^{-1} \circ f = \mathrm{id}_V \\ \Rightarrow\quad & \mathcal{M}_{B_V \leftarrow B_V}(f^{-1} \circ f) = \mathcal{M}_{B_V \leftarrow B_V}(\mathrm{id}_V) \\ \Rightarrow\quad & \mathcal{M}_{B_V \leftarrow B_W}(f^{-1})\, \mathcal{M}_{B_W \leftarrow B_V}(f) = I_n \qquad \text{nach Satz 19.9 und Beispiel 19.4} \\ \Rightarrow\quad & \mathcal{M}_{B_V \leftarrow B_W}(f^{-1}) \qquad A \qquad = I_n. \end{aligned}$$

Das ist nach Satz 15.47 (Rechtsinverse quadratischer Matrizen sind Linksinverse und umgekehrt) bereits ausreichend, um (19.10) zu bestätigen. □

Wir können außerdem eine Variante von Satz 19.13 angeben, die die Links- bzw. Rechtsinvertierbarkeit linearer Abbildungen und ihrer Darstellungsmatrizen betrifft:[39]

Satz 19.14 (Links- bzw. Rechtsinvertierbarkeit linearer Abbildungen und ihrer Darstellungsmatrizen)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W und $f \colon V \to W$ ein Homomorphismus. Schließlich sei $A \coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f) \in K^{n \times m}$ die Darstellungsmatrix von f bzgl. dieser Basen.

(i) Es sind äquivalent:
   (a) f ist surjektiv, also Rang$(f) = n$.
   (b) Es existiert eine Rechtsinverse von f, also eine lineare Abbildung $f_r \colon W \to V$, sodass $f \circ f_r = \mathrm{id}_W$ gilt. f_r ist notwendig injektiv.
   (c) A besitzt vollen Zeilenrang, also Rang$(A) = n$.
   (d) Es existiert eine Rechtsinverse von A, also eine Matrix $R \in K^{m \times n}$ mit $A\,R = I_n$ gilt. R besitzt notwendig vollen Spaltenrang.

(ii) Es sind äquivalent:
   (a) f ist injektiv, also Defekt$(f) = 0$, d. h. Rang$(f) = m$.
   (b) Es existiert eine Linksinverse von f, also eine lineare Abbildung $f_\ell \colon W \to V$, sodass $f_\ell \circ f = \mathrm{id}_V$ gilt. f_ℓ ist notwendig surjektiv. Ihre Einschränkung $f_\ell|_{f(V)}$ auf das Bild von f ist eindeutig.
   (c) A besitzt vollen Spaltenrang, also Rang$(A) = m$.
   (d) Es existiert eine Linksinverse von A, also eine Matrix $L \in K^{m \times n}$ mit $L\,A = I_m$ gilt. L besitzt notwendig vollen Zeilenrang.

Beweis. Der Beweis ist Gegenstand der Übung. □

[39] Aufgrund der vorausgesetzen endlichen Dimension der Vektorräume und der Linearität der Abbildung benötigen wir hier im Unterschied zu Satz 6.44 das Auswahlaxiom nicht.

Aufgrund von Satz 19.14 wird manchmal auch die Sprechweise verwendet, eine **Matrix** A sei **surjektiv** oder **injektiv**. Gemeint ist immer die Aussage, dass jede durch A repräsentierte lineare Abbildung (z. B. $x \mapsto A\,x$) surjektiv bzw. injektiv ist.

§ 19.4 Transformationsmatrizen des Basiswechsels

Die Darstellung eines Vektors v eines endlich-dimensionalen Vektorraumes V mit Hilfe seines Koordinatenvektors $x = \Phi_{B_V}^{-1}(v) \in K^n$ hängt von der Wahl der Basis B_V ab. Ebenso hängt die Beschreibung von Homomorphismen über Darstellungsmatrizen von der Wahl der Basen im Definitions- und Zielraum ab. Es stellen sich daher folgende Fragen:

(1) Wie transformiert sich ein Koordinatenvektor beim Wechsel der Basis?

(2) Wie transformiert sich die Darstellungsmatrix eines Homomorphismus, wenn wir eine oder beide Basen wechseln?

(3) In welchen Basen hat die Darstellungsmatrix eines Homomorphismus eine besonders einfache Gestalt?

Die Frage (3) ist mit Lemma 19.5 bereits beantwortet: Die Darstellung $\left[\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right]$ ist möglich, und einfacher geht es nicht, weil alle Darstellungsmatrizen einer linearen Abbildung nach Satz 19.10 denselben Rang wie die Abbildung haben.

Die Beantwortung der ersten beiden Fragen gelingt im Folgenden mit Hilfe von **Transformationsmatrizen** für den Basiswechsel.

Definition 19.15 (Transformationsmatrix)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ zwei Basen von V.[40] Dann heißt

$$\mathcal{T}_{B_V \leftarrow \widehat{B}_V} \coloneqq \mathcal{M}_{B_V \leftarrow \widehat{B}_V}(\mathrm{id}_V) \in K^{n \times n} \tag{19.11}$$

die **Transformationsmatrix des Basiswechsels**, **Übergangsmatrix** oder **Basiswechselmatrix von $\widehat{B}_V$ nach B_V** (englisch: transformation matrix, transition matrix, change-of-basis matrix). △

Die Transformationsmatrix $\mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ ist also nichts anderes als die Darstellung der Identitätsabbildung bzgl. der „neuen“ Basis $\widehat{B}_V$ im Definitionsraum V und der „alten“ Basis B_V im Zielraum V. Die Einträge t_{ij} der Transformationsmatrix $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ bestimmen sich daher aus den Bedingungen

$$\underbrace{\mathrm{id}_V(\widehat{v}_j) = \widehat{v}_j}_{\text{Bild des Basisvektors } \widehat{v}_j \text{ im Definitionsraum unter der identischen Abbildung}} = \sum_{i=1}^n t_{ij}\, v_i \quad \text{für } j = 1, \ldots, n. \tag{19.12}$$

Die j-te Spalte von $T = \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ enthält also die Koeffizienten für die Darstellung des „neuen“ Basisvektors $\widehat{v}_j$ als Linearkombination der „alten“ Basis B_V.

[40] Wir verwenden die Konvention, dass wir die „alte“ Basis in blau und die „neue“ Basis in rot kennzeichnen. Warum wir in Definition 19.15 die Übergangsmatrix von „neu“ nach „alt“ in den Vordergrund stellen, wird später in Satz 19.20 noch deutlich.

Das folgende Resultat beantwortet nun die Frage (1) zur Transformation von Koordinatenvektoren bei einem Basiswechsel.

Lemma 19.16 (Eigenschaften von Transformationsmatrizen).
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ zwei Basen von V. Dann gilt:

(i) Ist $\widehat{x} \in K^n$ der Koordinatenvektor eines Vektors $v \in V$ bzgl. der Basis $\widehat{B}_V$, dann ist $x = \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \widehat{x}$ der Koordinatenvektor von v bzgl. der Basis B_V.[41]

(ii) Die Transformationsmatrix $\mathcal{T}_{B_V \leftarrow \widehat{B}_V} \in K^{n \times n}$ ist invertierbar.

(iii) $\left(\mathcal{T}_{B_V \leftarrow \widehat{B}_V}\right)^{-1} = \mathcal{T}_{\widehat{B}_V \leftarrow B_V}$.

(iv) Die von der Transformationsmatrix $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ induzierte lineare Abbildung $f_T \colon K^n \to K^n$ ist $f_T = \Phi_{B_V}^{-1} \circ \Phi_{\widehat{B}_V} \in \operatorname{Auto}(K^n)$.

Beweis. Wir setzen zur Abkürzung $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ wie in (19.12).

Aussage (i): Es gilt

$$\begin{aligned} v &= \sum_{j=1}^n \widehat{x}_j \, \widehat{v}_j && \text{nach Voraussetzung} \\ &= \sum_{j=1}^n \widehat{x}_j \sum_{i=1}^n t_{ij} \, v_i && \text{nach (19.12)} \\ &= \sum_{i=1}^n \underbrace{\sum_{j=1}^n (t_{ij} \, \widehat{x}_j)}_{\text{Koeffizient } x_i \text{ bzgl. der Basis } B_V = (v_1, \ldots, v_n)} v_i && \text{wegen Distributivität und Kommutativität im Körper } K. \end{aligned}$$

Nach Definition des Matrix-Vektor-Produkts (Bemerkung 15.9) gilt also $x = \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \widehat{x}$.

Aussage (ii): Die Invertierbarkeit von $\mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ folgt aus Satz 19.13, da $\mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ eine Darstellungsmatrix des bijektiven Homomorphismus $\mathrm{id}_V \colon V \to V$ ist.

Aussage (iii): Es gilt

$$\begin{aligned} \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \mathcal{T}_{\widehat{B}_V \leftarrow B_V} &= \mathcal{M}_{B_V \leftarrow \widehat{B}_V}(\mathrm{id}_V) \mathcal{M}_{\widehat{B}_V \leftarrow B_V}(\mathrm{id}_V) \\ &= \mathcal{M}_{B_V \leftarrow B_V}(\mathrm{id}_V \circ \mathrm{id}_V) && \text{nach Satz 19.9} \\ &= \mathcal{M}_{B_V \leftarrow B_V}(\mathrm{id}_V) \\ &= I_n && \text{nach Beispiel 19.4.} \end{aligned}$$

Aussage (iv): Nach Satz 19.8 und wegen $T = \mathcal{T}_{B_V \leftarrow \widehat{B}_V} = \mathcal{M}_{B_V \leftarrow \widehat{B}_V}(\mathrm{id}_V)$ gilt für die durch T induzierte Abbildung $K^n \to K^n$ die Gleichung

[41] Beachte das Muster $x = \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \widehat{x}$ mit „alter“ und „neuer“ Basis und zugehörigen Koordinatenvektoren.

$$f_T = \Phi_{B_V}^{-1} \circ \mathrm{id}_V \circ \Phi_{\widehat{B}_V} = \underbrace{\Phi_{B_V}^{-1}}_{\text{„alte" Koordinaten} \leftarrow\!\!\mapsto \text{ Vektor in } V} \circ \underbrace{\Phi_{\widehat{B}_V}}_{\text{Vektor in } V \leftarrow\!\!\mapsto \text{ „neue" Koordinaten}}.$$

$$\begin{tikzcd} V \arrow[d, "\Phi_{B_V}^{-1}"'] & V \arrow[l, "\mathrm{id}_V"'] \\ K^n & K^n \arrow[l, "f_T"] \arrow[u, "\Phi_{\widehat{B}_V}"'] \end{tikzcd}$$

□

Beispiel 19.17 (Transformationsmatrix)**.**
Auf dem Vektorraum $\mathbb{R}^\mathbb{R}$ der Funktionen $\mathbb{R} \to \mathbb{R}$ betrachten wir den zweidimensionalen Unterraum V der Funktionen, der von der „alten" Basis $B_V = (t \mapsto e^t,\ t \mapsto e^{-t})$ aufgespannt wird. Als „neue" Basis verwenden wir $\widehat{B}_V = (\sinh, \cosh)$. Es gilt

$$\sinh(t) = \frac{e^t - e^{-t}}{2} \quad \text{und} \quad \cosh(t) = \frac{e^t + e^{-t}}{2} \quad \text{für alle } t \in \mathbb{R}.$$

Die Transformationsmatrix $\mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ ist also gegeben durch

$$\mathcal{T}_{B_V \leftarrow \widehat{B}_V} = \begin{array}{c} \begin{array}{cc} \sinh & \cosh \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array} \begin{array}{c} e^t \\ e^{-t} \end{array} \quad \text{mit inverser Matrix} \quad \mathcal{T}_{\widehat{B}_V \leftarrow B_V} = \begin{array}{c} \begin{array}{cc} e^t & e^{-t} \end{array} \\ \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \end{array} \begin{array}{c} \sinh \\ \cosh \end{array}.$$

Wollen wir beispielsweise die Funktion $7\ \sinh - 3\ \cosh$ mit dem „neuen" Koordinatenvektor $\widehat{x} = \binom{7}{-3}$ bzgl. der „alten" Basis darstellen, dann ist der zugehörige „alte" Koordinatenvektor gegeben durch

$$x = \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \widehat{x} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{pmatrix} 7 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ -5 \end{pmatrix},$$

also gilt $7\ \sinh(t) - 3\ \cosh(t) = 2\,e^t - 5\,e^{-t}$. Wollen wir umgekehrt beispielsweise die Funktion $t \mapsto e^t + 4\,e^{-t}$ mit dem „alten" Koordinatenvektor $x = \binom{1}{4}$ bzgl. der „neuen" Basis darstellen, dann ist der zugehörige „neue" Koordinatenvektor gegeben durch

$$\widehat{x} = \mathcal{T}_{\widehat{B}_V \leftarrow B_V} x = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \end{pmatrix},$$

also gilt $e^t + 4\,e^{-t} = -3\ \sinh(t) + 5\ \cosh(t)$. △

Bemerkung 19.18 (Eigenschaften von Transformationsmatrizen in K^n)**.**
Im Fall $V = K^n$ stimmen Vektoren mit ihren Koordinatenvektoren bzgl. der Standardbasis $(e_1, \ldots, e_n)$ überein. Ist $B_V = (v_1, \ldots, v_n)$ eine Basis von $V = K^n$, so hat die Transformationsmatrix $\mathcal{T}_{(e_1,\ldots,e_n) \leftarrow B_V}$ die einfache Gestalt

$$\mathcal{T}_{(e_1,\ldots,e_n) \leftarrow B_V} = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}.$$

Um nun den Übergang von einer Basis $\widehat{B}_V$ zu einer anderen Basis B_V zu beschreiben, können wir daher wie folgt vorgehen:

$$\mathcal{T}_{B_V \leftarrow \widehat{B}_V} = \mathcal{T}_{B_V \leftarrow (e_1,\ldots,e_n)} \mathcal{T}_{(e_1,\ldots,e_n) \leftarrow \widehat{B}_V}$$

$$= \left(\mathcal{T}_{(e_1,\dots,e_n)\leftarrow B_V}\right)^{-1} \mathcal{T}_{(e_1,\dots,e_n)\leftarrow \widehat{B}_V} = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}^{-1} \begin{bmatrix} | & & | \\ \widehat{v}_1 & \cdots & \widehat{v}_n \\ | & & | \end{bmatrix}. \tag{19.13}$$

△

Beispiel 19.19 (Transformationsmatrizen in K^n)**.**
Wir wollen im Raum $V = \mathbb{Q}^2$ über dem Körper $\mathbb{Q}$ die Transformationsmatrix $\mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ von der Basis $\widehat{B}_V = \left(\binom{2}{-1}, \binom{1}{-1}\right)$ zur Basis $B_V = \left(\binom{-1}{1}, \binom{1}{1}\right)$ finden. Nach (19.13) gilt

$$\mathcal{T}_{B_V \leftarrow \widehat{B}_V} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}.$$

Wir führen noch die Probe durch. Die erste Spalte der Transformationsmatrix sollte die Koeffizienten für die Darstellung des ersten „neuen“ Basisvektors $\widehat{v}_1$ als Linearkombination der „alten“ Basisvektoren enthalten:

$$\frac{-3}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \widehat{v}_1.$$

Analog gilt für den zweiten „neuen“ Basisvektor $\widehat{v}_2$:

$$\frac{-2}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{0}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \widehat{v}_2.$$

△

§ 19.5 Transformation der Darstellungsmatrizen von Homomorphismen

Wir kommen nun zur Frage (2), wie sich die Darstellungsmatrix einer linearen Abbildung bei einem Wechsel einer oder beider Basen transformiert.

Satz 19.20 (Transformation der Darstellungsmatrix eines Homomorphismus beim Wechsel der Basen)**.**
Es seien V, W endlich-dimensionale Vektorräume über demselben Körper K. Weiter seien B_V und $\widehat{B}_V$ Basen von V sowie B_W und $\widehat{B}_W$ Basen von W. Dann gilt für die Darstellungsmatrix eines Homomorphismus $f\colon V \to W$:[42]

$$\mathcal{M}_{\widehat{B}_W \leftarrow \widehat{B}_V}(f) = \mathcal{T}_{\widehat{B}_W \leftarrow B_W}\, \mathcal{M}_{B_W \leftarrow B_V}(f)\, \mathcal{T}_{B_V \leftarrow \widehat{B}_V}. \tag{19.14}$$

Beweis. Wir setzen zur Abkürzung

$$\begin{aligned}
T &\coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V} && \text{Übergang von } \widehat{B}_V \text{ nach } B_V,\\
A &\coloneqq \mathcal{M}_{B_W \leftarrow B_V}(f) && \text{„alte“ Darstellungsmatrix von } f,\\
S &\coloneqq \mathcal{T}_{B_W \leftarrow \widehat{B}_W} && \text{Übergang von } \widehat{B}_W \text{ nach } B_W,\\
\text{also } S^{-1} &= \mathcal{T}_{\widehat{B}_W \leftarrow B_W} && \text{Übergang von } B_W \text{ nach } \widehat{B}_W,\\
\widehat{A} &\coloneqq \mathcal{M}_{\widehat{B}_W \leftarrow \widehat{B}_V}(f) && \text{„neue“ Darstellungsmatrix von } f.
\end{aligned}$$

Wir betrachten das Diagramm

[42]Die Farben deuten wieder den Übergang von den „alten“ Basen zu den „neuen“ Basen an.

$$\begin{tikzcd}
K^n & & & K^m \arrow[lll, "f_{\widehat{A}}"'] \arrow[dl, "\Phi_{\widehat{B}_V}"'] \arrow[dd, "f_T = \Phi^{-1}_{B_V} \circ \Phi_{\widehat{B}_V}"] \\
& W \arrow[ul, "\Phi^{-1}_{\widehat{B}_W}"'] \arrow[dl, shift left, "\Phi^{-1}_{B_W}"] & V \arrow[l, "f"'] \arrow[dr, shift left, "\Phi^{-1}_{B_V}"] & \\
K^n \arrow[uu, "f_S^{-1} = \Phi^{-1}_{\widehat{B}_W} \circ \Phi_{B_W}"] \arrow[ur, shift left, "\Phi_{B_W}"] & & & K^m \arrow[lll, "f_A"] \arrow[ul, shift left, "\Phi_{B_V}"]
\end{tikzcd}$$

Das obere und das untere Trapez sind kommutativ nach Satz 19.8. Das linke Dreieck und das rechte Dreieck sind kommutativ nach Lemma 19.16. Damit kommutiert auch das äußere Rechteck, d. h., es gilt

$$\begin{aligned}
f_{\widehat{A}} &= f_S^{-1} \circ f_A \circ f_T \\
&= f_{S^{-1}} \circ f_A \circ f_T && \text{nach Lemma 17.12} \\
&= f_{S^{-1} A T} && \text{nach Lemma 17.12.}
\end{aligned}$$

Aus Lemma 17.12 folgt dann auch

$$\widehat{A} = S^{-1} A T, \tag{19.15}$$

also die Behauptung (19.14). □

Beispiel 19.21 (Transformation der Darstellungsmatrix eines Homomorphismus beim Wechsel der Basen).
Wir betrachten die Abbildung $f_A \colon \mathbb{R}^2 \to \mathbb{R}^3$, die durch die Matrix

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

induziert ist. Das heißt, A ist gleich der Darstellungsmatrix $\mathcal{M}_{(e_1,e_2,e_3) \leftarrow (e_1,e_2)}(f_A)$ bzgl. der Standardbasen in $\mathbb{R}^2$ und $\mathbb{R}^3$. Wir wollen die Darstellungsmatrix nun in die neuen Basen $\widehat{B}_V = \left(\left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \right)$ und $\widehat{B}_W = \left(\left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right) \right)$ umrechnen, also

$$\widehat{A} = \mathcal{M}_{\widehat{B}_W \leftarrow \widehat{B}_V}(f_A) = S^{-1} A T$$

bestimmen. Es gilt

$$T \coloneqq \mathcal{T}_{(e_1,e_2) \leftarrow \widehat{B}_V} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

und

$$S^{-1} \coloneqq \mathcal{T}_{\widehat{B}_W \leftarrow (e_1,e_2,e_3)} = \left(\mathcal{T}_{(e_1,e_2,e_3) \leftarrow \widehat{B}_W} \right)^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Wir erhalten also die transformierte Darstellungsmatrix bzgl. der neuen Basen gemäß (19.14)

$$\widehat{A} = S^{-1} A T = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 7 \\ 1 & 15 \end{bmatrix}.$$

△

Definition 19.22 (Äquivalenztransformation, äquivalente Matrizen)**.**
Es seien K ein Körper und $n, m \in \mathbb{N}_0$. Zwei Matrizen $A, \widehat{A} \in K^{n \times m}$ heißen **äquivalent** (englisch: equivalent), wenn es invertierbare Matrizen $S \in K^{n \times n}$ und $T \in K^{m \times m}$ gibt, sodass gilt:

$$\widehat{A} = S^{-1} A\, T. \tag{19.16}$$

Der Übergang von A zu $S^{-1} A\, T$, also die Multiplikation von links und von rechts mit invertierbaren Matrizen, heißt auch eine **Äquivalenztransformation** (englisch: equivalence transformation) von A. △

Beachte: Die Äquivalenz von Matrizen ist eine Äquivalenzrelation auf der Menge $K^{n \times m}$. Aufgrund von Folgerung 15.46 (Multiplikation mit invertierbaren Matrizen ändert den Rang nicht) besteht die Äquivalenzklasse von $A \in K^{n \times m}$ genau aus denjenigen $n \times m$-Matrizen, die denselben Rang wie A besitzen:

$$[A] = \left\{ \widehat{A} \in K^{n \times m} \,\middle|\, \operatorname{Rang}(\widehat{A}) = \operatorname{Rang}(A) \right\}. \tag{19.17}$$

In der Äquivalenzklasse $[A]$ einer Matrix $A \in K^{n \times m}$ der ranggleichen Matrizen gibt es einen natürlichen Repräsentanten, der eine besonders einfache Gestalt besitzt und der die **Rang-Normalform** (englisch: rank normal form) genannt wird:

Lemma 19.23 (Rang-Normalform einer Matrix)**.**
Es seien K ein Körper, $n, m \in \mathbb{N}_0$ und $A \in K^{n \times m}$ mit $\operatorname{Rang}(A) = r$. Dann existieren reguläre Matrizen $S \in K^{n \times n}$ und $T \in K^{m \times m}$, sodass gilt:

$$S^{-1} A\, T = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \in K^{n \times m}. \tag{19.18}$$

Beweis. A stimmt überein mit der Darstellungsmatrix der von A induzierten Abbildung $f_A \colon K^m \to K^n$ bzgl. der Standardbasen $B_{K^m} = (e_1, \ldots, e_m)$ und $B_{K^n} = (e_1, \ldots, e_n)$, siehe Beispiel 19.4, also

$$A = \mathcal{M}_{B_{K^n} \leftarrow B_{K^m}}(f_A).$$

Wählen wir jetzt weitere Basen $\widehat{B}_{K^m}$ von K^m und $\widehat{B}_{K^n}$ von K^n wie in Lemma 19.5, so gilt einerseits

$$\widehat{A} = \mathcal{M}_{\widehat{B}_{K^n} \leftarrow \widehat{B}_{K^m}}(f_A) = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \in K^{n \times m}.$$

Andererseits gilt aufgrund des Transformationssatzes 19.20

$$\widehat{A} = S^{-1} A\, T$$

mit den Transformationsmatrizen $T \coloneqq \mathcal{T}_{B_{K^m} \leftarrow \widehat{B}_{K^m}}$ und $S \coloneqq \mathcal{T}_{B_{K^n} \leftarrow \widehat{B}_{K^n}}$. □

Der folgende Satz fasst die Bedeutung äquivalenter Matrizen zusammen:

Satz 19.24 (über äquivalente Matrizen)**.**
Es seien K ein Körper, $n, m \in \mathbb{N}_0$ und $A, \widehat{A} \in K^{n \times m}$. Dann sind äquivalent:

(*i*) A und $\widehat{A}$ sind äquivalente Matrizen.

(*ii*) Es gilt $\operatorname{Rang}(A) = \operatorname{Rang}(\widehat{A})$.

(*iii*) Sind V, W Vektorräume über K mit $\dim(V) = m$ und $\dim(W) = n$ und Basen B_V bzw. B_W, ist $f\colon V \to W$ ein Homomorphismus und gilt $A = \mathcal{M}_{B_W \leftarrow B_V}(f)$, dann gibt es Basen $\widehat{B}_V$ und $\widehat{B}_W$, sodass $\widehat{A} = \mathcal{M}_{\widehat{B}_W \leftarrow \widehat{B}_V}(f)$ gilt.

Beweis. Aussage (*i*) $\Rightarrow$ Aussage (*ii*): Es sei $\widehat{A}$ äquivalent zu A, d. h., es existieren invertierbare Matrizen $S \in K^{n \times n}$ und $T \in K^{m \times m}$, sodass $\widehat{A} = S^{-1} A\, T$ gilt. Nach Folgerung 15.46 (Multiplikation mit invertierbaren Matrizen ändert den Rang nicht) gilt $\operatorname{Rang}(\widehat{A}) = \operatorname{Rang}(A)$.

Aussage (*ii*) $\Rightarrow$ Aussage (*i*): Die Äquivalenzklassen der Äquivalenzrelation partitionieren die Menge $K^{n \times m}$. Die Matrizen A und $\widehat{A}$ gehören also zu jeweils genau einer der Äquivalenzklassen. Sie liegen aber wegen Folgerung 15.46 notwendigerweise in derselben Äquivalenzklasse.

Aussage (*i*) $\Rightarrow$ Aussage (*iii*): Es sei $\widehat{A}$ äquivalent zu A, d. h., es existieren invertierbare Matrizen $S \in K^{n \times n}$ und $T \in K^{m \times m}$, sodass $\widehat{A} = S^{-1} A\, T$ gilt. Wir können T als Transformationsmatrix eines Basiswechsels $T = \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ auffassen, indem wir die „neue" Basis von V durch

$$\widehat{v}_j = \sum_{i=1}^{m} t_{ij}\, v_i \quad \text{für } j = 1, \ldots, m$$

definieren, vgl. (19.12). Ebenso können wir S als Transformationsmatrix eines Basiswechsels $S = \mathcal{T}_{B_W \leftarrow \widehat{B}_W}$ auffassen, indem wir die „neue" Basis von W durch

$$\widehat{w}_j = \sum_{i=1}^{n} s_{ij}\, w_i \quad \text{für } j = 1, \ldots, n$$

definieren. Wir sehen jetzt

$$\begin{aligned}
f_{\widehat{A}} &= f_{S^{-1} A\, T} && \text{nach Voraussetzung} \\
&= f_{S^{-1}} \circ f_A \circ f_T && \text{nach Lemma 17.12} \\
&= f_{S^{-1}} \circ \Phi_{B_W}^{-1} \circ f \circ \Phi_{B_V} \circ f_T && \text{nach Satz 19.8} \\
&= \Phi_{\widehat{B}_W}^{-1} \circ \Phi_{B_W} \circ \Phi_{B_W}^{-1} \circ f \circ \Phi_{B_V} \circ \Phi_{B_V}^{-1} \circ \Phi_{\widehat{B}_V} && \text{nach Lemma 19.16} \\
&= \Phi_{\widehat{B}_W}^{-1} \circ f \circ \Phi_{\widehat{B}_V}.
\end{aligned}$$

Das bedeutet wiederum nach Satz 19.8 aber, dass $\widehat{A}$ die Darstellungsmatrix von f bzgl. der „neuen" Basen $\widehat{V}$ und $\widehat{W}$ ist.

Aussage (*iii*) $\Rightarrow$ Aussage (*i*): Das folgt sofort aus Satz 19.20. □

Bemerkung 19.25 (Matrizen stellen immer lineare Abbildungen dar)**.**
Matrizen stellen (immer) lineare Abbildungen zwischen Vektorräumen dar. Wir verwenden Matrizen also nicht zum „Selbstzweck". Vielmehr sind sie ein praktisches Hilfsmittel, um lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen zu untersuchen. Bisher haben wir beispielsweise den Rang, das Bild und den Kern von Abbildungen $f\colon V \to W$ mit Hilfe ihrer Darstellungsmatrizen bestimmt (Satz 19.10, Beispiel 19.12). △

Expertenwissen: über äquivalente Matrizen

Es seien K ein Körper sowie V und W zwei endlich-dimensionale Vektorräume über K mit $\dim(V) = m$ und $\dim(W) = n$ für $m, n \in \mathbb{N}_0$. Der Satz 19.24 bedeutet, dass folgendes Diagramm für jede feste Wahl von Basen B_V und B_W kommutiert:

$$\begin{tikzcd}
\operatorname{Homo}(V,W) \arrow[r, "\mathcal{M}_{B_W \leftarrow B_V}", "\text{isomorph}"'] \arrow[d, "\pi"'] & K^{n \times m} \arrow[d, "\widetilde{\pi}"] \\
\operatorname{Homo}(V,W) \,/ \sim_{\text{Rang}} \arrow[r, "\text{bijektiv}", "\mathcal{M}_{B_W \leftarrow B_V}"'] & K^{n \times m} \,/ \sim_{\text{Rang}}
\end{tikzcd}$$

Dabei ist auf der linken Seite des Diagramms $\sim_{\text{Rang}}$ die Äquivalenzrelation „gleicher Rang“ auf $\operatorname{Homo}(V, W)$ mit den Äquivalenzklassen

$$[f] \coloneqq \{g \in \operatorname{Homo}(V,W) \mid \operatorname{Rang}(g) = \operatorname{Rang}(f)\}.$$

Weiter ist $\operatorname{Homo}(V, W) \,/ \sim_{\text{Rang}}$ die zugehörige Faktormenge (Menge aller Äquivalenzklassen) und π die kanonische Surjektion $f \mapsto [f]$. Übrigens gibt es wegen $0 \leqslant \operatorname{Rang}(f) \leqslant \min\{m, n\}$ höchstens $\min\{m, n\} + 1$ dieser Äquivalenzklassen, und tatsächlich sind genau $\min\{m, n\} + 1$ Äquivalenzklassen, da man leicht einen Homomorphismus des entsprechenden Ranges angeben kann.

Auf der rechten Seite des Diagramms steht das Symbol $\sim_{\text{Rang}}$ für die Äquivalenzrelation „gleicher Rang“ auf $K^{n,m}$ mit den Äquivalenzklassen

$$\begin{aligned}
[A] &\coloneqq \{\widehat{A} \in K^{n \times m} \mid \operatorname{Rang}(\widehat{A}) = \operatorname{Rang}(A)\} \\
&= \{\widehat{A} \in K^{n \times m} \mid \widehat{A} \text{ ist äquivalent zu } A\} \\
&= \{S^{-1} A\, T \mid S \in K^{n \times n} \text{ und } T \in K^{m \times m} \text{ sind invertierbar}\}.
\end{aligned}$$

Weiter ist $K^{n \times m} \,/ \sim_{\text{Rang}}$ die zugehörige Faktormenge und $\widetilde{\pi}$ die kanonische Surjektion $A \mapsto [A]$.

Nach Satz 19.24 bildet $\mathcal{M}_{B_W \leftarrow B_V}$ eine gesamte Äquivalenzklasse von $\operatorname{Homo}(V, W)$ auf eine Äquivalenzklasse von $K^{n \times m}$ ab und verschiedene Äquivalenzklassen von $\operatorname{Homo}(V, W)$ auf verschiedene Äquivalenzklassen von $K^{n \times n}$. Allerdings sind die Äquivalenzklassen hier (bis auf die für Rang 0) keine Unterräume; $[f] \mapsto [A]$ kann also keine lineare Zuordnung sein. Stattdessen ist jede Äquivalenzklasse $[f]$ von $\operatorname{Homo}(V, W)$ eine sogenannte **glatte Mannigfaltigkeit** (englisch: smooth manifold). Über solche Strukturen erfahren Sie mehr in Vorlesungen über Differentialgeometrie.

Ebenso ist jede Äquivalenzklasse $[A]$ von $K^{n \times m}$ eine glatte Mannigfaltigkeit, und die Abbildung $\mathcal{M}_{B_W \leftarrow B_V}$ ist ein **Diffeomorphismus** zwischen $[f]$ und $[A]$ mit $\operatorname{Rang}(f) = \operatorname{Rang}(A)$. Ein Diffeomorphismus ist eine (beliebig oft) differenzierbare Abbildung (hier: zwischen glatten Mannigfaltigkeiten), deren inverse Abbildung ebenfalls (beliebig oft) differenzierbar ist.

Ende der Vorlesung 27

Ende der Woche 13

Kapitel 5 Dualräume und duale Abbildungen

§ 20 Dualräume

Literatur: Beutelspacher, 2014, Kapitel 5.4; Bosch, 2014, Kapitel 5.1; Fischer, Springborn, 2020, Kapitel 7.1

§ 20.1 Der Dualraum eines Vektorraumes

Definition 20.1 (Dualraum, Linearform)**.**
Es sei V ein Vektorraum über dem Körper K. Der Vektorraum

$$V^* \coloneqq \operatorname{Homo}(V, K) \tag{20.1}$$

der linearen Abbildungen $V \to K$ heißt der **(algebraische) Dualraum** (englisch: (algebraic) dual space) von V. Die Elemente von V^* heißen auch **lineare Funktionale** (englisch: linear functionals) oder **Linearformen** (englisch: linear forms) auf V. △

Zur Unterscheidung bezeichnen wir den Vektorraum V auch als **primalen Vektorraum** (englisch: primal space) und die Elemente von V als **primale Vektoren** (englisch: primal vector). Entsprechend heißt V^* auch der **duale Vektorraum**, und die Elemente von V^* können als **duale Vektoren** (englisch: dual vector) oder als **Covektoren** (englisch: covectors) bezeichnet werden.

Linearformen sind also lineare Abbildungen eines Vektorraumes V in den zugrunde liegenden Skalarkörper K, aufgefasst als 1-dimensionaler Vektorraum über sich selbst. Nach Satz 17.14 ist V^* tatsächlich ein Vektorraum. Der Nullvektor in V^* ist die Nullabbildung $0\colon V \to K$, die jeden Vektor in V auf das Nullelement $0 \in K$ abbildet. Die Nullabbildung wird auch als die **Nullform** (englisch: zero form) auf V bezeichnet.

Beispiel 20.2 (Dualraum, Linearform)**.**

(i) Die Projektion auf die i-te Koordinate (Beispiel 17.5)

$$\pi_i\colon K^n \ni x \mapsto x_i \in K$$

mit $1 \leqslant i \leqslant n$ und $n \in \mathbb{N}$ ist eine Linearform auf dem Standardvektorraum K^n über dem Körper K.

(*ii*) Es sei K ein Körper und X eine Menge. Wir betrachten den Vektorraum $V = K^X$ (Beispiel 11.3). Für jedes feste $x_0 \in X$ ist die **Auswertungsabbildung** (englisch: evaluation map) **an der Stelle** x_0

$$E_{x_0} \colon K^X \ni f \mapsto f(x_0) \in K$$

eine Linearform auf K^X.

(*iii*) Ist $A \in K^{1\times n}$ eine Matrix, dann ist die von A induzierte lineare Abbildung $f_A \colon K^n \to K^1 \cong K$, also die Abbildung $x \mapsto A\,x$, eine Linearform auf dem Standardvektorraum K^n über dem Körper K.

(*iv*) Ist $B = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$ eine Basis des endlich-dimensionalen Vektorraumes V und verwenden wir im Zielraum K die Standardbasis (1), so hat eine Linearform $v^* \in V^*$ die Darstellungsmatrix

$$\mathcal{M}_{(1)\leftarrow B}(v^*) = \begin{bmatrix} v^*(v_1) & \cdots & v^*(v_n) \end{bmatrix} \in K^{1\times n}.$$

(*v*) Ist V ein Vektorraum und $f \colon V \to \mathbb{R}$ eine differenzierbare Funktion, dann ist die Ableitung $f'(v) \colon V \to \mathbb{R}$ an jeder Stelle $v \in V$ eine Linearform auf V:

$$f'(v) \colon V \ni \delta v \mapsto \underbrace{f'(v)(\delta v)}_{\text{Richtungsableitung an der Stelle } v \text{ in Richtung } \delta v} \in \mathbb{R}. \qquad \triangle$$

Wir werden die Vektoren eines Vektorraumes V typischerweise mit v bezeichnen und Linearformen auf V (also Elemente von V^*) mit v^*.

Bemerkung 20.3 (zur Sprechweise im Umgang mit Linearformen)**.**

(*i*) Ausgehend von einer Linearform $v^* \in V^*$ bezeichnen wir die Auswertung von $v^*(v)$ für einen Vektor $v \in V$ auch als das **Einsetzen des Vektors v** in die Linearform v^*.

(*ii*) Stellen wir dagegen den Vektor $v \in V$ in den Vordergrund, so bezeichnen wir die Auswertung von $v^*(v)$ auch als die **Anwendung der Linearform v^*** auf den Vektor v. $\triangle$

Bemerkung 20.4 (zur Bedeutung von Linearformen)**.**

(*i*) Wir können eine Linearform (Covektor) interpretieren als eine Art „Maßstab" für Vektoren. Jeder Vektor wird mit einer „Zahl" bewertet, also einem Element des zugrundeliegenden Skalarkörpers K.

(*ii*) Vektoren, also Elemente eines Vektorraumes, haben oft eine physikalische Bedeutung, z. B. Geschwindigkeit. Als solche tragen sie auch eine Dimension (im Sinne des internationalen Einheitensystems) wie etwa „Länge geteilt durch Zeit" im Falle der Geschwindigkeit, die sich auch in den zugehörigen Einheiten widerspiegelt wie $\frac{\text{m}}{\text{s}}$ oder $\frac{\text{km}}{\text{h}}$.

Elemente des Vektorraumes V^* tragen dann ebenfalls eine Dimension. Für die Dimensionen gilt $[V^*] \cdot [V] = [K]$. Da auch der Skalarkörper eine Dimension tragen kann und nicht notwendig dimensionslos ist, hat V^* nicht notwendig die zu V reziproke Dimension.

(*iii*) Ist $f: V \to \mathbb{R}$ eine differenzierbare Funktion, dann ist die Ableitung $f'(v): V \to \mathbb{R}$ an jeder Stelle $v \in V$ eine Linearform auf V:

$$f'(v): V \ni \delta v \mapsto f'(v)(\delta v) \in \mathbb{R}.$$

Tragen dabei die Funktionswerte $f(v)$ beispielsweise die physikalische Einheit $\mathrm{J} = \mathrm{N\,m}$ (Energie) und die Vektoren in V die Einheit m (Länge), dann tragen die Linearformen und insbesondere die Ableitung $f'(v) \in V^*$ die Einheit N (Kraft). △

Bemerkung 20.5 (auch Vektoren „sind" lineare Abbildungen)**.**
Covektoren (Linearformen) sind genauso natürliche Objekte wie Vektoren. Übrigens können auch primale Vektoren als lineare Abbildungen aufgefasst werden. Dem Vektor $v \in V$ entspricht dabei die Abbildung

$$K \ni \alpha \mapsto \alpha\, v \in V,$$

also ein Element von $\operatorname{Homo}(K, V)$, genauer: Durch diese Zuordnung ist ein Isomorphismus zwischen V und $\operatorname{Homo}(K, V)$ gegeben. Wir können also Vektoren in V mit Elementen in $\operatorname{Homo}(K, V)$ identifizieren, während Covektoren Elemente von $\operatorname{Homo}(V, K)$ sind. △

Definition 20.6 (kanonische duale Paarung)**.**
Es seien V ein Vektorraum über dem Körper K und V^* sein Dualraum. Die Abbildung[1]

$$\langle \cdot\,, \cdot \rangle : V^* \times V \ni (v^*, v) \mapsto \langle v^*, v \rangle \coloneqq v^*(v) \in K \tag{20.2}$$

heißt die **(kanonische) duale Paarung** (englisch: (canonical) dual pairing) der Räume V^* und V. △

Wenn wir die Räume betonen wollen, so schreiben wir auch $\langle v^*, v \rangle_{V^*,V}$.

Die duale Paarung dient zunächst nur dazu, die Lesbarkeit von Ausdrücken im Umgang mit Linearformen zu erleichtern. Wir halten eine wichtige Eigenschaft fest:

Lemma 20.7 (die duale Paarung ist linear in beiden Argumenten)**.**
Es seien V ein Vektorraum und v_1^* sein Dualraum. Die duale Paarung (20.2) ist linear in beiden Argumenten, d. h., für $v^*, v_1^*, v_2^* \in V^*$ und $v, v_1, v_2 \in V$ sowie $\alpha, \beta \in K$ gilt

$$\langle \alpha\, v_1^* + \beta\, v_2^*, v \rangle = \alpha\, \langle v_1^*, v \rangle + \beta\, \langle v_2^*, v \rangle \tag{20.3a}$$

$$\langle v^*, \alpha\, v_1 + \beta\, v_2 \rangle = \alpha\, \langle v^*, v_1 \rangle + \beta\, \langle v^*, v_2 \rangle. \tag{20.3b}$$

Beweis. Es gilt

$$\begin{aligned}
\langle \alpha\, v_1^* + \beta\, v_2^*, v \rangle &= (\alpha\, v_1^* + \beta\, v_2^*)(v) && \text{nach Definition von } \langle \cdot\,, \cdot \rangle \\
&= \alpha\, v_1^*(v) + \beta\, v_2^*(v) && \text{nach Definition der Addition und S-Multiplikation} \\
& && \text{von Homomorphismen (Satz 17.14)} \\
&= \alpha\, \langle v_1^*, v \rangle + \beta\, \langle v_2^*, v \rangle && \text{nach Definition von } \langle \cdot\,, \cdot \rangle.
\end{aligned}$$

Die Gleichung (20.3b) ist nichts anderes als die Linearität von $v^* \in V^* = \operatorname{Homo}(V, K)$. □

[1] Manchmal wird die **duale Paarung** auch in der anderen Reihenfolge definiert, also als $\langle \cdot\,, \cdot \rangle : V \times V^* \ni (v, v^*) \mapsto \langle v, v^* \rangle \coloneqq v^*(v) \in K$.

Ist $B = (v_i)_{i \in I}$ eine Basis von V und ist $(\alpha_i)_{i \in I}$ eine Familie in K mit derselben Indexmenge I, dann gibt es nach Satz 17.10 genau eine Linearform $v^* \in V^*$ mit der Eigenschaft $v^*(v_i) = \langle v^*, v_i \rangle = \alpha_i$ für alle $i \in I$.

Satz 20.8 (die Zuordnung zur Familie der Bilder ist ein Vektorraumisomorphismus, vgl. Satz 19.6)**.**

Es sei V ein Vektorraum über dem Körper K. Weiter sei $B = (v_i)_{i \in I}$ eine Basis von V. Die Zuordnung

$$I_B : V^* \ni v^* \mapsto \langle v^*, v_i \rangle_{i \in I} \in K^I \tag{20.4}$$

einer Linearform zur Familie der Bilder auf der Basis B ist ein Isomorphismus von Vektorräumen.

Beachte: Der Dualraum V^* ist folglich isomorph zu K^I, während V isomorph zum Vektorraum der endlich getragenen Funktionen $(K^I)_{00}$ ist (Beispiel 17.5).

Beweis. Der Beweis folgt dem Beweis von Satz 19.6, der für endlich-dimensionale Vektorräume formuliert war.

Schritt 1: Wir zeigen zunächst die Linearität von I_B.

Es seien dazu $v^*, w^* \in V^*$ und $\alpha, \beta \in K$, dann gilt nach (20.3a)

$$\langle \alpha\, v^* + \beta\, w^*, v_i \rangle = \alpha\, \langle v^*, v_i \rangle + \beta\, \langle w^*, v_i \rangle,$$

also $I_B(\alpha\, v^* + \beta\, w^*) = \alpha\, I_B(v^*) + \beta\, I_B(w^*)$.

Schritt 2: Wir zeigen: I_B ist injektiv.

Es sei dazu $v^* \in V^*$ so, dass $I_B(v^*) = 0 \in K^I$ (die Nullfamilie) ergibt. Das heißt,

$$\langle v^*, v_i \rangle = 0$$

gilt für alle $i = 1, \ldots, m$. Da v^* durch diese Werte aber eindeutig festgelegt ist, kann $v^* : V \to K$ nur die Nullabbildung sein, also der Nullvektor in V^*. Daher gilt $\operatorname{Kern}(I_B) = \{0\}$, und nach Lemma 17.9 ist I_B injektiv.

Schritt 3: Wir zeigen: I_B ist surjektiv.

Es sei dazu ein beliebiges Element $(\alpha_i)_{i \in I} \in K^I$ gegeben. Nach Satz 17.10 gibt es (genau) einen Homomorphismus $v^* : V \to K$, der $\langle v^*, v_i \rangle = \alpha_i$ als Bilder hat, also $I_B(v^*) = (\alpha_i)_{i \in I}$. □

Wir betrachten zunächst Dualräume endlich-dimensionaler Vektorräume genauer.

Folgerung 20.9 (Dimension des Dualraumes, vgl. Folgerung 19.7)**.**
Es sei V ein endlich-dimensionaler Vektorraum. Dann gilt $V \cong V^*$ und damit $\dim(V) = \dim(V^*)$.

Beweis. Es seien $\dim(V) = n \in \mathbb{N}_0$ und $B = (v_1, \ldots, v_n)$ eine Basis von V. Aus Satz 20.8 folgt, dass

$$I_B : V^* \ni v^* \mapsto \big(\langle v^*, v_1 \rangle, \ldots, \langle v^*, v_n \rangle \big) \in K^{[\![1,n]\!]}$$

ein Isomorphismus ist.Andererseits gilt nach Satz 18.1 auch $V \cong K^{[\![1,n]\!]}$, also auch $V^* \cong V$, und damit $\dim(V) = \dim(V^*) = n$. □

Beachte: Der in Folgerung 20.9 konstruierte Isomorphismus zwischen V und V^* hängt von der gewählten Basis B in V ab. Es gibt insbesondere keinen kanonischen (natürlichen) Isomorphismus zwischen den endlich-dimensionalen Vektorräumen V und V^*.

Der folgende Satz beantwortet die Frage, wie wir aus einer Basis B eines endlich-dimensionalen Vektorraumes V eine darauf abgestimmte Basis B^* des Dualraumes V^* erhalten können.

Satz 20.10 (Basis des Dualraumes).
Es seien K ein Körper und V ein endlich-dimensionaler Vektorraum über K mit Basis $B = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$.

(i) Die Linearformen $v_i^* \in V^*$, $i = 1, \dots, n$, definiert durch die folgende Wahl der Bilder auf der Basis $B = (v_j)_{j \in [\![1,n]\!]}$,

$$\langle v_i^*, v_j \rangle := \delta_{ij} = \begin{cases} 1 & \text{im Fall } i = j, \\ 0 & \text{im Fall } i \neq j, \end{cases} \tag{20.5}$$

bilden eine Basis von V^*, die wir mit $B^* := (v_1^*, \dots, v_n^*)$ bezeichnen.

(ii) Für den Koordinatenvektor $x = \Phi_B^{-1}(v)$ eines Vektors $v \in V$ gilt

$$x_i = \langle v_i^*, v \rangle, \quad i = 1, \dots, n, \tag{20.6a}$$

also hat jedes $v \in V$ die eindeutige Darstellung

$$v = \sum_{i=1}^n \langle v_i^*, v \rangle \, v_i. \tag{20.6b}$$

(iii) Für den Koordinatenvektor $\xi = \Phi_{B^*}^{-1}(v^*)$ einer Linearform $v^* \in V^*$ gilt

$$\xi_i = \langle v^*, v_i \rangle, \quad i = 1, \dots, n, \tag{20.7a}$$

also hat jedes $v^* \in V^*$ die eindeutige Darstellung

$$v^* = \sum_{i=1}^n \langle v^*, v_i \rangle \, v_i^*. \tag{20.7b}$$

Beweis. Zunächst stellen wir fest, dass die Linearformen $v_1^*, \dots, v_n^*$ durch (20.5) eindeutig definiert sind, da wir für jedes v_i^* die Bilder der Basisvektoren $v_1, \dots, v_n$ angegeben haben, die nach Satz 17.10 die lineare Abbildung $v_i^* \in \operatorname{Homo}(V, K) = V^*$ eindeutig festlegen.

Wir zeigen zunächst gemeinsam die Aussagen (i) und (iii) und anschließend Aussage (ii).

Schritt 1: Die Vektoren $v_1^*, \dots, v_n^*$ sind linear unabhängig:
Es sei $\sum_{i=1}^n \alpha_i \, v_i^* = 0_{V^*}$ der Nullvektor in V^*, also die Nullform. Durch Einsetzen von v_j in diese Linearkombination erhalten wir unter Benutzung von (20.3a)

$$0 = \langle 0_{V^*}, v_j \rangle = \Big\langle \sum_{i=1}^n \alpha_i \, v_i^*, v_j \Big\rangle = \sum_{i=1}^n \alpha_i \underbrace{\langle v_i^*, v_j \rangle}_{=\delta_{ij}} = \alpha_j$$

für alle $j = 1, \dots, n$, also sind alle Koeffizienten notwendigerweise gleich Null.

Schritt 2: Die Vektoren $v_1^*, \ldots, v_n^*$ erzeugen den ganzen Raum V^*:

Es sei dazu $v^* \in V^*$ beliebig. Wir müssen zeigen, dass v^* als Linearkombination von $v_1^*, \ldots, v_n^*$ darstellbar ist. Dazu überprüfen wir die Darstellung (20.7), also

$$v^* = \sum_{i=1}^n \langle v^*, v_i \rangle \, v_i^*.$$

Beide Seiten der Gleichung sind Elemente von V^*. Wir zeigen, dass diese Linearformen auf den Bildern der Basisvektoren $v_1, \ldots, v_n$ von V (und damit überall in V) übereinstimmen. In der Tat gilt

$$\Big\langle \sum_{i=1}^n \langle v^*, v_i \rangle \, v_i^*, v_j \Big\rangle = \sum_{i=1}^n \langle v^*, v_i \rangle \, \underbrace{\langle v_i^*, v_j \rangle}_{=\delta_{ij}} = \langle v^*, v_j \rangle.$$

Schritt 3: Wir zeigen noch die Darstellung (20.6), also Aussage *(ii)*.

Für jeden der Basisvektoren $v_j \in V$ ist die Darstellung wegen

$$\sum_{i=1}^n \underbrace{\langle v_i^*, v_j \rangle}_{=\delta_{ij}} \, v_i = v_j$$

korrekt. Jedes $v \in V$ hat eine eindeutige Darstellung der Form $v = \sum_{j=1}^n \alpha_j \, v_j$ mit noch unbekannten Koeffizienten $\alpha_j \in K$. Es gilt also

$$\begin{aligned} v &= \sum_{j=1}^n \alpha_j \, v_j \\ &= \sum_{j=1}^n \alpha_j \sum_{i=1}^n \langle v_i^*, v_j \rangle \, v_i && \text{wie gerade bestätigt} \\ &= \sum_{i=1}^n \Big\langle v_i^*, \sum_{j=1}^n \alpha_j \, v_j \Big\rangle \, v_i && \text{wegen der Linearität der dualen Paarung} \\ &= \sum_{i=1}^n \langle v_i^*, v \rangle \, v_i && \text{aufgrund der Darstellung von } v. \end{aligned}$$

□

Definition 20.11 (duale Basis)**.**
Es sei V ein endlich-dimensionaler Vektorraum mit Basis $B = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$.

(*i*) Die durch (20.5) definierte Basis $B^* \coloneqq (v_1^*, \ldots, v_n^*)$ von V^* heißt die zur Basis B **duale Basis** (englisch: dual basis).

(*ii*) Wir bezeichnen das Element v_i^* der zu B dualen Basis B^* auch die ***i*-te Koordinatenform** (englisch: *i*-th coordinate form) der Basis B. △

Bemerkung 20.12 (zur dualen Basis)**.**

(i) Jeder Vektor v_i^* der dualen Basis hängt von der Gesamtheit der Wahl der Basisvektoren $(v_1, \dots, v_n)$ der Basis von V ab, nicht nur von v_i, da die dualen Basisvektoren gemeinsam das System (20.5) erfüllen müssen.

(ii) Die Elemente $v_1^*, \dots, v_n^*$ der dualen Basis B^* dienen als **„Koordinatenermittler"** für primale Vektoren bzgl. der Basis B, siehe (20.6). Wenn wir die duale Basis B^* kennen, müssen wir also kein lineares Gleichungssystem mehr lösen, um die Koordinatenvektor $x = \Phi_B^{-1}(v)$ eines Vektors $v \in V$ zu bestimmen.

Umgekehrt dienen die Elemente $v_1, \dots, v_n$ der primalen Basis B als **„Koordinatenermittler"** für duale Vektoren (Linearformen) bzgl. der dualen Basis B^*, siehe (20.7).

(iii) Es seien V und W Vektorräume über demselben Körper K mit Basen $B_V = (v_1, \dots, v_m)$ bzw. $B_W = (w_1, \dots, w_n)$ und $m, n \in \mathbb{N}_0$. Die Darstellungsmatrix $A = \mathcal{M}_{B_W \leftarrow B_V}(f) \in K^{n \times m}$ eines Homomorphismus $f\colon V \to W$ hat die Eigenschaft, dass sie spaltenweise die Koordinaten der Bilder der Basisvektoren v_j bzgl. der Basis B_W enthält:

$$f(v_j) = \sum_{i=1}^n a_{ij}\, w_i \quad \text{für alle } j = 1, \dots, m,$$

vgl. (19.3). Wir hatten bisher aber keine Möglichkeit, beispielsweise einzelne Einträge der Darstellungsmatrix A anzugeben.[2]

Ist jedoch die duale Basis $B_{W^*} = (w_1^*, \dots, w_n^*)$ von W^* bekannt, so können wir einzelne oder auch alle Einträge der Darstellungsmatrix A wie in (20.6) leicht bestimmen:

$$a_{ij} = \langle w_i^*, f(v_j) \rangle. \tag{20.8}$$

△

Beispiel 20.13 (duale Basis).

(i) Es sei K ein Körper, $n \in \mathbb{N}_0$ und $B = (e_1, \dots, e_n)$ die Standardbasis von $V = K^n$ für $n \in \mathbb{N}_0$. Dann besteht die zu B duale Basis B^* aus den Projektionen auf die Koordinaten (Beispiel 20.2), also $B^* = (\pi_1, \dots, \pi_n)$, denn es gilt

$$\langle \pi_i, e_j \rangle = i\text{-te Koordinate von } e_j = \delta_{ij}$$

für alle $i, j = 1, \dots, n$, also (20.5).

Wie in (20.6) festgestellt, ermittelt das Element π_i der dualen Basis die i-te Koordinate eines Vektors bzgl. der Standardbasis in K^n.

(ii) Es sei $X = \{x_1, \dots, x_n\}$ eine endliche Menge mit $n \in \mathbb{N}_0$ paarweise verschiedenen Elementen. Für jeden Körper K bilden die charakteristischen Funktionen $B \coloneqq (e_{x_1}, \dots, e_{x_n})$ (Beispiel 12.2) eine Basis des Vektorraumes $V \coloneqq K^X$. Die zugehörige duale Basis ist gegeben durch die Auswertungsabbildungen (Beispiel 20.2) $B^* \coloneqq (E_{x_1}, \dots, E_{x_n})$, die eine gegebene Funktion $X \to K$ an jeweils einer der Stellen $x_i \in X$ auswerten.

(iii) Wir betrachten ähnlich wie in Beispiel 19.12 einen dreidimensionalen Unterraum V von $\mathbb{Q}^{\mathbb{Q}}$, der durch die Basis

$$B = \big(t \mapsto 1,\ t \mapsto t,\ t \mapsto t^2\big)$$

[2] I. A. muss zur Bestimmung der Spalte $a_{\bullet j} = \Phi_{B_W}^{-1}(f(v_j))$ ein lineares Gleichungssystem für die Linearkombinationskoeffizienten gelöst werden.

aufgespannt wird. Die duale Basis kann dann beispielsweise in Form von Punktauswertungsfunktionalen dargestellt werden:

$$B^* = \Big(p \mapsto p(0),\ p \mapsto \frac{1}{2}p(1) - \frac{1}{2}p(-1),\ p \mapsto \frac{1}{2}p(1) + \frac{1}{2}p(-1) - p(0)\Big)$$

für $p \in V$. (**Quizfrage 20.1:** Wie kommt man auf eine solche Darstellung der dualen Basis?)

Wollen wir nun beispielsweise die Linearform $v^* \coloneqq p \mapsto p(2) \in V^*$ als Linearkombination der Basisvektoren von B^* darstellen, so ergeben sich die Koeffizienten nach (20.7) dadurch, dass wir v^* auf die Elemente der primalen Basis B wirken lassen (also die „Koordinatenermittler"-Eigenschaft ausnutzen):

$$v^*(p_1) = 1, \quad v^*(p_2) = 2, \quad v^*(p_3) = 8.$$

Für die Linearform v^* gilt also, dass sie $p \in V$ abbildet auf

$$\begin{aligned} &p(0) + 2\Big(\frac{1}{2}p(1) - \frac{1}{2}p(-1)\Big) + 8\Big(\frac{1}{2}p(1) + \frac{1}{2}p(-1) - p(0)\Big) \\ &\quad = -7\,p(0) + 5\,p(1) + 3\,p(-1). \end{aligned}$$

Wir erhalten also die Darstellung $p(2) = -7\,p(0) + 5\,p(1) + 3\,p(-1)$ für alle p aus dem Unterraum V von $\mathbb{Q}^\mathbb{Q}$. △

Abschließend halten wir noch fest, dass die Verwendung dualer Basen auch zur Auswertung der dualen Paarung nützlich ist:

Lemma 20.14 (Darstellung der dualen Paarung bei Verwendung von Basis und dualer Basis)**.** Es sei V ein endlich-dimensionaler Vektorraum mit Basis $B = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$, und zugehöriger dualer Basis $B^* = (v_1^*, \ldots, v_n^*)$. Sind

$$v^* = \sum_{i=1}^n \xi_i\, v_i^* \quad \text{und} \quad v = \sum_{j=1}^n x_j\, v_j$$

die Darstellungen von $v^* \in V^*$ bzw. $v \in V$ bzgl. dieser Basen, so gilt

$$\langle v^*, v\rangle = \sum_{j=1}^n \xi_j\, x_j = \xi^\mathsf{T} x. \tag{20.9}$$

Beweis. Unter Ausnutzung der Linearität der dualen Paarung in beiden Argumenten (Lemma 20.7) erhalten wir

$$\langle v^*, v\rangle = \Big\langle \sum_{i=1}^n \xi_i\, v_i^*, \sum_{j=1}^n x_j\, v_j \Big\rangle = \sum_{i=1}^n \sum_{j=1}^n \xi_i\, x_j \underbrace{\langle v_i^*, v_j\rangle}_{\delta_{ij}} = \sum_{j=1}^n \xi_j\, x_j = \xi^\mathsf{T} x. \qquad \square$$

Durch die Verwendung von Basis und dualer Basis wird die duale Paarung also zurückgeführt auf eine Art Standardform (20.9), das Matrix-Matrix-Produkt $\xi^\mathsf{T} x \in K$ zwischen den zugehörigen Koordinatenvektoren $\xi \in K^n \cong K^{n\times 1}$ und $x \in K^n \cong K^{n\times 1}$.

Beispiel 20.15 (Darstellung der dualen Paarung bei Verwendung von Basis und dualer Basis).

(*i*) Im Vektorraum K^n stimmen Vektoren v mit ihren Koordinatenvektoren x überein, wenn wir die Standardbasis $(e_1, \ldots, e_n)$ wählen. Eine Linearform $v^* \in (K^n)^*$ können wir durch ihren Koordinatenvektor $\xi \in K^n$ bzgl. der dualen Basis $(\pi_1, \ldots, \pi_n)$ darstellen, bestehend aus den den Projektionen auf die Koordinaten (Beispiel 20.13). Die duale Paarung $\langle v^*, v\rangle$ hat dann die Gestalt $\xi^\intercal x$, siehe (20.9).

(*ii*) Wir betrachten wie in Beispiel 20.13 eine endliche Menge $X = \{x_1, \ldots, x_n\}$ mit $n \in \mathbb{N}_0$ paarweise verschiedenen Elementen. Für jeden Körper K bilden die charakteristischen Funktionen $B \coloneqq (e_{x_1}, \ldots, e_{x_n})$ eine Basis des Vektorraumes $V \coloneqq K^X$. Die zugehörige duale Basis ist gegeben durch die Auswertungsabbildungen $B^* \coloneqq (E_{x_1}, \ldots, E_{x_n})$.

Jede Linearform auf V wird eindeutig durch ihren Koordinatenvektor bzgl. der dualen Basis dargestellt. Es sei beispielsweise $X = \{\text{rot}, \text{grün}, \text{blau}\}$ und $V = \mathbb{Q}^X$ mit der Basis $B \coloneqq (e_{\text{rot}}, e_{\text{grün}}, e_{\text{blau}})$ ausgestattet. Wir betrachten als Beispiel die Linearform $v^* \in V^*$, die wie folgt auf Funktionen $f\colon X \to \mathbb{Q}$ (Vektoren in V) wirkt:

$$\langle v^*, f\rangle \coloneqq \frac{3}{2} f(\text{rot}) - \frac{1}{4} f(\text{grün}) + 2 f(\text{blau}).$$

Die Linearform v^* wird bzgl. der dualen Basis B^* durch den Koordinatenvektor $\xi = \frac{1}{4}\begin{pmatrix} 6 \\ -1 \\ 8 \end{pmatrix}$ dargestellt. Hat also f den Koordinatenvektor $x = \begin{pmatrix} f(\text{rot}) \\ f(\text{grün}) \\ f(\text{blau}) \end{pmatrix}$ bzgl. der Basis B, dann können wir die duale Paarung in der Form $\langle v^*, f\rangle = \xi^\intercal x$ schreiben. △

Bemerkung 20.16 (Identifikation von $(K^n)^*$ mit K^n).

Wir vereinbaren, dass wir im Folgenden im Vektorraum K^n wann immer möglich die Standardbasis $B_n = (e_1, \ldots, e_n)$ verwenden, sodass Vektoren $v \in K^n$ mit ihren Koordinatenvektoren $x = \Phi_{B_n}^{-1}(v)$ übereinstimmen.

Im Dualraum $(K^n)^*$ verwenden wir dann die zur Standardbasis duale Basis $B_n^* = (\pi_1, \ldots, \pi_n)$, also die Projektionen auf die Koordinaten. Wie im obigen Beispiel 20.15 können wir dann Linearformen v^* auf K^n über ihre Koordinatenvektoren $\xi = \Phi_{B_n^*}^{-1}(v^*) \in K^n$ darstellen.

Es wird sich als nützlich erweisen, den Dualraum $(K^n)^*$ über diesen Isomorphismus $\Phi_{B_n^*}^{-1}$ mit K^n zu identifizieren. Das heißt konkret, dass wir im Folgenden Koordinatenvektoren $\xi \in K^n$ als Elemente von $(K^n)^*$ auffassen. Koordinatenvektoren, die primale Vektoren in K^n darstellen, werden notationell unterschieden und weiterhin mit x usw. bezeichnet. Die duale Paarung wird durch $\xi^\intercal x$ realisiert. △

Wir wenden uns nun noch kurz den Dualräumen unendlich-dimensionaler Vektorräume zu. (**Quizfrage 20.2:** Warum funktioniert der Beweis der Isomorphie von V und V^* in Folgerung 20.9 nicht auch für unendlich-dimensionale Vektorräume?)

Bemerkung 20.17 (Dualräume unendlich-dimensionaler Vektorräume).

(i) Ist $B = (v_i)_{i \in I}$ eine Basis des Vektorraumes V, dann gilt $V \cong (K^I)_{00}$ (Beispiel 17.5). Für den Dualraum gilt nach Satz 20.8 jedoch $V^* \cong K^I$.

Besitzt V unendliche Dimension und würden wir unendliche Dimensionen von Vektorräumen mittels Kardinalzahlen genauer unterscheiden (Bemerkung 13.15), so könnten wir hier $\dim(V^*) > \dim(V)$ feststellen.

Insbesondere erhalten wir im Falle $\dim(V) = \infty$ mit der Konstruktion (20.5) keine Basis des Dualraumes V^*: Ist $B = (v_i)_{i \in I}$ eine Basis von V mit unendlicher Indexmenge I und definieren wir wie in (20.5) eine gleichmächtige Familie $B^* \coloneqq (v_i^*)_{i \in I}$ an Elementen von V^* durch

$$\langle v_i^*, v_j \rangle \coloneqq \delta_{ij} = \begin{cases} 1 & \text{im Fall } i = j, \\ 0 & \text{im Fall } i \neq j, \end{cases}$$

dann ist B^* zwar eine linear unabhängige Familie in V^*, sie erzeugt aber nicht den gesamten Vektorraum V^*, denn die durch

$$\langle v^*, v_i \rangle = 1 \quad \text{für alle } i \in I$$

definierte Linearform kann nicht als (endliche!) Linearkombination von Elementen aus B^* geschrieben werden.

(ii) Jeder Vektorraum V besitzt einen Dualraum, jedoch ist nicht jeder Vektorraum selbst ein Dualraum. Beispielweise kann ein Dualraum keine abzählbar unendliche Basis besitzen. Damit ist der Vektorraum $(K^{\mathbb{N}})_{00}$ der endlich getragenen Folgen über einem Körper K (siehe Beispiel 13.2) von keinem Vektorraum der Dualraum. △

Ende der Vorlesung 1

§ 20.2 Transformationsmatrizen des Basiswechsels

Das folgende Lemma klärt, wie sich die duale Basis bei einem Wechsel der primalen Basis (§ 19.4) verhält.

Lemma 20.18 (Zusammenhang der Basiswechsel zwischen primaler und dualer Basis)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter seien $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ zwei Basen von V. Wir betrachten $B_{V^*} = (v_1^*, \ldots, v_n^*)$ und $\widehat{B}_{V^*} = (\widehat{v}_1^*, \ldots, \widehat{v}_n^*)$, die jeweiligen dualen Basen. Ist

$$T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V} = \mathcal{M}_{B_V \leftarrow \widehat{B}_V}(\mathrm{id}_V)$$

die Transformationsmatrix des Basiswechsels von $\widehat{B}_V$ nach B_V (Definition 19.15), dann ist

$$T^{-\mathsf{T}} = \mathcal{T}_{B_{V^*} \leftarrow \widehat{B}_{V^*}} = \mathcal{M}_{B_{V^*} \leftarrow \widehat{B}_{V^*}}(\mathrm{id}_{V^*}) \tag{20.10}$$

die Transformationsmatrix des Basiswechsels von $\widehat{B}_{V^*}$ nach B_{V^*}. Ist $\widehat{\xi} \in K^n$ der Koordinatenvektor der Linearform $v^* \in V^*$ bzgl. der Basis $\widehat{B}_{V^*}$, dann ist $\xi = T^{-\mathsf{T}} \widehat{\xi}$ der Koordinatenvektor von v^* bzgl. der Basis B_{V^*}.[3]

[3]Vergleiche mit der Transformation von Koordinatenvektoren $x = \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \widehat{x} = T \widehat{x}$ von Vektoren in V, siehe Lemma 19.16.

Beweis. Mit der Transformationsmatrix $T = \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ des Basiswechsels von $\widehat{B}_V$ nach B_V gilt

$$\widehat{v}_j = \sum_{i=1}^n t_{ij}\, v_i,$$

siehe (19.12). Die gesuchte Transformationsmatrix $S = \mathcal{T}_{B_{V^*} \leftarrow \widehat{B}_{V^*}}$ des Basiswechsels von $\widehat{B}_{V^*}$ nach B_{V^*} erfüllt analog

$$\widehat{v}_j^* = \sum_{i=1}^n s_{ij}\, v_i^*.$$

Wir betrachten nun

$$\begin{aligned}
\langle \widehat{v}_i^*, \widehat{v}_j \rangle &= \Big\langle \sum_{k=1}^n s_{ki}\, v_k^*, \sum_{\ell=1}^n t_{\ell j}\, v_\ell \Big\rangle \\
&= \sum_{k=1}^n \sum_{\ell=1}^n s_{ki}\, t_{\ell j}\, \langle v_k^*, v_\ell \rangle && \text{wegen der beidseitigen Linearität der dualen Paarung} \\
&= \sum_{k=1}^n s_{ki}\, t_{kj} && \text{denn die Dualität der Basen } B_{V^*} \text{ und } B_V \text{ besagt } \langle v_k^*, v_\ell \rangle = \delta_{k\ell} \\
&= (S^\intercal T)_{ij} && \text{nach Definition (15.9) der Matrix-Matrix-Multiplikation.}
\end{aligned}$$

Andererseits ist auch $\widehat{B}_{V^*}$ die zu $\widehat{B}_V$ duale Basis, also muss $\langle \widehat{v}_i^*, \widehat{v}_j \rangle = \delta_{ij}$ gelten. Das heißt, wir haben $I = S^\intercal T$, also $S = T^{-\intercal}$.

