Vorlesungsskript
Lineare Algebra

Wintersemester 2023 - Sommersemester 2024

Roland Herzog*

2024-07-11

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

Dieses Skript orientiert sich an früheren Vorlesungen von Jan Johannes und Alexander Schmidt (Universität Heidelberg).

Änderungen gegenüber bereits veröffentlichten Versionen werden in dieser Farbe gekennzeichnet.

Material für 29 Vorlesungen (Lineare Algebra I). Material für 28 Vorlesungen (Lineare Algebra II).

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 versteht den Begriff Algebra deutlich weiter, es geht jedoch immer um Strukturen, Abbildungen zwischen Strukturen und die in ihnen geltenden „Rechen“regeln. 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, 2022b, Kapitel 1.1, Magnus u. a., 2023, Kapitel 1–14

Definition 1.1 (Aussage, Wahrheitswert)**.**
Eine **Aussage** (englisch: statement) 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 $\perp$, 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 A, B usw.

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

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

(*ii*) B: Am 17.10.2023 ist London die Hauptstadt von Frankreich.
Dieses ist eine falsche Aussage.

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

(*iv*) D: Das Team des VfL Wolfsburg ist in der Saison 2023/24 deutscher Meister in der Frauen-Fußball-Bundesliga.
Dieses ist eine Aussage, deren Wahrheitswert wir im Moment aber nicht kennen.

(*v*) E: 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 definieren einen Junktor über seine **Wahrheitswerttabelle** (auch: **Wahrheitstafel**, englisch: truth table).

Definition 1.3 (Junktoren)**.**

Im Folgenden seien A und B Aussagen. Wir definieren folgende wichtige ein- und zweistellige Junktoren.

(*i*) **Negation** (**Verneinung**, englisch: negation) $\neg$

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

A	$\neg A$
W	F
F	W

(*ii*) **Konjunktion**[2] (**Und-Verknüpfung**, englisch: conjunction) $\wedge$

Die Aussage $A \wedge B$ (sprich: „A und B“) ist dann wahr, wenn A und B beide wahr sind, ansonsten falsch.

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

[1] siehe Primzahlzwillingsvermutung
[2] lateinisch: coniungere: verbinden

(*iii*) **Disjunktion**[3] (**Oder-Verknüpfung**, englisch: disjunction) $\vee$

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

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

(*iv*) **Implikation**[4] (**Konditional**[5], **Wenn-Dann-Verknüpfung**, englisch: implication) $\longrightarrow$

Die Aussage $A \longrightarrow B$ ist über die nebenstehende Wahrheitswerttabelle definiert. Man benennt die Aussage auch als „A ist **hinreichend** für B“ (englisch: „A is sufficient for B“), „B ist **notwendig** für A“ (englisch: „B is necessary for A“), „A impliziert B“ (englisch: „A implies B“) oder „Wenn A, dann B“ (englisch: „If A, then B“). In einer Implikation $A \longrightarrow B$ nennt man A auch das **Antezedens** (englisch: antecedent, lateinisch: antecedens: das Vorausgehende) und B das **Konsequens** (englisch: consequent, lateinisch: consequentis: folgerichtig).

A	B	$A \longrightarrow B$
W	W	W
W	F	F
F	W	W
F	F	W

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

(*v*) **Äquivalenz**[6] (**Bikonditional**, **Genau-Dann-Wenn-Verkn.**, englisch: equivalence) $\longleftrightarrow$

Die Aussage $A \leftrightarrow B$ ist wahr, wenn entweder A und B beide wahr oder beide falsch sind, ansonsten falsch. Man benennt die Aussage auch als „A ist **notwendig und hinreichend** für B“ (englisch: „A is necessary and sufficient for B“), „A ist äquivalent zu B“ (englisch: „A is equivalent to B“), „A genau dann, wenn B“ oder „A dann und nur dann, wenn B“ (englisch: „A if and only if B“, „A iff B“).

A	B	$A \leftrightarrow B$
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 A und B aussagt. Man spricht auch von **materialer Äquivalenz** (englisch: material equivalence). △

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

Quizfrage 1.2: Können Sie alle zweistelligen Junktoren aus den oben genannten, also aus $\neg$ sowie $\wedge$, $\vee$, $\longrightarrow$ und $\leftrightarrow$, zusammensetzen? Reicht evtl. sogar eine Teilmenge davon aus?

Beispiel 1.4 (Symbolisierung von Sätzen der Umgangssprache[7]).
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.

[3] lateinisch: disiungere: trennen, unterscheiden
[4] lateinisch: implicare: verwickeln
[5] lateinisch: conditio: Bedingung
[6] lateinisch: aequivalens: gleichwertig
[7] angelehnt an Beispiele aus Magnus u. a., 2023, Kapitel 5, genutzt unter der Lizenz CC-BY 4.0

(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 A und B Aussagen.

(i) Die Aussagen

- $A \rightarrow B$
- $(\neg A) \vee B$
- $(\neg B) \rightarrow (\neg A)$

haben dieselben Wahrheitstafeln.

(ii) Die Aussagen

- $A \leftrightarrow B$
- $(A \rightarrow B) \wedge (B \rightarrow A)$

haben dieselben Wahrheitstafeln.

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

A	B	$A \to B$	$(\neg A) \lor B$	$\neg B$	$\neg A$	$(\neg B) \to (\neg A)$
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 Teil von Hausaufgabe I-1.3. □

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

$$\neg \text{ bindet stärker als } \land \text{ bindet stärker als } \lor \text{ bindet stärker als } \to \text{ bindet stärker als } \leftrightarrow . \tag{1.1}$$

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

$$\begin{aligned} (\neg A) \land B &\quad \text{dasselbe wie} \quad \neg A \land B \\ \text{und} \quad (\neg(A \land B)) \to (B \lor \neg B) &\quad \text{dasselbe wie} \quad \neg(A \land B) \to B \lor \neg B. \end{aligned}$$

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

Wir berechnen jetzt die Wahrheitstafeln einiger zusammengesetzter Aussagen.

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

(i) $\neg(\neg A \land \neg B)$

A	B	$\neg A$	$\neg B$	$\neg A \land \neg B$	$\neg(\neg A \land \neg B)$
W	W	F	F	F	W
W	F	F	W	F	W
F	W	W	F	F	W
F	F	W	W	W	F

Diese Wahrheitstafel ist offenbar dieselbe wie die von $A \lor B$.

(ii) $A \lor B \to B \land C$

A	B	C	$A \lor B$	$B \land C$	$A \lor B \to B \land C$
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) $\neg(A \land B) \leftrightarrow \neg A \lor \neg B$

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

△

Die letzte Aussage $\neg(A \wedge B) \leftrightarrow \neg A \vee \neg B$ hat also immer den Wahrheitswert W, unabhängig von den Wahrheitswerten der Aussagen A und B. Eine solche Aussage nennt man **Tautologie**[8] (englisch: tautology) oder **logisches Gesetz**. Tautologien spielen eine entscheidende Rollen in mathematischen Beweisen, siehe § 3.

Definition 1.7 (logische Implikation, logische Äquivalenz)**.**

Es seien A und B Aussagen.

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

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

Beachte: Die logische Implikation und die logische Äquivalenz sind *Aussagen über Aussagen*. Sie sind von den Junktoren „Implikation" (Konditional) und „Äquivalenz" (Bikonditional) zu unterscheiden!

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

Beispiel 1.8 (logische Implikationen und Äquivalenzen)**.**

(i) Die Aussage $(A \rightarrow B) \wedge A$ impliziert die Aussage B, kurz: $(A \rightarrow B) \wedge A \Rightarrow B$, denn $(A \rightarrow B) \wedge A \rightarrow B$ ist eine Tautologie:

A	B	$A \rightarrow B$	$(A \rightarrow B) \wedge A$	$(A \rightarrow B) \wedge A \rightarrow B$
W	W	W	W	W
W	F	F	F	W
F	W	W	F	W
F	F	W	F	W

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

[8] altgriechisch: *ταυτο*: dasselbe

Satz 1.9 (logische Implikationen und Äquivalenzen).
Es seien A, B und C Aussagen. Es gelten die folgenden Implikationen und Äquivalenzen.

$$\neg(\neg A) \Leftrightarrow A \qquad \text{doppelte Verneinung[9]} \tag{1.2}$$

$$A \Rightarrow \top \qquad \text{„Aus Beliebigem folgt Wahres."[10]} \tag{1.3a}$$
$$\bot \Rightarrow A \qquad \text{„Aus Falschem folgt Beliebiges."[11]} \tag{1.3b}$$

$$A \wedge A \Leftrightarrow A \qquad \textbf{Idempotenz}\text{[12]} \tag{1.4a}$$
$$A \vee A \Leftrightarrow A \qquad \textbf{Idempotenz} \tag{1.4b}$$

$$A \wedge \top \Leftrightarrow A \qquad \textbf{Neutralität}\text{[13]} \tag{1.5a}$$
$$A \vee \bot \Leftrightarrow A \qquad \textbf{Neutralität} \tag{1.5b}$$

$$A \wedge \bot \Leftrightarrow \bot \qquad \textbf{Absorption}\text{[14]} \tag{1.6a}$$
$$A \vee \top \Leftrightarrow \top \qquad \textbf{Absorption} \tag{1.6b}$$

$$A \wedge \neg A \Leftrightarrow \bot \qquad \textbf{Komplementarität}\text{[15]} \tag{1.7a}$$
$$A \vee \neg A \Leftrightarrow \top \qquad \textbf{Komplementarität}\text{[16]} \tag{1.7b}$$

$$A \wedge B \Leftrightarrow B \wedge A \qquad \textbf{Kommutativität von } \wedge\text{[17]} \tag{1.8a}$$
$$A \vee B \Leftrightarrow B \vee A \qquad \textbf{Kommutativität von } \vee \tag{1.8b}$$

$$(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C) \qquad \textbf{Assoziativität von } \vee\text{[18]} \tag{1.9a}$$
$$(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C) \qquad \textbf{Assoziativität von } \wedge \tag{1.9b}$$

$$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B \qquad \textbf{De Morgansches Gesetz}\text{[19]} \tag{1.10a}$$
$$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B \qquad \textbf{De Morgansches Gesetz} \tag{1.10b}$$

$$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C) \qquad \textbf{Distributivität}\text{[20]} \tag{1.11a}$$
$$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C) \qquad \textbf{Distributivität} \tag{1.11b}$$

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

Ende der Vorlesung 1

[9] lateinisch: duplex negatio affirmat
[10] lateinisch: verum ex quolibet
[11] lateinisch: ex falso quodlibet
[12] englisch: idempotence
[13] englisch: neutrality
[14] englisch: absorption
[15] englisch: complementarity
[16] Gesetz vom ausgeschlossenen Dritten, lateinisch: tertium non datur
[17] englisch: commutativity, lateinisch: commutare: tauschen, vertauschen
[18] englisch: associativity, lateinisch: associare: verbinden, beigesellen
[19] englisch: De Morgan's law
[20] englisch: distributivity, lateinisch: distribuere: verteilen, aufteilen

§ 2 Prädikatenlogik

Literatur: 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 nicht wie erforderlich symbolisieren. Die Schwierigkeit ist, dass wir in der Aussagenlogik keine Aussagen mit Variablen zur Verfügung haben. Wir benötigen dazu die **Prädikatenlogik**[21], 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

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

  Beispiele:

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

  Die Anzahl der Variablen einer Aussageform heißt deren **Stelligkeit** (englisch: arity).

- **Quantoren** (englisch: quantifier), und zwar

  $$\begin{aligned} \forall &\quad \text{für alle (\textbf{Allquantor}, englisch: universal quantifier),} \\ \exists &\quad \text{es existiert (mindestens) ein (\textbf{Existenzquantor}, englisch: existential quantifier),} \\ \exists! &\quad \text{es existiert genau ein (\textbf{Eindeutigkeitsquantor}, englisch: uniqueness quantifier).} \end{aligned}$$

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

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

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

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

[21] genauer: Prädikatenlogik erster Stufe, englisch: first order logic

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

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

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

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 $\exists x \; (E(x) \wedge S(x))$ und $\exists y \; (E(y) \wedge S(y))$ 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 *Analysis*)

Die Funktion $f\colon (a,b) \rightarrow \mathbb{R}$ 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 $f\colon (a,b) \rightarrow \mathbb{R}$ ist gleichmäßig stetig:
$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in (a,b) \; \forall y \in (a,b) \; \big(|x-y| < \delta \rightarrow |f(x) - f(y)| < \varepsilon\big).$$

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

Satz 2.2 (logische Implikationen und Äquivalenzen von Aussagen mit Quantoren)**.**
Es seien A, B einstellige Aussageformen mit gemeinsamem Grundbereich und C eine zweistellige Aussageform. Es gelten die folgenden Implikationen und Äquivalenzen.[22]

$$\neg(\forall x \; A(x)) \Leftrightarrow \exists x \; (\neg A(x)) \quad \text{Negation des Allquantors} \tag{2.1a}$$
$$\neg(\exists x \; A(x)) \Leftrightarrow \forall x \; (\neg A(x)) \quad \text{Negation des Existenzquantors} \tag{2.1b}$$
$$\forall x \; \forall y \; C(x,y) \Leftrightarrow \forall y \; \forall x \; C(x,y) \quad \textbf{Kommutativität gleicher Quantoren} \tag{2.2a}$$
$$\exists x \; \exists y \; C(x,y) \Leftrightarrow \exists y \; \exists x \; C(x,y) \quad \textbf{Kommutativität gleicher Quantoren} \tag{2.3a}$$
$$\forall x \; (A(x) \wedge B(x)) \Leftrightarrow \forall x \; A(x) \wedge \forall x \; B(x) \quad \textbf{Distributivität} \tag{2.4a}$$
$$\exists x \; (A(x) \vee B(x)) \Leftrightarrow \exists x \; A(x) \vee \exists x \; B(x) \quad \textbf{Distributivität} \tag{2.4b}$$
$$(\forall x \; A(x)) \vee (\forall x \; B(x)) \Rightarrow \forall x \; (A(x) \vee B(x)) \tag{2.5a}$$
$$\exists x \; (A(x) \wedge B(x)) \Rightarrow (\exists x \; A(x)) \wedge (\exists x \; B(x)) \tag{2.5b}$$
$$\forall x \; (A(x) \rightarrow B(x)) \Rightarrow (\forall x \; A(x)) \rightarrow (\forall x \; B(x)) \tag{2.6a}$$

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

$$\exists x\ (A(x) \to B(x)) \quad \Leftrightarrow \quad (\forall x\ A(x)) \to (\exists x\ B(x)) \tag{2.6b}$$

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

§ 3 Beweismuster

Literatur: Deiser, 2022b, Kapitel 1.1, Magnus u. a., 2023, Kapitel 15–21

In einem Beweis versuchen wir in der Regel, für gegebene Aussagen A, B die Implikation $A \Rightarrow B$ nachzuweisen. Das heißt, wir müssen nachweisen, dass $A \to B$ eine Tautologie ist. Meistens besteht die Prämisse A selbst aus einer Konjunktion (Und-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) 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. Vielmehr werden wir Schlussregeln anwenden, die auf Tautologien beruhen. Solche Tautologien haben wir in Satz 1.9 und Satz 2.2 bereits aufgeführt. Dazu kommen die weiteren Tautologien

$$(A \to B) \land A \quad \Rightarrow \quad B \qquad \textbf{modus ponendo ponens}, \tag{3.1a}$$

$$(A \to B) \land \neg B \quad \Rightarrow \quad \neg A \qquad \textbf{modus tollendo tollens}, \tag{3.1b}$$

$$(A \to \neg B) \land A \quad \Rightarrow \quad \neg B \qquad \textbf{modus ponendo tollens}^{23}, \tag{3.1c}$$

$$(\neg A \to B) \land \neg A \quad \Rightarrow \quad B \qquad \textbf{modus tollendo ponens}^{24}, \tag{3.1d}$$

$$(A \to B) \land (B \to C) \quad \Rightarrow \quad (A \to C) \qquad \textbf{Kettenschluss}\ \text{(englisch: chain inference)}. \tag{3.2}$$

Quizfrage 3.1: Können Sie einfache Beispiele in Alltagssprache für die Argumentation gemäß der vier Argumentationsmuster in (3.1a)–(3.1d) angeben?

Folgende Beweismuster für Implikationen $A \Rightarrow B$ werden häufig verwendet:

(1) Beim **direkten Beweis** (englisch: direct proof) wird $A \Rightarrow B$, typischerweise unter Verwendung von Axiomen und bereits bewiesenen Sätzen, direkt mit Hilfe von Schlussregeln hergeleitet.

(2) Beim **indirekten Beweis** oder **Beweis durch Kontraposition** (englisch: indirect proof, proof by contrapositive) nutzen wir die Äquivalenz $(A \to B) \Leftrightarrow (\neg B \to \neg A)$ aus. Wir führen also einen direkten Beweis für $\neg B \Rightarrow \neg A$.

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

[23] Der modus ponendo tollens wird häufig als $\neg(A \land B) \land A \Rightarrow \neg B$ geschrieben.
[24] Der modus tollendo ponens wird häufig als $(A \lor B) \land \neg A \Rightarrow B$ geschrieben.

(4) Beim **Beweis durch Fallunterscheidung** (englisch: proof by distinction of cases) nutzen wir die Äquivalenz $(A \land C \to B) \land (A \land \neg C \to B) \Leftrightarrow A \to B$. Dabei ist C irgendeine weitere Aussage. Wir nehmen also zunächst die Aussagen A und C als wahr an und zeigen, dass dann auch die Aussage B wahr ist. Anschließend nehmen wir die Aussage A weiterhin als wahr aber die Aussage C als falsch an und zeigen, dass dann wiederum die Aussage B wahr ist.

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:

$$A(m,n) : m^2 < n^2$$
$$B(m,n) : m < n$$

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

$$\forall m \in \mathbb{N}\ \forall n \in \mathbb{N}\ (A(m,n) \Rightarrow B(m,n)).$$

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

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

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 A und B 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**[25]

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 $(\sin x)^2 + (\cos x)^2 = 1$

[25] Dieses Beispiel ist Thiele, 1979 entnommen.

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.

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 $n \in \mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ 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.

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 $A_1 \Rightarrow A_2$, $A_2 \Rightarrow A_3$ usw. bis $A_{n-1} \Rightarrow A_n$ und $A_n \Rightarrow A_1$, was dann wiederum die gewünschten Äquivalenzen zur Folge hat. Das erfordert n Beweisschritte. Wir können sogar allgemeiner solange verschiedene Implikationen $A_i \Rightarrow A_j$ 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 3.2: Wieviele Implikationen wären zu zeigen, wenn man die Äquivalenz der Aussagen A_i und A_j für $i, j = 1, \dots, n$ 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 $A(n)$ für alle ganzen Zahlen $n \in \mathbb{Z}$ ab einem gewissen Startindex $n_0 \in \mathbb{Z}$ zeigen wollen, also für $n \geqslant n_0$. In diesem Fall zeigen wir am **Induktionsanfang** (englisch: base case) die Wahrheit der Aussage $A(n_0)$. Oft wird der Induktionsanfang bei $n_0 = 0$ oder $n_0 = 1$ gesetzt.

Im **Induktionsschritt** (englisch: induction step) wird $A(n) \Rightarrow A(n+1)$ gezeigt. Dabei heißt $A(n)$ die **Induktionsannahme** oder **Induktionsvoraussetzung** (englisch: induction hypothesis). Bei Bedarf kann sogar auf alle vorgehenden Aussagen $A(n_0), \dots, A(n)$ zurückgegriffen werden, also $A(n_0) \wedge \cdots \wedge A(n) \Rightarrow A(n + n_0)$ gezeigt werden.

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 $\frac{1}{2}n\,(n+1)$.

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

Induktionsanfang bei $n_0 = 1$: $A(1)$ lautet: $\sum_{j=1}^{1} j = \frac{1}{2} \cdot 1 \cdot 2$, was eine wahre Aussage ist. Wir zeigen nun im Induktionsschritt, dass $A(n)$ auch $A(n+1)$ impliziert:

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

was $A(n+1)$ entspricht. △

Ende der Vorlesung 2

Ende der Woche 1

§ 4 Mengenlehre

Literatur: Deiser, 2022b, Kapitel 1.2, Jänich, 2008, Kapitel 1.1, Jänich, 2008, Kapitel 6

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:

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

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

$\mathbb{N} \coloneqq \{1, 2, 3, \ldots\}$	Menge der **natürlichen Zahlen**[26],
$\mathbb{N}_0 \coloneqq \{0, 1, 2, 3, \ldots\}$	Menge der **natürlichen Zahlen mit Null**,
$\mathbb{Z} \coloneqq \{0, 1, -1, 2, -2, \ldots\}$	Menge der **ganzen Zahlen**[27],
$\widetilde{\mathbb{Q}} \coloneqq \left\{\frac{m}{n} \,\middle|\, m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\}\right\}$	(vorläufige) Menge der **rationalen Zahlen**[28],
$\mathbb{R}$	Menge der **reellen Zahlen**[29],
$\mathbb{C} \coloneqq \{a + b\,i \mid a, b \in \mathbb{R}\}$	Menge der **komplexen Zahlen**[30],

die hier nur informell definiert werden. Für die wirkliche Definition der rationalen Zahlen $\mathbb{Q}$ verweisen wir auf das Ende von § 5. Elementare Eigenschaften der komplexen Zahlen $\mathbb{C}$ werden in Anhang A besprochen.

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

$$Y \coloneqq \left\{x \in X \,\middle|\, A(x)\right\} \tag{4.1}$$

betrachten, bestehend aus den Elementen von X, für die $A(x)$ 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 zu, X als die Menge aller Mengen zu definieren. Wählen wir dann $A(x)$ als die Aussageform „enthält sich nicht selbst“, so definiert

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

also die „Menge aller Mengen, die sich nicht selbst enthalten“. Stellen wir jetzt die Frage, ob R sich selbst enthält, so erkennen wir das Problem:

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

- Falls R sich nicht selbst enthält ($R \notin R$), dann erfüllt R die Komprehensionsbedingung $R \notin R$ nicht, also gilt $R \in R$.

[26] englisch: natural numbers
[27] englisch: integer numbers, lateinisch: integer: ganz, unversehrt
[28] englisch: rational numbers, lateinisch: ratio: Verhältnis
[29] englisch: real numbers
[30] englisch: complex numbers

In Kurzform erhalten wir den Widerspruch $R \in R \Leftrightarrow R \notin R$. 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.[31]

Die Auflösung in der modernen, 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[32] nicht behandeln und verweisen auf spätere Spezialveranstaltungen. Wir weisen aber darauf hin, dass die Mengenkomprehension (4.1) in Form des sogenannten Aussonderungsaxioms als Konstruktionsprinzip von Mengen weiterhin vorkommt. Wesentlich ist nur eben, dass der Grundbereich X der Aussageform A eine Menge im Sinne der ZF-Axiome sein muss.[33]

Intervalle lassen sich beispielsweise über Mengenkomprehension definieren:

Beispiel 4.1 (Mengenkomprehension).
Es seien $a, b \in \mathbb{R}$. Dann heißt

$[a, b] := \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ **abgeschlossenes Intervall**[34],

$(a, b] := \{x \in \mathbb{R} \mid a < x \leqslant b\}$ **links offenes, rechts abgeschlossenes Intervall**[35],

$[a, b) := \{x \in \mathbb{R} \mid a \leqslant x < b\}$ **links abgeschlossenes, rechts offenes Intervall**[36],

$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$ **offenes Intervall**[37],

$[a, \infty) := \{x \in \mathbb{R} \mid a \leqslant x\}$ **rechtsseitig unendliches abgeschlossenes Intervall**[38],

$(a, \infty) := \{x \in \mathbb{R} \mid a < x\}$ **rechtsseitig unendliches offenes Intervall**[39],

$(-\infty, b] := \{x \in \mathbb{R} \mid x \leqslant b\}$ **linksseitig unendliches abgeschlossenes Intervall**[40],

$(-\infty, b) := \{x \in \mathbb{R} \mid x < b\}$ **linksseitig unendliches offenes Intervall**[41],

$(-\infty, \infty) := \{x \in \mathbb{R} \mid \top\} = \mathbb{R}$ **beidseitig unendliches Intervall**[42].

Dabei ist $\{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ eine gebräuchliche Kurzschreibweise für $\{x \in \mathbb{R} \mid a \leqslant x \ \wedge\ x \leqslant b\}$. 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 sind leer, wenn $b < a$ bzw. $b \leqslant a$ gilt. Die Bedeutung der Eigenschaften **offen** (englisch: open) und **abgeschlossen** (englisch: closed) wird in der Lehrveranstaltung *Analysis* behandelt.

Wir definieren für $a, b \in \mathbb{Z}$ auch

$[\![a, b]\!] := [a, b] \cap \mathbb{Z}$ **ganzzahliges Intervall** (englisch: integer interval). △

[31] 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?

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

[33] Ist der Grundbereich keine Menge, so landet man beim Begriff der **Klasse** (englisch: class), siehe etwa Deiser, 2022a, Kapitel 3. Ein wichtiges Beispiel ist die **Klasse aller Mengen** (englisch: class of all sets).

[34] englisch: closed interval

[35] englisch: left-open, right-closed interval

[36] englisch: left-closed, right-open interval

[37] 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.

[38] englisch: unbounded above, closed interval

[39] englisch: unbounded above, open interval

[40] englisch: unbounded below, closed interval

[41] englisch: unbounded below, open interval

[42] englisch: unbounded above and below interval

Definition 4.2 (Teilmenge, Obermenge).
Für Mengen A und B definieren wir:

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

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

Die Teilmengenbeziehung $\subseteq$ zwischen Mengen heißt auch **Inklusion** (englisch: inclusion).[43] △

Beispielsweise erzeugt die Mengenkomprehension (4.1) immer eine Teilmenge $Y \subseteq X$. Außerdem gelten die echten Inklusionen

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

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) $\emptyset$.

Definition 4.3 (Schnitt, disjunkte Mengen, Vereinigung, Differenz, symmetrische Differenz).

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

$$\bigcap \mathcal{U} \coloneqq \big\{x \,\big|\, \forall U \in \mathcal{U} \; (x \in U)\big\} \tag{4.2}$$

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

$$\bigcap_{i \in I} U_i \coloneqq \big\{x \,\big|\, \forall i \in I \; (x \in U_i)\big\}. \tag{4.3}$$

Besteht speziell $\mathcal{U} = \{U_1, U_2\}$ aus nur zwei Elementen, so schreiben wir auch

$$U_1 \cap U_2 \coloneqq \big\{x \,\big|\, x \in U_1 \;\wedge\; x \in U_2\big\}. \tag{4.4}$$

Gilt $\bigcap \mathcal{U} = \emptyset$ bzw. $\bigcap_{i \in I} U_i = \emptyset$ bzw. $U_1 \cap U_2 = \emptyset$, so heißen die Elemente von $\mathcal{U}$ bzw. die Mengen U_i bzw. die Mengen U_1 und U_2 **disjunkt** (englisch: disjoint).

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

$$\bigcup \mathcal{U} \coloneqq \big\{x \,\big|\, \exists U \in \mathcal{U} \; (x \in U)\big\} \tag{4.5}$$

die **Vereinigungmenge** oder die **Vereinigung** (englisch: union) von $\mathcal{U}$. Sind die Elemente von $\mathcal{U}$ über eine Indexmenge I indiziert, gilt also $\mathcal{U} = \{U_i \,|\, i \in I\}$, so schreiben wir auch

$$\bigcup_{i \in I} U_i \coloneqq \big\{x \,\big|\, \exists i \in I \; (x \in U_i)\big\}. \tag{4.6}$$

Besteht speziell $\mathcal{U} = \{U_1, U_2\}$ aus nur zwei Elementen, so schreiben wir auch

$$U_1 \cup U_2 \coloneqq \big\{x \,\big|\, x \in U_1 \;\vee\; x \in U_2\big\}. \tag{4.7}$$

[43] lateinisch: includere: einschließen

$\triangle$

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

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

$$X \setminus Y := \{x \in X \mid x \notin Y\}, \tag{4.8}$$

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

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

$$X \triangle Y := (X \setminus Y) \cup (Y \setminus X). \tag{4.9}$$

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

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

$$A^c := X \setminus A = \{x \in X \mid x \notin A\}. \tag{4.10}$$

Da die Menge X im Symbol A^c nicht angegeben wird, muss sie dabei aus dem Zusammenhang klar sein. $\triangle$

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

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

$$X \cap Y = Y \cap X \qquad \textbf{Kommutativität von } \cap \tag{4.11a}$$
$$X \cup Y = Y \cup X \qquad \textbf{Kommutativität von } \cup \tag{4.11b}$$

$$(X \cap Y) \cap Z = X \cap (Y \cap Z) \qquad \textbf{Assoziativität von } \cap \tag{4.12a}$$
$$(X \cup Y) \cup Z = X \cup (Y \cup Z) \qquad \textbf{Assoziativität von } \cup \tag{4.12b}$$

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z) \qquad \textbf{Distributivität} \tag{4.13a}$$
$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z) \qquad \textbf{Distributivität} \tag{4.13b}$$

$$X \setminus Y = X \setminus (X \cap Y) \tag{4.14}$$

$$X \cap Y = X \quad \Leftrightarrow \quad X \subseteq Y \tag{4.15a}$$
$$X \cup Y = Y \quad \Leftrightarrow \quad X \subseteq Y \tag{4.15b}$$

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

$$(A \cap B)^c = A^c \cup B^c \qquad \textbf{De Morgansches Gesetz} \tag{4.16a}$$
$$(A \cup B)^c = A^c \cap B^c \qquad \textbf{De Morgansches Gesetz} \tag{4.16b}$$

$$(A^c)^c = A \qquad \text{Komplementbildung ist } \textbf{involutorisch}^{44} \tag{4.17}$$

$$A \subseteq B \quad \Leftrightarrow \quad B^c \subseteq A^c \tag{4.18}$$

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

Zur Vereinfachung der Notation vereinbaren wir auch hier wieder Bindungsregeln:

$$\cdot^c \text{ bindet stärker als } \setminus \text{ bindet stärker als } \cap \text{ bindet stärker als } \cup, \tag{4.19}$$

wodurch wir beispielsweise das erste Distributivgesetz auch als $X \cap (Y \cup Z) = X \cap Y \cup X \cap Z$ schreiben könnten.

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

$$\mathcal{P}(A) := \{B \mid B \subseteq A\} \tag{4.20}$$

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

In der axiomatischen Mengenlehre nach Zermelo und Fraenkel gibt es das Potenzmengenaxiom, das garantiert, dass jede Menge eine Potenzmenge besitzt.

Beispiel 4.7 (Potenzmenge)**.**

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

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

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

Definition 4.8 (kartesisches Produkt endlich vieler Mengen)**.**

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

$$A \times B := \{(a, b) \mid a \in A \ \wedge\ b \in B\}. \tag{4.21}$$

Die Elemente des kartesischen Produkts heißen **geordnete Paare** (englisch: ordered pairs) oder einfach **Paare** (englisch: pairs) (a, b).

(ii) Analog können wir auch das kartesische Produkt von mehr als zwei Mengen definieren, etwa $A \times B \times C$, dessen Elemente **Tripel** (englisch: triplets) (a, b, c) sind. Allgemeiner heißen die Elemente $(a_1, a_2, \ldots, a_n)$ des Produkts $\bigtimes_{i=1}^n A_i$ von $n \geqslant 2$ Mengen **n-Tupel** (englisch: n-tuples). Dabei gilt $a_i \in A_i$ für $i = 1, \ldots, n$.

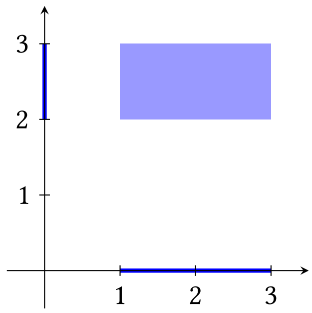
(iii) Wir schreiben $A^2 = A \times A$ und allgemeiner $A^n = \bigtimes_{i=1}^n A$ für das kartesische Produkt einer Menge A mit sich selbst. △

Beispiel 4.9 (kartesisches Produkt)**.**

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

[44]auch: selbst-invers, englisch: involutory, self-inverse

(*ii*) Für Intervalle $A = [1,3]$ und $B = [2,3]$ können wir das **mehrdimensionale Intervall** (englisch: multi-dimensional interval) $A \times B = \{(x_1, x_2) \,|\, 1 \leqslant x_1 \leqslant 3 \wedge 2 \leqslant x_2 \leqslant 3\} \subseteq \mathbb{R} \times \mathbb{R}$ wie folgt illustrieren:



△

Ende der Vorlesung 3

§ 5 Relationen

Literatur: Deiser, 2022b, Kapitel 1.3

Relationen[45] geben Beziehungen zwischen Objekten an wie beispielsweise $1 \leqslant 3$ oder $5 \in \mathbb{N}$ oder $3 \mid 756$ („3 teilt 756").

Definition 5.1 (Relation)**.**

Es seien X und Y Mengen. Ist $R \subseteq X \times Y$, so heißt (R, X, Y) eine **Relation** (englisch: relation) **zwischen** X und Y. Die Menge R heißt der **Graph der Relation** (englisch: graph of a relation). 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 selbst sei die Relation. Statt $(x, y) \in R$ schreiben wir auch $x \, R \, y$, um die Lesart „x steht in Relation zu y" zu erleichtern.

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) „Die Zahl x teilt die Zahl y." auf der Menge $X = Y = \{0, 1, 2, \dots, 10\} = [\![0, 10]\!]$ dar:

[45] lateinisch: relatio: Verhältnis, Beziehung

$x \mid y$	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 $R = \big\{(x, y) \in \mathbb{R}^2 \,\big|\, x \leqslant y\big\}$ 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 $R = \big\{(A, B) \in \mathcal{P}(X) \times \mathcal{P}(X) \,\big|\, A \subseteq B\big\}$ die **Inklusionsrelation** (englisch: inclusion relation).

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

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

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

$$S \circ R \coloneqq \big\{(x, z) \in X \times Z \,\big|\, \exists y \in Y \text{ mit } (x, y) \in R \text{ und } (y, z) \in S\big\} \tag{5.2}$$

die **Komposition** (englisch: composition, lateinisch: componere: zusammenstellen), die **Hintereinanderausführung**, die **Verknüpfung** 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 „Onkel sein von“ ausdrücken?

Definition 5.4 (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

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

definiert ist. △

Quizfrage 5.3: Wie bezeichnet man die Umkehrrelationen von „kleiner oder gleich sein als“, „Teilmenge sein von“ bzw. „Teiler sein von“?

Quizfrage 5.4: Wie könnte man die Umkehrrelationen der Teilbarkeitsrelation auf $\mathbb{Z}$ bezeichnen?

Wir definieren nun einige wichtige Eigenschaften, die Relationen auf einer Menge besitzen können.

Definition 5.5 (Eigenschaften homogener Relationen)**.**
Es sei X eine Menge und (R, X, X) 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 **symmetrisch** (englisch: symmetric), wenn gilt:

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

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

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

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

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

(v) 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 5.5: Die Reflexivität von R kann man auch als $\mathrm{id}_X \subseteq R$ ausdrücken. Wie sieht das für die anderen Eigenschaften aus?

Beispiel 5.6 (Eigenschaften homogener Relationen)**.**

- Die Teilbarkeitsrelation $\mid$ auf $\mathbb{Z}$ ist reflexiv und transitiv, aber nicht symmetrisch, antisymmetrisch oder total.
- Die Teilbarkeitsrelation $\mid$ auf $\mathbb{N}_0$ ist reflexiv, antisymmetrisch und transitiv, aber nicht symmetrisch oder total.
- Die Relation „x liebt y“ auf einer Menge von Personen hat in der Regel keine der fünf genannten Eigenschaften. △

§ 5.1 Ordnungsrelationen

Definition 5.7 (Ordnungsrelation)**.**
Es sei X eine Menge.

(i) Eine Relation (R, X, X) auf X heißt eine **Ordnungsrelation**, **Halbordnung** oder **partielle Ordnung** (englisch: partial ordering), wenn sie reflexiv, antisymmetrisch und transitiv ist. Das Paar (X, R) heißt dann eine **halbgeordnete Menge** (englisch: partially ordered set).

(ii) Ist die Relation R zusätzlich total, dann heißt sie eine **Totalordnung** (englisch: total ordering). Das Paar (X, R) heißt dann eine **totalgeordnete Menge** (englisch: totally ordered set). △

Ordnungsrelationen werden oft mit Symbolen wie $\leqslant$, $\preccurlyeq$ 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:

$$x \preccurlyeq x, \tag{5.3a}$$
$$x \preccurlyeq y \text{ und } y \preccurlyeq x \quad \Rightarrow \quad x = y, \tag{5.3b}$$
$$x \preccurlyeq y \text{ und } y \preccurlyeq z \quad \Rightarrow \quad x \preccurlyeq z. \tag{5.3c}$$

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 unbedingt.

Beispiel 5.8 (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.

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

(iv) Die Inklusionsrelation $\subseteq$ ist eine Halbordnung auf der Potenzmenge $\mathcal{P}(X)$ jeder beliebigen Menge X. 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 $\mathbb{N}$. △

Lemma 5.9 (Halbordnungen $\preccurlyeq$ und $\succcurlyeq$)**.**
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 Teil von Hausaufgabe I-2.3. □

Definition 5.10 (Vergleichbarkeit, obere und untere Schranken, Supremum und Infimum, maximale und minimale Elemente, Maximum und Minimum)**.**
Es sei X mit der Relation $\preccurlyeq$ eine halbgeordnete Menge.

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

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

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

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

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

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

$$b \in A, \text{und für alle } x \in A \text{ gilt: } b \preccurlyeq x \implies x = b. \qquad \text{(„Kein Element von } A \text{ ist größer.“)}$$

(*v*) $b \in X$ heißt ein **Maximum** (englisch: maximum) von $A \subseteq X$, wenn gilt:

$$b \in A, \text{und für alle } x \in A \text{ gilt: } x \preccurlyeq b. \qquad \text{(„Ganz } A \text{ ist höchstens so groß.“)}$$

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

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

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

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

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

$$a \in A, \text{ und für alle } x \in A \text{ gilt: } x \preccurlyeq a \implies x = a. \qquad \text{(„Kein Element von } A \text{ ist kleiner.“)}$$

(*ix*) $a \in X$ heißt ein **Minimum** (englisch: minimum) von $A \subseteq X$, wenn gilt:

$$a \in A, \text{und für alle } x \in A \text{ gilt: } a \preccurlyeq x. \qquad \text{(„Ganz } A \text{ ist mindestens so groß.“)}$$

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

Wir zeigen nun einige ausgewählte Eigenschaften.

Lemma 5.11 (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.

(*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.

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, $b \in X$ und $\overline{b} \in X$ seien beides Maxima von A. Dann gehören b und $\overline{b}$ beide zu A. Da b ein Maximum von A ist, gilt $\overline{b} \preccurlyeq b$. Da $\overline{b}$ ein Maximum von A ist, gilt $b \preccurlyeq \overline{b}$. Aufgrund der Antisymmetrie von $\preccurlyeq$ folgt nun $b = \overline{b}$

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

Beispiel 5.12 (Schranken, extremale Elemente, Maxima und Minima, Suprema und Infima)**.**

(*i*) In den natürlichen Zahlen $\mathbb{N}$ mit der gewöhnlichen Totalordnung $\preccurlyeq$ ist die Zahl 1 das Minimum und damit das Infimum. Eine obere Schranke existiert nicht.

(*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 $A = \mathcal{P}(X) \setminus \{\emptyset\}$ das Infimum $\emptyset$, aber kein Minimum. Die minimalen Elemente von A sind genau die einelementigen Teilmengen von X. △

Quizfrage 5.6: Können Sie sich eine Menge mit einer Halbordnung oder einer totalen Ordnung vorstellen, die kein maximales Element besitzt?

§ 5.2 Äquivalenzrelation

Definition 5.13 (Äquivalenzrelation)**.**
Es sei X eine Menge. Eine Relation (R, X, X) auf X heißt eine **Äquivalenzrelation** (englisch: equivalence relation), wenn sie reflexiv, symmetrisch und transitiv ist. Elemente $x, y \in X$, die $x \, R \, y$ erfüllen, heißen **(zueinander) äquivalent** (englisch: equivalent). △

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:

$$x \sim x, \tag{5.4a}$$

$$x \sim y \quad \Rightarrow \quad y \sim x, \tag{5.4b}$$

$$x \sim y \text{ und } y \sim z \quad \Rightarrow \quad x \sim z. \tag{5.4c}$$

Die Idee von Äquivalenzrelationen ist es, die Elemente einer Menge, die eine bestimmte Eigenschaft gemeinsam haben, zusammenzugruppieren und als gleichwertig zu betrachten.

Beispiel 5.14 (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 $X = \mathbb{Z}$ ist durch

$$x \overset{m}{\equiv} y \quad \Leftrightarrow \quad \exists n \in \mathbb{Z}\, (x - y = n\, m) \tag{5.5}$$

eine Äquivalenzrelation erklärt (**Quizfrage 5.7:** Details?). Anders ausgedrückt, x und y unterscheiden sich nur um ein Vielfaches von m, also, $m \mid (x - y)$. Diese Relation heißt **Kongruenzrelation modulo** m (englisch: congruence relation modulo m).[46] △

Definition 5.15 (Äquivalenzklasse, Repräsentant, Repräsentantensystem).
Es sei X eine Menge mit der Äquivalenzrelation $\sim$.

(*i*) Für $x \in X$ heißt die Menge

$$[x] \coloneqq \{y \in X \mid y \sim x\} \tag{5.6}$$

die **Äquivalenzklasse** (englisch: equivalence class) von x bzgl. $\sim$. Statt $[x]$ schreibt man manchmal auch $[x]_\sim$ oder auch $x / \sim$.

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

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

Beispiel 5.16 (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 einzige 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 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).[47] Die Restklasse von $a \in \mathbb{Z}$ modulo m ist also

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

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

[46] Oft wird diese Relation statt $x \overset{m}{\equiv} y$ als $x \equiv y \pmod m$ geschrieben.
[47] 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.

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

$$\begin{aligned}[0] &= \{y \in \mathbb{Z} \,|\, y \overset{2}{\equiv} 0\} \\ &= \{y \in \mathbb{Z} \,|\, \exists n \in \mathbb{Z}\; (y - 0 = 2\,n)\} \\ &= \{y \in \mathbb{Z} \,|\, y \text{ ist gerade}\}, \\[1ex] [1] &= \{y \in \mathbb{Z} \,|\, y \overset{2}{\equiv} 1\} \\ &= \{y \in \mathbb{Z} \,|\, \exists n \in \mathbb{Z}\; (y - 1 = 2\,n)\} \\ &= \{y \in \mathbb{Z} \,|\, y \text{ ist ungerade}\}.\end{aligned}$$

Das natürliche Repräsentantensystem ist $\{0, 1\}$, ein anderes ist $\{-2, 4339\}$. △

Satz 5.17 (Äquivalenzklassen sind 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 $\bar{x}$ ein beliebiges Element aus $[x]$. Dann gilt

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

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

Definition 5.18 (Partition)**.**
Es sei X eine nichtleere Menge und $\mathcal{U}$ eine Menge von Teilmengen von X, also $\mathcal{U} \subseteq \mathcal{P}(X)$. $\mathcal{U}$ heißt eine **Partition** (englisch: partition) oder **disjunkte Zerlegung** von X, wenn gilt:

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

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

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

Zu Eigenschaft (*i*) sagen wir auch, dass die Mengen in $\mathcal{U}$ die Menge X **überdecken** (englisch: to cover) oder eine **Überdeckung** (englisch: cover, covering) von X darstellen. Zu Eigenschaft (*ii*) sagen wir, dass die Mengen in $\mathcal{U}$ **paarweise disjunkt** (englisch: pairwise disjoint) sind.

Satz 5.19 (Partitionen werden genau durch Äquivalenzrelationen erzeugt)**.**

(*i*) Es sei X eine nichtleere Menge mit der Äquivalenzrelation $\sim$. Dann bildet die Menge der Äquivalenzklassen $\{[x] \,|\, x \in X\}$ eine Partition von X.

(*ii*) Es sei X eine nichtleere Menge und $\mathcal{U}$ eine Partition von X. Dann gibt es eine eindeutig bestimmte Äquivalenzrelation $\sim$, sodass $\mathcal{U}$ genau aus den Äquivalenzklassen von $\sim$ besteht.

Wir könnten diesen Satz etwas ungenau auch so ausdrücken, dass die Partition einer Menge X „dasselbe“ ist wie eine Äquivalenzrelationen auf X.

Beweis. Aussage (i): Zur Abkürzung sei $\mathcal{U} := \big\{[x] \,\big|\, x \in X\big\}$ die Menge der Äquivalenzklassen. Wir weisen die Eigenschaften der Definition 5.18 nach. Zunächst ist jedes $x \in X$ Element seiner Äquivalenzklasse $[x]$, da ja $x \sim x$ gilt. Das zeigt Eigenschaft (i). Nach Satz 5.17 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 Teil von Hausaufgabe I-2.3. □

Definition 5.20 (Quotientenmenge, Invarianz).
Es sei X eine nichtleere Menge mit der Äquivalenzrelation $\sim$.

(i) Die Menge der Äquivalenzklassen

$$X / \sim := \big\{[x] \,\big|\, x \in X\big\} \tag{5.7}$$

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

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

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

Beispiel 5.21 (wohldefinierte Aussageformen).

(i) Die Aussageform „x ist eine gerade ganze Zahl“ ist unter der Kongruenzrelation $\overset{2}{\equiv}$ wohldefiniert, da die Restklassen $[0]$ und $[1]$ jeweils nur aus geraden bzw. nur aus ungeraden ganzen Zahlen bestehen.

(ii) Dieselbe Aussageform 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 zu Beginn von § 4 vorläufig als

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

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

$$\frac{m_1}{n_1} \sim \frac{m_2}{n_2} \quad \Leftrightarrow \quad m_1 \cdot n_2 = m_2 \cdot n_1. \tag{5.8}$$

Das führt uns zur Definition

$$\mathbb{Q} := \left\{\frac{m}{n} \,\middle|\, m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\}\right\} \Big/ \sim \tag{5.9}$$

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

Ende der Vorlesung 4

Ende der Woche 2

§ 6 Abbildungen

Literatur: Deiser, 2022b, Kapitel 1.3, Deiser, 2022b, 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 auch 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. Statt $x f y$ schreiben wir $y = f(x)$ oder $x \mapsto f(x)$ und sagen, x werde **abgebildet auf** $f(x)$. Auch die kompakten Schreibweisen

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

für die Definition einer Funktion sind üblich.

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

Die Menge

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

heißt der **Graph** (englisch: graph) der Funktion $f \colon X \to Y$.[48]

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.

[48] Der Begriff des Graphen einer Funktion stimmt also überein mit dem Begriff des Graphen der Funktion als Relation, vgl. Definition 5.1.

(*ii*) Es seien X und Y Mengen mit $X \subseteq Y$. Dann heißt die Abbildung $i_{X\to Y}$ mit

$$X \ni x \mapsto i_{X\to Y}(x) := x$$

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.

(*iii*) Im Fall $X = Y$ heißt die kanonische Einbettung auch die **Identität** (englisch: identity) oder **identische Abbildung** (englisch: identity map) von X in Y und wird mit id_X bezeichnet, also

$$X \ni x \mapsto \mathrm{id}_X(x) := x.$$

Der Graph von id_X ist also gerade die Diagonale Δ_X, siehe (5.1). △

Definition 6.4 (Bild, Einschränkung, Fortsetzung).
Es sei $f\colon X \to Y$ eine Funktion.[49]

(*i*) Für $A \subseteq X$ heißt

$$f(A) := \{f(x) \,|\, x \in A\} \tag{6.2}$$

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 $f|_A$

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

die **Einschränkung** oder **Restriktion** (englisch: restriction, lateinisch: restringere: zurückziehen) von f auf A.

(*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

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

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

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

bei der gegenüber f nur die Zielmenge ersetzt wird.

(*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. △

An Stelle von $f|_A$ schreibt man auch manchmal $f \restriction A$.