Wie in Lemma 19.16 transformiert sich der Koordinatenvektor ξ wie folgt:

$$\xi = \mathcal{T}_{B_{V^*} \leftarrow \widehat{B}_{V^*}}\, \widehat{\xi} = T^{-\intercal} \widehat{\xi}. \qquad \square$$

Diese Transformationsvorschrift können wir uns auch aus (20.9) herleiten, denn diese Gleichung ist ja unabhängig davon, welche Basis und zugeörige duale Basis wir wählen. Wir wissen also einerseits

$$\langle v^*, v \rangle = \xi^\intercal x = \widehat{\xi}^\intercal \widehat{x}$$

und andererseits $x = T\,\widehat{x}$, also muss $\xi = T^{-\intercal} \widehat{\xi}$ gelten.

Bemerkung 20.19 (kovariante und kontravariante Transformationen).
Wir haben in Lemma 19.16 gesehen, dass sich beim Übergang von $\widehat{B}_V$ zu B_V mittels $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ mit

$$\widehat{v}_j = \sum_{i=1}^n T_{ij}\, v_i \quad \text{die Koordinatenvektoren gemäß} \quad \widehat{x}_i = \sum_{j=1}^n (T^{-1})_{ij}\, x_j$$

$$\text{bzw. nach Tausch } i \leftrightarrow j \text{ gemäß} \quad \widehat{x}_j = \sum_{i=1}^n (T^{-\intercal})_{ij}\, x_i \tag{20.11a}$$

transformieren. Demgegenüber transformieren sich

$$\text{duale Koordinatenvektoren gemäß} \quad \widehat{\xi}_i = \sum_{j=1}^n (T^\intercal)_{ij}\, \xi_j$$

$$\text{bzw. nach Tausch } i \leftrightarrow j \text{ gemäß} \qquad \widehat{\xi_j} = \sum_{i=1}^{n} T_{ij}\, \xi_i. \tag{20.11b}$$

Eine historische, aber immer noch verbreitete Sprechweise für diese Beobachtung ist, dass sich

die Koordinaten eines Vektors in V **kontravariant transformieren**[4] in Bezug auf die Transformation der Basisvektoren von V, nämlich mit der Matrix $T^{-\intercal}$ gegenüber T, siehe (20.11a).[5]

die Koordinaten einer Linearform in V^* **kovariant transformieren**[6] in Bezug auf die Transformation der Basisvektoren von V, nämlich mit derselben Matrix T, siehe (20.11b).[5]

△

§ 20.3 Annihilatoren und Prä-Annihilatoren

Definition 20.20 (Annihilator und Prä-Annihilator)**.**

Es seien K ein Körper und V ein Vektorraum über K.

(i) Ist $M \subseteq V$, dann heißt

$$\begin{aligned} M^0 &\coloneqq \{v^* \in V^* \mid \langle v^*, v\rangle = 0 \text{ für alle } v \in M\} \\ &= \{v^* \in V^* \mid M \subseteq \operatorname{Kern}(v^*)\} \\ &= \bigcap_{v \in M} \{v\}^0 \subseteq V^* \end{aligned} \tag{20.12}$$

der **Annihilator** von M (englisch: annihilator, lateinisch: annihilare: vernichten, auslöschen).[8]

(ii) Ist $F \subseteq V^*$, dann heißt

$$\begin{aligned} {}^0F &\coloneqq \{v \in V \mid \langle v^*, v\rangle = 0 \text{ für alle } v^* \in F\} \\ &= \bigcap_{v^* \in F} \operatorname{Kern}(v^*) \subseteq V \end{aligned} \tag{20.13}$$

der **Prä-Annihilator** von F (englisch: pre-annihilator).[9] △

Wir geben zunächst die (Prä-)Annihilatoren bestimmter Teilmengen von V und V^* an, nämlich die zu den trivialen Unterräumen eines Vektorraumes V bzw. seines Dualraumes V^*.

[4] englisch: contravariant, lateinisch: contra: gegen, lateinisch: variare: verändern, verwandeln

[5] Die einfachste Ausprägung dieser Beobachtung ergibt sich, wenn etwa v_1 in der neuen Basis durch $\widehat{v_1} = 2\, v_1$ ersetzt wird. Dadurch muss für den Koordinatenvektor bzgl. der neuen Basis $\widehat{x_1} = \frac{1}{2} x_1$ gelten (kontravariant). Folglich muss sich der Koordinatenvektor einer Linearform mit $\widehat{\xi_1} = 2\, \widehat{\xi_2}$ transformieren (kovariant).

[6] englisch: covariant, lateinisch: co- (Vorsilbe): mit

[8] Der Annihilator einer Menge $M \subseteq V$ besteht also aus denjenigen Linearformen auf V, die auf der Menge M verschwinden, also dort den Funktionswert Null haben.

[9] Der Prä-Annihilator einer Menge $F \subseteq V^*$ besteht also aus denjenigen Vektoren in V, auf denen alle Linearformen $v^* \in F$ verschwinden.

Lemma 20.21 (Annihilatoren und Prä-Annihilatoren trivialer Unterräume[10] AoC)**.**
Es seien K ein Körper und V ein Vektorraum über K. Dann gilt:

$$\text{für die Annihilatoren} \quad \{0_V\}^0 = V^* \quad \text{und} \quad V^0 = \{0_{V^*}\}, \tag{20.14a}$$
$$\text{für die Prä-Annihilatoren} \quad {}^0\{0_{V^*}\} = V \quad \text{und} \quad {}^0(V^*) = \{0_V\}. \tag{20.14b}$$

Beweis. Wir zeigen zunächst (20.14a). Jede Linearform $v^* \in V^*$ bildet 0_V auf $0 \in K$ ab (Lemma 17.6). Das bedeutet
$$\{0_V\}^0 = \{v^* \in V^* \mid \langle v^*, 0_V \rangle = 0\} = V^*.$$
Die Menge
$$V^0 = \{v^* \in V^* \mid \langle v^*, v \rangle = 0 \text{ für alle } v \in V\}$$
enthält nach Definition nur die Nullform, also gilt $V^0 = \{0_{V^*}\}$.

Nun zu (20.14b): Nach Definition ist
$${}^0\{0_{V^*}\} = \{v \in V \mid \langle 0_{V^*}, v \rangle = 0\} = V.$$
Weiter gilt
$${}^0(V^*) = \{v \in V \mid \langle v^*, v \rangle = 0 \text{ für alle } v^* \in V^*\}.$$
Es sei $v \in V \setminus \{0_V\}$ beliebig. Dann existiert nach Satz 17.10 (ii) (für dessen Beweis der Basisergänzungssatz 13.5 benötigt wird) eine Linearform $v^* \in V^*$ mit der Eigenschaft $v^*(v) = 1$. (**Quizfrage 20.3:** Klar?) Das bedeutet, dass ein solches $v \in V \setminus \{0_V\}$ nicht zu ${}^0(V^*)$ gehören kann. Andererseits gehört 0_V sicher zu ${}^0(V^*)$, da $v^*(0_V) = 0$ für alle $v^* \in V^*$ gilt. Das bedeutet aber ${}^0(V^*) = \{0_V\}$. □

Die Annihilatoren bzw. Prä-Annihilatoren der trivialen Unterräume aus Lemma 20.21 sind selbst wieder Unterräume. Das gilt auch allgemein für die (Prä-)Annihilatoren beliebiger Mengen:

Lemma 20.22 (Annihilatoren und Prä-Annihilatoren sind Unterräume)**.**
Es seien K ein Körper und V ein Vektorraum über K. Dann gilt:

(i) Ist $M \subseteq V$, dann ist der Annihilator M^0 ein Unterraum von V^*.
Weiter gilt $M^0 = \langle M \rangle^0$.

(ii) Ist $F \subseteq V^*$, dann ist der Prä-Annihilator ${}^0F$ ein Unterraum von V.
Weiter gilt ${}^0F = {}^0\langle F \rangle$.

Beweis. Aussage (i): Wir benutzen das Unterraumkriterium (Satz 11.11). Weil die Nullform 0_{V^*} in M^0 liegt, ist $M^0 \neq \emptyset$. Sind nun $v^*, w^* \in M^0$ und $\alpha, \beta \in K$, dann gilt
$$\langle \alpha\, v^* + \beta\, w^*, v \rangle = \alpha\, \langle v^*, v \rangle + \beta\, \langle w^*, v \rangle = \alpha\, 0 + \beta\, 0 = 0$$
für alle $v \in M$, also gehört auch $\alpha\, v^* + \beta\, w^*$ zu M^0. Das heißt, M^0 ist ein Unterraum.

[10] Nur die letzte Aussage ${}^0(V^*) = \{0_V\}$ hängt im Fall, dass V unendlich-dimensional ist, über den Basisergänzungssatz 13.5 vom Auswahlaxiom ab.

Außerdem gilt $M \subseteq \langle M \rangle$, also auch $\langle M \rangle^0 \subseteq M^0$. Ist jedoch $v^* \in M^0$, dann annihiliert v^* auch alle Linearkombinationen von M, also $\langle M \rangle$ (Satz 11.16). Das heißt $M^0 \subseteq \langle M \rangle^0$.

Aussage (ii): Der Beweis geht analog. □

Der folgende Satz klärt, wie wir in endlich-dimensionalen Vektorräumen eine Basis eines Annihilator- bzw. eines Prä-Annihilator-Unterraumes bestimmen können und was die Dimensionen dieser Unterräume sind.

Satz 20.23 (Basen von Annihilatoren und Prä-Annihilatoren)**.**
Es seien V ein endlich-dimensionaler Vektorraum mit Basis $B = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$, und $B^* = (v_1^*, \dots, v_n^*)$ die duale Basis.

(i) Ist $(v_1, \dots, v_k)$ für $k \in \mathbb{N}_0$ eine Basis des Unterraumes $U \subseteq V$, dann ist $(v_{k+1}^*, \dots, v_n^*)$ eine Basis des Annihilators U^0. Insbesondere gilt

$$\dim(U^0) = \dim(V) - \dim(U) = \operatorname{codim}(U). \tag{20.15}$$

(ii) Ist $(v_1^*, \dots, v_k^*)$ für $k \in \mathbb{N}_0$ eine Basis des Unterraumes $F \subseteq V^*$, dann ist $(v_{k+1}, \dots, v_n)$ eine Basis des Prä-Annihilators ${}^0F$. Insbesondere gilt

$$\dim({}^0F) = \dim(V^*) - \dim(F) = \dim(V) - \dim(F) = \operatorname{codim}(F). \tag{20.16}$$

Beweis. Wir beweisen nur Aussage (i). Als Teilfamilie der Basis B^* ist $(v_{k+1}^*, \dots, v_n^*)$ linear unabhängig (Lemma 12.4). Zu zeigen ist also nur, dass $(v_{k+1}^*, \dots, v_n^*)$ eine erzeugende Menge von U^0 ist.

Schritt 1: $\langle v_{k+1}^*, \dots, v_n^* \rangle \subseteq U^0$:

Aufgrund der Eigenschaften der dualen Basis gilt $v_i^*(v_j) = 0$ für alle $i = k+1, \dots, n$ und alle $j = 1, \dots, k$. Also verschwindet v_i^* sogar auf dem ganzen Unterraum $U = \langle v_1, \dots, v_k \rangle$, d. h., wir haben $v_i^* \in U^0$ für alle $i = k+1, \dots, n$. Da U^0 aber ein Unterraum (von V^*) ist, gilt auch $\langle v_{k+1}^*, \dots, v_n^* \rangle \subseteq U^0$.

Schritt 2: $\langle v_{k+1}^*, \dots, v_n^* \rangle \supseteq U^0$:

Es sei $v^* \in U^0 \subseteq V^*$. Die Linearform v^* ist eindeutig durch ihre Bilder auf der Basis $B = (v_1, \dots, v_n)$ festgelegt (Satz 17.10). Es gilt

$$\begin{aligned}
v^* &= \sum_{i=1}^n v^*(v_i)\, v_i^* && \text{nach (20.7b)} \\
&= \sum_{i=1}^k \underbrace{v^*(v_i)}_{=0}\, v_i^* + \sum_{i=k+1}^n v^*(v_i)\, v_i^* && \text{wegen } v^* \in U^0 \text{ und } v_i \in U \text{ für } i = 1, \dots, k \\
&= \sum_{i=k+1}^n v^*(v_i)\, v_i^*.
\end{aligned}$$

Das bedeutet nach Satz 11.16 aber $v^* \in \langle v_{k+1}^*, \dots, v_n^* \rangle$. □

Beispiel 20.24 (Basen von Annihilatoren und Prä-Annihilatoren).

(i) Wir betrachten den Vektorraum $V \coloneqq K^{2\times 2}$ der 2×2-Matrizen über dem Körper K. Es sei $U \coloneqq K^{2\times 2}_{\text{sym}} \subseteq V$ der Unterraum der symmetrischen Matrizen, also

$$U = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \,\middle|\, a_{12} - a_{21} = 0 \right\}.$$

Der Annihilator U^0 besteht aus allen Linearformen $v^*\colon V \to K$, die auf U verschwinden. Das sind gerade genau die Vielfachen der durch $v^*(A) = a_{12} - a_{21}$ für $A \in K^{2\times 2}$ definierten Linearform v^*, also

$$U^0 = \{v^* \in V^* \mid v^*(A) = \alpha\,(a_{12} - a_{21}) \text{ für ein } \alpha \in K\}.$$

Wir können das mit Hilfe von Satz 20.23 bestätigen. Im Falle von $\operatorname{char}(K) \neq 2$ ist

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

eine Basis von V, wobei die ersten $\dim(U) = 3$ Basisvektoren den Unterraum U aufspannen. Die duale Basis ist

$$B^* = \Big(A \mapsto a_{11},\ A \mapsto a_{22},\ A \mapsto \frac{1}{2}(a_{12} + a_{21}),\ A \mapsto \frac{1}{2}(a_{12} - a_{21})\Big).$$

Nach Satz 20.23 spannt der letzte Basisvektor von B^* den Annihilator U^0 auf. (**Quizfrage 20.4:** Wie sind die Basen B und B^* zu verändern im Fall $\operatorname{char} K = 2$?)

(ii) Wir betrachten den Vektorraum $V \coloneqq K^{3\times 3}$ der 3×3-Matrizen über dem Körper K. Es sei $F \subseteq V^*$ der Unterraum, der durch die Linearformen

$$a_{12} + a_{21}, \quad a_{13} + a_{31}, \quad a_{23} + a_{32}, \quad a_{11}, \quad a_{22} \quad \text{und} \quad a_{33}$$

aufgespannt wird. Dann besteht der Prä-Annihilator ${}^0F$ gerade aus allen Elementen von $K^{3\times 3}$, sodass diese sechs Linearformen (und damit auch alle Linearkombinationen) dort verschwinden, also

$${}^0F = \left\{ A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in K^{3\times 3} \,\middle|\, \begin{matrix} a_{12} + a_{21} = 0, & a_{11} = 0 \\ a_{13} + a_{31} = 0, & a_{22} = 0 \\ a_{23} + a_{32} = 0, & a_{33} = 0 \end{matrix} \right\}.$$

Wir betrachten wieder den Fall $\operatorname{char}(K) \neq 2$, dann gilt ${}^0F = K^{3\times 3}_{\text{skew}}$ (die antisymmetrischen 3×3-Matrizen). Wählen wir als Basis des Dualraumes V^*

$$\begin{aligned} B^* = \big(&A \mapsto a_{11},\ A \mapsto a_{22},\ A \mapsto a_{33},\ A \mapsto a_{12} + a_{21},\ A \mapsto a_{13} + a_{31},\ A \mapsto a_{23} + a_{32}\big) \\ &\parallel \big(A \mapsto a_{12} - a_{21},\ A \mapsto a_{13} - a_{31},\ A \mapsto a_{23} - a_{32}\big), \end{aligned}$$

so spannen die ersten $\dim({}^0F) = 6$ Basisvektoren den Unterraum ${}^0F$ auf. Diejenige Basis B von V, sodass B^* die zugehörige duale Basis ist, lautet dann

$$\begin{aligned} B = \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \right) \\ \parallel \left(\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \right). \end{aligned}$$

Nach Satz 20.23 spannen also die letzten $\dim({}^0F) = \dim(V) - \dim(F) = 9 - 6 = 3$ Basisvektoren von B den Prä-Annihilator ${}^0F$ auf. △

Das folgende Lemma zeigt, dass jeder Unterraum ein Prä-Annihilator ist, nämlich der Prä-Annihilator seines Annihilators.

Lemma 20.25 (Unterräume sind Prä-Annihilatoren[AoC 11]**).**
Es sei V ein Vektorraum über dem Körper K. Für jeden Unterraum $U \subseteq V$ gilt

$$U = \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*) = {}^0(U^0). \tag{20.17}$$

Beweis. Nach Definition des Prä-Annihilators (20.13) gilt

$${}^0(U^0) = \left\{ v \in V \,\middle|\, \langle v^*, v \rangle_{V^*,V} = 0 \text{ für alle } v^* \in U^0 \right\} = \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*).$$

Es ist also nur noch das erste Gleichheitszeichen in (20.17) zu zeigen.

Schritt 1: Wir zeigen zunächst $U \subseteq \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*)$:
Für beliebiges $v^* \in U^0$ gilt nach Definition (20.12) $\langle v^*, u \rangle = 0$ für alle $u \in U$, also $U \subseteq \operatorname{Kern}(v^*)$. Da v^* beliebig war, folgt $U \subseteq \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*)$.

Schritt 2: Umgekehrt zeigen wir nun $\bigcap_{v^* \in U^0} \operatorname{Kern}(v^*) \subseteq U$, und zwar durch Kontraposition:
Es sei also $u \in V$, aber $u \notin U$. Wir beginnen mit einer Basis $\widehat{B} = (v_i)_{i \in \widehat{I}}$ von U.[AoC] Weil u nicht in U liegt, bleibt $\widehat{B}$ nach Hinzufügen von u eine linear unabhängige Familie. Wir ergänzen nun[AoC] $\widehat{B}$ plus u zu einer Basis $B = (v_i)_{i \in I}$ von V.
Nun definieren wir eine Linearform $v^* \in V^*$ durch Festlegung der Werte auf der Basis B (Satz 17.10). Wir setzen $\langle v^*, u \rangle \coloneqq 1$ und $\langle v^*, v_i \rangle = 0$ für alle anderen Basiselemente. Insbesondere ist also v^* die Nullform auf U, also $v^* \in U^0$. Aber wegen $v^*(u) = 1$ gilt $u \notin \operatorname{Kern}(v^*)$, also auch $u \notin \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*)$. □

Satz 20.23 und Beispiel 20.24 haben gezeigt, wie wir den Annihilator eines Unterraumes bestimmen können, wenn wir eine an den Unterraum angepasste Basis und die zugehörige duale Basis verwenden (und analog für Prä-Annihilatoren). Sofern solche Basen nicht vorliegen, so können wir auch eine beliebige Basis und zugehörige duale Basis verwenden, mit Koordinatenvektoren arbeiten und so die Aufgabe der Bestimmung von Annihilatoren und Prä-Annihilatoren im Standardvektorraum K^n ausführen:

Lemma 20.26 (Bestimmung von Annihilatoren und Prä-Annihilatoren mit Hilfe von Koordinatenvektoren)**.**
Es sei V ein endlich-dimensionaler Vektorraum mit Basis $B = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$, und zugehöriger dualer Basis $B^* = (v_1^*, \ldots, v_n^*)$.

[11] Die Aussage $U \supseteq {}^0(U^0)$ hängt im Fall, dass V unendlich-dimensional ist, über den Basisexistenzsatz 13.6 und den Basisergänzungssatz 13.8 vom Zornschen Lemma und damit vom Auswahlaxiom ab.

(*i*) Es seien $U \subseteq V$ ein Unterraum und $X \coloneqq \Phi_B^{-1}(U) \subseteq K^n$ der Unterraum der zugehörigen Koordinatenvektoren. Dann gilt für den Annihilator U^0

$$U^0 = \Phi_{B^*}(X^0) \quad \text{mit} \quad X^0 = \left\{\xi \in K^n \,\middle|\, \xi^\intercal x = 0 \text{ für alle } x \in X\right\}.$$

(Hier identifizieren wir den Annihilator X^0 von $X \subseteq K^n$ wie in Bemerkung 20.16 mit einem Unterraum von K^n.)

(*ii*) Es seien $F \subseteq V^*$ ein Unterraum und $\Xi \coloneqq \Phi_{B^*}^{-1}(F) \subseteq K^n$ der Unterraum der zugehörigen Koordinatenvektoren. Dann gilt für den Prä-Annihilator ${}^0F$

$${}^0F = \Phi_B({}^0\Xi) \quad \text{mit} \quad {}^0\Xi = \left\{x \in K^n \,\middle|\, \xi^\intercal x = 0 \text{ für alle } \xi \in \Xi\right\}.$$

(Hier identifizieren wir den Prä-Annihilator ${}^0\Xi$ von $\Xi \subseteq K^n$ ähnlich wie in Bemerkung 20.16 mit einem Unterraum von K^n.)

Beweis. Aussage (*i*): Nach Definition 20.20 gilt $U^0 = \left\{v^* \in V^* \,\middle|\, \langle v^*, u\rangle = 0 \text{ für alle } u \in U\right\}$. Wir stellen Vektoren $u \in U$ bzgl. der Basis B als Koordinatenvektoren $x = \Phi_B^{-1}(U) \in X$ dar und Linearformen $v^* \in V^*$ bzgl. der dualen Basis B^* als Koordinatenvektoren $\xi = \Phi_{B^*}^{-1}(v^*) \in K^n$. Es gilt $\langle v^*, u\rangle = \xi^\intercal x$ für alle $u \in U$ und alle $v^* \in V^*$. Es folgt

$$U^0 = \left\{v^* \in V^* \,\middle|\, \langle v^*, u\rangle = 0 \text{ für alle } u \in U\right\} = \left\{\Phi_{B^*}(\xi) \in V^* \,\middle|\, \xi^\intercal x = 0 \text{ für alle } x \in X\right\}.$$

Der Beweis von Aussage (*ii*) ist Übung. □

Beispiel 20.27 (Bestimmung von Annihilatoren und Prä-Annihilatoren mit Hilfe von Koordinatenvektoren)**.**

(*i*) Wir betrachten einen dreidimensionalen Vektorraum (einen Unterraum von $\mathbb{Q}^\mathbb{Q}$), der durch die Basis $B = (t \mapsto 1,\ t \mapsto t,\ t \mapsto t^2,\ t \mapsto t^3)$ aufgespannt wird. Die zugehörige duale Basis ist

$$B^* = \left(p \mapsto p(0),\ p \mapsto \frac{1}{2}p(1) - \frac{1}{2}p(-1),\ p \mapsto \frac{1}{2}p(1) + \frac{1}{2}p(-1) - p(0)\right),$$

siehe Beispiel 20.13. Gesucht ist der Annihilator U^0 desjenigen Unterraumes U, dessen Vektoren Koordinaten im Unterraum $X = \left\langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \right\rangle$ besitzen. (U besteht also aus den Funktionen mit der Darstellung $\alpha_1\,(1+t) + \alpha_2\, t^2$ für $\alpha_1, \alpha_2 \in \mathbb{Q}$.)

Der Annihilator X^0 besteht aus allen Vektoren $\xi \in \mathbb{Q}^3$, die $\xi^\intercal \begin{pmatrix}1\\1\\0\end{pmatrix} = 0$ sowie $\xi^\intercal \begin{pmatrix}0\\0\\1\end{pmatrix} = 0$ erfüllen. Das ist gerade der Kern der Matrix $\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}$. Da diese schon in reduzierter Zeilenstufenform vorliegt, können wir den Annihilator X^0 sofort ablesen: $\Xi = \left\langle \begin{pmatrix}-1\\1\\0\end{pmatrix} \right\rangle$.

Nach Lemma 20.26 müssen wir diese Koordinatendarstellung nun noch bzgl. der dualen Basis interpretieren. Wir erhalten also

$$U^0 = \Phi_{B^*}(X^0) = \left\langle p \mapsto -p(0) + \frac{1}{2}p(1) - \frac{1}{2}p(-1) \right\rangle.$$

Die Probe ergibt, dass die gefundene Linearform tatsächlich die Funktionen $t \mapsto 1 + t$ sowie $t \mapsto t^2$ und damit auch alle Linearkombinationen annihiliert.

(*ii*) In den gleichen Vektorräumen V und V^* wie oben betrachten wir den Unterraum $F \subseteq V^*$, der durch die Linearform $p \mapsto p(1)$ aufgespannt wird. Bezüglich der dualen Basis entspricht F dem Unterraum $\Xi = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$ von $\mathbb{Q}^3$.

Der Prä-Annihilator ${}^0\Xi$ besteht aus allen Vektoren $x \in \mathbb{Q}^3$, die $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}^\intercal x = 0$ erfüllen. Das ist gerade der Kern der Matrix $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, also ${}^0\Xi = \left\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle$.

Nach Lemma 20.26 müssen wir diese Koordinatendarstellung nun noch bzgl. der primalen Basis interpretieren. Wir erhalten also

$$ {}^0F = \Phi_B({}^0\Xi) = \left\langle t \mapsto -1 + t,\; t \mapsto -1 + t^2 \right\rangle. $$

Die Probe ergibt, dass die gefundenen Funktionen $t \mapsto -1+t$ und $t \mapsto -1+t^2$ tatsächlich von der Linearform $p \mapsto p(1)$ annihiliert werden. △

Beachte: Koeffizientenvektoren für primale und duale Vektoren sehen formal gleich aus. Dennoch ist es wichtig zu unterscheiden, ob sie Vektoren in V oder in V^* repräsentieren!

Abschließend betrachten wir, welche Darstellung der Annihilator einer Teilmenge von K^n bzw. der Prä-Annihilator einer Teilmenge von $(K^n)^*$ hat, wenn wir wie in Bemerkung 20.16 $(K^n)^*$ mit K^n identifizieren:

Beispiel 20.28 ((Prä-)Annihilator von Teilmengen in K^n bzw. $(K^n)^*$)**.**

(*i*) Ist $M \subseteq K^n$, dann gilt

$$ M^0 = \left\{ \xi \in K^n \,\middle|\, \xi^\intercal x = 0 \text{ für alle } x \in M \right\}. \tag{20.18a} $$

(*ii*) Wird $F \subseteq K^n$ interpretiert als Teilmenge von $(K^n)^*$, dann gilt

$$ {}^0F = \left\{ x \in K^n \,\middle|\, \xi^\intercal x = 0 \text{ für alle } \xi \in F \right\}. \tag{20.18b} $$

△

Beachte: Das Beispiel 20.28 zeigt, dass die Konzepte von Annihilator und Prä-Annihilator formal identisch aussehen. Dennoch ist es auch hier wichtig zu unterscheiden, ob eine Menge von Vektoren in K^n tatsächlich Objekte in K^n oder im Dualraum $(K^n)^*$ repräsentiert.

Bemerkung 20.29 (homogene lineare Gleichungssysteme als Prä-Annihilatoren)**.**
Die Gleichung (20.18b) weist auf einen Zusammenhang zwischen Prä-Annihilatoren und homogenen linearen Gleichungssystemen hin: Für $A \in K^{n \times m}$ hat die Bedingung $x \in \operatorname{Kern}(A)$ die Gestalt

$$ Ax = \begin{bmatrix} \text{———} & a_{1\bullet} & \text{———} \\ & \vdots & \\ \text{———} & a_{n\bullet} & \text{———} \end{bmatrix} x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{20.19} $$

Das lässt sich auch so lesen, dass die durch die Zeilenvektoren $a_{1\bullet}, \ldots, a_{n\bullet}$ von A definierten Linearformen $K^m \ni x \mapsto a_{i\bullet} x \in K$ auf x allesamt verschwinden. Mit anderen Worten, x liegt im Prä-Annihilator des von den (transponierten) Zeilenvektoren von A aufgespannten Unterraumes, also $x \in {}^0\operatorname{Bild}(A^\intercal)$! Wir kommen auf diese Beobachtung in Folgerung 21.12 zurück. △

Ende der Vorlesung 2

§ 21 Duale Homomorphismen

Zu jeder linearen Abbildung $f \in \operatorname{Homo}(V, W)$ gibt es eine zugehörige **duale Abbildung** $f^* \in \operatorname{Homo}(W^*, V^*)$, die wir nun definieren und untersuchen:

Definition 21.1 (dualer Homomorphismus)**.**

Es seien V und W Vektorräume über demselben Körper. Weiter sei $f \in \operatorname{Homo}(V, W)$ eine lineare Abbildung. Dann heißt

$$f^* \colon W^* \ni w^* \mapsto v^* \coloneqq w^* \circ f \in V^* \tag{21.1}$$

der zu f **duale** (englisch: dual homomorphism) oder **transponierte Homomophismus** (englisch: transpose homomorphism) bzw. die zu f **duale** oder **transponierte lineare Abbildung** (englisch: dual linear map, transpose linear map).[12] △

Die zu f duale Abbildung f^* ist also nichts Anderes als eine Abkürzung für den Ausdruck „$\cdot \circ f$“, in den man an der Stelle $\cdot$ Elemente aus W^* einsetzen kann und dadurch Elemente aus V^* erhält. Die lineare Abbildung $f \in \operatorname{Homo}(V, W)$ wird benutzt, um Linearformen w^* auf W zu Linearformen v^* auf V „zurückzuziehen“. Man bezeichnet daher das Ergebnis $v^* = w^* \circ f$ auch als den **Pullback** (englisch: pullback) von w^* durch f.

$$\begin{tikzcd} K & W \arrow[l, "w^*"'] & V \arrow[l, "f"'] \arrow[ll, bend left, "v^* = w^* \circ f"] \end{tikzcd}$$

Weil nach Definition 21.1 $f^*(w^*) = w^* \circ f$ gilt, ist $f^* \colon W^* \to V^*$ diejenige Abbildung, die durch die Bedingung $f^*(w^*)(v) = (w^* \circ f)(v)$ für alle $v \in V$ definiert wird. In der Schreibweise mit der dualen Paarung wird f^* also durch die Bedingung

$$\langle f^*(w^*) \, , v \rangle_{V^*,V} = \langle w^* \, , f(v) \rangle_{W^*,W} \tag{21.2}$$

für alle $v \in V$ und alle $w^* \in W^*$ festgelegt.

Wir zeigen nun, dass die duale Abbildung f^* wieder linear ist, sodass die Bezeichnungen aus Definition 21.1 gerechtfertigt sind.

Lemma 21.2 (die duale Abbildung ist wieder linear)**.**

Es seien V und W Vektorräume über demselben Körper. Weiter sei $f \in \operatorname{Homo}(V, W)$ eine lineare Abbildung. Dann ist $f^* \in \operatorname{Homo}(W^*, V^*)$.

[12] Die Notation für die duale Abbildung ist in der Literatur nicht einheitlich. Man findet auch Bezeichnungen wie $f^\intercal$, $f^\star$, f', $f^\vee$ und weitere. Manche Autoren, vor allem in Texten zur Funktionalanalysis (Analysis in unendlich-dimensionalen Vektorräumen) benutzen auch den Begriff **adjungierter Homomorphismus** oder **adjungierte lineare Abbildung** (englisch: adjoint homomorphism, adjoint linear map). Wir verwenden das Attribut „adjungiert“ später aber für etwas anderes.

🅭🅯🄏

Beweis. f^* ist additiv, denn es gilt

$$
\begin{aligned}
f^*(w_1^* + w_2^*) &= (w_1^* + w_2^*) \circ f && \text{nach Definition von } f^* \\
&= w_1^* \circ f + w_2^* \circ f && \text{nach Definition der Addition von Abbildungen} \\
& && \text{von Homomorphismen (Satz 17.14)} \\
&= f^*(w_1) + f^*(w_2) && \text{nach Definition von } f^*.
\end{aligned}
$$

Außerdem ist f^* homogen wegen

$$
\begin{aligned}
f^*(\alpha\, w^*) &= (\alpha\, w^*) \circ f && \text{nach Definition von } f^* \\
&= \alpha\,(w^* \circ f) && \text{nach Definition der S-Multiplikation von Abbildungen} \\
& && \text{von Homomorphismen (Satz 17.14)} \\
&= \alpha\, f^*(w^*) && \text{nach Definition von } f^*.
\end{aligned}
$$ □

Beispiel 21.3 (dualer Homomorphismus)**.**

(i) Es sei V ein Vektorraum. Dann gilt $\mathrm{id}_V^* = \mathrm{id}_{V^*}$, denn id_V^* ist eindeutig festgelegt durch die Bedingungen

$$\big\langle \mathrm{id}_V^*(v^*)\,,v\big\rangle_{V^*,V} = \big\langle v^*\,, \mathrm{id}_V(v)\big\rangle_{V^*,V} = \langle v^*\,,v\rangle_{V^*,V}$$

für alle $v^* \in V^*$ und alle $v \in V$. Diese werden von $\mathrm{id}_V^* = \mathrm{id}_{V^*}$ erfüllt.

(ii) Es seien K ein Körper, $V = K^n$ und $\pi_i \colon K^n \to K$ die Projektion auf die i-te Koordinate, siehe (17.2). Dann ist $\pi_i^* \colon K^* \to (K^n)^*$ definiert durch die Bedingungen

$$\langle \pi_i^*(y^*)\,,x\rangle_{(K^n)^*,K^n} = \langle y^*\,,\pi_i(x)\rangle_{K^*,K} = \langle y^*\,,x_i\rangle_{K^*,K}$$

für alle $x \in K^n$ und $y^* \in K^*$.

Durch die Identifikation von K^* mit K und von $(K^n)^*$ mit K^n wie in Bemerkung 20.16 können wir π_i^* als Abbildung $K \to K^n$ auch darstellen durch

$$\pi_i^*(y) = \begin{array}{c@{}l} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y \\ 0 \\ \vdots \\ 0 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{array} \end{array} \quad \text{mit } y \in K \text{, denn es gilt} \quad \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y \\ 0 \\ \vdots \\ 0 \end{pmatrix}^{\mathsf{T}} x = y\,x_i.$$

(iii) Es sei $V = K^{[\![1,5]\!]}$ der Vektorraum der endlichen Folgen der Länge 5 über einem Körper K und

$$f \colon V \ni (y_1, y_2, y_3, y_4, y_5) \mapsto (0, y_1, y_2, y_3, y_4) \in W = V$$

die Shift-Abbildung nach rechts (Beispiel 19.4). Dann ist die duale Abbildung $f^* \colon W^* \ni w^* \mapsto v^* \coloneqq f^*(w^*) \in V^*$ gegeben durch die Bedingungen

$$\big\langle v^*\,, (y_1, y_2, y_3, y_4, y_5)\big\rangle_{V^*,V} = \big\langle w^*\,, (0, y_1, y_2, y_3, y_4)\big\rangle_{W^*,W}$$

für beliebige $(y_1, y_2, y_3, y_4, y_5) \in V$. Beispielsweise für die durch

$$\big\langle w^*\,, (z_1, z_2, z_3, z_4, z_5)\big\rangle = z_3 + z_4$$

definierte Linearform w^* ist $v^* = f^*(w^*)$ gegeben durch

$$\big\langle v^*\,, (y_1, y_2, y_3, y_4, y_5)\big\rangle = \big\langle w^*\,, (0, y_1, y_2, y_3, y_4)\big\rangle = y_2 + y_3.$$ △

Lemma 21.4 (Dualisieren der Synthese-Abbildung)**.**
Es sei V ein endlich-dimensionaler Vektorraum mit Basis $B = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$, und zugehöriger dualer Basis $B^* = (v_1^*, \ldots, v_n^*)$. Bei Identifikation von $(K^n)^*$ mit K^n gilt:

(*i*) Die duale Abbildung der Synthese-Abbildung $\Phi_B \colon K^n \to V$ ist die Analyse-Abbildung $\Phi_{B^*}^{-1} \colon V^* \to K^n$, also $\Phi_B^* = \Phi_{B^*}^{-1}$.

(*ii*) Die duale Abbildung der Analyse-Abbildung $\Phi_B^{-1} \colon V \to K^n$ ist die Synthese-Abbildung $\Phi_{B^*} \colon K^n \to V^*$, also $\Phi_B^{-*} = \Phi_{B^*}$.

Beweis. Wir beweisen zunächst Aussage (*i*). Es seien $x \in K^n$ und $v \coloneqq \Phi_B(x) \in V$ sowie $\xi \in K^n$ und $v^* \coloneqq \Phi_{B^*}(\xi) \in V^*$. Dann gilt

$$\begin{aligned}
\Phi_B^*(v^*)^\intercal x &= \langle \Phi_B^*(v^*)\,, x\rangle_{(K^n)^*,K^n} && \text{gemäß der Identifikation von } (K^n)^* \text{ mit } K^n \\
&= \langle v^*\,, \Phi_B(x)\rangle_{V^*,V} && \text{nach Definition (21.2) der dualen Abbildung} \\
&= \langle v^*\,, v\rangle_{V^*,V} && \text{wegen } v = \Phi_B(x) \\
&= \xi^\intercal x && \text{nach Lemma 20.14.}
\end{aligned}$$

Da $x \in K^n$ beliebig war, muss $\Phi_B^*(v^*) = \xi = \Phi_{B^*}^{-1}(v^*)$ gelten, und da auch $v^* \in V^*$ beliebig war, folgt $\Phi_B^* = \Phi_{B^*}^{-1}$.

Aussage (*ii*) folgt einfach durch Invertierung der Abbildungen aus Aussage (*i*). □

Lemma 21.5 (Dualisieren einer Komposition von Homomorphismen)**.**
Es seien U, V und W Vektorräume über demselben Körper. Ist $f \in \operatorname{Homo}(V, W)$ und $g \in \operatorname{Homo}(U, V)$, dann gilt für die duale Abbildung der Komposition $f \circ g \in \operatorname{Homo}(U, W)$:

$$(f \circ g)^* = g^* \circ f^*. \tag{21.3}$$

Beweis. Der Beweis ist Gegenstand der Übung. □

Wir zeigen nun, dass verschiedene Homomorphismen auch verschiedene duale Abbildungen besitzen.

Satz 21.6 (die Dualisierungsabbildung ist ein injektiver[AoC] Vektorraumhomomorphismus)**.**
Es seien V und W Vektorräume über demselben Körper.

(*i*) Die **Dualisierungsabbildung** (englisch: dualization map), also die Zuordnung

$$D\colon\ \operatorname{Homo}(V, W) \ni f \mapsto f^* \in \operatorname{Homo}(W^*, V^*)$$

einer linearen Abbildung zu ihrer dualen Abbildung, ist ein injektiver[AoC] Homomorphismus von Vektorräumen.[13]

[13] Die Injektivität („Zwei verschiedene Homomorphismen haben auch verschiedene duale Homomorphismen.“) hängt im Fall, dass W unendlich-dimensional ist, über die letzte Aussage ${}^0(W^*) = \{0_W\}$ aus Lemma 20.21 vom Zornschen Lemma und damit vom Auswahlaxiom ab.

(*ii*) Sind V und W endlich-dimensional, dann ist D auch surjektiv, also ein Isomorphismus von Vektorräumen. Es gilt dann

$$\dim(\operatorname{Homo}(W^*, V^*)) = \dim(\operatorname{Homo}(V, W)) = \dim(V)\ \dim(W). \tag{21.4}$$

Beweis. Aussage (*i*):

Schritt 1: Wir untersuchen zunächst die Linearität von D.

D ist additiv, denn für $f, g \in \operatorname{Homo}(V, W)$ und $w^* \in W^*$ gilt

$$\begin{aligned}\big(D(f+g)\big)(w^*) &= w^* \circ (f+g) && \text{nach Definition von } D\\ &= w^* \circ f + w^* \circ g && \text{denn } w^* \text{ ist linear}\\ &= \big(D(f)\big)(w^*) + \big(D(g)\big)(w^*) && \text{nach Definition von } D\\ &= \big(D(f) + D(g)\big)(w^*).\end{aligned}$$

Außerdem ist D homogen, denn es gilt

$$\begin{aligned}\big(D(\alpha f)\big)(w^*) &= w^* \circ (\alpha f) && \text{nach Definition von } D\\ &= \alpha\, w^* \circ (f) && \text{denn } w^* \text{ ist linear}\\ &= \alpha \big(D(f)(w^*)\big) && \text{nach Definition von } D\\ &= \big((\alpha D)(f)\big)(w^*).\end{aligned}$$

Schritt 2: Zum Nachweis der Injektivät untersuchen wir $\operatorname{Kern}(D)$.

Die Eigenschaft $f \in \operatorname{Kern}(D)$ heißt, dass $f^* = D(f) \in \operatorname{Homo}(W^*, V^*)$ die Nullabbildung ist, d. h.

$$f^*(w^*) = w^* \circ f = 0_{V^*}$$

für alle $w^* \in W^*$, oder auch

$$w^*(f(v)) = 0$$

für alle $w^* \in W^*$ und alle $v \in V$. Anders ausgedrückt, es gilt (vgl. (20.13))

$$f(V) \subseteq \bigcap_{w^* \in W^*} \operatorname{Kern}(w^*) = {}^0(W^*).$$

Nach Lemma 20.21 gilt aber (falls W unendlich-dimensional ist unter der Annahme des Auswahlaxioms) ${}^0(W^*) = \{0_W\}$. Das heißt aber, f ist die Nullabbildung. Also gilt $\operatorname{Kern}(D) = \{0_W\}$, und D ist nach Lemma 17.9 injektiv.

Aussage (*ii*): Wenn V und W endlich-dimensional sind, dann ist auch $\operatorname{Homo}(V, W)$ endlich-dimensional, und es gilt nach Folgerung 19.7 $\dim(\operatorname{Homo}(V, W)) = \dim(V)\ \dim(W)$. Außerdem gilt wegen Folgerung 20.9 auch $\dim(V) = \dim(V^*)$ und $\dim(W) = \dim(W^*)$, also

$$\dim(\operatorname{Homo}(V, W)) = \dim(V)\ \dim(W) = \dim(V^*)\ \dim(W^*) = \dim(\operatorname{Homo}(W^*, V^*)).$$

Die Dualisierungsabbildung D bildet also zwischen zwei Vektorräumen derselben endlichen Dimension ab. Da D in Aussage (*i*) bereits als injektiv gezeigt wurde (wofür wegen der endlichen Dimension von W das Auswahlaxiom nicht benötigt wird), ergibt sich aus Folgerung 18.9, dass D auch surjektiv ist, also ein Isomorphismus. □

Schließlich erhalten wir folgende Aussage über den Zusammenhang von Injektivität und Surjektivität von Homomorphismen und den zugehörigen dualen Abbildungen:

Satz 21.7 (Injektivität und Surjektivität des dualen Homomorphismus[AoC 14]).
Es seien V und W zwei Vektorräume über demselben Körper. Weiter sei $f \in \operatorname{Homo}(V, W)$ und $f^* \in \operatorname{Homo}(W^*, V^*)$ die zugehörige duale Abbildung. Dann gilt:

(*i*) $f \in \operatorname{Homo}(V, W)$ ist genau dann injektiv, wenn $f^* \in \operatorname{Homo}(W^*, V^*)$ surjektiv ist.[15]

(*ii*) $f \in \operatorname{Homo}(V, W)$ ist genau dann surjektiv, wenn $f^* \in \operatorname{Homo}(W^*, V^*)$ injektiv ist.[16]

(*iii*) $f \in \operatorname{Homo}(V, W)$ ist genau dann bijektiv, wenn $f^* \in \operatorname{Homo}(W^*, V^*)$ bijektiv ist.[17]

(*iv*) Falls f und f^* beide bijektiv sind, gilt $(f^*)^{-1} = (f^{-1})^*$.
Aus diesem Grund können wir statt $(f^*)^{-1}$ auch einfach f^{-*} schreiben.

Beweis. Wir zeigen zunächst Aussage (*i*).

Schritt 1: Es sei zunächst f injektiv. Wir definieren $U \coloneqq \operatorname{Bild}(f) \subseteq W$. Dann ist die Einschränkung $f|^U : V \to U$ bijektiv. Die Umkehrabbildung sei $g \in \operatorname{Homo}(U, V)$. Es sei weiter $\widehat{U}$ ein zu U komplementärer Unterraum von W.[AoC]

Wir wollen zeigen, dass $f^* \in \operatorname{Homo}(W^*, V^*)$ surjektiv ist. Es sei dazu $v^* \in V^*$ beliebig. Wir definieren das Element $w^* \in W^*$ dadurch, dass wir auf dem Unterraum U die Funktionswerte

$$w^*(u) \coloneqq \langle v^*, g(u) \rangle \in K$$

setzen und mit $w^*(u) \coloneqq 0$ für $u \in \widehat{U}$ ergänzen. (**Quizfrage 21.1:** Warum definiert das w^* eindeutig?)

Damit gilt nun für alle $v \in V$

$$\begin{aligned}
\langle f^*(w^*), v \rangle &= \langle w^*, f(v) \rangle && \text{nach Definition von } f^* \\
&= \langle v^*, g(f(v)) \rangle && \text{nach Definition von } w^* \text{ und wegen } f(v) \in U \\
&= \langle v^*, v \rangle && \text{da } g \text{ die Umkehrabbildung von } f|^U \text{ ist.}
\end{aligned}$$

Das bedeutet aber gerade $f^*(w^*) = v^*$, also ist f^* surjektiv.

Schritt 2: Umgekehrt sei nun $f^* \in \operatorname{Homo}(W^*, V^*)$ surjektiv. Wir führen einen Widerspruchsbeweis und nehmen an, f sei nicht injektiv. Dann besteht also $\operatorname{Kern}(f)$ aus mehr als dem Nullunterraum von V. Unter Verwendung des Basisergänzungssatzes 13.5[AoC] können wir leicht zeigen, dass es eine Linearform $v^* \in V^*$ gibt, deren Einschränkung

[14] Dieses Resultat hängt im Fall, dass V bzw. W unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab. Die genauen Abhängigkeiten sind bei den einzelnen Aussagen angegeben.

[15] Die Aussage „$\Rightarrow$“ hängt im Fall, dass W unendlich-dimensional ist, über die Folgerung 14.10 zur Existenz eines komplementären Unterraumes vom Auswahlaxiom ab. Die Aussage „$\Leftarrow$“ hängt im Fall, dass V unendlich-dimensional ist, über den Basisergänzungssatz 13.5 vom Auswahlaxiom ab.

[16] Die Aussage „$\Leftarrow$“ hängt im Fall, dass W unendlich-dimensional ist, über die Folgerung 14.10 zur Existenz eines komplementären Unterraumes vom Auswahlaxiom ab.

[17] Interessanterweise kann die Aussage „$\Rightarrow$“ auch ohne Verwendung des Auswahlaxioms gezeigt werden. Die Aussage „$\Leftarrow$“ hängt im Fall, dass V oder W unendlich-dimensional sind, wie in den Aussagen (*i*) und (*ii*) vom Auswahlaxiom ab.

auf $\operatorname{Kern}(f)$ nicht die Nullform ist. (**Quizfrage 21.2:** Wie geht das genau?) Das bedeutet aber, dass v^* nicht von der Form $f^*(w^*) = w^* \circ f$ sein kann für irgendein $w^* \in W^*$, denn jede Abbildung der Form $w^* \circ f$ ist auf $\operatorname{Kern}(f)$ die Nullform. Das steht im Widerspruch zur angenommenen Surjektivität von f^*. Also muss f injektiv sein.

Nun zu Aussage *(ii)*.

Schritt 1: Es sei zunächst f surjektiv. Es seien $w_1^*, w_2^* \in W^*$ mit der Eigenschaft $f^*(w_1^*) = f^*(w_2^*)$ gegeben, d. h., es gilt $w_1^* \circ f = w_2^* \circ f$.

Es sei $w \in W$. Da f surjektiv ist, existiert ein $v \in V$ mit $w = f(v)$. Wir haben also

$$\langle w_1^*, w\rangle = \langle w_1^*, f(v)\rangle = (w_1^* \circ f)(v) = (w_2^* \circ f)(v) = \langle w_2^*, f(v)\rangle = \langle w_2^*, w\rangle$$

für alle $w \in W$ und damit $w_1^* = w_2^*$. Das heißt aber, f^* ist injektiv.

Schritt 2: Umgekehrt sei nun $f^* \in \operatorname{Homo}(W^*, V^*)$ injektiv. Wir führen einen Widerspruchsbeweis und nehmen an, f sei nicht surjektiv. Wir definieren $U \coloneqq \operatorname{Bild}(f) \subseteq W$. Es sei weiter $\widehat{U}$ ein zu U komplementärer Unterraum von W.[AoC] Da f nicht surjektiv ist, ist $\widehat{U}$ nicht der Nullraum.

Wie im Beweis von Aussage *(i)* können mit Hilfe des Basisergänzungssatzes 13.5 eine Linearform $w^* \in W^*$ konstruieren mit der Eigenschaft $w^*|_U = 0$ (Nullform) und $w^*|_{\widehat{U}} \neq 0$. Dann gilt aber $f^*(w^*) = w^* \circ f = 0_{V^*}$, also liegt w^* in $\operatorname{Kern}(f^*)$, im Widerspruch zur angenommenen Injektivität von f^*. Also muss f surjektiv sein.

Aussage *(iii)* ist eine Kombination der Aussagen *(i)* und *(ii)*. Die Richtung „$\Rightarrow$“ kann jedoch mit einem anderen Beweis auch ohne Rückgriff auf das Auswahlaxiom gezeigt werden: Wenn $f \in \operatorname{Homo}(V, W)$ bijektiv ist, dann existiert die Umkehrkehrfunktion $f^{-1} \in \operatorname{Homo}(W, V)$. Wir betrachten deren duale Abbildung $(f^{-1})^* \in \operatorname{Homo}(V^*, W^*)$. Für diese gilt

$$\begin{aligned} f^* \circ (f^{-1})^* &= (f^{-1} \circ f)^* && \text{nach Lemma 21.5} \\ &= \mathrm{id}_V^* \\ &= \mathrm{id}_{V^*} && \text{nach Beispiel 21.3} \end{aligned}$$

und außerdem

$$\begin{aligned} (f^{-1})^* \circ f^* &= (f \circ f^{-1})^* && \text{nach Lemma 21.5} \\ &= \mathrm{id}_W^* \\ &= \mathrm{id}_{W^*} && \text{nach Beispiel 21.3.} \end{aligned}$$

Das zeigt, dass f^* bijektiv und $(f^{-1})^*$ die Inverse von f^* ist.

Schließlich zu Aussage *(iv)*: Sind f und f^* beide bijektiv, dann zeigt die obige Rechnung, dass $(f^{-1})^*$ die Inverse von f^* ist. □

§ 21.1 Darstellungsmatrizen dualer Homomorphismen

Wir untersuchen jetzt, wie die Darstellungsmatrizen einer linearen Abbildung $f \in \operatorname{Homo}(V, W)$ und ihrer dualen linearen Abbildung $f^* \in \operatorname{Homo}(W^*, V^*)$ zusammenhängen. Dazu seien V und W endlich-dimensionale Vektorräume über demselben Körper K und $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W. Wir erinnern uns, dass die Darstellungsmatrix $A = \mathcal{M}_{B_W \leftarrow B_V}(f) = (a_{ij})$ von f bzgl. dieser Basen durch (19.3) gegeben ist, also

$$f(v_j) = \sum_{i=1}^{n} a_{ij}\, w_i \quad \text{für alle } j = 1, \ldots, m. \tag{21.5a}$$

Für die Darstellungsmatrix von f^* werden wir die dualen Basen verwenden. Die gesuchte Matrix $B = \mathcal{M}_{B_{V^*} \leftarrow B_{W^*}}(f^*) \in K^{m \times n}$ mit Einträgen b_{ij} erfüllt dann

$$f^*(w_j^*) = \sum_{i=1}^{m} b_{ij}\, v_i^* \quad \text{für alle } j = 1, \ldots, n. \tag{21.5b}$$

Um die Einträge der Matrizen A und B zu vergleichen, wenden wir auf beide Seiten von (21.5a) das Element $w_k^* \in W^*$ der zu B_W dualen Basis an, das ja nach Satz 20.10 als Koordinatenermittler wirkt. Es folgt

$$\langle w_k^*, f(v_j)\rangle = \left\langle w_k^*, \sum_{i=1}^{n} a_{ij}\, w_i \right\rangle = \sum_{i=1}^{n} a_{ij}\, \underbrace{\langle w_k^*, w_i\rangle}_{=\delta_{ki}} = a_{kj} \quad \text{für alle } j = 1, \ldots, m. \tag{21.6a}$$

Analog ergibt das Einsetzen des Elements v_k der primalen Basis B_V in (21.5b)

$$\underbrace{\langle f^*(w_j^*), v_k\rangle}_{=\langle w_j^*, f(v_k)\rangle} = \left\langle \sum_{i=1}^{m} b_{ij}\, v_i^*, v_k \right\rangle = \sum_{i=1}^{m} b_{ij}\, \underbrace{\langle v_i^*, v_k\rangle}_{=\delta_{ik}} = b_{kj} \quad \text{für alle } j = 1, \ldots, n. \tag{21.6b}$$

Nach Definition von f^* ist die linke Seite von (21.6b) aber auch gleich $\langle w_j^*, f(v_k)\rangle$. Ein Vergleich der linken Seiten von (21.6a) und (21.6b) zeigt also den Zusammenhang $b_{kj} = a_{jk}$, also $B = A^\mathsf{T}$!

Wir haben damit den folgenden Satz bewiesen:

Satz 21.8 (Darstellungsmatrix des dualen Homomorphismus)**.**
Es seien V und W endlich-dimensionale Vektorräume über demselben Körper und B_V und B_W Basen von V bzw. von W. Weiter sei $f \in \operatorname{Homo}(V, W)$ eine lineare Abbildung mit dualer Abbildung $f^* \in \operatorname{Homo}(W^*, V^*)$. Ist $A = \mathcal{M}_{B_W \leftarrow B_V}(f)$ die Darstellungsmatrix von f bzgl. der Basen B_V und B_W, dann ist $A^\mathsf{T} = \mathcal{M}_{B_{V^*} \leftarrow B_{W^*}}(f^*)$ die Darstellungsmatrix von f^* bzgl. der dualen Basen B_{W^*} und B_{V^*}.

Transponierte Matrizen weisen also darauf hin, dass eine duale Abbildung im Spiel ist.

Folgerung 21.9 (Rang der dualen Abbildung)**.**
Es seien V und W endlich-dimensionale Vektorräume über demselben Körper. Weiter sei $f \in \operatorname{Homo}(V, W)$ eine lineare Abbildung mit dualer Abbildung $f^* \in \operatorname{Homo}(W^*, V^*)$. Dann gilt $\operatorname{Rang}(f^*) = \operatorname{Rang}(f)$.

Beweis. Wir wählen irgendwelche Basen B_V und B_W von V bzw. W. Dann haben die Darstellungsmatrizen $A = \mathcal{M}_{B_W \leftarrow B_V}(f)$ und $A^\intercal = \mathcal{M}_{B_{V^*} \leftarrow B_{W^*}}(f^*)$ denselben Rang (Lemma 15.29). Nach Satz 19.10 haben wir also $\operatorname{Rang}(f) = \operatorname{Rang}(A) = \operatorname{Rang}(A^\intercal) = \operatorname{Rang}(f^*)$. □

Folgerung 21.10 (duale Abbildung der von einer Matrix induzierten linearen Abbildung)**.**

Es seien K ein Körper und $A \in K^{n \times m}$. Bei Identifikation von $(K^n)^*$ bzw. $(K^m)^*$ mit K^n bzw. K^m wie in Bemerkung 20.16 gilt für die duale Abbildung $(f_A)^* \in \operatorname{Homo}(K^n, K^m)$ der durch A induzierten Abbildung $f_A \in \operatorname{Homo}(K^m, K^n)$

$$(f_A)^* = f_{A^\intercal}. \tag{21.7}$$

Beweis. Die Abbildung auf der linken Seite in (21.7) ist wegen der Identifikation von K^n und K^m mit dem jeweiligen Dualraum beschrieben durch die Bedingung

$$\big[(f_A)^*(y)\big]^\intercal x = y^\intercal f_A(x) = y^\intercal A\, x \quad \text{für alle } x \in K^m \text{ und } y \in K^n.$$

Das heißt aber $(f_A)^*(y) = A^\intercal y$ für alle $y \in K^n$, was (21.7) zeigt. □

Abschließend untersuchen wir noch, wie sich die Darstellungsmatrix eines dualen Homomorphismus bei einem Basiswechsel transformiert. Wir kennen aus Lemma 20.18 bereits den Zusammenhang[18]

$$\mathcal{T}_{\widehat{B}_V^* \leftarrow B_{V^*}} = [\mathcal{T}_{B_V \leftarrow \widehat{B}_V}]^\intercal = [\mathcal{T}_{\widehat{B}_V \leftarrow B_V}]^{-\intercal}$$

der Transformationsmatrizen für Koeffizientenvektoren bei einem Basiswechsel im Raum V und dem zugehörigen Wechsel der dualen Basen im Raum V^*. Während für die Transformation der Darstellungsmatrizen eines Homomorphismus $f \in \operatorname{Homo}(V, W)$ nach (19.14)

$$\begin{aligned} \mathcal{M}_{\widehat{B}_W \leftarrow \widehat{B}_V}(f) &= \mathcal{T}_{\widehat{B}_W \leftarrow B_W}\, \mathcal{M}_{B_W \leftarrow B_V}(f)\, \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \\ \widehat{A} \quad &= \quad S^{-1} \qquad\quad A \qquad\quad T \end{aligned} \tag{21.8a}$$

gilt, erhalten wir für die Transformation des dualen Homomorphismus $f^* \in \operatorname{Homo}(W^*, V^*)$ also die Beziehung

$$\begin{aligned} \mathcal{M}_{\widehat{B}_V^* \leftarrow \widehat{B}_W^*}(f^*) &= \mathcal{T}_{\widehat{B}_V^* \leftarrow B_{V^*}}\, \mathcal{M}_{B_{V^*} \leftarrow B_{W^*}}(f^*)\, \mathcal{T}_{B_{W^*} \leftarrow \widehat{B}_W^*} \\ \widehat{A}^\intercal \quad &= \quad T^\intercal \qquad\quad A^\intercal \qquad\quad S^{-\intercal} \end{aligned} \tag{21.8b}$$

Ende der Vorlesung 3

Ende der Woche 1

[18] Dieser Zusammenhang kann jetzt angesichts $\mathcal{T}_{B_V \leftarrow \widehat{B}_V} = \mathcal{M}_{B_V \leftarrow \widehat{B}_V}(\mathrm{id}_V)$ und damit (nach Satz 21.8) $\mathcal{T}_{\widehat{B}_V^* \leftarrow B_{V^*}} = \mathcal{M}_{\widehat{B}_V^* \leftarrow B_{V^*}}(\mathrm{id}_{V^*}) = [\mathcal{M}_{B_V \leftarrow \widehat{B}_V}(\mathrm{id}_V)]^\intercal = [\mathcal{T}_{B_V \leftarrow \widehat{B}_V}]^\intercal$ nochmal bestätigt werden.

§ 21.2 Die vier fundamentalen Unterräume zu einer linearen Abbildung

In diesem Abschnitt zeigen wir einen wichtigen Zusammenhang zwischen den Unterräumen Kern und Bild von f sowie deren Annihilatoren. Man spricht auch von den **vier fundamentalen Unterräumen** (englisch: four fundamental subspaces) zu einer linearen Abbildung f.

Satz 21.11 (vier fundamentale Unterräume zu einer linearen Abbildung[AoC19]).
Es seien V und W Vektorräume über demselben Körper. Weiter sei $f \in \operatorname{Homo}(V, W)$ eine lineare Abbildung mit dualer Abbildung $f^* \in \operatorname{Homo}(W^*, V^*)$. Dann gilt:

$$\operatorname{Bild}(f^*) = \operatorname{Kern}(f)^0 \quad \text{in } V^*,\text{[20]} \tag{21.9a}$$

$$\operatorname{Kern}(f^*) = \operatorname{Bild}(f)^0 \quad \text{in } W^*, \tag{21.9b}$$

$$\operatorname{Bild}(f) = {}^0\operatorname{Kern}(f^*) \quad \text{in } W,\text{[21]} \tag{21.9c}$$

$$\operatorname{Kern}(f) = {}^0\operatorname{Bild}(f^*) \quad \text{in } V.\text{[22]} \tag{21.9d}$$

Beweis. Wir zeigen zunächst (21.9a). Die beteiligten Unterräume von V^* sind

$$\operatorname{Bild}(f^*) = \{f^*(w^*) \in V^* \mid w^* \in W^*\}$$
$$\operatorname{Kern}(f)^0 = \{v^* \in V^* \mid \langle v^*, v\rangle = 0 \text{ für alle } v \in \operatorname{Kern}(f)\}.$$

Schritt 1: $\operatorname{Bild}(f^*) \subseteq \operatorname{Kern}(f)^0$:

Es sei $v^* \in \operatorname{Bild}(f^*)$, also existiert ein $w^* \in W^*$ mit $v^* = f^*(w^*)$. Für beliebiges $v \in \operatorname{Kern}(f)$ gilt

$$\langle v^*, v\rangle = \langle f^*(w^*), v\rangle = \langle w^*, f(v)\rangle = \langle w^*, 0_W\rangle = 0.$$

Also liegt v^* in $\operatorname{Kern}(f)^0$.

Schritt 2: $\operatorname{Kern}(f)^0 \subseteq \operatorname{Bild}(f^*)$:

Es sei $v^* \in \operatorname{Kern}(f)^0$. Wir definieren ein passendes Element $w^* \in W^*$ mit der Eigenschaft $v^* = f^*(w^*)$. Dazu setzen wir die Werte von w^* zunächst auf dem Unterraum $\operatorname{Bild}(f) \subseteq W$ durch

$$\langle w^*, f(v)\rangle \coloneqq \langle v^*, v\rangle \quad \text{für alle } v \in V$$

fest. Diese Vorschrift ist wohldefiniert, denn $f(v_1) = f(v_2)$ bedeutet $v_1 - v_2 \in \operatorname{Kern}(f)$. Da v^* in $\operatorname{Kern}(f)^0$ liegt, gilt $\langle v^*, v_1 - v_2\rangle = 0$, also $\langle v^*, v_1\rangle = \langle v^*, v_2\rangle$.

Wir können w^* nun beliebig auf ganz W fortsetzen. (**Quizfrage 21.3:** Wie geht das genau?) Schließlich gilt per Definition

$$\langle f^*(w^*), v\rangle = \langle w^*, f(v)\rangle = \langle v^*, v\rangle$$

für alle $v \in V$, also $f^*(w^*) = v^*$.

[19]Dieses Resultat hängt im Fall, dass V bzw. W unendlich-dimensional ist, vom Auswahlaxiom ab. Die genauen Abhängigkeiten sind bei den einzelnen Aussagen angegeben.

Nun adressieren wir (21.9b). Die beteiligten Unterräume von W^* sind

$$\begin{aligned}\operatorname{Kern}(f^*) &= \{w^* \in W^* \mid f^*(w^*) = 0_{V^*}\}\\ &= \{w^* \in W^* \mid w^* \circ f = 0_{V^*}\}\\ \operatorname{Bild}(f)^0 &= \{w^* \in W^* \mid \langle w^*, w\rangle = 0 \text{ für alle } w \in \operatorname{Bild}(f)\}\\ &= \{w^* \in W^* \mid \langle w^*, f(v)\rangle = 0 \text{ für alle } v \in V\}.\end{aligned}$$

Es gilt

$$\begin{aligned}& w^* \in \operatorname{Bild}(f)^0\\ \Leftrightarrow\quad & \langle w^*, f(v)\rangle = 0 \text{ für alle } v \in V\\ \Leftrightarrow\quad & \langle f^*(w^*), v\rangle = 0 \text{ für alle } v \in V\\ \Leftrightarrow\quad & f^*(w^*) = 0_{V^*}\\ \Leftrightarrow\quad & w^* \in \operatorname{Kern}(f^*).\end{aligned}$$

Der Beweis von (21.9c) ist nun einfach. Die beteiligten Unterräume von W sind

$$\begin{aligned}\operatorname{Bild}(f) &= \{f(v) \in W \mid v \in V\}\\ {}^0\operatorname{Kern}(f^*) &= \{w \in W \mid \langle w^*, w\rangle = 0 \text{ für alle } w^* \in \operatorname{Kern}(f^*)\}\\ &= \{w \in W \mid \langle w^*, w\rangle = 0 \text{ für alle } w^* \in W^* \text{ mit } f^*(w^*) = 0_{V^*}\}.\end{aligned}$$

Wir haben

$$\begin{aligned}\operatorname{Bild}(f) &= {}^0\operatorname{Bild}(f)^0 && \text{nach Lemma 20.25}\\ &= {}^0\operatorname{Kern}(f^*) && \text{wegen (21.9b).}\end{aligned}$$

Die Teilaussage $\operatorname{Bild}(f) \supseteq {}^0\operatorname{Bild}(f)^0$ aus Lemma 20.25 hängt dabei im Fall, dass W unendlich-dimensional ist, vom Auswahlaxiom ab.

Um schließlich (21.9d) zu zeigen, betrachten wir die beteiligten Unterräume von V:

$$\begin{aligned}\operatorname{Kern}(f) &= \{v \in V \mid f(v) = 0\}\\ {}^0\operatorname{Bild}(f^*) &= \{v \in V \mid \langle v^*, v\rangle = 0 \text{ für alle } v^* \in \operatorname{Bild}(f^*)\}\\ &= \{v \in V \mid \langle f^*(w^*), v\rangle = 0 \text{ für alle } w^* \in W^*\}\\ &= \{v \in V \mid \langle w^*, f(v)\rangle = 0 \text{ für alle } w^* \in W^*\}.\end{aligned}$$

Wir könnten hier auch mit Hilfe von Lemma 20.25 argumentieren (**Quizfrage 21.4:** Wie nämlich?), würden uns aber im Fall, dass V unendlich-dimensional ist, eine weitere Abhängigkeit vom Auswahlaxiom einhandeln. Stattdessen führen wir den Beweis direkt:

Schritt 1: $\operatorname{Kern}(f) \subseteq {}^0\operatorname{Bild}(f^*)$:

Es sei $v \in \operatorname{Kern}(f)$. Dann gilt

$$\langle f^*(w^*), v\rangle = \langle w^*, f(v)\rangle = \langle w^*, 0_W\rangle = 0$$

für alle $w^* \in W^*$, also $v \in {}^0\operatorname{Bild}(f^*)$.

Schritt 2: ${}^0\mathrm{Bild}(f^*) \subseteq \mathrm{Kern}(f)$:

Es sei $v \in {}^0\mathrm{Bild}(f^*)$. Angenommen $f(v) \neq 0_W$, dann können wir eine Linearform $w^* \in W^*$ finden[AoC] mit der Eigenschaft $\langle w^*, f(v)\rangle \neq 0$, im Widerspruch zu $v \in {}^0\mathrm{Bild}(f^*)$. Also gilt $v \in \mathrm{Kern}(f)$. □

Abschließend betrachten wir eine Version von Satz 21.11, wobei $V = K^m$, $W = K^n$ und $f = f_A$ durch die Matrix-Vektor-Multiplikation mit der Matrix $A \in K^{n\times m}$ gegeben ist. In diesem Fall können wir die Aussage von Satz 21.11 mit Hilfe der Matrizen A und $A^\intercal$ ausdrücken.

Folgerung 21.12 (vier fundamentale Unterräume zu einer Matrix)**.**
Es seien K ein Körper und $A \in K^{n\times m}$. Weiter sei $r \coloneqq \mathrm{Rang}(A)$. Bei Identifikation[23] von $(K^n)^*$ bzw. $(K^m)^*$ mit K^n bzw. K^m wie in Bemerkung 20.16 gilt für die Matrix A und ihre Transponierte $A^\intercal$

$$\mathrm{Bild}(A^\intercal) = \mathrm{Kern}(A)^0 \quad \text{hat Dimension } r \text{ in } K^m, \tag{21.10a}$$
$$\mathrm{Kern}(A^\intercal) = \mathrm{Bild}(A)^0 \quad \text{hat Dimension } n-r \text{ in } K^n, \tag{21.10b}$$
$$\mathrm{Bild}(A) = {}^0\mathrm{Kern}(A^\intercal) \quad \text{hat Dimension } r \text{ in } K^n, \tag{21.10c}$$
$$\mathrm{Kern}(A) = {}^0\mathrm{Bild}(A^\intercal) \quad \text{hat Dimension } m-r \text{ in } K^m. \tag{21.10d}$$

Beachte: Wir könnten nach Beispiel 20.28 die Beziehungen (21.10c) und (21.10d) auch als $\mathrm{Bild}(A) = \mathrm{Kern}(A^\intercal)^0$ bzw. als $\mathrm{Kern}(A) = \mathrm{Bild}(A^\intercal)^0$ schreiben.