Beispiel 6.5 (Bild, Einschränkung, Fortsetzung).
Wir betrachten die Funktionen[50]

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

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

[49] Wir sagen damit insbesondere, dass X und Y Mengen sind.
[50] Hierbei bedeutet $\frac{\pi}{2}\mathbb{Z}$ die Menge der ganzzahligen Vielfachen von $\frac{\pi}{2}$.

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

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 $f\colon X \longrightarrow Y$ 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:

$$f\Big(\bigcup_{i\in I} X_i\Big) = \bigcup_{i\in I} f(X_i) \tag{6.4a}$$

$$f\Big(\bigcap_{i\in I} X_i\Big) \subseteq \bigcap_{i\in I} f(X_i) \tag{6.4b}$$

$$f^{-1}\Big(\bigcup_{j\in J} Y_j\Big) = \bigcup_{j\in J} f^{-1}(Y_j) \tag{6.4c}$$

$$f^{-1}\Big(\bigcap_{j\in J} Y_j\Big) = \bigcap_{j\in J} f^{-1}(Y_j) \tag{6.4d}$$

Beweis. Wir beweisen hier nur (6.4a) und (6.4c). Die Aussagen (6.4b) und (6.4d) sind Teil von Hausaufgabe I-3.1.

Zum Beweis von (6.4a):

$$\begin{aligned} & y \in f\Big(\bigcup_{i\in I} X_i\Big) \\ \Leftrightarrow\quad & \exists i \in I\ \exists x \in X_i\ (y = f(x)) && \text{nach Definition (4.6) der Vereinigungsmenge} \\ \Leftrightarrow\quad & \exists i \in I\ (y \in f(X_i)) && \text{nach Definition (6.2) des Bildes } f(X_i) \\ \Leftrightarrow\quad & y \in \bigcup_{i\in I} f(X_i) && \text{nach Definition (4.6) der Vereinigungsmenge.} \end{aligned}$$

Zum Beweis von (6.4c):

$$\begin{aligned} & x \in f^{-1}\Big(\bigcup_{j\in J} Y_j\Big) \\ \Leftrightarrow\quad & \exists y \in \bigcup_{j\in J} Y_j\ \big(y = f(x)\big) && \text{nach Definition (6.3) des Urbildes} \\ \Leftrightarrow\quad & \exists j \in J\ \exists y \in Y_j\ \big(y = f(x)\big) && \text{nach Definition (4.6) der Vereinigungsmenge} \\ \Leftrightarrow\quad & \exists j \in J\ \big(x \in f^{-1}(Y_j)\big) && \text{nach Definition (6.3) des Urbildes} \\ \Leftrightarrow\quad & x \in \bigcup_{j\in J} f^{-1}(Y_j) && \text{nach Definition (4.6) der Vereinigungsmenge.} \end{aligned}$$ □

Beispiel 6.9 (Bilder und Urbilder von Vereinigungen und Durchschnitten).
In (6.4b) gilt i. A. nicht die Gleichheit, wie folgendes Beispiel zeigt: Es sei

$$\mathbb{R}^2 \ni (x, y) \mapsto f(x, y) := x \in \mathbb{R}.$$

Für die Mengen $A := \{(0, 0)\}$ und $B = \{(0, 1)\}$ gilt

$$\begin{aligned} f(A \cap B) &= f(\emptyset) = \emptyset, \\ \text{aber} \quad f(A) \cap f(B) &= \{0\} \cap \{0\} = \{0\}. \end{aligned}$$

△

§ 6.1 Injektivität und Surjektivität

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

(i) **surjektiv** (englisch: surjective, onto) oder **Surjektion** oder **rechtstotal** (englisch: right-total), wenn $f(X) = Y$ gilt.[51] Man sagt auch, f bilde X **auf** Y ab.

Äquivalent dazu ist
$$\forall y \in Y \ (f^{-1}(\{y\}) \neq \emptyset)$$

(ii) **injektiv** (englisch: injective, one-to-one) oder **Injektion** oder **linkseindeutig** (englisch: left-unique), wenn für alle $x_1, x_2 \in X$ gilt: $f(x_1) = f(x_2) \ \Rightarrow \ x_1 = x_2$.[52]

Äquivalent dazu ist

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

(iii) **bijektiv** (englisch: bijective), wenn f surjektiv und injektiv ist.[53]

Äquivalent dazu ist
$$\forall y \in Y \ (f^{-1}(\{y\}) \text{ hat genau ein Element})$$

△

Als Substantive sind die Bezeichnungen **Surjektion** (englisch: surjection), **Injektion** (englisch: injection) und **Bijektion** (englisch: bijection) geläufig.

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

Lemma 6.11 (Bijektiv-Machen einer injektiven Funktion).
Es sei $f : X \to Y$ eine injektive Funktion. Dann ist $f|^{f(X)}$ (also die Einschränkung der Zielmenge auf die tatsächliche Bildmenge) bijektiv.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-3.2. □

Beispiel 6.12 (Injektivität, Surjektivität, Bijektivität).

[51] Die Surjektivität von f wird manchmal auch durch die Schreibweise $f : X \twoheadrightarrow Y$ ausgedrückt.
[52] Die Injektivität von f wird manchmal auch durch die Schreibweise $f : X \rightarrowtail Y$ ausgedrückt.
[53] Die Bijektivität von f wird manchmal auch durch die Schreibweise $f : X \rightarrowtail\!\!\!\!\twoheadrightarrow Y$ ausgedrückt.

(*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*) Sind X und Y Mengen mit $X \subseteq Y$, dann ist die kanonische Injektion $i_{X \to Y}$ injektiv. △

Definition 6.13 (Komposition von Funktionen)**.**
Es seien $f\colon X \to Y$ und $g\colon Y \to Z$ Funktionen. Die Funktion

$$X \ni x \mapsto h(x) \coloneqq g(f(x)) \in Z$$

heißt die **Komposition** (englisch: composition, lateinisch: componere: zusammenstellen), die **Hintereinanderausführung**, die **Verknüpfung** oder die **Verkettung** von f und g. Sie wird auch mit $h = g \circ f$ bezeichnet. Um die Reihenfolge klar zu benennen, sagt man auch „g **nach** f“. △

Wir können den Sachverhalt aus Definition 6.13 auch durch

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

illustrieren.

Die Voraussetzung, dass die Zielmenge von f mit der Definitionsmenge von g übereinstimmt, kann relaxiert werden. Die Komposition $g \circ f$ ist definiert, sofern $f(X) \subseteq Y$ gilt.

Beispiel 6.14 (Komposition von Funktionen)**.**
Es seien

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

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

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

△

Bemerkung 6.15 (Komposition mit der Identität und mit der kanonischen Einbettung).
Es sei $f\colon X \to Y$ eine Funktion. Dann gilt

$$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f, \tag{6.5}$$
$$f|_A = f \circ \iota_{A \to X} \qquad \text{für } A \subseteq X, \tag{6.6}$$
$$f = (\iota_{Y \to B} \circ f)|^Y \qquad \text{für } B \supseteq Y. \tag{6.7}$$

△

Lemma 6.16 (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. Für $x \in X$ gilt

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

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. □

Lemma 6.17 (Komposition injektiver und surjektiver Funktionen).
Es seien $f\colon X \to Y$ und $g\colon Y \to Z$ 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 $g \circ f$ surjektiv, so ist g surjektiv.

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 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. □

Folgerung 6.18 (Komposition zur Identität).
Es seien $f\colon X \to Y$ und $g\colon Y \to X$ 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 6.17, Aussagen (*iii*) und (*iv*) folgt daher, dass f injektiv und g surjektiv ist. □

Ende der Vorlesung 5

§ 6.2 Umkehrabbildung

Lemma 6.19 (Charakterisierung der Bijektivität)**.**
Es sei $f\colon X \to Y$ eine Funktion. Dann sind äquivalent:

(*i*) f ist bijektiv.

(*ii*) Für alle $y \in Y$ gibt es genau ein $x \in X$ mit der Eigenschaft $f(x) = y$.

(*iii*) Es existiert eine Abbildung $g\colon Y \to X$ mit der Eigenschaft $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$. Die Abbildung g ist eindeutig bestimmt und notwendig bijektiv.

Beweis. Aussage (*i*) $\Rightarrow$ Aussage (*ii*): Es sei f bijektiv, also surjektiv und injektiv. Ist $y \in Y$ beliebig, dann folgt aus der Surjektivität die Existenz eines $x_1 \in X$ mit $f(x_1) = y$. Ist $x_2 \in X$ ein weiteres Element mit $f(x_2) = y$, dann folgt aus der Injektivität $x_1 = x_2$.

Aussage (*ii*) $\Rightarrow$ Aussage (*iii*): Wir definieren die Abbildung $g\colon Y \to X$ wie folgt: Wir setzen für $y \in Y$ als $g(y)$ das nach Voraussetzung eindeutig definierte $x \in X$, für das $y = f(x)$ gilt. Für diese Funktion haben wir also $g(y) = x \Leftrightarrow f(x) = y$.

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

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.

Es sei nun $\widehat{g}\colon Y \to X$ eine weitere Funktion mit der Eigenschaft $f \circ \widehat{g} = \mathrm{id}_Y$. Dann gilt

$$\begin{aligned} g &= g \circ \mathrm{id}_Y && \text{nach (6.5)} \\ &= g \circ (f \circ \widehat{g}) && \text{nach Voraussetzung} \\ &= (g \circ f) \circ \widehat{g} && \text{nach Lemma 6.16} \\ &= \mathrm{id}_X \circ \widehat{g} && \text{nach Voraussetzung} \\ &= \widehat{g} && \text{nach (6.5).} \end{aligned}$$

Aussage (*iii*) $\Rightarrow$ Aussage (*i*): Die Abbildung $g \circ f = \mathrm{id}_X$ ist bijektiv, insbesondere injektiv. Aus Lemma 6.17 (*iii*) folgt also, dass f injektiv ist. Die Abbildung $f \circ g = \mathrm{id}_Y$ ist bijektiv, insbesondere surjektiv. Aus Lemma 6.17 (*iv*) folgt also, dass f surjektiv ist. □

Die Funktion $g\colon Y \to X$ aus Aussage (*iii*) heißt die **Umkehrfunktion**, **Umkehrabbildung**, **inverse Funktion** oder **inverse Abbildung** (englisch: inverse map) von f. Sie wird mit $f^{-1}\colon Y \to X$ bezeichnet. Für ihre Abbildungsvorschrift gilt $f^{-1}(y) = x \Leftrightarrow y = f(x)$. Die Funktion $f\colon X \to Y$ heißt **invertierbar** (englisch: invertible), wenn die Umkehrfunktion existiert. Nach Lemma 6.19 ist das genau dann der Fall, wenn f bijektiv ist.

Bemerkung 6.20 (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 6.21 (Umkehrfunktion der Komposition)**.**
Es seien $f: X \to Y$ und $g: Y \to Z$ bijektive Funktionen. Dann ist auch $g \circ f$ bijektiv, und die Umkehrfunktion ist gegeben durch

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}. \tag{6.8}$$

Quizfrage 6.2: Wie erklärt man sich anschaulich, dass sich bei der Umkehrfunktion die Reihenfolge ändert?

Beweis. Die Bijektivität von $g \circ f$ folgt sofort aus Lemma 6.17, Aussagen (i) und (ii). Wir wissen über die Abbildungsvorschrift

$$\begin{aligned} & (g \circ f)^{-1}(z) = x \\ \Leftrightarrow\quad & (g \circ f)(x) = z \\ \Leftrightarrow\quad & g(f(x)) = z \\ \Leftrightarrow\quad & f(x) = g^{-1}(z) \\ \Leftrightarrow\quad & x = f^{-1}(g^{-1}(z)) \\ \Leftrightarrow\quad & x = (f^{-1} \circ g^{-1})(z) \end{aligned}$$

für alle $x \in X$ und $z \in Z$. Das bedeutet aber $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

Lemma 6.22 (Charakterisierung der Injektivität)**.**
Es sei $f: X \to Y$ eine Funktion und $X \neq \emptyset$. Dann sind äquivalent:

(i) f ist injektiv.

(ii) Es existiert eine Abbildung $g: Y \to X$ 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.

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

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

Damit ist $\overline{g} \circ f = \mathrm{id}_X$ gezeigt. Aufgrund von Folgerung 6.18 ist $\overline{g}$ surjektiv. Wir setzen nun $\overline{g}: f(X) \to X$ zu $g: X \to X$ fort. Dazu wählen wir irgendein $x_0 \in X$ und setzen $g(y) \coloneqq \overline{g}(y)$ 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 $\overline{g}$.

Angenommen, $h: Y \to X$ 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. □

Eine analoge Charakterisierung der Surjektivität folgt erst in Satz 6.34, weil wir dafür interessanterweise das Auswahlaxiom benötigen.

§ 6.3 Mächtigkeit von Mengen

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

Definition 6.23 (Gleichmächtigkeit von Mengen).
Es seien X und Y Mengen. Wir sagen, X sei **gleichmächtig** (englisch: equinumerous) zu Y, wenn es eine bijektive Abbildung $f\colon X \to Y$ gibt. Wir schreiben in diesem Fall $X \sim Y$. △

Die Gleichmächtigkeit von Mengen ist eine Äquivalenzrelation auf der Klasse aller Mengen, siehe Hausaufgabe I-3.3. Die Äquivalenzklassen heißen **Kardinalzahlen** (englisch: cardinal numbers).

Definition 6.24 (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: $\#X = n$.[54]

(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 leere Menge $\emptyset$ ist nur zu sich selbst gleichmächtig. Sie ist die einzige Menge mit Mächtigkeit 0.

Beispiel 6.25 (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.
(Beweis in der Lehrveranstaltung *Analysis*)

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

(iv) Die Menge der reellen Zahlen $\mathbb{R}$ ist überabzählbar.
(Beweis in der Lehrveranstaltung *Analysis*) △

Lemma 6.26 (Veränderung der Kardinalität um 1).
Es sei X eine endliche Menge und $x \in X$. Dann gilt

$$\#X = \#(X \setminus \{x\}) + 1. \tag{6.9}$$

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

[54] In dieser Lehrveranstaltung verwenden wir das Symbol # nur für endliche Mengen.

Satz 6.27 (Funktionen auf endlichen Mengen)**.**

Es seien X und Y endliche, gleichmächtige Mengen und $f\colon X \to Y$ eine Funktion. Dann sind äquivalent:

(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 $n = \#X = \#Y$. 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 6.26 $\#X \setminus \{x\} = \#Y \setminus \{y\} = n$.

Wir bezeichnen mit $\widehat{f}\colon X \setminus \{x\} \to Y \setminus \{y\}$ 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 $\widehat{f}$ die Injektivität von f. Nach Induktionsvoraussetzung ist $\widehat{f}$ daher auch surjektiv, alle Elemente von $Y \setminus \{y\}$ liegen also im Bild von $\widehat{f}$ 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 nur eine mögliche Abbildung, diese ist ebenfalls 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 $\overline{y} \in Y$, sodass das Urbild $f^{-1}(\{\overline{y}\})$ (mindestens) aus zwei verschiedenen Elemente besteht, sagen wir $\overline{x}, \overline{\overline{x}} \in f^{-1}(\{\overline{y}\})$ und $\overline{x} \neq \overline{\overline{x}}$. Wir wählen außerdem ein $\widehat{y} \in Y \setminus \{\overline{y}\}$ aus, was wegen $\#Y = n+1 \geqslant 2$ möglich ist. Dazu existiert ein $\widehat{x}$ mit $f(\widehat{x}) = \widehat{y}$. Wegen $\widehat{y} \neq \overline{y}$ ist $\widehat{x} \neq \overline{x}$ und $\widehat{x} \neq \overline{\overline{x}}$.

Wir konstruieren nun eine Funktion $\widehat{f}\colon X \setminus \{\overline{x}\} \to Y \setminus \{\overline{y}\}$ durch

$$\widehat{f}(x) \coloneqq \begin{cases} f(x) & \text{im Fall } f(x) \neq \overline{y}, \\ \widehat{y} & \text{im Fall } f(x) = \overline{y}. \end{cases}$$

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

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

(2) Außerdem gilt $f(\overline{\overline{x}}) = \overline{y}$, also $\widehat{f}(\overline{\overline{x}}) = \widehat{y}$.

Aufgrund von Lemma 6.26 gilt wieder $\#X \setminus \{\overline{x}\} = \#Y \setminus \{\overline{y}\} = n$. Nach Induktionsvoraussetzung ist $\widehat{f}$ daher auch injektiv. Jedoch enthält $\widehat{f}^{-1}(\{\widehat{y}\})$ neben $\widehat{x}$ auch noch mindestens das weitere Element $\overline{\overline{x}} \in f^{-1}(\{\overline{y}\})$. Das steht im Widerspruch zur Injektivität von $\widehat{f}$.

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 6.27 ist falsch, wenn X und Y zwar gleichmächtig, aber nicht endlich sind, siehe Hausaufgabe I-3.3.

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

Definition 6.28 (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 bijektive Abbildung von X auf eine Teilmenge von Y gibt. Wir schreiben in diesem Fall $X \preccurlyeq Y$. △

Die Reflexivität und Transitivität der Relation $\preccurlyeq$ sind leicht einzusehen. (**Quizfrage 6.3:** Details?) Der Beweis der Antisymmetrie ist jedoch aufwändig und erfordert den Satz von Cantor-Bernstein-Schröder, der äquivalent zum Auswahlaxiom (siehe § 6.5) ist. Unter Zuhilfenahme des Auswahlaxioms kann man außerdem zeigen, dass zwei Mengen bzgl. $\preccurlyeq$ stets vergleichbar sind. Es folgt, dass $\preccurlyeq$ sogar eine totale Ordnung auf der Menge aller Kardinalzahlen ist.

§ 6.4 Familien und Folgen

Definition 6.29 (Familie von Elementen, Teilfamilie, Oberfamilie, Folge, endliche Folge)**.**
Es seien I und Y Mengen.

(i) Eine Abbildung
$$I \ni i \mapsto y(i) := y_i \in Y$$
heißt eine **Familie von Elementen** (englisch: family of elements) aus Y mit der **Indexmenge** (englisch: index set) I. Kurz wird diese auch mit $(y_i)_{i \in I}$ bezeichnet.

(ii) 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 $(y_i)_{i \in I_0}$.

(iii) 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). 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.

(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) der Kardinalität n. Ist speziell $I = [\![1, n]\!]$, so heißt $(y_i)_{i \in I}$ eine **endliche Folge** (englisch: finite sequence) in Y der Kardinalität n. △

Bemerkung 6.30 (Familien und Mengen)**.**

(i) Im Unterschied zu einer Menge kann eine Familie $(y_i)_{i \in I}$ Elemente mehrfach enthalten.

(ii) Jeder Familie $(y_i)_{i \in I}$ von Elementen aus Y können wir eine Menge $\{y_i \mid i \in I\} \subseteq Y$ zuordnen.

(iii) Wir können eine endliche Folge auch als **n-Tupel** (englisch: n-tuple) $(y_1, y_2, \ldots, y_n)$ notieren. Während $(y_i)_{i \in I}$ keine Reihenfolge hat (da I als Menge ungeordnet ist), hat ein n-Tupel jedoch eine festgelegte Reihenfolge. △

Beispiel 6.31 (Folge)**.**
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 $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ bezeichnet. △

§ 6.5 Das Auswahlaxiom

Das **Auswahlaxiom** (englisch: axiom of choice) der axiomatischen Mengenlehre besagt: Ist $\mathcal{U}$ eine Menge von nichtleeren Mengen, dann gibt es eine Funktion $F\colon \mathcal{U} \to \bigcup \mathcal{U}$, sodass gilt:

$$\forall U \in \mathcal{U} \; \big(F(U) \in U\big).$$

Eine solche Funktion F heißt **Auswahlfunktion** (englisch: choice function) für $\mathcal{U}$, weil sie aus jedem Element U von $\mathcal{U}$ irgendein Element auswählt. Das Auswahlaxiom besagt also, dass es möglich ist, aus jedem Element von $\mathcal{U}$ ein Element auszuwählen, selbst wenn $\mathcal{U}$ überabzählbar viele Mengen als Elemente enthält und man daher nicht in der Lage ist, ein Verfahren anzugeben, nach dem die Auswahl geschehen soll.

Das Auswahlaxiom ist kein fester Bestandteil der axiomatischen Mengenlehre nach Zermelo und Fraenkel, sondern es kann dazugenommen werden oder auch nicht.[55] Es wird aber wohl von den meisten Mathematiker:innen akzeptiert. In Fällen, in denen $\mathcal{U}$ nur endlich viele Mengen enthält, wird das Auswahlaxiom nicht benötigt, weil seine Aussage bereits aus den anderen Axiomen folgt. Wir werden in der Vorlesung darauf hinweisen, wenn ein Resultat von der Hinzunahme des Auswahlaxioms abhängt. Einige Beispiele folgen bereits in diesem Abschnitt, siehe Satz 6.34.

Definition 6.32 (allgemeines kartesisches Produkt)**.**
Es sei I eine Indexmenge und $(A_i)_{i\in I}$ eine Familie von Mengen. Dann ist das **kartesische Produkt** (englisch: Cartesian product) dieser Familie von Mengen gegeben durch

$$\bigtimes_{i\in I} A_i \coloneqq \Big\{F\colon I \to \bigcup_{i\in I} A_i \;\Big|\; F(i) \in A_i \text{ für alle } i \in I\Big\}. \tag{6.10}$$

△

Das kartesische Produkt einer Familie von Mengen besteht also aus *Funktionen* auf der Indexmenge I, deren Funktionswerte jeweils im richtigen Faktor liegen. Im Fall $I = \emptyset$ besteht das kartesische Produkt (6.10) aus dem einzigen Element $F\colon \emptyset \to \emptyset$.

Bemerkung 6.33 (allgemeine kartesische Produkte)**.**
Das **kartesische Produkt** hatten wir bisher nur für endlich viele Mengen definiert, siehe Definition 4.8. Die allgemeine Definition 6.32 erfordert den Funktionenbegriff, der nun zur Verfügung steht. Die Definition 6.32 lässt sich als Verallgemeinerung der Definition 4.8 verstehen: Ist nämlich die Indexmenge $I = [\![1, n]\!]$, so ist $\bigtimes_{i\in I} A_i$ nach (6.10) die Menge aller n-elementigen Folgen. Wenn wir eine solche endliche Folge gemäß der natürlichen Kleiner-Gleich-Ordnung der Indexmenge I als n-Tupel $(a_1, a_2, \ldots, a_n)$ schreiben, so haben wir ein Element aus $\bigtimes_{i\in I} A_i$ gemäß Definition 4.8. Diese Zuordnung ist bijektiv.

[55] Man spricht von den ZF-Axiomen (ohne das Auswahlaxiom) und von den ZFC-Axiomen (mit Auswahlaxiom).

Wenn alle Mengen $A_i = A$ sind, so schreiben wir statt $\bigtimes_{i \in I} A$ auch A^I. Es ist also beispielsweise

$$\begin{aligned} \mathbb{R}^{\mathbb{N}} \quad & \text{die Menge aller Folgen mit Werten in } \mathbb{R}, \\ \{0,1\}^A \quad & \text{die Menge aller } \{0,1\}\text{-wertigen (binären) Funktionen auf einer Menge } A. \end{aligned}$$

Letztere wird manchmal als Schreibweise für die Potenzmenge $\mathcal{P}(A)$ verwendet. (**Quizfrage 6.4:** Inwiefern ist diese Schreibweise gerechtfertigt?) △

Das Auswahlaxiom hat eine ganze Menge äquivalenter, teilweise überraschender Charakterisierungen, von denen der nächste Satz (ohne Beweis) einige angibt.

Satz 6.34 (zum Auswahlaxiom äquivalente Aussagen)**.**
Folgende Aussagen sind in der Mengenlehre von Zermelo-Fraenkel äquivalent:

(i) Es gilt das Auswahlaxiom.

(ii) Ist I eine beliebige Menge und $(A_i)_{i \in I}$ eine Familie nichtleerer Mengen, so ist das kartesische Produkt $\bigtimes_{i \in I} A_i$ eine nichtleere Menge.

(iii) Jede Äquivalenzrelation besitzt ein vollständiges Repräsentantensystem.

(iv) Es sei $f: X \to Y$ eine beliebige Funktion. Dann sind äquivalent:

   (a) f ist surjektiv.

   (b) Es existiert eine Abbildung $h: Y \to X$ mit der Eigenschaft $f \circ h = \mathrm{id}_Y$. Eine solche Abbildung heißt eine **Rechtsinverse** (englisch: right inverse) von f. Sie ist notwendig injektiv.

(v) Es gilt das Lemma von Zorn 6.35.

Lemma 6.35 (**Lemma von Zorn**[56])**.**
Es sei X mit der Relation $\preccurlyeq$ eine halbgeordnete Menge. Weiter besitze jede totalgeordnete Teilmenge $A \subseteq X$ eine obere Schranke in X.[57] Dann existiert in X ein maximales Element.

Wir werden das Auswahlaxiom in Gestalt des Lemmas von Zorn 6.35 später noch verwenden. Wie angekündigt werden wir darauf hinweisen, wenn ein Resultat von der Hinzunahme des Auswahlaxioms oder der Verwendung eines zu ihm äquivalenten Resultats abhängt.

Die Schwierigkeiten in der intuitiven Erfassung des Auswahlaxioms und des äquivalenten Lemmas von Zorn 6.35 (sowie des ebenfalls äquivalenten Wohlordnungssatzes (englisch: well-ordering theorem), den wir hier nicht angeben) 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 6

Ende der Woche 3

[56] englisch: Zorn's lemma

[57] 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, 2022b, Kapitel 3.4, Fischer, Springborn, 2020, Kapitel 2.2

Definition 7.1 (Verknüpfung)**.**
Es sei X eine Menge. Eine **(innere) Verknüpfung** (englisch: (inner) operation) auf X ist eine Abbildung $\star\colon X \times X \to X$. △

Wir schreiben $a \star b$ statt $\star(a, b)$.

Beispiel 7.2 (Verknüpfung)**.**

(i) Ist X endlich, so können wir eine Verknüpfung auf X mit Hilfe einer **Verknüpfungstafel** oder **Vernüpfungstabelle** (englisch: Cayley table) definieren, zum Beispiel

$$+_2\colon \{0,1\} \times \{0,1\} \to \{0,1\} \quad \text{mit der Verknüpfungstafel} \quad \begin{array}{c|cc} +_2 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

$$\cdot_2\colon \{0,1\} \times \{0,1\} \to \{0,1\} \quad \text{mit der Verknüpfungstafel} \quad \begin{array}{c|cc} \cdot_2 & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Quizfrage 7.1: Wo kommen die Definitionen dieser Verknüpfungen her?

Beachte: Die Konvention ist, dass die Zeile das erste Argument (a) und die Spalte das zweite Argument (b) einer Verknüpfung ($a \star b$) angibt.

(ii) Die bekannten Verknüpfungen + und · in $\mathbb{N}$

$$+\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N} \quad \text{mit } (x, y) \mapsto x + y$$
$$\cdot\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N} \quad \text{mit } (x, y) \mapsto x \cdot y$$

sind Verknüpfungen auf $\mathbb{N}$. Analoges gilt für die Mengen $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$.

(*iii*) Es sei X eine Menge und $\mathbb{R}^X = \{f \mid f\colon X \to \mathbb{R}\}$. Dann sind durch die punktweise Addition und die punktweise Multiplikation

$$\begin{aligned} +\colon \mathbb{R}^X \times \mathbb{R}^X &\to \mathbb{R}^X \quad \text{mit } (f,g) \mapsto f+g, \text{definiert durch } (f+g)(x) \coloneqq f(x)+g(x) \\ \cdot\colon \mathbb{R}^X \times \mathbb{R}^X &\to \mathbb{R}^X \quad \text{mit } (f,g) \mapsto f \cdot g, \text{definiert durch } (f \cdot g)(x) \coloneqq f(x) \cdot g(x) \end{aligned}$$

Verknüpfungen auf der Menge der Funktionen $X \to \mathbb{R}$ definiert. (**Quizfrage 7.2:** Was benötigt man als Minimalvoraussetzung, um die Menge Y^X der Funktionen $X \to Y$ mit einer Verknüpfung ausstatten zu können?)

(*iv*) Es sei X eine Menge und $X^X = \{f \mid f\colon X \to X\}$. Dann ist durch die Komposition

$$\circ\colon X^X \times X^X \to X^X \quad \text{mit } (f,g) \mapsto f \circ g, \text{definiert durch } (f \circ g)(x) \coloneqq f(g(x))$$

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

$$(x \star y) \star z = x \star (y \star z) \quad \text{für alle } x, y, z \in H. \tag{7.1}$$

△

Wegen der Assoziativität von $\star$ dürfen wir für die Verknüpfung von drei oder mehr Elementen wie bei $x \star y \star z$ die Klammern weglassen.

Beispiel 7.4 (Halbgruppen).
Beispiele für Halbgruppen sind:

(*i*) $(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$, $(\mathbb{N}_0, +)$, $(\mathbb{N}_0, \cdot)$, $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, +)$, $(\mathbb{C}, \cdot)$

(*ii*) $(\{0,1\}, +_2)$ und $(\{0,1\}, \cdot_2)$ aus Beispiel 7.2

(*iii*) $(\mathbb{R}^X, +)$ und $(\mathbb{R}^X, \cdot)$. Sie erben die Assoziativität von $(\mathbb{R}, +)$ und $(\mathbb{R}, \cdot)$.

(*iv*) $(X^X, \circ)$. Die Assoziativität von $\circ$ wurde in Lemma 6.16 gezeigt.

(*v*) Ist X eine Menge, dann sind $(\mathcal{P}(X), \cap)$, $(\mathcal{P}(X), \cup)$ und $(\mathcal{P}(X), \triangle)$ Halbgruppen.

(*vi*) 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 $\circ$ auf Σ^* durch die Konkatenation von Tupeln:

$$(x_1, \ldots, x_n) \circ (y_1, \ldots, y_m) \coloneqq (x_1, \ldots, x_n, y_1, \ldots, y_m).$$

Dann ist $(\Sigma^*, \circ)$ eine Halbgruppe.[1] △

Beispiel 7.5 (Gegenbeispiele).
Keine Halbgruppen sind:

[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 in der Regel ohne die Klammern notiert, also etwa $ab \circ ba = abba$.

(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 $\mathbb{Z}$, ist aber nicht assoziativ.

(iii) $(\mathbb{N}, \wedge)$ mit $a \wedge b := a^b$. 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:
$$e \star x = x \star e = x \quad \text{für alle } x \in H. \tag{7.2}$$
Falls in $(H, \star)$ ein neutrales Element existiert, dann heißt $(H, \star)$ auch ein **Monoid** (englisch: monoid).
△

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)**.**

(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.

(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) $(\mathcal{P}(X), \cap)$ besitzt das neutrale Element X.

(vii) $(\mathcal{P}(X), \cup)$ besitzt das neutrale Element $\emptyset$.

(viii) $(\mathcal{P}(X), \triangle)$ besitzt das neutrale Element $\emptyset$.

(ix) $(\Sigma^*, \circ)$ 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 (Rechts- und Linkstranslation)**.**
Es sei $(H, \star)$ eine Halbgruppe. Für festes $a \in H$ heißt die Abbildung

$$\star_a\colon H \ni x \mapsto x \star a \in H \quad \text{die **Rechtstranslation** (englisch: right translation) mit } a, \tag{7.3a}$$
$$_a\star\colon H \ni x \mapsto a \star x \in H \quad \text{die **Linkstranslation** (englisch: left translation) mit } a. \quad \triangle \tag{7.3b}$$

Beispiel 7.10 (Rechts- und Linkstranslation)**.**

(i) In $(\mathbb{R}, +)$ ist die Rechtstranslation mit $a = \sqrt{2}$ gegeben durch die Abbildung $x \mapsto x + \sqrt{2}$. Sie ist wegen der Kommutativität von $+$ identisch zur Linkstranslation mit a.

(*ii*) In $(\mathbb{R}^\mathbb{R}, \circ)$ und $g = x \mapsto 2\,x$ ist die Rechtstranslation mit g gegeben durch

$$\circ_g \colon \mathbb{R}^\mathbb{R} \ni f \mapsto f \circ g \in \mathbb{R}^\mathbb{R}, \qquad \text{wobei } f(g(x)) = f(2\,x),$$

während die Linkstranslation mit g gegeben ist durch

$$_g\circ \colon \mathbb{R}^\mathbb{R} \ni f \mapsto g \circ f \in \mathbb{R}^\mathbb{R}, \quad \text{wobei } g(f(x)) = 2\,f(x). \quad \triangle$$

Quizfrage 7.3: Wie lässt sich der Begriff **neutrales Element** in einer Halbgruppe mit Hilfe der Begriffe **Rechtstranslation** und **Linkstranslation** ausdrücken?

Definition 7.11 (invertierbare Elemente)**.**
Es sei $(H, \star)$ eine Halbgruppe mit neutralem Element e. Ein Element $a \in H$ heißt **invertierbar** (englisch: invertible) oder eine **Einheit** (englisch: unit) von $(H, \star)$, wenn ein $b \in H$ existiert mit

$$a \star b = b \star a = e. \tag{7.4}$$

In diesem Fall heißt b ein zu a **inverses Element** (englisch: inverse element) oder ein **Inverses** zu a. △

Beachte: b ist Inverses zu a genau dann, wenn a Inverses zu b ist!

Lemma 7.12 (inverse Elemente sind eindeutig)**.**
Es sei $(H, \star)$ eine Halbgruppe mit neutralem Element e. Ist $a \in H$ invertierbar und sind b_1 und b_2 beides Inverse zu a, dann gilt $b_1 = b_2$.

Beweis. Es gilt

$$\begin{aligned} b_1 &= b_1 \star e \\ &= b_1 \star (a \star b_2) \\ &= (b_1 \star a) \star b_2 \\ &= e \star b_2 \\ &= b_2. \end{aligned}$$ □

Quizfrage 7.4: Welches Element eines Monoids ist immer invertierbar? Was ist sein Inverses?

Bemerkung 7.13 (abkürzende Schreibweisen)**.**

(*i*) Das inverse Element von a wird oft mit a' bezeichnet.

(*ii*) Bezeichnet man die Verknüpfung $\star$ einer Halbgruppe H als „Addition“ und notiert sie als „+“ o. ä., so nennt man ein eventuell existierendes neutrales Element auch **Nullelement** (englisch: additive identity) „0_H“.

Für $n \in \mathbb{N}$ und $a \in H$ ist $n\,a$ eine Abkürzung für $a + \cdots + a$ (n-mal). (**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 die Inverse als $-a$. Dann ist auch $n\,a$ invertierbar für $n \in \mathbb{N}_0$, und wir setzen $(-n)\,a \coloneqq -(n\,a)$. Insbesondere ist $(-1)\,a \coloneqq -a$ und $(-0)\,a = -(0\,a) = -0_H = 0_H$.

Es gilt

$$n\,(m\,a) = (n \cdot m)\,a \quad \text{und} \quad (n+m)\,a = n\,a + m\,a \tag{7.5}$$

für alle $n, m \in \mathbb{N}$ bzw. $n, m \in \mathbb{N}_0$ bzw. $n, m \in \mathbb{Z}$.

Die Bezeichnung $a - b$ steht für $a + (-b)$, vorausgesetzt, b ist invertierbar.

(*iii*) Bezeichnet man die Verknüpfung dagegen als „Multiplikation" und notiert sie als „$\cdot$", so nennt man ein eventuell existierendes neutrales Element auch **Einselement** (englisch: multiplicative identity) „1_H".

Für $n \in \mathbb{N}$ und $a \in H$ ist a^n eine Abkürzung für $a \cdot \dots \cdot a$ (n-mal).

Besitzt H das neutrale Element 1_H, so definieren wir auch $a^0 := 1_H$.

Ist weiter $a \in H$ invertierbar, so notieren wir die Inverse als a^{-1}. Dann ist auch a^n invertierbar für $n \in \mathbb{N}_0$, und wir setzen $a^{-n} = (a^n)^{-1}$. Insbesondere ist $a^{-0} = (a^0)^{-1} = 1_H^{-1} = 1_H$.

Es gilt

$$(a^n)^m = a^{n \cdot m} \quad \text{und} \quad a^{n+m} = a^n \cdot a^m \tag{7.6}$$

für alle $n, m \in \mathbb{N}$ bzw. $n, m \in \mathbb{N}_0$ bzw. $n, m \in \mathbb{Z}$.

(*iv*) Bezeichnet man die Verknüpfung dagegen als „Komposition" und notiert sie als „$\circ$", so nennt man ein eventuell existierendes neutrales Element auch **Identität** (englisch: identity) „id". In diesem Fall verwenden wir ebenfalls die multiplikative Notation, z. B. ist a^n eine Abkürzung für $a \circ \dots \circ a$ (n-mal). △

Beispiel 7.14 (invertierbare Elemente).

(*i*) In $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind alle Elemente invertierbar. Das Inverse von a wird mit $-a$ bezeichnet.

(*ii*) 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 $1/a$ bezeichnet.

Die Bezeichnung $\frac{a}{b}$ steht für ab^{-1}, vorausgesetzt, b ist invertierbar.

(*iii*) In $(\mathbb{N}_0, +)$ ist nur das Element 0 invertierbar. Die Inverse von 0 ist wiederum 0.

(*iv*) In $(\mathbb{Z}, \cdot)$ sind nur 1 und -1 invertierbar. Beide sind zu sich selbst invers.

(*v*) In $(\{0, 1\}, +_2)$ aus Beispiel 7.2 sind beide Elemente invertierbar. Beide sind zu sich selbst invers.

(*vi*) In $(\{0, 1\}, \cdot_2)$ aus Beispiel 7.2 ist nur das Element 1 invertierbar. Es ist zu sich selbst invers.

(*vii*) In $(X^X, \circ)$ sind genau die bijektiven Funktionen $X \to X$ invertierbar. △

Quizfrage 7.6: Welches sind die invertierbaren Elemente in den Monoiden $(\mathcal{P}(X), \cap)$, $(\mathcal{P}(X), \cup)$ und $(\mathcal{P}(X), \triangle)$?

Ende der Vorlesung 7

§ 7.2 Gruppen

Definition 7.15 (Gruppe)**.**
Es sei $(H, \star)$ ein Monoid. $(H, \star)$ heißt **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 7.16 (Gruppen und Gegenbeispiele)**.**

(i) Es sei $(H, \star)$ ein Monoid. Dann ist die Teilmenge der invertierbaren Elemente

$$E(H, \star) \coloneqq \{a \in H \mid a \text{ ist invertierbar}\} \tag{7.7}$$

eine Gruppe, genannt die **Einheitengruppe** (englisch: unit group, group of units) $E(H, \star)$ von $(H, \star)$.

(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 $1/a \in \mathbb{Q}_{\neq 0}$. Dasselbe gilt für $(\mathbb{R}_{\neq 0}, \cdot)$ und $(\mathbb{C}_{\neq 0}, \cdot)$.

(v) Für $m \in \mathbb{N}$ bildet die Menge $\mathbb{Z}_m \coloneqq \{0, 1, \dots, m-1\}$ mit der Verknüpfung $+_m$ eine abelsche Gruppe (siehe Definition 7.19). Dabei ist $+_m$ die **Addition modulo m** (englisch: addition modulo m) definiert als[2]

$$\begin{aligned} a +_m b &\coloneqq \begin{cases} a + b, & \text{falls } a + b \leqslant m - 1 \\ a + b - m, & \text{falls } a + b \geqslant m \end{cases} \\ &= \text{der natürliche Repräsentant von } a + b \text{ in der Restklasse } [a + b] \text{ modulo } m \\ &= \text{Rest von } a + b \text{ bei ganzzahliger Division durch } m. \end{aligned} \tag{7.8}$$

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.

(vi) Für $m \in \mathbb{N}$, $m \geqslant 2$, bildet die Menge $\mathbb{Z}_m$ mit der Verknüpfung $\cdot_m$ ein kommutatives Monoid. Dabei ist $\cdot_m$ die **Multiplikation modulo m** (englisch: multiplication modulo m) definiert als[3]

$$\begin{aligned} a \cdot_m b &\coloneqq \text{der natürliche Repräsentant von } a \cdot b \text{ in der Restklasse } [a \cdot b] \text{ modulo } m \\ &= \text{Rest von } a \cdot b \text{ bei ganzzahliger Division durch } m. \end{aligned} \tag{7.9}$$

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. In diesem Fall ist $(\mathbb{Z}_1, \cdot_1)$ isomorph (Definition 8.1) zu $(\mathbb{Z}_1, +_1)$.

Den Fall $m = 2$ kennen wir bereits als $(\{0, 1\}, \cdot_2)$ aus Beispiel 7.2.

[2] Beispielsweise ist $3 +_6 5 = 2$, weil $3 + 5 = 8$ ist und $8 \overset{6}{\equiv} 2$ gilt.

[3] Beispielsweise ist $3 \cdot_6 5 = 3$, weil $3 \cdot 5 = 15$ ist und $15 \overset{6}{\equiv} 3$ gilt.

(*vii*) $(\mathbb{R}^X, +)$ ist eine Gruppe.

(*viii*) $(\mathbb{R}^X, \cdot)$ 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.

(*ix*) $(X^X, \circ)$ 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. △

Quizfrage 7.7: Können Sie die Additions- und Multiplikationstabellen für $\mathbb{Z}_m$ im Fall $m = 5$ und $m = 8$ aufstellen? Haben Sie eine Vermutung, welche Elemente in $(\mathbb{Z}_m, \cdot_m)$ invertierbar sind?

Satz 7.17 (Rechenregeln für Inverse)**.**
Es sei $(G, \star)$ eine Gruppe mit neutralem Element e.

(*i*) Es gelten die **Kürzungsregeln** (englisch: cancellation rules)

$$a \star b_1 = a \star b_2 \quad \Rightarrow \quad b_1 = b_2 \tag{7.10a}$$
$$b_1 \star a = b_2 \star a \quad \Rightarrow \quad b_1 = b_2 \tag{7.10b}$$

für $a, b_1, b_2 \in G$.

(*ii*) In einer Gruppe reicht es für den Nachweis, dass $a \in G$ und $b \in G$ Inverse voneinander sind, aus, diese in einer der beiden Reihenfolgen miteinander zu verknüpfen:

$$a \star b = e \quad \Rightarrow \quad b = a', \tag{7.11a}$$
$$a \star b = e \quad \Rightarrow \quad a = b'. \tag{7.11b}$$

(*iii*) Die Invertierung ist **involutorisch** (englisch: involutory), d. h., für alle $a \in G$ gilt

$$(a')' = a. \tag{7.12}$$

(*iv*) Für das inverse Element zu $a \star b$ für $a, b \in G$ gilt

$$(a \star b)' = b' \star a'. \tag{7.13}$$

Beweis. Aussage (*i*):

$$\begin{aligned}
& a \star b_1 = a \star b_2 && \\
\Rightarrow \quad & a' \star (a \star b_1) = a' \star (a \star b_2) && a' \text{ existiert in der Gruppe } (G, \star) \\
\Rightarrow \quad & (a' \star a) \star b_1 = (a' \star a) \star b_2 && \text{wegen der Assoziativität von } \star \\
\Rightarrow \quad & e \star b_1 = e \star b_2 && \text{da } a' \text{ invers zu } a \text{ ist} \\
\Rightarrow \quad & b_1 = b_2 && \text{wegen der Eigenschaften von } e.
\end{aligned}$$

Die Aussage (7.10b) folgt analog.

Aussage (*ii*): Es gilt $a \star b = e$ und ebenso $a \star a' = e$. Nach (7.10a) muss also $b = a'$ gelten. Weiter gilt $a \star b = e$ und ebenso $b' \star b = e$. Nach (7.10b) muss also $a = b'$ gelten.

Aussage (*iii*): Wir müssen nachweisen, dass a die Inverse zu a' ist. Wegen $a \star a' = a' \star a = e$ ist das aber der Fall.

Aussage (*iv*): Wir müssen nachweisen, dass $b' \star a'$ die Inverse zu $a \star b$ ist. Wir haben

$$\begin{aligned}(a \star b) \star (b' \star a') &= a \star (b \star b') \star a' && \text{wegen der Assoziativität von } \star \\ &= a \star e \star a' && \text{da } b' \text{ invers zu } b \text{ ist} \\ &= a \star a' && \text{wegen der Eigenschaften von } e \\ &= e && \text{da } a' \text{ invers zu } a \text{ ist.}\end{aligned}$$

□

Das folgende Lemma gibt mit Hilfe von Rechts- und Linkstranslationen eine notwendige und eine hinreichende Bedingung dafür an, wann eine Halbgruppe sogar eine Gruppe ist.

Lemma 7.18 (Gruppenkriterium mit Rechts- und Linkstranslationen („**Sudoku-Kriterium**")).

(*i*) Ist $(G, \star)$ eine Gruppe und ist $a \in G$ beliebig, so sind die Rechts- und Linkstranslation $\star_a$ und ${}_a\star$ bijektive Abbildungen $G \to G$.

(*ii*) Ist $(H, \star)$ eine nichtleere Halbgruppe und gilt für alle $a \in H$, dass die Rechts- und Linkstranslationen $\star_a$ und ${}_a\star$ surjektive Abbildungen sind, dann ist $(H, \star)$ eine Gruppe.

Beweis. Dieser Beweis ist Teil von Hausaufgabe I-4.3. □

Definition 7.19 (kommutative Halbgruppe, kommutatives Monoid, kommutative Gruppe).
Eine Halbgruppe bzw. ein Monoid bzw. eine Gruppe $(H, \star)$ heißt **kommutativ** (englisch: commutative) oder **abelsch** (englisch: Abelian), wenn

$$x \star y = y \star x \quad \text{für alle } x, y \in H \tag{7.14}$$

gilt. △

Beispiel 7.20 (kommutative Halbgruppen und Gruppen).
$(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$, $(\mathbb{N}_0, +)$, $(\mathbb{N}_0, \cdot)$, $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, +)$, $(\mathbb{C}, \cdot)$ sind alle kommutativ. Beispielsweise ist $(\mathbb{N}, +)$ eine kommutative Halbgruppe (aber kein Monoid), $(\mathbb{N}_0, +)$ ein kommutatives Monoid (aber keine Gruppe) und $(\mathbb{Z}, +)$ eine kommutative Gruppe. △

Weitere Beispiele folgen in der Übung.

§ 7.3 Die symmetrische Gruppe

Definition 7.21 (symmetrische Gruppe).
Es sei X eine nichtleere Menge und $S(X) \coloneqq \{f \colon X \to X \mid f \text{ ist bijektiv}\}$. Dann heißt $(S(X), \circ)$ die **symmetrische Gruppe** (englisch: symmetric group) auf X. Jedes Element von $S(X)$ heißt eine **Permutation** (englisch: permutation) von X.

Ist $X = [\![1, n]\!]$ für $n \in \mathbb{N}$, so schreiben wir 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 7.14 *(vii)* ist $S(X)$ tatsächlich eine Gruppe. Das neutrale Element ist id_X. Wenn X drei oder mehr Elemente enthält, dann ist $(S(X), \circ)$ nicht kommutativ, ansonsten kommutativ.

Eine Permutation $\sigma \in S_n$ können wir 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 $n \in \mathbb{N}$ ist gleich $n!$ („n Fakultät").

Beispiel 7.22 (symmetrische Gruppe vom Grad 3)**.**
Die symmetrische Gruppe S_3 hat $3! = 6$ Elemente:

$$\begin{aligned} \sigma_0 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \sigma_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, & \sigma_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} & &\text{(Drehungen)}, \\ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, & \sigma_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, & \sigma_5 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} & &\text{(Spiegelungen)}. \end{aligned}$$

Sie lassen sich identifizieren mit den Kongruenzabbildungen, die ein gleichseitiges Dreieck mit den Eckpunkten 1, 2 und 3 auf sich selbst überführen. 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 erwartet tatsächlich nicht kommutativ. △

Definition 7.23 (Transposition)**.**
Eine Permutation $\sigma \in S_n$, $n \in \mathbb{N}$, heißt eine **Transposition** (englisch: transposition), 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 $\tau(i, j)$. △

Eine Transposition vertauscht also genau zwei Elemente von $[\![1, n]\!]$ und lässt den Rest unverändert. Offenbar gilt für jede Transposition

$$\tau^2 = \tau \circ \tau = \mathrm{id}, \quad \text{also } \tau^{-1} = \tau. \tag{7.16}$$

Transpositionen sind also selbstinvers.