Beweis. Die Matrix A induziert die Abbildung $f_A \in \mathrm{Homo}(K^m, K^n)$, und nach Folgerung 21.10 induziert $A^\intercal$ die Abbildung $(f_A)^*$. Nach Definition von Bild und Kern für Matrizen (siehe § 19.3) folgt das Resultat sofort aus Satz 21.11. □

Definition 21.13 (Co-Bild und Co-Bild einer Matrix)**.**
Es sei K ein Körper und $A \in K^{n\times m}$. Neben den aus § 19.3 bereits bekannten Begriffen

Bild von A	$\mathrm{Bild}(A) \coloneqq \mathrm{SR}\,A = \{Ax \mid x \in K^m\}$
Kern von A	$\mathrm{Kern}(A) \coloneqq \{x \in K^m \mid Ax = 0\}$

definieren wir

das Co-Bild von A[24]	$\mathrm{coBild}(A) \coloneqq (\mathrm{ZR}\,A)^\intercal = \{A^\intercal y \mid y \in K^n\} = \mathrm{Bild}(A^\intercal)$
den Co-Kern von A[25]	$\mathrm{coKern}(A) \coloneqq \{y \in K^n \mid A^\intercal y = 0\} = \mathrm{Kern}(A^\intercal)$. △

[20] Die Teilaussage $\mathrm{Bild}(f^*) \supseteq \mathrm{Kern}(f)^0$ hängt im Fall, dass W unendlich-dimensional ist, über den Basisergänzungssatz 13.5 vom Auswahlaxiom ab.

[21] Die Teilaussage $\mathrm{Bild}(f) \supseteq {}^0\mathrm{Kern}(f^*)$ hängt im Fall, dass W unendlich-dimensional ist, über Lemma 20.25 vom Auswahlaxiom ab.

[22] Die Teilaussage $\mathrm{Kern}(f) \supseteq {}^0\mathrm{Bild}(f^*)$ hängt im Fall, dass W unendlich-dimensional ist, über den Basisergänzungssatz 13.5 vom Auswahlaxiom ab.

[23] Ohne die Identifikation (also formal ganz korrekt) lauten die Beziehungen aus (21.10)

$$\mathrm{Bild}(A^\intercal) = \Phi_{B_m^*}^{-1}(\mathrm{Kern}(A)^0) \quad \text{in } K^m, \qquad \mathrm{Bild}(A) = {}^0(\Phi_{B_n^*}(\mathrm{Kern}(A^\intercal))) \quad \text{in } K^n,$$
$$\mathrm{Kern}(A^\intercal) = \Phi_{B_n^*}^{-1}(\mathrm{Bild}(A)^0) \quad \text{in } K^n, \qquad \mathrm{Kern}(A) = {}^0(\Phi_{B_m^*}(\mathrm{Bild}(A^\intercal))) \quad \text{in } K^m.$$

Wir wollen Folgerung 21.12 nochmals von einem übergeordneten Standpunkt aus interpretieren. Wir erinnern uns an die Aussage von Bemerkung 20.29, dass Prä-Annihilatoren Unterräume sind, die durch homogene lineare Gleichungssysteme beschrieben werden. Haben wir nun eine explizite Darstellung eines Unterraumes durch ein Erzeugendensystem oder sogar eine Basis gegeben, so können wir mit Hilfe von Folgerung 21.12 eine implizite Darstellung durch Gleichungen erhalten, und umgekehrt. Der Unterraum $U \subseteq K^n$ sei beispielsweise explizit durch die Vektoren $a_{\bullet 1}, \ldots, a_{\bullet m}$ erzeugt, also $U = \langle a_{\bullet 1}, \ldots, a_{\bullet m} \rangle = \operatorname{Bild}(A)$. Dann können wir mit Hilfe von (21.10c), also $\operatorname{Bild}(A) = {}^0\operatorname{Kern}(A^\intercal)$, eine implizite Beschreibung von U erhalten: Wir bestimmen wie in § 16 eine Basis von $\operatorname{Kern}(A^\intercal) \subseteq K^m$ und verwenden die Basisvektoren als die Zeilen einer Matrix. Die Lösungsmenge des zu dieser Matrix gehörenden homogenen linearen Gleichungssystems ist dann genau U.

Es sei umgekehrt ein Unterraum U implizit als Lösungsmenge eines homogenen linearen Gleichungssystems $Ax = 0$ mit der Koeffizientenmatrix $A \in K^{n \times m}$ gegeben. Dann können wir wie in § 16 eine Basis von $\operatorname{Kern}(A) \subseteq K^m$ bestimmen, um eine explizite Darstellung von U zu erhalten. Indirekt verwenden wir dabei (21.10d), also $\operatorname{Kern}(A) = {}^0\operatorname{Bild}(A^\intercal)$, denn nach Bemerkung 20.29 können wir ja die Bedingung $Ax = 0$ auch als $x \in {}^0\operatorname{Bild}(A^\intercal)$ interpretieren.

Wir illustrieren das Vorgehen an einem Beispiel.

Beispiel 21.14 (explizite und implizite Beschreibung von Unterräumen).

(*i*) Gegeben sei der Unterraum

$$U = \left\langle \begin{pmatrix} 1 \\ 3 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 4 \\ 0 \end{pmatrix} \right\rangle$$

von $\mathbb{Q}^4$. Das ist eine explizite Beschreibung von U durch Erzeugendensystem. Gesucht ist eine alternative, implizite Beschreibung von U durch Gleichungen, also als Lösungsmenge eines linearen Gleichungssystems. Wir setzen dazu

$$A := \begin{bmatrix} 1 & 0 \\ 3 & 1 \\ -2 & 4 \\ 1 & 0 \end{bmatrix},$$

sodass $U = \operatorname{Bild}(A)$ gilt. Nach (21.10c) ist $\operatorname{Bild}(A) = {}^0\operatorname{Kern}(A^\intercal)$. Wie in § 16 können wir eine Basis von $\operatorname{Kern}(A^\intercal)$ bestimmen. Dazu bringen wir $A^\intercal$ auf reduzierte Zeilenstufenform:

$$A^\intercal = \begin{bmatrix} 1 & 3 & -2 & 1 \\ 0 & 1 & 4 & 0 \end{bmatrix} \quad \rightsquigarrow \quad \begin{bmatrix} 1 & 0 & -14 & 1 \\ 0 & 1 & 4 & 0 \end{bmatrix}.$$

Hieraus können wir die Basis

$$\left(\begin{pmatrix} 14 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

[24] englisch: co-image
[25] englisch: co-kernel

von Kern($A^\intercal$) ablesen.[26] Wegen (21.10c) gilt nun also

$$U = \operatorname{Bild}(A) = {}^0\operatorname{Kern}(A^\intercal) = \left\{ y \in \mathbb{Q}^4 \,\middle|\, \begin{bmatrix} 14 & -4 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} y = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$
$$= \left\{ y \in \mathbb{Q}^4 \,\middle|\, \begin{aligned} 14\, y_1 - 4\, y_2 + y_3 &= 0 \\ -y_1 + y_4 &= 0 \end{aligned} \right\}.$$

Somit haben wir wie gewünscht U durch Gleichungen beschrieben.

(ii) Es sei umgekehrt der Unterraum U des Vektorraumes $\mathbb{Q}^5$ implizit, also durch Gleichungen beschrieben, etwa

$$U = \left\{ x \in \mathbb{Q}^5 \,\middle|\, \begin{aligned} x_1 + 3\,x_3 - x_5 &= 0 \\ x_1 + x_2 + x_4 &= 0 \end{aligned} \right\} = \left\{ x \in \mathbb{Q}^5 \,\middle|\, \begin{bmatrix} 1 & 0 & 3 & 0 & -1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Gesucht ist eine explizite Beschreibung von U mit Hilfe einer Basis. Dazu bringen wir die Koeffizientenmatrix A in reduzierte Zeilenstufenform:

$$\begin{bmatrix} 1 & 0 & 3 & 0 & -1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 3 & 0 & -1 \\ 0 & 1 & -3 & 1 & 1 \end{bmatrix}.$$

Daraus können wir eine Basis von $U = \operatorname{Kern}(A)$ ablesen:

$$\left(\begin{pmatrix} -3 \\ 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Somit haben wir wie gewünscht U mit Hilfe einer Basis beschrieben. △

§ 21.3 Zusammenspiel zwischen Dualräumen und Faktorräumen

In diesem Abschnitt untersuchen wir Dualräume von Faktorräumen und Faktorräume von Dualräumen.

Lemma 21.15 (Dualraum eines Faktorraumes)**.**
Es seien V ein Vektorraum und U ein Unterraum von V. Weiter sei $\pi\colon V \to V/U$ die kanonische Surjektion auf den Faktorraum. Dann gilt: Das Bild der zu π dualen Abbildung $\pi^*\colon (V/U)^* \to V^*$ ist U^0. Die Einschränkung

$$\pi^*|^{U^0} \colon \begin{cases} (V/U)^* \to U^0 \\ q^* \mapsto \pi^*(q^*) \end{cases} \quad \text{ist ein Isomorphismus} \quad (V/U)^* \cong U^0. \tag{21.11}$$

Beachte: Der Vektorraum U^0 der Linearformen auf V, die auf dem Unterraum U verschwinden, ist also isomorph zum Vektorraum der Linearformen, die nur auf dem gröberen Faktorraum V/U wirken.

[26]Wie in § 16 kennzeichnen wir vorübergehend unabhängige und abhängige Variable farbig.

Beweis. Wir zerlegen den Beweis in mehrere Schritte:

Schritt 1: Die kanonische Surjektion $\pi\colon V \to V/U$ ist eine surjektive lineare Abbildung (Satz 17.17). Die duale Abbildung $\pi^*\colon (V/U)^* \to V^*$ ist nach Satz 21.7 also injektiv; dieses Resultat verwendet nicht das Auswahlaxiom.

Wir müssen nun das Bild von π^* identifizieren.

Schritt 2: Wir zeigen zunächst $\pi^*((V/U)^*) \subseteq U^0$.

Dazu sehen wir uns an, wie π^* auf ein $q^* \in (V/U)^*$ wirkt. Für beliebiges $v \in V$ gilt

$$\begin{aligned} \langle \pi^*(q^*), v\rangle &= \langle q^*, \pi(v)\rangle && \text{nach Definition (21.2) von } \pi^* \\ &= \langle q^*, [v]\rangle && \text{nach Definition (17.14) von } \pi. \end{aligned}$$

Setzen wir für v speziell ein Element $u \in U$ ein, so ist $[u] = [0]$ der Nullvektor in V/U, und wir erhalten

$$\begin{aligned} \langle \pi^*(q^*), u\rangle &= \langle q^*, [0]\rangle \\ &= 0 \in K && \text{denn } q^* \text{ ist linear.} \end{aligned}$$

Das bedeutet aber, dass $\pi^*(q^*)$ für beliebiges $q^* \in (V/U)^*$ den Unterraum U annulliert, dass also $\pi^*(q^*)$ zum Unterraum U^0 von V^* gehört.

Schritt 3: Wir zeigen nun $U^0 \subseteq \pi^*((V/U)^*)$, dass also jedes Element von U^0 tatsächlich als Bild eines $q^* \in (V/U)^*$ unter π^* angenommen wird.

Zu gegebenem $v^* \in U^0 \subseteq V^*$ definieren wir $q^* \in (V/U)^*$ durch

$$\langle q^*, \underbrace{[v]}_{\text{beliebiges Element von } V/U}\rangle \coloneqq \langle v^*, v\rangle.$$

Diese Vorschrift ist wohldefiniert, also unabhängig von der Wahl des Repräsentanten $v \in [v] = v + U$. (**Quizfrage 21.5:** Warum?) Es gilt nun also für alle $v \in V$

$$\langle \pi^*(q^*), v\rangle = \langle q^*, \pi(v)\rangle = \langle q^*, [v]\rangle = \langle v^*, v\rangle,$$

was wie gewünscht $\pi^*(q^*) = v^*$ zeigt.

Da π^* in Schritt 1 als injektiv erkannt wurde, ist nach Lemma 6.12 die Einschränkung auf das Bild $\pi^*|^{U^0}\colon (V/U)^* \to U^0$ tatsächlich ein Isomorphismus. □

Beachte: Lemma 21.15 liefert ein Kriterium, wann eine Linearform $v^* \in V^*$ mit der Struktur des Faktorraumes V/U verträglich ist. Es gilt nämlich, dass $\langle v^*, v\rangle$ genau dann für alle Repräsentanten $v \in [v]$ eines Elements $[v] \in V/U$ den gleichen Wert annimmt, wenn v^* auf dem Unterraum U verschwindet, also $v^* \in U^0$ gilt. In diesem Fall ordnet Lemma 21.15 dem Element $[v] \in V/U$ die durch $\langle q^*, [v]\rangle \coloneqq \langle v^*, v\rangle$ definierte Linearform auf dem Faktorraum V/U zu.

Beispiel 21.16 (Dualraum eines Faktorraumes)**.**
Wir betrachten wie in Beispiel 17.19 den Folgenraum $V \coloneqq K^{\mathbb{N}}$ über einem Körper K und den Unterraum U derjenigen Folgen, deren Einträge an allen ungeraden Indizes $n \in \mathbb{N}$ gleich 0 sind.

$$U \quad \boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\boxed{0}\;\boxed{}\;\cdots$$

V / U 

Die mit gekennzeichneten Einträge legen dabei die Äquivalenzklasse (also das Element von V / U) fest.

Wir betrachten beispielsweise die Linearform $v^* \in V^*$, die den Vektor $(x_n)_{n\in\mathbb{N}} \in V$ auf den Wert $\langle v^*, (x_n)_{n\in\mathbb{N}}\rangle \coloneqq x_1 - x_3$ abbildet. Da v^* auf dem Unterraum U verschwindet, ordnet π^{-*} der Linearform $v^* \in V^*$ die Linearform $q^* \in (V / U)^*$ zu, die ein Element $[(x_n)_{n\in\mathbb{N}}] \in V / U$ auf den Wert $\langle q^*, [(x_n)_{n\in\mathbb{N}}]\rangle \coloneqq \langle v^*, (x_n)_{n\in\mathbb{N}}\rangle = x_1 - x_3$ abbildet. △

Wir kommen nun zu Faktorräumen von Dualräumen. Wir interessieren uns dabei insbesondere für das Ausfaktorisieren von Unterräumen, die Annihilatoren sind.

Lemma 21.17 (Faktorraum eines Dualraumes[AoC 27]).
Es seien V ein Vektorraum und U ein Unterraum von V. Weiter sei $i\colon U \to V$ die kanonische Injektion. Dann gilt: Der Kern der zu i dualen Abbildung $i^*\colon V^* \to U^*$ ist U^0. Es gilt also nach dem Homomorphiesatz für Vektorräume 17.21:

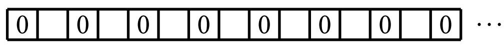
$$I\colon \begin{cases} V^* / U^0 \to U^* \\ [v^*] \mapsto i^*(v^*) = v^*|_U \end{cases} \quad \text{ist ein Isomorphismus} \quad V^* / U^0 \cong U^*. \tag{21.12}$$

Beachte: Der Vektorraum V^* / U^0 besteht aus Äquivalenzklassen von Linearformen auf V. Zwei Linearformen $v_1^*, v_2^* \in V^*$ gehören dabei genau dann zur selben Äquivalenzklasse, wenn $v_1^* - v_2^* \in U^0$ liegt, wenn also v_1^* und v_2^* auf dem Unterraum U übereinstimmen. Der Faktorraum V^* / U^0 dieser Äquivalenzklassen ist isomorph zum Vektorraum der Linearformen, die nur auf dem Unterraum U wirken.

Beweis. Es sei $i\colon U \to V$ die kanonische Injektion. Nach Satz 21.11 gilt $\operatorname{Kern}(i^*) = \operatorname{Bild}(i)^0 = U^0$. Weiter ist $i\colon U \to V$ injektiv, sodass nach Satz 21.7 $i^*\colon V^* \to U^*$ surjektiv ist[AoC]. Es gilt $\langle i^*(v^*), u\rangle = \langle v^*, i(u)\rangle = \langle v^*, u\rangle$ für alle $v^* \in V^*$ und $u \in U$, also $i^*(v^*) = v^*|_U$.

Der Homomorphiesatz für Vektorräume 17.21, angewendet auf i^*, zeigt nun, dass die Abbildung I, die eine ganze Nebenklasse $[v^*]$ von $v^* \in V^*$ auf $i^*(v^*)$ abbildet, ein Isomorphismus $V^* / U^0 \cong U^*$ ist. □

Beispiel 21.18 (Faktorraum eines Dualraumes).
Wir betrachten wie in Beispiel 17.19 den Folgenraum $V \coloneqq K^{\mathbb{N}}$ über einem Körper K und den Unterraum U derjenigen Folgen, deren Einträge an allen ungeraden Indizes $n \in \mathbb{N}$ gleich 0 sind.

U 

Wir betrachten beispielsweise die Linearform $u^* \in U^*$, die den Vektor $(x_n)_{n\in\mathbb{N}} \in U$ auf den Wert $\langle u^*, (x_n)_{n\in\mathbb{N}}\rangle \coloneqq x_2 - x_4$ abbildet. Jede beliebige Fortsetzung von u^* zu einer Linearform $v^* \in V^*$ gehört zur selben Nebenklasse $I^{-1}(u^*) = [v^*] \in V^* / U^0$. △

[27] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, über Satz 21.7 vom Auswahlaxiom ab.

§ 22 Der Bidualraum

In diesem Abschnitt betrachten wir den Dualraum eines Dualraumes.

Definition 22.1 (Bidualraum)**.**
Es seien V ein Vektorraum und V^* sein Dualraum. Der Dualraum von V^* heißt der **Bidualraum** (englisch: bidual space) von V. Er wird statt mit $(V^*)^*$ auch einfach mit V^{**} bezeichnet. △

Der Bidualraum V^{**} ist also der Vektorraum der Linearformen auf V^*. Wir können uns fragen, wie die Elemente von V^{**} aussehen. Dabei fällt auf, dass für jedes fest gewählte $v \in V$ die Abbildung

$$V^* \ni v^* \mapsto \langle v^* , v \rangle \in K$$

eine Linearform auf V^*, also ein Element von V^{**}, ist. Diese Linearform können wir mit $\langle \cdot , v \rangle$ bezeichnen.

Satz 22.2 (kanonische Injektion$^{\text{AoC}}$)**.**
Es seien V ein Vektorraum über dem Körper K und V^* sein Dualraum.

(i) Die Abbildung

$$i_V \coloneqq V \ni v \mapsto \langle \cdot , v \rangle \in V^{**} \tag{22.1}$$

ist ein injektiver$^{\text{AoC}}$ Homomorphismus, genannt die **kanonische Injektion** (englisch: canonical injection) von V in V^{**}.[28]

(ii) Ist V endlich-dimensional, dann ist i_V auch surjektiv, also ein Isomorphismus. In diesem Fall gilt $\dim(V) = \dim(V^{**})$.

Beweis. Aussage (i): Wir weisen zunächst die Linearität von i_V nach. Für $v_1, v_2 \in V^*$ haben wir

$$\begin{aligned}
\big(i_V(v_1 + v_2)\big)(v^*) &= \langle v^* , v_1 + v_2 \rangle && \text{nach Definition von } i_V \\
&= \langle v^* , v_1 \rangle + \langle v^* , v_2 \rangle && \text{denn } v^* \text{ ist eine Linearform auf } V \\
&= i_V(v_1)(v^*) + i_V(v_2)(v^*) && \text{nach Definition von } i_V \\
&= \big(i_V(v_1) + i_V(v_2)\big)(v^*) && \text{nach Definition der Addition in } V^{**}.
\end{aligned}$$

Das bedeutet $i_V(v_1 + v_2) = i_V(v_1) + i_V(v_2)$, d. h., i_V ist additiv. Außerdem ist i_V homogen, denn mit $\alpha \in K$ und $v^* \in V^*$ gilt

$$\begin{aligned}
\big(i_V(\alpha\, v)\big)(v^*) &= \langle v^* , \alpha\, v \rangle && \text{nach Definition von } i_V \\
&= \alpha \, \langle v^* , v \rangle && \text{denn } v^* \text{ ist eine Linearform auf } V \\
&= \alpha \, i_V(v) && \text{nach Definition von } i_V \\
&= (\alpha \, i_V)(v) && \text{nach Definition der S-Multiplikation in } V^{**}.
\end{aligned}$$

[28] Die Injektivität („Zwei verschiedene Elemente von V erzeugen auch verschiedene Elemente von V^{**}“) hängt im Fall, dass V unendlich-dimensional ist, über die Aussage ${}^0(V^*) = \{0_V\}$ aus Lemma 20.21 vom Auswahlaxiom ab.

Um die Injektivität von i_V zu prüfen, betrachten wir ein $v \in \operatorname{Kern}(i_V)$. Es gilt also

$$i_V(v) = 0_{V^{**}}, \quad \text{d. h., } \big(i_V(v)\big)(v^*) = \langle v^*, v\rangle = 0$$

für alle $v^* \in V^*$. Mit anderen Worten haben wir $v \in \operatorname{Kern}(v^*)$ für alle $v^* \in V^*$ oder kurz: $v \in {}^0(V^*)$. Nach Lemma 20.21[AoC] bedeutet das aber $v = 0$. Also ist i_V injektiv.

Aussage (ii): Im Fall $\dim(V) = n \in \mathbb{N}_0$ gilt auch $\dim(V^*) = n$, siehe Folgerung 20.9. Wenden wir dieses Resultat erneut an, so folgt $\dim(V^{**}) = \dim(V^*)$, also auch $\dim(V^{**}) = \dim(V) = n$. Nun können wir mit Folgerung 18.9 schließen, dass i_V nicht nur injektiv (Aussage (i)), sondern sogar bijektiv ist, also ein Isomorphismus. □

Beachte: In Folgerung 20.9 hatten wir für endliche-dimensionale Vektorräume V einen Isomorphismus zwischen V und seinem Dualraum V^* etabliert, der aber von der Wahl einer Basis abhängig war. Im Unterschied dazu ist die kanonische Injektion i_V (wie das Attribut „kanonisch“ bereits andeutet) ein natürlicher Isomorphismus zwischen V und seinem Doppeldual V^{**}, also nicht von der Wahl einer Basis abhängig.

Wir klären jetzt, wie Annihilatoren im Dualraum aussehen.

Folgerung 22.3 (Eigenschaften von Annihilatoren im Dualraum)**.**

Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Dann gilt

(i) Für jede Teilmenge $F \subseteq V^*$ ist $F^0 = i_V({}^0F)$.

(ii) Für jeden Unterraum $U \subseteq V$ ist $(U^0)^0 = i_V(U)$.

Wenn wir wie in Bemerkung 20.16 $(K^n)^*$ mit K^n identifizieren, dann gilt:

(i) Für jede Teilmenge $F \subseteq K^n$ ist $F^0 = {}^0F$.

(ii) Für jeden Unterraum $U \subseteq K^n$ ist $(U^0)^0 = U$.

Beweis. Aussage (i): Nach Definition (20.12) gilt für den Annihilator

$$\begin{aligned} F^0 &= \{f \in V^{**} \mid \langle f, v^*\rangle_{V^{**},V^*} = 0 \text{ für alle } v^* \in F\} \\ &= \{i_V(v) \in V^{**} \mid \langle v^*, v\rangle_{V^*,V} = 0 \text{ für alle } v^* \in F\}, \end{aligned}$$

wobei die bijektive Ersetzung $f = i_V(v)$ (Satz 22.2) verwendet wurde. Für den Prä-Annihilator ergibt Definition (20.13)

$$\begin{aligned} {}^0F &= \{v \in V \mid \langle v^*, v\rangle_{V^*,V} = 0 \text{ für alle } v^* \in F\}, \\ \text{also} \quad i_V({}^0F) &= \{i_V(v) \in V^{**} \mid \langle v^*, v\rangle_{V^*,V} = 0 \text{ für alle } v^* \in F\} \\ &= F^0. \end{aligned}$$

Aussage (ii): Nach Satz 22.2 ist die kanonische Injektion $i_V \in \operatorname{Homo}(V, V^{**})$ bijektiv. Es gilt

$$\begin{aligned} U &= {}^0(U^0) && \text{nach Lemma 20.25} \\ &= i_V^{-1}\Big(i_V\big({}^0(U^0)\big)\Big) && \text{wegen } i_V^{-1} \circ i_V = \mathrm{id}_V \\ &= i_V^{-1}\big((U^0)^0\big) && \text{nach Aussage } (i) \text{ mit } F = U^0. \end{aligned}$$

Nochmaliges Anwenden von i_V auf beide Seiten zeigt schließlich $i_V(U) = (U^0)^0$. □

Wie einen Vektorraum können wir auch eine lineare Abbildung $f \in \operatorname{Homo}(V, W)$ zweimal dualisieren. Das führt zum Begriff des **bidualen Homomorphismus**:

Definition 22.4 (bidualer Homomorphismus, vgl. Definition 21.1)**.**
Es seien V und W Vektorräume über demselben Körper. Weiter sei $f \in \operatorname{Homo}(V, W)$ eine lineare Abbildung. Dann heißt

$$f^{**} \colon V^{**} \ni v^{**} \mapsto w^{**} \coloneqq v^{**} \circ f^* \in W^{**} \tag{22.2}$$

der zu f **biduale Homomorphismus** (englisch: bidual homomorphism) bzw. die zu f **biduale lineare Abbildung** (englisch: bidual linear map). △

Mit anderen Worten: $f^{**} \in \operatorname{Homo}(V^{**}, W^{**})$ ist durch die Bedingungen

$$\langle f^{**}(v^{**}) \, , w^* \rangle_{W^{**},W^*} = \langle v^{**} \, , f^*(w^*) \rangle_{V^{**},V^*} \tag{22.3}$$

für alle $v^{**} \in V^{**}$ und alle $w^* \in W^*$ festgelegt. Für den Zusammenhang zwischen f und f^{**} gilt:

Lemma 22.5 (Zusammenhang zwischen f und f^{**})**.**
Es seien V und W Vektorräume über demselben Körper sowie $f \in \operatorname{Homo}(V, W)$ ein Homomorphismus. Dann gilt

$$i_W \circ f = f^{**} \circ i_V . \tag{22.4}$$

Mit anderen Worten, folgendes Diagramm kommutiert:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "i_V"'] & W \arrow[d, "i_W"] \\ V^{**} \arrow[r, "f^{**}"] & W^{**} \end{tikzcd}$$

Beweis. Beide Seiten von (22.4) sind per Definition lineare Abbildungen $V \to W^{**}$. Um die Gleichheit zu zeigen, setzen wir ein beliebiges Element $v \in V$ in beide Abbildungen ein und erhalten dadurch Bildelemente in W^{**}. Diese vergleichen wir, indem wir sie wirken lassen auf ein beliebiges Element $w^* \in W^*$.

Wenn wir zeigen können, dass für alle $v \in V$ und $w^* \in W^*$ die Gleichheit $\langle (i_W \circ f)(v) \, , w^* \rangle_{W^{**},W^*} = \langle (f^{**} \circ i_V)(v) \, , w^* \rangle_{W^{**},W^*}$ gilt, so ist die Behauptung (22.4) bewiesen. (**Quizfrage 22.1:** Klar?) Es gilt

$$\begin{aligned}
&\langle (f^{**} \circ i_V)(v) \, , w^* \rangle_{W^{**},W^*} \\
&\qquad = \langle f^{**}(i_V(v) \, , w^* \rangle_{W^{**},W^*} && \text{nach Definition der Komposition} \\
&\qquad = \langle i_V(v) \circ f^* \, , w^* \rangle_{W^{**},W^*} && \text{nach Definition (22.2) von } f^{**} \\
&\qquad = \langle i_V(v) \, , f^*(w^*) \rangle_{V^{**},V^*} && \text{wegen der Assoziativität der Komposition} \\
&\qquad = \langle i_V(v) \, , w^* \circ f \rangle_{V^{**},V^*} && \text{nach Definition (21.1) von } f^* \\
&\qquad = \langle w^* \circ f \, , v \rangle_{V^*,V} && \text{nach Definition (22.1) von } i_V
\end{aligned}$$

$$
\begin{aligned}
&= \langle w^*, f(v) \rangle_{W^*,W} && \text{wegen der Assoziativität der Komposition} \\
&= \langle i_W(f(v)), w^* \rangle_{W^{**},W^*} && \text{nach Definition (22.1) von } i_V \\
&= \langle (i_W \circ f)(v), w^* \rangle_{W^{**},W^*} && \text{nach Definition der Komposition.}
\end{aligned}
$$

Damit ist in der Tat $\langle (i_W \circ f)(v), w^* \rangle_{W^{**},W^*} = \langle (f^{**} \circ i_V)(v), w^* \rangle_{W^{**},W^*}$ gezeigt, also (22.4). □

Ende der Vorlesung 4

Kapitel 6 Multilineare Abbildungen und Tensorprodukträume

§ 23 Bilineare Abbildungen und das Tensorprodukt von zwei Vektorräumen

Literatur: Fischer, Springborn, 2020, Kapitel 7.3–7.4; Hackbusch, 2019, Chapter 3

§ 23.1 Bilineare Abbildungen

Definition 23.1 (bilineare Abbildung, Bilinearform)**.**
Es seien U, V und W Vektorräume über demselben Körper K.

(i) Eine Abbildung

$$b\colon U \times V \to W$$

heißt **bilinear** (englisch: bilinear map), wenn[1]

$$\begin{aligned} &\text{für jedes } v \in V \text{ die Abbildung} \quad b(\cdot, v)\colon U \ni u \mapsto b(u, v) \in W \quad \text{linear ist} && (23.1a)\\ &\text{und für jedes } u \in U \text{ die Abbildung} \quad b(u, \cdot)\colon V \ni v \mapsto b(u, v) \in W \quad \text{linear ist}^2. && (23.1b)\end{aligned}$$

Die Funktionen in (23.1) kommen durch **partielles Einsetzen** (englisch: partial application) zustande.[3]

(ii) Die Menge aller bilinearen Abbildungen $U \times V \to W$ bezeichnen wir mit $\operatorname{Bil}(U, V; W)$.

(iii) Eine bilineare Abbildung in den Vektorraum $W = K$ heißt eine **Bilinearform** (englisch: bilinear form) auf $U \times V$.[4]

(iv) Die Menge aller Bilinearformen $U \times V \to K$ bezeichnen wir mit $\operatorname{Bil}(U, V; K)$ oder kurz mit $\operatorname{Bil}(U, V)$. △

[1]Wir schreiben die Funktionswerte eines bilinearen Abbildung in der Form $b(u, v)$ und nicht als $b((u, v))$.
[2]Kurz gesagt ist $b\colon U \times V \to W$ also bilinear, wenn für alle $u \in U$ und alle $v \in V$ gilt: $b(\cdot, v) \in \operatorname{Homo}(U, W)$ und $b(u, \cdot) \in \operatorname{Homo}(V, W)$.
[3]Wir vermeiden den Begriff **partielle Funktion**, weil dieser häufig für „nicht notwendigerweise überall definierte Funktionen" verwendet wird, also für rechtseindeutige Relationen, die nicht notwendigerweise linkstotal sind.
[4]Kurz gesagt ist $b\colon U \times V \to K$ also eine Bilinearform, wenn für alle $u \in U$ und alle $v \in V$ gilt: $b(u, \cdot) \in V^*$ und $b(\cdot, v) \in U^*$ liegt.

Ausgeschrieben besagt die Bilinearitätsbedingung (23.1) gerade

$$\begin{aligned} b(u_1+u_2,v) &= b(u_1,v)+b(u_2,v) \text{ und } b(\alpha\, u,v) = \alpha\, b(u,v) && \text{(Linearität im ersten Argument)} \\ b(u,v_1+v_2) &= b(u,v_1)+b(u,v_2) \text{ und } b(u,\beta\, v) = \beta\, b(u,v) && \text{(Linearität im zweiten Argument)} \end{aligned} \tag{23.2}$$

für alle $u, u_1, u_2 \in U$, alle $v, v_1, v_2 \in V$ und alle $\alpha, \beta \in K$. Mit Hilfe vollständiger Induktion sieht man leicht, dass (23.1) auch äquivalent ist zu

$$b\Big(\sum_{i=1}^n \alpha_i\, u_i, \sum_{j=1}^m \beta_j\, v_j\Big) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i\, \beta_j\, b(u_i, v_j) \tag{23.3}$$

für alle $n, m \in \mathbb{N}_0$, $u_1, \ldots, u_n \in U$, $v_1, \ldots, v_m \in V$ und $\alpha_1, \ldots, \alpha_n \in K$ sowie $\beta_1, \ldots, \beta_m \in K$.

Beispiel 23.2 (bilineare Abbildung, Bilinearform)**.**

(i) Für jeden Körper ist die Abbildung

$$K \times K \ni (\alpha, \beta) \mapsto \alpha\, \beta \in K$$

eine Bilinearform auf $K \times K$.

(ii) Es seien K ein Körper und $n, m \in \mathbb{N}_0$. Für jede Matrix $A \in K^{n \times m}$ ist die Abbildung

$$K^m \times K^n \ni (x, y) \mapsto y^\intercal A\, x \in K$$

eine Bilinearform auf $K^m \times K^n$.

(iii) Es seien K ein Körper und $n, m, \ell \in \mathbb{N}_0$. Das Matrix-Matrix-Produkt

$$K^{n \times m} \times K^{m \times \ell} \ni (A, B) \mapsto A\, B \in K^{n \times \ell}$$

ist eine bilineare Abbildung, also ein Element von $\operatorname{Bil}(K^{n \times m}, K^{m \times \ell}; K^{n \times \ell})$, siehe Lemma 15.8.

(iv) Die duale Paarung (20.2) eines Vektorraumes V und seines Dualraumes V^*, also

$$\langle \cdot\,, \cdot \rangle \colon V^* \times V \ni (v^*, v) \mapsto \langle v^*\,, v \rangle \in K,$$

ist eine Bilinearform auf $V^* \times V$, siehe Lemma 20.7.

(v) Es seien U, V Vektorräume über demselben Körper K. Für gegebene $u^* \in U^*$ und $v^* \in V^*$ ist die Abbildung

$$U \times V \ni (u, v) \mapsto \langle u^*\,, u \rangle\, \langle v^*\,, v \rangle \in K$$

eine Bilinearform auf $U \times V$.

(vi) Für jede Menge X und jeden Körper $(K, +, \cdot)$ ist die punktweise Multiplikation zweier Funktionen $X \to K$, also

$$K^X \times K^X \ni (f, g) \mapsto f \cdot g \in K^X,$$

eine bilineare Abbildung, also ein Element von $\operatorname{Bil}(K^X, K^X; K^X)$. △

So wie die Menge der linearen Abbildungen $\operatorname{Homo}(U, V)$ zwischen Vektorräumen U und V wieder einen Vektorraum bzgl. der punktweisen Addition und punktweisen S-Multiplikation bildet (Satz 17.14), so trifft das auch auf die Menge der bilinearen Abbildungen zu:

Satz 23.3 (bilineare Abbildungen bilden einen Vektorraum, vgl. Satz 17.14)**.**
Es seien U, V und W Vektorräume über demselben Körper. Dann ist $\operatorname{Bil}(U, V; W)$ ein Unterraum des Vektorraumes $W^{U\times V} = \{f : U \times V \to W\}$ aller Abbildungen $U \times V \to W$.

Beweis. Der Beweis ist Gegenstand der Übung. □

Genau wie lineare Abbildungen zwischen Vektorräumen können wir auch bilineare Abbildung dadurch eindeutig bestimmen, dass wir ihre Bilder auf einer geeigneten Menge (oder Familie) festlegen. Das kartesische Produkt von Basen ist eine solche Menge (oder Familie):

Satz 23.4 (Existenz- und Eindeutigkeitssatz für bilineare Abbildungen, vgl. Satz 17.10)**.**
Es seien U, V und W Vektorräume über demselben Körper K. Weiter sei $(u_i)_{i\in I}$ eine Basis von U und $(v_j)_{j\in J}$ eine Basis von V. Außerdem sei $(w_{ij})_{(i,j)\in I\times J}$ eine Familie von Vektoren in W. Dann gibt es genau eine bilineare Abbildung $b : U \times V \to W$ mit der Eigenschaft $b(u_i, v_j) = w_{ij}$ für alle $(i, j) \in I \times J$.

Beweis. Es sei $(u, v) \in U \times V$ beliebig. Wegen der Basiseigenschaft (Satz 13.3) gibt es (bis auf Summanden mit Nullkoeffizienten) eindeutige Darstellungen

$$u = \sum_{i\in I_0} \alpha_i\, u_i \quad \text{und} \quad v = \sum_{j\in J_0} \beta_j\, v_j$$

mit endlichen Teilfamilien I_0 von I und J_0 von J.

Wenn es eine bilineare Abbildung $b : U \times V \to W$ mit der gesuchten Eigenschaft $b(u_i, v_j) = w_{ij}$ gibt, so muss diese notwendigerweise

$$\begin{aligned} b(u, v) &= b\Big(\sum_{i\in I_0} \alpha_i\, u_i, \sum_{j\in J_0} \beta_j\, v_j\Big) \\ &= \sum_{(i,j)\in I_0\times J_0} \alpha_i\, \beta_j\, b(u_i, v_j) && \text{wegen der Bilinearität} \\ &= \sum_{(i,j)\in I_0\times J_0} \alpha_i\, \beta_j\, w_{ij} && \text{wegen der gewünschten Bilder} \end{aligned}$$

erfüllen. Wegen der Eindeutigkeit der Linearkombinationen (bis auf Summanden mit Nullkoeffizienten) ist also die gesuchte bilineare Abbildung, wenn sie existiert, tatsächlich eindeutig bestimmt.

Andererseits definiert die gefundene Vorschrift

$$b(u, v) \coloneqq \sum_{(i,j)\in I_0\times J_0} \alpha_i\, \beta_j\, w_{ij}$$

für Vektoren $u = \sum_{i \in I_0} \alpha_i \, u_i$ und $v = \sum_{j \in J_0} \beta_j \, v_j$ tatsächlich eine bilineare Abbildung. Um das zu sehen, halten wir zunächst $v \in V$ fest und betrachten beliebige $u, \widehat{u} \in U$. Diese haben die Darstellungen $u = \sum_{i \in I_0} \alpha_i \, u_i$ und $\widehat{u} = \sum_{i \in I_0} \widehat{\alpha}_i \, u_i$. (**Quizfrage 23.1:** Warum können wir von derselben endlichen Indexmenge I_0 für beide Vektoren ausgehen?) Dann gilt

$$b(u + \widehat{u}, v) = \sum_{(i,j) \in I_0 \times J_0} (\alpha_i + \widehat{\alpha}_i) \, \beta_j \, w_{ij} = \sum_{(i,j) \in I_0 \times J_0} \alpha_i \, \beta_j \, w_{ij} + \sum_{(i,j) \in I_0 \times J_0} \widehat{\alpha}_i \, \beta_j \, w_{ij} = b(u, v) + b(\widehat{u}, v)$$

und für $\alpha \in K$ auch

$$b(\alpha \, u, v) = \sum_{(i,j) \in I_0 \times J_0} \alpha \, \alpha_i \, \beta_j \, w_{ij} = \alpha \sum_{(i,j) \in I_0 \times J_0} \alpha_i \, \beta_j \, w_{ij} = \alpha \, b(u, v).$$

Das heißt aber: $U \ni u \mapsto b(u, v) \in W$ ist für jedes $v \in V$ linear. Analog können wir zeigen, dass auch $V \ni v \mapsto b(u, v) \in W$ für jedes $u \in U$ linear ist. □

Bemerkung 23.5 (Bilinearität vs. Linearität).

(i) Bilineare Abbildungen $U \times V \to W$ sind zwar auf dem kartesischen Produkt $U \times V$ (als Menge) definiert, sie sind aber — anders als es lineare Abbildungen $U \times V \to W$ wären — inkompatibel mit der Struktur von $U \times V$ als Vektorraum (Definition 11.4). Beispielsweise erfüllt eine bilineare Abbildung ja

$$f(\alpha \, (u, v)) = f(\alpha \, u, \alpha \, v) = \alpha \, f(u, \alpha \, v) = \alpha^2 f(u, v)$$

und nicht etwa $f(\alpha \, (u, v)) = \alpha \, f(u, v)$, wie es für eine lineare Abbildung $U \times V \to W$ der Fall wäre.[5]

(ii) Ein weiterer Hinweis darauf, dass die Vektorraumstruktur von $U \times V$ für bilineare Abbildungen keine Rolle spielt, zeigt sich in Satz 23.4: Während für die Dimension des Vektorraumes $U \times V$ nach Lemma 13.18 $\dim(U \times V) = \dim(U) + \dim(V)$ gilt, wird eine bilineare Abbildung durch $\dim(U) \cdot \dim(V)$ viele Bilder festgelegt. △

Wir halten fest: Um bilineare Abbildungen $U \times V \to W$ zu definieren, verwenden wir von $U \times V$ nur die Struktur als Menge. Das ist auch bereits ausreichend dafür, dass $\operatorname{Bil}(U, V; W)$ zu einem Vektorraum wird, denn dafür ist nur die Vektorraumstruktur auf W ausschlaggebend (Satz 23.3). Dennoch gibt es gute Gründe dafür, zu versuchen, bilineare Abbildungen $U \times V \to W$ als lineare Abbildungen $\mathcal{T} \to W$ auf einem noch zu definierenden Vektorraum $\mathcal{T}$ zu verstehen, der der **Tensorproduktraum** von U und V genannt wird. Das hat u. a. den Vorteil, dass wir für bilineare (und später multilineare) Abbildungen keine parallele Infrastruktur aufbauen müssen, sondern dass wir stattdessen die uns bereits bekannten Konzepte für lineare Abbildungen sofort wieder anwenden können.

Wir werden in § 23.2 unsere Anforderungen an den Tensorproduktraum formulieren. Zum ersten Mal werden wir dieses neue Objekt dabei nicht sofort konkret konstruieren, sondern es stattdessen durch seine **universelle Eigenschaft** charakterisieren. Diese ist aber ausreichend, um den Tensorproduktraum bis auf Isomorphie eindeutig zu bestimmen (Satz 23.8), wodurch alle seine Eigenschaften festgelegt sind. Erst im Anschluss werden wir uns in § 23.4 davon überzeugen, dass es Tensorprodukträume tatsächlich gibt.

[5] Tatsächlich kann man zeigen, dass eine Abbildung $U \times V \to W$, die gleichzeitig linear und multilinear ist, die Nullabbildung sein muss.

§ 23.2 Das Tensorprodukt von zwei Vektorräumen

Wir beginnen mit der Formulierung unserer Anforderungen an Tensorprodukträume. Gegeben seien zwei Vektorräume U, V über demselben Körper K. Unser Ziel ist es, jede beliebige bilineare Abbildung $b\colon U \times V \to W$ in irgendeinen weiteren K-Vektorraum W in zwei Teilabbildungen zerlegen zu können:

$$\begin{tikzcd} U \times V \arrow[r, "\otimes"] \arrow[dr, "b"'] & \mathcal{T} \arrow[d, "f"] \\ & W \end{tikzcd}$$

Die erste Teilabbildung $\otimes\colon U \times V \to \mathcal{T}$ ist dabei allein für die Bilinearität verantwortlich. Sie soll **universell** sein in dem Sinne, dass sie für alle bilinearen Abbildungen $b\colon U \times V \to W$ in irgendeinen K-Vektorraum W verwendbar ist. Die wesentliche Information, um welche bilineare Abbildung es sich handelt, steckt dann ausschließlich in der linearen Folgeabbildung $f\colon \mathcal{T} \to W$.

Zu gegebenen Vektorräumen U und V wünschen wir uns also einen K-Vektorraum $\mathcal{T}$, zusammen mit einer bilinearen Abbildung $\otimes\colon U \times V \to \mathcal{T}$, sodass die Zuordnung $b \mapsto f$ bijektiv ist und einen Isomorphismus von Vektorräumen

$$\operatorname{Bil}(U, V; W) \cong \operatorname{Homo}(\mathcal{T}, W)$$

ergibt. In der Gesamtschau bedeutet das, dass wir die Eigenschaft der Bilinearität mit Hilfe der immer gleichen bilinearen Abbildung $\otimes\colon U \times V \to \mathcal{T}$ „abtrennen" und so bilineare Abbildungen $b\colon U \times V \to W$ mit linearen Abbildungen $f\colon \mathcal{T} \to W$ identifizieren können.

Wir halten die gewünschten Eigenschaften in einer Definition fest:

Definition 23.6 (Tensorprodukt, Tensorproduktraum, universelle bilineare Abbildung)**.**
Es seien U und V Vektorräume über demselben Körper K. Weiter seien $\mathcal{T}$ ein weiterer K-Vektorraum und $\otimes\colon U \times V \to \mathcal{T}$ eine bilineare Abbildung.[6]

(i) Das Paar $(\mathcal{T}, \otimes)$ heißt ein[7] **Tensorprodukt** (englisch: tensor product) von U und V, wenn die folgende **universelle Eigenschaft** (englisch: universal property) erfüllt ist:

Für jede bilineare Abbildung $b\colon U \times V \to W$ in irgendeinen K-Vektorraum W gibt es eine eindeutig bestimmte lineare Abbildung $f \in \operatorname{Homo}(\mathcal{T}, W)$ mit der Eigenschaft $b = f \circ \otimes$, also $b(u, v) = f(u \otimes v)$ für alle $u \in U$ und $v \in V$.

(ii) Dabei heißt $\mathcal{T}$ der **Tensorproduktraum** (englisch: tensor product space) des Tensorprodukts $(\mathcal{T}, \otimes)$. Die Elemente von $\mathcal{T}$ heißen **Tensoren** (englisch: tensors). Der Nullvektor in $\mathcal{T}$ heißt der **Nulltensor** (englisch: zero tensor).

[6] Für die Bilder dieser **äußeren Verknüpfung** (englisch: outer operation) schreiben wir $u \otimes v$ an Stelle von $\otimes(u, v)$.
[7] zur Eindeutigkeit des Tensorprodukts gleich mehr

(*iii*) $\otimes\colon U \times V \to \mathcal{T}$ heißt die **universelle bilineare Abbildung** (englisch: universal bilinear map) des Tensorprodukts $(\mathcal{T}, \otimes)$. Tensoren der Form $u \otimes v$ mit $u \in U$ und $v \in V$ — also die Bilder der universellen bilinearen Abbildung — heißen **Elementartensoren** (englisch: elementary tensors) oder **einfache Tensoren** (englisch: simple tensors). $u \otimes v$ heißt auch das **Tensorprodukt** (englisch: tensor product) der Vektoren u und v oder kurz: „u Tensor v". △

Bemerkung 23.7 (zum Tensorprodukt)**.**

(*i*) Wir vereinbaren, dass das Tensorproduktsymbol $\otimes$ als „multiplikatives" Symbol Vorrang vor der Addition bekommt, sodass die Bedingung der Bilinearität von $\otimes$ ausgeschrieben lautet:

$$u_1 \otimes v \, + \, u_2 \otimes v = (u_1 + u_2) \otimes v \tag{23.4a}$$

$$u \otimes v_1 \, + \, u \otimes v_2 = u \otimes (v_1 + v_2) \tag{23.4b}$$

$$(\alpha\, u) \otimes v = \alpha\, (u \otimes v) \tag{23.4c}$$

$$u \otimes (\beta\, v) = \beta\, (u \otimes v) \tag{23.4d}$$

für alle $u, u_1, u_2 \in U$ und $v, v_1, v_2 \in V$ sowie $\alpha, \beta \in K$.

(*ii*) Wegen (23.4c) und (23.4d) gilt $(\alpha\, u) \otimes v = \alpha\, (u \otimes v) = u \otimes (\alpha\, v)$. Wir können also einfach $\alpha\, u \otimes v$ schreiben, ohne Klammern verwenden zu müssen.

(*iii*) Wie für alle bilinearen Abbildungen gilt nach (23.3)

$$\Big(\sum_{i=1}^n \alpha_i\, u_i\Big) \otimes \Big(\sum_{j=1}^m \beta_j\, v_j\Big) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i\, \beta_j\, u_i \otimes v_j \tag{23.5}$$

für alle $n, m \in \mathbb{N}_0$, $u_1, \ldots, u_n \in U$, $v_1, \ldots, v_m \in V$ und $\alpha_1, \ldots, \alpha_n \in K$ sowie $\beta_1, \ldots, \beta_m \in K$.

(*iv*) Das Tensorprodukt $u \otimes v$ ist i. A. nicht kommutativ, selbst wenn $U = V$ gilt.

(*v*) Der Tensorproduktraum $\mathcal{T}$ eines Tensorprodukts $(\mathcal{T}, \otimes)$ der Vektorräume U und V wird in der Regel mit dem Symbol $U \otimes V$ bezeichnet.[8] Das werden wir ab sofort auch tun. △

Wir zeigen jetzt, dass mit Hilfe eines Tensorprodukts tatsächlich ein kanonischer[9] Isomorphismus von Vektorräumen $\operatorname{Bil}(U, V; W) \cong \operatorname{Homo}(U \otimes V, W)$ entsteht. Bilineare Abbildungen sind dann also nur noch lineare Abbildungen in anderer Gestalt!

Satz 23.8 (universelle Eigenschaft des Tensorprodukts $(U \otimes V, \otimes)$)**.**
Es seien U und V Vektorräume über demselben Körper K und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Außerdem sei auch W ein K-Vektorraum. Dann gilt:

(*i*) Zu jeder bilinearen Abbildung $b\colon U \times V \to W$ existiert genau eine lineare Abbildung $f\colon U \otimes V \to W$ mit der Eigenschaft $b = f \circ \otimes$, also $b(u, v) = f(u \otimes v)$ für alle $u \in U$ und $v \in V$. Mit anderen Worten, das folgende Diagramm kommutiert:

[8] sprich: „U Tensor V"
[9] „Kanonisch" bedeutet im Kontext von Vektorräumen immer „unabhängig von der Wahl einer Basis".

$$\begin{tikzcd} U \times V \arrow[r, "\otimes"] \arrow[dr, "b"'] & U \otimes V \arrow[d, "f"] \\ & W \end{tikzcd}$$

(ii) Umgekehrt wird für jede lineare Abbildung $f\colon U \otimes V \longrightarrow W$ durch die Komposition mit der universellen bilinearen Abbildung $\otimes\colon U \times V \longrightarrow U \otimes V$ eine bilineare Abbildung $f \circ \otimes\colon U \times V \ni (u, v) \mapsto f(u \otimes v) \in W$ definiert.

(iii) Die Zuordnung

$$\operatorname{Bil}(U, V; W) \ni b \mapsto f \in \operatorname{Homo}(U \otimes V, W)$$

ist ein kanonischer Isomorphismus von Vektorräumen.

Beweis. Aussage (i) folgt sofort aus der Definition 23.6.

Aussage (ii): Die Komposition der universellen bilinearen Abbildung $\otimes\colon U \times V \longrightarrow U \otimes V$ mit $f \in \operatorname{Homo}(U \otimes V, W)$ ergibt eine bilineare Abbildung $U \times V \longrightarrow W$, denn

$$(f \circ \otimes)(u, v) = f(\otimes(u, v)) = f(u \otimes v)$$

erfüllt

$$\begin{aligned} f((u_1 + u_2) \otimes v) &= f(u_1 \otimes v + u_2 \otimes v) = f(u_1 \otimes v) + f(u_2 \otimes v) \\ f((\alpha\, u) \otimes v) &= f(\alpha\,(u \otimes v)) = \alpha\, f(u_1 \otimes v) \end{aligned}$$

für $u_1, u_2 \in U$, $v \in V$ und $\alpha \in K$, ist also linear im ersten Argument. Analog können wir auch die Linearität im zweiten Argument zeigen.

Aussage (iii): Wir unterteilen den Beweis in zwei Schritte:

Schritt 1: Wir müssen zunächst beweisen, dass die Zuordnung $b \mapsto f$ bijektiv ist. Dazu zeigen wir, dass $b \mapsto f \mapsto b'$ die Identität auf $\operatorname{Bil}(U, V; W)$ und dass $f \mapsto b \mapsto f'$ die Identität auf $\operatorname{Homo}(U \otimes V, W)$ ist (vgl. Definition 6.22 und Lemma 6.24).

Es sei zuerst $f \in \operatorname{Homo}(U \otimes V, W)$ die zu $b \in \operatorname{Bil}(U, V; W)$ gehörige lineare Abbildung, die also die Eigenschaft $b(u, v) = f(u \otimes v)$ erfüllt. Dann gilt für $b' = f \circ \otimes$ und alle $u \in U$ und $v \in V$: $b'(u, v) = f(u \otimes v) = b(u, v)$, also $b' = b$.

Umgekehrt sei $b \in \operatorname{Bil}(U, V; W)$ die zu $f \in \operatorname{Homo}(U \otimes V, W)$ gehörige bilineare Abbildung, definiert durch $b(u, v) = f(u \otimes v)$ für alle $u \in U$ und $v \in V$. Dann gilt für die zu b gehörige lineare Abbildung $f' \in \operatorname{Homo}(U \otimes V, W)$ und alle $u \in U$ und $v \in V$: $f'(u \otimes v) = b(u, v) = f(u \otimes v)$, also $f' = f$.

Schritt 2: Es bleibt noch die Linearität zu zeigen. Es seien also $b_1, b_2\colon U \times V \longrightarrow W$ bilineare Abbildungen. Weiter seien $f_1, f_2 \in \operatorname{Homo}(U \otimes V, W)$ die zugehörigen linearen Abbildungen. Dann gilt

$$\begin{aligned} &(b_1 + b_2)(u, v) \\ &\quad = b_1(u, v) + b_2(u, v) && \text{wegen der Def. der Addition in } \operatorname{Bil}(U, V; W) \\ &\quad = f_1(u \otimes v) + f_2(u \otimes v) && \text{da } f_i \text{ zu } b_i \text{ gehört, also } b_i(u, v) = f_i(u \otimes v) \text{ gilt} \\ &\quad = (f_1 + f_2)(u \otimes v) && \text{wegen der Def. der Addition in } \operatorname{Homo}(U \otimes V, W) \end{aligned}$$

für alle $(u,v) \in U \times V$. Für bilineares $b\colon U \times V \to W$ und die zugehörige lineare Abbildung $f \in \operatorname{Homo}(U \otimes V, W)$ sowie $\alpha \in K$ gilt außerdem

$$\begin{aligned}
&(\alpha\, b)(u,v) \\
&\qquad = \alpha\, b(u,v) && \text{wegen der Def. der S-Multiplikation in } \operatorname{Bil}(U,V;W) \\
&\qquad = \alpha\, f(u \otimes v) && \text{da } f \text{ zu } b \text{ gehört} \\
&\qquad = (\alpha\, f)(u \otimes v) && \text{wegen der Def. der S-Multiplikation in } \operatorname{Homo}(U \otimes V, W)
\end{aligned}$$

für alle $(u,v) \in U \times V$. Das zeigt, dass $f_1 + f_2$ die zu $b_1 + b_2$ gehörige lineare Abbildung und dass $\alpha\, f$ die zu $\alpha\, b$ gehörige lineare Abbildung ist. □

Bemerkung 23.9 (Angabe linearer Abbildungen auf Tensorprodukträumen)**.**
Der Satz 23.8 eröffnet uns die Möglichkeit, lineare Abbildungen $f\colon U \otimes V \to W$ durch ihre Bilder auf den Elementartensoren $u \otimes v$ festzulegen, sofern die zugehörige Abbildungsvorschrift bilinear ist. Das liest sich dann häufig in etwa so: „Es sei $f\colon U \otimes V \to W$ diejenige lineare Abbildung, deren Bilder auf Elementartensoren $u \otimes v$ durch die bilineare Abbildungsvorschrift …festgelegt sind." Wir werden von dieser Möglichkeit im folgenden Beispiel bereits Gebrauch machen. △

Beispiel 23.10 (universelle Eigenschaft der universellen bilinearen Abbildung $\otimes$, vgl. Beispiel 23.2)**.**

(i) Es sei K ein Körper. Wir betrachten die bilineare Abbildung $b\colon K \times K \to K$ mit $b(\alpha,\beta) = \alpha\,\beta$. Die zugehörige lineare Abbildung $f\colon K \otimes K \to K$ ist durch die Werte $f(\alpha \otimes \beta) = \alpha\,\beta$ auf den Elementartensoren $\alpha \otimes \beta$ für $\alpha, \beta \in K$ festgelegt.

(ii) Es seien K ein Körper und $A \in K^{n \times m}$, $n, m \in \mathbb{N}_0$. Wir betrachten die Bilinearform $b\colon K^m \times K^n \to K$ mit $b(x,y) = y^\intercal A\, x$. Die zugehörige Linearform $f\colon K^m \otimes K^n \to K$ ist durch die Werte $f(x \otimes y) = y^\intercal A\, x$ auf den Elementartensoren $x \otimes y$ für $x \in K^m$ und $y \in K^n$ festgelegt.

(iii) Es seien K ein Körper und $n, m, \ell \in \mathbb{N}_0$. Wir betrachten die bilineare Abbildung $b\colon K^{n \times m} \times K^{m \times \ell} \to K^{n \times \ell}$ mit $b(A,B) = A\,B$. Die zugehörige lineare Abbildung $f\colon K^{n \times m} \otimes K^{m \times \ell} \to K^{n \times \ell}$ ist durch die Werte $f(A \otimes B) = A\,B$ auf den Elementartensoren $A \otimes B$ für $A \in K^{n \times m}$ und $B \in K^{m \times \ell}$ festgelegt.

(iv) Es sei V ein Vektorraum über einem Körper K und V^* sein Dualraum. Wir betrachten die Bilinearform (duale Paarung) $b\colon V^* \times V \to K$ mit $b(v^*, v) = \langle v^*, v\rangle$. Die zugehörige Linearform $f\colon V^* \otimes V \to K$ ist durch die Werte $f(v^* \otimes v) = \langle v^*, v\rangle$ auf den Elementartensoren $v^* \otimes v$ für $v^* \in V^*$ und $v \in V$ festgelegt.

(v) Es seien U, V Vektorräume über demselben Körper K sowie $u^* \in U^*$ und $v^* \in V^*$ gegeben. Wir betrachten die Bilinearform $b\colon U \times V \to K$ mit $b(u,v) = \langle u^*, u\rangle\,\langle v^*, v\rangle$. Die zugehörige Linearform $f\colon U \otimes V \to K$ ist durch die Werte $f(u \otimes v) = \langle u^*, u\rangle\,\langle v^*, v\rangle$ auf den Elementartensoren $u \otimes v$ für $u \in U$ und $v \in V$ festgelegt.

(vi) Es seien X eine Menge und K ein Körper. Wir betrachten die bilineare Abbildung $b\colon K^X \times K^X \to K^X$ mit $b(f,g) = f \cdot g$ (punktweise Multiplikation). Die zugehörige lineare Abbildung $F\colon K^X \otimes K^X \to K^X$ ist durch die Werte $F(f \otimes g) = f \cdot g$ auf den Elementartensoren $f \otimes g$ für $f, g \in K^X$ festgelegt. △

Man sieht relativ leicht, dass es zu Vektorräumen U, V verschiedene Tensorprodukte gibt. Beispielsweise können wir die bilineare Abbildung $\otimes \colon U \times V \to U \otimes V$ durch $\alpha \otimes$ mit $\alpha \in K \setminus \{0\}$ ersetzen und erhalten wieder eine universelle bilineare Abbildung. Zudem können wir den Tensorproduktraum $U \otimes V$ durch ein beliebiges isomorphes Abbild ersetzen und $\otimes$ entsprechend anpassen.

Allerdings sind alle Tensorprodukte von U und V im Wesentlichen, d. h. bis auf Isomorphie, gleich:

Satz 23.11 (Eindeutigkeit des Tensorprodukts)**.**
Es seien U und V Vektorräume über demselben Körper K.

(i) Sind $(U \otimes V, \otimes)$ und $(U \widetilde{\otimes} V, \widetilde{\otimes})$ zwei Tensorprodukte von U und V, dann gibt es einen Isomorphismus $I \in \operatorname{Homo}(U \otimes V, U \widetilde{\otimes} V)$ mit der Eigenschaft $\widetilde{\otimes} = I \circ \otimes$, also $I(u \otimes v) = u \widetilde{\otimes} v$ für alle $u \in U$ und $v \in V$. Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} & U \times V \arrow[dl, "\otimes"'] \arrow[dr, "\widetilde{\otimes}"] & \\ U \otimes V \arrow[rr, "I"'] & & U \widetilde{\otimes} V \end{tikzcd}$$

(ii) Ist $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V und ist $I \in \operatorname{Homo}(U \otimes V, U \widetilde{\otimes} V)$ ein Isomorphismus mit einem weiteren K-Vektorraum $U \widetilde{\otimes} V$, dann ist auch $(U \widetilde{\otimes} V, \widetilde{\otimes})$ ein Tensorprodukt von U und V mit $\widetilde{\otimes} \coloneqq I \circ \otimes \colon U \times V \to U \widetilde{\otimes} V$.

Beweis. Aussage (i): Wir betrachten die universelle bilineare Abbildung $\widetilde{\otimes} \colon U \times V \to U \widetilde{\otimes} V$ des zweiten Tensorprodukts $(U \widetilde{\otimes} V, \widetilde{\otimes})$. Da dies eine bilineare Abbildung von $U \times V$ in einen Vektorraum (nämlich $U \widetilde{\otimes} V$) ist, können wir sie unter Ausnutzung der Universilitätseigenschaft des ersten Tensorprodukts mit Hilfe einer linearen Abbildung $f_1 \in \operatorname{Homo}(U \otimes V, U \widetilde{\otimes} V)$ in der Form $\widetilde{\otimes} = f_1 \circ \otimes$ darstellen. Mit demselben Argument und vertauschten Rollen erhalten wir eine lineare Abbildung $f_2 \in \operatorname{Homo}(U \widetilde{\otimes} V, U \otimes V)$ mit der Eigenschaft $\otimes = f_2 \circ \widetilde{\otimes}$.

Wir müssen zeigen, dass f_1 und f_2 Inverse voneinander sind; dann ist $I = f_1$ der gesuchte Isomorphismus. Durch Einsetzen ergibt sich

$$\widetilde{\otimes} = f_1 \circ \otimes = f_1 \circ (f_2 \circ \widetilde{\otimes}) = (f_1 \circ f_2) \circ \widetilde{\otimes}.$$

Andererseits gilt natürlich auch

$$\widetilde{\otimes} = \mathrm{id}_{U \widetilde{\otimes} V} \circ \widetilde{\otimes}.$$

Das sind nun aber zwei konkurrierende Darstellungen der bilinearen Abbildung $\widetilde{\otimes} \colon U \times V \to U \widetilde{\otimes} V$ der Form „$\widetilde{\otimes}$ gefolgt von einer linearen Abbildung“. Aufgrund der Eindeutigkeit der Darstellung muss $f_1 \circ f_2 = \mathrm{id}_{U \widetilde{\otimes} V}$ gelten.

Durch Vertauschen der Rollen erhalten wir ganz analog auch $f_2 \circ f_1 = \mathrm{id}_{U \otimes V}$. Damit ist gezeigt, dass f_1 und f_2 tatsächlich Inverse voneinander sind.

Der Beweis von Aussage (ii) ist Gegenstand der Übung. □

Aufgrund von Satz 23.11 können wir weiter von „einem“ Tensorprodukt der Vektorräume U und V sprechen, insbesondere um zu betonen, dass es verschiedene Realisierungen von $(U \otimes V, \otimes)$ gibt. Da diese aber alle isomorph sind, ist es auch legitim, von „dem“ Trensorprodukt von U und V zu sprechen, vor allem, wenn wir das abstrakte Konzept betonen wollen, unabhängig von einer konkreten Realisierung.

Ende der Vorlesung 5

Ende der Woche 2

§ 23.3 Eigenschaften des Tensorprodukts

In diesem Abschnitt wollen wir Eigenschaften von Tensorprodukten (also der Kombination eines Tensorproduktraumes und der zugehörigen universellen bilinearen Abbildung) untersuchen und auch das „Rechnen“ mit Tensoren einüben. Dass wir noch gar keine Realisierung des Tensorprodukts konstruiert haben, wird uns davon nicht abhalten, denn wir werden nur die charakteristischen Eigenschaften des Tensorprodukts nutzen. Das ist in etwa vergleichbar mit unserem gewohnten Umgang mit reellen Zahlen: Wir nutzen ihre Eigenschaften und „Rechenregeln“ routiniert, ohne ständig an die konkrete Realisierung, etwa mit Dedekindschen Schnitten (Anhang A) zu denken oder zu erinnern.

Um die Gestalt der Elemente in einem Tensorproduktraum $U \otimes V$ besser verstehen zu können, zeigen wir jetzt, dass sich jeder Tensor als Linearkombination (und sogar als Summe) von Elementartensoren schreiben lässt.

Lemma 23.12 (die Elementartensoren bilden ein Erzeugendensystem)**.**
Es seien U und V Vektorräume über demselben Körper und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann gilt: Die Elementartensoren bilden ein Erzeugendensystem von $U \otimes V$. Es gilt also

$$U \otimes V = \big\langle (u \otimes v)_{(u,v) \in U \times V} \big\rangle.$$

Folglich lässt sich jeder Tensor in $U \otimes V$ als Linearkombination von Elementartensoren schreiben. Für jedes $t \in U \otimes V$ existieren also $n \in \mathbb{N}_0$, Vektoren $u_1, \dots, u_n \in U$ und $v_1, \dots, v_n \in V$ sowie $\alpha_1, \dots, \alpha_n \in K$ mit

$$t = \sum_{i=1}^{n} \alpha_i \, u_i \otimes v_i. \tag{23.6}$$

Die Koeffizienten α_i können o. B. d. A. alle als 1 gewählt werden.

Beweis. Wir definieren den durch die Elementartensoren aufgespannten Unterraum

$$W \coloneqq \big\langle (u \otimes v)_{(u,v) \in U \times V} \big\rangle \subseteq U \otimes V.$$

Unser Ziel ist es, $W = U \otimes V$ zu zeigen. Wir betrachten dazu den Faktorraum mit seiner kanonischen Surjektion

$$\pi \colon U \otimes V \to (U \otimes V) \,/\, W.$$

Wegen $u \otimes v \in W$ gilt $\pi(u \otimes v) = 0$ für alle $u \in U$ und $v \in V$. Die Abbildung $\pi \circ \otimes \colon U \times V \to (U \otimes V) \,/\, W$ ist also eine bilineare Abbildung von $U \times V$ in einen Vektorraum, die mit der bilinearen Nullabbildung $0 \circ \otimes \colon U \times V \to (U \otimes V)\,/\,W$ übereinstimmt. Aufgrund der universellen Eigenschaft des Tensorprodukts muss $\pi = 0$ gelten. Das heißt aber $W = U \otimes V$. (**Quizfrage 23.2:** Klar?)

Jeder Tensor $t \in U \otimes V$ ist also darstellbar als Linearkombination von Elementartensoren wie in (23.6). Wegen $\alpha_i\, u_i \otimes v_i = (\alpha_i\, u_i) \otimes v_i$ können wir die Koeffizienten α_i o. B. d. A. alle als 1 wählen. □

Wir zeigen nun noch, dass die aus Basisvektoren von U und V gebildeten Elementartensoren eine Basis von $U \otimes V$ ergeben:

Lemma 23.13 (Basis und Dimension des Tensorproduktraumes $U \otimes V$)**.**
Es seien U und V Vektorräume über dem Körper K mit Basen $B_U = (u_i)_{i \in I}$ bzw. $B_V = (v_j)_{j \in J}$. Weiter sei $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann gilt:

(i) $B_{U \otimes V} \coloneqq (u_i \otimes v_j)_{(i,j) \in I \times J}$ ist eine Basis von $U \otimes V$. Sie wird die von B_U und B_V **induzierte Basis** (englisch: induced basis) oder die **Tensorproduktbasis** (englisch: tensor product basis) zu B_U und B_V von $U \otimes V$ genannt.

(ii) $\dim(U \otimes V) = \dim(U) \cdot \dim(V)$.
(Hierbei ist $0 \cdot \infty = \infty \cdot 0 = 0$ zu verstehen, $\infty \cdot n = n \cdot \infty = \infty$ für alle $n \in \mathbb{N}$ und $\infty \cdot \infty = \infty$.)

Beweis. Wir zeigen zunächst die Aussage (i).

Schritt 1: $B_{U \otimes V}$ ist eine erzeugende Familie.