In S_1 gibt es keine Transpositionen. (**Quizfrage 7.8:** Wieviele verschiedene Transpositionen gibt es in S_n?)

Satz 7.24 (Darstellung von Permutationen als Komposition von Transpositionen)**.**
Es sei $n \in \mathbb{N}$. Jede Permutation $\sigma \in S_n$ lässt sich als Komposition von $0 \leqslant r \leqslant n - 1$ Transpositionen schreiben.

Beweis. Wir zeigen die Behauptung für $n \geqslant 1$ durch vollständige Induktion. Induktionsanfang: Das einzige Element von

$$S_1 = \{\mathrm{id}_{\{1\}}\}$$

ist 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 $0 \leqslant r \leqslant n-1$. Schließlich zeigt $\sigma = \tau(m,n+1) \circ \widetilde{\sigma} = \tau(m,n+1) \circ \tau_1 \circ \cdots \circ \tau_r$ die Behauptung.

□

Die Darstellung einer Permutation als Komposition von Transpositionen ist nicht eindeutig. Jedoch ist Anzahl der benötigten Transpositionen entweder immer gerade oder immer ungerade, wie wir gleich beweisen werden (Folgerung 7.30).

Definition 7.25 (Fehlstand, Signum einer Permutation)**.**
Es sei $n \in \mathbb{N}$ 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[4]

$$\operatorname{sgn} \sigma \coloneqq \prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} \tag{7.17}$$

heißt das **Signum** (englisch: sign, lateinisch: signum: Zeichen) von σ. △

Beispiel 7.26 (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

$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

[4]Definitionsgemäß wird das im Fall $n = 1$ leere Produkt als 1 interpretiert.

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 7.27 (zu Definition 7.25)**.**
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. △

Lemma 7.28 (Signum einer Transposition)**.**
Ist $\tau \in S_n$, $n \in \mathbb{N}$, 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 $\operatorname{sgn}\tau(i,j) = (-1)^{2\,(j-i)-1} = -1.$ □

Satz 7.29 (Signum ist verträglich mit Komposition von Permutationen)**.**
Es sei $n \in \mathbb{N}$ 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.

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 ist die Anzahl 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 haben o. B. d. A. $\ell > k$, daher können wir $\tau(k,\ell)$ in der Form

$$\underbrace{\tau(k,\ell) = \tau(k,k+1)\circ\cdots\circ\tau(\ell-2,\ell-1)}\circ\underbrace{\tau(\ell,\ell-1)\circ\cdots\circ\tau(k+1,k)},$$

also als Komposition von $(2\,(\ell-k)-1)$ Transpositionen benachbarter Elemente schreiben. Aufgrund von **Schritt** 1 und der Assoziativität der Komposition haben wir nun also

$$\begin{aligned}
&\operatorname{sgn}\big(\tau(k,\ell)\circ\sigma_2\big)\\
&\quad=\operatorname{sgn}\big(\tau(k,k+1)\circ\cdots\circ\tau(\ell-2,\ell-1)\circ\tau(\ell,\ell-1)\circ\cdots\circ\tau(k+1,k)\circ\sigma_2\big)\\
&\quad=\big(\operatorname{sgn}\tau(k,k+1)\big)\cdot\operatorname{sgn}\big(\cdots\circ\tau(\ell-2,\ell-1)\circ\tau(\ell,\ell-1)\circ\cdots\circ\tau(k+1,k)\circ\sigma_2\big)\\
&\quad=\cdots\\
&\quad=\big(\operatorname{sgn}\tau(k,k+1)\big)\cdots\big(\operatorname{sgn}\tau(\ell-2,\ell-1)\big)\big(\operatorname{sgn}\tau(\ell,\ell-1)\big)\cdots\big(\operatorname{sgn}\tau(k+1,k)\big)(\operatorname{sgn}\sigma_2)\\
&\quad=\cdots\\
&\quad=\big(\operatorname{sgn}\tau(k,\ell)\big)\cdot(\operatorname{sgn}\sigma_2).
\end{aligned}$$

Schritt 3: Schließlich können wir den allgemeinen Fall zeigen.

Ist $\sigma_1\in S_n$ eine beliebige Permutation, so können wir sie nach Satz 7.24 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:

$$\begin{aligned}
&\operatorname{sgn}(\sigma_1\circ\sigma_2)\\
&\quad=\operatorname{sgn}\big(\tau_1\circ\cdots\circ\tau_r\circ\sigma_2\big)\\
&\quad=(\operatorname{sgn}\tau_1)\cdot\operatorname{sgn}\big(\cdots\circ\tau_r\circ\sigma_2\big)\\
&\quad=\cdots\\
&\quad=(\operatorname{sgn}\tau_1)\cdots(\operatorname{sgn}\tau_r)(\operatorname{sgn}\sigma_2)\\
&\quad=\cdots\\
&\quad=\operatorname{sgn}\big(\tau_1\circ\cdots\circ\tau_r\big)\cdot\operatorname{sgn}(\sigma_2).
\end{aligned}$$

□

Folgerung 7.30 (zu Satz 7.29)**.**
Es sei $n\in\mathbb{N}$ und σ eine Permutation in S_n.

(i) Ist $\sigma=\sigma_1\circ\cdots\circ\sigma_s$ dargestellt als Komposition[5] von $s\in\mathbb{N}_0$ Permutationen $\sigma_i\in S_n$, so gilt $\operatorname{sgn}\sigma=(\operatorname{sgn}\sigma_1)\cdots(\operatorname{sgn}\sigma_s)$.

(ii) Ist insbesondere $\sigma=\tau_1\circ\cdots\circ\tau_r$ dargestellt als Komposition von $r\in\mathbb{N}$ Transpositionen in S_n, so gilt $\operatorname{sgn}\sigma=(-1)^r$.

(iii) Es gilt $\operatorname{sgn}\mathrm{id}=1$.

(iv) Es gilt $\operatorname{sgn}\sigma=\operatorname{sgn}\sigma^{-1}$.

[5] Vereinbarungsgemäß ist die Verknüpfung von null Permutationen das neutrale Element in S_n, also die identische Abbildung id.

Beweis. Nach Satz 7.29 ist das Signum einer Komposition von zwei Permutationen gleich dem Produkt der Signa der beiden Faktoren. Wie im Beweis von Satz 7.29 können wir die Aussage leicht auf mehr als zwei Faktoren ausdehnen. Die Fälle $s = 0$ und $s = 1$ sind trivial. Das zeigt Aussage (i).

Das Signum einer Transposition ist nach Lemma 7.28 gleich -1. Aussage (ii) folgt damit aus Aussage (i).

Die identische Abbildung ist Produkt von null Transpositionen, also gilt $\operatorname{sgn} \mathrm{id} = (-1)^0 = 1$, also Aussage (iii).

Schließlich gilt

$$1 = \operatorname{sgn} \mathrm{id} = \operatorname{sgn}(\sigma \circ \sigma^{-1}) = (\operatorname{sgn} \sigma) \cdot (\operatorname{sgn} \sigma^{-1}),$$

also $\operatorname{sgn} \sigma^{-1} = 1/\operatorname{sgn} \sigma = \operatorname{sgn} \sigma$, da $\operatorname{sgn} \sigma \in \{\pm 1\}$ ist. Das zeigt Aussage (iv). □

Ende der Vorlesung 8

Ende der Woche 4

§ 7.4 Untergruppen

Definition 7.31 (Untergruppe).
Es sei $(G, \star)$ eine Gruppe.

(i) Eine Teilmenge $U \subseteq G$ heißt **abgeschlossen** (englisch: closed) bzgl. $\star$, wenn $\star \colon G \times G \to G$ eingeschränkt werden kann zu $\star_U \colon U \times U \to U$. In diesem Fall heißt $\star_U$ die auf U **induzierte (innere) Verknüpfung** (englisch: induced operation, lateinisch: inducere: hineinführen).

(ii) Eine bzgl. $\star$ abgeschlossene Teilmenge $U \subseteq G$ heißt eine **Untergruppe** (englisch: subgroup) von $(G, \star)$, wenn $(U, \star_U)$ selbst wieder eine Gruppe ist. Manchmal schreibt man dies als $(U, \star_U) \leqslant (G, \star)$.

(iii) Eine Untergruppe $(U, \star_U)$ von $(G, \star)$ heißt **echt** (englisch: proper subgroup), wenn $U \subsetneq G$ gilt. △

Beachte: Die Assoziativität wird von $\star$ auf $\star_U$ vererbt. Ist $(G, \star)$ abelsch, dann auch $(U, \star)$.

Lemma 7.32 (neutrale und inverse Elemente in einer Untergruppe).
Es sei $(U, \star_U)$ eine Untergruppe der Gruppe $(G, \star)$. Dann ist das neutrale Element e_U von $(U, \star_U)$ 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 von Hausaufgabe I-5.1. □

Aufgrund dieser Erkenntnis benötigen wir also keine neue Notation für das neutrale Element und die Inversen in einer Untergruppe. Außerdem schreiben wir ab jetzt einfach $\star$ statt $\star_U$.

Die Prüfung einer Teilmenge $U \subseteq G$ auf die Untergruppen-Eigenschaft lässt sich mit folgendem Kriterium abkürzen:

Satz 7.33 (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 $a \star b' \in U$.

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 7.32 auch das neutrale Element in $(U, \star)$ ist. Für $a, b \in U$ gilt $b' \in U$ nach Lemma 7.32. 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 $U \subseteq$ eine Untergruppe von $(G, \star)$. □

Quizfrage 7.9: Könnte man an Stelle von Aussage (*ii*) in Satz 7.33 äquivalent auch $U \neq \emptyset$, und für alle $a, b \in U$ gilt $a' \star b \in U$ fordern?

Beispiel 7.34 (Untergruppen).

(*i*) Es sei $(G, \star)$ eine Gruppe mit neutralem Element e. Dann sind $(\{e\}, \star)$ und $(G, \star)$ Untergruppen von $(G, \star)$. Diese heißen die **trivialen Untergruppe** (englisch: trivial subgroups).

(*ii*) $(\mathbb{R}_{>0}, \cdot)$ ist eine Untergruppe von $(\mathbb{R}_{\neq 0}, \cdot)$.

(*iii*) Für jede Zahl $m \in \mathbb{N}$ ist $m\mathbb{Z} \coloneqq \{m\,z \mid z \in \mathbb{Z}\}$ mit der Verknüpfung $+$ eine Untergruppe der Gruppe $(\mathbb{Z}, +)$.

(*iv*) Für $n \geqslant 2$ 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. Sie hat $\frac{1}{2}n!$ Elemente.

(*v*) In S_3 besteht die alternierende Untergruppe A_3 in der Notation von Beispiel 7.22 gerade aus $\{\sigma_0, \sigma_1, \sigma_2\}$. Diese entsprechen bei Interpretation als Kongruenzabbildungen eines gleichseitigen Dreiecks auf sich selbst gerade den Drehungen. △

Quizfrage 7.10: Können Sie eine Gruppe finden, die außer den trivialen Untergruppen keine weiteren Untergruppen besitzt?

Beachte: Die Menge der Untergruppen einer Gruppe $(G, \star)$ sind bzgl. der Eigenschaft „ist Untergruppe von" partiell geordnet.

Lemma 7.35 (Durchschnitt von Untergruppen)**.**
Es sei $(G, \star)$ eine Gruppe und $(U_i, \star)_{i \in I}$ eine Familie von Untergruppen mit der nichtleeren Indexmenge I. Dann ist auch $\bigcap_{i \in I} U_i$ mit $\star$ eine Untergruppe von $(G, \star)$.

Beweis. Dieser Beweis ist Gegenstand von Hausaufgabe I-5.1. □

Definition 7.36 (erzeugte Untergruppe, Erzeugendensystem, zyklische Gruppe, Ordnung eines Elements)**.**
Es sei $(G, \star)$ eine Gruppe und $E \subseteq G$.

(i) Dann heißt

$$\langle E \rangle := \bigcap \big\{ U \,\big|\, (U, \star) \text{ ist Untergruppe von } (G, \star) \text{ und } E \subseteq U \big\} \tag{7.21}$$

die von E **erzeugte Untergruppe** (englisch: subgroup generated by E) in $(G, \star)$.

Beachte: Bezeichnen wir mit $\mathcal{R}$ die Menge auf rechten Seite von (7.21), über die der Durchschnitt gebildet wird, dann ist $\langle E \rangle$ das Minimum der Menge $\mathcal{R}$ bzgl. der Inklusionshalbordnung und sogar das Minimum der Menge $\mathcal{R}$ bzgl. der Halbordnung „ist Untergruppe von".

Ist speziell $E = \{a\}$ für ein $a \in G$, so schreiben wir auch $\langle a \rangle$ statt $\langle \{a\} \rangle$ und nennen $\langle a \rangle$ die **von a erzeugte zyklische Untergruppe** (englisch: cyclic subgroup) von $(G, \star)$.

(ii) 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).

(iii) Die Gruppe $(G, \star)$ heißt **zyklisch** (englisch: cyclic), wenn es ein $a \in G$ gibt, sodass gilt: $\langle a \rangle = G$. In diesem Fall heißt a ein **Erzeuger** (englisch: generator) von G.

(iv) Ein Element $a \in G$ heißt von **Ordnung $n \in \mathbb{N}$** (englisch: order), wenn $n \in \mathbb{N}$ die kleinste Zahl ist, für die (in multiplikativer Notation) $a^n = 1$ gilt. Falls kein $n \in \mathbb{N}$ existiert, sodass $a^n = 1$ ist, so heißt a von **unendlicher Ordnung** (englisch: infinite order). Wir schreiben $\operatorname{ord}(a) = n$ bzw. $\operatorname{ord}(a) = \infty$. △

Satz 7.37 (Darstellung der erzeugten Untergruppe)**.**
Es sei $(G, \star)$ eine Gruppe und $E \subseteq G$. Dann gilt für die von E erzeugte Untergruppe:

$$\langle E \rangle = \big\{ a_1 \star \cdots \star a_n \,\big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; (a_i \in E \cup E') \big\}, \tag{7.22}$$

wobei E' die Menge der Inversen von E bezeichnet. (Im Fall $n = 0$ interpretieren wir wie üblich die Verknüpfung von null Elementen in der rechten Menge als das neutrale Element e. Insbesondere im Fall $E = \emptyset$ ist 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 7.33). 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 7.11:** 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 7.38 (erzeugte Untergruppe, Erzeugendensystem, zyklische Gruppe, Ordnung eines Elements)**.**

(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 7.22, also

$$\sigma_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \qquad \text{(Drehungen)}$$
$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \qquad \text{(Spiegelungen)}$$

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 7.34. 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. △

Bemerkung 7.39 (abkürzende Schreibweisen)**.**
Es sei $(H, \star)$ eine Halbgruppe, $a \in H$ sowie $A, B \subseteq H$. Zur Abkürzung vereinbaren wir folgende Schreibweisen:

$$a \star B := \{a \star b \,|\, b \in B\}, \tag{7.23a}$$
$$B \star a := \{b \star a \,|\, b \in B\}, \tag{7.23b}$$
$$A \star B := \{a \star b \,|\, a \in A,\ b \in B\}, \tag{7.23c}$$
$$A' := \{a' \,|\, a \in A \text{ ist invertierbar}\}. \tag{7.23d}$$

Wenn A bzw. B die leere Menge ist, ist das Ergebnis in allen obigen Fällen die leere Menge. Mit Hilfe dieser Abkürzungen können wir z. B. das Untergruppenkriterium Satz 7.33 (ii) als $U \neq \emptyset$ und $U \star U' \subseteq U$ formulieren. △

Ende der Vorlesung 9

§ 7.5 Untergruppen induzieren Äquivalenzrelationen

Lemma 7.40 (von Untergruppe induzierte Äquivalenzrelationen)**.**
Es sei $(G, \star)$ eine Gruppe und $(U, \star)$ eine Untergruppe. Dann sind durch

$$a \sim^U b \quad \Leftrightarrow \quad b \in a \star U \quad \Leftrightarrow \quad a' \star b \in U \tag{7.24a}$$

$$a \, {}^U{\sim} \, b \quad \Leftrightarrow \quad a \in U \star b \quad \Leftrightarrow \quad a \star b' \in U \tag{7.24b}$$

für $a, b \in G$ zwei Äquivalenzrelationen auf G erklärt.[6] Für die Äquivalenzklassen gilt:

$$[a]_{\sim^U} = a \star U \tag{7.25a}$$

$$[a]_{{}^U\sim} = U \star a. \tag{7.25b}$$

Jede der Äquivalenzklassen $[a]_{\sim^U}$ und $[a]_{{}^U\sim}$ ist gleichmächtig zu U.

Beweis. Wir zeigen zunächst, dass die beiden angegebenen Bedingungen in (7.24) tatsächlich äquivalent sind. Wir haben

$$\begin{aligned}
& b \in a \star U \\
\Leftrightarrow \quad & \exists c \in U \; (b = a \star c) \\
\Leftrightarrow \quad & \exists c \in U \; (a' \star b = a' \star a \star c) \quad \text{„}{\Leftarrow}\text{“ folgt aus der Kürzungsregel (7.10a)} \\
\Leftrightarrow \quad & \exists c \in U \; (a' \star b = c) \\
\Leftrightarrow \quad & a' \star b \in U.
\end{aligned}$$

Wir weisen nun für $a \sim^U b$ die Eigenschaften einer Äquivalenzrelation nach. Das neutrale Element von U und G wird wieder mit e bezeichnet.

Schritt 1: $\sim^U$ ist reflexiv:

Es sei $a \in G$, dann ist $a' \star a = e \in U$, da U Untergruppe ist.

Schritt 2: $\sim^U$ ist symmetrisch:

Es gelte $a \sim^U b$, also $a' \star b \in U$. Dann ist auch das Inverse $(a' \star b)' \in U$, da U Untergruppe ist. Für das Inverse gilt nach Satz 7.17 *(iii)* und *(iv)*:

$$(a' \star b)' = b' \star (a')' = b' \star a \in U.$$

Das heißt aber $b \sim^U a$.

Schritt 3: $\sim^U$ ist transitiv:

Es gelte $a \sim^U b$ und $b \sim^U c$, also $a' \star b \in U$ und $b' \star c \in U$. Aufgrund der Untergruppeneigenschaft von U ist auch $a' \star b \star b' \star c = a' \star c \in U$. Das heißt aber $a \sim^U c$.

Die Darstellung der Äquivalenzklasse (7.25a) folgt sofort aus (7.24a).

Um zu zeigen, dass U und $[a]_{\sim^U} = a \star U$ gleichmächtig sind (Definition 6.23), betrachten wir die Abbildung $U \ni b \mapsto a \star b \in a \star U$. Diese Abbildung ist nach Definition von $a \star U$ surjektiv. Außerdem ist sie injektiv, denn aus $a \star b = a \star c$ folgt mit der Kürzungsregel (7.10a) $b = c$.

Der Beweis für (7.24b) und (7.25b) geht analog. □

[6] Für diese Relationen gibt es in der Literatur keine einheitliche Notation.

Die Äquivalenzklasse $[a]_{\sim^U} = a \star U$ heißt auch die **Linksnebenklasse** (englisch: left coset) von U nach a.[7] Weil $\sim^U$ eine Äquivalenzrelation ist, bilden die Linksnebenklassen der Untergruppe U eine Partition der Gruppe G (Satz 5.19). Man notiert die Quotientenmenge als $G \,/ \sim^U$ oder auch als $G \,/\, U$.

Die Äquivalenzklasse $[a]_{{}^U\!\sim} = U \star a$ heißt auch die **Rechtsnebenklasse** (englisch: right coset) von U nach a. Weil auch ${}^U\!\sim$ eine Äquivalenzrelation ist, bilden auch die Rechtsnebenklassen der Untergruppe U eine Partition der Gruppe G. Man notiert die Quotientenmenge als $G \,/\, {}^U\!\sim$ oder auch als $U \backslash G$.

Folgerung 7.41 (zu Lemma 7.40)**.**

Es sei $(G, \star)$ eine abelsche Gruppe und $(U, \star)$ eine Untergruppe. Dann sind die Äquivalenzrelationen $a \sim^U b$ und $a \,{}^U\!\sim b$ identisch. Entsprechend gilt für die Nebenklassen $a \star U = U \star a$ für alle $a \in G$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-5.3. □

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) $[a]_{\overset{U}{\sim}} = U \star a = a \star U$ von U.

Beispiel 7.42 (Nebenklassen)**.**

(i) In der abelschen Gruppe $(\mathbb{Z}, +)$ erzeugt die Untergruppe $m\,\mathbb{Z}$ für $m \in \mathbb{N}$ gerade die Kongruenzrelation modulo m, d. h., $\overset{m\mathbb{Z}}{\sim}$ und $\overset{m}{\equiv}$ stimmen überein. Die Nebenklassen von $m\,\mathbb{Z}$ gilt (auch **Restklassen modulo** m genannt, vgl. Beispiel 5.16)

$$[a] = \{a + n\,m \mid n \in \mathbb{Z}\} = a + m\,\mathbb{Z}$$

partitionieren die ganzen Zahlen $\mathbb{Z}$ in m gleichmächtige Restklassen, $[0], [1], \ldots, [m-1]$.

(ii) 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. △

Aus Lemma 7.40 folgt der folgende wichtige **Satz von Lagrange** (englisch: Lagrange's theorem) der Gruppentheorie:

Satz 7.43 (Satz von Lagrange)**.**

Es sei $(G, \star)$ eine endliche Gruppe und $(U, \star)$ 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 von Hausaufgabe I-5.3. □

[7] Merke: Bei den *Links*nebenklassen $a \star U$ steht der Repräsentant a *links* vom U. Im Relationszeichen $\sim^U$ steht die Tilde $\sim$ ebenfalls links vom U.

§ 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 (Halbgruppenhomomorphismus)**.**
Es seien $(H_1, \star)$ und $(H_2, \square)$ zwei Halbgruppen.

(i) Eine Abbildung $f \colon H_1 \to H_2$ heißt **strukturverträglich** oder ein **(Halbgruppen-)Homomorphismus** (englisch: semigroup homomorphism) von $(H_1, \star)$ in $(H_2, \square)$, wenn gilt:

$$f(a \star b) = f(a) \mathbin{\square} f(b) \quad \text{für alle } a, b \in H_1. \tag{8.1}$$

(ii) Im Fall $(H_1, \star) = (H_2, \square)$ sprechen wir auch von einem **(Halbgruppen-)Endomorphismus** (englisch: semigroup endomorphism, altgriechisch: *ένδον*: innen).

(iii) Ist $f \colon H_1 \to H_2$ bijektiv, so heißt f auch **strukturerhaltend** oder ein **(Halbgruppen-)Isomorphismus** (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) Im Fall $(H_1, \star) = (H_2, \square)$ und $f \colon H_1 \to H_2$ bijektiv sprechen wir auch von einem **(Halbgruppen-) Automorphismus** (englisch: semigroup automorphism, altgriechisch: *αυτος*: selbst).[8] △

Quizfrage 8.1: Welche Art von Relation ist die Isomorphie auf der Klasse aller Halbgruppen?

Bemerkung 8.2 (Halbgruppenhomomorphismus als kommutatives Diagramm)**.**
Wir können den Sachverhalt, dass $f \colon (H_1, \star) \to (H_2, \square)$ ein Halbgruppenhomomorphismus ist, durch das folgende **kommutative Diagramm** (englisch: commutative diagram) ausdrücken:[9]

$$\begin{tikzcd}
H_1 \times H_1 \arrow[r, "f \times f"] \arrow[d, "\star"] & H_2 \times H_2 \arrow[d, "\square"] \\
H_1 \arrow[r, "f"] & H_2
\end{tikzcd}$$

Ein solches Diagramm heißt **kommutativ** (englisch: commutative diagram), wenn alle Pfade mit demselben Ausgangs- und demselben Endpunkt dasselbe Ergebnis produzieren. △

[8] Ein Automorphismus ist somit ein bijektiver Endomorphismus oder auch ein Isomorphismus von einer Halbgruppe/Monoid/Gruppe auf sich selbst.