Aufgrund von Lemma 23.12 können wir jeden Tensor $t \in U \otimes V$ in der Form

$$t = \sum_{(u,v) \in M} u \otimes v$$

mit einer endlichen Teilmenge $E \subseteq U \times V$ darstellen. Wegen der Basiseigenschaft von $(u_i)_{i \in I}$ können wir jedes $u \in U$ wiederum als Linearkombination der Form $u = \sum_{i \in I^{(u)}} \alpha_i^{(u)} u_i$ schreiben, wobei $I^{(u)} \subseteq I$ endlich ist und $\alpha_i^{(u)} \in K$ für alle $i \in I^{(u)}$ gilt. Ganz analog erhalten wir auch $v = \sum_{j \in J^{(v)}} \beta_j^{(v)} v_j$ für jedes $v \in V$. Einsetzen ergibt

$$\begin{aligned} t &= \sum_{(u,v) \in E} \Big(\sum_{i \in I^{(u)}} \alpha_i^{(u)} u_i \Big) \otimes \Big(\sum_{j \in J^{(v)}} \beta_j^{(v)} v_j \Big) \\ &= \sum_{(u,v) \in E} \sum_{i \in I^{(u)}} \sum_{j \in J^{(v)}} \alpha_i^{(u)} \beta_j^{(v)} u_i \otimes v_j. \end{aligned}$$

Somit haben wir t wie gewünscht als Linearkombination der Elementartensoren $u_i \otimes v_j$ mit $(i, j) \in I \times J$ dargestellt.

Schritt 2: $B_{U\otimes V}$ ist linear unabhängig.

Dazu sei $M \subseteq I \times J$ eine endliche Teilmenge und $\alpha_{ij} \in K$ für alle $(i,j) \in M$ sowie

$$\sum_{(i,j)\in M} \alpha_{ij}\, u_i \otimes v_j = 0 \in U \otimes V$$

der Nulltensor. Wir zeigen, dass alle Koeffizienten α_{ij} gleich 0 sein müssen.

Wir fixieren dazu ein Indexpaar $(i_0, j_0) \in M$ und definieren eine bestimmte Linearform auf $U \otimes V$, die den Term $u_{i_0} \otimes v_{j_0}$ in der obigen Linearkombination isoliert.

Dazu definieren wir zunächst eine Linearform u^* auf U, indem wir ihre Bilder auf der Basis $(u_i)_{i\in I}$ von U durch $\langle u^*, u_i\rangle \coloneqq \delta_{ii_0}$ für alle $i \in I$ festlegen. Außerdem definieren wir eine Linearform v^* auf V durch $\langle v^*, v_j\rangle \coloneqq \delta_{jj_0}$ für alle $j \in J$.

Wir kombinieren diese beiden Linearformen zu einer Bilinearform

$$U \times V \ni (u,v) \mapsto \langle u^*, u\rangle \, \langle v^*, v\rangle \in K.$$

Aufgrund der universellen Eigenschaft des Tensorprodukts existiert dann eine eindeutig bestimmte Linearform $f \in \operatorname{Homo}(U \otimes V, K)$ mit der Eigenschaft $f(u \otimes v) = \langle u^*, u\rangle \, \langle v^*, v\rangle$ für alle $u \in U$ und $v \in V$. Es folgt

$$\begin{aligned}
0 &= f\Big(\sum_{(i,j)\in M} \alpha_{ij}\, u_i \otimes v_j\Big) && \text{wegen der Linearität von } f\\
&= \sum_{(i,j)\in M} \alpha_{ij}\, f(u_i \otimes v_j) && \text{wegen der Linearität von } f\\
&= \sum_{(i,j)\in M} \alpha_{ij}\, \langle u^*, u_i\rangle \, \langle v^*, v_j\rangle && \text{nach Definition von } f\\
&= \sum_{(i,j)\in M} \alpha_{ij}\, \delta_{ii_0}\, \delta_{jj_0} && \text{nach Definition von } u^* \text{ und } v^*\\
&= \alpha_{i_0 j_0}.
\end{aligned}$$

Da das Indexpaar $(i_0, j_0) \in M$ beliebig war, ist B in der Tat linear unabhängig und damit eine Basis von $U \otimes V$.

Die Aussage *(ii)* folgt sofort aus Aussage *(i)* und der Tatsache, dass die Anzahl der Elemente von B gleich $\#(I \times J) = \#I \cdot \#J = \dim(U) \cdot \dim(V)$ ist. □

Wir haben nun genügend Informationen, um ein erstes Beispiel eines Tensorproduktraumes anzugeben. Wir wählen dazu einen endlichen Körper und Vektorräume U und V endlicher Dimension, sodass auch $U \otimes V$ endliche Dimension und damit nur endlich viele Elemente besitzt, die wir explizit aufzählen können.

Beispiel 23.14 (Tensorproduktraum).
Wir betrachten den Körper $K = \mathbb{Z}_2$ und das Tensorprodukt $U \otimes V$ der Vektorräume $U = V = K^2$. Dieser Raum hat nach Lemma 23.13 die Dimension 4, also $2^4 = 16$ Elemente, die wir nun alle angeben. Insbesondere stellen wir jeden Tensor auf zwei Arten dar: einmal als die eindeutige

Linearkombination der Tensoren der (aus der Standardbasis gebildeten) Tensorproduktbasis $\binom{1}{0} \otimes \binom{1}{0}$, $\binom{1}{0} \otimes \binom{0}{1}$, $\binom{0}{1} \otimes \binom{1}{0}$ und $\binom{0}{1} \otimes \binom{0}{1}$ und nochmal in Form einer Linearkombination von Elementartensoren mit möglichst wenigen Summanden.

Neben dem

$$\text{Nulltensor} = 0 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1}$$

gibt es die vier Basistensoren

$$\begin{aligned}
1 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{1}{0}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{0}{1}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{0}{1} \otimes \binom{1}{0}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{0}{1} \otimes \binom{0}{1}
\end{aligned}$$

sowie die fünf weiteren Elementartensoren

$$\begin{aligned}
1 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{1}{1}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{0}{1} \otimes \binom{1}{1}\\
1 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{1} \otimes \binom{1}{0}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{1} \otimes \binom{0}{1}\\
1 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{1} \otimes \binom{1}{1}
\end{aligned}$$

und schließlich die sechs nicht-einfachen Tensoren

$$\begin{aligned}
1 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{1}{0} + \binom{0}{1} \otimes \binom{0}{1}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{0}{1} + \binom{0}{1} \otimes \binom{1}{0}\\
1 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 0 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{1}{1} + \binom{0}{1} \otimes \binom{1}{0}\\
1 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 0 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{0} \otimes \binom{1}{1} + \binom{0}{1} \otimes \binom{0}{1}\\
1 \binom{1}{0} \otimes \binom{1}{0} + 0 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{1} \otimes \binom{1}{0} + \binom{0}{1} \otimes \binom{0}{1}\\
0 \binom{1}{0} \otimes \binom{1}{0} + 1 \binom{1}{0} \otimes \binom{0}{1} + 1 \binom{0}{1} \otimes \binom{1}{0} + 1 \binom{0}{1} \otimes \binom{0}{1} &= \binom{1}{1} \otimes \binom{0}{1} + \binom{0}{1} \otimes \binom{1}{0}. \qquad \triangle
\end{aligned}$$

Bemerkung 23.15 (zu Tensorprodukten)**.**

(*i*) Die Darstellung eines Tensors wie in (23.6) ist nicht eindeutig. Beispielsweise gilt

$$\begin{aligned}
\binom{1}{0} \otimes \binom{1}{1} + \binom{0}{1} \otimes \binom{1}{0} &= \binom{1}{0} \otimes \binom{1}{0} + \binom{1}{0} \otimes \binom{0}{1} + \binom{0}{1} \otimes \binom{1}{0}\\
&= \binom{1}{1} \otimes \binom{1}{0} + \binom{1}{0} \otimes \binom{0}{1}.
\end{aligned}$$

(*ii*) Das Symbol $\otimes$ („**Tensorprodukt**“) tritt in zwei Bedeutungen auf. Erstens dient es zur Bezeichnung des Vektorraumes $U \otimes V$ und zweitens als Name für die universelle bilineare Abbildung $\otimes\colon U \times V \to U \otimes V$, also die Verknüpfung von zwei Vektoren zu einem Tensor. Um Missverständnisse zu vermeiden, werden wir darauf verzichten, $\otimes$ im Sinne der Bemerkung 7.20 für die Verknüpfung von Mengen von Tensoren zu verwenden. Andernfalls könnte $U \otimes V$ auch als

$$\{u \otimes v \mid u \in U,\ v \in V\}$$

verstanden werden, was i. A. aber eine echte Teilmenge des Vektorraumes $U \otimes V$ ist, denn i. A. ist nicht jeder Tensor in $U \otimes V$ einfach. △

Wie das Beispiel 23.14 bereits zeigt, sind i. A. nicht alle Tensoren eines Tensorproduktraumes $U \otimes V$ einfache Tensoren.[10] Manche benötigen also mehr als einen Summanden in der Darstellung (23.6).[11] Das führt uns auf die folgende Definition:

Definition 23.16 (Rang eines Tensors)**.**
Es seien U und V Vektorräume über demselben Körper und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Der **Rang** eines Tensors (englisch: rank of a tensor) $t \in U \otimes V$, geschrieben $\operatorname{Rang}(t) \in \mathbb{N}_0$, ist die minimale Anzahl von Summanden, mit denen eine Darstellung der Form (23.6) möglich ist. △

Bemerkung 23.17 (Rang eines Tensors)**.**

(i) Da die Elementartensoren ein Erzeugendensystem von $U \otimes V$ bilden (Lemma 23.12), kann jeder Tensor $t \in U \otimes V$ als Linearkombination von Elementartensoren dargestellt werden. Daher ist der Rang eines Tensors tatsächlich immer endlich.

(ii) In der Quantenmechanik bezeichnet man Tensoren vom Rang $\geqslant 2$ als **verschränkte Tensoren** (englisch: entangled tensor). △

Wir werden in § 23.6 sehen, wie wir den Rang eines Tensors bestimmen können. Zunächst geht es nur um den Rang von Elementartensoren:

Lemma 23.18 (Rang von Elementartensoren)**.**
Es seien U und V Vektorräume mit Basen über demselben Körper und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann gilt:

(i) Der Nulltensor ist der einzige Tensor vom Rang 0.

(ii) Für $u \in U$ und $v \in V$ gilt

$$u \otimes v = 0 \quad \text{(Nulltensor)} \quad \Leftrightarrow \quad u = 0 \text{ oder } v = 0 \quad \text{(Nullvektor)}. \tag{23.7}$$

(iii) Jeder Elementartensor $u \otimes v$ mit $u, v \neq 0$ ist vom Rang 1.

Beweis. Aussage (i): Der Nulltensor kann durch die leere Summe in (23.6) dargestellt werden, die ja per Definition immer das neutrale Element — hier in der abelschen Gruppe $(U \otimes V, +)$ — ergibt. Damit ist $\operatorname{Rang}(0) = 0$ bestätigt und auch gezeigt, dass alle anderen Tensoren mindestens Rang 1 haben.

Aussage (ii): Wenn $u = 0$ oder $v = 0$ gilt, dann folgt $u \otimes v = 0$ aus der Bilinearität der universellen bilinearen Abbildung. (**Quizfrage 23.3:** Wie nämlich?)

Es seien nun $u \neq 0$ sowie $v \neq 0$. Wir benötigen jetzt Basen von U und V, die wir mit $(u_i)_{i \in I}$ bzw. $(v_j)_{j \in J}$ bezeichnen und deren Existenz wir angenommen haben. Wir stellen u und v als Linearkombinationen der Basisvektoren dar:

$$u = \sum_{i \in I_0} \alpha_i \, u_i \quad \text{und} \quad v = \sum_{j \in J_0} \beta_j \, v_j.$$

[10] Mit anderen Worten, die universelle bilineare Abbildung $\otimes \colon U \times V \to U \otimes V$ ist i. A. nicht surjektiv.
[11] Tatsächlich sind i. A. sogar „die meisten" Tensoren vom Rang $\geqslant 2$.

Aufgrund der Bilinearität des Tensorprodukts folgt

$$u \otimes v = \sum_{(i,j) \in I_0 \times J_0} \alpha_i \, \beta_j \, u_i \otimes v_j.$$

Mindestens ein Koeffizient $\alpha_i \, \beta_j$ ist dabei ungleich Null, da $u \neq 0$ und $v \neq 0$ gilt. Da die Familie der Elementartensoren $(u_i \otimes v_j)_{(i,j) \in I \times J}$ eine Basis von $U \otimes V$ bildet (Lemma 23.13), ist $u \otimes v$ nicht der Nulltensor.

Wir beweisen schließlich Aussage *(iii)*: Für $u, v \neq 0$ ist $u \otimes v$ nach Aussage *(ii)* nicht der Nulltensor. Nach Aussage *(i)* besitzt er also mindestens Rang 1. Andererseits besitzt $u \otimes v$ offenbar eine Darstellung von der Form (23.6) mit einem Summanden, hat also höchstens Rang 1. □

Beispiel 23.19 (Rang eines Tensors, vgl. Beispiel 23.14)**.**
Wir betrachten wie in Beispiel 23.14 das Tensorprodukt $\mathbb{Z}_2^2 \otimes \mathbb{Z}_2^2$.

$$\begin{aligned} &\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \otimes \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) + \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \otimes \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \otimes \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right) \quad && \text{hat Rang } 1 \\ &\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \otimes \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) + \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) \otimes \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) && \text{hat Rang } 2. \end{aligned}$$

Insgesamt hat der Tensorproduktraum $\mathbb{Z}_2^2 \otimes \mathbb{Z}_2^2$ einen Tensor vom Rang 0, neun Tensoren vom Rang 1 und sechs Tensoren vom Rang 2. Tensoren von einem höheren Rang als 2 gibt es im Tensorproduktraum $\mathbb{Z}_2^2 \otimes \mathbb{Z}_2^2$ nicht. △

Die Frage, wie groß der Rang eines Tensors in $U \otimes V$ i. A. maximal sein kann, beantworten wir in Folgerung 23.39.

Wir kommen noch einmal zu sprechen auf die Isomorphie von Tensorprodukten. Wie Satz 23.11 besagt, kann in einem Tensorprodukt der Tensorproduktraum immer ausgetauscht werden durch einen isomorphen Vektorraum, also einen Vektorraum derselben Dimension. Dabei ist auch die universelle bilineare Abbildung entsprechend auszutauschen.

Beispiel 23.20 (isomorphe Tensorprodukte)**.**

(i) Es sei K ein Körper $(K \otimes K, \otimes)$ ein Tensorprodukt von K mit sich selbst. Der Vektorraum $K \otimes K$ besitzt nach Lemma 23.13 die Dimension 1, er ist also isomorph zu K.

Jeder Isomorphismus ist nach Bemerkung 23.9 eindeutig festgelegt durch die Bilder auf den Elementartensoren, sofern die Abbildungsvorschrift bilinear ist. Ein möglicher Isomorphismus ist also $\alpha \otimes \beta \mapsto \alpha \, \beta$. Dieser wird als kanonisch bezeichnet, weil keine weitere Wahl von Daten getroffen werden muss. (**Quizfrage 23.4:** Gibt es noch weitere Isomorphismen?)

(ii) Allgemeiner gibt es für jeden K-Vektorraum U einen kanonischen Isomorphismus $K \otimes U \cong U$. (**Quizfrage 23.5:** Wie lautet die Abbildungsvorschrift für diesen Isomorphismus?)

(iii) Für je zwei K-Vektorräume U und V gibt es einen kanonischen Isomorphismus $U \otimes V \cong V \otimes U$. (**Quizfrage 23.6:** Wie ist der definiert?) △

Wir geben als ein weiteres Beispiel noch an, wie man mit Hilfe des Tensorprodukts einen Vektorraum über dem Körper $\mathbb{R}$ zu einem Vektorraum über dem Körper $\mathbb{C}$ erweitern kann.

Beispiel 23.21 (Komplexifizierung eines $\mathbb{R}$-Vektorraumes).
Es sei V ein beliebiger Vektorraum über dem Körper $\mathbb{R}$ mit der Basis $(v_j)_{j \in J}$. Weiter sei $U = \mathbb{C}$, aufgefasst als zweidimensionaler Vektorraum über $\mathbb{R}$ mit der Basis $(1, i)$. Dann bildet nach Lemma 23.13 die Familie $(1 \otimes v_j)_{j \in J} \parallel (i \otimes v_j)_{j \in J}$ eine Basis des $\mathbb{R}$-Vektorraumes $\mathbb{C} \otimes V$.

Jedes Element von $\mathbb{C} \otimes V$ hat eine (bis auf Nullkoeffizienten eindeutige) Darstellung

$$\begin{aligned}\widehat{v} &= \sum_{j \in J_0} \alpha_j \, (1 \otimes v_j) + \sum_{j \in J_0} \beta_j \, (i \otimes v_j) = \sum_{j \in J_0} \alpha_j \otimes v_j + \sum_{j \in J_0} (\beta_j \, i) \otimes v_j \\ &= \Big(\sum_{j \in J_0} (\alpha_j + \beta_j \, i) \Big) \otimes v_j \eqqcolon \sum_{j \in J_0} \gamma_j \otimes v_j\end{aligned}$$

mit einer endlichen Indexmenge J_0 sowie Koeffizienten $\alpha_j, \beta_j \in \mathbb{R}$ bzw. $\gamma_j \coloneqq \alpha_j + \beta_j \, i \in \mathbb{C}$.

Wir können nun im $\mathbb{R}$-Vektorraum $\mathbb{C} \otimes V$ die S-Multiplikation mit Skalaren in $\mathbb{R}$ auf Skalare in $\mathbb{C}$ erweitern, indem wir für $\widehat{v}$ mit obiger Darstellung und $\mu \in \mathbb{C}$ definieren:

$$\mu \, \widehat{v} \coloneqq \sum_{j \in J_0} (\mu \, \gamma_j) \otimes v_j.$$

Wir können mit dieser Definition nun nachprüfen, dass $\mathbb{C} \otimes V$ nicht nur ein Vektorraum über $\mathbb{R}$, sondern auch ein Vektorraum über $\mathbb{C}$ ist und als solcher die Basis $(1 \otimes v_j)_{j \in J}$ besitzt. Dieser $\mathbb{C}$-Vektorraum $\mathbb{C} \otimes V$ heißt die **Komplexifizierung** (englisch: complexification) des $\mathbb{R}$-Vektorraumes V. Der ursprüngliche Raum V wird in seiner isomorphen Gestalt $1 \otimes V$ zu einem $\mathbb{R}$-Unterraum von $\mathbb{C} \otimes V$. △

§ 23.4 Konstruktion eines Tensorprodukts

Wir wollen uns in diesem Abschnitt davon überzeugen, dass es Tensorprodukte mit den in Definition 23.6 festgelegten Eigenschaften tatsächlich gibt. Dazu werden wir, gegeben zwei beliebige K-Vektorräume U und V, einen konkreten K-Vektorraum $U \otimes_/ V$ und die zugehörige universelle bilineare Abbildung $\otimes_/ \colon U \times V \to U \otimes_/ V$ konstruieren, also eine konkrete Realisierung des abstrakten Konzepts „Tensorprodukt". Wir wissen nach Satz 23.11 bereits, dass auch andere (isomorphe) Konstruktionen möglich sind.

Wir erinnern daran, dass wir im Tensorprodukt zwar die Vektorraumstruktur von U und V einzeln nutzen, dass die Vektorraum-Eigenschaft des Produktraumes $U \times V$ jedoch inkompatibel mit der Bilinearität ist und wir $U \times V$ deshalb nur als Menge betrachten können (Bemerkung 23.5). Daher wird es bei der Konstruktion des Tensorproduktraumes notwendig sein, zunächst einen Vektorraum zu konstruieren, der die Elemente von $U \times V$ als Vektoren enthält. Das ist für jede Menge X möglich und führt auf das Konzept des **freien Vektorraumes**.

Der freie Vektorraum über einer Menge

Der Gedanke hinter der Konstruktion eines **freien Vektorraumes** ist, dass wir eine beliebige Menge X in einen Vektorraum einbetten wollen, sodass die Elemente der Menge als Vektoren interpretiert werden können.

Definition 23.22 (freier Vektorraum über einer Menge)**.**
Es sei X eine Menge und K ein Körper. Der **freie K-Vektorraum über der Menge X** (englisch: free K-vector space) ist der K-Vektorraum

$$(K^X)_{00} \coloneqq \{f\colon X \to K \mid \operatorname{supp} f \text{ ist endlich}\} \tag{23.8}$$

der endlich getragenen Funktionen $X \to K$. △

Die Familie der charakteristischen Funktionen (Beispiel 12.2) $(e_y)_{y\in X}$ bildet eine Basis von $(K^X)_{00}$. Es gilt somit $\dim((K^X)_{00}) = \#X$. Zur Erinnerung: Die charakteristische Funktion $e_y\colon X \to K$ ist durch die Werte

$$e_y(x) \coloneqq \begin{cases} 1, & \text{falls } x = y, \\ 0, & \text{falls } x \neq y \end{cases}$$

definiert.

Beachte: Die Abbildung $e_\bullet\colon X \ni y \mapsto e_y \in (K^X)_{00}$ ist eine Einbettung (also eine injektive Funktion) von X in $(K^X)_{00}$. Dadurch gelingt es, die Elemente einer beliebigen Menge X mit Vektoren in einem Vektorraum zu identifizieren. Jedes Element von X steht dabei für eine Dimension im Vektorraum $(K^X)_{00}$.

Beispiel 23.23 (freier Vektorraum)**.**

(*i*) Der freie Vektorraum $(K^\emptyset)_{00}$ über der leeren Menge $\emptyset$ ist der Nullvektorraum, denn er hat nur ein Element (die leere Funktion $\emptyset \to K$).

(*ii*) Für jede einelementige Menge $\{x\}$ ist der freie Vektorraum $(K^{\{x\}})_{00}$ über $\{x\}$ isomorph zum Körper K.

(*iii*) Für jede endliche Menge $X = \{x_1, \ldots, x_n\}$ mit $n \in \mathbb{N}_0$ Elementen ist der freie Vektorraum $(K^X)_{00}$ über X isomorph zum Produktraum $\bigtimes_{i=1}^n K$ sowie zum Standardvektorraum K^n. △

Mit Hilfe des freien Vektorraumes $(K^X)_{00}$ über der Menge X können wir nun beliebige Funktionen $X \to W$ in einen K-Vektorraum W als lineare Abbildungen $(K^X)_{00} \to W$ verstehen:

Satz 23.24 (universelle Eigenschaft des freien Vektorraumes)**.**
Es sei X eine Menge, K ein Körper und W ein Vektorraum über K.

(*i*) Zu jeder Funktion $f\colon X \to W$ existiert genau eine lineare Abbildung $g\colon (K^X)_{00} \to W$ mit der Eigenschaft $f = g \circ e_\bullet$, also $f(x) = g(e_x)$ für alle $x \in X$. Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} X \arrow[r, "e_\bullet"] \arrow[dr, "f"'] & (K^X)_{00} \arrow[d, "g"] \\ & W \end{tikzcd}$$

Man sagt auch: Die Funktion f **faktorisiert durch** (englisch: factors through) die Einbettung $e_\bullet\colon X \to (K^X)_{00}$. Die Funktion $g\colon (K^X)_{00} \to W$ wird auch die **lineare Fortsetzung** (englisch: linear extension) von $f\colon X \to W$ genannt.

(ii) Umgekehrt wird für jede lineare Abbildung $g\colon (K^X)_{00} \to W$ durch $f(x) \coloneqq g(e_x)$ für alle $x \in X$ eine Funktion $f\colon X \to W$ definiert.

(iii) Die Zuordnung

$$W^X \ni f \mapsto g \in \operatorname{Homo}((K^X)_{00}, W)$$

ist ein kanonischer Isomorphismus von Vektorräumen.

Beachte: Wir können also in der Tat beliebige Funktionen $X \to W$ mit Werten in beliebigen K-Vektorräumen W mit linearen Abbildungen von $(K^X)_{00}$ in W im Sinne von Vektorräumen identifizieren!

Beweis. Aussage (i): Es sei $f\colon X \to W$ eine beliebige Funktion. Wir definieren die lineare Abbildung $g\colon (K^X)_{00} \to W$ durch Festlegung der Bilder auf der Basis $(e_x)_{x\in X}$ (Satz 17.10) und setzen

$$g(e_x) \coloneqq f(x) \quad \text{für alle } x \in X.$$

Dann gilt $f = g \circ e_\bullet$, und diese Wahl von g ist die einzig mögliche.

Die Aussage (ii) ist offensichtlich.

Aussage (iii):

Schritt 1: Wir müssen zunächst beweisen, dass die Zuordnung $f \mapsto g$ bijektiv ist. Dazu zeigen wir, dass $f \mapsto g \mapsto f'$ die Identität auf W^X und dass $g \mapsto f \mapsto g'$ die Identität auf $\operatorname{Homo}((K^X)_{00}, W)$ ist (vgl. Definition 6.22 und Lemma 6.24).

Es sei zuerst $g \in \operatorname{Homo}((K^X)_{00}, W)$ die zu $f \in W^X$ gehörige lineare Abbildung, definiert durch $g(e_x) = f(x)$ für alle $x \in X$. Dann ist $f'(x) = g(e_x) = f(x)$ für alle $x \in X$, also $f' = f$. Umgekehrt sei $f \in W^X$ die zu $g \in \operatorname{Homo}((K^X)_{00}, W)$ gehörige Funktion, definiert durch $f(x) = g(e_x)$ für alle $x \in X$. Dann gilt $g'(e_x) = f(x) = g(e_x)$ für alle $x \in X$.

Schritt 2: Es bleibt noch die Linearität zu zeigen. Es seien also $f_1, f_2\colon X \to W$ Funktionen. Weiter seien $g_1, g_2 \in \operatorname{Homo}((K^X)_{00}, W)$ die zugehörigen linearen Abbildungen. Dann gilt

$$\begin{aligned}
&(f_1 + f_2)(x) \\
&\quad = f_1(x) + f_2(x) && \text{wegen der Def. der Addition in } W^X \\
&\quad = g_1(e_x) + g_2(e_x) && \text{da } g_i \text{ zu } f_i \text{ gehört, also } f_i(x) = g_i(e_x) \text{ gilt} \\
&\quad = (g_1 + g_2)(e_x) && \text{wegen der Def. der Addition in } \operatorname{Homo}((K^X)_{00}, W)
\end{aligned}$$

für alle $x \in X$. Für $f\colon X \to W$ und die zugehörige lineare Abbildung $g \in \operatorname{Homo}((K^X)_{00}, W)$ sowie $\alpha \in K$ gilt außerdem

$$\begin{aligned}
&(\alpha f)(x) \\
&\quad = \alpha f(x) && \text{wegen der Def. der S-Multiplikation in } W^X
\end{aligned}$$

$$
\begin{aligned}
&= \alpha\, g(e_x) && \text{da } g \text{ zu } f \text{ gehört} \\
&= (\alpha\, g)(e_x) && \text{wegen der Def. der S-Multiplikation in } \operatorname{Homo}((K^X)_{00}, W)
\end{aligned}
$$

für alle $x \in X$. Das zeigt, dass $g_1 + g_2$ die zu $f_1 + f_2$ gehörige lineare Abbildung und dass $\alpha\, g$ die zu $\alpha\, f$ gehörige lineare Abbildung ist. □

Beispiel 23.25 (universelle Eigenschaft des freien Vektorraumes)**.**
Es sei beispielsweise $X = \{\text{rot}, \text{grün}, \text{blau}\}$, und wir betrachten die Funktion $f\colon X \to W \coloneqq \mathbb{Q}^2$ mit den Funktionswerten

$$
f(\text{rot}) = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad f(\text{grün}) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \quad \text{und} \quad f(\text{blau}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.
$$

Für den freien $\mathbb{Q}$-Vektorraum über X gilt $(\mathbb{Q}^X)_{00} = \mathbb{Q}^X$, da X endlich ist. $\mathbb{Q}^X$ sei mit der Basis $B \coloneqq (e_{\text{rot}}, e_{\text{grün}}, e_{\text{blau}})$ ausgestattet, wobei beispielsweise $e_{\text{rot}}\colon X \to \mathbb{Q}$ durch die Werte

$$
e_{\text{rot}}(\text{rot}) = 1, \quad e_{\text{rot}}(\text{grün}) = 0, \quad e_{\text{rot}}(\text{blau}) = 0
$$

festgelegt ist. Wir definieren die lineare Abbildung $g\colon \mathbb{Q}^X \to W = \mathbb{Q}^2$ durch folgende Bilder auf der Basis B:

$$
g(e_{\text{rot}}) = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad g(e_{\text{grün}}) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \quad \text{und} \quad g(e_{\text{blau}}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.
$$

Dann gilt $f = g \circ e_\bullet$, und wir können die Funktion f mit der linearen Abbildung g identifizieren. Da g auf einem Vektorraum definiert ist, können wir g auch an anderen Vektoren als den Basisvektoren $e_{\text{rot}}, e_{\text{grün}}, e_{\text{blau}}$ auswerten, beispielsweise

$$
g(2\, e_{\text{rot}} + 4\, e_{\text{blau}}) = 2\, g(e_{\text{rot}}) + 4\, g(e_{\text{blau}}) = \begin{pmatrix} 2 \\ -4 \end{pmatrix}.
$$

Die Funktion f kann dagegen nur auf den drei Elementen von X ausgewertet werden. △

Bei der Konstruktion des Tensorproduktraumes werden wir gleich nicht beliebige Funktionen $X \to W$ betrachten (mit $X = U \times V$), sondern nur solche aus einem bestimmten Unterraum (nämlich dem Unterraum der bilinearen Funktionen). Diese zusätzliche Eigenschaft können wir dadurch kodieren, dass wir nicht den gesamten freien Vektorraum $(K^X)_{00}$ verwenden, sondern nur einen geeigneten Faktorraum.

Ist also S ein Unterraum des Vektorraumes der Funktionen $X \to W$, dann können wir einen passenden Unterraum N von $(K^X)_{00}$ finden, sodass jedes $f \in S$ nicht nur durch $(K^X)_{00}$, sondern sogar durch den kleineren Faktorraum $(K^X)_{00} / N$ faktorisiert. Das bedeutet, dass das obenstehende kommutative Diagramm wie folgt erweitert werden kann:

$$
\begin{tikzcd}
X \arrow[r, "e_\bullet"] \arrow[dr, "f"'] & (K^X)_{00} \arrow[r, "\pi"] \arrow[d, "g"] & (K^X)_{00} / N \arrow[dl, "\widehat{g}"] \\
& W &
\end{tikzcd}
$$

Wir wollen nun begründen, wie man ausgehend vom Unterraum S einen passenden Unterraum N von $(K^X)_{00}$ bestimmen kann. Dabei sind wir an einem möglichst großen Unterraum N interessiert, sodass der „Flaschenhals" der Faktorisierung — der Faktorraum $(K^X)_{00} / N$ — möglichst klein wird. Das ist auch erforderlich dafür, dass die Zuordnung $f \mapsto \widehat{g}$ von S in $\operatorname{Homo}((K^X)_{00} / N, W)$ wieder bijektiv wird und nicht etwa mehrere Funktionen $\widehat{g}$ dieselbe Funktion f ergeben.

Eine lineare Abbildung $g \in \operatorname{Homo}((K^X)_{00}, W)$ faktorisiert genau dann durch den Faktorraum $(K^X)_{00} / N$, wenn der Unterraum $N \subseteq \operatorname{Kern}(g)$ erfüllt.[12] Wir benötigen diese Eigenschaft nun aber nicht nur für eine bestimmte Abbildung g, sondern für alle diejenigen g, die zu Funktionen $f \in S$ gehören (die also $f = g \circ e_\bullet$ für ein $f \in S$ erfüllen). Da wir an einem möglichst großen Unterraum N interessiert sind, setzen wir

$$N \coloneqq \bigcap \big\{\operatorname{Kern}(g) \,\big|\, g \in \operatorname{Homo}((K^X)_{00}, W),\ g \circ e_\bullet \in S\big\}.$$

Wie das obenstehende Diagramm andeutet, können wir dann f nicht nur durch $(K^X)_{00}$, sondern sogar durch den kleineren Faktorraum $(K^X)_{00} / N$ faktorisieren. Für jedes $f \in S$ existiert ein eindeutig bestimmtes $\widehat{g} \in \operatorname{Homo}((K^X)_{00} / N, W)$ mit der Eigenschaft $f = \widehat{g} \circ \pi \circ e_\bullet$.

Beispiel 23.26 (Faktorräume des freien Vektorraumes kodieren zusätzliche Eigenschaften, vgl. Beispiel 23.25).
Es sei K und Körper und W ein beliebiger K-Vektorraum. Wir betrachten die dreielementige Menge $X = \{\text{rot}, \text{grün}, \text{blau}\}$. Anstelle beliebiger Funktionen $X \to W$ betrachten wir den Unterraum

$$S \coloneqq \{f \colon X \to W \mid f(\text{rot}) + f(\text{grün}) - f(\text{blau}) = 0\}.$$

Nach Satz 23.24 existiert zu jedem $f \in W^X$ und insbesondere zu jedem $f \in S$ eine eindeutig bestimmte lineare Abbildung $g \colon (K^X)_{00} \to W$ mit $f = g \circ e_\bullet$, wie in Beispiel 23.25 illustiert.

Wenn wir uns jedoch auf Funktionen $f \in S$ beschränken, so können wir die zusätzlichen Eigenschaften von f dadurch kodieren, dass wir einen geeigneten Faktorraum $(K^X)_{00} / N$ nach einem passenden Unterraum N verwenden. Der größtmögliche Unterraum ist gegeben durch

$$N \coloneqq \bigcap \big\{\operatorname{Kern}(g) \,\big|\, g \in \operatorname{Homo}((K^X)_{00}, W),\ g \circ e_\bullet \in S\big\}.$$

Die Bedingung $g \circ e_\bullet \in S$ bedeutet hier

$$g(e_{\text{rot}}) + g(e_{\text{grün}}) - g(e_{\text{blau}}) = g(e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}}) = 0.$$

Das heißt, dass jedes solche g notwendigerweise $\langle e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}} \rangle$ im Kern hat. Also gilt $\langle e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}} \rangle \subseteq N$. Andererseits können wir ein g mit der Eigenschaft $\operatorname{Kern}(g) = \langle e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}} \rangle$ angeben.[13] Wir haben damit

$$N = \langle e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}} \rangle$$

[12] Diese Aussage ist eine Konsequenz der universellen Eigenschaft von Faktorräumen, die im Expertenblock am Ende von § 17 behandelt wurde. Zumindest die eine Richtung sehen wir leicht, denn der Kern von $g = \widehat{g} \circ \pi$ enthält notwendigerweise den Kern von π, also N.

[13] Dazu können wir beispielsweise g durch die Bilder auf der Basis $(e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}}, e_{\text{rot}}, e_{\text{grün}})$ von $(K^X)_{00}$ festlegen, etwa durch $g(e_{\text{rot}} + e_{\text{grün}} - e_{\text{blau}}) = 0$ sowie $g(e_{\text{rot}}) = g(e_{\text{grün}}) = w$ für einen beliebigen Vektor $w \in W \setminus \{0\}$.

bestimmt. Es gilt nun: Für jedes $f \in S$ existiert ein eindeutig bestimmtes $\widehat{g} \in \operatorname{Homo}((K^X)_{00} / N, W)$ mit der Eigenschaft $f = \widehat{g} \circ \pi \circ e_\bullet$, wobei $\pi\colon (K^X)_{00} \to (K^X)_{00} / N$ die kanonische Surjektion ist.

Die zusätzliche Eigenschaft von $f \in S$ bewirkt also, dass f nicht nur durch den 3-dimensionalen Vektorraum $(K^X)_{00}$, sondern sogar durch den 2-dimensionalen Faktorraum $(K^X)_{00} / N$ faktorisiert werden kann. Das ist auch erforderlich dafür, dass die Zuordnung $f \mapsto \widehat{g}$ von S in $\operatorname{Homo}((K^X)_{00} / N, W)$ wieder bijektiv wird. △

Konstruktion eines Tensorprodukts

Wir können nun eine konkrete Realisierung des abstrakten Konzepts Tensorprodukt aus Definition 23.6 konstruieren. Dazu beginnen wir mit dem freien K-Vektorraum[14] über der Menge $U \times V$, also

$$(K^{U\times V})_{00} = \{f\colon U \times V \to K \mid \operatorname{supp} f \text{ ist endlich}\}.$$

Mit Hilfe von Satz 23.24 können wir jetzt beliebige Abbildungen $U \times V \to W$ in irgendeinen K-Vektorraum W mit linearen Abbildungen $(K^{U\times V})_{00} \to W$ identifizieren. Wir sind aber nur an bilinearen Abbildungen $U \times V \to W$ interessiert, also am Unterraum $S \coloneqq \operatorname{Bil}(U, V; W)$ von $W^{U\times V}$. Dieser Unterraum kann für beliebiges W wie folgt durch Bedingungen beschrieben werden, vgl. (23.2):

$$S \coloneqq \left\{ b\colon U \times V \to W \;\middle|\; \begin{aligned} & b(u_1+u_2, v) - b(u_1, v) - b(u_2, v) = 0 && \text{für alle } u_1, u_2 \in U,\ v \in V \\ & b(u, v_1+v_2) - b(u, v_1) - b(u, v_2) = 0 && \text{für alle } u \in U,\ v_1, v_2 \in V \\ & b(\alpha\, u, v) - \alpha\, b(u, v) = 0 && \text{für alle } u \in U,\ v \in V,\ \alpha \in K \\ & b(u, \alpha\, v) - \alpha\, b(u, v) = 0 && \text{für alle } u \in U,\ v \in V,\ \alpha \in K \end{aligned} \right\}. \tag{23.9}$$

Analog zu Beispiel 23.26 erwarten wir (und werden das in Satz 23.8 auch bestätigen), dass der (maximale) auszufaktorisierende Unterraum $N = \langle E \rangle$ als die lineare Hülle der Funktionen

$$\begin{aligned} E \coloneqq & \left\{ e_{(u_1+u_2, v)} - e_{(u_1, v)} - e_{(u_2, v)} \;\middle|\; u_1, u_2 \in U,\ v \in V \right\} \\ \cup & \left\{ e_{(u, v_1+v_2)} - e_{(u, v_1)} - e_{(u, v_2)} \;\middle|\; u \in U,\ v_1, v_2 \in V \right\} \\ \cup & \left\{ e_{(\alpha\, u, v)} - \alpha\, e_{(u, v)} \;\middle|\; u \in U,\ v \in V,\ \alpha \in K \right\} \\ \cup & \left\{ e_{(u, \beta\, v)} - \beta\, e_{(u, v)} \;\middle|\; u \in U,\ v \in V,\ \beta \in K \right\} \subseteq (K^{U\times V})_{00} \end{aligned} \tag{23.10}$$

geschrieben werden kann.

Wir definieren jetzt durch

$$U \otimes_{/} V \coloneqq (K^{U\times V})_{00} / \langle E \rangle \tag{23.11a}$$

$$\text{und} \quad u \otimes_{/} v \coloneqq [e_{(u,v)}] = e_{(u,v)} + \langle E \rangle \quad \text{für alle } u \in U, v \in V \tag{23.11b}$$

unseren Kandidaten für das Tensorprodukt der K-Vektorräume U und V. Um zu bestätigen, dass $(U \otimes_{/} V, \otimes_{/})$ tatsächlich die gewünschten Eigenschaften eines Tensorprodukts (Definition 23.6) erfüllt, benötigen wir zunächst die Bilinearität von $\otimes_{/}$.

[14] Der freie Vektorraum $(K^{U\times V})_{00}$ hat eine gewaltige Dimension, nämlich $\dim((K^{U\times V})_{00}) = \#(U \times V)$. Ist beispielsweise $K = \mathbb{Z}_2$ und gilt $\dim(U) = 2$ und $\dim(V) = 3$, so gilt $\dim (K^{U\times V})_{00} = \#(U \times V) = 4 \cdot 8 = 32$. Der Raum $(K^{U\times V})_{00}$ hat also 2^{32} (über 4 Milliarden) Elemente, während U nur $2^2 = 4$ und V nur $2^3 = 8$ Elemente hat.

Lemma 23.27 ($\otimes_/$ ist bilinear)**.**
Es seien U und V Vektorräume über demselben Körper. Dann gilt

$$u_1 \otimes_/ v \;+\; u_2 \otimes_/ v = (u_1 + u_2) \otimes_/ v \tag{23.12a}$$
$$u \otimes_/ v_1 \;+\; u \otimes_/ v_2 = u \otimes_/ (v_1 + v_2) \tag{23.12b}$$
$$(\alpha\, u) \otimes_/ v = \alpha\, (u \otimes_/ v) \tag{23.12c}$$
$$u \otimes_/ (\beta\, v) = \beta\, (u \otimes_/ v) \tag{23.12d}$$

für alle $u, u_1, u_2 \in U$ und $v, v_1, v_2 \in V$ sowie $\alpha, \beta \in K$.

Beweis. Wir erinnern daran, dass $u \otimes_/ v$ nicht anderes ist als die Nebenklasse von $e_{(u,v)}$ bzgl. $\langle E \rangle$, also $u \otimes_/ v = [e_{(u,v)}] = e_{(u,v)} + \langle E \rangle$ mit E aus (23.10).

Wir zeigen zunächst (23.12a). Die Funktion $e_{(u_1,v)}$ ist ein Repräsentant von $u_1 \otimes_/ v$, und die Funktion $e_{(u_2,v)}$ ist ein Repräsentant von $u_2 \otimes_/ v$. Nach Definition der Addition in Faktorräumen (Satz 17.17) ist also $e_{(u_1,v)} + e_{(u_2,v)}$ ein Repräsentant von $u_1 \otimes_/ v + u_2 \otimes_/ v$. Da $e_{(u_1+u_2,v)} - e_{(u_1,v)} - e_{(u_2,v)} \in E$ gilt, ist

$$e_{(u_1,v)} + e_{(u_2,v)} + e_{(u_1+u_2,v)} - e_{(u_1,v)} - e_{(u_2,v)} = e_{(u_1+u_2,v)}$$

ein weiterer Repräsentant der Nebenklasse $u_1 \otimes_/ v + u_2 \otimes_/ v$, aber offenbar gleichzeitig auch ein Repräsentant der Nebenklasse $(u_1 + u_2) \otimes_/ v$. Damit sind beide Nebenklassen identisch.

Der Beweis von (23.12b) läuft analog.

Wir zeigen noch (23.12c). Die Funktion $e_{(u,v)}$ ist ein Repräsentant von $u \otimes_/ v$. Nach Definition der S-Multiplikation in Faktorräumen (Satz 17.17) ist also $\alpha\, e_{(u,v)}$ ein Repräsentant von $\alpha\, (u \otimes_/ v)$. Da $e_{(\alpha\, u,v)} - \alpha\, e_{(u,v)} \in E$ gilt, ist

$$\alpha\, e_{(u,v)} + e_{(\alpha\, u,v)} - \alpha\, e_{(u,v)} = e_{(\alpha\, u,v)}$$

ein weiterer Repräsentant der Nebenklasse $\alpha\, (u \otimes_/ v)$, aber offenbar gleichzeitig auch ein Repräsentant der Nebenklasse $(\alpha\, u) \otimes_/ v$. Damit sind beide Nebenklassen identisch.

Der Beweis von (23.12d) läuft wieder analog. □

Satz 23.28 ($(U \otimes_/ V, \otimes_/)$ ist ein Tensorprodukt)**.**
Es seien U und V Vektorräume über demselben Körper K. Dann ist (23.11) ein Tensorprodukt von U und V.

Beweis. Wir müssen nur noch die universelle Eigenschaft von $(U \otimes_/ V, \otimes_/)$ bestätigen (vgl. Definition 23.6). Dazu sei W ein beliebiger K-Vektorraum und $b\colon U \times V \to W$ eine bilineare Abbildung. Wir müssen also zeigen, dass es eine eindeutig bestimmte lineare Abbildung $f \in \operatorname{Homo}(U \otimes_/ V, W)$ mit der Eigenschaft $b(u, v) = f(u \otimes_/ v)$ gibt.

Schritt 1: Wir definieren zunächst eine lineare Abbildung $\varphi\colon (K^{U \times V})_{00} \to W$ durch Angabe der Bilder auf der Basis $(e_{(u,v)})_{(u,v) \in U \times V}$:

$$\varphi(e_{(u,v)}) \coloneqq b(u, v) \quad \text{für alle } (u, v) \in U \times V.$$

Schritt 2: Wir zeigen, dass φ auf $\langle E\rangle$ verschwindet.

Beispielsweise gilt für alle $u_1, u_2 \in U$ und $v \in V$

$$\begin{aligned}
&\varphi(e_{(u_1+u_2,v)} - e_{(u_1,v)} - e_{(u_2,v)}) \\
&\quad = \varphi(e_{(u_1+u_2,v)}) - \varphi(e_{(u_1,v)}) - \varphi(e_{(u_2,v)}) && \text{wegen der Linearität von } \varphi \\
&\quad = b(u_1+u_2,v) - b(u_1,v) - b(u_2,v) && \text{nach Definition von } \varphi \\
&\quad = 0 && \text{wegen der Bilinearität von } b.
\end{aligned}$$

Analog können wir auch $\varphi(e_{(u,v_1+v_2)} - e_{(u,v_1)} - e_{(u,v_2)}) = 0$ zeigen sowie $\varphi(e_{(\alpha\, u,v)} - \alpha\, e_{(u,v)}) = 0$ und $\varphi(e_{(u,\beta\, v)} - \beta\, e_{(u,v)}) = 0$.

Damit verschwindet φ auf allen Elementen der erzeugenden Menge E und damit auf der gesamten linearen Hülle $\langle E\rangle$.

Schritt 3: Wir konstruieren eine lineare Abbildung $f \in \operatorname{Homo}(U \otimes_/ V, W)$ mit der Eigenschaft $b = f \circ \otimes_/$.

Es sei $t \in U \otimes_/ V = (K^{U\times V})_{00} \,/\, \langle E\rangle$ beliebig und $s \in (K^{U\times V})_{00}$ irgendein Repräsentant der Nebenklasse von T, also $T = [s] = s + \langle E\rangle$. Dann gilt

$$s = \sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, e_{(u,v)}.$$

Wir setzen

$$f(T) \coloneqq \varphi(s) = \sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, \varphi(e_{(u,v)}) = \sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, b(u,v).$$

Diese Vorschrift ist wohldefiniert, denn ist $\widetilde{s} \in (K^{U\times V})_{00}$ ein weiterer Repräsentant von T, dann gilt $\widetilde{s} - s \in \langle E\rangle$. Wegen Schritt 2 ist aber $\varphi(\widetilde{s} - s) = 0$ und damit $\varphi(\widetilde{s}) = \varphi(s)$.

Insbesondere gilt

$$f(u \otimes_/ v) = f([e_{(u,v)}]) = \varphi(e_{(u,v)}) = b(u,v),$$

also $b = f \circ \otimes_/$.

Schritt 4: Wir zeigen, dass f eindeutig bestimmt ist.

Nehmen wir an, $f_1, f_2 \in \operatorname{Homo}(U \otimes_/ V, W)$ seien zwei Abbildungen mit der Eigenschaft $b(u,v) = f_1(u \otimes_/ v) = f_2(u \otimes_/ v)$ für alle $u \in U$ und $v \in V$. Ist $t \in U \otimes_/ V$ beliebig, so gilt $T = [s]$ für ein $s \in (K^{U\times V})_{00}$, und damit wie oben

$$\begin{aligned}
f_1(T) &= f_1([s]) \\
&= f_1\Big(\Big[\sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, e_{(u,v)}\Big]\Big) \\
&= \sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, f_1([e_{(u,v)}]) && \text{wegen der Linearität von } f_1 \\
&= \sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, f_1(u \otimes_/ v) && \text{wegen } [e_{(u,v)}] = u \otimes_/ v
\end{aligned}$$

$$= \sum_{(u,v)\in\operatorname{supp}(s)} s(u,v)\, b(u,v) \qquad \text{nach Voraussetzung.}$$

Dasselbe gilt aber für f_2, und damit gilt $f_1(T) = f_2(T)$ für alle $t \in U \otimes_/ V$. □

Es gibt noch andere gängige Konstruktionen eines Tensorprodukts. Zwei davon werden im folgenden Beispiel genannt.

Beispiel 23.29 (alternative Konstruktionen eines Tensorprodukts).

(*i*) Sind U und V Vektorräume über demselben Körper K mit Basen $B_U = (u_i)_{i\in I}$ bzw. $B_V = (v_j)_{j\in J}$, so kann ein Tensorproduktraum durch

$$U \otimes_B V \coloneqq (K^{I\times J})_{00} = \{f : I \times J \to K \mid \operatorname{supp}(f) \text{ endlich}\}$$

definiert werden. Die zugehörige universelle bilineare Abbildung wird durch die Werte

$$u_i \otimes_B v_j \coloneqq e_{(i,j)} \quad \text{für alle } i \in I \text{ und } j \in J$$

eindeutig festgelegt (Satz 23.4).

(*ii*) Sind U und V endlich-dimensionale Vektorräume über demselben Körper K, so kann ein Tensorproduktraum durch

$$U \otimes_{\text{dual}} V \coloneqq \operatorname{Bil}(U^*, V^*; K)$$

definiert werden. Die zugehörige universelle bilineare Abbildung wird als

$$u \otimes_{\text{dual}} v : U^* \times V^* \ni (u^*, v^*) \mapsto (u \otimes_{\text{dual}} v)(u^*, v^*) \coloneqq \langle u^*, u\rangle \, \langle v^*, v\rangle \in K$$

definiert. △

Ende der Vorlesung 6

§ 23.5 Das Tensorprodukt linearer Abbildungen

Es wird im Folgenden nützlich sein, zwei lineare Abbildungen zu einer linearen Abbildungen zwischen den betreffenden Tensorprodukträumen zu kombinieren.

Definition 23.30 (Tensorprodukt linearer Abbildungen).
Es seien U_1, U_2, V_1 und V_2 Vektorräume über demselben Körper und $(U_1 \otimes U_2, \otimes)$ ein Tensorprodukt von U_1 und U_2 sowie $(V_1 \otimes V_2, \otimes)$ ein Tensorprodukt von V_1 und V_2. Weiter seien $f_1 \in \operatorname{Homo}(U_1, V_1)$ und $f_2 \in \operatorname{Homo}(U_2, V_2)$. Dann heißt die für $u_1 \in U_1$ und $u_2 \in U_2$ durch

$$f_1 \boxtimes f_2 : \begin{cases} U_1 \otimes U_2 \to V_1 \otimes V_2 \\ u_1 \otimes u_2 \mapsto f_1(u_1) \otimes f_2(u_2) \end{cases} \tag{23.13}$$

eindeutig bestimmte lineare Abbildung das **Tensorprodukt** der Abbildungen f_1 und f_2. △

Beachte: Wir machen in (23.13) Gebrauch von Bemerkung 23.9 und nutzen, dass eine lineare Abbildung auf einem Tensorproduktraum durch ihre Bilder auf Elementartensoren eindeutig bestimmt ist, sofern diese durch eine bilineare Abbildungsvorschrift gegeben sind. Das ist aber der Fall; beispielsweise gilt

$$\begin{aligned}
(u_1 + u_1') \otimes u_2 \mapsto f_1(u_1 + u_1') \otimes f_2(u_2) \quad & \text{nach Definition der Abbildungsvorschrift} \\
= (f_1(u_1) + f_1(u_1')) \otimes f_2(u_2) \quad & \text{wegen der Linearität von } f_1 \\
= f_1(u_1) \otimes f_2(u_2) + f_1(u_1') \otimes f_2(u_2) \quad & \text{wegen der Bilinearität von } \otimes.
\end{aligned}$$

Analog zeigt man auch $(\alpha\, u_1) \otimes u_2 \mapsto \alpha\,(f_1(u_1) \otimes f_2(u_2))$ sowie die Linearität im zweiten Argument.

Beachte: Viele Autoren bezeichnen die Abbildung $f_1 \boxtimes f_2$ auch mit $f_1 \otimes f_2$. Das wäre aber bereits die dritte Bedeutung für das Symbol $\otimes$.

Beispiel 23.31 (Tensorprodukt linearer Abbildungen)**.**

(i) Sind U und V Vektorräume über demselben Körper, dann gilt $\mathrm{id}_U \boxtimes \mathrm{id}_V = \mathrm{id}_{U \otimes V}$.

(ii) Es seien $U_1, U_2, V_1, V_2, W_1, W_2$ Vektorräume über demselben Körper. Weiter seien $f_1 \in \operatorname{Homo}(V_1, W_1)$, $f_2 \in \operatorname{Homo}(U_1, V_1)$, $g_1 \in \operatorname{Homo}(V_2, W_2)$ und $g_2 \in \operatorname{Homo}(U_2, V_2)$. Dann gilt

$$(f_1 \circ f_2) \boxtimes (g_1 \circ g_2) = (f_1 \boxtimes g_1) \circ (f_2 \boxtimes g_2). \qquad \triangle$$

Wir wollen nun untersuchen, wie sich die Darstellungsmatrix des Tensorprodukts $f_1 \boxtimes f_2$ zweier linearer Abbildungen zusammensetzt aus den Darstellungsmatrizen der beiden einzelnen Abbildungen.

Lemma 23.32 (Darstellungsmatrix des Tensorprodukts linearer Abbildungen)**.**
Es seien U_1, U_2, V_1 und V_2 endlich-dimensionale Vektorräume über demselben Körper K mit Basen $B_{U_1} = (u_1^{(1)}, \ldots, u_{m_1}^{(1)})$, $B_{U_2} = (u_1^{(2)}, \ldots, u_{m_2}^{(2)})$, $B_{V_1} = (v_1^{(1)}, \ldots, v_{n_1}^{(1)})$ und $B_{V_2} = (v_1^{(2)}, \ldots, v_{n_2}^{(2)})$. Weiter seien

$$\begin{aligned}
f_1 &\in \operatorname{Homo}(U_1, V_1) \quad \text{mit Darstellungsmatrix} \quad A = \mathcal{M}_{B_{V_1} \leftarrow B_{U_1}}(f_1) \\
f_2 &\in \operatorname{Homo}(U_2, V_2) \quad \text{mit Darstellungsmatrix} \quad B = \mathcal{M}_{B_{V_2} \leftarrow B_{U_2}}(f_2).
\end{aligned}$$

Ordnen wir die Basis $B_{U_1 \otimes U_2} = (u_i^{(1)} \otimes u_j^{(2)})_{(i,j) \in [\![1,m_1]\!] \times [\![1,m_2]\!]}$ von $U_1 \otimes U_2$ in der lexikographischen Reihenfolge, also

$$\begin{aligned}
&B_{U_1 \otimes U_2} \\
&= \big(u_1^{(1)} \otimes u_1^{(2)}, \ldots, u_1^{(1)} \otimes u_{m_2}^{(2)}, \quad u_2^{(1)} \otimes u_1^{(2)}, \ldots, u_2^{(1)} \otimes u_{m_2}^{(2)}, \quad \ldots, \quad u_{m_1}^{(1)} \otimes u_1^{(2)}, \ldots, u_{m_1}^{(1)} \otimes u_{m_2}^{(2)}\big),
\end{aligned}$$

und verfahren mit der Basis $B_{V_1 \otimes V_2}$ analog, dann gilt für die Darstellungsmatrix von $f_1 \boxtimes f_2$:

$$\mathcal{M}_{B_{V_1 \otimes V_2} \leftarrow B_{U_1 \otimes U_2}}(f_1 \boxtimes f_2) = A \boxtimes B \in K^{(n_1 n_2) \times (m_1 m_2)}. \tag{23.14}$$

Dabei heißt die Matrix

$$A \boxtimes B := \begin{bmatrix} a_{11}\, B & a_{12}\, B & \cdots & a_{1m_1}\, B \\ a_{21}\, B & a_{22}\, B & \cdots & a_{2m_1}\, B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_1 1}\, B & a_{n_1 2}\, B & \cdots & a_{n_1 m_1}\, B \end{bmatrix} \tag{23.15}$$

das **Kroneckerprodukt** (englisch: Kronecker product) der Matrizen A und B.[15]

Beweis. Da A und B die Darstellungsmatrizen von f_1 und f_2 bzgl. Basen B_{U_1} und B_{V_1} bzw. B_{U_2} und B_{V_2} sind, können wir die Bilder der Basisvektoren und f_1 bzw. f_2 nach Satz 19.3 schreiben als

$$f_1(u_k^{(1)}) = \sum_{i=1}^{n_1} a_{ik}\, v_i^{(1)} \quad \text{und} \quad f_2(u_\ell^{(2)}) = \sum_{j=1}^{n_2} b_{j\ell}\, v_j^{(2)}$$

für alle $k \in [\![1, m_1]\!]$ und $\ell \in [\![1, m_2]\!]$. Nach Definition des Tensorprodukts der linearen Abbildungen f_1 und f_2 gilt damit

$$\begin{aligned} &(f_1 \boxtimes f_2)(u_k^{(1)} \otimes u_\ell^{(2)}) \\ &\quad = f_1(u_k^{(1)}) \otimes f_2(u_\ell^{(2)}) && \text{nach Definition von } f_1 \boxtimes f_2 \\ &\quad = \Big(\sum_{i=1}^{n_1} a_{ik}\, v_i^{(1)}\Big) \otimes \Big(\sum_{j=1}^{n_2} b_{j\ell}\, v_j^{(2)}\Big) \\ &\quad = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} a_{ik}\, b_{j\ell}\, (v_i^{(1)} \otimes v_j^{(2)}) && \text{wegen der Bilinearität des Tensorprodukts.} \end{aligned}$$

Die Spalte in der Darstellungsmatrix von $f_1 \boxtimes f_2$, die zum Indexpaar (k, ℓ) bzw. zum linearen Index $m_2(k-1)+\ell$ gehört, besteht also aus den Einträgen $a_{ik}\, b_{j\ell}$ für alle $i \in [\![1, n_1]\!]$ und $j \in [\![1, n_2]\!]$. Dabei wird der Eintrag zum Indexpaar (i, j) in der Zeile $n_2(i-1)+j$ stehen.

Beispielsweise enthält die erste Spalte ($k = 1$ und $\ell = 1$) der Darstellungsmatrix zuoberst die Einträge $a_{11}\, b_{\bullet 1}$ (zu $i = 1$ und $j = 1, \ldots, n_2$), darunter ($i = 2$ und $j = 1, \ldots, n_2$) die Einträge $a_{21}\, b_{\bullet 1}$ usw. bis zu den letzten Einträgen ($i = n_1$ und $j = 1, \ldots, n_2$) $a_{n_1 1}\, b_{\bullet 1}$. In der zweiten Spalte wiederholt sich dieses Muster, aber mit $k = 1$ und $\ell = 2$ usw.:

$$\begin{bmatrix} a_{11}\, b_{\bullet 1} & a_{11}\, b_{\bullet 2} & \cdots \\ a_{21}\, b_{\bullet 1} & a_{21}\, b_{\bullet 2} & \cdots \\ \vdots & \vdots & \\ a_{n_1 1}\, b_{\bullet 1} & a_{n_1 1}\, b_{\bullet 2} & \cdots \end{bmatrix}.$$

Das sind aber genau die Einträge von $A \boxtimes B$. □

Wir geben abschließend noch einige Eigenschaften des Kronecker-Produkts von Matrizen an, die jeweils Entsprechungen in Eigenschaften des Tensorprodukts linearer Abbildungen haben. (**Quizfrage 23.7:** Wie sehen die Zusammenhänge aus?)

[15] Das Kronecker-Produkt von Matrizen wird meistens als $A \otimes B$ geschrieben.

Lemma 23.33 (Eigenschaften des Kronecker-Produkts von Matrizen)**.**
Es sei K ein Körper. Für das Kronecker-Produkt gelten folgende Eigenschaften jeweils für Matrizen passender Dimension über K:

(i) Das Kronecker-Produkt ist bilinear:

$$\begin{aligned}(A_1 + A_2) \boxtimes B &= A_1 \boxtimes B + A_2 \boxtimes B \\ A \boxtimes (B_1 + B_2) &= A \boxtimes B_1 + A \boxtimes B_2 \\ \alpha\,(A \boxtimes B) &= (\alpha\,A) \boxtimes B = A \boxtimes (\alpha\,B).\end{aligned}$$

(ii) Das Kronecker-Produkt ist assoziativ:

$$(A \boxtimes B) \boxtimes C = A \boxtimes (B \boxtimes C).$$

(iii) Das Kronecker-Produkt ist verträglich mit der Matrixmultiplikation:

$$(A_1\,A_2) \boxtimes (B_1\,B_2) = (A_1 \boxtimes B_1)\,(A_2 \boxtimes B_2).$$

(iv) Das Kronecker-Produkt ist verträglich mit der Transposition:

$$(A \boxtimes B)^\intercal = A^\intercal \boxtimes B^\intercal.$$

(v) Das Kronecker-Produkt ist verträglich mit der Invertierung: Sind A und B invertierbar, dann auch $A \boxtimes B$, und es gilt

$$(A \boxtimes B)^{-1} = A^{-1} \boxtimes B^{-1}.$$

(vi) Das Kronecker-Produkt ist verträglich mit dem Rang:

$$\operatorname{Rang}(A \boxtimes B) = \operatorname{Rang}(A)\ \operatorname{Rang}(B).$$

Beweis. □

§ 23.6 Darstellung von Tensoren

Unser Ziel ist es nun, Elemente eines Tensorproduktraumes $U \otimes V$ für K-Vektorräume U und V in eine Standarddarstellung zu bringen, die unabhängig von den konkreten Vektorräumen U und V ist. Ab sofort sind bis zum Ende von § 23 alle Vektorräume endlich-dimensional.

Ist $\dim(U) = n$ und $\dim(V) = m$ mit $n, m \in \mathbb{N}_0$, dann gilt nach Lemma 23.13 $\dim(U \otimes V) = m\,n$. Es bietet sich also an, die Elemente von $U \otimes V$ in Form von Matrizen in $K^{n \times m}$ darzustellen. Ähnlich wie bei der Koordinatendarstellung von Vektoren in einem Vektorraum bzgl. einer gewählten Basis (§ 19.1) wird auch diese Darstellung abhängig sein von der Wahl von Basen in U und V.

Satz 23.34 (Komponentenmatrix, Syntheseabbildung, Analyseabbildung, vgl. Satz 19.1)**.**
Es seien U und V Vektorräume über demselben Körper K mit $\dim(U) = n \in \mathbb{N}_0$ und $\dim(V) = m \in \mathbb{N}_0$. Weiter seien die Familien $B_U = (u_1, \dots, u_n)$ bzw. $B_V = (v_1, \dots, v_m)$ Basen von U bzw. V und $B_{U\otimes V} \coloneqq (u_i \otimes v_j)_{(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ die zugehörige Tensorproduktbasis von $U \otimes V$ (Lemma 23.13). Dann gilt:

(i) Die Abbildung

$$\Phi_{B_{U\otimes V}} : K^{n\times m} \ni A \mapsto t \coloneqq \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\,(u_i \otimes v_j) \in U \otimes V \tag{23.16}$$

ist ein linearer Isomorphismus $K^{n\times m} \to U \otimes V$. Diese erzeugt aus einer **Komponentenmatrix** (englisch: component matrix) $A \in K^{n\times m}$ die zugehörige Linearkombination der Basistensoren $u_i \otimes v_j$. Wir bezeichnen die Abbildung $\Phi_{B_{U\otimes V}}$ daher auch als die **Syntheseabbildung** (englisch: synthesis map) bzgl. der Basis $B_{U\otimes V}$.

(ii) Der zu $\Phi_{B_{U\otimes V}}$ inverse Isomorphismus ist die Abbildung

$$\Phi^{-1}_{B_{U\otimes V}} : U \otimes V \ni t \mapsto A \in K^{n\times m}, \tag{23.17}$$

die jedem Tensor $t \in U \otimes V$ seine eindeutige **Komponentenmatrix** $A \in K^{n\times m}$ bzgl. der Basis $B_{U\otimes V}$ zuordnet. Wir nennen daher $\Phi^{-1}_{B_{U\otimes V}}$ auch die **Komponentenabbildung** (englisch: component map) oder die **Analyseabbildung** (englisch: analysis map) bzgl. der Basis $B_{U\otimes V}$. Der Eintrag $a_{ij} \in K$ heißt die (i,j)-te **Komponente** (englisch: component) des Tensors $t \in U \otimes V$ bzgl. der Basis $B_{U\otimes V}$.

Beweis. Die Abbildung $\Phi_{B_{U\otimes V}}$ ist linear und bildet die Standardbasis $(E_{ij})_{(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ von $K^{n\times m}$ (Satz 15.3) bijektiv auf die Basis $B_{U\otimes V} = (u_i \otimes v_j)_{(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ von $U \otimes V$ ab. Nach Satz 17.10 ist $\Phi_{B_{U\otimes V}}$ damit ein Isomorphismus. □

Beispiel 23.35 (Komponentenmatrix)**.**
Wir betrachten ein Tensorprodukt von K^n und K^m über einem Körper K. Der Satz 23.34 etabliert einen (basisabhängigen) Isomorphismus zwischen $K^n \otimes K^m$ und $K^{n\times m}$.

Wir wählen hier in beiden Räumen die Standardbasis $(e_1, \dots, e_n)$ bzw. $(e_1, \dots, e_m)$. Dann ist die Komponentenmatrix des Basistensors $e_i \otimes e_j$ gerade

$$\Phi^{-1}_{B_{K^n\otimes K^m}}(e_i \otimes e_j) = e_i\,e_j^\intercal = \begin{array}{cl} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{array}{l} \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \end{array} \\ \qquad\quad \uparrow j\text{-te Spalte} & \end{array}$$

Allgemeiner hat der Elementartensor $x \otimes y$ die Komponentenmatrix $x y^\intercal$, denn es gilt

$$\Phi^{-1}_{B_{K^n\otimes K^m}}(x \otimes y) = \Phi^{-1}_{B_{K^n\otimes K^m}}\Big(\sum_{i=1}^{n}\sum_{j=1}^{m} x_i\,y_j\,(e_i \otimes e_j)\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i\,y_j\,\Phi^{-1}_{B_{K^n\otimes K^m}}(e_i \otimes e_j) = x y^\intercal.$$

Ist
$$A = B\,C = \sum_{j=1}^{r} b_{\bullet j}\, c_{j\bullet}$$
irgendeine Faktorisierung von $A \in K^{n\times m}$ mit $B \in K^{n\times r}$ und $C \in K^{r\times m}$, so ist A die Komponentenmatrix des Tensors
$$\Phi_{B_{K^n\otimes K^m}}(A) = \sum_{k=1}^{r} b_{\bullet k} \otimes c_{k\bullet}.$$
Insbesondere ergibt sich (durch die Wahl von $B = I_n$ und $C = A$ bzw. durch die Wahl von $B = A$ und $C = I_m$), dass A die Komponentenmatrix von
$$\Phi_{B_{K^n\otimes K^m}}(A) = \sum_{j=1}^{n} a_{\bullet j} \otimes e_j = \sum_{i=1}^{m} e_i \otimes a_{i\bullet}$$
ist. (**Quizfrage 23.8:** Wie kann das sein, dass A offenbar die Komponentenmatrix mehrerer Tensoren ist, wo doch $\Phi_{B_{K^n\otimes K^m}}$ als Isomorphismus bijektiv ist?) △

Bemerkung 23.36 (Tensoren und Komponentenmatrizen)**.**

Obwohl Satz 23.34 einen Isomorphismus zwischen Tensoren und ihren Komponentenmatrizen herstellt, sollten die beiden Konzepte nicht gleichgesetzt werden. Wir sollten also einen Tensor nicht mit seiner Komponentenmatrix (bzgl. einer bestimmten Basis) verwechseln, sondern die Komponentenmatrix als eine Darstellung des Tensors verstehen. △

Im Beispiel 23.35 konnten wir aufgrund der Wahl der Standardbasen leicht zwischen Tensoren in $K^n \otimes K^m$ und ihren Komponentenmatrizen in $K^{n\times m}$ hin- und herwechseln. Es stellt sich aber die Frage, wie wir auch in allgemeinen Vektorräumen U, V und bei beliebigen Basen B_U, B_V die Komponenten eines Tensors t bzgl. der Tensorproduktbasis $B_{U\otimes V}$ bestimmen können. Wie wir aus Bemerkung 20.12 wissen, können die Koeffizienten in einer Linearkombination von Basisvektoren in beliebigen endlich-dimensionalen Vektorräumen mit Hilfe der dualen Basis ermittelt werden. Hier geht es um die zu $B_{U\otimes V}$ duale Basis, also um eine Familie von Elementen des Dualraumes $(U \otimes V)^*$.

Das folgende Resultat untersucht die Gestalt des Dualraumes $(U \otimes V)^*$ und einer dualen Basis.

Lemma 23.37 (Dualraum des Tensorprodukts und duale Basis)**.**

Es seien U und V endlich-dimensionale Vektorräume über demselben Körper, $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V sowie $(U^* \otimes V^*, \otimes)$ ein Tensorprodukt der Dualräume U^* und V^*. Dann gilt:

(i) Die Abbildung
$$\begin{aligned} U^* \otimes V^* &\to (U \otimes V)^* \\ u^* \otimes v^* &\mapsto u^* \boxtimes v^* = \langle u^*, \cdot\rangle\, \langle v^*, \cdot\rangle \end{aligned} \tag{23.18a}$$

definiert einen kanonischen linearen Isomorphismus. Dabei ist $u^* \boxtimes v^* \in (U \otimes V)^*$ gegeben durch die Bilder auf den Elementartensoren, nämlich durch[16]

$$(u^* \boxtimes v^*)(u \otimes v) \coloneqq \langle u^*, u\rangle \, \langle v^*, v\rangle \tag{23.18b}$$

für alle $u \in U$ und $v \in V$.

(*ii*) Es seien $B_U = (u_1, \dots, u_n)$ bzw. $B_V = (v_1, \dots, v_m)$ mit $n, m \in \mathbb{N}_0$ Basen von U bzw. von V. Weiter seien $B_{U^*} = (u_1^*, \dots, u_n^*)$ bzw. $B_{V^*} = (v_1^*, \dots, v_m^*)$ die zugehörigen dualen Basen von U^* bzw. von V^*. Dann ist

$$(u_i^* \boxtimes v_j^*)_{(i,j) \in [\![1,n]\!] \times [\![1,m]\!]} \tag{23.19}$$

die zu $B_{U \otimes V} = (u_i \otimes v_j)_{(i,j) \in [\![1,n]\!] \times [\![1,m]\!]}$ duale Basis von $(U \otimes V)^*$.

Beweis. Wir beweisen die Aussagen (*i*) und (*ii*) gemeinsam.

Schritt 1: Wir halten zunächst fest, dass (23.18) überhaupt eindeutig eine lineare Abbildung $U^* \otimes V^* \longrightarrow (U \otimes V)^*$ definiert. Dabei machen wir Gebrauch von Bemerkung 23.9: Die für Elementartensoren $u \otimes v$ durch (23.18b) erklärte Abbildung $u^* \boxtimes v^*$ besitzt eine eindeutige Fortsetzung zu einer linearen Abbildung $U \otimes V \longrightarrow K$, da die Abbildungsvorschrift auf der rechten Seite von (23.18b) bilinear in (u, v), also linear in u und linear in v ist.

Weiter ist die Abbildung $U^* \otimes V^* \longrightarrow (U \otimes V)^*$ durch (23.18a), also durch ihre Bilder auf den Elementartensoren eindeutig bestimmt, da die Abbildungsvorschrift wiederum bilinear in (u^*, v^*) ist.

Schritt 2: Um zu zeigen, dass (23.18) tatsächlich ein Isomorphismus ist, betrachten wir Basen $B_U = (u_1, \dots, u_n)$ und $B_V = (v_1, \dots, v_m)$ von U und V sowie deren duale Basen $B_{U^*} = (u_1^*, \dots, u_n^*)$ und $B_{V^*} = (v_1^*, \dots, v_m^*)$. Desweiteren benötigen wir

die induzierte Basis	$B_{U \otimes V} = (u_k \otimes v_\ell)_{(k,\ell) \in [\![1,n]\!] \times [\![1,m]\!]}$
die induzierte Basis	$B_{U^* \otimes V^*} = (u_i^* \otimes v_j^*)_{(i,j) \in [\![1,n]\!] \times [\![1,m]\!]}$
sowie die zu $B_{U \otimes V}$ duale Basis	$B_{(U \otimes V)^*} = ((u_i \otimes v_j)^*)_{(i,j) \in [\![1,n]\!] \times [\![1,m]\!]}$.

Wir zeigen jetzt, dass die Abbildung (23.18) die Basis $B_{U^* \otimes V^*}$ von $U^* \otimes V^*$ bijektiv auf die Basis $B_{(U \otimes V)^*}$ von $(U \otimes V)^*$ abbildet. Aus Satz 17.10 folgt dann, dass (23.18) ein Isomorphismus ist.

In der Tat gilt

$$\begin{aligned}
(u_i^* \boxtimes v_j^*)(u_k \otimes v_\ell) &= \langle u_i^*, u_k\rangle \, \langle v_j^*, v_\ell\rangle && \text{wegen der Abbildungsvorschrift (23.18)} \\
&= \delta_{ik} \, \delta_{j\ell} && \text{wegen der Dualität der Basen} \\
&= (u_i \otimes v_j)^*(u_k \otimes v_\ell) && \text{wegen der Dualität der Basen.}
\end{aligned}$$

Das bedeutet aber gerade, dass wie behauptet $u_i^* \boxtimes v_j^* = (u_i \otimes v_j)^*$ für alle i und j gilt. Außerdem zeigt die Rechnung, dass $(u_i^* \boxtimes v_j^*)_{(i,j) \in [\![1,n]\!] \times [\![1,m]\!]}$ die zu $B_{U \otimes V}$ duale Basis ist. □

[16]Nach Definition 23.30 ist das Tensorprodukt $u^* \boxtimes v^*$ der linearen Abbildungen $u^* \in U^* = \operatorname{Homo}(U, K)$ und $v^* \in V^* = \operatorname{Homo}(V, K)$ eigentlich eine lineare Abbildung $U \otimes V \longrightarrow K \otimes K$. Wir können aber $K \otimes K \cong K$ identifizieren (Beispiel 23.20). Dabei wird das Tensorprodukt $\alpha \otimes \beta$ zur Multiplikation $\alpha \, \beta$ in K.

Nachdem wir jetzt die zu einer Basis $B_{U\otimes V}$ duale Basis — also die „Komponentenermittler" — kennen (vgl. Bemerkung 20.12), können wir die Einträge der Komponentenmatrix A eines Tensors T angeben. Mit den Bezeichnungen aus Satz 23.34 gilt

$$a_{ij} = (u_i^* \boxtimes v_j^*)(t). \tag{23.20}$$

Die Komponentendarstellung eines Tensors ermöglicht es uns übrigens, insbesondere den Rang des Tensors zu bestimmen:

Satz 23.38 (Rang eines Tensors ist Rang seiner Komponentenmatrizen[17]).
Es seien U und V Vektorräume über demselben Körper K mit $\dim(U) = n \in \mathbb{N}_0$ und $\dim(V) = m \in \mathbb{N}_0$ sowie $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Weiter seien die Familien B_U und B_V Basen von U bzw. V und $B_{U\otimes V}$ die induzierte Basis von $U \otimes V$. Schließlich sei $t \in U \otimes V$ ein Tensor und $A \in K^{n\times m}$ seine Komponentenmatrix bzgl. der Basis $B_{U\otimes V}$, also $A = \Phi_{B_{U\otimes V}}^{-1}(t)$. Dann gilt:

$$\operatorname{Rang}(t) = \operatorname{Rang}(A).$$

Beweis. Wir bezeichnen die Basisfamilien von U bzw. V mit $B_U = (u_1, \dots, u_n)$ bzw. $B_V = (v_1, \dots, v_m)$. Wir führen den Beweis in zwei Schritten:

Schritt 1: Wir zeigen $\operatorname{Rang}(t) \leqslant \operatorname{Rang}(A)$.

Es sei $r = \operatorname{Rang}(A)$ und $A = BC$ eine Rangfaktorisierung (Folgerung 15.16) von A mit $B \in K^{n\times r}$ und $C \in K^{r\times m}$. Dann gilt $a_{ij} = \sum_{k=1}^r b_{ik}\, c_{kj}$ für alle $1 \leqslant i \leqslant n$ und $1 \leqslant j \leqslant m$ und daher

$$\begin{aligned}
T &= \sum_{i=1}^n \sum_{j=1}^m a_{ij}\,(u_i \otimes v_j) && \text{denn } A \text{ ist die Komponentenmatrix von } T \text{ bzgl. der Basis } B_{U\otimes V}\\
&= \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r b_{ik}\, c_{kj}\,(u_i \otimes v_j) && \text{wegen } A = BC\\
&= \sum_{k=1}^r \sum_{i=1}^n \sum_{j=1}^m b_{ik}\, c_{kj}\,(u_i \otimes v_j)\\
&= \sum_{k=1}^r \Big(\sum_{i=1}^n b_{ik} u_i\Big) \otimes \Big(\sum_{j=1}^m c_{kj} v_j\Big) && \text{wegen der Bilinearität von } \otimes\text{, vgl. (23.5).}
\end{aligned}$$

Das zeigt, dass T geschrieben werden kann als eine Summe von r Elementartensoren. Daher gilt $\operatorname{Rang}(t) \leqslant r$.

Schritt 2: Wir zeigen $\operatorname{Rang}(A) \leqslant \operatorname{Rang}(t)$.

Es sei $r = \operatorname{Rang}(t)$ und $t = \sum_{k=1}^r u_k' \otimes v_k'$ eine Darstellung von t als Summe von r Elementartensoren. Dann gilt

$$T = \sum_{k=1}^r u_k' \otimes v_k'$$

[17] Anders ausgedrückt besagt dieser Satz, dass der Rang eines Tensors unter einem Isomorphismus invariant ist. Das gilt auch allgemeiner für Tensorprodukte unendlich-dimensionaler Vektorräume.

$$
\begin{aligned}
&= \sum_{k=1}^{r} \Big(\sum_{i=1}^{n} b_{ik}\, u_i \Big) \otimes \Big(\sum_{j=1}^{m} c_{kj}\, v_j \Big) && \text{mit geeigneten Koeffizienten } b_{ik} \text{ und } c_{kj}, \\
& && \text{da } B_U \text{ und } B_V \text{ Basen von } U \text{ und } V \text{ sind} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{r} b_{ik}\, c_{kj}\, (u_i \otimes v_j) && \text{wegen der Bilinearität von } \otimes, \text{ vgl. (23.5)} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, (u_i \otimes v_j) && \text{mit } a_{ij} \coloneqq \sum_{k=1}^{r} b_{ik}\, c_{kj}.
\end{aligned}
$$

Das zeigt, dass A geschrieben werden kann als das Produkt der Matrizen $B \in K^{n \times r}$ und $C \in K^{r \times m}$. Nach Satz 15.17 über den Rang des Produkts von Matrizen gilt $\operatorname{Rang}(A) \leqslant r$. □

Mithilfe von Satz 23.38 können nun auch die noch offene Frage beantworten, wie groß der Rang eines Tensors sein kann:

Folgerung 23.39 (maximaler Rang eines Tensors).
Es seien U und V Vektorräume über demselben Körper K und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann gilt:

(*i*) Für alle $t \in U \otimes V$ gilt $\operatorname{Rang}(t) \in \mathbb{N}_0$.

(*ii*) Hat mindestens einer der Räume U und V endliche Dimension, so gilt für alle $t \in U \otimes V$

$$\operatorname{Rang}(t) \leqslant \min\{\dim(U), \dim(V)\}. \tag{23.21}$$

Für jedes $r \in \mathbb{N}_0$ mit $r \leqslant \min\{\dim(U), \dim(V)\}$ existiert ein Tensor $t \in U \otimes V$ mit $\operatorname{Rang}(t) = r$.

(*iii*) Haben U und V beide unendliche Dimension, dann existiert für jedes $r \in \mathbb{N}_0$ ein Tensor $t \in U \otimes V$ mit $\operatorname{Rang}(t) = r$.

Beweis. □

§ 23.7 Transformation von Komponentenmatrizen bei Basiswechsel

In diesem Abschnitt untersuchen wir, wie sich die Komponentenmatrix eines Tensors t ändert, wenn wir die Basen der endlich-dimensionalen Vektorräume U und V ändern. Wir benutzen dafür analoge Bezeichnungen wie in § 19.4 und § 20.2. Es seien also

$$
\begin{aligned}
B_U &= (u_1, \ldots, u_n) \quad \text{und} \quad \widehat{B}_U = (\widehat{u}_1, \ldots, \widehat{u}_n) \quad \text{die „alte“ bzw. „neue“ Basis von } U \\
B_V &= (v_1, \ldots, v_m) \quad \text{und} \quad \widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_m) \quad \text{die „alte“ bzw. „neue“ Basis von } V.
\end{aligned}
$$

Außerdem seien $S \coloneqq \mathcal{T}_{B_U \leftarrow \widehat{B}_U} \in K^{n \times n}$ und $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \in K^{m \times m}$ die Transformationsmatrizen des Basiswechsels von $\widehat{B}_U$ zu B_U bzw. von $\widehat{B}_V$ zu B_V. Es gilt also

$$\widehat{u}_j = \sum_{i=1}^{n} s_{ij}\, u_i \quad \text{bzw.} \quad u_i = \sum_{k=1}^{n} (S^{-1})_{ki}\, \widehat{u}_k$$

$$\text{und} \quad \widehat{v}_j = \sum_{i=1}^{m} t_{ij}\, v_i \quad \text{bzw.} \quad v_i = \sum_{\ell=1}^{m} (T^{-1})_{\ell i}\, \widehat{v}_\ell.$$

Setzen wir diese Beziehungen in die Syntheseabbildung ein, so erhalten wir für die Komponentenmatrix eines Tensors $t \in U \otimes V$ die Beziehung

$$\begin{aligned}
t &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, (u_i \otimes v_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \left(\sum_{k=1}^{n} (S^{-1})_{ki}\, \widehat{u}_k \right) \otimes \left(\sum_{\ell=1}^{m} (T^{-1})_{\ell j}\, \widehat{v}_\ell \right) \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{m} \sum_{i=1}^{n} \sum_{j=1}^{m} (S^{-1})_{ki}\, (A)_{ij}\, (T^{-\intercal})_{j\ell}\, \widehat{u}_k \otimes \widehat{v}_\ell \qquad \text{wegen der Bilinearität von } \otimes \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{m} \big(S^{-1} A\, T^{-\intercal}\big)_{k\ell}\, (\widehat{u}_k \otimes \widehat{v}_\ell).
\end{aligned}$$

Das zeigt, dass die Komponentenmatrix von t bzgl. der neuen Basen $\widehat{B}_U$ und $\widehat{B}_V$ gerade gegeben ist durch

$$\widehat{A} = S^{-1} A\, T^{-\intercal}. \tag{23.22}$$

Ein Vergleich mit (20.10) zeigt: Die bei der Transformation der Komponentenmatrix eines Tensors vorkommenden Matrizen sind

$$S^{-1} = \mathcal{T}_{\widehat{B}_U \leftarrow B_U} \quad \text{und} \quad T^{-\intercal} = \mathcal{T}_{B_{V^*} \leftarrow \widehat{B}_V^*}.$$

Die Komponentenmatrix A eines Tensors in $U \otimes V$ transformiert sich also nach denselben Regeln, wie es die Darstellungsmatrix A einer Abbildung $f \in \operatorname{Homo}(V^*, U)$ tun würde, nämlich gemäß

$$\begin{matrix}
\mathcal{M}_{\widehat{B}_U \leftarrow \widehat{B}_V^*}(f) & = & \mathcal{T}_{\widehat{B}_U \leftarrow B_U} & \mathcal{M}_{B_U \leftarrow B_{V^*}}(f) & \mathcal{T}_{B_{V^*} \leftarrow \widehat{B}_V^*} \\
\widehat{A} & = & S^{-1} & A & T^{-\intercal} & .
\end{matrix} \tag{23.23}$$

Dieser Beobachtung werden wir im folgenden Abschnitt nachgehen.

§ 23.8 Tensoren als lineare Abbildungen

Wie wir gerade in (23.23) gesehen haben, transformiert sich die Komponentenmatrix eines Tensors in $U \otimes V$ beim Wechsel der Basen von U und V nach denselben Regeln wie die Darstellungsmatrix einer Abbildung $f \in \operatorname{Homo}(V^*, U)$. Das liegt daran, dass Tensoren über endlichdimensionalen Vektorräumen tatsächlich i. W. dasselbe sind wie lineare Abbildungen. Genauer: Es gibt einen kanonischen Isomorphismus zwischen $U \otimes V$ und $\operatorname{Homo}(V^*, U)$:

Satz 23.40 (Tensoren „sind“ lineare Abbildungen)**.**
Es seien U und V endlich-dimensionale Vektorräume über demselben Körper K und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann besteht der folgende kanonische Isomorphismus, der durch seine Bilder auf den Elementartensoren eindeutig bestimmt ist:

$$I\colon U \otimes V \ni u \otimes v \mapsto \langle \cdot\,, v \rangle\, u \in \operatorname{Homo}(V^*, U). \tag{23.24}$$

Sind B_U bzw. B_V Basen von U bzw. von V mit den zugehörigen dualen Basen B_{U^*} bzw. B_{V^*}, so stimmt die Komponentenmatrix eines Tensors $t \in U \otimes V$ mit der Darstellungsmatrix der ihm zugeordneten Abbildung $I(t) \in \operatorname{Homo}(V^*, U)$ überein:

$$\Phi^{-1}_{B_{U\otimes V}}(t) = \mathcal{M}_{B_U \leftarrow B_{V^*}}(I(t)). \tag{23.25}$$

Beweis. Die Abbildungsvorschrift

$$U \times V \ni (u, v) \mapsto \langle \cdot\,, v\rangle\, u \in \operatorname{Homo}(V^*, U)$$

ist bilinear. Also ist I nach Bemerkung 23.9 wohldefiniert.

Schritt 1: I ist ein Isomorphismus.

Die Räume $U\otimes V$ und $\operatorname{Homo}(V^*, U)$ haben dieselbe Dimension. Wir zeigen: $\operatorname{Kern}(I) = \{0\}$ (der Nulltensor in $U\otimes V$). In der Tat ist $\langle \cdot\,, v\rangle\, u$ die Nullabbildung in $\operatorname{Homo}(V^*, U)$ genau dann, wenn $u = 0$ oder $v = 0$ ist. (**Quizfrage 23.9:** Klar?) Daher ist I injektiv und wegen Folgerung 18.9 auch surjektiv, also ein Isomorphismus.

Schritt 2: Wir zeigen die Gleichheit der Komponenten- und der Darstellungsmatrix (23.25).

Dazu sei B_U bzw. B_V Basen von U bzw. von V mit den zugehörigen dualen Basen B_{U^*} bzw. B_{V^*}.

Es sei $t \in U \otimes V$ ein Tensor und A seine Komponentenmatrix bzgl. der Tensorproduktbasis $(u_i \otimes v_j)_{(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ von $U \otimes V$, also

$$t = \Phi_{B_{U\otimes V}}(A) = \sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k\ell}\,(u_k \otimes v_\ell).$$

Wir betrachten die Darstellungsmatrix $B \coloneqq \mathcal{M}_{B_U\leftarrow B_{V^*}}(I(t))$. Deren j-te Spalte enthält die Koeffizienten des Bildes $I(t)(v_j^*)$ bzgl. der Basis B_U. Der Eintrag B_{ij} ist also gerade

$$\begin{aligned}
B_{ij} &= \langle u_i^*, I(t)(v_j^*)\rangle && \text{denn } u_i^* \text{ ist der } i\text{-te Koordinatenermittler des Bildes von } v_j^* \text{ unter } I(t)\\
&= \Big\langle u_i^*, I\Big(\sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k\ell}\,(u_k\otimes v_\ell)\Big)(v_j^*)\Big\rangle && \text{wegen der Darstellung von } t\\
&= \Big\langle u_i^*, \Big(\sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k\ell}\, I(u_k\otimes v_\ell)\Big)(v_j^*)\Big\rangle && \text{aufgrund der Linearität von } I\\
&= \Big\langle u_i^*, \Big(\sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k\ell}\,\langle \cdot\,, v_\ell\rangle\, u_k\Big)(v_j^*)\Big\rangle && \text{wegen der Definition von } I\\
&= \Big\langle u_i^*, \sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k\ell}\,\langle v_j^*, v_\ell\rangle\, u_k\Big\rangle\\
&= \sum_{k=1}^{n}\sum_{\ell=1}^{m} a_{k\ell}\,\underbrace{\langle v_j^*, v_\ell\rangle}_{=\delta_{j\ell}}\,\underbrace{\langle u_i^*, u_k\rangle}_{=\delta_{ik}} && \text{wegen der Linearität der dualen Paarung}
\end{aligned}$$

$$= a_{ij}.$$

Damit ist (23.25) gezeigt. □

Beachte: Wir könnten einen Tensor analog auch als eine lineare Abbildung in $\operatorname{Homo}(U^*, V)$ auffassen, indem wir den Isomorphismus

$$U \otimes V \ni u \otimes v \mapsto \langle \cdot\,, u \rangle\, v \in \operatorname{Homo}(U^*, V)$$

verwenden, bei den gegenüber (23.24) die Rollen von U und V vertauscht werden. (**Quizfrage 23.10:** Wie wird die Darstellungsmatrix dieser Abbildung aussehen?)

Was sagt uns die Erkenntnis aus Satz 23.40, dass Tensoren in $U \otimes V$ dasselbe (in anderer Gestalt) sind wie lineare Abbildungen in $\operatorname{Homo}(V^*, U)$ und auch wie lineare Abbildungen in $\operatorname{Homo}(U, V^*)$?

Wir können uns einen Tensor in $U \otimes V$ vorstellen als eine flexibel verwendbare „Maschine" mit zwei Eingängen. Der erste Eingang akzeptiert Vektoren aus U^*. Wird dieser Eingang mit einem Vektor aus U^* belegt, so wird er vom Tensor „konsumiert", und wir erhalten als Ergebnis einen Vektor aus V. Der Faktor U im Tensorproduktraum $U \otimes V$ wird dabei „verbraucht". Hier wirkt der Tensor als lineare Abbildung $U^* \longrightarrow V$.

Wird dagegen der zweite Eingang mit einem Vektor aus V^* belegt, so erhalten wir als Ergebnis einen Vektor aus U. Der Tensor wirkt dann als lineare Abbildung $V^* \longrightarrow U$. Werden sogar beide Eingänge gleichzeitig belegt, so ergibt sich als Ergebnis ein Skalar aus dem Körper K. Der Tensor wirkt dann als lineare Abbildung $U^* \otimes V^* \longrightarrow K$.

Diese Funktionsweisen können wir am besten anhand der Komponentenmatrix eines Tensors erläutern (Abbildung 23.1). Es seien dazu $(u_1, \dots, u_n)$ und $(v_1, \dots, v_m)$ Basen von U und V mit den zugehörigen dualen Basen $(u_1^*, \dots, u_n^*)$ und $(v_1^*, \dots, v_m^*)$.[18] Der Tensor habe die Komponentendarstellung

$$t = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, (u_i \otimes v_j).$$

Setzen wir jetzt einen Vektor $u^* = \sum_{i=1}^{n} \xi_i\, u_i^* \in U^*$ in den ersten Eingang ein, so ergibt sich als Resultat der Vektor

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, \underbrace{\langle u^*, u_i \rangle}_{=\xi_i}\, v_j = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, \xi_i\, v_j \in V$$

mit dem Koordinaten(zeilen)vektor $\sum_{i=1}^{n} \xi_i\, a_{i\bullet} \in K_m$. Hier wirkt der Tensor t also als Element von $\operatorname{Homo}(U^*, V)$.

Setzen wir dagegen einen Vektor $v^* = \sum_{j=1}^{m} \eta_j\, v_j^* \in V^*$ in den zweiten Eingang ein, so erhalten wir

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, u_i\, \underbrace{\langle v^*, v_j \rangle}_{=\eta_j} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, u_i\, \eta_j \in U$$

[18] Wir kennzeichnen hier Vektoren aus den primalen Räumen U und V und Vektoren aus den Dualräumen U^* und V^* farbig, sodass leicht zu erkennen ist, wo sich ein dualer und ein primaler Vektor treffen und aus ihnen ein Skalar entsteht. Alle anderen Größen sind Skalare.

mit dem Koordinaten(spalten)vektor $\sum_{j=1}^m \eta_j \, a_{\bullet j} \in K^n$. Hier wirkt der Tensor t jetzt als Element von $\operatorname{Homo}(V^*, U)$.

Wir können auch beide Eingänge gleichzeitig belegen und erhalten dann

$$\sum_{i=1}^n \sum_{j=1}^m a_{ij} \underbrace{\langle u^*, u_i \rangle}_{=\xi_i} \underbrace{\langle v^*, v_j \rangle}_{=\eta_j} = \sum_{i=1}^n \sum_{j=1}^m a_{ij} \, \xi_i \, \eta_j \in K.$$

Hier agiert der Tensor t als Bilinearform, also als ein Element von $\operatorname{Bil}(U, V; K)$ bzw. als Element von $\operatorname{Homo}(U^* \otimes V^*, K)$.

Diese flexible Verwendbarkeit eines Tensors in $U \otimes V$ äußert sich mathematisch darin, dass es — bei endlich-dimensionalen Vektorräumen U und V — die folgenden kanonischen Isomorphismen gibt:

$$\begin{tikzcd}
 & \begin{array}{c} U \otimes V \cong \\ \operatorname{Homo}(K, U \otimes V) \end{array} \arrow[dl, dash] \arrow[dr, dash] & \\
\operatorname{Homo}(U^*, V) \arrow[dr, dash] & & \operatorname{Homo}(V^*, U) \arrow[dl, dash] \\
 & \operatorname{Homo}(U^* \otimes V^*, K) &
\end{tikzcd}$$

Bemerkung 23.41 (Umgang mit Komponentenmatrizen)**.**

(i) In der Sprache der Komponentenmatrizen bedeutet das Einsetzen eines Vektors, dass — wie bei der Matrix-Vektor-Multiplikation — entlang der betreffenden Achse eine mit den Einträgen des Koeffizientenvektors gewichtete Summe gebildet wird. Dabei verschwindet die betreffende Achse, und aus der Komponentenmatrix wird ein Koordinatenvektor, vgl. Abbildung 23.1.

(ii) Die eingesetzten Koordinatenvektoren ξ bzw. η haben zwar die feste Dimension n bzw. m, aber die Frage, ob sie als Zeilen- oder Spaltenvektoren anzusehen sind, ist vollkommen unerheblich. Letztlich ist ja auch die Anordnung der Achsen des Tensors (erste Achse = U-Achse: vertikal, zweite Achse = V-Achse: horizontal) willkürlich gewählt. △

Ende der Vorlesung 7

Ende der Woche 3

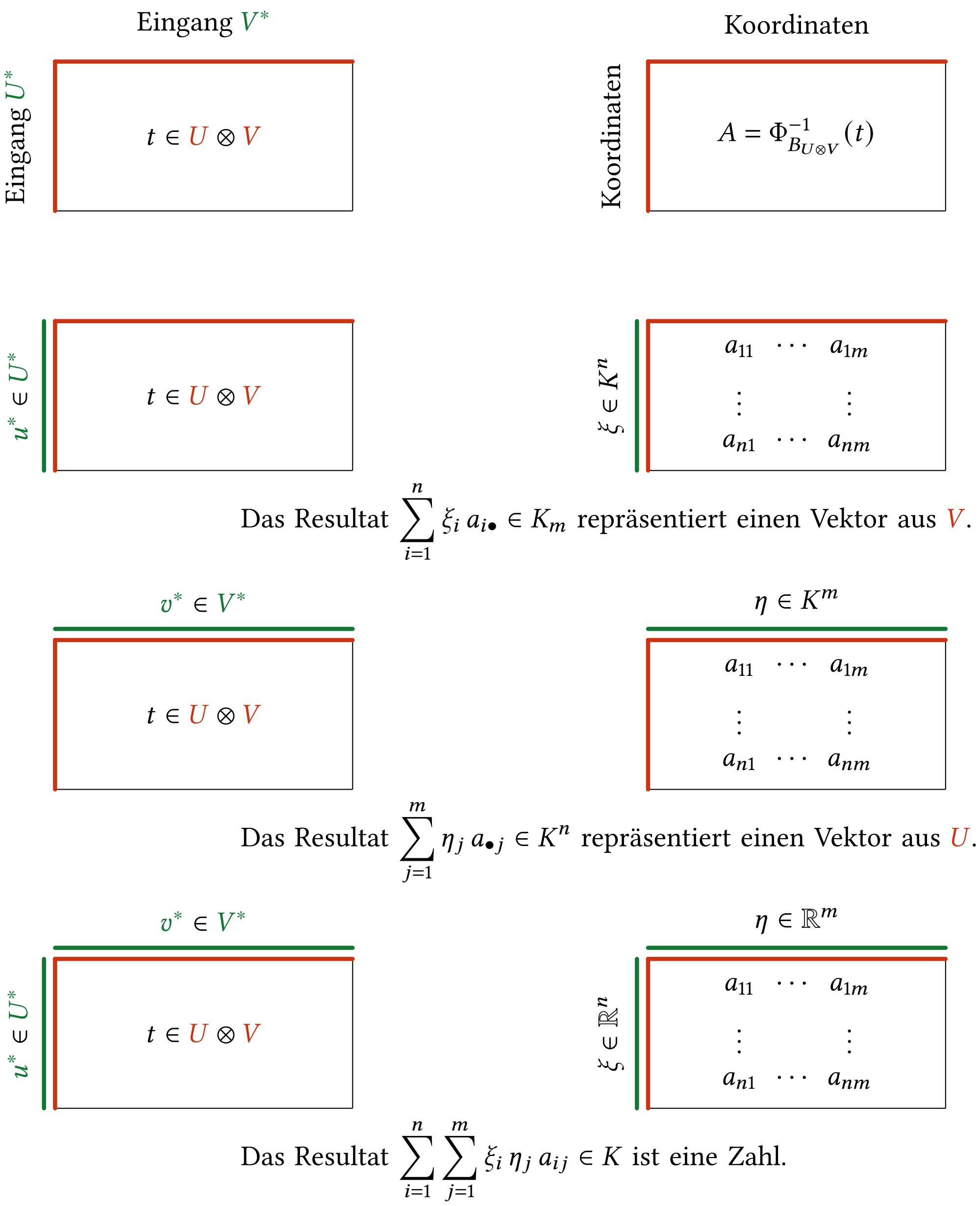


Abbildung 23.1.: Verschiedene Nutzungsmöglichkeiten eines Tensors $t \in U \otimes V$ (linke Spalte) und deren Realisierung mittels Komponentenmatrix A (rechte Spalte). In der zweiten Zeile „konsumiert" der Tensor einen Vektor aus U^* und liefert als Ergebnis einen Vektor aus V. Hier agiert der Tensor also als ein Element von $\operatorname{Homo}(U^*, V)$. In der dritten Zeile „konsumiert" der Tensor einen Vektor aus V^* und liefert als Ergebnis einen Vektor aus U. Hier agiert der Tensor als ein Element von $\operatorname{Homo}(V^*, U)$. In der vierten Zeile „konsumiert" der Tensor gleichzeitig einen Vektor aus U^* und einen Vektor aus V^* und liefert als Ergebnis einen Skalar aus K. Hier agiert der Tensor als eine Bilinearform auf $U^* \times V^*$ bzw. als Element von $\operatorname{Homo}(U^* \otimes V^*, K)$.

§ 24 Multilineare Abbildungen und das mehrfache Tensorprodukt von Vektorräumen

Bisher haben wir Tensorprodukte von zwei Vektorräumen betrachtet, mit deren Hilfe wir bilineare Abbildungen als lineare Abbildungen darstellen konnten. Allgemeiner können wir auch multilineare Abbildungen und mehrfache Tensorprodukte zulassen. Da die Argumente dieselben sind wie bei Tensorprodukten von zwei Vektorräumen, geben wir in diesem Abschnitt alle Resultate ohne Beweise an.

Definition 24.1 (multilineare Abbildung, Multilinearform)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ sowie W Vektorräume über demselben Körper K.

(i) Eine Abbildung

$$m\colon \bigtimes_{k=1}^{N} V_k = V_1 \times \cdots \times V_N \to W$$

heißt **multilinear** (englisch: multilinear map) oder genauer **N-linear** (englisch: N-linear map), wenn für jedes $i \in [\![1, N]\!]$ und alle fest gewählten $v_j \in V_j$, $j \in [\![1, N]\!] \setminus \{i\}$, die Abbildung

$$m(v_1, \ldots, v_{i-1}, \cdot, v_{i+1}, \ldots, v_N)\colon V_i \ni v_i \mapsto m(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_N) \in W \tag{24.1}$$

linear ist.

(ii) Die Menge aller multilinearen Abbildungen $V_1 \times \cdots \times V_N \to W$ bezeichnen wir mit dem Symbol $\operatorname{Mult}(V_1, \ldots, V_N; W)$.

(iii) Eine multilineare Abbildung in den Vektorraum $W = K$ nennen wir eine **Multilinearform** (englisch: multilinear form) auf $V_1 \times \cdots \times V_N$.

(iv) Die Menge aller Multilinearformen $V_1 \times \cdots \times V_N \to K$ bezeichnen wir mit $\operatorname{Mult}(V_1, \ldots, V_N; K)$ oder kurz mit $\operatorname{Mult}(V_1, \ldots, V_N)$. △

Bemerkung 24.2 (multilineare Abbildungen)**.**

(i) Im Fall $N = 0$ ist das kartesische Produkt $V_1 \times \cdots \times V_N$ die Menge, die nur aus dem leeren Tupel $()$ besteht, vgl. Definitionen 4.8 und 6.42. Jede Abbildung $\{()\} \to W$ ist multilinear.

(ii) Im Fall $N = 1$ sind die multilinearen Abbildungen $V_1 \to W$ gerade die linearen Abbildungen, es gilt also $\operatorname{Mult}(V_1; W) = \operatorname{Homo}(V_1, W)$.

(iii) Der Fall $N = 2$ entspricht den bilinearen Abbildungen $V_1 \times V_2 \to W$ aus § 23, es gilt also $\operatorname{Mult}(V_1, V_2; W) = \operatorname{Bil}(V_1, V_2; W)$. △

Es gelten die mehrdimensionalen Analoga von Satz 23.3 und Satz 23.4:

Satz 24.3 (multilineare Abbildungen bilden einen Vektorraum, vgl. Satz 23.3)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ sowie W Vektorräume über demselben Körper. Dann ist $\operatorname{Mult}(V_1, \ldots, V_N; W)$ ein Unterraum des Vektorraumes $W^{V_1 \times \cdots \times V_N} = \{f\colon V_1 \times \cdots \times V_N \to W\}$ aller Abbildungen $V_1 \times \cdots \times V_N \to W$.

Satz 24.4 (Existenz- und Eindeutigkeitssatz für multilineare Abbildungen, vgl. Satz 23.4)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ sowie W Vektorräume über demselben Körper. Weiter seien $(v_{k,i_k})_{i_k \in I_k}$ Basen von V_k für $k = 1, \ldots, N$. Außerdem sei $(w_{i_1,\ldots,i_N})_{(i_1,\ldots,i_N) \in I_1 \times \cdots \times I_N}$ eine Familie von Vektoren in W. Dann gibt es genau eine N-ltilineare Abbildung $m\colon V_1 \times \cdots \times V_N \to W$ mit der Eigenschaft $m(v_{i_1}, \ldots, v_{i_N}) = w_{i_1,\ldots,i_N}$ für alle $(i_1, \ldots, i_N) \in I_1 \times \cdots \times I_N$.

In Erweiterung von Definition 23.6 können wir angeben:

Definition 24.5 (Tensorprodukt, Tensorproduktraum, universelle multilineare Abbildung, vgl. Definition 23.6)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ Vektorräume über demselben Körper K. Weiter sei $\bigotimes_{k=1}^N V_k$ ein weiterer K-Vektorraum[19] und $\otimes\colon \bigtimes_{k=1}^N V_k \to \bigotimes_{k=1}^N V_k$ eine N-lineare Abbildung.[20]

(i) Das Paar $(\bigotimes_{k=1}^N V_k, \otimes)$ heißt ein **Tensorprodukt** von $V_1, \ldots, V_N$, wenn die folgende **universelle Eigenschaft** erfüllt ist:

Für jede N-lineare Abbildung $m\colon \bigtimes_{k=1}^N V_k \to W$ in irgendeinen K-Vektorraum W gibt es eine eindeutig bestimmte lineare Abbildung $f \in \operatorname{Homo}(\bigotimes_{k=1}^N V_k, W)$ mit der Eigenschaft $m = f \circ \otimes$, also $m(v_1, \ldots, v_N) = f(v_1 \otimes \cdots \otimes v_N)$ für alle $v_k \in V_k$, $k = 1, \ldots, N$.

(ii) Dabei heißt $\bigotimes_{k=1}^N V_k$ der **Tensorproduktraum** des Tensorprodukts $(\bigotimes_{k=1}^N V_k, \otimes)$. Die Elemente von $\bigotimes_{k=1}^N V_k$ heißen **Tensoren** der **Stufe** (englisch: degree) oder **Ordnung** (englisch: order) N oder auch **N-achsige Tensoren** (englisch: N-axis tensors) oder kurz **N-Tensoren** (englisch: N-tensors).[21] Der Nullvektor in $\bigotimes_{k=1}^N V_k$ heißt der **Nulltensor**.

(iii) $\otimes\colon \bigtimes_{k=1}^N V_k \to \bigotimes_{k=1}^N V_k$ heißt die **universelle multilineare Abbildung** des Tensorprodukts $(\bigotimes_{k=1}^N V_k, \otimes)$. Tensoren der Form $v_1 \otimes \cdots \otimes v_N$ mit $v_k \in V_k$ für alle $k = 1, \ldots, N$ — also die Bilder der universellen multilinearen Abbildung — heißen **Elementartensoren** oder **einfache Tensoren**. $v_1 \otimes \cdots \otimes v_N$ heißt auch das **Tensorprodukt** der Vektoren $v_1, \ldots v_N$ oder kurz: „v_1 Tensor $\cdots$ Tensor v_N". △

In Analogie zu Satz 23.8 gilt, dass wir jede multilineare Abbildung durch die universelle multilineare Abbildung eines Tensorprodukts faktorisieren können. Die Eigenschaft der Multilinearität wird also „abgetrennt", und wir können multilineare Abbildungen mit linearen Abbildungen auf dem Tensorproduktraum identifizieren:

Satz 24.6 (universelle Eigenschaft des Tensorprodukts $(\bigotimes_{k=1}^N V_k, \otimes)$, vgl. Satz 23.8)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ Vektorräume über demselben Körper K und $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1, \ldots, V_N$. Außerdem sei auch W ein K-Vektorraum. Dann gilt:

[19] Genau wie $U \otimes V$ ist auch $\bigotimes_{k=1}^N V_k$ nur ein Name für den im Tensorprodukt auftauchenden Vektorraum. Als alternative Bezeichnung werden wir auch das Symbol $V_1 \otimes \cdots \otimes V_N$ benutzen.

[20] Für die Bilder dieser **äußeren Verknüpfung** von N Elementen schreiben wir $v_1 \otimes \cdots \otimes v_N$ an Stelle von $\otimes(v_1, \ldots, v_N)$. Ausnahme: Im Fall $N = 1$ schreiben wir $\otimes(v_1)$ für $v_1 \in V_1$, und im Fall $N = 0$ schreiben wir $\otimes(())$ für das einzige Element $()$ von $\bigtimes_{k=1}^N V_k$ (vgl. Definitionen 4.8 und 6.42).

[21] Manche Autoren verwenden auch den Begriff **Rang** anstelle von **Stufe** bzw.**Ordnung**. Das ist aber missverständlich, siehe Definition 24.12.

(*i*) Zu jeder N-linearen Abbildung $m\colon \bigtimes_{k=1}^N V_k \to W$ existiert genau eine lineare Abbildung $f\colon \bigotimes_{k=1}^N V_k \to W$ mit der Eigenschaft $m = f \circ \otimes$, also $m(v_1, \ldots, v_N) = f(v_1 \otimes \cdots \otimes v_N)$ für alle $v_k \in V_k$, $k = 1, \ldots, N$. Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} \bigtimes_{k=1}^N V_k \arrow[r, "\otimes"] \arrow[dr, "m"'] & \bigotimes_{k=1}^N V_k \arrow[d, "f"] \\ & W \end{tikzcd}$$

(*ii*) Umgekehrt wird für jede lineare Abbildung $f\colon \bigotimes_{k=1}^N V_k \to W$ durch die Komposition mit der universellen N-linearen Abbildung $\otimes\colon \bigtimes_{k=1}^N V_k \to \bigotimes_{k=1}^N V_k$ eine N-lineare Abbildung $f \circ \otimes\colon \bigtimes_{k=1}^N V_k \ni (v_1, \ldots, v_N) \mapsto f(v_1 \otimes \cdots \otimes v_N) \in W$ definiert.

(*iii*) Die Zuordnung

$$\operatorname{Mult}(V_1, \ldots, V_N; W) \ni m \mapsto f \in \operatorname{Homo}(V_1 \otimes \cdots \otimes V_N, W)$$

ist ein kanonischer Isomorphismus von Vektorräumen.

Bemerkung 24.7 (Angabe linearer Abbildungen auf Tensorprodukträumen, vgl. Bemerkung 23.9).

Wie in Bemerkung 23.9 können wir mit Hilfe von Satz 24.6 lineare Abbildungen $f\colon \bigotimes_{k=1}^N V_k \to W$ durch ihre Bilder auf den Elementartensoren $v_1 \otimes \cdots \otimes v_N$ festlegen, sofern die zugehörige Abbildungsvorschrift N-linear ist. △

Satz 24.8 (Eindeutigkeit des Tensorprodukts).
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ Vektorräume über demselben Körper K.

(*i*) Sind $(\bigotimes_{k=1}^N V_k, \otimes)$ und $(\widetilde{\bigotimes}_{k=1}^N V_k, \widetilde{\otimes})$ zwei Tensorprodukte von $V_1, \ldots, V_N$, dann gibt es einen Isomorphismus $I \in \operatorname{Homo}(\bigotimes_{k=1}^N V_k, \widetilde{\bigotimes}_{k=1}^N V_k)$ mit der Eigenschaft $\widetilde{\otimes} = I \circ \otimes$, also $I(v_1 \otimes \cdots \otimes v_N) = v_1 \widetilde{\otimes} \cdots \widetilde{\otimes} v_N$ für alle $v_k \in V_k$, $k = 1, \ldots, N$. Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} & \bigtimes_{k=1}^N V_k \arrow[dl, "\otimes"'] \arrow[dr, "\widetilde{\otimes}"] & \\ \bigotimes_{k=1}^N V_k \arrow[rr, "I"'] & & \widetilde{\bigotimes}_{k=1}^N V_k \end{tikzcd}$$

(*ii*) Ist $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1, \ldots, V_N$ und ist $I \in \operatorname{Homo}(\bigotimes_{k=1}^N V_k, \widetilde{\bigotimes}_{k=1}^N V_k)$ ein Isomorphismus mit einem weiteren K-Vektorraum $\widetilde{\bigotimes}_{k=1}^N V_k$, dann ist auch $(\widetilde{\bigotimes}_{k=1}^N V_k, \widetilde{\otimes})$ ein Tensorprodukt von $V_1, \ldots, V_N$ mit $\widetilde{\otimes} \coloneqq I \circ \otimes\colon \bigtimes_{k=1}^N V_k \to \widetilde{\bigotimes}_{k=1}^N V_k$.

Beispiel 24.9 (mehrfaches Tensorprodukt).

(*i*) Ein 0-faches Tensorprodukt ($N = 0$) von K-Vektorräumen ist isomorph zum Körper K, d. h., es gilt $\bigotimes_{k=1}^0 V_k \cong K$, und $\otimes\colon \{()\} \to K$ mit $\otimes(()) = 1$ ist eine zugehörige universelle 0-lineare Abbildung.

Um diese Behauptung zu bestätigen, prüfen wir die universelle Eigenschaft von $(K, \otimes)$ nach. Zunächst besteht das kartesische Produkt $\bigtimes_{k=1}^{0} V_k = \{()\}$ nur aus dem leeren Tupel $()$ (vgl. Bemerkung 24.2). Es sei nun $m\colon \{()\} \to W$ eine beliebige (automatisch 0-lineare) Abbildung in irgendeinen K-Vektorraum W. Eine solche Abbildung m ist eindeutig durch ihren einzigen Funktionswert $m(())$ bestimmt. Die universelle Eigenschaft von $(K, \otimes)$ verlangt die Existenz und Eindeutigkeit einer linearen Abbildung $f\colon K \to W$, sodass $m = f \circ \otimes$ gilt, also $m(()) = f(\otimes(())) = f(1)$. Durch diese Beziehung ist f tatsächlich eindeutig bestimmt.

(ii) Ein 1-faches Tensorprodukt ($N = 1$) eines K-Vektorraumes V ist isomorph zum Vektorraum V selbst, und $\otimes\colon V \to V$ mit $\otimes(v) = v$ für alle $v \in V$ ist eine zugehörige universelle 1-lineare Abbildung. (**Quizfrage 24.1:** Klar?)

(iii) Die 2-fachen Tensorprodukte $V_1 \otimes V_2$ von K-Vektorräumen V_1 und V_2 wurden in § 23 besprochen. △

Auch beim mehrfachen Tensorprodukt gilt, dass die Elementartensoren eines Tensorproduktraumes den gesamten Vektorraum aufspannen:

Lemma 24.10 (die Elementartensoren bilden ein Erzeugendensystem, vgl. Lemma 23.12)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \dots, V_N$ Vektorräume über demselben Körper und $(\bigotimes_{k=1}^{N} V_k, \otimes)$ ein Tensorprodukt von $V_1, \dots, V_N$. Dann gilt: Die Elementartensoren bilden ein Erzeugendensystem von $\bigotimes_{k=1}^{N} V_k$. Es gilt also

$$\bigotimes_{k=1}^{N} V_k = \left\langle (v_1 \otimes \cdots \otimes v_N)_{(v_1,\dots,v_N) \in \bigtimes_{k=1}^{N} V_k} \right\rangle.$$

Folglich lässt sich jeder Tensor in $\bigotimes_{k=1}^{N} V_k$ als Linearkombination von Elementartensoren schreiben. Für jedes $t \in \bigotimes_{k=1}^{N} V_k$ existieren also $n \in \mathbb{N}_0$, Vektoren $v_{k,1}, \dots, v_{k,n} \in V_k$ sowie $\alpha_1, \dots, \alpha_n \in K$ mit

$$T = \sum_{i=1}^{n} \alpha_i \, (v_{1,i} \otimes \cdots \otimes v_{N,i}). \tag{24.2}$$

Die Koeffizienten α_i können o. B. d. A. alle als 1 gewählt werden.

In Analogie zu Lemma 23.13 können wir aus Basen der Vektorräume $V_1, \dots, V_N$ eine Basis von $\bigotimes_{k=1}^{N} V_k$ zusammenstellen:

Lemma 24.11 (Basis und Dimension des Tensorproduktraumes $\bigotimes_{k=1}^{N} V_k$, vgl. Lemma 23.13)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \dots, V_N$ Vektorräume über demselben Körper mit Basen $B_{V_k} = (v_{k,i_k})_{i_k \in I_k}$ für $k = 1, \dots, N$. Weiter sei $(\bigotimes_{k=1}^{N} V_k, \otimes)$ ein Tensorprodukt von $V_1, \dots, V_N$. Dann gilt:

(i) $B_{V_1 \otimes \cdots \otimes V_N} \coloneqq (v_{1,i_1} \otimes \cdots \otimes v_{N,i_N})_{(i_1,\dots,i_N) \in I_1 \times \cdots \times I_N}$ ist eine Basis von $\bigotimes_{k=1}^{N} V_k$. Sie wird die von $B_{V_1}, \dots, B_{V_N}$ **induzierte Basis** oder die **Tensorproduktbasis** zu $B_{V_1}, \dots, B_{V_N}$ von $\bigotimes_{k=1}^{N} V_k$ genannt.

(*ii*) $\dim(\bigotimes_{k=1}^N V_k) = \prod_{k=1}^N \dim(V_k)$.
(Im Fall $N = 0$ ist das leere Produkt als $1 \in \mathbb{N}$ zu verstehen. Im Fall $N \geqslant 1$ ist das Produkt als 0 zu verstehen, sobald $\dim(V_k) = 0$ für mindestens ein $k \in [\![1, N]\!]$ gilt. Ist das nicht der Fall und ist eine der Dimensionen gleich ∞, dann ist das Produkt als ∞ zu verstehen.)

Definition 24.12 (Rang eines Tensors, vgl. Definition 23.16)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \dots, V_N$ Vektorräume über demselben Körper und $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1, \dots, V_N$. Der **Rang** eines Tensors (englisch: rank of a tensor) $t \in \bigotimes_{k=1}^N V_k$ ist die minimale Anzahl von Summanden, mit denen eine Darstellung der Form (24.2) möglich ist. △

Bemerkung 24.13 (Rang eines Tensors)**.**

(*i*) Im Fall $N = 0$ haben alle Tensoren in $\bigotimes_{k=1}^0 V_k \cong K$ bis auf den Nulltensor (die Null in K) den Rang 1.

(*ii*) Im Fall $N = 1$ haben ebenfalls alle Tensoren in $\bigotimes_{k=1}^1 V_k \cong V_1$ bis auf den Nulltensor (der Nullvektor in V_1) den Rang 1.

(*iii*) Im Fall $N = 2$ gibt es Tensoren $t \in \bigotimes_{k=1}^2 V_k \cong V_1 \otimes V_2$ mit $\operatorname{Rang}(t) = r$ für alle $r \in \mathbb{N}_0$ mit $r \leqslant \min\{\dim(V_1), \dim(V_2)\}$, siehe Folgerung 23.39. Sind beide Dimensionen endlich, dann kann der Rang über den Rang einer Komponentenmatrix von t bestimmt werden (Satz 23.38).

(*iv*) Im Fall $N \geqslant 3$ ist die Bestimmung des Ranges eines N-stufigen Tensors über vielen Körpern K ein NP-schweres Problem und über endlichen Körpern K ein NP-vollständiges Problem im Sinne der Komplexitätstheorie, siehe Håstad, 1990.

(*v*) Selbst der maximal auftretende Rank eines N-stufigen Tensors in $\bigotimes_{k=1}^N V_k$ ist im Fall $N \geqslant 3$ i. A. nicht bekannt, nicht einmal dann, wenn alle V_k dieselbe endliche Dimension besitzen. Bekannt sind im Fall $N = 3$ zur Zeit (Anfang 2026):

$$\begin{aligned} \dim(V_1) = \dim(V_2) = \dim(V_3) = 2 \quad &\Rightarrow \quad \max\{\operatorname{Rang}(t) \mid t \in V_1 \otimes V_2 \otimes V_3\} = 3 \\ \dim(V_1) = \dim(V_2) = \dim(V_3) = 3 \quad &\Rightarrow \quad \max\{\operatorname{Rang}(t) \mid t \in V_1 \otimes V_2 \otimes V_3\} = 5. \end{aligned}$$

Die Bestimmung weiterer scharfer Schranken für den maximalen Rang eines N-stufigen Tensors ($N \geqslant 3$) ist ein aktives Forschungsgebiet. △

Lemma 24.14 (Rang von Elementartensoren, vgl. Lemma 23.18)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \dots, V_N$ Vektorräume mit Basen über demselben Körper und $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1, \dots, V_N$. Dann gilt:

(*i*) Der Nulltensor ist der einzige Tensor vom Rang 0.

(*ii*) Für $v_k \in V_k$ mit $k = 1, \dots, N$ gilt

$$v_1 \otimes \cdots \otimes v_N = 0 \quad \text{(Nulltensor)} \quad \Leftrightarrow \quad \text{mindestens ein } v_k \text{ ist der Nullvektor.} \tag{24.3}$$

(*iii*) Jeder Elementartensor $v_1 \otimes \cdots \otimes v_N$, wobei alle $v_k \in V_k$ für $k = 1, \dots, N$ ungleich dem Nullvektor sind, hat Rang 1.

Analog zu Definition 23.30 können wir das Tensorprodukt von $N \in \mathbb{N}_0$ linearen Abbildungen bilden:

Definition 24.15 (Tensorprodukt linearer Abbildungen, vgl. Definition 23.30)**.**

Es seien $N \in \mathbb{N}_0$ und $U_1, \ldots, U_N$ sowie $V_1, \ldots, V_N$ Vektorräume über demselben Körper und $(\bigotimes_{k=1}^N U_k, \otimes)$ ein Tensorprodukt von $U_1, \ldots, U_N$ sowie $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1, \ldots, V_N$. Dann heißt die für $u_k \in U_k$, $k = 1, \ldots, N$ durch

$$f_1 \boxtimes \cdots \boxtimes f_N \colon \begin{cases} U_1 \otimes \cdots \otimes U_N \to V_1 \otimes \cdots \otimes V_N \\ u_1 \otimes \cdots \otimes u_N \mapsto f_1(u_1) \otimes \cdots \otimes f_N(u_N) \end{cases} \tag{24.4}$$

eindeutig bestimmte lineare Abbildung das **Tensorprodukt** der Abbildungen $f_1, \ldots, f_N$. △

Ähnlich wie in § 23.6 geht es nun um eine Standarddarstellung von N-achsigen Tensoren in einem Tensorproduktraum der endlich-dimensionalen Vektorräume $V_1, \ldots, V_N$. Gilt $\dim(V_k) = n_k \in \mathbb{N}_0$ für $k = 1, \ldots, N$, dann ist nach Lemma 24.11 $\dim(\bigotimes_{k=1}^N V_k) = \prod_{k=1}^N n_k$. Zur Darstellung von N-Tensoren bieten sich demnach N-achsige Verallgemeinerungen von Matrizen an, die **Hypermatrizen** genannt werden.

Definition 24.16 (Hypermatrix, vgl. Definition 15.1)**.**

Es seien K ein Körper, $N \in \mathbb{N}_0$ sowie $n_1, \ldots, n_N \in \mathbb{N}_0$.

(*i*) Eine N-fach indizierte Familie in K mit der Indexmenge $[\![1, n_1]\!] \times \cdots \times [\![1, n_N]\!]$ heißt eine **Hypermatrix** (englisch: hypermatrix) der **Stufe** (englisch: degree) oder **Ordnung** (englisch: order) N oder auch eine **N-achsige Hypermatrix** (englisch: N-axis hypermatrix) der **Dimension $n_1 \times \cdots \times n_N$ über dem Körper K**.[22] Wir schreiben die Hypermatrix in der Form

$$A = (a_{i_1,\ldots,i_N})_{(i_1,\ldots,i_N) \in [\![1,n_1]\!] \times \cdots \times [\![1,n_N]\!]}. \tag{24.5}$$

(*ii*) Die Menge aller Hypermatrizen der Dimension $n_1 \times \cdots \times n_N$ wird mit $K^{n_1 \times \cdots \times n_N}$ bezeichnet.[23]

(*iii*) Sind $x_1 \in K^{n_1}, \ldots, x_N \in K^{n_N}$ Spaltenvektoren, so heißt diejenige N-achsige Hypermatrix A der Dimension $n_1 \times \cdots \times n_N$, deren Einträge durch die Produkte

$$a_{i_1,\ldots,i_N} = \underbrace{x_{1,i_1}}_{i_1\text{-Koordinate von } x_1} \cdots \underbrace{x_{N,i_N}}_{i_N\text{-Koordinate von } x_N} \tag{24.6}$$

gegeben sind, das **äußere Produkt** (englisch: outer product) der Vektoren $x_1, \ldots, x_N$, kurz: $A = x_1 \boxtimes \cdots \boxtimes x_N$.[24]

[22] Wir können eine Hypermatrix also als eine Abbildung $[\![1, n_1]\!] \times \cdots \times [\![1, n_N]\!] \to K$ auffassen.
[23] Alternative Bezeichnungen sind $K^{n_1,\ldots,n_N}$ oder $M_{n_1,\ldots,n_M}(K)$.
[24] Im Fall $N = 2$ ist das äußere Produkt der Vektoren x_1 und x_2 gerade die Matrix $x_1\, x_2^\intercal$.

(*iv*) Der **Rang**[25] einer Hypermatrix $A \in K^{n_1 \times \cdots \times n_N}$ ist die minimale Anzahl von Summanden, mit denen eine Darstellung der Form

$$A = \sum_{i=1}^{n} x_1^{(i)} \boxtimes \cdots \boxtimes x_N^{(i)} \tag{24.7}$$

mit Vektoren $x_k^{(i)} \in K^{n_k}$ für $k = 1, \ldots, N$ und $i = 1, \ldots, n$ möglich ist. △

Bemerkung 24.17 (zu Hypermatrizen).

(*i*) Die Hypermatrizen der Dimension $n_1 \times \cdots \times n_N$ über einem Körper K bilden einen Vektorraum über K bezüglich der komponentenweisen Addition und S-Multiplikation.

(*ii*) Die Dimension dieses Vektorraumes $K^{n_1 \times \cdots \times n_N}$ ist $\prod_{k=1}^{N} n_k$.

(*iii*) Die Matrizen $E_{i_1,\ldots,i_N}$ mit $i_k \in [\![1, n_k]\!]$ für $k = 1, \ldots, N$ bilden eine Basis von $K^{n_1 \times \cdots \times n_N}$, die die **Standardbasis** genannt wird. Dabei ist $E_{i_1,\ldots,i_N}$ die Hypermatrix, deren einziges von Null verschiedenes Element an der Stelle $(i_1, \ldots, i_N)$ den Wert 1 besitzt. Es gilt

$$E_{i_1,\ldots,i_N} = e_{i_1} \boxtimes \cdots \boxtimes e_{i_N}$$

mit den Standardbasisvektoren $e_{i_k} \in K^{n_k}$ für $k = 1, \ldots, N$.

(*iv*) Eine 0-achsige Hypermatrix ist eine Abbildung $\{()\} \to K$, kann also als ein Skalar aus K aufgefasst werden.

(*v*) Eine 1-achsige Hypermatrix kann als Spaltenvektor in K^{n_1} (oder als Zeilenvektor in K_{n_1}) aufgefasst werden.

(*vi*) Eine 2-achsige Hypermatrix kann als Matrix in $K^{n_1 \times n_2}$ aufgefasst werden. Der Rang-Begriff aus Definition 24.16 stimmt mit dem bekannt Rang-Begriff für Matrizen als Definition 15.14 überein.

(*vii*) Allgemein kann eine N-achsige Hypermatrix als ein N-dimensionales Array von Elementen des Körpers K verstanden werden. △

Wir können nun die Tensoren in einem Tensorproduktraum endlich-dimensionaler Vektorräume durch Hypermatrizen darstellen:

Satz 24.18 (Komponentenhypermatrix, Syntheseabbildung, Analyseabbildung, vgl. Satz 23.34). Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ Vektorräume über demselben Körper mit $\dim(V_k) = n_k \in \mathbb{N}_0$ für $k = 1, \ldots, N$. Weiter seien die Familien $B_{V_k} = (v_{k,i_k})_{i_k \in I_k}$ für $k = 1, \ldots, N$ Basen von $V_1, \ldots, V_N$ und $B_{V_1 \otimes \cdots \otimes V_N} \coloneqq (v_{1,i_1} \otimes \cdots \otimes v_{N,i_N})_{(i_1,\ldots,i_N) \in I_1 \times \cdots \times I_N}$ die zugehörige Tensorproduktbasis von $\bigotimes_{k=1}^{N} V_k$ (Lemma 24.11). Dann gilt:

(*i*) Die Abbildung

$$\Phi_{B_{V_1 \otimes \cdots \otimes V_N}} : K^{n_1 \times \cdots \times n_N} \ni A \mapsto t \coloneqq \sum_{i_1=1}^{n_1} \cdots \sum_{i_N=1}^{n_N} a_{i_1,\ldots,i_N} \, (v_{1,i_1} \otimes \cdots \otimes v_{N,i_N}) \in \bigotimes_{k=1}^{N} V_k \tag{24.8}$$

[25] Zur Unterscheidung von anderen Rangbegriffen wird dieser Rangbegriff auch als **Tensor-Rang** (englisch: tensor rank) oder **CP-Rang** (englisch: canonical polyadic rank) der Hypermatrix bezeichnet. Wir werden jedoch in dieser Lehrveranstaltung keine anderen Rangbegriffe für Hypermatrizen bzw. Tensoren benötigen.

ist ein linearer Isomorphismus $K^{n_1\times\cdots\times n_N} \to \bigotimes_{k=1}^N V_k$. Diese erzeugt aus einer **Komponentenhypermatrix** (englisch: component hypermatrix) $A \in K^{n_1\times\cdots\times n_N}$ die zugehörige Linearkombination der Basistensoren $v_{1,i_1} \otimes \cdots \otimes v_{N,i_N}$. Wir bezeichnen die Abbildung $\Phi_{B_{V_1\otimes\cdots\otimes V_N}}$ daher auch als die **Syntheseabbildung** bzgl. der Basis $B_{V_1\otimes\cdots\otimes V_N}$.

(ii) Der zu $\Phi_{B_{V_1\otimes\cdots\otimes V_N}}$ inverse Isomorphismus ist die Abbildung

$$\Phi^{-1}_{B_{V_1\otimes\cdots\otimes V_N}} : \bigotimes_{k=1}^N V_k \ni t \mapsto A \in K^{n_1\times\cdots\times n_N}, \tag{24.9}$$

die jedem Tensor $t \in \bigotimes_{k=1}^N V_k$ seine eindeutige **Komponentenhypermatrix** $A \in K^{n_1\times\cdots\times n_N}$ bzgl. der Basis $B_{V_1\otimes\cdots\otimes V_N}$ zuordnet. Wir nennen daher $\Phi^{-1}_{B_{V_1\otimes\cdots\otimes V_N}}$ auch die **Komponentenabbildung** oder die **Analyseabbildung** bzgl. der Basis $B_{V_1\otimes\cdots\otimes V_N}$. Der Eintrag $a_{i_1,\ldots,i_N} \in K$ heißt die $(i_1,\ldots,i_N)$-te **Komponente** des Tensors $t \in \bigotimes_{k=1}^N V_k$ bzgl. der Basis $B_{V_1\otimes\cdots\otimes V_N}$.

Beispiel 24.19 (Komponentenhypermatrizen für Tensoren verschiedener Stufe)**.**

(i) 0-achsige Tensoren werden durch 0-achsige Hypermatrizen (Skalare) dargestellt.

(ii) 1-achsige Tensoren werden durch 1-achsige Hypermatrizen (Spaltenvektoren) dargestellt.

(iii) 2-achsige Tensoren werden durch 2-achsige Hypermatrizen (gewöhnliche Matrizen) dargestellt.

(iv) N-achsige Tensoren mit $N \geqslant 3$ werden durch N-achsige Hypermatrizen (N-dimensionale Arrays) dargestellt. △

Auch für mehrfache Tensorprodukte können wir die Gestalt des Dualraumes $(\bigotimes_{k=1}^N V_k)^*$ untersuchen und die zu einer Tensorproduktbasis duale Basis angeben:

Lemma 24.20 (Dualraum des Tensorprodukts und duale Basis, vgl. Lemma 23.37)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1,\ldots,V_N$ endlich-dimensionale Vektorräume über demselben Körper, $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1,\ldots,V_N$ sowie $(\bigotimes_{k=1}^N V_k^*, \otimes)$ ein Tensorprodukt der Dualräume $V_1^*,\ldots,V_N^*$. Dann gilt:

(i) Die Abbildung

$$\begin{aligned} \bigotimes_{k=1}^N V_k^* &\to \Big(\bigotimes_{k=1}^N V_k\Big)^* \\ v_1^* \otimes \cdots \otimes v_N^* &\mapsto v_1^* \boxtimes \cdots \boxtimes v_N^* = \langle v_1^*\,,\cdot\rangle \cdots \langle v_N^*\,,\cdot\rangle \end{aligned} \tag{24.10a}$$

definiert einen kanonischen linearen Isomorphismus. Dabei ist $v_1^* \boxtimes \cdots \boxtimes v_N^* \in (\bigotimes_{k=1}^N V_k)^*$ gegeben durch die Bilder auf den Elementartensoren, nämlich durch[26]

$$(v_1^* \boxtimes \cdots \boxtimes v_N^*)(v_1 \otimes \cdots \otimes v_N) \coloneqq \langle v_1^*\,, v_1\rangle \cdots \langle v_N^*\,, v_N\rangle \tag{24.10b}$$

für alle $v_k \in V_k$ und $k = 1,\ldots,N$.

[26]Nach Definition 24.15 ist das Tensorprodukt $v_1^* \boxtimes v_N^*$ der linearen Abbildungen $v_1^* \in V_1^* = \operatorname{Homo}(V_1,K)$ usw. bis $v_N^* \in V_N^* = \operatorname{Homo}(V_N,K)$ eigentlich eine lineare Abbildung $\bigotimes_{k=1}^N V_k \to K \otimes \cdots \otimes K$. Wir können aber $K \otimes \cdots \otimes K \cong K$ identifizieren. Dabei wird das Tensorprodukt $\alpha_1 \otimes \cdots \otimes \alpha_N$ zur Multiplikation $\alpha_1 \cdots \alpha_N$ in K.

(*ii*) Es seien $B_{V_k} = (v_{k,i_1}, \ldots, v_{k,n_k})$ für $k = 1, \ldots, N$ Basen von $V_1, \ldots, V_N$ und $B^*_{V_k} = (v^*_{k,1}, \ldots, v^*_{k,n_k})$ die zugehörigen dualen Basen von $V_1^*, \ldots, V_N^*$. Dann ist

$$(v^*_{1,i_1} \boxtimes \cdots \boxtimes v^*_{N,i_N})_{(i_1,\ldots,i_N) \in [\![1,n_1]\!] \times \cdots \times [\![1,n_N]\!]} \tag{24.11}$$

die zu $B_{V_1 \otimes \cdots \otimes V_N} = (v_{1,i_1} \otimes \cdots \otimes v_{N,i_N})_{(i_1,\ldots,i_N) \in [\![1,n_1]\!] \times \cdots \times [\![1,n_N]\!]}$ duale Basis von $(V_1 \otimes \cdots \otimes V_N)^*$.

Die Elemente der dualen Basis dienen wieder als „Komponentenermittler" eines Tensors. Mit den Bezeichnungen aus Satz 24.18 gilt

$$a_{i_1,\ldots,i_N} = (v^*_{1,i_1} \boxtimes \cdots \boxtimes v^*_{N,i_N})(t). \tag{24.12}$$

Auch bei mehrfachen Tensorprodukten gilt wieder, dass der Rang eines Tensors mit dem Rang seiner Komponentenhypermatrix übereinstimmt:[27]

Satz 24.21 (Rang eines Tensors ist Rang seiner Komponentenhypermatrizen, vgl. Satz 23.38)**.** Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ Vektorräume über demselben Körper mit $\dim(V_k) = n_k \in \mathbb{N}_0$ für $k = 1, \ldots, N$ sowie $(\bigotimes_{k=1}^N V_k, \otimes)$ ein Tensorprodukt von $V_1, \ldots, V_N$. Weiter seien die Familien B_{V_k} für $k = 1, \ldots, N$ Basen von $V_1, \ldots, V_N$ und $B_{V_1 \otimes \cdots \otimes V_N}$ die induzierte Basis von $\bigotimes_{k=1}^N V_k$. Schließlich sei $t \in \bigotimes_{k=1}^N V_k$ ein Tensor und $A \in K^{n_1 \times \cdots \times n_N}$ seine Komponentenhypermatrix bzgl. der Basis $B_{V_1 \otimes \cdots \otimes V_N}$, also $A = \Phi^{-1}_{B_{V_1 \otimes \cdots \otimes V_N}}(t)$. Dann gilt:

$$\operatorname{Rang}(t) = \operatorname{Rang}(A).$$

Wir könnten jetzt wie in § 23.7 das Transformationsverhalten beliebiger Tensoren in $\bigotimes_{k=1}^N V_k$ bei einem Wechsel der Basen B_{V_k} für V_k für $k = 1, \ldots, N$ untersuchen. Wir werden das aber auf § 25 verschieben und dort nur den praktisch besonders relevanten Fall betrachten, dass jeder Faktor V_k entweder derselbe Vektorraum V ist oder sein Dualraum V^*. Vorher können wir jedoch wie in § 23.8 noch untersuchen, wie flexibel Tensoren als lineare Abbildungen genutzt werden können.

Zur Einstimmung betrachten wir ein Tensorprodukt von drei Vektorräumen U, V, W über demselben Körper. Ein Tensor in $U \otimes V \otimes W$ besitzt drei Eingänge, die Vektoren aus den jeweiligen Dualräumen U^*, V^*, W^* akzeptieren. Wird ein Eingang mit einem solchen Vektor belegt, so wird der Vektor vom Tensor „konsumiert". Der entsprechende Faktor aus dem Tensorproduktraum $U \otimes V \otimes W$ wird dabei „verbraucht" und verschwindet. Dieser Prozess kann auch mit mehreren Eingängen gleichzeitig geschehen. Das Ergebnis ist ein Tensor im Tensorproduktraum der nicht belegten Eingänge.