[9] Die Abbildung $f \times f$ ist dabei definiert durch $f \times f \colon H_1 \times H_1 \ni (a, b) \mapsto (f(a), f(b)) \in H_2 \times H_2$.

Wenn $(M_1, \star)$ und $(M_2, \square)$ beides Monoide sind, so können wir ganz analog zu Definition 8.1 die Begriffe **(Monoid-)Homomorphismus**, **-Endomorphismus**, **-Isomorphismus** und **-Automorphismus** (englisch: monoid homomorphism, endomorphism, isomorphism, automorphism) definieren. Zusätzlich zu (8.1) fordert man dabei aber noch, dass für die Einselemente gilt:

$$f(e_1) = e_2. \tag{8.2}$$

Die Monoide $(M_1, \star)$ und $(M_2, \square)$ heißen zueinander **isomorph**, wenn es zwischen ihnen einen Monoidisomorphismus gibt. Wir schreiben dann $(M_1, \star) \cong (M_2, \square)$.

Sind $(G_1, \star)$ und $(G_2, \square)$ beides Gruppen, so ergeben sich die Begriffe **(Gruppen-)Homomorphismus**, **-Endomorphismus**, **-Isomorphismus** und **-Automorphismus** (englisch: group homomorphism, endomorphism, isomorphism, automorphism). Hier wiederum muss man die Bedingung (8.2) nicht separat fordern, denn sie folgt aus (8.1); siehe Lemma 8.5. Die Gruppen $(G_1, \star)$ und $(G_2, \square)$ heißen zueinander **isomorph**, wenn es zwischen ihnen einen Gruppenisomorphismus gibt. Wir schreiben dann $(G_1, \star) \cong (G_2, \square)$.

Bemerkung 8.3 (zu Definition 8.1)**.**
Zwei zueinander isomorphe Halbgruppen/Monoide/Gruppen können und müssen, was ihre algebraischen Eigenschaften als Halbgruppen/Monoide/Gruppen angeht, nicht unterschieden werden. △

Beispiel 8.4 (Homomorphismen von Halbgruppen und Gruppen)**.**

(i) Es sei Σ eine nichtleere Menge und $(\Sigma^*, \circ)$ die Halbgruppe der Tupel über Σ mit der Konkatenation $\circ$, siehe Beispiel 7.4. Die Abbildung $\#\colon (\Sigma^*, \circ) \to (\mathbb{N}_0, +)$, die die Kardinalität eines Tupels angibt, ist ein Monoidhomomorphismus, denn es gilt

$$\begin{aligned} &\#\big((x_1, \ldots, x_n) \circ (y_1, \ldots, y_m)\big) = \#(x_1, \ldots, x_n) + \#(y_1, \ldots, y_m) = n + m \\ \text{und} \quad &\#() = 0. \end{aligned}$$

Genau dann, wenn Σ einelementig ist, ist $\#$ auch bijektiv, also ein Monoidisomorphismus.

(ii) Für $a > 0$, $a \neq 1$ ist $\log_a\colon (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +)$ ist wegen

$$\log_a(x \cdot y) = \log_a(x) + \log_a(y)$$

ein Gruppenhomomorphismus. Weiter ist $\log_a$ bijektiv, also sogar ein Gruppenisomorphismus.

(iii) 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.

(iv) Für festes $n \in \mathbb{Z}$ ist die Abbildung (vgl. (7.6))

$$G \ni a \mapsto a^n \in G$$

in einer *abelschen* Gruppe $(G, \cdot)$ ein Gruppenendomorphismus.

(v) 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 7.29

$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = (\operatorname{sgn} \sigma_1) \cdot (\operatorname{sgn} \sigma_2).$$

Genau für $n = 2$ ist sgn auch bijektiv, also ein Gruppenisomorphismus.

(*vi*) Die in Beispiel 7.22 vorgenommene „Identifikation“ der symmetrischen Gruppe S_3 mit der Gruppe der Kongruenzabbildungen eines gleichseitigen Dreiecks stellt einen Gruppenisomorphismus dar.

(*vii*) Die Abbildung

$$\mathbb{R} \ni f(x) := \exp(ix) = \cos(x) + i\sin(x) \in \mathbb{C}$$

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 die **Additionstheoreme** für die Winkelsumme

$$\cos(x+y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y) \tag{8.3a}$$

$$\sin(x+y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y). \tag{8.3b}$$

△

Lemma 8.5 (Eigenschaften von Gruppenhomomorphismen)**.**
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. Dann gilt:

(*i*) $f(e_1) = e_2$.

(*ii*) $(f(a))' = f(a')$.

Beweis. Es gilt

$$\begin{aligned} f(e_1) \,\square\, e_2 &= f(e_1) && \text{da } e_2 \text{ neutrales Element in } (H_2, \square) \text{ ist} \\ &= f(e_1 \star e_1) && \text{da } e_1 \text{ neutrales Element in } (H_1, \star) \text{ ist} \\ &= f(e_1) \,\square\, f(e_1) && \text{da } f \text{ Homomorphismus ist.} \end{aligned}$$

Die Verknüpfung dieses Ausdrucks von links mit dem Inversen von $f(e_1)$, also die Anwendung der Kürzungsregel (7.10a), zeigt $e_2 = f(e_1)$, also Aussage (*i*).

Die Aussage (*ii*) folgt aus

$$\begin{aligned} f(a') \,\square\, f(a) &= f(a' \star a) && \text{da } f \text{ Homomorphismus ist} \\ &= f(e_1) && \text{da } e_1 \text{ neutrales Element in } (H_1, \star) \text{ ist} \\ &= e_2 && \text{wegen Aussage } (i). \end{aligned}$$

Aus (7.11b) folgt nun $f(a') = (f(a))'$. □

Beachte: Gruppenhomomorphismen bilden neutrale Elemente auf neutrale Element ab und inverse Elemente auf inverse Elemente. Das Lemma 8.5 gilt i. A. nicht, wenn $(G_2, \square)$ keine Gruppe, sondern nur ein Monoid ist!

Quizfrage 8.2: Kann $f\colon (\mathbb{Z}, +) \ni n \mapsto n+1 \in (\mathbb{Z}, +)$ ein Gruppenhomomorphismus sein?

Wir wollen nun Gruppenhomomorphismen genauer studieren.

Definition 8.6 (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

$$\operatorname{Bild}(f) \coloneqq \{f(x) \in G_2 \,|\, x \in G_1\} = f(G_1). \tag{8.4}$$

(ii) Der **Kern** (englisch: kernel) von f ist definiert als

$$\operatorname{Kern}(f) \coloneqq \{x \in G_1 \,|\, f(x) = e_2\} = f^{-1}(\{e_2\}). \tag{8.5}$$

△

Lemma 8.7 (Bild und Kern sind Untergruppen).
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) $\operatorname{Bild}(f)$ ist eine Untergruppe von $(G_2, \square)$.

(ii) $\operatorname{Kern}(f)$ ist eine Untergruppe von $(G_1, \star)$.

Beweis. Aussage (i): Wir überprüfen das Untergruppenkriterium (Satz 7.33). Es gilt $e_2 = f(e_1)$ nach Lemma 8.5, 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: $a_2 \square b_2' \in \operatorname{Bild}(f)$.

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

$$\begin{aligned}
a_2 \square b_2' &= f(a_1) \square (f(b_1))' && \text{nach Voraussetzung} \\
&= f(a_1) \square f(b_1') && \text{nach Lemma 8.5} \\
&= f(a_1 \star b_1') && \text{da } f \text{ Homomorphismus ist}
\end{aligned}$$

und damit $a_2 \square b_2' \in \operatorname{Bild}(f)$.

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, b irgendwelche Elemente in $\operatorname{Kern}(f)$. Wir müssen zeigen: $a \star b' \in \operatorname{Kern}(f)$.

$$\begin{aligned}
f(a \star b') &= f(a) \square f(b') && \text{da } f \text{ Homomorphismus ist} \\
&= f(a) \square (f(b))' && \text{nach Lemma 8.5} \\
&= e_2 \square e_2' && \text{da } a, b \in \operatorname{Kern}(f) \text{ liegen} \\
&= e_2 && \text{da } e_2' = e_2 \text{ ist.}
\end{aligned}$$

Damit ist $a \star b' \in \operatorname{Kern}(f)$ gezeigt. □

Beispiel 8.8 (Bild und Kern sind Untergruppen).

(i) Für die Abbildung $\#\colon (\Sigma^*, \circ) \to (\mathbb{N}_0, +)$ aus Beispiel 8.4 gilt:

$$\operatorname{Bild}(\#) = \mathbb{N}_0 \quad \text{und} \quad \operatorname{Kern}(\#) = \{()\},$$

wobei $()$ das leere Tupel kennzeichnet.

(ii) Für die Abbildung $\operatorname{sgn}\colon (S_n, \circ) \to (\{\pm 1\}, \cdot)$ aus Beispiel 8.4 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).

(iii) Für die Abbildung $f\colon (\mathbb{R}_{\neq 0}, \cdot) \ni x \mapsto x^2 (\mathbb{R}_{\neq 0}, \cdot)$ gilt

$$\operatorname{Bild}(f) = \mathbb{R}_{>0} \quad \text{und} \quad \operatorname{Kern}(f) = \{\pm 1\}. \qquad \triangle$$

Lemma 8.9 (Charakterisierung der Injektivität)**.**

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. 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 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 ist, vereinfacht sich dieser Nachweis erheblich. Wir müssen dann nur noch zeigen, dass neben dem neutralen Element e_1 kein weiteres Element auf das neutrale Element e_2 abgebildet wird.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Nach Lemma 8.5 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

$$\begin{aligned} f(a \star b') &= f(a) \mathbin{\square} f(b') \\ &= f(a) \mathbin{\square} (f(b))' \\ &= f(a) \mathbin{\square} (f(a))' \\ &= e_2, \end{aligned}$$

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 8.5 $f(e_1) = e_2$ gilt. $\square$

Ende der Vorlesung 10

Ende der Woche 5

§ 8.1 Normalteiler

Wir hatten in Lemma 7.40 gesehen, dass jede Untergruppe $(U, \star)$ einer Gruppe $(G, \star)$ zwei Äquivalenzrelationen $\sim^U$ und ${}^U\!\sim$ auf G induziert, deren Äquivalenzklassen durch $a \star U$ bzw. $U \star a$ gegeben und die i. A. verschieden sind.

Definition 8.10 (Normalteiler)**.**
Es sei $(G, \star)$ eine Gruppe. Eine Untergruppe $(N, \star)$ heißt eine **normale Untergruppe** (englisch: normal subgroup) oder **Normalteiler** von $(G, \star)$, wenn gilt:

$$a \star N = N \star a \quad \text{für alle } a \in G. \tag{8.6}$$

Manchmal notiert man die Eigenschaft, dass $(N, \star)$ ein Normalteiler der Gruppe $(G, \star)$ ist, als $(N, \star) \trianglelefteq (G, \star)$. △

Anders ausgedrückt ist $(N, \star)$ genau dann eine normale Untergruppe, wenn die durch sie induzierten Äquivalenzrelationen $\sim^N$ und ${}^N\!\sim$ (siehe Lemma 7.40) übereinstimmen.

Beachte: Die Relation „ist Normalteiler von" ist zwar reflexiv und antisymmetrisch, aber im Gegensatz zur Relation „ist Untergruppe von" i. A. nicht transitiv!

Beispiel 8.11 (Normalteiler)**.**

(*i*) In jeder Gruppe $(G, \star)$ sind die trivialen Untergruppen $(\{e\}, \star)$ und $(G, \star)$ Normalteiler.

(*ii*) In einer abelschen Gruppe $(G, \star)$ ist jede Untergruppe ein Normalteiler (Folgerung 7.41). △

Lemma 8.12 (Kerne von Gruppenhomomorphismen sind Normalteiler)**.**
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. Dann gilt:

(*i*) Für alle $a \in G_1$ gilt:
$$f^{-1}(\{f(a)\}) = a \star \operatorname{Kern}(f) = \operatorname{Kern}(f) \star a.$$

(*ii*) $\operatorname{Kern}(f)$ ist ein Normalteiler von G_1.

Beachte: Das Urbild eines Elements in der Bildmenge $f(G_1)$ ist also immer eine Nebenklasse von $\operatorname{Kern}(f)$.

Beweis. Wir zeigen zunächst die Aussage (*i*) in mehreren Schritten.

Schritt 1: $f^{-1}(\{f(a)\}) \subseteq \operatorname{Kern}(f) \star a$:

Es sei $b \in f^{-1}(\{f(a)\})$, also $f(b) = f(a)$. Dann gilt also

$$\begin{aligned} e_2 &= f(b) \square (f(a))' \\ &= f(b) \square f(a') && \text{nach Lemma 8.5} \\ &= f(b \star a') && \text{da } f \text{ Homomorphismus ist.} \end{aligned}$$

Das heißt aber, dass $b \star a' \in f^{-1}(\{e_2\}) = \operatorname{Kern}(f)$ liegt. Mit anderen Worten: $b \in \operatorname{Kern}(f) \star a$.

Schritt 2: $f^{-1}(\{f(a)\}) \subseteq a \star \operatorname{Kern}(f)$:

Ganz analog zu **Schritt 1** gilt auch

$$\begin{aligned} e_2 &= (f(a))' \sqcap f(b) \\ &= f(a') \sqcap f(b) && \text{nach Lemma 8.5} \\ &= f(a' \star b) && \text{da } f \text{ Homomorphismus ist.} \end{aligned}$$

Das heißt aber $a' \star b \in f^{-1}(\{e_2\}) = \operatorname{Kern}(f)$ und daher $b \in a \star \operatorname{Kern}(f)$.

Schritt 3: $\operatorname{Kern}(f) \star a \subseteq f^{-1}(\{f(a)\})$:

Es sei $b \in \operatorname{Kern}(f)$. Wir müssen $b \star a \in f^{-1}(\{f(a)\})$ zeigen, also $f(b \star a) = f(a)$. Das folgt aber sofort aus

$$\begin{aligned} f(b \star a) &= f(b) \sqcap f(a) && \text{da } f \text{ Homomorphismus ist} \\ &= e_2 \sqcap f(a) && \text{da } b \in \operatorname{Kern}(f) \text{ ist} \\ &= f(a). \end{aligned}$$

Schritt 4: $a \star \operatorname{Kern}(f) \subseteq f^{-1}(\{f(a)\})$:

Es sei $b \in \operatorname{Kern}(f)$. Wir müssen $a \star b \in f^{-1}(\{f(a)\})$ zeigen, also $f(a \star b) = f(a)$. Das folgt aber sofort aus

$$\begin{aligned} f(a \star b) &= f(a) \sqcap f(b) && \text{da } f \text{ Homomorphismus ist} \\ &= f(a) \sqcap e_2 && \text{da } b \in \operatorname{Kern}(f) \text{ ist} \\ &= f(a). \end{aligned}$$

Aus Lemma 8.7 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). □

Wenn $(N, \star)$ ein Normalteiler einer Gruppe $(G, \star)$ ist, dann können wir die Faktormenge $G / \overset{N}{\sim} = G / N$ mit einer Gruppenverknüpfung $\tilde{\star}$ ausstatten. Aus der Faktormenge wird damit die **Faktorgruppe** (englisch: factor group) oder **Quotientengruppe** (englisch: quotient group) von G nach N. Man sagt auch: „Aus der Gruppe $(G, \star)$ wird der Normalteiler N ausfaktorisiert."

Satz 8.13 (Faktorgruppe)**.**
Es sei $(G, \star)$ eine Gruppe mit neutralem Element e und $(N, \star)$ einer ihrer Normalteiler. Dann gilt:

(i) Auf der Faktormenge

$$G / N = \{[a] = a \star N \mid a \in G\}$$

ist $\tilde{\star}$, definiert als

$$[a] \mathbin{\tilde{\star}} [b] := [a \star b] \quad \text{für } a, b \in G, \tag{8.7}$$

eine assoziative Verknüpfung, bzgl. der $(G / N, \tilde{\star})$ eine Gruppe bildet. Das neutrale Element ist $[e] = N$, und für die Inversen gilt $[a]' = [a']$.

(*ii*) Die Abbildung

$$\pi\colon \begin{cases} G \to G \,/\, N \\ a \mapsto [a], \end{cases} \tag{8.8}$$

die jedem Element $a \in G$ seine Nebenklasse $[a]$ zuordnet, ist ein surjektiver Gruppenhomomorphismus. Sie heißt die **kanonische Surjektion** (englisch: canonical surjection) von G auf $G \,/\, N$. Es gilt $\operatorname{Kern}(\pi) = N$.

(*iii*) Wenn $(G, \star)$ abelsch ist, dann auch $(G \,/\, N, \tilde{\star})$.

Beweis. Aussage (*i*): Wir müssen zunächst zeigen, dass $\tilde{\star}$ überhaupt eine Verknüpfung auf $G \,/\, N$ darstellt, also dass (8.7) wohldefiniert ist, da wir dort ja Bezug auf konkrete Repräsentanten $a, b \in G$ nehmen. Es seien also $a_1, a_2, b_1, b_2 \in G$ gegeben, wobei $a_1 \overset{N}{\sim} a_2$ und $b_1 \overset{N}{\sim} b_2$ angenommen wird, d. h., $a_1 \star N = a_2 \star N$ und $b_1 \star N = b_2 \star N$. Dann gilt

$$\begin{aligned}
[a_2] \,\tilde{\star}\, [b_2] &= [a_2 \star b_2] \\
&= (a_2 \star b_2) \star N && \text{nach (7.25)} \\
&= a_2 \star (b_2 \star N) && \text{da } \star \text{ assoziativ ist} \\
&= a_2 \star (N \star b_2) && \text{da } N \text{ Normalteiler ist} \\
&= a_2 \star N \star b_2 && \text{da } \star \text{ assoziativ ist} \\
&= a_2 \star N \star N \star b_2 && \text{da } N \text{ Untergruppe ist} \\
&= (a_2 \star N) \star (N \star b_2) && \text{da } \star \text{ assoziativ ist} \\
&= (a_1 \star N) \star (N \star b_1) && \text{da } a_1 \overset{N}{\sim} a_2 \text{ und } b_1 \overset{N}{\sim} b_2 \\
&= (a_1 \star N) \star (b_1 \star N) && \text{da } N \text{ Normalteiler ist} \\
&= [a_1] \,\tilde{\star}\, [b_1].
\end{aligned}$$

Damit ist $\tilde{\star}$ als Verknüpfung auf $G \,/\, N$ wohldefiniert. Die Assoziativität von $\tilde{\star}$ ergibt sich aus der Assoziativität von $\star$ und der Normalteilereigenschaft, denn es gilt:

$$\begin{aligned}
([a] \,\tilde{\star}\, [b]) \,\tilde{\star}\, [c] &= [a \star b] \,\tilde{\star}\, [c] = (a \star b \star N) \star (c \star N) = (a \star b \star N) \star (N \star c) \\
&= a \star b \star N \star c = a \star b \star c \star N \\
[a] \,\tilde{\star}\, ([b] \,\tilde{\star}\, [c]) &= [a] \,\tilde{\star}\, [b \star c] = (a \star N) \star (b \star c \star N) = (a \star N) \star (N \star b \star c) \\
&= a \star N \star b \star c = a \star b \star c \star N
\end{aligned}$$

Damit haben wir zunächst $(G \,/\, N, \tilde{\star})$ als Halbgruppe bestätigt.

Als nächstes zeigen wir, 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 (8.7)

$$[e] \,\tilde{\star}\, [a] = [e \star a] = [a] \quad \text{sowie} \quad [a] \,\tilde{\star}\, [e] = [a \star e] = [a].$$

Also ist $(G \,/\, N, \tilde{\star})$ ein Monoid mit neutralem Element $[e]$.

Nun zeigen wir, dass jedes $[a] \in G \,/\, N$ invertierbar ist mit Inverser $[a]' = [a']$:

$$[a] \,\tilde{\star}\, [a'] = [a \star a'] = [e] \quad \text{sowie} \quad [a'] \,\tilde{\star}\, [a] = [a' \star a] = [e].$$

Aussage (*ii*): Die Eigenschaft, ein Gruppenhomomorphismus zu sein, bedeutet $\pi(a \star b) = \pi(a) \,\tilde{\star}\, \pi(b)$. Nach Definition von π heißt das aber gerade $[a \star b] = [a] \,\tilde{\star}\, [b]$, 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 (iii): 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. □

Bemerkung 8.14 (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 $a \star b' \in N$ gilt) nicht mehr unterschieden werden. Die Faktorgruppe $(G\,/\,N, \tilde{\star})$ ist also eine „gröbere Version" der Gruppe $(G, \star)$. △

Beispiel 8.15 (Faktorgruppe)**.**

(i) Es sei $(G, \star)$ eine beliebige Gruppe. Dann ist die triviale Untergruppe $\{e\}$ nach Beispiel 8.11 ein Normalteiler. Die zugehörige Faktorgruppe $(G\,/\,\{e\}, \tilde{\star})$ ist isomorph zur Ausgangsgruppe $(G, \star)$ selbst.

(ii) Es sei $(G, \star)$ eine beliebige Gruppe. Dann ist die triviale Untergruppe G nach Beispiel 8.11 ein Normalteiler. Die zugehörige Faktorgruppe $(G\,/\,G, \tilde{\star})$ ist isomorph zu $(\{e\}, \star)$.

(iii) In der abelschen Gruppe $(\mathbb{Z}, +)$ ist jede Untergruppe der Form $m\,\mathbb{Z}$ mit $m \in \mathbb{N}$ 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 7.42. Es gilt

$$[a] \mathbin{\tilde{+}} [b] = [a + b].$$

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 7.16 mittels des Isomorphismus $[a] \mapsto$ natürlicher Repräsentant von a in $\mathbb{Z}_m$. Beispielsweise können wir für $m = 5$ wie folgt rechnen:

$$\begin{tikzcd}[row sep=large] \text{in } (\mathbb{Z}\,/\,5\,\mathbb{Z}, \tilde{+}) & {[-21]} \arrow[d, mapsto] & \tilde{+} & {[9]} \arrow[d, mapsto] & = & {[-12]} \arrow[d, mapsto] \\ \text{in } (\mathbb{Z}_5, +_5) & 4 & +_5 & 4 & = & 3 \end{tikzcd}$$

(iv) In der abelschen Gruppe $(\mathbb{R}_{\neq 0}, \cdot)$ ist die Untergruppe $(\{\pm 1\}, \cdot)$ ein Normalteiler. Die Elemente der Faktorgruppe sind die Nebenklassen

$$[a] = a \cdot \{\pm 1\} = \{a, -a\}$$

für $a \in \mathbb{R}_{\neq 0}$. Ein mögliches Repräsentantensystem ist $\mathbb{R}_{>0}$. △

Bemerkung 8.16 (Normalteiler sind genau die Kerne von Gruppenhomomorphismen)**.**
Es sei $(G_1, \star)$ eine Gruppe. Nach Lemma 8.12 ist für jeden beliebigen Gruppenhomomorphismus $f\colon G_1 \to G_2$ in irgendeine Gruppe $(G_2, \square)$ die Untergruppe $\operatorname{Kern}(f)$ immer ein Normalteiler von $(G_1, \star)$.

Umgekehrt kann man zeigen, dass jeder Normalteiler von dieser Form ist. Also gilt: Jeder Normalteiler von $(G_1, \star)$ ist der Kern eines geeignet gewählten Gruppenhomomorphismus von $(G_1, \star)$ in eine geeignet gewählte Gruppe $(G_2, \square)$. △

§ 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. 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 die Untergruppe $\operatorname{Kern}(f)$.

Satz 8.17 (Homomorphiesatz für Gruppen[10]).
Es seien $(G_1, \star)$ und $(G_2, \square)$ Gruppen. Weiter sei $f\colon G_1 \to G_2$ ein Homomorphismus. Dann gilt

$$G_1 \,/\, \operatorname{Kern}(f) \cong \operatorname{Bild}(f) \tag{8.9a}$$

mit dem Isomorphismus

$$I([a]) \coloneqq f(a) \quad \text{für } [a] = a \star \operatorname{Kern}(f) \in G_1 \,/\, \operatorname{Kern}(f). \tag{8.9b}$$

Beweis. Wir bezeichnen die neutralen Elemente von G_1 und G_2 mit e_1 bzw. e_2.