Durch die Belegung bzw. Nicht-Belegung der drei Eingänge gibt es acht verschiedene Möglichkeiten, einen Tensor in $U \otimes V \otimes W$ als lineare Abbildung zu nutzen. Wir sortieren diese Möglichkeiten nach der Auswahl und Anzahl der belegten Eingänge:

[27] Das ändert aber nichts an der Tatsache, dass die Bestimmung des Ranges für $N \geqslant 3$ schwierig ist, siehe Bemerkung 24.13.

$$\mathrm{Homo}(K, U \otimes V \otimes W)$$

$$\mathrm{Homo}(U^*, V \otimes W) \qquad \mathrm{Homo}(V^*, U \otimes W) \qquad \mathrm{Homo}(W^*, U \otimes V)$$

$$\mathrm{Homo}(U^* \otimes V^*, W) \qquad \mathrm{Homo}(U^* \otimes W^*, V) \qquad \mathrm{Homo}(V^* \otimes W^*, U)$$

$$\mathrm{Homo}(U^* \otimes V^* \otimes W^*, K)$$

Wir geben nun einige Beispiele dafür, wie der Isomorphismus jeweils aussieht, der die jeweilige Interpretation bzw. Nutzung eines Tensors in $U \otimes V \otimes W$ als lineare Abbildung realisiert. Insbesondere geben wir an, wie wir die Bilder der jeweiligen linearen Abbildung mit Hilfe der Komponentenhypermatrix bestimmen können. Kurz gesagt wählen wir diejenigen Achsen aus, die als Eingänge (Definitionsraum) fungieren sollen, und **kontrahieren** (lateinisch: contrahere: zusammenziehen) sie mit den Komponentenvektoren der einzusetzenden Elementartensoren.

Wir erinnern dazu an die Verwendung von Darstellungsmatrizen A linearer Abbildungen. Die Matrix-Vektor-Multiplikation $y = A\,x$ mit dem Koordinatenvektor x ergibt den Koordinatenvektor y des Bildes. Durch die Matrix-Vektor-Multiplikation wird die 2-achsige Matrix A entlang der zweiten Achse (den Spalten) mit den Koeffizienten des Koordinatenvektors x linearkombiniert, also kontrahiert.

Beispiel 24.22 (Nutzung von Tensoren als lineare Abbildung, vgl. Abbildung 23.1)**.**

Es seien U, V, W endlich-dimensionale Vektorräume über demselben Körper K mit Basen B_U, B_V bzw. B_W und zugehörigen dualen Basen B_{U^*}, B_{V^*} bzw. B_{W^*}. Wir betrachten einen Tensor in $U \otimes V \otimes W$ und dessen (3-achsige) Komponentenhypermatrix A bzgl. der Tensorproduktbasis $B_{U\otimes V\otimes W}$.

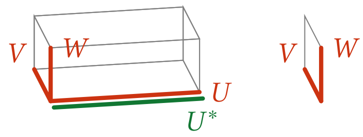
(i) Wird kein Eingang belegt (oberste Zeile im obigen Diagramm), so können wir den Tensor dennoch als lineare Abbildung vom Körper K in den Tensorproduktraum $U \otimes V \otimes W$ auffassen (vgl. Bemerkung 20.5):

$$\begin{aligned} U \otimes V \otimes W &\to \mathrm{Homo}(K, U \otimes V \otimes W) \\ u \otimes v \otimes w &\mapsto \cdot\,(u \otimes v \otimes w). \end{aligned}$$

Für $\alpha \in K$ besitzt das Bild $\alpha\,(u \otimes v \otimes w)$ die 3-achsige Komponentenhypermatrix $\alpha\,A$.

(ii) Wird nur der erste Eingang belegt (zweite Zeile links), so agiert der Tensor als lineare Abbildung von U^* nach $V \otimes W$:

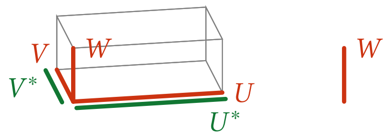
$$\begin{aligned} U \otimes V \otimes W &\to \operatorname{Homo}(U^*, V \otimes W) \\ u \otimes v \otimes w &\mapsto \langle \cdot\,, u \rangle\, (v \otimes w). \end{aligned}$$



Enthält ξ die Koordinaten eines Vektors $u^* \in U^*$ bzgl. der Basis B_{U^*}, dann besitzt das Bild $\langle u^*, u \rangle\, (v \otimes w)$ von u^* die 2-achsige Komponentenhypermatrix $\sum_i \xi_i\, a_{i\bullet\bullet}$.

(iii) Werden die ersten beiden Eingänge belegt (dritte Zeile links), so nimmt der Tensor die Funktion als lineare Abbildung von $U^* \otimes V^*$ in W ein:

$$\begin{aligned} U \otimes V \otimes W &\to \operatorname{Homo}(U^* \otimes V^*, W) \\ u \otimes v \otimes w &\mapsto \langle \cdot\,, u \otimes v \rangle\, w. \end{aligned}$$



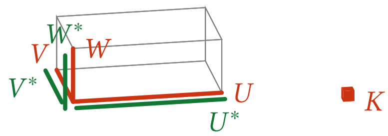
Der Ausdruck $\langle \cdot\,, u \otimes v \rangle \in \operatorname{Homo}(U^* \otimes V^*, K)$ ist dabei durch die folgende bilineare Abbildungsvorschrift auf Elementartensoren $v^* \otimes w^* \in V^* \otimes W^*$ eindeutig bestimmt:

$$\langle \cdot\,, u \otimes v \rangle = \big[u^* \otimes v^* \mapsto \langle u^*, u \rangle\, \langle v^*, v \rangle \big].$$

Enthalten ξ und η die Koordinaten von Vektoren $u^* \in U^*$ bzw. $v^* \in V^*$ bzgl. der Basis B_{U^*} bzw. B_{V^*}, dann besitzt das Bild $\langle u^*, u \rangle\, \langle v^*, v \rangle\, w$ des Elementartensors $u^* \otimes v^*$ die 1-achsige Komponentenhypermatrix $\sum_i \sum_j \xi_i\, \eta_j\, a_{ij\bullet}$.

(iv) Werden schließlich alle drei Eingänge belegt (untere Zeile), so agiert der Tensor als lineare Abbildung von $U^* \otimes V^* \otimes W^*$ in K:

$$\begin{aligned} U \otimes V \otimes W &\to \operatorname{Homo}(U^* \otimes V^* \otimes W^*, K) \\ u \otimes v \otimes w &\mapsto \langle \cdot\,, u \otimes v \otimes w \rangle. \end{aligned}$$



Enthalten nun ξ, η und μ die Koordinaten von Vektoren $u^* \in U^*$, $v^* \in V^*$ bzw. $w^* \in W^*$ bzgl. der Basis B_{U^*}, B_{V^*} bzw. B_{W^*}, dann besitzt das Bild $\langle u^*, u \rangle\, \langle v^*, v \rangle\, \langle w^*, w \rangle$ des Elementartensors $u^* \otimes v^* \otimes w^*$ die 0-achsige Komponentenhypermatrix $\sum_i \sum_j \sum_k \xi_i\, \eta_j\, \mu_k\, a_{ijk} \in K$. △

Beachte: Alle in Beispiel 24.22 genannten Abbildungsvorschriften sind 3-linear, sodass sie tatsächlich eindeutig lineare Abbildungen auf $U \otimes V \otimes W$ festlegen.

Die in Beispiel 24.22 gezeigten Operationen bezeichnet man als das **Einsetzen von Vektoren in einen Tensor** (englisch: inserting vectors into a tensor) oder auch als **Kontraktion** (englisch: contraction) **eines Tensors mit einem oder mehreren Vektoren**. Auf Ebene der Hypermatrix von Komponenten spricht man auch von der **Tensor-Vektor-Multiplikation** (englisch: tensor-vector multiplication) oder genauer: **Hypermatrix-Vektor-Multiplikation** (englisch: hypermatrix-vector multiplication).

Wir werden für das Einsetzen von Vektoren in Tensoren in Zukunft gelegentlich die Notation $t(\cdot, \cdots, \cdot)$ verwenden. Beispielsweise ist für einen Elementartensor $t = u \otimes v \otimes w \in U \otimes V \otimes W$ das Einsetzen von zwei Vektoren $u^* \in U^*$ und $v^* \in V^*$ in die ersten beiden Eingänge von t

erklärt durch

$$t(u^*, v^*, \cdot) \coloneqq \langle u^*, u\rangle \, \langle v^*, v\rangle \, w,$$

siehe Beispiel 24.22. Diese Schreibweise vermeidet, dass wir für jede Verwendungsart eines Tensors einen neuen Namen für die zugehörige lineare Abbildung einführen müssen.

Die obigen Vorüberlegungen führen zu der folgenden allgemeinen Aussage:

Satz 24.23 (Tensoren „sind" lineare Abbildungen, vgl. Satz 23.40)**.**
Es seien $N \in \mathbb{N}_0$ und $V_1, \ldots, V_N$ endlich-dimensionale Vektorräume über demselben Körper K. Weiter sei $B \subseteq [\![1, N]\!]$.[28] Weiter sei $s \in [\![0, N]\!]$ und π eine Permutation der Menge $[\![1, N]\!]$ mit der Eigenschaft $\pi(1) < \cdots < \pi(s)$ und $\pi(s+1) < \cdots < \pi(N)$.[29]. Dann gilt:

(*i*) Es besteht der folgende kanonische Isomorphismus, der durch seine Bilder auf den Elementartensoren eindeutig bestimmt ist:[30]

$$\begin{aligned} \bigotimes_{k=1}^{N} V_k &\to \operatorname{Homo}\Big(\bigotimes_{k=1}^{s} V_{\pi(k)}^*, \bigotimes_{k=s+1}^{N} V_{\pi(k)}\Big) \\ v_1 \otimes \cdots \otimes v_N &\mapsto \langle \cdot\,, v_{\pi(1)} \otimes \cdots \otimes v_{\pi(s)}\rangle \, (v_{\pi(s+1)} \otimes \cdots \otimes v_{\pi(N)}). \end{aligned} \tag{24.13}$$

(*ii*) Es seien B_{V_k} für $k = 1, \ldots, N$ Basen von $V_1, \ldots, V_N$ und $B_{V_k}^*$ die zugehörigen dualen Basen von $V_1^*, \ldots, V_N^*$. Weiter seien A die Komponentenhypermatrix eines Tensors $t \in \bigotimes_{k=1}^N V_k$ bzgl. der Tensorproduktbasis $B_{V_1 \otimes \cdots \otimes V_N}$ und $\xi_{\pi(k)}$ der Koordinatenvektor eines Vektors aus $V_{\pi(k)}^*$ bzgl. der Basis $B_{V_{\pi(k)}^*}$ für $k = 1, \ldots, s$. Dann ist das Bild des Elementartensors $v_{\pi(1)}^* \otimes \cdots \otimes v_{\pi(s)}^*$ unter der linearen Abbildung auf der rechten Seite von (24.13) gegeben durch die Komponentenhypermatrix[31]

$$\Big(\sum_{i_{\pi(1)}} \cdots \sum_{i_{\pi(s)}} \xi_{\pi(1)} \cdots \xi_{\pi(s)} \, a_{i_1,\ldots,i_N}\Big)_{(i_{\pi(s+1)},\ldots,i_{\pi(N)})}. \tag{24.14}$$

Wir könnten uns jetzt auch noch fragen, wie die Darstellungsmatrix der einem Tensor zugeordneten linearen Abbildung $\bigotimes_{k=1}^s V_{\pi(k)}^* \to \bigotimes_{k=s+1}^N V_{\pi(k)}$ aus (24.13) aussieht. Diese wird (vgl. Satz 23.40) aus den Einträgen der Komponententypermatrix des Tensors bestehen, wobei die Eingangsachsen (Spaltenindizes) in der Reihenfolge $\pi(1), \ldots, \pi(s)$ abgezählt werden und die Ausgangsachsen (Zeilenindizes) in der Reihenfolge $\pi(s+1), \ldots, \pi(N)$. Allerdings ist es praktischer und natürlicher, die Darstellungsmatrix nicht zu verwenden und bei der Komponentenhypermatrix zu bleiben, die dieselben Informationen enthält und keiner Abzählung bedarf.

Insbesondere haben die Achsen einer Komponentenhypermatrix zwar eine bestimmte Reihenfolge, aber keine bestimmte grafische Anordnung. Daher gibt es bei den Koordinatenvektoren,

[28] B steht für die Indizes der mit Vektoren belegten Eingänge (Achsen). In Beispiel 24.22 (*iii*) ist $B = (1, 2)$.

[29] Die Indizes $\pi(1), \ldots, \pi(s)$ bezeichnen die belegten Eingänge (sortiert in aufsteigender Reihenfolge), während die übrigen Indizes $\pi(s+1), \ldots, \pi N$ die unbelegten Eingänge (in aufsteigender Reihenfolge) angeben. In Beispiel 24.22 (*iii*) haben wir beispielweise $\pi = \mathrm{id}$ auf $\{1, 2, 3\}$ und $s = 2$.

[30] Zur Erinnerung: Ein leeres Tensorprodukt kann als der Körper K aufgefasst werden (Beispiel 24.9).

[31] Aus Gründen der Lesbarkeit geben wir hier die Laufbereiche der Indizes $i_{\pi(k)}$ für $k = 1, \ldots, s$ nicht an. Diese ergeben sich aus den Dimensionen der Vektorräume $V_{\pi(k)}$.

die an den Eingängen anliegen können, ebenfalls keine Anordnung; es gibt also kein Konzept von Zeilen- und Spaltenvektoren.

Abschließend geben zur Übung noch einige lineare Abbildungen zwischen Vektorräumen von Matrizen in Form von Tensoren an.

Beispiel 24.24 (Tensordarstellung von Abbildungen zwischen Vektorräumen von Matrizen)**.**

(*i*) Wir betrachten zunächst die identische Abbildung $\mathrm{id}\colon K^{n\times m} \to K^{n\times m}$ mit $n, m \in \mathbb{N}_0$. Wir identifizieren $K^{n\times m}$ dabei mit $K^n \otimes K^m$ über den von den Standardbasen induzierten Isomorphismus (siehe Beispiel 23.35). Wir sind also interessiert an der identischen Abbildung $\mathrm{id}\colon K^n \otimes K^m \to K^n \otimes K^m$. Diese können wir nach Satz 24.23 als einen Tensor in $(K^n)^* \otimes (K^m)^* \otimes K^n \otimes K^m$ auffassen, von dem dann die ersten beiden Eingänge belegt werden.

Wir betrachten die Komponentenhypermatrix A dieses Tensors bzgl. der Tensorproduktbasis $B_{(K^n)^*\otimes(K^m)^*\otimes K^n\otimes K^m}$, wobei in K^n und K^m die Standardbasen verwendet werden. Um die Einträge $a_{ijk\ell}$ zu ermitteln, setzen wir die Elemente der dualen Basis als „Komponentenermittler“ ein, siehe (24.12). Mit anderen Worten: Wir setzen $e_i \otimes e_j$ (was der Standardbasismatrix E_{ij} entspricht) in die identische Abbildung ein und ermitteln die Komponenten des Bildes mit Hilfe der Elemente der dualen Basis $\pi_k \boxtimes \pi_\ell$. Es gilt $\langle \pi_k \boxtimes \pi_\ell\, , e_i \otimes e_j\rangle = 1$ genau dann, wenn $i = k$ und $j = \ell$ ist. Andernfalls ergibt sich der Wert 0. Die gesuchte Komponentenhypermatrix hat also die Einträge

$$a_{ijk\ell} = \delta_{ik}\, \delta_{j\ell}$$

mit $i, k \in [\![1, n]\!]$ und $j, \ell \in [\![1, m]\!]$. (**Quizfrage 24.2:** Wieviele Einträge sind also gleich 1?)

(*ii*) Die Transpositionsabbildung $K^{n\times m} \ni X \mapsto X^\intercal \to K^{m\times n}$ können wir als einen Tensor in $(K^n)^*\otimes(K^m)^*\otimes K^m\otimes K^n$ auffassen, von dem die ersten beiden Eingänge belegt werden. Mit den gleichen Überlegungen wie oben erhalten wir, dass die Komponentenhypermatrix A dieses Tensors die Darstellung

$$a_{ijk\ell} = \delta_{i\ell}\, \delta_{jk}$$

mit $i, \ell \in [\![1, n]\!]$ und $j, k \in [\![1, m]\!]$ besitzt.

(*iii*) Die Matrix-Matrix-Multiplikation $K^{n\times m} \times K^{m\times \ell} \ni (X, Y) \mapsto XY \in K^{n\times \ell}$ mit $n, m, \ell \in \mathbb{N}_0$ können wir als einen Tensor in $(K^n)^* \otimes (K^m)^* \otimes (K^m)^* \otimes (K^\ell)^* \otimes K^n \otimes K^\ell$ auffassen, von dem die ersten vier Eingänge belegt werden. Das Bild der Matrix-Matrix-Multiplikationsabbildung der Basismatrizen E_{ab} und E_{cd} ist die Matrix $\delta_{bc}\, E_{ad}$. Die „Koordinatenermittler“ $\pi_e \boxtimes \pi_f$ liefern — angewendet auf $e_a \otimes e_d$ — genau dann den Wert 1, wenn $e = a$ und $f = d$ gilt. Daraus ergibt sich die Darstellung

$$a_{abcdef} = \delta_{bc}\, \delta_{ae}\, \delta_{df}$$

der Komponentenhypermatrix des Matrix-Matrix-Multiplikationstensors mit $a, e \in [\![1, n]\!]$, $b, c \in [\![1, m]\!]$ und $d, f \in [\![1, \ell]\!]$. △

Ende der Vorlesung 8

§ 25 Tensoren über einem Vektorraum

In diesem Abschnitt betrachten wir eine für die Praxis besonders relevante Klasse von Tensorprodukten noch genauer. Bei diesen treten als Faktoren anstelle verschiedener Vektorräume $V_1, \dots, V_N$ nur ein einziger Vektorraum V und sein Dualraum V^* auf.

Definition 25.1 (Tensoren über einem Vektorraum).
Es seien V ein Vektorraum über einem Körper K und V^* sein Dualraum. Für $(r,s) \in \mathbb{N}_0 \times \mathbb{N}_0$ heißt

$$\mathcal{T}^{(r,s)}(V) \coloneqq \underbrace{V \otimes \cdots \otimes V}_{r\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s\text{-mal}} \tag{25.1}$$

der **Tensorproduktraum vom Typ** (r,s) (englisch: tensor product space of type (r,s)) **über dem Vektorraum** V.[32] Die Elemente dieses Tensorproduktraumes heißen **Tensoren vom Typ** (r,s) über dem Vektorraum V.[33] △

Wir werden uns im Folgenden auf den Fall beschränken, dass V endlich-dimensional ist. Es gelte also $\dim(V) = \dim(V^*) = n \in \mathbb{N}_0$.

Wir führen nun die für den Umgang mit solchen Tensoren übliche Notation ein:[34]

- Die Vektoren einer Basis des „primalen" Raumes V werden mit unteren Indizes nummeriert. Wir bezeichnen diese typischerweise mit $B_V = (v_1, \dots, v_n)$.

  Die Koordinaten von Vektoren $v \in V$ bzgl. dieser Basis werden mit oberen Indizes notiert, also $v = \sum_{i=1}^n x^i\, v_i$.
- Die Covektoren der zu B_V dualen Basis von V^* werden mit oberen Indizes nummeriert. Wir bezeichnen diese typischerweise mit $B_{V^*} = (v^1, \dots, v^n)$. Es gilt also $\langle v^i, v_j \rangle = \delta_{ij}$.

  Die Koordinaten von Covektoren in V^* bzgl. dieser Basis werden mit unteren Indizes notiert, also $v^* = \sum_{i=1}^n x_i\, v^i$.

Jeder Tensor in $\mathcal{T}^{(r,s)}(V)$ ist eine Linearkombination der Elemente der Tensorproduktbasis $B_{V\otimes\cdots\otimes V\otimes V^*\cdots\otimes V^*}$ (Lemma 24.11)

$$v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes v^{j_1} \otimes \cdots \otimes v^{j_s} \quad \text{mit } 1 \leqslant i_1, \dots, i_r \leqslant n \text{ und } 1 \leqslant j_1, \dots, j_s \leqslant n. \tag{25.2}$$

Er kann also dargestellt werden in der Form

$$t = \sum_{i_1=1}^n \cdots \sum_{i_r=1}^n \sum_{j_1=1}^n \cdots \sum_{j_s=1}^n a^{i_1,\dots,i_r}_{j_1,\dots,j_s}\, v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes v^{j_1} \otimes \cdots \otimes v^{j_s}, \tag{25.3}$$

siehe Satz 24.18. Die Hypermatrix der **Komponenten** besitzt $r+s$ Achsen und damit n^{r+s} Einträge. Im Einklang mit der oben eingeführten Konvention werden in der Hypermatrix der

[32] Die Notation für diese Tensorprodukträume ist in der Literatur leider nicht einheitlich. Manchmal ist die Bedeutung von r und s gerade vertauscht.

[33] Historisch spricht man auch von r-fach **kontravarianten** und s-fach **kovarianten** Tensoren, vgl. Bemerkung 20.19 und Satz 25.4.

[34] Wir kennzeichnen Vektoren aus dem primalen Raum V und dem dualen Raum V^* hier wieder farbig.

Komponenten die ersten r Indizes $(i_1, \ldots, i_r)$, die zu den primalen Faktoren gehören, als obere Indizes notiert, während die letzten s Indizes $(j_1, \ldots, j_s)$, die zu den dualen Faktoren gehören, als untere Indizes geschrieben werden.

Es ist außerdem üblich, die Summenzeichen in (25.3) nicht zu notieren, sondern sich diese implizit dazuzudenken. Das nennt sich die **Einsteinsche Summenkonvention** (englisch: Einstein summation convention): „Über jeden Index, der in einem Ausdruck zweifach (und zwar notwendigerweise einmal oben und einmal unten) auftritt, wird summiert.“[35] Wir werden die Einsteinsche Summenkonvention aber nicht verwenden.

Nach Satz 24.18 ist die Zuordnung

$$\mathcal{T}^{(r,s)}(V) \ni t \mapsto \Phi^{-1}_{B_{V \otimes \cdots \otimes V \otimes V^* \cdots \otimes V^*}}(t) = A = \left(a^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}\right) \in K^{n \times \cdots \times n}$$

eines Tensors t zu seiner Komponentenhypermatrix A bzgl. der durch $(v_1, \ldots, v_n)$ und die dazugehörige duale Basis festgelegten Tensorproduktbasis von $\mathcal{T}^{(r,s)}(V)$ ein Isomorphismus von Vektorräumen.

Bemerkung 25.2 (Tensoren und ihre Komponenten)**.**

Oft begegnet man der Aussage, das Zahlenschema der Komponenten $a^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$ *sei* der eigentliche Tensor t.[36] Dieser Sichtweise schließen wir uns hier nicht an. Jedoch kann man — wegen des obigen Isomorphismus! — mit einem Tensor „arbeiten“ (z. B. den Typ und die Anzahl der Achsen ablesen oder den Rang bestimmen, Vektoren und Covektoren in den Tensor einsetzen etc.), indem man die entsprechenden Aktionen an der Komponentenhypermatrix ausführt. △

Aus Satz 24.23 folgt, dass wir die Tensoren vom Typ (r, s) über einem endlich-dimensionalen Vektorraum V durch Einsetzen einiger oder aller Argumente wie folgt verwenden können:

$$\begin{aligned}
&\mathcal{T}^{(0,0)}(V) = K && \cong \operatorname{Homo}(K, K) && && && (25.4a)\\
&\mathcal{T}^{(0,1)}(V) = V^* && \cong \operatorname{Homo}(K, V^*) && \cong \operatorname{Homo}(V, K) && && (25.4b)\\
&\mathcal{T}^{(1,0)}(V) = V && \cong \operatorname{Homo}(K, V) && \cong \operatorname{Homo}(V^*, K) && && (25.4c)\\
&\mathcal{T}^{(0,2)}(V) = V^* \otimes V^* && \cong \operatorname{Homo}(K, V^* \otimes V^*) && \cong \operatorname{Homo}(V, V^*) && \cong \operatorname{Homo}(V \otimes V, K) && (25.4d)\\
&\mathcal{T}^{(1,1)}(V) = V \otimes V^* && \cong \operatorname{Homo}(K, V \otimes V^*) && \cong \operatorname{Homo}(V, V) && \cong \operatorname{Homo}(V^* \otimes V, K) && (25.4e)\\
&\mathcal{T}^{(2,0)}(V) = V \otimes V && \cong \operatorname{Homo}(K, V \otimes V) && \cong \operatorname{Homo}(V^*, V) && \cong \operatorname{Homo}(V^* \otimes V^*, K) && (25.4f)
\end{aligned}$$

und so weiter. Dabei haben wir den Dualraum V^{**} von V^* mit V identifiziert (Satz 22.2), also die endliche Dimension von V ausgenutzt. Insbesondere gilt also:

[35] Ein häufigeres als zweifaches Auftreten eines Index ist nicht zulässig.

[36] Diese Aussage steht auf einer Stufe mit „Der Koordinatenvektor eines Vektors *ist* der Vektor.“ oder „Die Darstellungsmatrix einer linearen Abbildung *ist* die Abbildung“.

Tensoren vom Typ ...	sind
$(0,0)$	Skalare in K
$(1,0)$	Vektoren in V
$(0,1)$	Covektoren in V^*
$(1,1)$	Endomorphismen in $\operatorname{Endo}(V)$
$(0,2)$	Bilinearformen in $\operatorname{Bil}(V,V;K)$

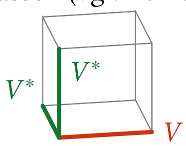
Anhand eines Tensors vom Typ $(1,2)$ illustrieren wir die verschiedenen Nutzungsformen und deren Realisierung anhand der Komponentenhypermatrix im folgenden Beispiel.

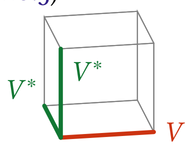
Beispiel 25.3 (Nutzung von Tensoren über einem Vektorraum als lineare Abbildung, vgl. Beispiel 24.22)**.**

Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K mit der Basis B_V und V^* sein Dualraum mit der zu B_V dualen Basis B_{V^*}. Wir betrachten einen Tensor in $\mathcal{T}^{(1,2)}(V)$ und dessen 3-achsige Komponentenhypermatrix A bzgl. der Tensorproduktbasis $B_{V\otimes V^*\otimes V^*}$. Die Vektoren $v_1, v_2, v_3 \in V$ und $v^1, v^2, v^3 \in V^*$ sind hier irgendwelche Vektoren (nicht notwendigerweise aus der jeweiligen Basis).

(i) Wird kein Eingang belegt, so können wir den Tensor als lineare Abbildung vom Körper K in den Tensorproduktraum $V \otimes V^* \otimes V^*$ auffassen (vgl. Bemerkung 20.5):

$$\begin{aligned} V \otimes V^* \otimes V^* &\to \operatorname{Homo}(K, V \otimes V^* \otimes V^*) \\ v_1 \otimes v^2 \otimes v^3 &\mapsto \cdot\,(v_1 \otimes v^2 \otimes v^3). \end{aligned}$$

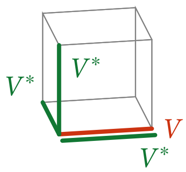


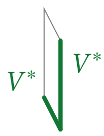


Für $\alpha \in K$ besitzt das Bild $\alpha\,(v_1 \otimes v^2 \otimes v^3)$ die 3-achsige Komponentenhypermatrix $\alpha\,A$.

(ii) Wird nur der erste Eingang belegt, so agiert der Tensor als lineare Abbildung von V^* nach $V^* \otimes V^*$:

$$\begin{aligned} V \otimes V^* \otimes V^* &\to \operatorname{Homo}(V^*, V^* \otimes V^*) \\ v_1 \otimes v^2 \otimes v^3 &\mapsto \langle \cdot\,, v_1\rangle\,(v^2 \otimes v^3). \end{aligned}$$

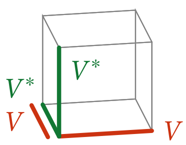


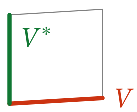


Sind $(x_1, \dots, x_n)$ die Koordinaten eines Vektors $v^1 \in V^*$ bzgl. der Basis B_{V^*}, dann besitzt das Bild $\langle v^1, v_1\rangle\,(v^2 \otimes v^3)$ von v^1 die 2-achsige Komponentenhypermatrix $\sum_i x_i\, a_{i\bullet\bullet}$.

(iii) Wird dagegen nur der zweite Eingang belegt, so agiert der Tensor als lineare Abbildung von V nach $V \otimes V^*$:

$$\begin{aligned} V \otimes V^* \otimes V^* &\to \operatorname{Homo}(V, V \otimes V^*) \\ v_1 \otimes v^2 \otimes v^3 &\mapsto \langle v^2, \cdot\rangle\,(v_1 \otimes v^3). \end{aligned}$$





Sind $(y^1, \dots, y^n)$ die Koordinaten eines Vektors $v_2 \in V$ bzgl. der Basis B_V, dann besitzt das Bild $\langle v^2, v_2\rangle\,(v_1 \otimes v^3)$ von v_2 die 2-achsige Komponentenhypermatrix $\sum_j y^j\, a_{\bullet j \bullet}$.

(*iv*) Werden die letzten beiden Eingänge belegt, so nimmt der Tensor die Funktion als lineare Abbildung von $V \otimes V$ in V ein:

$$\begin{aligned} V \otimes V^* \otimes V^* &\longrightarrow \operatorname{Homo}(V \otimes V, V) \\ v_1 \otimes v^2 \otimes v^3 &\mapsto \langle v^2 \otimes v^3, \cdot \rangle \, v_1. \end{aligned}$$

Der Ausdruck $\langle v^2 \otimes v^3, \cdot \rangle \in \operatorname{Homo}(V \otimes V, V)$ ist dabei durch die folgende bilineare Abbildungsvorschrift auf Elementartensoren $v_2 \otimes v_3 \in V \otimes V$ eindeutig bestimmt:

$$\langle v^2 \otimes v^3, \cdot \rangle = \big[v_2 \otimes v_3 \mapsto \langle v^2, v_2 \rangle \, \langle v^3, v_3 \rangle \big].$$

Sind $(y^1, \ldots, y^n)$ und $(z^1, \ldots, z^n)$ die Koordinaten von Vektoren $v_2, v_3 \in V$ bzgl. der Basis B_V, dann besitzt das Bild $\langle v^2, v_2 \rangle \, \langle v^3, v_3 \rangle \, v_1$ des Elementartensors $v_2 \otimes v_3$ die 1-achsige Komponentenhypermatrix $\sum_j \sum_k y^j \, z^k \, a_{\bullet jk}$. △

Wie angekündigt untersuchen wir jetzt noch das Transformationsverhalten der Komponentenhypermatrix eines Tensors vom Typ (r, s) über einem endlich-dimensionalen Vektorraum V.

Satz 25.4 (Transformation der Komponentenhypermatrix eines Tensors vom Typ (r, s) beim Wechsel der Basen, vgl. (23.23)).
Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Weiter seien B_V und $\widehat{B}_V$ Basen von V und B_{V^*} und $\widehat{B}_{V^*}$ die dazu gehörigen dualen Basen. Schließlich sei $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ die Transformationsmatrix für den Wechsel von $\widehat{B}_V$ zu B_V, siehe (19.11). Dann gilt für die Komponentenhypermatrizen $A = \Phi^{-1}_{B_{V \otimes \cdots \otimes V \otimes V^* \cdots \otimes V^*}}(t)$ und $\widehat{A} = \Phi^{-1}_{\widehat{B}_{V \otimes \cdots \otimes V \otimes V^* \cdots \otimes V^*}}(t)$ eines Tensors t vom Typ (r, s) über V die folgende Transformationsvorschrift:

$$\widehat{a}^{i_1, \ldots, i_r}_{j_1, \ldots, j_s} = \sum_{k_1=1}^n \cdots \sum_{k_r=1}^n \sum_{\ell_1=1}^n \cdots \sum_{\ell_s=1}^n \underbrace{(T^{-1})_{i_1 k_1} \cdots (T^{-1})_{i_r k_r}}_{\text{kontravariante Achsen}} \, \underbrace{(T^\intercal)_{j_1 \ell_1} \cdots (T^\intercal)_{j_s \ell_s}}_{\text{kovariante Achsen}} \, a^{k_1, \ldots, k_r}_{\ell_1, \ldots, \ell_s}. \tag{25.5}$$

Beweis. Es sei $T \coloneqq \mathcal{T}_{B_V \leftarrow \widehat{B}_V} \in K^{n \times n}$ die Transformationsmatrix von $\widehat{B}_V$ zu B_V, es gilt also nach (19.12)

$$\widehat{v}_j = \sum_{i=1}^m (T)_{ij} \, v_i \quad \text{bzw.} \quad v_i = \sum_{j=1}^n (T^{-1})_{ji} \, \widehat{v}_j.$$

Für die duale Basis gilt wegen (20.10)

$$\widehat{v}^j = \sum_{i=1}^n (T^{-\intercal})_{ij} \, v^i \quad \text{bzw.} \quad v^i = \sum_{j=1}^n (T^\intercal)_{ji} \, \widehat{v}^j \quad \text{bzw.} \quad v^j = \sum_{i=1}^n (T^\intercal)_{ij} \, \widehat{v}^i.$$

Aus Gründen der Übersichtlichkeit betrachten wir hier nur den Fall $r = s = 1$. Der Tensor t besitzt also die Darstellung

$$t = \sum_{i=1}^n \sum_{j=1}^n a^i_j \, v_i \otimes v^j$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a^i_j \Big(\sum_{k=1}^{n} (T^{-1})_{ki}\, \widehat{v}_k \Big) \otimes \Big(\sum_{\ell=1}^{n} (T^\intercal)_{\ell j}\, \widehat{v}^\ell \Big) && \text{durch Einsetzen der Transformationsvorschriften} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{\ell=1}^{n} (T^{-1})_{ki}\, (T^\intercal)_{\ell j}\, a^i_j\, \widehat{v}_k \otimes \widehat{v}^\ell && \text{wegen der Bilinearität des Tensorprodukts} \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \Big[\sum_{i=1}^{n} \sum_{j=1}^{n} (T^{-1})_{ki}\, (T^\intercal)_{\ell j}\, a^i_j \Big]\, \widehat{v}_k \otimes \widehat{v}^\ell && \text{durch Umsortieren.}
\end{aligned}
$$

Das zeigt, dass die transformierte Hypermatrix der Komponenten von t bzgl der neuen Basen $\widehat{B}_V$ und $\widehat{B}_{V^*}$ die Gestalt

$$\widehat{a}^k_\ell = \sum_{i=1}^{n} \sum_{j=1}^{n} (T^{-1})_{ki}\, (T^\intercal)_{\ell j}\, a^i_j$$

besitzt oder aber — nach Umbenennung der Indizes $k \leftrightarrow i$ und $\ell \leftrightarrow j$ —

$$\widehat{a}^i_j = \sum_{k=1}^{n} \sum_{\ell=1}^{n} (T^{-1})_{ik}\, (T^\intercal)_{j\ell}\, a^k_\ell.$$

Die Verallgemeinerung auf beliebige $r, s \in \mathbb{N}_0$ erfolgt durch entsprechend häufige Anwendung der obigen Transformationen. □

Bemerkung 25.5 (zur Sprechweise bei der Transformation von Komponenten eines Tensors)**.**
Der Satz 25.4 macht eine Aussage dazu, wie sich die Hypermatrix der Komponenten eines Tensors $t \in \mathcal{T}^{(r,s)}(V)$ transformiert, wenn wir die Basis von V (und damit auch die dazu gehörige duale Basis) von V^* wechseln. Die weit verbreitete Sprechweise, Tensoren vom Typ (r, s) „transformieren sich r-fach kontraviant und s-fach kovariant", ist irreführend. Der Tensor t selbst bleibt vom Wechsel der Basen unberührt, es ändert sich lediglich seine basisabhängige Darstellung! △

§ 26 Symmetrische, schiefsymmetrische und alternierende Tensoren

Wir betrachten in diesem Abschnitt nur Tensoren vom Typ $(r, 0)$ mit $r \in \mathbb{N}_0$ über einem Vektorraum V. Dennoch werden auch weiterhin duale Größen auftauchen, sodass wir auch hier primale und duale Vektorräume und Vektoren farblich unterscheiden wollen.[37] Zur Vereinfachung der Notation schreiben wir für diesen Raum auch einfach

$$V^{\otimes r} \coloneqq \mathcal{T}^{(r,0)}(V) = \bigotimes_{k=1}^{r} V.$$

Wir interessieren uns dabei insbesondere für solche Tensoren, die gewisse Symmetrieeigenschaften erfüllen.

[37] Wir verwenden auch weiterhin die Konventionen aus § 25, sodass typische primale Vektoren v_j und typische duale Vektoren v^j heißen.

Wir werden uns wieder auf den Fall beschränken, dass V endlich-dimensional ist. Weil dann Tensoren in $V^{\otimes r}$ kanonisch mit Multilinearformen $V^* \times \cdots \times V^*$ identifiziert werden können (Satz 24.23), sprechen wir gleichzeitig auch immer über entsprechende Eigenschaften von Multilinearformen.

Wir erinnern uns an die symmetrische Gruppe vom Grad $r \in \mathbb{N}_0$, bestehend aus den Permutationen auf $[\![1, r]\!]$, also den bijektiven Abbildungen $[\![1, r]\!] \to [\![1, r]\!]$, siehe Definition 7.30.

Definition 26.1 (Permutation eines Tensors vom Typ $(r, 0)$)**.**
Es sei V ein Vektorraum über einem Körper K und $r \in \mathbb{N}_0$. Weiter sei $\sigma \in S_r$ eine Permutation auf $[\![1, r]\!]$. Dann heißt die durch

$$P_\sigma \colon \begin{cases} V^{\otimes r} \to V^{\otimes r} \\ v_1 \otimes \cdots \otimes v_r \mapsto v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(r)} \end{cases} \tag{26.1}$$

eindeutig definierte lineare Abbildung die **durch σ induzierte Permutationsabbildung** (englisch: induced permutation map) auf $V^{\otimes r}$. △

Beachte: In der Definition (26.1) wird die inverse Permutation σ^{-1} verwendet.

Lemma 26.2 (Permutationen eines Tensors vom Typ $(r, 0)$)**.**
Es sei V ein endlich-dimensionaler Vektorraum mit $\dim(V) = n \in \mathbb{N}_0$ über dem Körper K und $r \in \mathbb{N}_0$. Dann gilt:

(i) Für jedes $\sigma \in S_r$ definiert die Abbildung P_σ aus (26.1) einen Automorphismus von $V^{\otimes r}$.

(ii) Es gilt

$$P_{\sigma_1} \circ P_{\sigma_2} = P_{\sigma_1 \circ \sigma_2} \tag{26.2}$$

für alle $\sigma_1, \sigma_2 \in S_r$.
(Die Abbildung $\sigma \mapsto P_\sigma$ ist also ein Gruppenhomomorphismus von S_r in die Automorphismengruppe von $V^{\otimes r}$, siehe Definition 17.16.)

(iii) Ist B_V eine Basis von V, $t \in V^{\otimes r}$ ein Tensor und $A = \Phi^{-1}_{B_{V \otimes \cdots \otimes V}}(t)$ seine Komponentenhypermatrix bzgl. der Tensorproduktbasis $B_{V \otimes \cdots \otimes V}$, dann gilt für jedes $\sigma \in S_r$:

$$\left(a^{i_{\sigma^{-1}(1)}, \ldots, i_{\sigma^{-1}(r)}}\right)_{(i_1, \ldots, i_r) \in [\![1,n]\!] \times \cdots \times [\![1,n]\!]} = \Phi^{-1}_{B_{V \otimes \cdots \otimes V}}(P_\sigma(t)) \tag{26.3}$$

ist die Komponentenhypermatrix von $P_\sigma(t)$.
(Eine Permutation der Achsen eines Tensors entspricht also einer Permutation der Achsen der Komponentenhypermatrix.)

(iv) Betrachten wir den Tensor $t \in V^{\otimes r}$ als Multilinearform in $\operatorname{Mult}(V^*, \ldots, V^*; K)$, dann gilt für jedes $\sigma \in S_r$:

$$P_\sigma(t)(v^1, \ldots, v^r) = t(v^{\sigma(1)}, \ldots, v^{\sigma(r)}). \tag{26.4}$$

Beweis. □

Definition 26.3 (symmetrische, schiefsymmetrische und alternierende Tensoren, vgl. Definition 15.30).
Es sei V ein Vektorraum über einem Körper K. Weiter sei $r \in \mathbb{N}_0$.

(i) Ein Tensor $t \in V^{\otimes r}$ heißt **(total) symmetrisch** (englisch: (totally) symmetric tensor), wenn für jede Permutation $\sigma \in S_r$ gilt:

$$P_\sigma(t) = t. \tag{26.5}$$

Die Menge aller symmetrischen Tensoren in $V^{\otimes r}$ bezeichnen wir mit $V^{\otimes r}_{\text{sym}}$.

(ii) Ein Tensor $t \in V^{\otimes r}$ heißt **(total) schiefsymmetrisch** (englisch: (totally) skew-symmetric tensor), wenn für jede Permutation $\sigma \in S_r$ gilt:

$$P_\sigma(t) = \operatorname{sgn}(\sigma)\, t. \tag{26.6}$$

Die Menge aller schiefsymmetrischen Tensoren in $V^{\otimes r}$ bezeichnen wir mit $V^{\otimes r}_{\text{skew}}$.

(iii) Ein Tensor $t \in V^{\otimes r}$ heißt **(total) alternierend** (englisch: (totally) alternating tensor), wenn für alle $v^1, \ldots, v^r \in V^*$ gilt:

$$v^i = v^j \text{ für ein } i \neq j \quad \Rightarrow \quad t(v^1, \ldots, v^r) = 0. \tag{26.7}$$

Die Menge aller alternierenden Tensoren in $V^{\otimes r}$ bezeichnen wir mit $V^{\otimes r}_{\text{alt}}$. △

Beachte: Das Attribut **total** steht zur Unterscheidung von Tensoren mit partiellen Symmetrie-, Schiefsymmetrie- oder alternierenden Eigenschaften, die wir nicht betrachten. Wir werden das Attribut **total** daher in Zukunft weglassen.

Wir verwenden die Begriffe **symmetrisch**, **schiefsymmetrisch** und **alternierend** analog auch für Multilinearformen in $\operatorname{Mult}(V^* \times \cdots \times V^*; K)$. Die Symmetrie von bedeutet gerade, dass die Reihenfolge der eingesetzten r Argumente unerheblich ist. Die Schiefsymmetrie bedeutet dagegen, dass jedes Mal das Vorzeichen wechselt, wenn zwei Argumente vertauscht werden. Die alternierende Eigenschaft bedeutet schließlich, dass das Bild gleich 0 ist, sobald zwei Argumente identisch sind. Die alternierende Eigenschaft hat wegen der Multilinearität aber eine noch weitreichendere Bedeutung, wie die folgende Charakterisierung zeigt:

Lemma 26.4 (Alternierende Tensoren erkennen lineare Abhängigkeit).
Es sei V ein Vektorraum über einem Körper K und $r \in \mathbb{N}_0$. Weiter sei $t \in V^{\otimes r}$. Dann sind äquivalent:

(i) Der Tensor t ist alternierend.

(ii) Für linear abhängige Familien $(v^1, \ldots, v^r)$ von Vektoren in V^* gilt $t(v^1, \ldots, v^r) = 0$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Im Fall $r = 0$ ist nichts zu zeigen, denn es gibt keine linear abhängigen Familien von $r = 0$ Vektoren (Bemerkung 12.3). Im Fall $r = 1$ ist ebenfalls nichts zu zeigen, denn die lineare Abhängigkeit der Familie (v^1) bedeutet gerade $v^1 = 0$, und damit gilt $t(v^1) = t(0) = 0$.

Es sei also $r \geqslant 2$ und $(v^1, \dots, v^r)$ linear abhängig, dann gibt es einen Index $1 \leqslant i \leqslant r$ und Koeffizienten $\alpha_j \in K$, sodass

$$v^i = \sum_{\substack{j=1\\ j\neq i}}^{r} \alpha_j \, v^j$$

gilt (Lemma 12.5). Wir erhalten

$$\begin{aligned} t(v^1, \dots, v^r) &= t\Big(v^1, \dots, v^{i-1}, \sum_{\substack{j=1\\ j\neq i}}^{r} \alpha_j \, v^j, v^{i+1}, \dots, v^r\Big) \\ &= \sum_{\substack{j=1\\ j\neq i}}^{r} \alpha_j \, \underbrace{t\big(v^1, \dots, v^{i-1}, v^j, v^{i+1}, \dots, v^r\big)}_{=0} && \text{wegen der Multilinearität} \\ &= 0 && \text{wegen der alternierenden Eigenschaft.} \end{aligned}$$

Aussage $(ii) \Rightarrow$ Aussage (i): Im Fall $r = 0$ und $r = 1$ ist nichts zu zeigen, denn dann ist jeder Tensor alternierend. (**Quizfrage 26.1:** Klar?) Es sei also $r \geqslant 2$. Wir betrachten $(v^1, \dots, v^r)$, wobei zwei der Vektoren identisch sind. Dann ist $(v^1, \dots, v^r)$ linear abhängig, also gilt nach Voraussetzung $t(v^1, \dots, v^r) = 0$. Damit ist t alternierend. □

Lemma 26.5 (Zusammenhang zwischen schiefsymmetrischen und alternierenden Tensoren, vgl. Lemma 15.32).
Es sei V ein Vektorraum über einem Körper K. Weiter sei $r \in \mathbb{N}_0$. Dann gilt:

(i) Im Fall $r = 0$ und $r = 1$ gilt $V^{\otimes r}_{\text{sym}} = V^{\otimes r}_{\text{skew}} = V^{\otimes r}_{\text{alt}} = V^{\otimes r}$.
(Skalare und Vektoren gelten also sowohl als symmetrische als auch als schiefsymmetrische und alternierende Tensoren.)

(ii) $V^{\otimes r}_{\text{alt}} \subseteq V^{\otimes r}_{\text{skew}}$.

(iii) Im Fall $\operatorname{char}(K) \neq 2$ gilt sogar $V^{\otimes r}_{\text{alt}} = V^{\otimes r}_{\text{skew}}$.

(iv) Im Fall $\operatorname{char}(K) = 2$ gilt $V^{\otimes r}_{\text{sym}} = V^{\otimes r}_{\text{skew}}$.

(v) $V^{\otimes r}_{\text{sym}}$, $V^{\otimes r}_{\text{skew}}$ und $V^{\otimes r}_{\text{alt}}$ sind Unterräume von $V^{\otimes r}$.

Beweis. Aussage (i): Die einzige Permutation in S_0 ist die leere Permutation, und die einzige Permutation in S_1 ist die Identität. Daher sind die Bedingungen (26.5) und (26.6) im Fall $r = 0$ und $r = 1$ für jede Permutation in S_r und jeden Tensor in $V^{\otimes r}$ erfüllt. Außerdem ist im Fall $r = 0$ und $r = 1$ jeder Tensor alternierend, da es nicht möglich ist, die Prämisse von (26.7) zu erfüllen.

Aussage (ii): Wegen Aussage (i) müssen wir nur den Fall $r \geqslant 2$ betrachten. Es seien $v^1, \dots, v^r \in V^*$ und $i, j \in [\![1, r]\!]$ mit $i < j$. Wir setzen $v^1, \dots, v^r$ in den Tensor $t \in V^{\otimes r}_{\text{alt}}$ ein; an die Positionen i und j schreiben wir allerdings $v^i + v^j$:

$$\begin{aligned} 0 &= t(v^1, \dots, v^i + v^j, \dots, v^i + v^j, \dots, v^r) \quad \text{wegen der alternierenden Eigenschaft von } t \\ &= t(v^1, \dots, v^i, \dots, v^i, \dots, v^r) \\ &\quad + t(v^1, \dots, v^i, \dots, v^j, \dots, v^r) \end{aligned}$$

$$\begin{aligned}
&+ t(v^1, \dots, v^j, \dots, v^i, \dots, v^r) \\
&+ t(v^1, \dots, v^j, \dots, v^j, \dots, v^r) && \text{wegen der Multilinearität} \\
= {}& t(v^1, \dots, v^i, \dots, v^j, \dots, v^r) \\
&+ t(v^1, \dots, v^j, \dots, v^i, \dots, v^r) && \text{wegen der alternierenden Eigenschaft von } t.
\end{aligned}$$

Das heißt aber, dass das Vertauschen der Einträge an den Positionen i und j das Vorzeichen von $t(v^1, \dots, v^r)$ ändert. Für jede Transposition $\tau \coloneqq \tau(i, j)$ haben wir also

$$\begin{aligned}
t(v^1, \dots, v^r) &= t(v^1, \dots, v^i, \dots, v^j, \dots, v^r) \\
&= -t(v^1, \dots, v^j, \dots, v^i, \dots, v^r) \\
&= \operatorname{sgn}(\tau)\, t(v^{\tau(1)}, \dots, v^{\tau(i)}, \dots, v^{\tau(j)}, \dots, v^{\tau(r)}) \\
&= \operatorname{sgn}(\tau)\, P_\tau(t)(v^1, \dots, v^r)
\end{aligned}$$

für alle $v^1, \dots, v^r \in V^*$. Das heißt aber $P_\tau(t) = \operatorname{sgn}(\tau)\, t$ für jede Transposition τ.

Da jede Permutation $\sigma \in S_r$ als Produkt von Transpositionen geschrieben werden kann (Satz 7.33), folgt daraus mit Lemma 26.2 aber auch, dass $P_\sigma(t) = \operatorname{sgn}(\sigma)\, t$ für alle $\sigma \in S_r$ gilt. (**Quizfrage 26.2:** Klar?) Jede alternierende Permutation ist also schiefsymmetrisch.

Aussage (*iii*): Es seien $\operatorname{char}(K) \neq 2$ und $t \in V^{\otimes r}_{\text{skew}}$. Auch hier müssen wir wegen Aussage (*i*) nur den Fall $r \geqslant 2$ betrachten. Es seien $v^1, \dots, v^r \in V^*$, und es gelte $v^i = v^j$ für ein $i < j$. Dann ist

$$\begin{aligned}
&t(v^1, \dots, v^i, \dots, v^j, \dots, v^r) \\
&\quad = -t(v^1, \dots, v^j, \dots, v^i, \dots, v^r) && \text{wegen der Schiefsymmetrie von } t \\
&\quad = -t(v^1, \dots, v^i, \dots, v^j, \dots, v^r) && \text{wegen } v^i = v^j.
\end{aligned}$$

Das heißt aber $2\, t(v^1, \dots, v^r) = 0$, und wegen $\operatorname{char}(K) \neq 2$ folgt $t(v^1, \dots, v^r) = 0$. Damit ist $t \in V^{\otimes r}_{\text{alt}}$.

Aussage (*iv*): Im Fall $\operatorname{char}(K) = 2$ gilt $-1 = 1$, also ist (26.6) äquivalent zu (26.5).

Aussage (*v*): Nach der definierenden Bedingung (26.5) ist

$$V^{\otimes r}_{\text{sym}} = \operatorname{Kern}(P_\sigma - \operatorname{id}_{V^{\otimes r}}).$$

Aus Lemma 17.8 folgt damit die Unterraumeigenschaft von $V^{\otimes r}_{\text{sym}}$. Analog ist

$$V^{\otimes r}_{\text{skew}} = \operatorname{Kern}(P_\sigma - \operatorname{sgn}(\sigma)\, \operatorname{id}_{V^{\otimes r}}),$$

also ist auch $V^{\otimes r}_{\text{skew}}$ ein Unterraum von $V^{\otimes r}$.

Für die Unterraumeigenschaft von $V^{\otimes r}_{\text{alt}}$ verwenden wir das Unterraumkriterium. Da der Nulltensor zu $V^{\otimes r}_{\text{alt}}$ gehört, ist $V^{\otimes r}_{\text{alt}} \neq \emptyset$. Sind $t_1, t_2 \in V^{\otimes r}_{\text{alt}}$ und $\alpha \in K$, so sind aber $\alpha\, t_1 \in V^{\otimes r}_{\text{alt}}$ und $t_1 + t_2 \in V^{\otimes r}_{\text{alt}}$ klar. □

Als nächstes untersuchen wir, wie die Symmetrieeigenschaften eines Tensors anhand einer Komponentenhypermatrix erkannt werden können.

Lemma 26.6 (Symmetrie, Schiefsymmetrie und alternierende Eigenschaft von Tensoren und Komponenten)**.**
Es seien V ein endlich-dimensionaler Vektorraum mit $\dim(V) = n \in \mathbb{N}_0$ über dem Körper K und B_V eine Basis von V. Weiter sei $r \in \mathbb{N}_0$. Wir betrachten einen Tensor $t \in V^{\otimes r}$ und seine Komponentenhypermatrix $A = \Phi^{-1}_{B_{V\otimes\cdots\otimes V}}(t)$.

Die folgenden Aussagen sind äquivalent:

(i) t ist symmetrisch, also $t \in V^{\otimes r}_{\text{sym}}$.

(ii) A ist **symmetrisch**, d. h., A erfüllt

$$a^{i_{\sigma(1)},\dots,i_{\sigma(r)}} = a^{i_1,\dots,i_r} \tag{26.8}$$

für alle $i_1, \dots, i_r \in [\![1, n]\!]$ und alle Permutationen $\sigma \in S_r$.

Weiter sind äquivalent:

(iii) t ist schiefsymmetrisch, also $t \in V^{\otimes r}_{\text{skew}}$.

(iv) A ist **schiefsymmetrisch**, d. h., A erfüllt

$$a^{i_{\sigma(1)},\dots,i_{\sigma(r)}} = (\operatorname{sgn} \sigma)\, a^{i_1,\dots,i_r} \tag{26.9}$$

für alle $i_1, \dots, i_r \in [\![1, n]\!]$ und alle Permutationen $\sigma \in S_r$.

Außerdem sind äquivalent:

(v) t ist alternierend, also $t \in V^{\otimes r}_{\text{alt}}$.

(vi) A ist **alternierend**, d. h., A erfüllt

$$a^{i_1,\dots,i_r} = 0 \text{ für alle } i_1, \dots, i_r \in [\![1, n]\!] \text{ mit } i_j = i_k \text{ für ein } j \neq k \tag{26.10}$$

und zusätzlich die Bedingungen aus (26.9).[38]

Beachte: Die oben genannten Eigenschaften der Komponentenhypermatrix sind entweder für alle Tensorproduktbasen oder für keine erfüllt. (**Quizfrage 26.3:** Woran liegt das?)

Beweis. Nach Lemma 26.2 gilt für die Komponentenhypermatrix eines Tensors $t \in V^{\otimes r}$ bezüglich der Tensorproduktbasis $B_{V\otimes\cdots\otimes V}$ und jede Permutation $\sigma \in S_r$:

$$\big(a^{i_{\sigma^{-1}(1)},\dots,i_{\sigma^{-1}(r)}}\big)_{(i_1,\dots,i_r)\in[\![1,n]\!]\times\cdots\times[\![1,n]\!]} = \Phi^{-1}_{B_{V\otimes\cdots\otimes V}}(P_\sigma(t)).$$

Daher gilt

$$\begin{aligned}
& t \in V^{\otimes r}_{\text{sym}} \\
\Leftrightarrow\quad & P_\sigma(t) = t \text{ für alle } \sigma \in S_r \\
\Leftrightarrow\quad & \Phi^{-1}_{B_{V\otimes\cdots\otimes V}}(P_\sigma(t)) = \Phi^{-1}_{B_{V\otimes\cdots\otimes V}}(t) \text{ für alle } \sigma \in S_r && \text{denn } \Phi^{-1}_{B_{V\otimes\cdots\otimes V}} \text{ ist bijektiv} \\
\Leftrightarrow\quad & a^{i_{\sigma(1)},\dots,i_{\sigma(r)}} = a^{i_1,\dots,i_r} \text{ für alle } i_1, \dots, i_r \in [\![1, n]\!] \text{ und } \sigma \in S_r && \text{wegen Lemma 26.2.}
\end{aligned}$$

[38] Im Fall $\operatorname{char}(K) \neq 2$ impliziert die Bedingung der Schiefsymmetrie (26.9) bereits (26.10), sodass die alternierende Eigenschaft nicht stärker ist als die Schiefsymmetrie.

Das zeigt die Äquivalenz der Aussage (i) und Aussage (ii). Ganz analog lässt sich die Äquivalenz von Aussage (iii) und Aussage (iv) zeigen.

Aussage (v) $\Rightarrow$ Aussage (vi): Wenn $t \in V_{\text{alt}}^{\otimes r}$ ist, so ist t insbesondere schiefsymmetrisch, also gilt (26.9). Um die Bedingung (26.10) zu sehen, setzen wir Vektoren $v^{i_1}, \dots, v^{i_r}$ aus der Basis B_{V^*} in den Tensor ein, die dual zur Basis B_V ist. Dann gilt gerade

$$t(v^{i_1}, \dots, v^{i_r}) = a^{i_1,\dots,i_r},$$

wir isolieren also einen bestimmten Eintrag aus der Hypermatrix der Komponenten. Bei Wiederholung von Indizes ergibt sich nach Voraussetzung $a^{i_1,\dots,i_r} = 0$, also (26.10).

Aussage (vi) $\Rightarrow$ Aussage (v): Wir zeigen zunächst: Wenn wir in die Eingänge $k < \ell$ des Tensors t beide Male denselben beliebigen Vektor $v \in V^*$ einsetzen und ansonsten nur Basisvektoren, so ergibt sich als Wert die Null. Wir stellen dazu $v = \sum_{i=1}^n \alpha_i \, v^i$ als Linearkombination der dualen Basisvektoren dar und erhalten

$$\begin{aligned}
&t(v^{i_1}, \dots, v^{i_{k-1}}, v, v^{i_{k+1}}, \dots, v^{i_{\ell-1}}, v, v^{i_{\ell+1}}, \dots, v^{i_r}) \\
&\quad = t\Big(v^{i_1}, \dots, v^{i_{k-1}}, \sum_{i=1}^n \alpha_i \, v^i, v^{i_{k+1}}, \dots, v^{i_{\ell-1}}, \sum_{j=1}^n \alpha_j \, v^j, v^{i_{\ell+1}}, \dots, v^{i_r}\Big) \\
&\quad = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \, \alpha_j \, t(v^{i_1}, \dots, v^{i_{k-1}}, v^i, v^{i_{k+1}}, \dots, v^{i_{\ell-1}}, v^j, v^{i_{\ell+1}}, \dots, v^{i_r}).
\end{aligned}$$

Wir teilen jetzt die Doppelsumme auf in die Fälle $i = j$ und $i \neq j$. Im letzteren Fall können wir die Summanden mit $i < j$ und $i > j$ jeweils paarweise zusammenfassen:

$$\begin{aligned}
&= \sum_{i=1}^n \alpha_i^2 \, t(v^{i_1}, \dots, v^{i_{k-1}}, v^i, v^{i_{k+1}}, \dots, v^{i_{\ell-1}}, v^i, v^{i_{\ell+1}}, \dots, v^{i_r}) \\
&\quad + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \big[\alpha_i \, \alpha_j \, t(v^{i_1}, \dots, v^{i_{k-1}}, v^i, v^{i_{k+1}}, \dots, v^{i_{\ell-1}}, v^j, v^{i_{\ell+1}}, \dots, v^{i_r}) \\
&\qquad\qquad + \alpha_j \, \alpha_i \, t(v^{i_1}, \dots, v^{i_{k-1}}, v^j, v^{i_{k+1}}, \dots, v^{i_{\ell-1}}, v^i, v^{i_{\ell+1}}, \dots, v^{i_r})\big] \\
&= \sum_{i=1}^n \alpha_i^2 \, a^{i_1,\dots,i_{k-1},i,i_{k+1},\dots,i_{\ell-1},i,i_{\ell+1},\dots,i_r} \\
&\quad + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \alpha_i \, \alpha_j \big[a^{i_1,\dots,i_{k-1},i,i_{k+1},\dots,i_{\ell-1},j,i_{\ell+1},\dots,i_r} + a^{i_1,\dots,i_{k-1},j,i_{k+1},\dots,i_{\ell-1},i,i_{\ell+1},\dots,i_r}\big].
\end{aligned}$$

Die erste Summe ist gleich Null wegen (26.10). Die zweite Summe ist gleich Null wegen der Schiefsymmetrie (26.9) von A. Das gilt auch im Fall $\operatorname{char}(K) = 2$. (**Quizfrage 26.4:** Klar?) Wir erhalten also insgesamt den Wert Null.

Die Multilinearität von t zeigt, dass das auch dann noch richtig ist, wenn wir an den Stellen ungleich k und ℓ nicht nur Basisvektoren, sondern beliebige Vektoren einsetzen. $\square$

Beispiel 26.7 (symmetrische, schiefsymmetrische und alternierende Tensoren)**.**

Wir betrachten in diesem Beispiel Tensoren in $V^{\otimes 3}$ (also mit $r = 3$ Achsen) über einem K-Vektorraum V der Dimension $n = 3$ bzw. $n = 4$. Wir stellen jeden Tensor als Hypermatrix $K^{n\times n\times n}$ (bezüglich irgendeiner Tensorproduktbasis, die hier nichts zur Sache tut) dar und geben diese in **Scheiben** (englisch: slices) in Form der Matrizen $a^{\bullet\bullet 1}$, $a^{\bullet\bullet 2}$ und $a^{\bullet\bullet 3}$ an.

Da alternierende bzw. schiefsymmetrische Tensoren viele Nulleinträge aufweisen, schreiben wir diese der besseren Lesbarkeit wegen als $\cdot$ und nicht als 0.

(i) In diesem Beispiel ist $K = \mathbb{Q}$ und $n = \dim(V) = 3$. Der durch die Komponenten

$$a^{\bullet\bullet 1} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \quad a^{\bullet\bullet 2} = \begin{bmatrix} 2 & 4 & 5 \\ 4 & 7 & 8 \\ 5 & 8 & 9 \end{bmatrix}, \quad a^{\bullet\bullet 3} = \begin{bmatrix} 3 & 5 & 6 \\ 5 & 8 & 9 \\ 6 & 9 & 10 \end{bmatrix}$$

repräsentierte Tensor ist symmetrisch. Die Komponentenhypermatrix ist durch die zehn markierten Einträge an den Positionen $(1,1,1)$, $(1,1,2)$, $(1,1,3)$, $(1,2,2)$, $(1,2,3)$, $(1,3,3)$, $(2,2,2)$, $(2,2,3)$, $(2,3,3)$ und $(3,3,3)$ eindeutig bestimmt, die unabhängig voneinander gewählt werden können. Es gilt hier also offenbar $\dim(V^{\otimes 3}_{\text{sym}}) = 10$.

(ii) In diesem Beispiel ist $K = \mathbb{Q}$ und $n = \dim(V) = 3$. Der durch die Komponenten

$$a^{\bullet\bullet 1} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 \\ \cdot & -1 & \cdot \end{bmatrix}, \quad a^{\bullet\bullet 2} = \begin{bmatrix} \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot \end{bmatrix}, \quad a^{\bullet\bullet 3} = \begin{bmatrix} \cdot & 1 & \cdot \\ -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

repräsentierte Tensor ist schiefsymmetrisch und wegen Lemma 26.5 auch alternierend. Er ist durch den einzigen markierten Eintrag an der Position $(1,2,3)$ bereits eindeutig bestimmt. Alle Positionen mit doppelten Indizes sind Null.

Alle weiteren schiefsymmetrischen Tensoren in $V^{\otimes 3}_{\text{skew}}$ sind Vielfache dieses Tensors! Es gilt hier also offenbar $\dim(V^{\otimes 3}_{\text{skew}}) = 1$.

(iii) In diesem Beispiel ist $K = \mathbb{Q}$ und $n = \dim(V) = 4$. Der durch die Komponenten

$$a^{\bullet\bullet 1} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 2 \\ \cdot & -1 & \cdot & 3 \\ \cdot & -2 & -3 & \cdot \end{bmatrix}, \quad a^{\bullet\bullet 2} = \begin{bmatrix} \cdot & \cdot & -1 & -2 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & 4 \\ 2 & \cdot & -4 & \cdot \end{bmatrix},$$

$$a^{\bullet\bullet 3} = \begin{bmatrix} \cdot & 1 & \cdot & -3 \\ -1 & \cdot & \cdot & -4 \\ \cdot & \cdot & \cdot & \cdot \\ 3 & 4 & \cdot & \cdot \end{bmatrix}, \quad a^{\bullet\bullet 4} = \begin{bmatrix} \cdot & 2 & 3 & \cdot \\ -2 & \cdot & 4 & \cdot \\ -3 & -4 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

repräsentierte Tensor ist schiefsymmetrisch und wegen Lemma 26.5 auch alternierend. Er ist durch die vier markierten Einträge an den Positionen $(1,2,3)$, $(1,2,4)$, $(1,3,4)$ und $(2,3,4)$ eindeutig bestimmt, die unabhängig voneinander gewählt werden können. Es gilt hier also offenbar $\dim(V^{\otimes 3}_{\text{skew}}) = 4$.

(iv) Im abschließenden Beispiel ist $K = \mathbb{Z}_2$ und $n = \dim(V) = 4$. Der durch die Komponenten

$$a^{\bullet\bullet 1} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 \\ \cdot & 1 & \cdot & 1 \\ \cdot & 1 & 1 & \cdot \end{bmatrix}, \quad a^{\bullet\bullet 2} = \begin{bmatrix} \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 \\ 1 & \cdot & 1 & \cdot \end{bmatrix},$$

$$a^{\bullet\bullet 3} = \begin{bmatrix} \cdot & 1 & \cdot & 1 \\ 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot \end{bmatrix}, \quad a^{\bullet\bullet 4} = \begin{bmatrix} \cdot & 1 & 1 & \cdot \\ 1 & \cdot & 1 & \cdot \\ 1 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

gegebene Tensor ist alternierend und wegen Lemma 26.5 auch schiefsymmetrisch. Er ist durch die vier Gruppen verschiedenfarbig markierter Einträge an den Positionen $(1, 2, 3)$, $(1, 2, 4)$, $(1, 3, 4)$ und $(2, 3, 4)$ eindeutig bestimmt.

In jeder dieser vier Gruppen können wir die Einträge unabhängig voneinander als 0 oder 1 wählen und erhalten so alle $2^4 = 16$ verschiedenen Tensoren in $V_{\text{alt}}^{\otimes 3}$. Es gilt hier also offenbar $\dim(V_{\text{alt}}^{\otimes 3}) = 4$.

Indem wir beispielsweise a^{111} von 0 auf 1 ändern, erhalten wir einen Tensor, der zwar schiefsymmetrisch, aber nicht alternierend ist. Das ist nur im Fall $\operatorname{char}(K) = 2$ möglich (Lemma 26.5). △

Wir untersuchen jetzt die Dimensionen der Unterräume $V_{\text{sym}}^{\otimes r}$, $V_{\text{skew}}^{\otimes r}$ und $V_{\text{alt}}^{\otimes r}$ genauer.

Satz 26.8 (Dimension der Unterräume $V_{\text{sym}}^{\otimes r}$, $V_{\text{skew}}^{\otimes r}$ und $V_{\text{alt}}^{\otimes r}$, vgl. Lemma 15.32)**.**
Es sei V ein Vektorraum über einem Körper K und $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $r \in \mathbb{N}_0$.

(i) Es gilt:[39]

$$\dim(V_{\text{sym}}^{\otimes r}) = \binom{n+r-1}{r} = \begin{cases} \frac{(n+r-1)!}{r!\,(n-1)!} & \text{falls } n \geqslant 1 \\ 0 & \text{falls } n = 0 \end{cases} \tag{26.11a}$$

$$\dim(V_{\text{alt}}^{\otimes r}) = \binom{n}{r} = \begin{cases} \frac{n!}{r!\,(n-r)!} & \text{falls } n \geqslant r \\ 0 & \text{falls } n < r. \end{cases} \tag{26.11b}$$

(ii) Insbesondere gilt $\dim(V_{\text{alt}}^{\otimes r}) = 1$ für $n = r$ und $\dim(V_{\text{alt}}^{\otimes r}) = 0$ für $n < r$.

(iii) Im Fall $\operatorname{char}(K) \neq 2$ gilt $\dim(V_{\text{skew}}^{\otimes r}) = \dim(V_{\text{alt}}^{\otimes r})$.

(iv) Im Fall $\operatorname{char}(K) = 2$ gilt $\dim(V_{\text{skew}}^{\otimes r}) = \dim(V_{\text{sym}}^{\otimes r})$.

Beweis. Aussage (i): Nach Satz 23.34 ist der Tensorproduktraum $V^{\otimes r}$ isomorph zum Vektorraum $K^{n\times\cdots\times n}$ der Hypermatrizen. Der Unterraum $V_{\text{sym}}^{\otimes r}$ ist dabei nach Lemma 26.6 isomorph zum Unterraum der symmetrischen Hypermatrizen. Diese Hypermatrizen sind durch die Einträge an den Positionen

$$I_{\text{sym}} \coloneqq \big\{(i_1, \ldots, i_r) \in [\![1, n]\!]^r \,\big|\, 1 \leqslant i_1 \leqslant i_2 \leqslant \cdots \leqslant i_r \leqslant n\big\}$$

eindeutig bestimmt; genauer: Es besteht eine Bijektion zwischen den symmetrischen Hypermatrizen und ihren Einträgen mit Indizes in I_{sym}, also gilt $\dim(V_{\text{sym}}^{\otimes r}) = \#I_{\text{sym}}$. (**Quizfrage 26.5:** Wie bestimmen sich die anderen Einträge?)

[39]In (26.11) ist $\binom{n}{r}$ der **Binomialkoeffizient** (englisch: binomial coefficient) „n über r“. Zu beachten ist $\binom{n}{r} = 0$ für $r > n$.

Zur Bestätigung der Kardinalität $\#I_{\text{sym}} = \binom{n+r-1}{r}$ führen wir einen Induktionsbeweis nach der Dimension $\dim(V) = n \in \mathbb{N}_0$. Im Fall $n = 0$ ist $I_{\text{sym}} = \emptyset$, also ist $\#I_{\text{sym}} = 0 = \binom{r-1}{r}$. Beim Übergang von $n \in \mathbb{N}_0$ nach $n+1$ vergrößert sich die Menge I_{sym} um die Tupel, die nach einem gewissen Index $k \in [\![1, r]\!]$ alle den Wert $n+1$ annehmen: $i_{k+1} = \cdots = i_r = n+1$. Die ersten k Indizes können dabei beliebige monoton aufsteigende Werte aus $[\![1, n]\!]$ annehmen.[40] Die Anzahl der Möglichkeiten dafür beträgt nach Induktionsvoraussetzung $\binom{n+k-1}{k}$. Für die Kardinalität der Menge I_{sym} mit $n+1$ anstelle von n gilt also

$$\sum_{k=0}^{r} \binom{n+k-1}{k} = \binom{(n+1)+r-1}{r}.$$

Diese Gleichheit ist eine bekannte Identität der Binomialkoeffizienten, siehe Hockey-Stick-Identity. Das zeigt (26.11a).

Nun zur alternierenden Eigenschaft. Diese liegt nach Lemma 26.6 genau dann vor, wenn die Komponentenhypermatrix alternierend ist. In diesem Fall ist diese bereits durch die Einträge mit den Indizes

$$I_{\text{alt}} \coloneqq \left\{(i_1, \dots, i_r) \in [\![1, n]\!]^r \,\middle|\, 1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n\right\}$$

eindeutig bestimmt; genauer: Es besteht eine Bijektion zwischen den alternierenden Hypermatrizen und ihren Einträgen mit Indizes in I_{alt}, also gilt $\dim(V_{\text{skew}}^{\otimes r}) = \#I_{\text{alt}}$. (**Quizfrage 26.6:** Wie bestimmen sich die anderen Einträge?)

Die Behauptung $\#I_{\text{alt}} = \binom{n}{r}$ ist leicht zu sehen, da $\binom{n}{r}$ die Anzahl der Möglichkeiten angibt, genau r verschiedene Indizes aus der Menge $[\![1, n]\!]$ auszuwählen, die dann aufsteigend sortiert werden. □

Bemerkung 26.9 (Dimension der Unterräume $V_{\text{sym}}^{\otimes r}$, $V_{\text{skew}}^{\otimes r}$ und $V_{\text{alt}}^{\otimes r}$)**.**

(i) Im Fall von Matrizen, die 2-achsige Tensoren aus $V^{\otimes 2}$ über endlich-dimensionalen Vektorräumen repräsentieren, konnten wir zeigen (Lemma 15.32), dass (zumindest im Fall $\operatorname{char}(K) \neq 2$)

$$K^{n \times n} = K_{\text{sym}}^{n \times n} \oplus K_{\text{skew}}^{n \times n}$$

gilt, sodass jede $n \times n$-Matrix eindeutig dargestellt werden kann als Summe ihres symmetrischen und ihres schiefsymmetrischen Anteils. Das stimmt im Fall von $V^{\otimes r}$ mit $r \geqslant 3$ bereits für kleine Werte von $n = \dim(V)$ nicht mehr, wie die folgende Tabelle zeigt:

	$r = 2$			$r = 3$		
n	$\dim(V^{\otimes 2})$	$\dim(V_{\text{sym}}^{\otimes 2})$	$\dim(V_{\text{skew}}^{\otimes 2})$	$\dim(V^{\otimes 3})$	$\dim(V_{\text{sym}}^{\otimes 3})$	$\dim(V_{\text{skew}}^{\otimes 3})$
0	0	0	0	0	0	0
1	1	1	0	1	1	0
2	4	3	1	8	4	0
3	9	6	3	27	10	1
4	16	10	6	64	20	4

[40] Der Wert $k = r$ beschreibt den Fall, dass keiner der Indizes den neuen Wert $n+1$ annimmt.

(*ii*) Aus Satz 26.8 folgt insbesondere, dass es in einem Vektorraum V mit $\dim(V) = n \in \mathbb{N}_0$ bis auf Skalierung nur einen einzigen alternierenden Tensor in $V_{\text{alt}}^{\otimes n}$ gibt. Dasselbe gilt natürlich auch für Tensoren $V_{\text{alt}}^{*\otimes n}$ über dem Dualraum V^*. Das heißt nach Satz 24.23 aber gleichzeitig auch, dass es — bis auf Skalierung — nur eine einzige alternierende Multilinearform $V \times \cdots \times V \to K$ gibt. Diese wird Gegenstand von § 27 sein. △

Wir zeigen jetzt abschließend noch, wie wir einen beliebigen $(r, 0)$-Tensor symmetrisieren bzw. „schiefsymmetrisieren" können.

Satz 26.10 (Symmetrisierung und Schiefsymmetrisierung von Tensoren)**.**
Es sei V ein Vektorraum über einem Körper K mit $\operatorname{char}(K) = 0$. Weiter sei $r \in \mathbb{N}_0$.

(*i*) Der Endomorphismus (genannt die **Symmetrisierung**, englisch: symmetrization)
$$V^{\otimes r} \ni t \mapsto \operatorname{Sym}(t) \coloneqq \frac{1}{r!} \sum_{\sigma \in S_r} P_\sigma(t) \in V_{\text{sym}}^{\otimes r} \subseteq V^{\otimes r} \tag{26.12}$$
bildet beliebige Tensoren in $V^{\otimes r}$ auf symmetrische Tensoren in $V_{\text{sym}}^{\otimes r}$ ab. Für $t \in V^{\otimes r}$ gilt $\operatorname{Sym}(t) = t$ genau dann, wenn $t \in V_{\text{sym}}^{\otimes r}$ ist.

(*ii*) Der Endomorphismus (genannt die **Schiefsymmetrisierung**, englisch: skew-symmetrization)
$$V^{\otimes r} \ni t \mapsto \operatorname{Skew}(t) \coloneqq \frac{1}{r!} \sum_{\sigma \in S_r} (\operatorname{sgn} \sigma)\, P_\sigma(t) \in V_{\text{skew}}^{\otimes r} \subseteq V^{\otimes r} \tag{26.13}$$
bildet beliebige Tensoren in $V^{\otimes r}$ auf schiefsymmetrische Tensoren in $V_{\text{skew}}^{\otimes r}$ ab. Für $t \in V^{\otimes r}$ gilt $\operatorname{Skew}(t) = t$ genau dann, wenn $t \in V_{\text{skew}}^{\otimes r}$ ist.

Quizfrage 26.7: Was bedeutet $\frac{1}{r!}$ in einem beliebigen Körper, und warum fordern wir $\operatorname{char}(K) = 0$?

Beweis. Der Beweis ist Gegenstand der Übung. □

Quizfrage 26.8: Wie sehen die Komponenten eines symmetrisierten bzw. schiefsymmetrisierten Tensors aus?

Expertenwissen: symmetrische Produkte

Wenn wir uns ausschließlich mit *symmetrischen* Multilinearformen auf $\bigtimes_{k=1}^r V$ auf dem K-Vektorraum V beschäftigen wollen, so könnten wir anstelle des Tensorprodukts $V^{\otimes r}$ auch einen darauf angepassten Vektorraum verwenden, der der **symmetrische Produktraum** (englisch: symmetric product space) von V genannt und oft als $V^{\odot r}$ notiert wird. Dieser Vektorraum kommt mit einer **universellen symmetrischen Abbildung** $\odot\colon \bigtimes_{k=1}^r V \to V^{\odot r}$, sodass diese zusammen die Eigenschaft haben, dass jede symmetrische multilineare Abbildung $m\colon \bigtimes_{k=1}^r V \to W$ in einen beliebigen K-Vektorraum W eindeutig durch eine lineare Abbildung $f\colon V^{\odot r} \to W$ dargestellt werden kann, sodass $m = f \circ \odot$ gilt, also $m(v_1, \ldots, v_r) = f(v_1 \odot \cdots \odot v_r)$ für alle $v_1, \ldots, v_r \in V$.

Das symmetrische Produkt $(V^{\odot r}, \odot)$ ist durch die genannte universelle Eigenschaft bis auf Isomorphie eindeutig bestimmt. Eine Möglichkeit der Konstruktion ist als Faktorraum $V^{\otimes r} \,/\, S$ nach dem Unterraum $S \subseteq V^{\otimes r}$, der von den Tensoren der Form $v_1 \otimes \cdots \otimes v_r - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(r)}$ für alle $v_1, \ldots, v_r \in V$ und alle Permutationen σ von $[\![1, r]\!]$ aufgespannt wird; das sind genau diejenigen Tensoren, die im Kern jeder linearen Abbildung $f \colon V^{\odot r} \to W$ liegen, die zu einer symmetrischen multilinearen Abbildung $m \colon \bigtimes_{k=1}^r V \to W$ gehört. Das symmetrische Produkt $\odot$ wird definiert als $\pi \circ \otimes$ mit der kanonischen Surjektion $\pi \colon V^{\otimes r} \to V^{\otimes r} \,/\, S$ und der universellen multilinearen Abbildung $\otimes \colon \bigtimes_{k=1}^r V \to V^{\otimes r}$ des Tensorprodukts.

Eine andere Möglichkeit der Realisierung ist, den Unterraum $V^{\otimes r}_{\text{sym}}$ zu verwenden und als universelle symmetrische Abbildung $\odot = \operatorname{Sym} \circ \otimes$ zu wählen, also $v_1 \odot \cdots \odot v_r \coloneqq \operatorname{Sym}(v_1 \otimes \cdots \otimes v_r)$.

Expertenwissen: äußere Produkte

Analog gibt es auch für *alternierende* Multilinearformen auf $\bigtimes_{k=1}^r V$ einen angepassten Vektorraum, der der **äußere Produktraum** (englisch: exterior product space), der **alternierende Produktraum** (englisch: alternating product space) oder der **Dachproduktraum** (englisch: wedge product space) von V genannt wird und oft als $V^{\wedge r}$ notiert wird. Zu diesem Vektorraum gibt es eine **universelle alternierende Abbildung** $\wedge \colon \bigtimes_{k=1}^r V \to V^{\wedge r}$, sodass diese zusammen die Eigenschaft haben, dass jede alternierende multilineare Abbildung $m \colon \bigtimes_{k=1}^r V \to W$ in einen beliebigen K-Vektorraum W eindeutig durch eine lineare Abbildung $f \colon V^{\wedge r} \to W$ dargestellt werden kann, sodass $m = f \circ \wedge$ gilt, also $m(v_1, \ldots, v_r) = f(v_1 \wedge \cdots \wedge v_r)$ für alle $v_1, \ldots, v_r \in V$.

Auch das alternierende Produkt $(V^{\wedge r}, \wedge)$ ist durch die genannte universelle Eigenschaft bis auf Isomorphie eindeutig bestimmt. Der äußere Produktraum kann als Faktorraum $V^{\otimes r} \,/\, A$ nach dem Unterraum $A \subseteq V^{\otimes r}$ konstruiert werden, der von den Tensoren der Bauart $v_1 \otimes \cdots \otimes v_r$ mit $v_i = v_j$ für $i \neq j$ aufgespannt wird; das sind genau diejenigen Tensoren, die im Kern jeder linearen Abbildung $f \colon V^{\wedge r} \to W$ liegen, die zu einer alternierenden multilinearen Abbildung $m \colon \bigtimes_{k=1}^r V \to W$ gehört. Das alternierende Produkt $\wedge$ wird dann definiert als $\pi \circ \otimes$ mit der kanonischen Surjektion $\pi \colon V^{\otimes r} \to V^{\otimes r} \,/\, A$ und der universellen multilinearen Abbildung $\otimes \colon \bigtimes_{k=1}^r V \to V^{\otimes r}$ des Tensorprodukts.

Solange $\operatorname{char}(K) \neq 2$ gilt, können wir zur Realisierung alternativ auch den Unterraum $V^{\otimes r}_{\text{alt}}$ verwenden und als universelle alternierende Abbildung $\wedge = \operatorname{Skew} \circ \otimes$ definieren, also $v_1 \wedge \cdots \wedge v_r \coloneqq \operatorname{Skew}(v_1 \otimes \cdots \otimes v_r)$.

Ende der Vorlesung 9

Ende der Woche 4

§ 27 Determinanten

Literatur: Fischer, Springborn, 2020, Kapitel 4; Bosch, 2014, Kapitel 4

In diesem Abschnitt ist V stets ein endlich-dimensionaler Vektorraum über einem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Die **Determinante** ist eine Maßzahl für Endomorphismen $f \in \operatorname{Endo}(V)$, die gewisse Eigenschaften festhält. Unser Vorgehen wird sein, zunächst Determinantenformen auf V zu definieren, eine bestimmte davon als *die* Determinante für Matrizen zu erklären und diesen Begriff dann mittels Darstellungsmatrizen auf Endomorphismen zu übertragen.

Definition 27.1 (Determinantenform auf einem endlich-dimensionalen Vektorraum V)**.**
Es sei V ein Vektorraum über einem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Eine Multilinearform

$$\Delta \in \operatorname{Mult}(V^n; K),$$

die alternierend ist, heißt eine **Determinantenform** (englisch: determinant form) **auf dem Vektorraum V**. Eine Determinantenform heißt **nicht-trivial** (englisch: non-trivial), wenn sie nicht die Nullform ist. △

Die folgende Bemerkung hält erste Eigenschaften von Determinantenformen fest, die sich unmittelbar aus der Definition 27.1 ergeben.

Bemerkung 27.2 (zur Definition von Determinantenformen)**.**

(i) Ist die Familie $(v_1, \ldots, v_n)$ von Vektoren in V linear abhängig, so gilt $\Delta(v_1, \ldots, v_n) = 0$ (Lemma 26.4).

(ii) Bei Skalierung eines der Argumente von Δ wird der Funktionswert entsprechend skaliert:

$$\Delta(v_1, \ldots, v_{i-1}, \alpha\, v_i, v_{i+1}, \ldots, v_n) = \alpha\, \Delta(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n).$$

Das folgt sofort aus der Multilinearität von Δ.

(iii) Bei Addition des Vielfachen eines Arguments zu einem anderen Argument ändert sich der Funktionswert nicht:

$$\Delta(v_1, \ldots, v_{i-1}, v_i + \alpha\, v_j, v_{i+1}, \ldots, v_n) = \Delta(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$$

für $i \neq j$. Das folgt aus der Multilinearität von Δ zusammen mit der alternierenden Eigenschaft.

(iv) Beim Vertauschen von zwei Argumenten von Δ ändert sich das Vorzeichen des Funktionswertes:

$$\Delta(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) = -\Delta(v_1, \ldots, v_n)$$

für $i \neq j$. Das folgt aus der alternierenden Eigenschaft von Δ. △

Als n-lineare Form auf V wird jede Determinantenform Δ wegen $\operatorname{Mult}(V^n; K) \cong \operatorname{Homo}(V^{\otimes n}, K)$ durch genau einen Tensor $t \in \mathcal{T}^{(0,n)}(V) = V^{*\otimes n}$ repräsentiert. Da sie alternierend ist, gehört dieser Tensor sogar zum Unterraum $V^{*\otimes n}_{\text{alt}}$.

Wegen $\dim(V^{*\otimes n}_{\text{alt}}) = 1$ (Satz 26.8) bilden die Determinantenformen einen eindimensionalen Unterraum von $\operatorname{Mult}(V^n; K)$. Das hat folgende Konsequenzen:

(1) Eine Determinantenform Δ wird bereits durch einen einzigen Funktionswert $\Delta(b_1, \dots, b_n)$ eindeutig festgelegt, wobei $(b_1, \dots, b_n)$ eine beliebige Basis von V ist. (Für Nicht-Basen ist $\Delta(b_1, \dots, b_n)$ wegen Lemma 26.4 gleich Null.)

(2) Für zwei verschiedene Determinantenformen gilt, dass eine ein skalares Vielfaches der anderen ist.

Das folgende Resultat klärt die Gestalt von Determinantenformen und zeigt, wie man den Funktionswert $\Delta(v_1, \dots, v_n)$ für beliebige Vektoren $v_1, \dots, v_n \in V$ berechnen kann.

Satz 27.3 (Gestalt und Auswertung von Determinantenformen)**.**

Es seien V ein endlich-dimensionaler Vektorraum mit $\dim(V) = n \in \mathbb{N}_0$ über dem Körper K. Weiter seien $B_V = (b_1, \dots, b_n)$ eine Basis von V und $B_{V^*} = (b_1^*, \dots, b_n^*)$ die zugehörige duale Basis. Dann gilt:

(i) Die Determinantenformen Δ auf V sind genau die Tensoren in $V^{*\otimes n}$ der Gestalt[41]

$$\Delta = \alpha \sum_{\sigma \in S_n} \underbrace{(\operatorname{sgn} \sigma)}_{=\pm 1 \in K} b^*_{\sigma(1)} \otimes \cdots \otimes b^*_{\sigma(n)} \tag{27.1}$$

mit $\alpha \in K$. Die Zuordnung $V^{*\otimes n}_{\text{alt}} \ni \Delta \mapsto \alpha \in K$ ist ein Isomorphismus von (eindimensionalen) Vektorräumen.

(ii) Sind $v_1, \dots, v_n \in V$ beliebig und haben diese Vektoren die Darstellung[42] $v_j = \sum_{i=1}^n a_{ij}\, b_i$, dann gilt

$$\Delta(v_1, \dots, v_n) = \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n} \Big) \underbrace{\Delta(b_1, \dots, b_n)}_{\text{Normierungswert } \alpha \in K} \tag{27.2}$$

Beweis. Aussage (i): Jeder Tensor der Gestalt (27.1) repräsentiert nach Satz 24.6 genau eine multilineare Abbildung $\bigtimes_{k=1}^n V \to K$.

Schritt 1: Wir zeigen: Δ ist alternierend.

Die Hypermatrix der Komponenten von Δ bzgl. der von $B_{V^*} = (b_1^*, \dots, b_n^*)$ induzierten Tensorproduktbasis ist alternierend, denn die Hypermatrix zum Summanden $\alpha\, (\operatorname{sgn} \sigma)\, b^*_{\sigma(1)} \otimes \cdots \otimes b^*_{\sigma(n)}$ weist nur einen einzigen Eintrag ungleich Null auf, nämlich den an der Stelle $(\sigma(1), \dots, \sigma(n))$ mit dem Wert $\alpha\, (\operatorname{sgn} \sigma)$. Verschiedene Summanden haben ihre Nicht-Nulleinträge an verschiedenen Positionen. Positionen mit wiederholten Indizes kommen nicht vor und bleiben bei Null. Daraus ergibt sich die alternierende Eigenschaft der Hypermatrix der Komponenten, und aus Lemma 26.6 folgt, dass auch der Tensor Δ alternierend ist.

[41] Die Komponentenhypermatrix von Δ hat also nur Einträge $\pm\alpha$ und Null (an Positionen mit wiederholten Indizes), vgl. Beispiel 26.7.

[42] In der j-ten Spalte von A stehen also die Koordinaten von v_j bzgl. der Basis $(b_1, \dots, b_n)$. Falls die $(v_1, \dots, v_n)$ ebenfalls eine Basis bilden, so ist A die Transformationsmatrix $A = \mathcal{T}_{(b_1,\dots,b_n) \leftarrow (v_1,\dots,v_n)} \in K^{n \times n}$, siehe Definition 19.15.

Schritt 2: Wir zeigen die Isomorphismuseigenschaft.

Wie in Schritt 1 gezeigt, erzeugt die rechte Seite von (27.1) für jedes $\alpha \in K$ eine Determinantenform $\Delta \in V^{*\otimes n}_{\text{alt}}$, und verschiedene α erzeugen verschiedene Determinantenformen. Die Abbildung $K \ni \alpha \mapsto \Delta \in V^{*\otimes n}_{\text{alt}}$ ist also injektiv und offensichtlich linear. Da $\dim(K) = \dim(V^{*\otimes n}_{\text{alt}}) = 1$ gilt, ist sie auch surjektiv (Folgerung 18.9) und damit ein Isomorphismus von Vektorräumen.

Damit ist Aussage (i) gezeigt.

Aussage (ii): Aufgrund der Multilinearität der Determinantenform gilt

$$\begin{aligned}\Delta(v_1, \ldots, v_n) &= \Delta\Big(\sum_{i_1=1}^{n} a_{i_1,1}\, b_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n,n}\, b_{i_n}\Big) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1,1}\, a_{i_2,2} \cdots a_{i_n,n}\, \Delta(b_{i_1}, b_{i_2}, \ldots, b_{i_n}).\end{aligned}$$

Da Δ als Determinantenform alternierend ist, gilt $\Delta(b_{i_1}, b_{i_2}, \ldots, b_{i_n}) = 0$, sobald zwei der Indizes übereinstimmen. Wir können die Summation daher auf diejenigen Indextupel $(i_1, \ldots, i_n)$ beschränken, deren Einträge sämtlich verschieden sind, also auf die Permutationen von $[\![1, n]\!]$:

$$= \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n}\, \Delta(b_{\sigma(1)}, \ldots, b_{\sigma(n)}).$$

Da Δ als alternierende Multilinearform auch schiefsymmetrisch ist, folgt schließlich

$$= \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}\Big)\, \Delta(b_1, \ldots, b_n),$$

was zu beweisen war. □

Beispiel 27.4 (Determinantenformen)**.**

(i) Im Fall $\dim(V) = n = 0$ ist jede Abbildung $\Delta\colon \{()\} \ni () \mapsto \alpha \in K$ multilinear und alternierend.

(ii) Im Fall $\dim(V) = n = 1$ sind die multilinearen Abbildungen $V \to K$ gerade die linearen Abbildungen, also $\operatorname{Mult}(V; K) = \operatorname{Homo}(V, K) = V^*$. Diese sind automatisch auch alternierend. Jede Determinantenform Δ auf V hat die Gestalt

$$\Delta(v) = \alpha\, \langle b^*, v\rangle$$

mit irgendeinem (Basis-)Vektor $b^* \in V^* \setminus \{0\}$ und $\alpha \in K$.

(iii) Im Fall $\dim(V) = n = 2$ betrachten wir als Beispiel den Vektorraum $V = K^2$ über einem Körper K mit der Standardbasis (e_1, e_2). Dann hat jede Determinantenform Δ auf V die Gestalt

$$\Delta = \alpha\, (e_1^* \otimes e_2^* - e_2^* \otimes e_1^*)$$

mit einem Normierungswert $\alpha \in K$. Für den Wert der Determinantenform auf beliebigen Vektoren $v_1, v_2 \in V$ mit der Darstellung

$$\begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 & e_2 \end{bmatrix},$$

also $v_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ und $v_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$, gilt dann nach (27.2)

$$\Delta(v_1, v_2) = \alpha \sum_{\sigma \in S_2} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} = \alpha\, \big(\underbrace{a_{11}\, a_{22}}_{\text{zu } \sigma = \text{id}} - \underbrace{a_{21}\, a_{12}}_{\text{zu } \sigma = \tau(1,2)}\big).$$ △

Expertenwissen: Rang von Determinantentensoren

Die Darstellung (27.1) zeigt, dass Determinantentensoren maximal Rang $n!$ haben. Tatsächlich kann aber zum Beispiel im Fall $n = 3$ der Determinantentensor bereits als Summe von 5 einfachen Tensoren dargestellt werden:

$$\begin{aligned} \Delta = {} & (b_2^* + b_3^*) \otimes (b_1^* + b_3^*) \otimes (b_1^* + b_2^*) - b_2^* \otimes b_1^* \otimes (b_1^* + b_2^* + b_3^*) \\ & - b_3^* \otimes (b_1^* + b_2^* + b_3^*) \otimes b_1^* - (b_1^* + b_2^* + b_3^*) \otimes b_3^* \otimes b_2^* + b_1^* \otimes b_2^* \otimes b_3^*, \end{aligned}$$

und die Anzahl 5 ist auch optimal. Bereits im Fall $n = 4$ ist der Rang von Determinantentensoren nicht bekannt (Anfang 2026) und aktuelles Forschungsgebiet.

§ 27.1 Die Determinante einer Matrix

Bei der Bestimmung des Funktionswertes einer Determinantenform Δ gemäß (27.2), also

$$\Delta(v_1, \ldots, v_n) = \underbrace{\Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}\Big)}_{\det(A)}\ \underbrace{\Delta(b_1, \ldots, b_n)}_{\text{Normierungswert}}$$

können wir zwei Einflüsse unterscheiden: Der Normierungswert $\Delta(b_1, \ldots, b_n)$ unterscheidet nur darüber, welche Determinantenform aus dem eindimensionalen Vektorraum aller Determinantenformen wir ausgewählt haben, also nur über die Skalierung. Der „interessantere“ Faktor ist der erste, der von der Koeffizientenmatrix A der Argumente $v_j = \sum_{i=1}^n a_{ij}\, b_i$ abhängt. Diese Größe heißt die **Determinante** der Matrix $A \in K^{n \times n}$:

Definition 27.5 (Determinante einer Matrix)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$. Die **Determinante** (englisch: determinant) einer Matrix $A \in K^{n \times n}$ ist definiert durch

$$\det(A) \coloneqq \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}. \tag{27.3}$$

Die Gleichung (27.3) heißt die **Leibniz-Formel** (englisch: Leibniz formula). △

Bemerkung 27.6 (zur Determinante einer Matrix).

(*i*) Wir können die Determinante einer Matrix als den Faktor verstehen, der zwischen dem Normierungswert $\Delta(b_1, \dots, b_n)$ einer Determinantenform und ihrem Wert $\Delta(v_1, \dots, v_n)$ für beliebige Argumente $v_1, \dots, v_n \in V$ steht mit $v_j = \sum_{i=1}^n a_{ij}\, b_i$.

Wir können also die Aussage (27.2) auch formulieren als

$$\Delta(v_1, \dots, v_n) = \det(A)\, \Delta(b_1, \dots, b_n).$$

(*ii*) Alternativ können wir die Determinante einer Matrix selbst auch als eine bestimmte Determinantenform (und zwar auf dem Vektorraum $V = K^n$) verstehen: Dazu identifizieren wir ein Tupel $(x_1, \dots, x_n)$ von Vektoren in K^n (also ein Element von $K^n \times \dots \times K^n$) mit der Matrix $X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \in K^{n \times n}$.

Für diejenige Determinantenform Δ, die auf der Standardbasis $(e_1, \dots, e_n)$ von K^n den Normierungswert $\Delta(e_1, \dots, e_n) = 1$ hat, gilt dann $\Delta(x_1, \dots, x_n) = \det(X)$. (**Quizfrage 27.1:** Klar?)

(*iii*) Manche Autoren schreiben statt $\det(A)$ auch $|A|$. Dieser Praxis folgen wir nicht. △

Beispiel 27.7 (Determinante einer Matrix).

(*i*) Im Fall $n = 0$ ist die Determinante der einzigen (leeren) $n \times n$-Matrix $\det(A) = 1$. (**Quizfrage 27.2:** Wie folgt das aus (27.3)?)

(*ii*) Im Fall $n = 1$ gibt es nur die Permutation $\sigma = \text{id}$. Daher folgt für $A = (a)$ aus (27.3) der Zusammenhang $\det(A) = a$. Die Determinante ist dann also gerade der einzige Eintrag der Matrix.

(*iii*) Im Fall $n = 2$ gibt es zwei Permutationen, $\sigma = \text{id}$ und $\sigma = \tau(1, 2)$. Daher folgt aus der Leibniz-Formel (27.3):[43]

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\, a_{22} - a_{12}\, a_{21}.$$

Grafisch können wir diese Regel wie folgt darstellen:

$$\begin{matrix} a_{11} & & a_{12} \\ & \overset{+}{\times}_{-} & \\ a_{21} & & a_{22} \end{matrix}$$

Das Produkt der Hauptdiagonalelemente geht also positiv in den Wert der Determinante ein, während das Produkt der Elemente der „Gegendiagonale" negativ gewichtet wird.

(*iv*) Im Fall $n = 3$ gibt es sechs Permutionen. Davon sind drei gerade Permutationen, nämlich $\sigma = \text{id} = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right)$, $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$ und $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right)$, und drei ungerade: $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right)$, $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$ und $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$. Das Ausschreiben der Summe in (27.3) ergibt

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} \; a_{11}\, a_{22}\, a_{33} + a_{21}\, a_{32}\, a_{13} + a_{31}\, a_{12}\, a_{23} \\ -\; a_{31}\, a_{22}\, a_{13} - a_{21}\, a_{12}\, a_{33} - a_{11}\, a_{32}\, a_{23}. \end{array}$$

[43]Das zweite Klammernpaar lassen wir bei Ausdrücken wie $\det\left(\left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]\right)$ häufig weg und schreiben einfach $\det\left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$.

Grafisch können wir diese sogenannte **Regel von Sarrus** (englisch: rule of Sarrus) wie folgt darstellen:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ + & + & + & & \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ - & - & - & & \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Dabei werden die ersten beiden Spalten der Matrix gedanklich nochmals rechts daneben kopiert, d. h., beim Durchgehen der Diagonalen werden die Einträge in jeder Zeile zyklisch fortgesetzt. Wieder werden die Produkte der Elemente der Haupt- und Nebendiagonalen positiv gewertet, während die Produkte der Elemente in Richtung der „Gegendiagonale“ negativ gewichtet werden.

(v) Im Fall $n = 4$ gibt es bereits 24 Permutionen und keine einfache Merkregel mehr. △

Um die Determinante einer $n \times n$-Matrix zu bestimmen, wird man im Normalfall nur für $n \leqslant 3$ gemäß der Definition (27.3) vorgehen. Für größere Matrizen betrachten wir in § 27.2 noch effizientere Möglichkeiten.

Lemma 27.8 (Eigenschaften der Determinante von Matrizen)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$. Weiter seien $A, B \in K^{n\times n}$. Die Determinante $\det\colon K^{n\times n} \to K$ hat folgende Eigenschaften:

(i) $\det(A)$ ist eine alternierende Multilinearform auf den Spalten $(a_{\bullet 1}, \dots, a_{\bullet n})$ von A.

(ii) $\det(\alpha\, A) = \alpha^n \det(A)$ für alle $\alpha \in K$.

(iii) $\det(A\, B) = \det(A)\ \det(B)$.

(iv) $\det(I) = 1$.

(v) $\det(A) \neq 0 \Leftrightarrow A$ ist regulär $\Leftrightarrow \operatorname{Rang}(A) = n \Leftrightarrow (a_{\bullet 1}, \dots, a_{\bullet n})$ ist linear unabhängig. (Die Determinante kann also als Test für die Invertierbarkeit einer Matrix verwendet werden.)

(vi) $\det(A^{-1}) = 1/\det(A)$, falls A invertierbar ist.

(vii) $\det(A^\intercal) = \det(A)$.

Beachte: Die Aussagen (iii) und (iv) zeigen, dass die Determinante ein Monoidhomomorphismus des Monoids $(K^{n\times n}, \cdot)$ in das Monoid $(K, \cdot)$ ist.

Beweis. Aussage (i) folgt aus Bemerkung 27.6.

Aussage (ii) folgt sofort aus der Multilinearität der Determinante.

Aussage (iii): Wir schreiben die Matrix B spaltenweise als

$$B = \left[\sum_{i_1=1}^{n} b_{i_1,1}\, e_{i_1} \quad \cdots \quad \sum_{i_n=1}^{n} b_{i_n,n}\, e_{i_n} \right],$$

also als Linearkombination der Standardbasisvektoren von K^n. Dann gilt nach Definition des Matrix-Matrix-Produkts (vgl. Bemerkung 15.7)

$$\begin{aligned}
A\,B &= \left[A \sum_{i_1=1}^{n} b_{i_1,1}\, e_{i_1} \quad \cdots \quad A \sum_{i_n=1}^{n} b_{i_n,n}\, e_{i_n}\right] && \text{denn die Spalten des rechten Faktors erzeugen die Spalten des Produkts}\\
&= \left[\sum_{i_1=1}^{n} b_{i_1,1}\,(A\, e_{i_1}) \quad \cdots \quad \sum_{i_n=1}^{n} b_{i_n,n}\,(A\, e_{i_n})\right] && \text{wegen der Linearität des Matrix-Vektor-Produkts.}
\end{aligned}$$

Daher gilt für die Determinante

$$\begin{aligned}
\det(A\,B) &= \det\left(\sum_{i_1=1}^{n} b_{i_1,1}\,(A\, e_{i_1}) \quad \cdots \quad \sum_{i_n=1}^{n} b_{i_n,n}\,(A\, e_{i_n})\right) && \text{aufgrund der Multilinearität}\\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1,1} \cdots b_{i_n,n} \det\left(A\, e_{i_1} \quad \cdots \quad A\, e_{i_n}\right).
\end{aligned}$$

Aus Lemma 26.4 folgt, dass die obige Determinante den Wert Null ergibt, sobald sich Indizes wiederholen. Wir können daher die Summation auf diejenigen Indextupel $(i_1, \ldots, i_n)$ beschränken, deren Einträge sämtlich verschieden sind, also auf die Permutationen von $[\![1, n]\!]$:

$$\begin{aligned}
&= \sum_{\sigma \in S_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} \det\left(A\, e_{\sigma(1)} \quad \cdots \quad A\, e_{\sigma(n)}\right)\\
&= \underbrace{\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, b_{\sigma(1),1} \cdots b_{\sigma(n),n}}_{\det(B)} \underbrace{\det\left(A\, e_1 \quad \cdots \quad A\, e_n\right)}_{\det(A)} && \text{wegen der alternierenden Eigenschaft.}\\
&= \det(A)\, \det(B).
\end{aligned}$$

Aussage (*iv*): Die Einträge der Einheitsmatrix $I \in K^{n \times n}$ sind δ_{ij} für $i, j = 1, \ldots, n$. Also produziert nur $\sigma = \operatorname{id}$ einen Nicht-Null-Beitrag in der Summe (27.3). Genauer:

$$\det(I) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, \delta_{\sigma(1),1} \cdots \delta_{\sigma(n),n} = (\operatorname{sgn} \operatorname{id})\, \delta_{\operatorname{id}(1),1} \cdots \delta_{\operatorname{id}(n),n} = 1.$$

Aussagen (*v*) und (*vi*): Bei linearer Abhängigkeit von $(a_{\bullet 1}, \ldots, a_{\bullet n})$ ist $\det(A) = 0$ nach Lemma 26.4. Sind die Spalten von A dagegen linear unabhängig, dann ist A invertierbar, und nach den Aussagen (*iii*) und (*iv*) gilt

$$1 = \det(I) = \det(A\,A^{-1}) = \det(A)\, \det(A^{-1}),$$

insbesondere ist $\det(A) \neq 0$; vgl. auch Lemma 8.8. Wir haben also

$$\begin{aligned}
& \det(A) \neq 0 \\
\Leftrightarrow \quad & (a_{\bullet 1}, \ldots, a_{\bullet n}) \text{ ist linear unabhängig} && \text{wie gerade gezeigt}\\
\Leftrightarrow \quad & \operatorname{Rang}(A) = n && \text{nach Satz 15.13}\\
\Leftrightarrow \quad & A \text{ ist regulär} && \text{nach Satz 15.45.}
\end{aligned}$$

Aussage (vii): Es gilt

$$\begin{aligned}
\det(A^\intercal) &= \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} && \text{nach der Leibniz-Formel (27.3)}\\
&= \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{\sigma^{-1}(1),1}\cdots a_{\sigma^{-1}(n),n} && \text{durch Umsortieren der Faktoren}\\
&= \sum_{\tau\in S_n} (\operatorname{sgn}\tau^{-1})\, a_{\tau(1),1}\cdots a_{\tau(n),n} && \text{durch Ersetzen von } \tau \coloneqq \sigma^{-1}\\
&= \sum_{\tau\in S_n} (\operatorname{sgn}\tau)\, a_{\tau(1),1}\cdots a_{\tau(n),n} && \text{wegen } \operatorname{sgn}\tau = \operatorname{sgn}\tau^{-1}\text{, siehe Folgerung 7.41}\\
&= \det(A).
\end{aligned}$$

(**Quizfrage 27.3:** Warum durfte die Summation über $\sigma \in S_n$ durch die Summation über $\tau \in S_n$ ersetzt werden?) □

Ende der Vorlesung 10

§ 27.2 Berechnung der Determinante

Die Berechnung der Determinante nach der Leibniz-Formel (27.3) erfordert $n!$ Produkte. Wir geben daher in diesem Abschnitt Resultate an, die bei der effizienteren Berechnung der Determinante helfen können.

Lemma 27.9 (Determinante einer Dreiecksmatrix)**.**
Es seien K ein Körper und $A \in K^{n\times n}$ mit $n \in \mathbb{N}_0$.

(i) Ist $A \in K^{n\times n}_{\triangledown}$ oder $A \in K^{n\times n}_{\triangle}$, dann gilt

$$\det(A) = a_{11}\cdots a_{nn}. \tag{27.4}$$

(ii) Ist A eine obere oder untere Blockdreiecksmatrix, also $A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22}\end{bmatrix}$ bzw. $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}\end{bmatrix}$, dann gilt

$$\det(A) = \det(A_{11})\,\det(A_{22}). \tag{27.5}$$

Beweis. Der Beweis ist Gegenstand der Übung. □

Aufgrund des vorherigen Lemmas ist es nützlich, eine Matrix z. B. durch Zeilenumformungen — also durch Multiplikation mit Elementarmatrizen von links — auf Zeilenstufenform (also insbesondere obere Dreiecksgestalt) zu bringen, um ihre Determinante zu berechnen. Aufgrund von Lemma 27.8 multiplizieren sich dabei die Determinanten.

Lemma 27.10 (Determinante von Elementarmatrizen)**.**
Es sei K ein Körper und $n \in \mathbb{N}_0$. Für die Determinanten der $n \times n$-Elementarmatrizen aus § 15.3 gilt:

$$\text{Typ I} \qquad \det(D_\alpha(i)) = \alpha \tag{27.6a}$$

$$\text{Typ II} \qquad \det(S_{\alpha,i}(j)) = 1 \tag{27.6b}$$

$$\text{Typ III} \qquad \det(T_{i,j}) = -1 \tag{27.6c}$$

mit $i, j \in [\![1, n]\!]$, $i \neq j$ und $\alpha \in K$.

Bei der Berechnung einer Zeilenstufenform multiplizieren wir die betreffende Matrix A wiederholt von links mit Matrizen vom Typ II oder III. (**Quizfrage 27.4:** Wofür braucht man nochmal die Matrizen vom Typ I?) Bei jedem Zeilentausch (Typ III) ändert sich dabei die Determinante um den Faktor -1. Wir müssen uns also nur merken, ob wir eine gerade oder ungerade Anzahl von Zeilentauschvorgängen durchgeführt haben. Sobald bei der Umformung ein Diagonalelement Null wird und damit ein Pivot-Element nicht auf der Diagonalen steht, können wir die Rechnung beenden, weil die Determinante dann Null ist.

Beispiel 27.11 (Berechnung der Determinante durch elementare Zeilenumformungen)**.**

(*i*) Wir berechnen die Determinante von $A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 2 & 2 \end{bmatrix}$, die wir als Matrix über dem Körper $\mathbb{Z}_5$ auffassen, vgl. Beispiel 16.10. Wir bringen A dazu auf Zeilenstufenform:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 2 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 2 \end{bmatrix} \quad \text{Tauschen der Zeilen 1 und 2}$$

$$2 \; \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 2 \end{bmatrix} \quad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements}\end{array}$$

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{Tauschen der Zeilen 2 und 3}$$

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \quad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements}\end{array}$$

Die Determinante der letzten Matrix ist nach Lemma 27.9 das Produkt der Diagonalelemente, also $1 \cdot 3 \cdot 2 = 1$ in $\mathbb{Z}_5$. Da wir auf dem Weg dahin zweimal Zeilen getauscht haben, gilt auch $\det(A) = 1$.

(*ii*) Wir betrachten jetzt die modifizierte Matrix $A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 4 & 2 \end{bmatrix}$:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 4 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 4 & 2 \end{bmatrix} \quad \text{Tauschen der Zeilen 1 und 2}$$

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 4 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$
Erzeugen von Nullen unterhalb des Pivot-Elements

An dieser Stelle können wir die Rechnung beenden, weil ein Pivot-Element nicht auf der Diagonalen steht. Damit gilt $\operatorname{Rang}(A) < n = 3$, also $\det(A) = 0$ (Lemma 27.8). △

Es folgen jetzt weitere Begriffe im Umfeld des Konzepts der Determinante.

Definition 27.12 (Streichungsmatrix, Minoren, Kofaktoren und Adjunkte)**.**
Es seien K ein Körper und $A \in K^{n \times n}$, $n \in \mathbb{N}_0$. Weiter seien $1 \leqslant i, j \leqslant n$.

(i) Die **Streichungsmatrix** von A bzgl. des Index (i, j) ist diejenige $(n-1) \times (n-1)$-Untermatrix, die aus A durch Weglassen der i-ten Zeile und der j-ten Spalte entsteht, also

$$(A)_{\neq i, \neq j} \coloneqq \begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{bmatrix}. \tag{27.7}$$

(ii) Die **Unterdeterminante** (englisch: subdeterminant) oder der **Minor** (englisch: minor) von A bzgl. des Index (i, j) ist die Determinante der zugehörigen Streichungsmatrix, also[44]

$$[A]_{ij} \coloneqq \det\big((A)_{\neq i, \neq j}\big) = \det \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix} \tag{27.8}$$

(iii) Die Größe

$$\widetilde{a}_{ij} \coloneqq (-1)^{i+j} \, [A]_{ij} \tag{27.9}$$

heißt der **Kofaktor** (englisch: cofactor) der Matrix A bzgl. des Index (i, j). Die Matrix $\operatorname{cof}(A) \coloneqq (\widetilde{a}_{ij})_{1 \leqslant i,j \leqslant n}$ heißt die **Kofaktormatrix** (englisch: cofactor matrix) von A.

(iv) Die **Adjunkte** (englisch: adjugate matrix, adjunct matrix) von A oder die zu A **komplementäre Matrix** ist die Transponierte der Kofaktormatrix:

$$\operatorname{adj}(A) \coloneqq \operatorname{cof}(A)^\intercal. \tag{27.10}$$

(v) Im Fall $n = 0$ gibt es keine Streichungsmatrizen, also auch keine Minoren und keine Kofaktoren. Die Kofaktormatrix und die Adjunkte sind dann wie A selbst die leere Matrix. △

[44] Für Minoren gibt es in der Literatur keine einheitliche Notation.

Für die Kofaktoren $\widetilde{a}_{ij}$ gibt es eine gleichwertige alternative Definition, bei der die Zeile i und Spalte j (oder auch nur die Spalte j) ersetzt werden, statt sie zu streichen:

Lemma 27.13 (alternative Definition der Kofaktoren)**.**
Es seien K ein Körper und $A \in K^{n \times n}$, $n \in \mathbb{N}_0$. Für die Kofaktoren von A gilt

$$
\begin{aligned}
\widetilde{a}_{ij} &= \det \begin{array}{cl} \begin{pmatrix}
a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n}
\end{pmatrix} & \leftarrow i\text{-te Zeile} \\ \uparrow j\text{-te Spalte} & \end{array} && (27.11a) \\
&= \det \begin{array}{cl} \begin{pmatrix}
a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n} \\
a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n}
\end{pmatrix} & \leftarrow i\text{-te Zeile} \\ \uparrow j\text{-te Spalte} & \end{array} && (27.11b) \\
&= \det(a_{\bullet 1}, \ldots, a_{\bullet j-1}, e_i, a_{\bullet j+1}, \ldots, a_{\bullet n}). && (27.11c)
\end{aligned}
$$

Beweis. Durch $i-1$ Vertauschungen benachbarter Zeilen (durch Multiplikation mit Elementarmatrizen vom Typ III von links) und $j-1$ Vertauschungen benachbarter Spalten (durch Multiplikation mit Elementarmatrizen vom Typ III von rechts) kann die Matrix in (27.11a) auf die Gestalt

$$
\begin{bmatrix}
1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots \\
\vdots & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
\vdots & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots \\
0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n}
\end{bmatrix}
$$

gebracht werden, deren Determinante nach (27.5) gleich $[A]_{ij}$ ist. Durch die Zeilen- und Spaltenvertauschungen kommt der Faktor $(-1)^{i+j}$ hinzu. Also ist die Determinante der Matrix in (27.11a) gleich $(-1)^{i+1} [A]_{ij}$, was nach (27.9) die Definition des Kofaktors $\widetilde{a}_{ij}$ bzgl. des Index (i, j) ist.

Die Matrix in (27.11a) geht dadurch aus (27.11b) hervor, dass geeignete Vielfache der j-ten Spalte zu den anderen Spalten addiert werden. Dadurch ändert sich die Determinante nicht. Also gilt die Gleichheit von (27.11b) und (27.11a).

Die Übereinstimmung von (27.11c) und (27.11b) ist klar. □

Beispiel 27.14 (Streichungsmatrix, Minoren, Kofaktoren und Adjunkte).
Die Matrix $A = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right] \in \mathbb{Q}^{2\times 2}$ hat die Kofaktoren

$$\begin{aligned}
\widetilde{a}_{11} &= (-1)^{1+1}[A]_{11} = 4 && \text{oder nach (27.11b)} \quad \widetilde{a}_{11} = \det\begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} = 4 \\
\widetilde{a}_{12} &= (-1)^{1+2}[A]_{12} = -3 && \text{oder nach (27.11b)} \quad \widetilde{a}_{12} = \det\begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix} = -3 \\
\widetilde{a}_{21} &= (-1)^{2+1}[A]_{21} = -2 && \text{oder nach (27.11b)} \quad \widetilde{a}_{21} = \det\begin{pmatrix} 0 & 2 \\ 1 & 4 \end{pmatrix} = -2 \\
\widetilde{a}_{22} &= (-1)^{2+2}[A]_{22} = 1 && \text{oder nach (27.11b)} \quad \widetilde{a}_{22} = \det\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} = 1.
\end{aligned}$$

Die Kofaktormatrix von A ist daher $\operatorname{cof}(A) = \left[\begin{smallmatrix} 4 & -3 \\ -2 & 1 \end{smallmatrix}\right]$, und die Adjunkte von A ist $\operatorname{adj}(A) = \left[\begin{smallmatrix} 4 & -2 \\ -3 & 1 \end{smallmatrix}\right]$. △

Lemma 27.15 (Bedeutung der Adjunkten).
Es seien K ein Körper und $A \in K^{n\times n}$, $n \in \mathbb{N}_0$. Dann gilt

$$\operatorname{adj}(A)\, A = A \operatorname{adj}(A) = \det(A)\, I. \tag{27.12}$$

Beachte: Bis auf den Faktor $\det(A)$ ist die Adjunkte also die Inverse von A.

Beweis. Wir betrachten den Eintrag (i, k) der Matrix $B \coloneqq \operatorname{adj}(A)\, A$. Es gilt

$$\begin{aligned}
b_{ik} &= \sum_{j=1}^n \operatorname{adj}(A)_{ij}\, a_{jk} && \text{nach Definition des Matrix-Matrix-Produkts} \\
&= \sum_{j=1}^n \widetilde{a}_{ji}\, a_{jk} && \text{nach Definition der Adjunkte} \\
&= \sum_{j=1}^n a_{jk} \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, e_j, a_{\bullet i+1}, \ldots, a_{\bullet n}) && \text{nach (27.11c)} \\
&= \det\Big(a_{\bullet 1}, \ldots, a_{\bullet i-1}, \sum_{j=1}^n a_{jk}\, e_j, a_{\bullet i+1}, \ldots, a_{\bullet n}\Big) && \text{wegen der Multilinearität der Determinante} \\
&= \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, a_{\bullet k}, a_{\bullet i+1}, \ldots, a_{\bullet n}).
\end{aligned}$$

Dieser Ausdruck ist gleich 0, wenn $i \neq k$ ist, weil die Determinante alternierend ist und sich dann einer der Vektoren wiederholt. Wenn jedoch $i = k$ gilt, dann ist der Ausdruck gleich $\det(A)$. Das bedeutet $B = \operatorname{adj}(A)\, A = \det(A)\, I$.

Die zweite Behauptung $A\,\mathrm{adj}(A) = \det(A)\,I$ folgt, weil $\frac{1}{\det(A)}\,\mathrm{adj}(A)$ als Linksinverse von A gleichzeitig notwendig auch Rechtsinverse von A ist, siehe Satz 15.47. □

Beispiel 27.16 (Inverse einer 2×2-Matrix)**.**
Die Inverse einer invertierbaren 2×2-Matrix über einem Körper K

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

ist nach Lemma 27.15 gegeben durch $A^{-1} = \frac{1}{\det(A)}\,\mathrm{adj}(A)$, also explizit durch

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \frac{1}{a_{11}\,a_{22} - a_{21}\,a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$ △

Neben der Transformation auf Zeilenstufenform (Beispiel 27.11) gibt es eine weitere Möglichkeit, Determinanten zu berechnen. Diese bietet sich insbesondere dann an, wenn in einer Zeile oder in einer Spalte der Matrix viele Nulleinträge stehen.

Satz 27.17 (**Entwicklungssatz von Laplace**[45])**.**
Es seien K ein Körper und $A \in K^{n \times n}$, $n \in \mathbb{N}_0$. Dann gelten die folgenden Identitäten:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\, \det\big((A)_{\neq i, \neq j}\big) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\, [A]_{ij}. \tag{27.13a}$$

Diese wird auch mit dem Begriff „**Entwicklung nach der i-ten Zeile**“ (englisch: expansion by the i-th row) bezeichnet. Weiter gilt

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij}\, \det\big((A)_{\neq i, \neq j}\big) = \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij}\, [A]_{ij}, \tag{27.13b}$$

was als **Entwicklung nach der j-ten Spalte** (englisch: expansion by the j-th column) bezeichnet wird.

Beachte: Der Entwicklungssatz ersetzt also die Berechnung der Determinante einer $n \times n$-Matrix durch die Berechnung der Determinanten von $(n-1) \times (n-1)$-Untermatrizen.

Beweis. Wir zeigen (27.13b). Für festes $1 \leqslant j \leqslant n$ gilt

$$\begin{aligned} \det(A) &= \det(a_{\bullet 1}, \ldots, a_{\bullet j-1}, a_{\bullet j}, a_{\bullet j+1}, \ldots, a_{\bullet n}) \\ &= \sum_{i=1}^{n} a_{ij}\, \det(a_{\bullet 1}, \ldots, a_{\bullet j-1}, e_i, a_{\bullet j+1}, \ldots, a_{\bullet n}) && \text{wegen der Multilinearität} \\ &= \sum_{i=1}^{n} a_{ij}\, \widetilde{a}_{ij} && \text{nach (27.11c)} \\ &= \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij}\, [A]_{ij} && \text{nach Definition (27.9) des Kofaktors.} \end{aligned}$$

Damit ist (27.13b) gezeigt. Gleichung (27.13a) folgt analog. □

[45] englisch: Laplace expansion theorem

Beispiel 27.18 (Entwicklungssatz von Laplace).
Wir betrachten die Matrix $A \in \mathbb{Q}^{3\times 3}$ mit

$$A = \begin{bmatrix} -2 & - & 2 & 3 \\ -1 & + & 1 & 1 \\ -4 & - & 0 & 1 \end{bmatrix}.$$

Wir entwickeln wie oben bereits angedeutet die Determinante nach der zweiten Spalte ($j = 2$), weil dort zumindest eine Null auftaucht:[46]

$$\begin{aligned}
\det(A) &= -a_{12} \det\big((A)_{\neq 1,\neq 2}\big) + a_{22} \det\big((A)_{\neq 2,\neq 2}\big) - a_{32} \det\big((A)_{\neq 3,\neq 2}\big) \\
&= -2 \det \begin{pmatrix} -1 & 1 \\ -4 & 1 \end{pmatrix} + 1 \det \begin{pmatrix} -2 & 3 \\ -4 & 1 \end{pmatrix} - 0 \det \begin{pmatrix} -2 & 3 \\ -1 & 1 \end{pmatrix} \\
&= -2\,(-1+4) + 1\,(-2+12) + 0 \\
&= -6 + 10 \\
&= 4.
\end{aligned}$$

△

Wir schließen diesen Abschnitt mit der Cramerschen Regel, die es erlaubt, einzelne Einträge des Lösungsvektors eines linearen Gleichungssystems zu bestimmen, was mit dem Gaußschen Eliminationsverfahren aus § 16 nicht möglich ist.

Satz 27.19 (**Cramersche Regel**[47]).
Es seien K ein Körper und $A \in K^{n\times n}$, $n \in \mathbb{N}_0$. Weiter sei $b \in K^n$. Ist A invertierbar, dann gilt für die Koordinaten der eindeutigen Lösung $x \in K^n$ des linearen Gleichungssystems $Ax = b$

$$x_i = \frac{\det(a_{\bullet 1}, \dots, a_{\bullet i-1}, b, a_{\bullet i+1}, \dots, a_{\bullet n})}{\det(A)} \quad \text{für } i = 1, \dots, n. \tag{27.14}$$

Beweis. Nach Voraussetzung ist A invertierbar, also $\det(A) \neq 0$ (Lemma 27.8) und das Gleichungssystem $Ax = b$ eindeutig lösbar (Satz 16.3). Wir müssen nur nachweisen, dass die behauptete Lösung stimmt, dass also $\det(A)\,A\,x = \det(A)\,b$ gilt mit dem Vektor $x \in K^n$ und den behaupteten Einträgen x_i. Es gilt

$$\begin{aligned}
&\det(A)\,A\,x \\
&\quad = \det(A) \sum_{i=1}^n a_{\bullet i}\, x_i && \text{nach Definition des Matrix-Vektor-Produkts} \\
&\quad = \sum_{i=1}^n a_{\bullet i}\, \det(a_{\bullet 1}, \dots, a_{\bullet i-1}, b, a_{\bullet i+1}, \dots, a_{\bullet n}) && \text{nach Definition (27.14) von } x_i \\
&\quad = \sum_{i=1}^n a_{\bullet i} \sum_{j=1}^n b_j \det(a_{\bullet 1}, \dots, a_{\bullet i-1}, e_j, a_{\bullet i+1}, \dots, a_{\bullet n}) && \text{wegen der Multilinearität} \\
&\quad = \sum_{i=1}^n a_{\bullet i} \sum_{j=1}^n b_j\, \widetilde{a}_{ji} && \text{nach Darstellung (27.11c) des Kofaktors}
\end{aligned}$$

[46] Alternative: Entwicklung nach der dritten Zeile ($i = 3$)
[47] englisch: Cramer's rule

$$
\begin{aligned}
&= \sum_{j=1}^{n} b_j \sum_{i=1}^{n} a_{\bullet i}\, \widetilde{a}_{ji} && \text{nach Umsortieren} \\
&= \sum_{j=1}^{n} b_j \sum_{i=1}^{n} a_{\bullet i}\, (\operatorname{adj}(A))_{ij} && \text{nach Definition (27.10) der Adjunkte} \\
&= \sum_{j=1}^{n} b_j\, e_j\, \det(A) && \text{da nach (27.12) } A \operatorname{adj}(A) = \det(A)\, I \text{ gilt} \\
&= \det(A)\, b.
\end{aligned}
$$

□

Beispiel 27.20 (Cramersche Regel)**.**

Wir betrachten das lineare Gleichungssystem $A\,x = b$ mit

$$
A = \begin{pmatrix} -2 & 2 & 3 \\ -1 & 1 & 1 \\ -4 & 0 & 1 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}
$$

über dem Körper $\mathbb{Q}$. Es gilt $\det(A) = 4$, die Matrix A ist also regulär, vgl. Beispiel 27.18.

Wir bestimmen die Lösung des Gleichungssystems mittels Cramerscher Regel:

$$
\begin{aligned}
x_1 &= \frac{1}{\det(A)} \det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{4}(1-4) = -\frac{3}{4} \\
x_2 &= \frac{1}{\det(A)} \det \begin{pmatrix} -2 & 1 & 3 \\ -1 & 2 & 1 \\ -4 & 0 & 1 \end{pmatrix} = \frac{1}{4}(-4-4+24+1) = \frac{17}{4} \\
x_3 &= \frac{1}{\det(A)} \det \begin{pmatrix} -2 & 2 & 1 \\ -1 & 1 & 2 \\ -4 & 0 & 0 \end{pmatrix} = \frac{1}{4}(-16+4) = -3.
\end{aligned}
$$

Wir erhalten also die Lösung $x = \frac{1}{4}\begin{pmatrix} -3 \\ 17 \\ -12 \end{pmatrix}$. △

Die Cramersche Regel ist zur praktischen Lösung linearer Gleichungssysteme (oder zur Bestimmung einzelner Koordinaten des Lösungsvektors) nur für geringe Dimensionen geeignet oder aber dann, wenn die Koeffizientenmatrix A in einer Zeile oder Spalte viele Nulleinträge aufweist, sodass die benötigten Determinanten günstig mit Hilfe des Laplaceschen Entwicklungssatzes 27.17 berechnet werden können. Abseits dieser Fälle findet die Cramersche Regel auch als Mittel in Beweisen weitere Anwendung.

§ 27.3 Die Determinante eines Endomorphismus

Wir kommen zurück zu der eingangs von § 27 angeführten Motivation, die Determinante als Maßzahl für Endomorphismen $f \in \operatorname{Endo}(V)$ auf endlich-dimensionalen Vektorräumen V einzuführen.

Definition 27.21 (Determinante auf dem Raum $\operatorname{Endo}(V)$ der Endomorphismen)**.**
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B_V = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$. Die **Determinante** eines Endomorphismus $f \in \operatorname{Endo}(V)$ ist definiert als

$$\det(f) \coloneqq \det(A) \tag{27.15}$$

für die Darstellungsmatrix $A = \mathcal{M}_{B_V \leftarrow B_V}(f)$. △

Es stellt sich sofort die Frage, ob die Determinante eines Endomorphismus wohldefiniert ist, also für jede Wahl der Basis B_V von V übereinstimmt. Das ist aber der Fall: Ist $\widehat{B}_V$ eine weitere Basis von V, dann gilt mit der Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V \leftarrow B_V}$

$$\begin{aligned} \widehat{A} = \mathcal{M}_{\widehat{B}_V \leftarrow \widehat{B}_V}(f) &= \mathcal{T}_{\widehat{B}_V \leftarrow B_V}\, \mathcal{M}_{B_V \leftarrow B_V}(f)\, \mathcal{T}_{B_V \leftarrow \widehat{B}_V} && \text{nach Satz 19.20} \\ &= T\,A\,T^{-1} && \text{nach Lemma 19.16.} \end{aligned}$$

Aus Lemma 27.8 folgt jetzt

$$\det(\widehat{A}) = \det(T\,A\,T^{-1}) = \det(T)\,\det(A)\,\det(T^{-1}) = \det(A).$$

Daher ist tatsächlich unerheblich, welche Basis für die Darstellungsmatrix in der Definition (27.15) der Determinante eines Endomorphismus verwendet wird.

Lemma 27.22 (Eigenschaften der Determinante für Endormorphismen, vgl. Lemma 27.8)**.**
Es seien K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$ und Basis $B = (v_1, \dots, v_n)$. Weiter seien $f, g \in \operatorname{Endo}(f)$. Die Determinante $\det\colon \operatorname{Endo}(f) \to K$ hat folgende Eigenschaften:

(i) $\det(f)$ ist eine alternierende Multilinearform auf den Vektoren $(f(v_1), \dots, f(v_n))$.

(ii) $\det(\alpha f) = \alpha^n \det(f)$ für alle $\alpha \in K$.

(iii) $\det(f \circ g) = \det(f)\,\det(g)$.

(iv) $\det(\operatorname{id}_V) = 1$.

(v) $\det(f) \neq 0 \Leftrightarrow f$ ist invertierbar $\Leftrightarrow \operatorname{Rang}(f) = n \Leftrightarrow (f(v_1), \dots, f(v_n))$ ist linear unabh. (Die Determinante kann also als Test für die Invertierbarkeit eines Endomorphismus verwendet werden.)

(vi) $\det(f^{-1}) = 1/\det(f)$, falls f invertierbar ist.

(vii) $\det(f^*) = \det(f)$ für die zu f duale Abbildung $f^* \in \operatorname{Endo}(V^*)$.

Beachte: Statt „f ist invertierbar“ können wir auch „f ist ein Automorphismus“, also „$f \in \operatorname{Auto}(V)$“ schreiben.

Beweis. Der Beweis ist Gegenstand der Übung. □

§ 27.4 Orientierung eines Vektorraumes

Definition 27.23 (orientierungstreuer Endomorphismus).

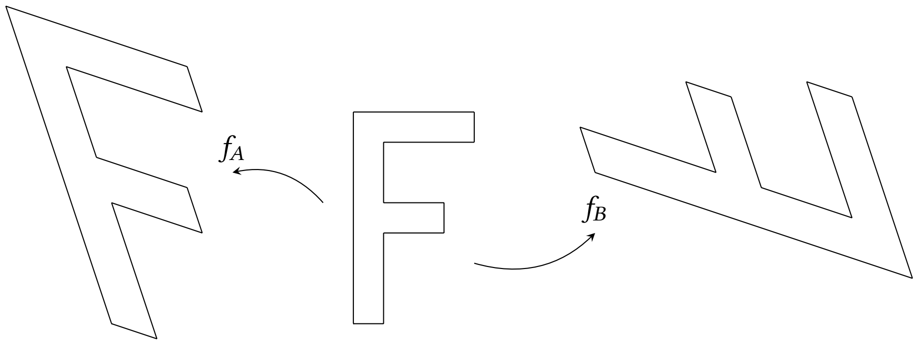
Es seien K ein geordneter Körper (Definition 10.19) und V ein endlich-dimensionaler Vektorraum über K. Ein Automorphismus $f \in \operatorname{Auto}(V)$ heißt **orientierungstreu** oder **orientierungserhaltend** (englisch: orientation preserving) im Fall $\det(f) > 0$ und **orientierungsumkehrend** (englisch: orientation reversing) im Fall $\det(f) < 0$. △

Beispiel 27.24 (orientierungstreuer Endomorphismus).

Wir betrachten die durch die Matrizen

$$A = \frac{1}{2}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \quad \text{und} \quad B = \frac{1}{2}\begin{pmatrix} -1 & 3 \\ 3 & -1 \end{pmatrix}$$

induzierten Automorphismen auf dem Vektorraum $\mathbb{R}^2$. Es gilt $\det(A) = 2 > 0$ und $\det(B) = -2 < 0$. Die Wirkung der orientierungstreuen Abbildung f_A und der orientierungsumkehrenden Abbildung f_B auf Punkte, die den Umriss des Buchstabens „F" ergeben, zeigt folgende Abbildung:



△

Definition 27.25 (gleich orientierte Basen).

Es seien K ein geordneter Körper und V ein endlich-dimensionaler Vektorraum über K.

(*i*) Zwei Basen $B_V = (v_1, \dots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \dots, \widehat{v}_n)$ heißen **gleich orientiert** (englisch: equally oriented), wenn die Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V \leftarrow B_V}$ die Bedingung $\det(T) > 0$ erfüllt.

(*ii*) Zwei Basen $B_V = (v_1, \dots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \dots, \widehat{v}_n)$ heißen **umgekehrt orientiert** (englisch: differently oriented), wenn die Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V \leftarrow B_V}$ die Bedingung $\det(T) < 0$ erfüllt. △

Lemma 27.26 (Gleichorientierung ist eine Äquivalenzrelation).

Es seien K ein geordneter Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Gleichorientierung ist eine Äquivalenzrelation auf der Menge aller Basen von V.

Beweis. Jede Basis ist zu sich selbst gleich orientiert, da die Transformationsmatrix dann die Einheitsmatrix ist und $\det(I) = 1 > 0$ gilt. Das ist die Reflexivität der Gleichorientierung. Die Symmetrie folgt, da mit $T = \mathcal{T}_{\widehat{B}_V \leftarrow B_V}$ und $\det(T) > 0$ auch $T^{-1} = \mathcal{T}_{B_V \leftarrow \widehat{B}_V}$ die Bedingung $\det(T^{-1}) = \det(T)^{-1} > 0$ erfüllt. Schließlich folgt die Transitivität aus der Tatsache, dass die Determinante des Produkts von Matrizen das Produkt der Determinanten ist. □

Beachte: Es gibt (für $n \geqslant 1$) genau zwei Äquivalenzklassen, da für geordnete Körper $\operatorname{char}(K) = 0$ und insbesondere $\operatorname{char}(K) \neq 2$ gilt (Lemma 10.20). Auch im Fall $n = 0$ definiert man aus Konsistenzgründen typischerweise zwei Orientierungen.

Jede der beiden Äquivalenzklassen von Basen eines Vektorraumes V mit $\dim(V) \in \mathbb{N}$ über einem geordneten Körper K wird als eine **Orientierung** (englisch: orientation) des Vektorraumes V bezeichnet. Oft wird eine der beiden die **positive Orientierung** (englisch: positive orientation) und die andere die **negative Orientierung** (englisch: negative orientation) des Vektorraumes genannt. Die Festlegung, welche welche ist, ist allerdings willkürlich, also nicht kanonisch.

Ende der Vorlesung 11

Ende der Woche 5

Kapitel A Zur Konstruktion der Zahlen

Literatur: Goldrei, 1996, Chapters 2–3

Die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$ oder $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ sind ausreichend, um Objekte zu zählen. Gleichungen wie $x + 2 = 1$ sind jedoch in $\mathbb{N}_0$ nicht lösbar. Dafür benötigen wir die Menge der ganzen Zahlen $\mathbb{Z}$. Gleichungen wie $2\,x = 1$ sind aber auch in $\mathbb{Z}$ nicht lösbar. Dafür benötigen wir die Menge der rationalen Zahlen $\mathbb{Q}$. Gleichungen wie $x^2 = 2$ sind aber auch in $\mathbb{Q}$ nicht lösbar. Dafür benötigen wir die Menge der reellen Zahlen $\mathbb{R}$. Gleichungen wie $x^2 = -1$ sind aber auch in $\mathbb{R}$ nicht lösbar. Dafür benötigen wir die Menge der komplexen Zahlen $\mathbb{C}$. Wir deuten in diesem Abschnitt kurz an, wie die die Zahlbereiche (4.1) von den natürlichen bis zu den komplexen Zahlen mitsamt ihren Verknüpfungen Addition und Multiplikation mathematisch fundiert konstruiert werden können.