Wir definieren $I\colon G_1 \,/\, \operatorname{Kern}(f) \to \operatorname{Bild}(f)$ wie in (8.9).

Schritt 1: Wir müssen zunächst zeigen, dass I als Abbildung wohldefiniert ist, da wir in der Definition (8.9b) Bezug auf den konkreten Repräsentanten $a \in G_1$ nehmen.

Es seien dazu $a, b \in G_1$ gegeben mit $a \overset{\operatorname{Kern}(f)}{\sim} b$, d. h., $a \star \operatorname{Kern}(f) = b \star \operatorname{Kern}(f)$. Dann gilt

$$\begin{aligned} f(a \star \operatorname{Kern}(f)) &= f(a) \square f(\operatorname{Kern}(f)) && \text{da } f \text{ Homomorphismus ist} \\ &= \{f(a)\} && \text{da } f(\operatorname{Kern}(f)) = \{e_2\} \text{ gilt} \end{aligned}$$

und analog $f(b \star \operatorname{Kern}(f)) = \{f(b)\}$. Aus $a \star \operatorname{Kern}(f) = b \star \operatorname{Kern}(f)$ folgt also $f(a) = f(b)$. Außerdem ist nach Definition von I klar, dass I in $\operatorname{Bild}(f)$ abbildet. Damit ist I wohldefiniert.

Schritt 2: Als nächstes zeigen wir, dass I ein Homomorphismus ist. In der Tat gilt

$$\begin{aligned} I([a] \mathbin{\tilde{\star}} [b]) &= I([a \star b]) && \text{nach Definition (8.7) von } \tilde{\star} \\ &= f(a \star b) && \text{nach Definition von } I \\ &= f(a) \square f(b) && \text{da } f \text{ Homomorphismus ist} \\ &= I([a]) \square I([b]) && \text{nach Definition von } I. \end{aligned}$$

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.

[10] englisch: fundamental theorem on group homomorphisms

Für die Injektivität genügt es nach Lemma 8.9 zu zeigen, dass $\operatorname{Kern}(I)$ nur aus dem neutralen Element des Definitionsbereiches $G_1 / \operatorname{Kern}(f)$ besteht, d. h., aus $[e_1] = \operatorname{Kern}(f)$, vgl. Satz 8.13. Es gilt

$$\begin{aligned}
\operatorname{Kern}(I) &= \{[a] \in G_1 / \operatorname{Kern}(f) \mid I([a]) = e_2\} && \text{nach Definition von } \operatorname{Kern}(I)\\
&= \{[a] \in G_1 / \operatorname{Kern}(f) \mid f(a) = e_2\} && \text{nach Definition von } I\\
&= \{[a] \in G_1 / \operatorname{Kern}(f) \mid a \in \operatorname{Kern}(f)\} && \text{nach Definition von } \operatorname{Kern}(f)\\
&= \{a \star \operatorname{Kern}(f) \mid a \in \operatorname{Kern}(f)\} && \text{wegen } [a] = a \star \operatorname{Kern}(f)\text{, siehe (7.25)}\\
&= \{\operatorname{Kern}(f)\} && \text{, denn } \operatorname{Kern}(f) \text{ ist Untergruppe von } (G_2, \square)\\
& && \text{nach Lemma 8.7.}
\end{aligned}$$

□

Beispiel 8.18 (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 8.4 und 8.8. 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 Nebenklassen

$$\begin{aligned}
[\mathrm{id}] &= \mathrm{id} \circ A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\} && \text{(gerade Permutationen),}\\
[\tau] &= \tau \circ A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = -1\} && \text{(ungerade Permutationen),}
\end{aligned}$$

wobei τ irgendeine Transposition in S_n ist. Gemäß Homomorphiesatz 8.17 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 8.17 besagt daher in diesem Fall

$$S_1 / \operatorname{Kern}(\operatorname{sgn}) = S_1 / A_1 \cong \operatorname{Bild}(\operatorname{sgn}) = \{1\}.$$

(*ii*) Für die Abbildung $f\colon (\mathbb{R}_{\neq 0}, \cdot) \ni x \mapsto x^2 \in (\mathbb{R}_{\neq 0}, \cdot)$ aus Beispiel 8.8 und Beispiel 8.15 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. △

Ende der Vorlesung 11

§ 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 die wichtigen Beispiele $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ mit den Verknüpfungen „Addition“ und „Multiplikation“ bezeichnen wir diese Verknüpfungen 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.[11] △

Wie üblich vereinbaren wir, dass $\cdot$ stärker bindet als $+$ („Punkt- vor Strichrechnung"), also könnten wir die rechte Seite in (9.1a) auch in der Form $a \cdot b + a \cdot c$ schreiben.

Wie in Gruppen in additiver Notation üblich (Bemerkung 7.13), bezeichnen wir das neutrale Element bzgl. $+$ als **Nullelement** (englisch: additive identity) und schreiben dafür zunächst „0_R". Außerdem bezeichnen wir das bzgl. $+$ inverse Element von $a \in R$ mit $-a$. Die Bezeichnung $a - b$ steht für $a + (-b)$.

Falls $(R, \cdot)$ 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). Existiert dann zu $a \in R$ bzgl. $\cdot$ ein inverses Element, so bezeichnen wir dieses mit a^{-1}.

Wir vereinbaren, dass $\cdot$ stärker bindet als $+$ und $-$, sodass wir beispielsweise statt $(a \cdot b) + (a \cdot c)$ auch $a \cdot b + a \cdot c$ schreiben können. Außerdem können wir $-a \cdot b$ schreiben statt $-(a \cdot b)$.

Beispiel 9.2 (Ring)**.**

(i) $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind kommutative Ringe mit Eins.

(ii) Der **Nullring** (englisch: zero ring) ist der (bis auf Isomorphie) eindeutig bestimmte Ring mit $R = \{0_R\}$ und den dadurch eindeutig bestimmten Verknüpfungen $0_R + 0_R = 0_R$ und $0_R \cdot 0_R = 0_R$. Da 0_R auch das neutrale Element bzgl. $\cdot$ ist, ist der Nullring ein Ring mit Eins, und es gilt $1_R = 0_R$. Er ist der einzige Ring, in dem das Nullelement und das Einselement identisch sind, siehe Lemma 9.3.

(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, denn nach Beispiel 7.16 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 $(\mathbb{Z}_m, +_m, \cdot_m)$ der Nullring.

[11] In diesem Fall fallen die beiden Distributivgesetze (9.1a) und (9.1b) zusammen. Es reicht also, eines von beiden zu prüfen.

(*v*) Es sei $(G, +)$ eine abelsche Gruppe. Wir definieren

$$\operatorname{End}(G) := \left\{ f\colon G \to G \,\middle|\, f \text{ ist Endomorphismus} \right\} \tag{9.2}$$

und statten $\operatorname{End}(G)$ mit den Verknüpfungen

$$\begin{aligned} +\colon\ & \operatorname{End}(G) \times \operatorname{End}(G) \to \operatorname{End}(G) \quad \text{mit } (f, g) \mapsto f + g, \text{definiert durch } (f + g)(x) := f(x) + g(x) \\ \circ\colon\ & \operatorname{End}(G) \times \operatorname{End}(G) \to \operatorname{End}(G) \quad \text{mit } (f, g) \mapsto f \circ g, \text{definiert durch } (f \circ g)(x) := f(g(x)) \end{aligned}$$

aus. Dann ist $(\operatorname{End}(G), +, \circ)$ ein Ring mit Einselement id_G, genannt der **Endomorphismenring** (englisch: ring of endomorphisms) der abelschen Gruppe $(G, +)$. Er ist i. A. nicht kommutativ.

Quizfrage 9.1: Warum definieren wir den Endomorphismenring nur für Endomorphismen auf abelschen Gruppen und nicht allgemeiner für Endomorphismen auf beliebigen Gruppen? △

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 nicht der Nullring, dann gilt $1_R \neq 0_R$.

Beachte: Hat R das Einselement 1_R, dann folgt aus Aussage (*ii*) insbesondere $-b = (-1_R) \cdot b$.

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 (7.10a) in der Gruppe $(R, +)$, also die Addition von $-(0_R \cdot a)$ zu beiden Seiten der Gleichung, zeigt $0_R \cdot a = 0_R \cdot a$. 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. Gemäß (7.11) 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

$$\begin{aligned} (-a) \cdot (-b) &= -(a \cdot (-b)) && \text{nach Aussage } (ii) \\ &= -(-a \cdot b) && \text{nach Aussage } (ii) \\ &= a \cdot b && \text{nach (7.12) (doppelte Invertierung).} \end{aligned}$$

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 der Nullring. □

Wir verwenden auch in Ringen $(R, +, \cdot)$ und insbesondere in der Gruppe $(R, +)$ die Schreibweise aus Bemerkung 7.13. Es gilt also für $n \in \mathbb{N}$

$$n\,a := a + \cdots + a.$$

Besitzt $(R, +, \cdot)$ das Einselement 1_R, dann gilt nach Distributivgesetz weiter

$$n\,a = a + \cdots + a = 1_R \cdot a + \cdots + 1_R \cdot a = (1_R + \cdots + 1_R) \cdot a = (n\,1_R) \cdot a.$$

Weiter ist $(-n)\,a := -(n\,a)$ und $0\,a := 0_R$.

Definition 9.4 (Charakteristik eines Ringes)**.**
Es sei R 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** (englisch: characteristic) 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*) Der Nullring ist (bis auf Isomorphie) der einzige Ring mit Charakteristik 1, also der einzige Ring, in dem $1_R = 0_R$ 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 8.15 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 $[0]$. $(\mathbb{Z}/m\mathbb{Z}, \tilde{+})$ bildet eine kommutative Gruppe.

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 Hausaufgabe I-6.1.

Schließlich können wir zeigen, dass die Distributivgesetze (9.1a) und (9.1b) gelten, denn:

$$
\begin{aligned}
[a] \mathbin{\tilde{\cdot}} ([b] \mathbin{\tilde{+}} [c]) &= [a] \mathbin{\tilde{\cdot}} [b+c] && \text{nach Definition von } \tilde{+} \\
&= [a \cdot (b+c)] && \text{nach Definition von } \tilde{\cdot} \\
&= [a \cdot b + a \cdot c] && \text{nach Distributivgesetz in } \mathbb{Z} \\
&= [a \cdot b] \mathbin{\tilde{+}} [a \cdot c] && \text{nach Definition von } \tilde{+} \\
&= [a] \mathbin{\tilde{\cdot}} [b] \mathbin{\tilde{+}} [a] \mathbin{\tilde{\cdot}} [c] && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}
$$

Das zweite Distributivgesetz (9.1b) ist wegen der Kommutativität der Halbgruppe $(\mathbb{Z}/m\mathbb{Z}, \tilde{\cdot})$ automatisch 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 $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ der 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 für $m = 4$ 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), wenn es ein $b \in R \setminus \{0_R\}$ gibt, sodass $a \cdot b = 0_R$ gilt. Das Element b heißt ein **Rechtsnullteiler** (englisch: right zero divisor), wenn es ein $a \in R \setminus \{0_R\}$ gibt, sodass $a \cdot b = 0_R$ gilt.

(ii) Der Ring $(R, +, \cdot)$ heißt **nullteilerfrei** (englisch: ring with no zero divisors), wenn es außer dem Nullelement 0_R keine weiteren Links- oder Rechtsnullteiler gibt, wenn also gilt:

$$\forall a, b \in R \; (a \cdot b = 0_R \;\Rightarrow\; a = 0_R \text{ oder } b = 0_R). \tag{9.3}$$

(Anders gesagt: Aus $a \neq 0_R$ und $b \neq 0_R$ folgt $a \cdot b \neq 0_R$.)

(*iii*) Ein Ring $(R, +, \cdot)$ heißt **Integritätsring** oder **Integritätsbereich** (englisch: integral domain), wenn gilt: $(R, +, \cdot)$ ist ein kommutativer, nullteilerfreier Ring mit Eins ungleich dem Nullring. △

Quizfrage 9.2: Ist der Nullring nullteilerfrei?

Beispiel 9.8 (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 9.9.

(*iii*) Es sei X eine Menge, $(R, +, \cdot)$ ein Ring und $R^X = \{f \mid f\colon X \to R\}$. Definieren wir ähnlich wie in Beispiel 7.2 die Verknüpfungen $+$ und $\cdot$ auf R^X als

$$\begin{aligned} +\colon R^X \times R^X \to R^X \quad &\text{mit } (f, g) \mapsto f + g, \text{definiert durch } (f + g)(x) \coloneqq f(x) + g(x), \\ \cdot\colon R^X \times R^X \to R^X \quad &\text{mit } (f, g) \mapsto f \cdot g, \text{definiert durch } (f \cdot g)(x) \coloneqq f(x) \cdot g(x), \end{aligned}$$

dann ist $(R^X, +, \cdot)$ ein Ring. $(R^X, +, \cdot)$ ist kommutativ genau dann, wenn $(R, +, \cdot)$ kommutativ ist. $(R^X, +, \cdot)$ besitzt ein Einselement genau dann, $(R, +, \cdot)$ ein Einselement besitzt.

Beachte: Wenn $(R, +, \cdot)$ nicht der Nullring ist, dann ist $(R^X, +, \cdot)$ nicht nullteilerfrei, sobald X zwei oder mehr Elemente enthält! (**Quizfrage 9.3:** Wie sieht man das?) △

Satz 9.9 (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 $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ der Nullring und damit kein Integritätsring. Wir betrachten also im Weiteren nur den Fall $m \geqslant 2$, für den $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ ein kommutativer Ring ungleich dem Nullring und mit dem Einselement $[1]$ ist. 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 $m \in \mathbb{N}$, $m \geqslant 4$, *keine* Primzahl, lässt sich also schreiben als $m = a \cdot b$ für Zahlen $a, b \in [\![2, m-1]\!]$. Die zugehörigen Restklassen $[a]$ und $[b]$ sind ungleich $[0]$ (**Quizfrage 9.4:** Warum?) Es gilt

$$\begin{aligned} [0] &= [m] && \text{da } 0 \overset{m}{\equiv} m \\ &= [a \cdot b] && \text{da } m = a \cdot b \\ &= [a] \mathbin{\tilde{\cdot}} [b] && \text{nach Definition von } \tilde{\cdot}. \end{aligned}$$

Damit ist $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ *nicht* nullteilerfrei.

Es sei nun umgekehrt $m \in \mathbb{N}$, $m \geqslant 2$, eine Primzahl. Wir nehmen an, $[a]$ und $[b]$ seien Elemente aus $\mathbb{Z}/m\mathbb{Z}$ mit $[0] = [a] \mathbin{\tilde{\cdot}} [b] = [a \cdot b]$. 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. □

Definition 9.10 (Unterring, vgl. Definition 7.31 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 ist und wenn $(U, +, \cdot)$ selbst wieder ein Ring ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, +)$ eine Untergruppe von $(R, +)$ ist und wenn U bzgl. $\cdot$ abgeschlossen ist.

(*ii*) Ist $(R, +, \cdot)$ ein Ring mit Einselement 1_R, dann fordern wir für einen **Unterring mit Eins** (englisch: subring with unity) $(U, +, \cdot)$ zusätzlich zu Eigenschaft (*i*), dass $1_R \in U$ liegt.[12]

Beachte: Es reicht nicht aus, zu fordern, dass $(U, \cdot)$ *irgendein* neutrales Element besitzt.

(*iii*) Ein Unterring $(U, +, \cdot)$ von $(R, +, \cdot)$ heißt **echt** (englisch: proper subring), wenn $U \subsetneq R$ gilt. △

Definition 9.11 (Ringhomomorphismus, vgl. Definition 8.1 eines Halbgruppenhomomorphismus)**.**
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 **(Ring-)Homomorphismus** (englisch: ring homomorphism) von $(R_1, +_1, \cdot_1)$ in $(R_2, +_2, \cdot_2)$, wenn gilt:

$$f(a +_1 b) = f(a) +_2 f(b) \quad \text{für alle } a, b \in R_1, \tag{9.4a}$$
$$f(a \cdot_1 b) = f(a) \cdot_2 f(b) \quad \text{für alle } a, b \in R_1. \tag{9.4b}$$

Besitzen beide Ringe ein Einselement 1_{R_1} bzw. 1_{R_2} und fordern wir zusätzlich

$$f(1_{R_1}) = 1_{R_2} \tag{9.4c}$$

dann nennen wir f genauer einen **Homomorphismus von Ringen mit Eins** (englisch: homomorphism of rings with unity).

(*ii*) Wie in Definition 8.1 sprechen wir im Fall $(R_1, +_1, \cdot_1) = (R_2, +_2, \cdot_2)$ von einem **(Ring-)Endomorphismus** (englisch: ring endomorphism) bzw. von einem **Endomorphismus eines Ringes mit Eins** (englisch: endomorphism of a ring with unity).

(*iii*) Ist $f\colon R_1 \to R_2$ bijektiv, so heißt f auch **strukturerhaltend** oder ein **(Ring-)Isomorphismus** (englisch: ring isomorphism) bzw. ein **Isomorphismus von Ringen mit Eins** (englisch: endomorphism 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*) Im Fall $(R_1, +_1, \cdot_1) = (R_2, +_2, \cdot_2)$ und $f\colon R_1 \to R_2$ bijektiv sprechen wir auch von einem **(Ring-)Automorphismus** (englisch: ring automorphism) bzw. von einem **Automorphismus eines Ringes mit Eins** (englisch: automorphism of a ring with unity).

(*v*) Das **Bild** (englisch: image) und der **Kern** (englisch: kernel) eines Ringhomomorphismus $f\colon R_1 \to R_2$ sind definiert als

$$\begin{aligned} \operatorname{Bild}(f) &:= \left\{ f(x) \in R_2 \,\middle|\, x \in R_1 \right\} = f(R_1), \\ \operatorname{Kern}(f) &:= \left\{ x \in R_1 \,\middle|\, f(x) = 0_{R_2} \right\} = f^{-1}(\{0_{R_2}\}). \end{aligned} \tag{9.5}$$
△

[12] Dadurch ist der Unterring $(U, +, \cdot)$ dann natürlich selbst wieder ein Ring mit dem Einselement 1_R.

Die Beziehung (9.4a) besagt, dass $f: (R_1, +_1) \to (R_2, +_2)$ ein Gruppenhomomorphismus ist. Aus Lemma 8.5 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 (9.4b), dass $f: (R_1, \cdot_1) \to (R_2, \cdot_2)$ ein Halbgruppenhomomorphismus ist. (9.4b) und (9.4c) zusammen bedeuten, dass $f: (R_1, \cdot_1) \to (R_2, \cdot_2)$ ein Monoidhomomorphismus ist.

Beispiel 9.12 (Ringhomomorphismen)**.**

(*i*) Die Abbildung

$$f: (\mathbb{Z}, +, \cdot) \ni a \mapsto [a] = a + m\mathbb{Z} \in (\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$$

ist ein surjektiver Ringhomomorphismus zwischen zwei kommutativen Ringen mit Eins, denn: f ist als kanonische Surjektion der Faktorgruppe nach Satz 8.13 und Beispiel 8.15 ein surjektiver Gruppenhomomorphismus $f: (\mathbb{Z}, +) \to (\mathbb{Z}/m\mathbb{Z}, \tilde{+})$, und außerdem ist $f: (\mathbb{Z}, \cdot) \to (\mathbb{Z}/m\mathbb{Z}, \tilde{\cdot})$ ein Monoidhomomorphismus, siehe Beispiel 9.6 und Hausaufgabe I-6.1.

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

$$f: (\mathbb{Z}_m, +_m, \cdot_m) \ni a \mapsto [a] = a + m\mathbb{Z} \in (\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$$

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 8.15 ein Gruppenisomorphismus $f: (\mathbb{Z}_m, +_m) \to (\mathbb{Z}/m\mathbb{Z}, \tilde{+})$, und außerdem ist $f: (\mathbb{Z}_m, \cdot_m) \to (\mathbb{Z}/m\mathbb{Z}, \tilde{\cdot})$ nach Hausaufgabe I-6.1 ein Monoidisomorphismus.

Es gilt

$$\begin{aligned} \operatorname{Bild}(f) &= \mathbb{Z}/m\mathbb{Z}, \\ \operatorname{Kern}(f) &= f^{-1}([0]) = \{0\}. \end{aligned}$$

△

Ende der Vorlesung 12

Ende der Woche 6

§ 10 Körper

Literatur: Beutelspacher, 2014, Kapitel 2, Bosch, 2014, Kapitel 1.3, Fischer, Springborn, 2020, Kapitel 2.3, Deiser, 2022b, 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 $\cdot$.

Definition 10.1 (Körper)**.**
Ein **Körper** (englisch: field) $(K, +, \cdot)$ ist eine Menge K mit zwei (inneren) Verknüpfungen $+$ („Addition") und $\cdot$ („Multiplikation"), die die folgenden Bedingungen erfüllen:

(*i*) $(K,+)$ ist eine abelsche Gruppe. Das Nullelement bezeichnen wir mit 0_K.

(*ii*) $(K \setminus \{0_K\}, \cdot)$ ist eine abelsche Gruppe. Das Einselement bezeichnen wir mit 1_K.

(*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 $a, b, c \in K$.[13] △

Oft wird $K \setminus \{0_K\}$ als K^* oder als $K^\times$ abgekürzt. Wir verwenden diese Bezeichnungen jedoch hier nicht.

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 7.16 mit den Verknüpfungstafeln aus Beispiel 7.2 ist ein Körper mit dem Nullelement 0 und dem Einselement 1.

(*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 ist und $[2] \mathbin{\tilde{\cdot}} [a] \neq [1]$ für alle $a \in \mathbb{Z}$ gilt und damit $[2]$ kein multiplikatives Inverses besitzt.

(*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, Monoid, Gruppe oder Ring 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 10.1:** Woran liegt das?) △

Lemma 10.3 (Eigenschaften eines Körpers)**.**
Es sei $(K,+,\cdot)$ ein Körper mit dem Nullelement 0_K und dem Einselement 1_K. Dann gilt:

(*i*) $0_K \neq 1_K$. Ein Körper hat also mindestens zwei Elemente.

(*ii*) $(K,+,\cdot)$ ist ein kommutativer, nullteilerfreier Ring mit dem Einselement 1_K ungleich dem Nullring, also ein Integritätsring.

(*iii*) Es gelten die **Kürzungsregeln** (englisch: cancellation rules)

$$a \cdot b_1 = a \cdot b_2 \quad \Rightarrow \quad b_1 = b_2 \tag{10.2a}$$
$$b_1 \cdot a = b_2 \cdot a \quad \Rightarrow \quad b_1 = b_2 \tag{10.2b}$$

für $a, b_1, b_2 \in K$ mit $a \neq 0_K$.

[13] Wie bereits in kommutativen Ringen fallen die beiden Distributivgesetze (10.1a) und (10.1b) zusammen. Es reicht also, eines von beiden zu prüfen.

Beweis. Aussage (i): Nach Definition 10.1 ist $K \setminus \{0_K\}$ eine Gruppe mit dem Einselement 1_K, also muss $0_K \neq 1_K$ gelten.

Aussage (ii): Nach Definition 10.1 ist $(K, +)$ eine abelsche Gruppe mit dem Nullelement 0_K. Wenn wir zeigen können, dass $(K, \cdot)$ ein abelsches Monoid mit dem Einselement 1_K ist, dann ist $(K, +, \cdot)$ per Definition 9.1 ein abelscher Ring mit dem Einselement 1_K. Dieser ist nach Aussage (i) nicht der Nullring. Wenn wir anschließend zeigen können, dass $(K, +, \cdot)$ nullteilerfrei ist, dann ist Aussage (ii) gezeigt.

Schritt 1: Wir zeigen: $0_K \cdot a = 0_K = a \cdot 0_K$ für alle $a \in K$.