Die natürlichen Zahlen

Die **natürlichen Zahlen** (englisch: natural numbers) können auf verschiedene Arten konstruiert werden, beispielsweise — unabhängig von den Axiomen der Zermelo-Fraenkel-Mengenlehre — mit Hilfe der Peano-Axiome. Alternativ können wir die natürlichen Zahlen $\mathbb{N}_0$ aber auch innerhalb der Zermelo-Fraenkel-Mengenlehre abbilden. Von Neumann setzt dazu

$$\begin{aligned}
0 &\coloneqq \emptyset \\
1 &\coloneqq 0 \cup \{0\} = \{\emptyset\} \\
2 &\coloneqq 1 \cup \{1\} = \{\emptyset, \{\emptyset\}\} \\
3 &\coloneqq 2 \cup \{2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}
\end{aligned}$$

usw. Gleichzeitig mit diesen „Zahlen" wird die Funktion **Nachfolger** (englisch: successor) $S(n) \coloneqq n \cup \{n\}$ definiert. Die Menge $\mathbb{N}_0$ ist dann der Durchschnitt aller Mengen, die die 0 (also $\emptyset$) enthalten und die die Eigenschaft haben, dass sie ihr Bild unter der Nachfolgerfunktion enthalten:

$$\mathbb{N}_0 \coloneqq \bigcap \{M \text{ ist Menge} \mid \emptyset \in M \text{ und } S(M) \subseteq M\}.$$

Da jede natürliche Zahl in $\mathbb{N}_0$ entweder gleich 0 oder der Nachfolger einer natürlichen Zahl ist, kann die **Addition** (englisch: addition) rekursiv durch

$$\begin{aligned}
m + 0 &\coloneqq m && \text{für } m \in \mathbb{N}_0 \\
m + S(n) &\coloneqq S(m + n) && \text{für } m, n \in \mathbb{N}_0
\end{aligned}$$

definiert werden. Es kann gezeigt werden, dass die Addition eine assioziative, kommutative Verknüpfung auf $\mathbb{N}_0$ mit neutralem Element 0 ist. Also ist $(\mathbb{N}_0, +)$ ein kommutatives Monoid. Beispielsweise gilt

$$\begin{aligned}
5+2 &= 5+S(1) = S(5+1) = S(5+S(0)) = S(S(5+0)) = S(S(5)) = S(6) = 7 \\
2+5 &= 2+S(4) = S(2+4) = S(2+S(3)) = S(S(2+3)) = S(S(2+S(2))) = S(S(S(2+2))) \\
&= S(S(S(2+S(1)))) = S(S(S(S(2+1)))) = S(S(S(S(2+S(0))))) = S(S(S(S(S(2+0))))) \\
&= S(S(S(S(S(2))))) = S(S(S(S(3)))) = S(S(S(4))) = S(S(5)) = S(6) = 7.
\end{aligned}$$

Aufbauend auf der Addition kann dann die **Multiplikation** (englisch: multiplication) rekursiv durch

$$\begin{aligned}
m \cdot 0 &\coloneqq 0 && \text{für } m \in \mathbb{N}_0 \\
m \cdot S(n) &\coloneqq (m \cdot n) + m && \text{für } m, n \in \mathbb{N}_0
\end{aligned}$$

definiert werden. Auch für die Multiplikation kann gezeigt werden, dass sie eine assioziative, kommutative Verknüpfung auf $\mathbb{N}_0$ ist, und zwar mit dem neutralen Element 1. Also ist auch $(\mathbb{N}_0, \cdot)$ ein kommutatives Monoid.

Nun kann noch das Distributivgesetz $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ für alle $a, b, c \in \mathbb{N}_0$ bestätigt werden.

Mit Hilfe der Addition kann die natürliche Totalordnung $\leqslant$ auf $\mathbb{N}_0$ definiert werden, und zwar durch

$$a \leqslant b \quad \Leftrightarrow \quad \exists n \in \mathbb{N}_0 \, (b = a + n).$$

Für alle $b \in \mathbb{N}_0$ gilt $b \geqslant 0$. Wie leicht zu sehen ist, ist diese Ordnung kompatibel mit der Addition und Multiplikation, es gilt also $a \leqslant b \Rightarrow a + c \leqslant b + c$ und $a \leqslant b \Rightarrow a \cdot c \leqslant b \cdot c$ für alle $c \in \mathbb{N}_0$.

Durch Entfernen der 0 aus der Menge $\mathbb{N}_0$ erhalten wir die kommutative Halbgruppe $(\mathbb{N}, +)$ bzw. das kommutative Monoid $(\mathbb{N}, \cdot)$.

Die ganzen Zahlen

Die **ganzen Zahlen** (englisch: integer numbers) können wir aus den natürlichen Zahlen $\mathbb{N}_0$ konstruieren, beispielsweise durch

$$\widetilde{\mathbb{Z}} \coloneqq \mathbb{N}_0 \times \mathbb{N}_0 / \sim$$

bzgl. der Äquivalenzrelation

$$(a, b) \sim (c, d) \quad \Leftrightarrow \quad a + d = b + c$$

auf $\mathbb{N}_0 \times \mathbb{N}_0$. Wir führen auf $\widetilde{\mathbb{Z}}$ durch

$$[(a, b)] + [(c, d)] \coloneqq [(a + c, b + d)]$$

die **Addition** ein. Für diese kann gezeigt werden, dass sie wohldefiniert, assoziativ und kommutativ ist und das neutrale Element $[(0,0)]$ besitzt. Damit ist $(\widetilde{\mathbb{Z}},+)$ ein kommutatives Monoid. Tatsächlich ist $(\widetilde{\mathbb{Z}},+)$ sogar eine abelsche Gruppe, denn $[(b,a)]$ ist wegen

$$[(a,b)] + [(b,a)] = [(a+b,b+a)] = [(a+b,a+b)] = [(0,0)]$$

das additive Inverse zu $[(a,b)]$.

Jede Äquivalenzklasse $[(a,b)]$ enthält entweder genau ein Element der Form $(c,0)$ mit $c \in \mathbb{N}$ oder genau ein Element der Form $(0,d)$ mit $d \in \mathbb{N}$ oder das Element $(0,0)$. Im ersten Fall bezeichnen wir die Äquivalenzklasse $[(c,0)]$ kurz mit c. Im zweiten Fall bezeichnen wir die Äquivalenzklasse $[(0,d)]$ kurz mit $-d$. Im dritten Fall bezeichnen wir die Äquivalenzklasse $[(0,0)]$ kurz mit 0. Durch diese bijektive Abbildung $\varphi\colon \widetilde{\mathbb{Z}} \to \mathbb{Z} \coloneqq \{0,1,-1,2,-2,\dots\}$ können wir auf der Repräsentantenmenge $\mathbb{Z}$ die Addition erklären, sodass $(\mathbb{Z},+)$ eine abelsche Gruppe mit dem neutralen Element 0 wird. Diese ist isomorph zu $\widetilde{\mathbb{Z}}$. Beispielsweise gilt

$$\begin{aligned}
2+3 &= [(2,0)] + [(3,0)] = [(5,0)] = 5 \\
2+(-3) &= [(2,0)] + [(0,3)] = [(2,3)] = [(0,1)] = -1 \\
2-(-3) &= [(2,0)] - [(0,3)] = [(2,0)] + [(3,0)] = [(5,0)] = 5.
\end{aligned}$$

Die zuvor definierte Addition in $\mathbb{N}_0$ stimmt dann mit der Addition in $\mathbb{Z}$, eingeschränkt auf $\mathbb{N}_0$, überein. Formal wird das Monoid $(\mathbb{N}_0,+)$ mit Hilfe des injektiven Monoidhomomorphismus $\mathbb{N}_0 \ni a \mapsto a \ni \mathbb{Z}$ in die Gruppe $(\mathbb{Z},+)$ eingebettet. So werden die natürlichen Zahlen zu einer Teilmenge der ganzen Zahlen und die Addition auf $\mathbb{Z}$ zu einer Fortsetzung der Addition auf $\mathbb{N}_0$.

Die **Multiplikation** kann nun ebenfalls zunächst wieder auf $\widetilde{\mathbb{Z}}$ definiert werden, und zwar (mit Hilfe der Addition und Multiplikation in $\mathbb{N}_0$) durch

$$[(a,b)] \cdot [(c,d)] \coloneqq [(a\cdot c + b\cdot d, a\cdot d + b\cdot c)].$$

Auch hier kann die Wohldefiniertheit, Assoziativität und Kommutativität gezeigt werden. Das neutrale Element ist $[(1,0)]$. Anschließend können wir wieder mit Hilfe der Bijektion $\varphi\colon \widetilde{\mathbb{Z}} \to \mathbb{Z}$ die Multiplikation auf $\mathbb{Z}$ erklären. Alternativ können wir die Multiplikation auch direkt nur auf $\mathbb{Z}$ definieren, indem wir für $a \cdot b$ eine Fallunterscheidung nach dem Vorzeichen von a und b vornehmen. (Damit ist die Fallunterscheidung $a \geqslant 0$ oder $a \leqslant 0$ bzw. $b \geqslant 0$ oder $b \leqslant 0$ bzgl. der gleich erläuterten Ordnung auf $\mathbb{Z}$ gemeint.) In jedem Fall ergibt sich, dass $(\mathbb{Z},\cdot)$ eine kommutative Halbgruppe mit neutralem Element 1 ist und dass $(\mathbb{N}_0,\cdot)$ in $(\mathbb{Z},\cdot)$ eingebettet ist.

Abschließend kann noch das Distributivgesetz $a\cdot(b+c) = (a\cdot b) + (a\cdot c)$ gezeigt werden, wodurch $(\mathbb{Z},+,\cdot)$ zu einem kommutativen Ring (Definition 9.1) mit dem Einselement 1 wird.

Die natürliche Totalordnung $\leqslant$ auf $\mathbb{N}_0$ kann auf $\mathbb{Z}$ fortgesetzt werden, und zwar wiederum durch

$$a \leqslant b \quad\Leftrightarrow\quad \exists n \in \mathbb{N}_0\,(b = a+n).$$

Es gilt $a \leqslant b \Rightarrow a+c \leqslant b+c$ für alle $c \in \mathbb{Z}$ sowie $a \leqslant b \Rightarrow a\cdot c \leqslant b\cdot c$ für alle $c \in \mathbb{N}_0$ und $a \leqslant b \Rightarrow a\cdot c \geqslant b\cdot c$ für alle $c \in -\mathbb{N}_0$.

Der **Betrag** $|a|$ einer ganzen Zahl a ist definiert als

$$|a| \coloneqq \begin{cases} a & \text{falls } a \geqslant 0, \\ -a & \text{falls } a < 0. \end{cases}$$

Die rationalen Zahlen

Die **rationalen Zahlen** (englisch: rational numbers) können wir aus den ganzen Zahlen $\mathbb{Z}$ konstruieren durch

$$\mathbb{Q} \coloneqq \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \,/\, \sim$$

mit der Äquivalenzrelation

$$(a, b) \sim (c, d) \quad \Leftrightarrow \quad a \cdot d = c \cdot b$$

auf $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. Auf $\mathbb{Q}$ führen wir durch

$$[(a, b)] + [(c, d)] \coloneqq [(a \cdot d + b \cdot c, b \cdot d)]$$

die **Addition** ein. Für diese kann gezeigt werden, dass sie wohldefiniert, assoziativ und kommutativ ist und das neutrale Element $[(0, 1)]$ besitzt. Damit ist $(\mathbb{Q}, +)$ ein kommutatives Monoid. Tatsächlich ist $(\mathbb{Q}, +)$ sogar eine abelsche Gruppe, denn $[(-a, b)]$ ist wegen

$$[(a, b)] + [(-a, b)] = [(a \cdot b + b \cdot (-a), b \cdot b)] = [(0, b \cdot b)] = [(0, 1)]$$

das additive Inverse zu $[(a, b)]$. Auch hier gilt, dass $\mathbb{Z}$ durch $a \mapsto [(a, 1)]$ in $\mathbb{Q}$ eingebettet ist und dass die Addition auf $\mathbb{Q}$, eingeschränkt auf $\mathbb{Z} \times \{1\}$, mit der Addition auf $\mathbb{Z}$ übereinstimmt.

Die **Multiplikation** auf $\mathbb{Q}$ wird definiert durch

$$[(a, b)] \cdot [(c, d)] \coloneqq [(a \cdot c, b \cdot d)].$$

Auch diese ist wohldefiniert, assoziativ und kommutativ mit neutralem Element $[(1, 1)]$. Damit ist $(\mathbb{Q}, \cdot)$ ein kommutatives Monoid. Tatsächlich ist $(\mathbb{Q} \setminus \{[(0, 1)]\}, \cdot)$ sogar eine abelsche Gruppe, denn $[(b, a)]$ ist wegen

$$[(a, b)] \cdot [(b, a)] = [(a \cdot b, b \cdot a)] = [(a \cdot b, a \cdot b)] = [(1, 1)]$$

das multiplikative Inverse zu $[(a, b)]$. Insbesondere ist $[(1, a)]$ das multiplikative Inverse zu $[(a, 1)]$ für alle $a \in \mathbb{Z} \setminus \{0\}$. Auch hier wird durch $a \mapsto [(a, 1)]$ die Menge $\mathbb{Z}$ in $\mathbb{Q}$ eingebettet und die Multiplikation auf $\mathbb{Q}$, eingeschränkt auf $\mathbb{Z} \times \{1\}$, stimmt mit der Multiplikation auf $\mathbb{Z}$ überein.

Weiterhin kann das Distributivgesetz gezeigt werden, wodurch $(\mathbb{Q}, +, \cdot)$ zu einem Körper wird (Definition 10.1). Es ist üblich, die Äquivalenzklasse $[(a, b)]$ in der Gestalt eines „Bruches“ $\frac{a}{b}$ zu notieren, vgl. (5.18). Durch „Kürzen“ und „Erweitern“ können wir jederzeit den Repräsentanten einer Äquivalenzklasse wechseln. Insbesondere kann ein Bruch immer in der Form $\frac{a}{b}$ mit „Zähler“ $a \in \mathbb{Z}$ und „Nenner“ $b \in \mathbb{N}$ dargestellt werden.

Die natürliche Totalordnung $\leqslant$ auf $\mathbb{Z}$ kann auf $\mathbb{Q}$ fortgesetzt werden, und zwar für die Darstellungen mit $b, d \in \mathbb{N}$ durch

$$\frac{a}{b} \leqslant \frac{c}{d} \quad \Leftrightarrow \quad a \cdot d \leqslant c \cdot b \quad \text{(im Sinne der Totalordnung $\leqslant$ auf $\mathbb{Z}$).}$$

Es gilt $a \leqslant b \Rightarrow a + c \leqslant b + c$ für alle $c \in \mathbb{Q}$ sowie $a \geqslant 0,\ b \geqslant 0 \Rightarrow a \cdot b \geqslant 0$. Damit wird $(\mathbb{Q}, +, \cdot)$ zu einem geordneten Körper (Definition 10.19).

Der **Betrag** $\left|\frac{a}{b}\right|$ einer rationalen Zahl $\frac{a}{b}$ mit $a \in \mathbb{Z}$ und $b \in \mathbb{Z} \setminus \{0\}$ ist definiert als

$$\left|\frac{a}{b}\right| \coloneqq \begin{cases} \frac{a}{b} & \text{falls } \frac{a}{b} \geqslant 0, \\ -\frac{a}{b} & \text{falls } \frac{a}{b} < 0. \end{cases}$$

Die reellen Zahlen

Es gibt verschiedene Möglichkeiten, die **reellen Zahlen** (englisch: real numbers) aus den rationalen Zahlen $\mathbb{Q}$ zu konstruieren. Wir skizzieren hier die Konstruktion über Dedekindsche Schnitte. Ein **Dedekindscher Schnitt** (englisch: Dedekind cut) ist ein Paar (A, B) von Mengen $A, B \subseteq \mathbb{Q}$ mit den folgenden Eigenschaften:

(1) $\{A, B\}$ bildet eine Partition von $\mathbb{Q}$,

(2) $y \in A$ und $x \in \mathbb{Q}$, $x \leqslant y$ impliziert $x \in A$,
(A ist nach unten abgeschlossen.)

(3) für alle $x \in A$ existiert $y \in A$ mit $y > x$.
(A besitzt kein Maximum.)

Da A und B eine Partition von $\mathbb{Q}$ bilden, kann Eigenschaft (2) auch in der Form $a < b$ für alle $a \in A$ und alle $b \in B$ ausgedrückt werden. (**Quizfrage A.1:** Klar?)

Wir definieren nun die reellen Zahlen als

$$\mathbb{R} \coloneqq \{A \subseteq \mathbb{Q} \mid (A, B) \text{ ist ein Dedekindscher Schnitt}\}.$$

Mit anderen Worten bestehen die reellen Zahlen also gerade aus den sogenannten **linken Abschnitten** A (englisch: Dedekind left sets) der Dedekindschen Schnitte (A, B).

Auf $\mathbb{R}$ führen wir durch

$$A_1 + A_2 \coloneqq \{a_1 + a_2 \mid a_1 \in A_1,\ a_2 \in A_2\}$$

die **Addition** ein. Für diese kann gezeigt werden, dass sie wohldefiniert, assoziativ und kommutativ ist und das neutrale Element $\mathbb{Q}_{<0} \coloneqq \{x \in \mathbb{Q} \mid x < 0\}$ („Null“) besitzt. Damit ist $(\mathbb{R}, +)$ ein kommutatives Monoid. Tatsächlich ist $(\mathbb{R}, +)$ sogar eine abelsche Gruppe, denn es kann gezeigt werden, dass

$$\begin{aligned} -A &= \mathbb{Q}_{<0} - (\mathbb{Q} \setminus A) \\ &= \{x - b \mid x \in \mathbb{Q}_{<0} \text{ und } b \in \mathbb{Q} \setminus A\} \\ &= \{y \in \mathbb{Q} \mid \exists b \in \mathbb{Q} \setminus A\, (y < -b)\} \end{aligned}$$

das additive Inverse zu A ist.

Bevor die Multiplikation auf $\mathbb{R}$ eingeführt werden kann, definieren wir die Ordnungsrelation $A_1 \leqslant A_2$ durch $A_1 \subseteq A_2$. Dadurch wird $\mathbb{R}$ zu einer totalgeordneten Menge. Ein Element $A \in \mathbb{R}$ heißt **negativ** (englisch: negative) im Fall $A \subsetneq \mathbb{Q}_{<0}$, **nicht-positiv** (englisch: non-positive) im Fall $A \subseteq \mathbb{Q}_{<0}$, **positiv** (englisch: positive) im Fall $\mathbb{Q}_{<0} \subsetneq A$ und **nicht-negativ** (englisch: non-negative) im Fall $\mathbb{Q}_{<0} \subseteq A$.

Die **Multiplikation** auf $\mathbb{R}$ wird nun fallweise definiert. Wenn A_1 und A_2 beide positiv sind, dann setzen wir

$$A_1 \cdot A_2 \coloneqq \{a_1 \cdot a_2 \mid a_1 \geqslant 0,\ a_2 \geqslant 0,\ a_1 \in A_1,\ a_2 \in A_2\} \cup \mathbb{Q}_{<0}.$$

Ist A_1 negativ und A_2 positiv, dann setzen wir $A_1 \cdot A_2 \coloneqq -((-A_1) \cdot A_2)$. Ist A_1 positiv und A_2 negativ, dann setzen wir $A_1 \cdot A_2 \coloneqq -(A_1 \cdot (-A_2))$. Sind A_1 und A_2 beide negativ, dann setzen wir $A_1 \cdot A_2 \coloneqq ((-A_1) \cdot (-A_2))$. Ist schließlich mindestens einer der beiden linken Abschnitte A_1 oder A_2 gleich $\mathbb{Q}_{<0}$ („Null"), dann setzen wir $A_1 \cdot A_2 \coloneqq \mathbb{Q}_{<0}$. Auch für die Multiplikation kann gezeigt werden, dass sie wohldefiniert, assoziativ und kommutativ ist und das neutrale Element $\{x \in \mathbb{Q} \mid x < 1\}$ besitzt. Damit ist $(\mathbb{R}, \cdot)$ ein kommutatives Monoid. Tatsächlich ist $(\mathbb{R} \setminus \mathbb{Q}_{<0}, \cdot)$ sogar eine abelsche Gruppe.

Weiterhin kann das Distributivgesetz gezeigt werden, wodurch $(\mathbb{R}, +, \cdot)$ zu einem Körper wird (Definition 10.1). Außerdem ist die oben definerte Totalordnung $\leqslant$ auf $\mathbb{R}$ kompatibel mit der Addition und Multiplikation, es gilt also $A_1 \leqslant A_2 \Rightarrow A_1 + A_3 \leqslant A_2 + A_3$ für alle $A_3 \in \mathbb{R}$, und für alle nicht-negativen A_1, A_2 ist auch $A_1 \cdot A_2$ nicht-negativ. Damit wird $(\mathbb{R}, +, \cdot)$ zu einem geordneten Körper (Definition 10.19).

Ist $\mathcal{A}$ eine nach oben beschränkte Teilmenge[1] von $\mathbb{R}$, so besitzt $\mathcal{A}$ ein Supremum in $\mathbb{R}$, und zwar $\sup \mathcal{A} = \bigcup_{A \in \mathcal{A}} A$. Ist $\mathcal{A}$ eine nach unten beschränkte Teilmenge von $\mathbb{R}$, so besitzt $\mathcal{A}$ ein Infimum in $\mathbb{R}$, und zwar $\inf \mathcal{A} = \bigcap_{A \in \mathcal{A}} A$. Die Existenz dieser Suprema und Infima ist der entscheidende Unterschied zwischen den rationalen und den reellen Zahlen.

Eine rationale Zahl $q \in \mathbb{Q}$ wird mit dem linken Abschnitt $A_q \coloneqq \{x \in \mathbb{Q} \mid x < q\}$ identifiziert. Man kann zeigen, dass diese Abbildung $q \mapsto A_q$ und verträglich mit der Addition und Multiplikation auf $\mathbb{Q}$ und $\mathbb{R}$ ist. Mit anderen Worten: Diese Abbildung ist ein Körperhomomorphismus von $(\mathbb{Q}, +, \cdot)$ nach $(\mathbb{R}, +, \cdot)$, durch den $\mathbb{Q}$ in $\mathbb{R}$ eingebettet wird. Dabei können wir ausnutzen, dass Körperhomomorphismen stets injektiv sind (Lemma 10.14).

Neben allen Elementen von $\mathbb{Q}$ besitzen auch bestimmte andere linke Abschnitte Kurzbezeichnungen, wie beispielsweise $\sqrt{2} \coloneqq \{x \in \mathbb{Q} \mid x \cdot x < 2 \text{ oder } x < 0\}$. Linke Abschnitte, die nicht von der Form $\{x \in \mathbb{Q} \mid x < q\}$ für $q \in \mathbb{Q}$ sind, heißen **irrationale Zahlen** (englisch: irrational numbers). Beispiele für (benannte) irrationale Zahlen sind $\sqrt{2}$, die Eulersche Zahl e und die Kreiszahl π.

An dieser Stelle halten wir einmal fest, dass die reellen Zahlen überlicherweise abstrakt als ein geordneter Körper mit der zusätzlichen Eigenschaft definiert werden, dass jede nach oben beschränkte Teilmenge ein Supremum besitzt. Letztere Eigenschaft nennt man auch die **Ordnungsvollständigkeit** (englisch: order completeness). Man kann dann zeigen, dass ein solcher Körper bis auf Isomorphie eindeutig ist. Insofern stellt die oben skizzierte Konstruktion der reellen Zahlen über Dedekindsche Schnitte ein mögliches Modell für die reellen Zahlen dar. Insbesondere zeigt es, dass eine algebraische Struktur mit den genannten Eigenschaften überhaupt existiert. In der Praxis arbeiten wir aber nur mit den abstrakten Eigenschaften des geordneten Körpers $\mathbb{R}$.

Die Ordnungsvollständigkeit der reellen Zahlen sorgt übrigens dafür, dass eine Wiederholung der Konstruktion durch Schnitte in $\mathbb{R}$, also $\{A \subseteq \mathbb{R} \mid (A, B) \text{ ist ein Dedekindscher Schnitt}\}$ wiederum auf einen Körper führt, der isomorph zu $\mathbb{R}$ ist.

Der **Betrag** $|a|$ einer reellen Zahl a ist definiert als

$$|a| \coloneqq \begin{cases} a & \text{falls } a \geqslant 0, \\ -a & \text{falls } a < 0. \end{cases}$$

[1] $\mathcal{A}$ ist also eine Menge linker Abschnitte, und es gibt einen linken Abschnitt A_0, sodass $A \subseteq A_0$ für alle $A \in \mathcal{A}$ gilt.

Schließlich können wir für $n \in \mathbb{N}$ die ***n*-te Wurzel** nicht-negativer reeller Zahlen $a \geqslant 0$ definieren:

$$\sqrt[n]{a} \coloneqq \{x \in \mathbb{Q} \mid \underbrace{x \cdot x \cdots x}_{n\text{-faches Produkt}} < a \text{ oder } x < 0\}.$$

Die **Exponentialfunktion** (englisch: exponential function)

$$\mathbb{R} \ni x \mapsto a^x \in \mathbb{R}_{>0}$$

zur Basis $a \in \mathbb{R}$, $a > 0$ können wir definieren, indem wir zunächst für rationale Exponenten $x = \frac{m}{n} \in \mathbb{Q}$ mit $m \in \mathbb{Z}$ und $n \in \mathbb{N}$

$$a^{\frac{m}{n}} \coloneqq \sqrt[n]{a^m} \in \mathbb{R}$$

und anschließend für beliebiges $x \in \mathbb{R}$

$$a^x \coloneqq \sup\left\{a^q \in \mathbb{R} \,\middle|\, q \in \mathbb{Q},\ q < x\right\}$$

setzen. Mit Mitteln der Analysis kann gezeigt werden, dass die Exponentialabbildung im Fall $a \neq 1$ bijektiv $\mathbb{R} \to \mathbb{R}_{>0}$ abbildet. Ihre Umkehrabbildung ist die **Logarithmusfunktion** (englisch: logarithm function) zur Basis a, notiert als $\log_a \colon \mathbb{R}_{>0} \to \mathbb{R}$. Häufig wird als Basis die **Eulersche Zahl** (englisch: Euler's number) verwendet, die als

$$\mathrm{e} \coloneqq \sup\left\{\left(1 + \frac{1}{n}\right)^n \,\middle|\, n \in \mathbb{N}\right\}$$

definiert werden kann. Die Exponentialfunktion $\mathbb{R} \ni x \mapsto \mathrm{e}^x \in \mathbb{R}_{>0}$ heißt auch die **natürliche Exponentialfunktion** (englisch: natural exponential function) und wird als exp notiert. Wir setzen also $\exp(x) \coloneqq \mathrm{e}^x$.

Die Definition trigonometrischer Funktionen wie der sin- und cos-Funktion erfolgt üblicherweise über Potenzreihen und wird ebenfalls in Lehrveranstaltungen zur *Analysis* behandelt.

Die komplexen Zahlen

Die **komplexen Zahlen** (englisch: complex numbers) können wir aus den rellen Zahlen $\mathbb{R}$ konstruieren durch

$$\mathbb{C} \coloneqq \mathbb{R} \times \mathbb{R}.$$

Auf $\mathbb{C}$ führen wir durch

$$(a, b) + (c, d) \coloneqq (a + c, b + d)$$

die **Addition** ein. Damit erbt $(\mathbb{C}, +)$ die Eigenschaften von $(\mathbb{R}, +)$ als Abelsche Gruppe und besitzt das neutrale Element $(0, 0)$.[2] Das zu (a, b) additive Inverse ist $(-a, -b)$.

Weiterhin führen wir durch

$$(a, b) \cdot (c, d) \coloneqq (a \cdot c - b \cdot d, a \cdot d + b \cdot c)$$

[2] Diese Konstruktion heißt das **(äußere) direkte Produkt** (englisch: (outer) direct product) der Gruppe $(\mathbb{R}, +)$ mit sich selbst.

die **Multiplikation** auf $\mathbb{C}$ ein. Auch diese ist assoziativ und kommutativ mit neutralem Element $(1, 0)$. Damit ist $(\mathbb{C}, \cdot)$ ein kommutatives Monoid. Tatsächlich ist $(\mathbb{C} \setminus \{(0, 0)\}, \cdot)$ sogar eine abelsche Gruppe, denn $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$ ist wegen $a^2 + b^2 > 0$ und

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right) = \left(\frac{a^2 + b^2}{a^2 + b^2}, \frac{a \cdot (-b) + b \cdot a}{a^2 + b^2}\right) = (1, 0)$$

das multiplikative Inverse zu (a, b).

Weiterhin kann das Distributivgesetz gezeigt werden, wodurch $(\mathbb{C}, +, \cdot)$ zu einem Körper wird (Definition 10.1). Es lässt sich jedoch keine Totalordnung auf $\mathbb{C}$ definieren, durch die $(\mathbb{C}, +, \cdot)$ zu einem geordneten Körper wird.

Durch den Körperhomomorphismus $a \mapsto (a, 0)$ in $(\mathbb{C}, +, \cdot)$ wird der Körper $(\mathbb{R}, +, \cdot)$ eingebettet.

Es ist üblich, das Element $(a, b) \in \mathbb{C}$ als $a+b\,i$ zu notieren, wobei $i \coloneqq (0, 1)$ ist und die **imaginäre Einheit** (englisch: imaginary unit) genannt wird. In dieser Darstellung heißt $a \eqqcolon \operatorname{Re}(z) \in \mathbb{R}$ der **Realteil** (englisch: real part) und $b \eqqcolon \operatorname{Im}(z) \in \mathbb{R}$ der **Imaginärteil** (englisch: imaginary part). Weiter gilt für die Addition

$$(a + b\,i) + (c + d\,i) \coloneqq (a + c) + (b + d)\,i$$

und für die Multiplikation

$$(a + b\,i) \cdot (c + d\,i) \coloneqq (a \cdot c - b \cdot d) + (a \cdot d + b \cdot c)\,i.$$

Formal rechnen wir wie in den reellen Zahlen mit dem zusätzlichen Symbol i, für das $i^2 = -1$ gilt.

Die Abbildung der **komplexen Konjugation** (englisch: complex conjugation

$$\mathbb{C} \ni z = a + b\,i \mapsto \overline{z} = a - b\,i \in \mathbb{C}$$

ist ein Körperautomorphismus $(\mathbb{C}, +, \cdot) \longrightarrow (\mathbb{C}, +, \cdot)$. Es gilt also insbesondere

$$\overline{z + w} = \overline{z} + \overline{w} \quad \text{und} \quad \overline{z \cdot w} = \overline{z} \cdot \overline{w}$$

für alle $z, w \in \mathbb{C}$. Dabei heißt $\overline{z}$ die zu z **konjugiert komplexe Zahl** (englisch: complex conjugate).

Mit Hilfe der komplexen Konjugation können wir auch nochmals das multiplikative Inverse von $z = a + b\,i$ bestätigen:

$$\frac{1}{a + b\,i} = \frac{1}{a + b\,i} \cdot \frac{a - b\,i}{a - b\,i} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}\,i.$$

Es gilt weiterhin

$$\begin{aligned}
\frac{z + \overline{z}}{2} &= \frac{a + b\,i + a - b\,i}{2} = a = \operatorname{Re}(z), \\
\frac{z - \overline{z}}{2\,i} &= \frac{a + b\,i - a + b\,i}{2\,i} = b = \operatorname{Im}(z).
\end{aligned}$$

Das Produkt $z \cdot \overline{z}$ einer komplexen Zahl $z = a + b\,i$ mit ihrer konjugiert komplexen Zahl ergibt

$$z \cdot \overline{z} = (a + b\,i) \cdot (a - b\,i) = a^2 + b^2 \in \mathbb{R}.$$

Der **Betrag** $|z|$ einer komplexen Zahl $z = a + b\,i$ ist definiert als

$$|z| \coloneqq \sqrt{z \cdot \overline{z}} = \sqrt{a^2 + b^2}.$$

Für die Beträge von $w, z \in \mathbb{C}$ gilt

$$|w \cdot z| = |w| \cdot |z|.$$

Die Exponentialfunktion $\mathbb{R} \ni x \mapsto \exp(x) \in \mathbb{R}_{>0}$ besitzt eine natürliche Fortsetzung auf die komplexen Zahlen. Ist $z = a + b\,i$ die Darstellung der komplexen Zahl z mit $a = \operatorname{Re}(z)$ und $b = \operatorname{Im}(z)$, so ist diese Fortsetzung gegeben durch

$$\mathbb{C} \ni z = a + b\,i \mapsto \exp(a + b\,i) \coloneqq \underbrace{\exp(a)}_{\in \mathbb{R}_{>0}} \cdot \underbrace{(\cos(b) + i \cdot \sin(b))}_{\in \mathbb{C}}.$$

Insbesondere gilt die **Eulersche Formel** (englisch: Euler's formula)

$$\exp(ix) = \cos(x) + i \cdot \sin(x) \quad \text{für } x \in \mathbb{R},$$

deren Spezialfall $\exp(i\pi) = -1$ als **Eulersche Identität** (englisch: Euler's identity) bekannt ist.

Jede komplexe Zahl besitzt eine **Polardarstellung** (englisch: polar form) der Form

$$z = r \cdot (\cos(\varphi) + i \cdot \sin(\varphi)) = r \cdot \exp(i\varphi)$$

mit $r = |z| \in \mathbb{R}_{\geqslant 0}$ und einem Winkel $\varphi \in \mathbb{R}$, der auch das **Argument** (englisch: argument) von z genannt wird. Im Fall $r > 0$ ist das Argument bis auf ganzzahlige Vielfache von $2\,\pi$ eindeutig bestimmt.

Kapitel B Liste algebraischer Strukturen

In der folgenden Tabelle ist X irgendeine Menge. Die Abkürzungen „komm." und „neut. E." stehen für „kommutativ" und „neutrales Element". Bei Ringen und Körpern bezieht sich die Kommutativität und die Angabe des neutrale Elements auf die zweite Verknüpfung. Die angegebenen Eigenschaften können in Einzelfällen abweichen, vor allem im Fall $m = 1$ oder wenn X die leere Menge oder eine einelementige Menge ist.

Symbol	Beschreibung	komm.	neut. E.	Referenz
Halbgruppen und Monoide ($m \in \mathbb{N}$)				
$(\mathbb{N}, +)$		✓	–	Beispiele 7.2, 7.4, 7.8, 7.17, 7.22
$(\mathbb{N}_0, +)$		✓	0	und 7.29
$(\mathbb{N}, \cdot)$		✓	1	
$(\mathbb{N}_0, \cdot)$		✓	1	
$(\mathbb{Z}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8, 7.17, 7.22
$(\mathbb{Q}, \cdot)$		✓	1	und 7.29
$(\mathbb{R}, \cdot)$		✓	1	
$(\mathbb{C}, \cdot)$		✓	1	
$(\mathbb{Z}_m, \cdot_m)$	multiplikatives Monoid $\mathbb{Z}$ modulo m	✓	1	Beispiele 7.17 und 7.22
$(H^X, +)$	Funktionen $X \to H$ in die Halbgruppe $(H, +)$	wie in $(H, +)$		
$(\mathbb{N}^X, +)$		✓	–	Beispiele 7.2, 7.4, 7.8 und 7.24
$(\mathbb{N}_0^X, +)$		✓	$x \mapsto 0$	
$(H^X, \cdot)$	Funktionen $X \to H$ in die Halbgruppe $(H, \cdot)$	wie in $(H, \cdot)$		
$(\mathbb{N}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{N}_0^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{Z}^X, \cdot)$		✓	$x \mapsto 1$	Beispiele 7.2, 7.4, 7.8 und 7.24
$(\mathbb{Q}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{R}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{C}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathcal{P}(X), \cap)$	Potenzmenge einer Menge X	✓	X	
$(\mathcal{P}(X), \cup)$		✓	$\emptyset$	Beispiele 7.4, 7.8 und 7.12
$(\mathcal{P}(X), \triangle)$		✓	$\emptyset$	
$(X^X, \circ)$		–	id_X	Beispiele 7.2, 7.4, 7.17, 7.22 und 7.24
$(\Sigma^*, \circ)$		–	$()$	Beispiele 7.4 und 7.8

Symbol	Beschreibung	komm.	neut. E.	Referenz
Gruppen	$(m \in \mathbb{N})$			
$(\mathbb{Z}, +)$		✓	0	Beispiele 7.2, 7.4, 7.8, 7.17, 7.22 und 7.29
$(\mathbb{Q}, +)$		✓	0	
$(\mathbb{R}, +)$		✓	0	
$(\mathbb{C}, +)$		✓	0	
$(\mathbb{Q}_{\neq 0}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8, 7.17, 7.22 und 7.29
$(\mathbb{R}_{\neq 0}, \cdot)$		✓	1	
$(\mathbb{C}_{\neq 0}, \cdot)$		✓	1	
$(m\,\mathbb{Z}, +)$	ganzzahlige Vielfache von m	✓	1	Beispiele 7.45 und 7.51
$(\mathbb{Z}_m, +_m)$	additive Gruppe $\mathbb{Z}$ modulo m	✓	0	Beispiele 7.22 und 8.23
$(\mathbb{Z} / m\,\mathbb{Z}, \tilde{+})$	Faktorgruppe, isomorph zu $(\mathbb{Z}_m, +_m)$	✓	$[0]$	
$(G^X, +)$	Funktionen $X \to G$ in die Gruppe $(G, +)$	wie in $(G, +)$		Beispiele 7.2, 7.4, 7.8 und 7.22
$(\mathbb{Z}^X, +)$		✓	$x \mapsto 0$	
$(\mathbb{Q}^X, +)$		✓	$x \mapsto 0$	
$(\mathbb{R}^X, +)$		✓	$x \mapsto 0$	
$(\mathbb{C}^X, +)$		✓	$x \mapsto 0$	
$(G^X, \cdot)$	Funktionen $X \to G$ in die Gruppe $(G, \cdot)$	wie in $(G, \cdot)$		Beispiele 7.2, 7.4, 7.8 und 7.22
$(\mathbb{Q}^X_{\neq 0}, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{R}^X_{\neq 0}, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{C}^X_{\neq 0}, \cdot)$		✓	$x \mapsto 1$	
$(S_n, \circ)$	symmetrische Gruppe auf $[\![1, n]\!]$	—	$\mathrm{id}_{[\![1,n]\!]}$	Definition 7.30 und Beispiele 7.31 und 7.45
$(A_n, \circ)$	alternierende Gruppe auf $[\![1, n]\!]$	—	$\mathrm{id}_{[\![1,n]\!]}$	

Symbol	Beschreibung	komm.	neut. E.	Referenz
Ringe	($m \in \mathbb{N}$, $n \in \mathbb{N}_0$)			
$(\{0_R\}, +, \cdot)$	Nullring	✓	0_R	Beispiele 9.2 und 9.5
$(\mathbb{Z}, +, \cdot)$		✓	1	Beispiele 9.2, 9.5 und 9.9
$(m\,\mathbb{Z}, +, \cdot)$	ganzzahlige Vielfache von m	✓	—	Beispiel 9.2
$(\mathbb{Z}_m, +_m, \cdot_m)$	$\mathbb{Z}$ modulo m	✓	1	Beispiele 9.2, 9.5, 9.6, 9.23, 9.32
$(\mathbb{Z} \,/\, m\,\mathbb{Z}, \tilde{+}, \tilde{\cdot})$	Restklassenring modulo m, isomorph zu $(\mathbb{Z}_m, +_m, \cdot_m)$	✓	$[1]$	und 10.2
$(R^X, +, \cdot)$	Funktionen $X \to R$ in den Ring $(R, +, \cdot)$	wie in $(R, +, \cdot)$		
$(\mathbb{Z}^X, +, \cdot)$		✓	1	
$(\mathbb{Q}^X, +, \cdot)$		✓	1	Beispiele 9.2, 9.9 und 10.2
$(\mathbb{R}^X, +, \cdot)$		✓	1	
$(\mathbb{C}^X, +, \cdot)$		✓	1	
$(\mathcal{P}(X), \triangle, \cap)$		✓	X	Beispiele 9.2 und 9.9
$(\operatorname{Endo}(G), +, \circ)$	Endomorphismenring der abelschen Gruppe $(G, +)$	—	id_G	Beispiel 9.2 und Satz 17.15
$(\operatorname{Endo}(V), +, \circ)$	Endomorphismenring des Vektorraumes $(V, +, \cdot)$	—	id_V	
$(K^{n \times n}, +, \cdot)$	quadratische $n \times n$-Matrizen über einem Körper K	—	I_n	
$(K^{n \times n}_{\text{◹}}, +, \cdot)$	Unterring der oberen Dreiecksmatrizen	—	I_n	Lemmata 15.33 und 15.38
$(K^{n \times n}_{\text{◺}}, +, \cdot)$	Unterring der unteren Dreiecksmatrizen	—	I_n	
$(K^{n \times n}_{\diagdown}, +, \cdot)$	Unterring der Diagonalmatrizen	✓	I_n	

Symbol	Beschreibung	komm.	neut. E.	Referenz
Körper	$(m \in \mathbb{N})$			
$(\mathbb{Q}, +, \cdot)$		✓	1	Beispiele 9.2, 9.5, 9.9, 10.2 und 10.21
$(\mathbb{R}, +, \cdot)$		✓	1	
$(\mathbb{C}, +, \cdot)$		✓	1	
$(\mathbb{Z}_m, +_m, \cdot_m)$	Körper von $\mathbb{Z}$ modulo m für Primzahlen m	✓	1	Beispiele 9.2, 10.2 und 10.8, Satz 9.11 und Folgerung 10.4
$(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$	Restklassenkörper mod. m für Primz. m, isomorph zu $(\mathbb{Z}_m, +_m, \cdot_m)$	✓	[1]	

Symbol	Beschreibung	Referenz
Vektorräume	$(n, m \in \mathbb{N}_0)$	
$(K_n, +, \cdot)$	Zeilenvektoren über einem Körper $(K, +, \cdot)$	Beispiel 11.3
$(K^n, +, \cdot)$	Spaltenvektoren über einem Körper $(K, +, \cdot)$	
$(K^X, +, \cdot)$	Funktionen $X \to K$ in den Körper $(K, +, \cdot)$	Beispiele 11.3 und 11.12
$(K^{\mathbb{N}}, +, \cdot)$	Folgen mit Werten im Körper $(K, +, \cdot)$	
$(K^{\mathbb{N}}, +, \cdot)_{00}$	Folgen mit endlichem Träger und Werten im Körper $(K, +, \cdot)$	
$(V^X, +, \cdot)$	Funktionen $X \to V$ in den Vektorraum $(V, +, \cdot)$	
$(V^{\mathbb{N}}, +, \cdot)$	Folgen mit Werten im Vektorraum $(V, +, \cdot)$	
$(V^{\mathbb{N}}, +, \cdot)_{00}$	Folgen mit endlichem Träger und Werten im Vektorraum $(V, +, \cdot)$	
$(K^{n \times m}, +, \cdot)$	$n \times m$-Matrizen über einem Körper $(K, +, \cdot)$	Sätze 15.3 und 19.6
$(\mathrm{Homo}(V, W), +, \cdot)$	Homomorphismen $V \to W$ mit VR $(V, +, \cdot)$ und $(W, +, \cdot)$ über demselben Körper	Satz 17.14

Lineare Algebra

Kapitel C Hüllenoperatoren

Wir haben im Verlauf der Vorlesung an verschiedenen Stellen die Bildung von *Hüllen* von Teilmengen gewisser Strukturen betrachtet. Diese sind in der untenstehenden Tabelle aufgeführt, ergänzt um einige weitere naheliegende Beispiele, die wir nicht explizit behandelt haben. Die Intention der Hüllenbildung war jeweils, eine gegebene Menge durch Vergrößerung mit einer gewünschten Zusatzeigenschaft auszustatten, die sie i. A. noch nicht besitzt. Beispielsweise können wir durch die Bildung der reflexiven Hülle von einer beliebigen zu einer reflexiven Relation übergehen.

Wir klären zunächst den Begriff des Hüllenoperators, der die Grundlage für alle Hüllenbegriffe bildet:

Definition C.1 (Hüllenoperator)**.**
Es sei X eine Menge. Ein **Hüllenoperator** auf X (englisch: closure operator) ist eine Abbildung $H\colon \mathcal{P}(X) \to \mathcal{P}(X)$ mit den folgenden Eigenschaften:

$$A \subseteq H(A) \qquad H \text{ ist } \textbf{extensiv}^1 \tag{C.1a}$$
$$A \subseteq B \quad \Rightarrow \quad H(A) \subseteq H(B) \qquad H \text{ ist } \textbf{monoton}^2 \tag{C.1b}$$
$$H(H(A)) = H(A) \qquad H \text{ ist } \textbf{idempotent}^3 \tag{C.1c}$$

für alle $A, B \subseteq X$. △

Wir könnten jetzt für die in der unten stehenden Tabelle aufgeführten Hüllen jeweils zeigen, dass sie Hüllenoperatoren im Sinne der Definition C.1 sind. Beispielsweise ist die reflexive Hülle $H(R) \coloneqq R^? = \bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\}$ ein Hüllenoperator auf $X \times X$, denn für jede Relation $R \subseteq X \times X$ gilt: R ist Teilmenge jeder Relation S, über die in der rechten Menge der Durchschnitt gebildet wird. Daher ist R auch Teilmenge des Durchschnitts $H(R)$, d. h., H ist extensiv. Bilden wir die reflexive Hülle von $H(R)$, so ist $H(R)$ Teilmenge jeder Relation S, über die in der rechten Menge der Durchschnitt gebildet wird. Daher ist $H(R)$ auch Teilmenge des Durchschnitts $H(H(R))$. Andererseits gilt wegen der Extensivität von H auch die umgekehrte Inklusion $H(R) \subseteq H(H(R))$, also ist H idempotent. Sind schließlich $R_1 \subseteq R_2$ zwei Relationen auf X, so ist jede Relation S, die in der rechten Menge der Durchschnittsbildung von $H(R_2)$ vorkommt, auch in der rechten Menge der Durchschnittsbildung von $H(R_1)$ enthalten. Daraus folgt $H(R_1) \subseteq H(R_2)$, d. h., H ist monoton.

Dieser Beweis würde für alle anderen Hüllenoperatoren in der Tabelle genau dieselben Argumente verwenden. Daher stellt sich die Frage nach einer Charakterisierung von Hüllenoperatoren.

[1] englisch: extensive
[2] englisch: monotone
[3] englisch: idempotent

Es stellt sich heraus, dass die entscheidende Eigenschaft ist, dass beispielsweise der Durchschnitt reflexiver Relationen wieder eine reflexive Relation ist. Diese Eigenschaft wird mit der folgenden Definition formalisiert:

Definition C.2 (Abschlusssystem)**.**
Es sei X eine Menge. Ein **Abschlusssystem** auf X (englisch: closure system) ist eine Teilmenge $\mathcal{A} \subseteq \mathcal{P}(X)$ mit der Eigenschaft $X \in \mathcal{A}$, sodass für jede nichtleere Menge $\mathcal{B} \subseteq \mathcal{A}$ gilt: $\bigcap \mathcal{B} \in \mathcal{A}$. △

Ein Abschlusssystem auf X ist also eine Menge von Teilmengen von X, die bezüglich der Durchschnittsbildung abgeschlossen ist. Beispielsweise ist die Menge $\mathcal{A}$ aller reflexiven Relationen auf einer Menge X ein Abschlusssystem auf $X \times X$, weil der Durchschnitt reflexiver Relationen wieder eine reflexive Relation ist (Lemma 5.13). Auch die Menge $\mathcal{A}$ aller Untergruppen einer Gruppe G bildet ein Abschlusssystem (auf G), weil der Durchschnitt von Untergruppen wieder eine Untergruppe ist (Lemma 7.47).

Der folgende Satz stellt nun klar, dass Hüllenoperatoren und Abschlusssysteme „dasselbe" sind, genauer: dass diese bijektiv aufeinander abgebildet werden können.

Satz C.3 (Zusammenhang zwischen Hüllenoperatoren und Abschlusssystemen)**.**
Es sei X eine Menge.

(i) Es sei $H\colon \mathcal{P}(X) \to \mathcal{P}(X)$ ein Hüllenoperator auf X. Dann ist

$$\mathcal{A}_H \coloneqq \{A \subseteq X \mid H(A) = A\}$$

ein Abschlusssystem auf X.
(Wir sammeln alle Teilmengen, die durch Hüllenbildung nicht größer werden.)

(ii) Es sei $\mathcal{A} \subseteq \mathcal{P}(X)$ ein Abschlusssystem auf X. Dann ist die für $A \subseteq X$ durch

$$H_{\mathcal{A}}(A) \coloneqq \bigcap \{B \in \mathcal{A} \mid A \subseteq B\}$$

definierte Abbildung $H_{\mathcal{A}}\colon \mathcal{P}(X) \to \mathcal{P}(X)$ ein Hüllenoperator auf X.
(Wir bilden den Durchschnitt über alle Obermengen im Abschlusssystem, die also die gewünschte Eigenschaft besitzen.)

(iii) Für jeden Hüllenoperator H auf X gilt: $H = H_{\mathcal{A}_H}$.

(iv) Für jedes Abschlusssystem $\mathcal{A}$ auf X gilt: $\mathcal{A} = \mathcal{A}_{H_{\mathcal{A}}}$.

Wir können nun auf einen Schlag und ohne weiteren Aufwand zeigen, dass alle in der Tabelle aufgeführten Begriffe tatsächlich Hüllenoperatoren im Sinne der Definition C.1 sind, denn: Wir wissen in jedem Fall bereits (oder können leicht zeigen), dass die jeweils relevante Menge $\mathcal{A}$ (z. B. die Menge aller reflexiven Relationen auf einer Menge X) ein Abschlusssystem ist. Im Einzelnen: Der Durchschnitt reflexiver Relationen ist wieder reflexiv, der Durchschnitt symmetrischer Relationen ist wieder symmetrisch usw. (Lemma 5.13). Der Durchschnitt von Unterhalbgruppen einer Halbgruppe ist wieder eine Unterhalbgruppe (leicht zu zeigen). Der Durchschnitt von Untergruppen einer Gruppe ist wieder eine Untergruppe (Lemma 7.47). Der Durchschnitt von

Normalteilern einer Gruppe ist wieder ein Normalteiler (leicht zu zeigen). Der Durchschnitt von Unterringen eines Ringes ist wieder eine Unterring (leicht zu zeigen). Der Durchschnitt von Idealen eines Ringes ist wieder ein Ideal (Lemma 9.34). Der Durchschnitt von Unterkörpern eines Körpers ist wieder ein Unterkörper (Lemma 10.10). Der Durchschnitt von Unterräumen eines Vektorraumes ist wieder ein Unterraum (Lemma 11.14).

Der Satz C.3 zeigt, dass wir für jede Eigenschaft, die unter Durchschnittsbildung abgeschlossen ist, einen zugehörigen Hüllenoperator definieren können. Die Hüllenbildung $H(A)$ wird zu einer Teilmenge $A \subseteq X$ genau dann nichts hinzufügen (also $H(A) = A$ liefern), wenn A die gewünschte Eigenschaft bereits besitzt.

Name	Definition	Referenz
homogene **Relation** R auf einer Menge X (Teilmenge R einer Menge $X \times X$)		
reflexive Hülle	$R^? \coloneqq \bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und } R \subseteq S\}$	
symmetrische Hülle	$R^{\text{sym}} \coloneqq \bigcap\{S \subseteq X \times X \mid S \text{ ist symmetrisch und } R \subseteq S\}$	
transitive Hülle	$R^+ \coloneqq \bigcap\{S \subseteq X \times X \mid S \text{ ist transitiv und } R \subseteq S\}$	Definition 5.9
reflexiv-transitive Hülle	$R^* \coloneqq \bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv und transitiv und } R \subseteq S\}$	
reflexiv-symmetrisch-transitive Hülle	$R^\sim \coloneqq \bigcap\{S \subseteq X \times X \mid S \text{ ist reflexiv, symmetrisch und transitiv und } R \subseteq S\}$	
Teilmenge E einer **Halbgruppe** H		
erzeugte Unterhalbgruppe	$\bigcap\{U \mid U \text{ ist Unterhalbgruppe von } H \text{ und } E \subseteq U\}$	—
Teilmenge E einer **Gruppe** G		
erzeugte Untergruppe	$\langle E \rangle \coloneqq \bigcap\{U \mid U \text{ ist Untergruppe von } G \text{ und } E \subseteq U\}$	Definition 7.48
erzeugter Normalteiler	$\bigcap\{U \mid U \text{ ist normale Untergruppe von } G \text{ und } E \subseteq U\}$	—
Teilmenge E eines **Ringes** R		
erzeugter Unterring	$\bigcap\{U \mid U \text{ ist Unterring von } R \text{ und } E \subseteq U\}$	—
erzeugtes Ideal	$(E) \coloneqq \bigcap\{J \mid J \text{ ist Ideal von } R \text{ und } E \subseteq J\}$	Definition 9.35
Teilmenge E eines **Körpers** K		
erzeugter Unterkörper	$\bigcap\{U \mid U \text{ ist Unterkörper von } R \text{ und } E \subseteq U\}$	—
Teilmenge E eines **Vektorraumes** V		
erzeugter Unterraum	$\langle E \rangle \coloneqq \bigcap\{U \mid U \text{ ist Unterraum von } V \text{ und } E \subseteq U\}$	Definition 11.10

Kapitel D Einige Algorithmen

Der folgende Algorithmus erzeugt aus einer Matrix $A \in K^{n\times m}$ eine Matrix in Zeilenstufenform $C \in K^{n\times m}$ (Definition 15.20) mit der Eigenschaft $\mathrm{ZR}(A) = \mathrm{ZR}(C)$ (Satz 15.23). Die Anzahl $r \in [\![0, n]\!]$ der Nicht-Nullzeilen von C ist der Rang von A. Diese Zeilen bilden eine Basis von $\mathrm{ZR}(A)$.

Algorithmus D.1 (Erzeugen einer Zeilenstufenform)**.**

Eingabe: Matrix $A \in K^{n\times m}$
Ausgabe: Matrix $C \in K^{n\times m}$ in Zeilenstufenform mit $\mathrm{ZR}(A) = \mathrm{ZR}(C)$
Ausgabe: $r = \mathrm{Rang}(A) = \mathrm{Rang}(C)$

```
Setze C ← A
Setze r ← 0                                                        // bisher festgestellter Rang
Setze j_0 ← 0                                                      // Spalte des letzten Pivot-Elements
while r < n und j_r < m und die Restmatrix (C)_{r+1≤i≤n, j_r+1≤j≤m} ist nicht die Nullmatrix do
    Setze j ← min{j_r + 1 ≤ j ≤ m | (C)_{r+1≤i≤n, j} }             // erste Nicht-Nullspalte der Restmatrix
    Setze i ← min{r + 1 ≤ i ≤ n | c_{ij} ≠ 0}                      // erster Nicht-Nulleintrag in dieser Spalte
    Setze r ← r + 1                                                // festgestellter Rang erhöht sich
    Setze j_r ← j                                                  // Spalte des neuen Pivot-Elements merken
    Tausche in der Matrix C die Zeilen i und r                     // Pivot-Element kommt nach oben (Typ III)
  // Erzeuge Nullen unterhalb des Pivot-Elements (r, j) durch Elementarmatrizen (Typ II):
    for i = r + 1, . . . , m do
        Setze c_{i•} ← c_{i•} − (c_{ij}/c_{rj}) c_{r•}              // Erzeuge eine Null an der Stelle (i, j) durch S_{i,r}(−c_{ij}/c_{rj})
    end for
end while
return C und r
```

Der folgende Algorithmus ist eine Erweiterung von Algorithmus D.1. Dieser erzeugt aus einer Matrix $A \in K^{n \times m}$ eine Faktorisierung $A = B\,C$ mit invertierbarer Matrix $B \in K^{n \times n}$ und derselben Matrix $C \in K^{n \times m}$ in Zeilenstufenform C wie in Algorithmus D.1. Durch Beschränkung auf die ersten r Spalten von B (mit $B_{\text{Rang}} \in K^{n \times r}$ bezeichnet) und die ersten r Zeilen von C (mit $C_{\text{Rang}} \in K^{r \times m}$ bezeichnet) erhalten wir eine Rangfaktorisierung $A = B\,C = B_{\text{Rang}}\,C_{\text{Rang}}$ (Folgerung 15.16, Bemerkung 15.25) der Matrix A. Die Spalten von B_{Rang} bilden eine Basis von $\operatorname{SR}(A)$. Die Zeilen von C_{Rang} bilden eine Basis von $\operatorname{ZR}(A)$.

Algorithmus D.2 (Erzeugen einer Rangfaktorisierung $A = B\,C = B_{\text{Rang}}\,C_{\text{Rang}}$).

Eingabe: Matrix $A \in K^{n \times m}$
Ausgabe: Matrix $B \in K^{n \times n}$ mit $\operatorname{SR}(A) = \operatorname{SR}(B)$
Ausgabe: Matrix $B_{\text{Rang}} \in K^{n \times r}$ mit $\operatorname{SR}(A) = \operatorname{SR}(B_{\text{Rang}})$
Ausgabe: Matrix $C \in K^{n \times m}$ in Zeilenstufenform mit $\operatorname{ZR}(A) = \operatorname{ZR}(C)$
Ausgabe: Matrix $C_{\text{Rang}} \in K^{r \times m}$ in Zeilenstufenform mit $\operatorname{ZR}(A) = \operatorname{ZR}(C_{\text{Rang}})$
Ausgabe: $r = \operatorname{Rang}(A) = \operatorname{Rang}(B) = \operatorname{Rang}(B_{\text{Rang}}) = \operatorname{Rang}(C) = \operatorname{Rang}(C_{\text{Rang}})$

Setze $B \leftarrow I_n$
Setze $C \leftarrow A$
Setze $r \leftarrow 0$ // bisher festgestellter Rang
Setze $j_0 \leftarrow 0$ // Spalte des letzten Pivot-Elements
while $r < n$ und $j_r < m$ und die Restmatrix $(C)_{r+1 \leqslant i \leqslant n,\ j_r+1 \leqslant j \leqslant m}$ ist nicht die Nullmatrix **do**
  Setze $j \leftarrow \min\{j_r + 1 \leqslant j \leqslant m \mid (C)_{r+1 \leqslant i \leqslant n,\ j}\}$ // erste Nicht-Nullspalte der Restmatrix
  Setze $i \leftarrow \min\{r + 1 \leqslant i \leqslant n \mid c_{ij} \neq 0\}$ // erster Nicht-Nulleintrag in dieser Spalte
  Setze $r \leftarrow r + 1$ // festgestellter Rang erhöht sich
  Setze $j_r \leftarrow j$ // Spalte des neuen Pivot-Elements merken
  Tausche in der Matrix C die Zeilen i und r // Pivot-Element kommt nach oben (Typ III)
  Tausche in der Matrix B die Spalten i und r
  // Erzeuge Nullen unterhalb des Pivot-Elements (r, j) durch Elementarmatrizen (Typ II):
  for $i = r + 1, \ldots, m$ **do**
    Setze $c_{i\bullet} \leftarrow c_{i\bullet} - \frac{c_{ij}}{c_{rj}} c_{r\bullet}$ // Erzeuge eine Null an der Stelle (i, j) durch $S_{i,r}(-\frac{c_{ij}}{c_{rj}})$
    Setze $b_{\bullet r} \leftarrow b_{\bullet r} + \frac{c_{ij}}{c_{rj}} b_{\bullet i}$ // Gehe von $A = B\,C$ über zu $A = B\,S_{i,r}(-\frac{c_{ij}}{c_{rj}})^{-1} S_{i,r}(-\frac{c_{ij}}{c_{rj}})\,C$
  end for
end while
Setze $B_{\text{Rang}} \leftarrow \begin{bmatrix} b_{\bullet 1} & \cdots & b_{\bullet r} \end{bmatrix}$
Setze $C_{\text{Rang}} \leftarrow \begin{bmatrix} c_{1\bullet} \\ \vdots \\ c_{r\bullet} \end{bmatrix}$
return B, B_{Rang}, C, C_{Rang} und r

Der folgende Algorithmus ist eine Erweiterung sowohl von Algorithmus D.1 als auch von Algorithmus D.2. Er erzeugt aus einer Matrix $A \in K^{n \times m}$ eine Faktorisierung $A = BC$ mit invertierbarer Matrix $B \in K^{n \times n}$ und Matrix in reduzierter Zeilenstufenform $C \in K^{n \times m}$ (Definition 16.5) mit der Eigenschaft $\mathrm{ZR}(A) = \mathrm{ZR}(C)$ (Satz 16.7). Durch Beschränkung auf die ersten r Spalten von B (mit $B_{\text{Rang}} \in K^{n \times r}$ bezeichnet) und die ersten r Zeilen von C (mit $C_{\text{Rang}} \in K^{r \times m}$ bezeichnet) erhalten wir eine Rangfaktorisierung $A = BC = B_{\text{Rang}}\, C_{\text{Rang}}$ (Folgerung 15.16, Bemerkung 15.25) der Matrix A. Die Spalten von B_{Rang} bilden eine Basis von $\mathrm{SR}(A)$. Die Zeilen von C_{Rang} bilden eine Basis von $\mathrm{ZR}(A)$.

Gilt $n = m$, dann ist A genau dann invertierbar, wenn $C = I_n$ gilt.

Algorithmus D.3 (Erzeugen einer Rangfaktorisierung $A = BC = B_{\text{Rang}}\, C_{\text{Rang}}$ mit reduzierter Zeilenstufenform).

Eingabe: Matrix $A \in K^{n \times m}$
Ausgabe: Matrix $B \in K^{n \times n}$ mit $\mathrm{SR}(A) = \mathrm{SR}(B)$
Ausgabe: Matrix $B_{\text{Rang}} \in K^{n \times r}$ mit $\mathrm{SR}(A) = \mathrm{SR}(B_{\text{Rang}})$
Ausgabe: Matrix $C \in K^{n \times m}$ in reduzierter Zeilenstufenform mit $\mathrm{ZR}(A) = \mathrm{ZR}(C)$
Ausgabe: Matrix $C_{\text{Rang}} \in K^{r \times m}$ in reduzierter Zeilenstufenform mit $\mathrm{ZR}(A) = \mathrm{ZR}(C_{\text{Rang}})$
Ausgabe: $r = \mathrm{Rang}(A) = \mathrm{Rang}(B) = \mathrm{Rang}(B_{\text{Rang}}) = \mathrm{Rang}(C) = \mathrm{Rang}(C_{\text{Rang}})$

Setze $B \leftarrow I_n$
Setze $C \leftarrow A$
Setze $r \leftarrow 0$ // bisher festgestellter Rang
Setze $j_0 \leftarrow 0$ // Spalte des letzten Pivot-Elements
while $r < n$ und $j_r < m$ und die Restmatrix $(C)_{r+1 \leqslant i \leqslant n,\, j_r+1 \leqslant j \leqslant m}$ ist nicht die Nullmatrix **do**
  Setze $j \leftarrow \min\{j_r + 1 \leqslant j \leqslant m \mid (C)_{r+1 \leqslant i \leqslant n,\, j}\}$ // erste Nicht-Nullspalte der Restmatrix
  Setze $i \leftarrow \min\{r + 1 \leqslant i \leqslant n \mid c_{ij} \neq 0\}$ // erster Nicht-Nulleintrag in dieser Spalte
  Setze $r \leftarrow r + 1$ // festgestellter Rang erhöht sich
  Setze $j_r \leftarrow j$ // Spalte des neuen Pivot-Elements merken
  Tausche in der Matrix C die Zeilen i und r // Pivot-Element kommt nach oben (Typ III)
  Tausche in der Matrix B die Spalten i und r
  // Erzeuge Nullen unterhalb des Pivot-Elements (r, j) durch Elementarmatrizen (Typ II):
  for $i = r + 1, \ldots, m$ **do**
    Setze $c_{i\bullet} \leftarrow c_{i\bullet} - \frac{c_{ij}}{c_{rj}} c_{r\bullet}$ // Erzeuge eine Null an der Stelle (i, j) durch $S_{i,r}(-\frac{c_{ij}}{c_{rj}})$
    Setze $b_{\bullet r} \leftarrow b_{\bullet r} + \frac{c_{ij}}{c_{rj}} b_{\bullet i}$ // Gehe von $A = BC$ über zu $A = B\, S_{i,r}(-\frac{c_{ij}}{c_{rj}})^{-1} S_{i,r}(-\frac{c_{ij}}{c_{rj}})\, C$
  end for
end while
for $k = r, \ldots, 1$ **do**
  Setze $j \leftarrow j_k$ // aktuelles Pivot-Element ist (k, j)
  // Erzeuge Nullen oberhalb des Pivot-Elements (k, j) durch Elementarmatrizen (Typ II):
  for $i = k - 1, \ldots, 1$ **do**
    Setze $c_{i\bullet} \leftarrow c_{i\bullet} - \frac{c_{ij}}{c_{kj}} c_{k\bullet}$ // Erzeuge eine Null an der Stelle (i, j) durch $S_{i,k}(-\frac{c_{ij}}{c_{kj}})$
    Setze $b_{\bullet k} \leftarrow b_{\bullet k} + \frac{c_{ij}}{c_{kj}} b_{\bullet i}$ // Gehe von $A = BC$ über zu $A = B\, S_{i,k}(-\frac{c_{ij}}{c_{kj}})^{-1} S_{i,k}(-\frac{c_{ij}}{c_{kj}})\, C$
  end for
  Setze $p \leftarrow c_{kj}$ // Hole das Pivot-Element
  Setze $c_{k\bullet} \leftarrow \frac{1}{p} c_{k\bullet}$ // Mache das Pivot-Element (k, j) zu 1 durch $D_k(\frac{1}{p})$

Setze $b_{\bullet k} \leftarrow p\, b_{\bullet k}$ // Gehe von $A = B\,C$ über zu $A = B\,D_k(\frac{1}{p})^{-1} D_k(\frac{1}{p})\,C$

end for

Setze $B_{\text{Rang}} \leftarrow \begin{bmatrix} b_{\bullet 1} & \cdots & b_{\bullet r} \end{bmatrix}$

Setze $C_{\text{Rang}} \leftarrow \begin{bmatrix} c_{1\bullet} \\ \vdots \\ c_{r\bullet} \end{bmatrix}$

return B, B_{Rang}, C, C_{Rang} und r

Kapitel E Das griechische Alphabet

Kleinbuchstabe	Großbuchstabe	Name
α	A	alpha
β	B	beta
γ	Γ	gamma
δ	Δ	delta
ϵ, ε	E	epsilon
ζ	Z	zeta
η	H	eta
θ, ϑ	Θ	theta
ι	I	iota
κ, $\varkappa$	K	kappa
λ	Λ	lambda
μ	M	mu
ν	N	nu
ξ	Ξ	xi
o	O	omikron
π, ϖ	Π	pi
ρ, ϱ	P	rho
σ, ς	Σ	sigma
τ	T	tau
υ	Υ	ypsilon
ϕ, φ	Φ	phi
χ	X	chi
ψ	Ψ	psi
ω	Ω	omega

Kapitel F Abkürzungen

Abkürzung	Bedeutung
bzgl.	bezüglich
d. h.	das heißt
etc.	et cetera
i. A.	im Allgemeinen
i. d. R.	in der Regel
i. W.	im Wesentlichen
o. B. d. A.	ohne Beschränkung der Allgemeinheit
o. ä.	oder ähnlich
usw.	und so weiter
vgl.	vergleiche

Literatur

Beutelspacher, A. (2014). *Lineare Algebra. Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen.* 8. Aufl. Springer Fachmedien Wiesbaden. DOI: 10.1007/978-3-658-02413-0.

Bosch, S. (2014). *Lineare Algebra.* 5. Aufl. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-55260-1.

Deiser, O. (2022). *Einführung in die Mathematik 2.1. Elementare Zahlentheorie und Graphentheorie.* URL: https://www.aleph1.info/?call=Puc&permalink=ema21.

Deiser, O. (2024a). *Einführung in die Mengenlehre. Die Mengenlehre Georg Cantors und ihre Axiomatisierung durch Ernst Zermelo.* URL: https://www.aleph1.info/?call=Puc&permalink=mengenlehre1.

Deiser, O. (2024b). *Grundbegriffe der Mathematik. Sprache, Zahlen und erste Erkundungen.* URL: https://www.aleph1.info/?call=Puc&permalink=grundbegriffe.

Fischer, G.; B. Springborn (2020). *Lineare Algebra.* 19. Aufl. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-61645-1.

Goldrei, D. (1996). *Classic Set Theory. For Guided Independent Study.* Chapman & Hall/CRC, Boca Raton, FL.

Hackbusch, W. (2019). *Tensor Spaces and Numerical Tensor Calculus.* Springer International Publishing. DOI: 10.1007/978-3-030-35554-8.

Håstad, J. (1990). „Tensor rank is NP-complete". *Journal of Algorithms* 11.4, S. 644–654. DOI: 10.1016/0196-6774(90)90014-6.

Jänich, K. (2008). *Lineare Algebra.* 11. Aufl. Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-75502-9.

Magnus, P. D.; T. Button; J. R. Loftis; R. Trueman; A. Thomas-Bolduc; R. Zach; S. Wimmer (2023). *forall x: Dortmund. Eine Einführung in die formale Logik.* URL: https://github.com/sbwimmer/forallx-do/.

Thiele, R. (1979). *Mathematische Beweise.* Bd. 99. Leipzig: B. G. Teubner Verlagsgesellschaft. URL: https://mathematikalpha.de/?smd_process_download=1&download_id=26662.

Velleman, D. J. (2019). *How to Prove It. A Structured Approach.* 3. Aufl. Cambridge University Press, Cambridge. DOI: 10.1017/9781108539890.

Index