Dieses Ergebnis folgt wie im Beweis von Lemma 9.3, Aussage (i): $0_K + 0_K \cdot a = 0_K \cdot a = (0_K + 0_K) \cdot a = 0_K \cdot a + 0_K \cdot a$ und daher $0_K \cdot a = 0_K$. Analog können wir $a \cdot 0_K = 0_K$ zeigen.

Schritt 2: Wir zeigen: $\cdot$ ist eine *assoziative* Verknüpfung auf ganz K:

Per Definition ist $\cdot$ eine Verknüpfung auf ganz K. Die Assoziativität $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ ist klar, wenn nur Elemente $a, b, c \in K \setminus \{0_K\}$ verknüpft werden, da $(K \setminus \{0_K\}, \cdot)$ eine Gruppe ist. Wenn eines oder mehrere Elemente a, b, c aber das Nullelement 0_K sind, dann gilt wegen **Schritt** 1, dass sowohl $a \cdot (b \cdot c)$ als auch $(a \cdot b) \cdot c$ gleich 0_K sind. Die Assoziativität von $\cdot$ gilt also auf ganz K.

Schritt 3: Wir zeigen: $\cdot$ ist eine *kommutative* Verknüpfung auf ganz K:

Die Kommutativität $a \cdot b = b \cdot a$ ist klar, wenn nur Elemente $a, b \in K \setminus \{0_K\}$ verknüpft werden, da $(K \setminus \{0_K\}, \cdot)$ eine kommutative Gruppe ist. Wenn eines oder mehrere Element a, b aber das Nullelement 0_K sind, dann gilt wegen **Schritt** 1, dass sowohl $a \cdot b$ als auch $b \cdot a$ gleich 0_K sind. Die Kommutativität von $\cdot$ gilt also auf ganz K.

Schritt 4: Wir zeigen: 1_K ist neutrales Element bzgl. $\cdot$ auf ganz K:

Wir wissen bereits $1_K \cdot a = a \cdot 1_K$ für alle $a \in K \setminus \{0_K\}$. Ist nun $a = 0_K$, so gilt wegen **Schritt** 1 $1_K \cdot a = a \cdot 1_K = 0_K$. Damit ist 1_K neutrales Element bzgl. $\cdot$ auf ganz K.

Damit haben wir bisher gezeigt, dass $(K, \cdot)$ ein abelsches Monoid mit dem Einselement 1_K ist, also ist $(K, +, \cdot)$ per Definition 9.1 ein abelscher Ring mit dem Einselement 1_K, der Aussage (i) nicht der Nullring ist.

Schritt 5: Wir zeigen: Der Ring $(K, +, \cdot)$ ist nullteilerfrei.

Wenn $a, b \in K \setminus \{0_K\}$ sind, dann ist auch $a \cdot b \in K \setminus \{0_K\}$, da per Definition $(K \setminus \{0\}, \cdot)$ eine Gruppe und damit insbesondere $K \setminus \{0\}$ abgeschlossen bzgl. $\cdot$ ist. Das heißt, $(K, +, \cdot)$ ist nullteilerfrei.

Aussage (iii): Für $a, b_1, b_2 \in K \setminus \{0\}$ sind die Kürzungsregeln (10.2) nichts anderes als die Kürzungsregeln (7.10) in der Gruppe $(K \setminus \{0\}, \cdot)$. Wir zeigen (10.2a) in den verbleibenden Fällen. Ist $b_1 = 0_K$, dann folgt aus Aussage (i) und der Voraussetzung $0_K = a \cdot b_1 = a \cdot b_2$. Wegen der Nullteilerfreiheit und $a \neq 0_K$ folgt weiter $b_2 = 0_K$, also $b_1 = b_2$. Ist andererseits $b_2 = 0_K$, dann folgt aus Aussage (i) und der Voraussetzung $0_K = a \cdot b_2 = a \cdot b_1$. Wegen der Nullteilerfreiheit und $a \neq 0_K$ folgt weiter $b_1 = 0_K$, also wiederum $b_1 = b_2$.

Die Aussage (10.2b) können wir analog beweisen. □

Beachte: Die Rechenregeln in Ringen aus Lemma 9.3 gelten also auch in Körpern.

Die Definition 9.4 der Charakteristik eines Ringes wird auch auf Körper angewendet. Für Körper ist die Charakteristik entweder 0 oder eine Primzahl.

Satz 10.4 (Wann ist ein Ring ein Körper?)**.**
Es sei K eine Menge mit zwei (inneren) Verknüpfungen $+$ und $\cdot$. Dann sind die folgenden Aussagen äquivalent:

(*i*) $(K, +, \cdot)$ ist ein Körper, dessen Nullelement mit 0_K und dessen Einselement mit 1_K bezeichnet werden.

(*ii*) $(K, +, \cdot)$ ist ein kommutativer Ring mit dem Einselement 1_K und dem Nullelement $0_K \neq 1_K$, wobei zu jedem $a \in K \setminus \{0_K\}$ ein Inverses bzgl. $\cdot$ in K existiert.

Beweis. Aussage $(i) \Rightarrow$ Aussage (ii): Nach Lemma 10.3 ist $(K, +, \cdot)$ ein kommutativer Ring mit dem Einselement 1_K und dem Nullelement $0_K \neq 1_K$. Da $(K \setminus \{0_K\}, \cdot)$ eine Gruppe ist, existiert zu jedem $a \in K \setminus \{0_K\}$ ein Inverses bzgl. $\cdot$.

Aussage $(ii) \Rightarrow$ Aussage (i): Nach Voraussetzung ist $(K, +)$ eine abelsche Gruppe mit dem Nullelement 0_K. Weiter gelten die Distributivgesetze (10.1) nach Voraussetzung. Es bleibt zu zeigen, dass $(K \setminus \{0_K\}, \cdot)$ eine abelsche Gruppe mit dem Einselement 1_K ist.

Nach Voraussetzung ist $(K, \cdot)$ ein abelsches Monoid mit neutralem Element 1_K. Aus der Eigenschaft $\forall a \in K \setminus \{0\}\ \exists a^{-1} \in K\ (a \cdot a^{-1} = a^{-1} \cdot a = 1_K)$ folgt die Nullteilerfreiheit:

$$a \cdot b = 0_K \quad \wedge \quad a \neq 0 \quad \Rightarrow \quad b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0_K = 0_K.$$

Also ist $K \setminus \{0_K\}$ bzgl. $\cdot$ abgeschlossen. Damit erbt $(K \setminus \{0_K\}, \cdot)$ die Eigenschaft, ein abelsches Monoid mit dem Einselement 1_K zu sein, von $(K, \cdot)$. Da jedes Element $a \in K \setminus \{0_K\}$ nach Voraussetzung ein Inverses a^{-1} bzgl. $\cdot$ mit $a^{-1} \in K$ besitzt und a^{-1} wegen $a \cdot 0_K = 0_K \neq 1_K$ sogar in $K \setminus \{0_K\}$ liegen muss, ist $(K \setminus \{0_K\}, \cdot)$ als abelsche Gruppe bestätigt. Das heißt, $(K, +, \cdot)$ ist ein Körper. □

Satz 10.5 (endliche Integritätsringe sind Körper)**.**
Es sei $(R, +, \cdot)$ ein Integritätsring mit endlich vielen Elementen. Dann ist $(R, +, \cdot)$ ein Körper.

Beweis. Es sei $(R, +, \cdot)$ ein Integritätsring, also ein kommutativer, nullteilerfreier Ring mit dem Einselement 1_R ungleich dem 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 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 7.18. 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 6.27, dass $\cdot_a$ auch surjektiv.

Ein analoges Argument zeigt, dass auch alle Linkstranslationen auf $R \setminus \{0\}$ surjektiv sind. Aus dem Gruppenkriterium Lemma 7.18 folgt nun, dass $(R \setminus \{0\}, \cdot)$ eine Gruppe ist. □

Folgerung 10.6 (Körpereigenschaft des Restklassenringes und des Ringes von $\mathbb{Z}$ modulo m, vgl. Satz 9.9).
Der Restklassenring modulo m $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ sowie der zu ihm isomorphe Ring (Beispiel 9.12) 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 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 9.9 haben wir gezeigt, dass $(\mathbb{Z}/m\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ genau dann ein Integritätsbereich ist, wenn $m \in \mathbb{N}$ eine Primzahl ist. Da aber $\mathbb{Z}/m\mathbb{Z}$ nur endlich viele (nämlich m) Elemente hat, ist Integritätsbereich zu sein gleichbedeutend mit der Körpereigenschaft.

Nach Beispiel 9.12 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 $(\mathbb{Z}_m, +_m, \cdot_m)$. □

Definition 10.7 (Unterkörper, vgl. Definition 9.10 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 ist und wenn $(U, +, \cdot)$ selbst wieder ein Körper ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, +)$ eine Untergruppe von $(K, +)$ ist 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. △

Beispiel 10.8 (Unterkörper).

(i) $(\mathbb{Q}, +, \cdot)$ ist ein Unterkörper von $(\mathbb{R}, +, \cdot)$.

(ii) $(\mathbb{R}, +, \cdot)$ ist ein Unterkörper von $(\mathbb{C}, +, \cdot)$. △

Definition 10.9 (Körperhomomorphismus, vgl. Definition 9.11 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 **(Körper-)Homomorphismus** (englisch: field homomorphism) von $(K_1, +_1, \cdot_1)$ in $(K_2, +_2, \cdot_2)$, wenn gilt:

$$f(a +_1 b) = f(a) +_2 f(b) \quad \text{für alle } a, b \in K_1, \tag{10.3a}$$
$$f(a \cdot_1 b) = f(a) \cdot_2 f(b) \quad \text{für alle } a, b \in K_1, \tag{10.3b}$$
$$f(1_{K_1}) = 1_{K_2}. \tag{10.3c}$$

(ii) Wie in Definition 8.1 sprechen wir im Fall $(K_1, +_1, \cdot_1) = (K_2, +_2, \cdot_2)$ von einem **(Körper-)Endomorphismus** (englisch: field endomorphism).

(iii) Ist $f\colon K_1 \to K_2$ bijektiv, so heißt f auch **strukturerhaltend** oder ein **(Körper-)Isomorphismus** (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) Im Fall $(K_1, +_1, \cdot_1) = (K_2, +_2, \cdot_2)$ und $f\colon K_1 \to K_2$ bijektiv sprechen wir auch von einem **(Körper-) Automorphismus** (englisch: field automorphism). △

Da die Bedingungen (10.3) mit denen aus (9.4) übereinstimmen, ist ein Körperhomomorphismus nichts anderes als ein Ringhomomorphismus, der speziell zwischen Körpern eingesetzt wird. Insbesondere haben wir auch hier wie in (9.6) wieder

$$f(0_{K_1}) = 0_{K_2}. \tag{10.4}$$

Interessanterweise gilt weiter, dass Körperhomomorphismen automatisch injektiv sind, denn nehmen wir $a \neq b$, aber $f(a) = f(b)$ an, so ergibt sich der Widerspruch

$$\begin{aligned}
1_{K_2} &= f(1_{K_1}) && \text{wegen (10.3c)}\\
&= f\big((a -_1 b)^{-1} \cdot_1 (a -_1 b)\big) && \text{da } a -_1 b \neq 0_{K_1} \text{ vorausgesetzt wurde}\\
&= f\big((a -_1 b)^{-1}\big) \cdot_2 f(a -_1 b) && \text{wegen (10.3a)}\\
&= f\big((a -_1 b)^{-1}\big) \cdot_2 \big(f(a) -_2 f(b)\big) && \text{wegen (10.3b)}\\
&= f\big((a -_1 b)^{-1}\big) \cdot_2 0_{K_2} && \text{da } f(a) = f(b) \text{ vorausgesetzt wurde}\\
&= 0_{K_2} && \text{nach Lemma 9.3.}
\end{aligned}$$

§ 11 Polynome

Literatur: Beutelspacher, 2014, Kapitel 6, Bosch, 2014, Kapitel 5, Fischer, Springborn, 2020, Kapitel 2.3, Jänich, 2008, Kapitel 9.4

Definition 11.1 (Polynom).
Es sei $(R, +, \cdot)$ ein kommutativer Ring.[14] Ein **Polynom** (englisch: polynomial, altgriechisch: *πολύ*: viel, altgriechisch: *όνομα*: Name) über R in der Variablen t ist ein formaler Ausdruck der Gestalt

$$a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \cdots + a_1 \cdot t + a_0 \quad \text{oder auch} \quad \sum_{i=0}^{n} a_i \cdot t^i. \tag{11.1}$$

Dabei ist $n \in \mathbb{N}_0$. Die Zahlen $a_i \in R$ heißen die **Koeffizienten** (englisch: coefficients) des Polynoms. △

Bemerkung 11.2 (Polynom).

(*i*) Die Variable t ist ein willkürlich gewähltes Symbol. Später werden wir für t geeignete Objekte einsetzen. Da zunächst unspezifiert ist, welche Objekte für die Variable t eingesetzt werden können, ist die Bedeutung der „Potenzen“ t^j, deren „Multiplikation“ mit den Koeffizienten $a_j \in R$ sowie die „Addition“ der daraus entstehenden Terme im Moment unklar. Daher verstehen wir (11.1) zunächst als formalen Ausdruck.

(*ii*) $a_1 \cdot t$ ist eine abkürzende Schreibweise für $a_1 \cdot t^1$, und a_0 ist eine abkürzende Schreibweise für $a_0 \cdot t^0$.

(*iii*) Die Reihenfolge der „Summanden“ in (11.1) ist unerheblich. Die Polynome $3 \cdot t^2 + 2 \cdot t$ und $2 \cdot t + 3 \cdot t^2$ werden also miteinander identifiziert.

(*iv*) Ist ein Koeffizient $a_i = 0_R \in R$, so lässt man häufig den entsprechenden „Summanden“ in der Darstellung (11.1) einfach weg. Die Polynome $3 \cdot t^2 + 0 \cdot t$ und $3 \cdot t^2$ werden also miteinander identifiziert.

[14] Tatsächlich ist R oft sogar ein Körper.

(*v*) Ist ein Koeffizient $a_i = 1_R$ in einem Ring mit dem Einselement 1_R, so lässt man den Koeffizienten 1_R in der Darstellung (11.1) manchmal weg, sofern es sich nicht um a_0 handelt. Die Polynome $1_R \cdot t^2$ und t^2 werden also miteinander identifiziert.

(*vi*) Die Menge aller Polynome über dem kommutativen Ring R in der Variablen t wird mit $R[t]$ bezeichnet.

(*vii*) Zwei Polynome sind gleich, wenn die entsprechenden Koeffizienten gleich sind.

(*viii*) Sind alle Koeffizienten gleich $0_R \in R$, so heißt das Polynom das **Nullpolynom** (englisch: zero polynomial), geschrieben 0_R.

(*ix*) Ist R ein Ring mit dem Einselement 1_R und sind alle Koeffizienten gleich $0_R \in R$ bis auf $a_0 = 1_R \in R$, so heißt das Polynom das **Einspolynom** (englisch: constant one polynomial), geschrieben 1_R.

(*x*) Ist R ein Ring mit dem Einselement 1_R, dann heißt ein Polynom der Form t^n das **Monom** (englisch: monomial) **vom Grad** $n \in \mathbb{N}_0$.[15]

(*xi*) Ein Polynom der Form $a_0 \in R \subseteq R[t]$ heißt ein **konstantes Polynom** (englisch: constant polynomial).

(*xii*) Ein Polynom der Form $a_0 + a_1 \cdot t \in R[t]$ mit $a_1 \neq 0_R$ heißt ein **lineares Polynom** (englisch: linear polynomial). △

Beispiel 11.3 (Polynome)**.**

(*i*) $p = \frac{3}{4} \cdot t^2 - 7 \cdot t + \frac{1}{2} = \frac{1}{2} + \frac{3}{4} \cdot t^2 - 7 \cdot t \in \mathbb{Q}[t]$ ist ein Polynom über dem Körper $\mathbb{Q}$.

(*ii*) $p = s^5 - \frac{\sqrt{2}}{3} \cdot s + \frac{2}{5} \in \mathbb{R}[s]$ ist ein Polynom über dem Körper $\mathbb{R}$.

(*iii*) $p = [1] \mathbin{\tilde{\cdot}} X^3 \mathbin{\tilde{+}} [3] \mathbin{\tilde{\cdot}} X^2 \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} X \in (\mathbb{Z}/4\mathbb{Z})[X]$ ist ein Polynom über dem Restklassenring $\mathbb{Z}/4\mathbb{Z}$. △

Wir definieren nun zwei Verknüpfungen + („Addition“) und · („Multiplikation“) auf der Menge $R[t]$ der Polynome über dem kommutativen Ring $(R, +, \cdot)$.[16] Es seien $p, q \in R[t]$ gegeben mit den Darstellungen

$$p = a_m \cdot t^m + \cdots + a_1 \cdot t + a_0 = \sum_{i=0}^{m} a_i \cdot t^i \tag{11.2a}$$

$$q = b_n \cdot t^n \quad + \cdots + b_1 \cdot t + b_0 = \sum_{i=0}^{n} b_i \cdot t^i \tag{11.2b}$$

und $m, n \in \mathbb{N}_0$. Zur Abkürzung setzen wir außerdem $N \coloneqq \max\{m, n\}$ und füllen in (11.2) nicht notierte Terme mit Nullkoeffizienten auf. Dann definieren wir die **Addition von Polynomen** (englisch: addition of polynomials)

$$p + q \coloneqq (a_N + b_N) \cdot t^N + \cdots + (a_1 + b_1) \cdot t + (a_0 + b_0) = \sum_{i=0}^{\max\{m,n\}} (a_i + b_i) \cdot t^i \tag{11.3a}$$

[15] Der Begriff des Grades für beliebige Polynome wird in Definition 11.8 eingeführt.

[16] Es ist Absicht, dass die Verknüpfungen in $R[t]$ genauso benannt werden wie die Verknüpfungen im Koeffizientenring R. Dadurch wird $(R, +, \cdot)$ zu einem Unterring von $(R[t], +, \cdot)$, nämlich dem Unterring der konstanten Polynome.

sowie die **Multiplikation von Polynomen** (englisch: multiplication of polynomials)

$$p \cdot q := c_{m+n} \cdot t^{m+n} + \cdots + c_1 \cdot t + c_0 \qquad = \sum_{k=0}^{m+n} c_k \cdot t^k, \tag{11.3b}$$

wobei c_k für $k \in \mathbb{N}_0$ als

$$c_k := \sum_{i=0}^{k} a_i \cdot b_{k-i} = \sum_{\substack{i,j=0 \\ i+j=k}}^{k} a_i \cdot b_j \tag{11.3c}$$

gesetzt wird. Man nennt die aus (11.3c) entstehende Folge $(c_k)_{k\in\mathbb{N}_0}$, also

$$\begin{aligned} c_0 &= a_0 \cdot b_0 \\ c_1 &= a_0 \cdot b_1 + a_1 \cdot b_0 \\ c_2 &= a_0 \cdot b_2 + a_1 \cdot b_1 + a_2 \cdot b_0 \quad \text{usw.} \end{aligned}$$

die **Faltung** (englisch: convolution, lateinisch: convolvere: zusammenrollen) der Folgen $(a_i)_{i\in\mathbb{N}_0}$ und $(b_i)_{i\in\mathbb{N}_0}$.

Quizfrage 11.1: Wie kann man sich die Faltung der Koeffizientenfolgen $(a_i)_{i\in\mathbb{N}_0}$ und $(b_i)_{i\in\mathbb{N}_0}$ grafisch vorstellen?

Definition 11.4 (Polynomring)**.**
Es sei $(R, +, \cdot)$ ein kommutativer Ring. Mit den zwei Verknüpfungen (11.3) wird $(R[t], +, \cdot)$ zu einem kommutativen Ring, genannt der **Polynomring** (englisch: polynomial ring) **über** R in der Variablen t. R heißt der **Koeffizientenring** (englisch: coefficient ring, ring of coefficients) von $R[t]$. △

Das Nullelement von $R[t]$ ist das Nullpolynom 0_R. Besitzt R das Einselement 1_R, dann ist $(R[t], +, \cdot)$ ebenfalls ein Ring mit Einselement 1_R, dem Einspolynom. Das Symbol $+$ aus (11.1) ist die gleichnamige Verknüpfung (11.3a) aus dem Polynomring. Das Symbol $\cdot$ aus (11.1) ist die gleichnamige Verknüpfung (11.3b) aus dem Polynomring, wobei ein Faktor ein Polynom der Form $a_i \in R$ ist und der andere Faktor ein Monom.

Bemerkung 11.5 (Polynomring als Erweiterung von $(R, +, \cdot)$)**.**
Wenn R ein Ring mit Eins, aber nicht der Nullring ist, so können wir den Polynomring $(R[t], +, \cdot)$ algebraisch auch verstehen als die (bis auf Ring-Isomorphie eindeutige) kleinstmögliche Erweiterung des kommutativen Ringes von $(R, +, \cdot)$ zu einem kommutativen Ring, der zusätzlich das freie Element $t \notin R$ enthält. Dadurch wird $(R, +, \cdot)$ zu ein Unterring von $(R[t], +, \cdot)$, dem Unterring der konstanten Polynome. △

Quizfrage 11.2: Wie sieht der Polynomring $R[t]$ aus, wenn R der Nullring ist?

Bemerkung 11.6 (Polynomring als Folgenring)**.**
Wir können ein Polynom identifizieren mit der Folge $\mathbb{N}_0 \to R$ (vgl. Definition 6.29) seiner in aufsteigender Reihenfolge der „Potenzen" sortierten Koeffizienten, von denen nur endlich viele ungleich $0_R \in R$ sind. Man sagt, eine solche Folge habe **endlichen Träger**[17] (englisch: finite support). Diese Teilmenge von $R^{\mathbb{N}_0}$ notieren wir in dieser Lehrveranstaltung als $(R^{\mathbb{N}_0})_{00}$.

[17]Der **Träger einer Folge** (englisch: support of a sequence) mit Werten in einem Ring (oder allgemeiner mit Werten in einer additiven Gruppe) ist die Menge derjenigen Indizes, deren Folgenglieder ungleich dem Nullelement sind.

Beispielsweise kann das Polynom $t - t^2 + 3 \in \mathbb{Z}[t]$ identifiziert werden mit der endlich getragenen Folge $(3, 1, -1, 0, 0, \ldots)$ seiner Koeffizienten in $\mathbb{Z}$. Das Polynom p aus (11.2a) wird identifiziert mit der endlich getragenen Folge $(a_0, a_1, \ldots, a_m, 0, 0, \ldots)$. Statten wir dann $(R^{\mathbb{N}_0})_{00}$ mit der elementweisen Addition + der Folgenelemente und der Faltung als Multiplikation $c = a \cdot b$ wie in (11.3c) aus, so wird $((R^{\mathbb{N}_0})_{00}, +, \cdot)$ ebenfalls zu einem kommutativen Ring, der zu $R[t]$ isomorph ist. △

Ende der Vorlesung 13

Beispiel 11.7 (Addition und Multiplikation von Polynomen).

(*i*) Für die Polynome

$$p = \frac{3}{4} \cdot t^2 - 7 \cdot t + \frac{1}{2}$$
$$q = -\frac{1}{2} \cdot t^3 - t + 1$$

über dem Körper $(\mathbb{Q}, +, \cdot)$ gilt

$$\begin{aligned}
p + q &= \left(0 - \frac{1}{2}\right) \cdot t^3 + \left(\frac{3}{4} + 0\right) \cdot t^2 + \big(-7 + (-1)\big) \cdot t + \left(\frac{1}{2} + 1\right) \\
&= -\frac{1}{2} \cdot t^3 + \frac{3}{4} \cdot t^2 - 8 \cdot t + \frac{3}{2} \\
p \cdot q &= \left(0 + 0 + \frac{3}{4} \cdot (-\frac{1}{2}) + 0 + 0 + 0\right) \cdot t^5 + \left(0 + (-7) \cdot (-\frac{1}{2}) + 0 + 0 + 0\right) \cdot t^4 \\
&\quad + \left(\frac{1}{2} \cdot (-\frac{1}{2}) + 0 + \frac{3}{4} \cdot (-1) + 0\right) \cdot t^3 + \left(0 + (-7) \cdot (-1) + \frac{3}{4} \cdot 1\right) \cdot t^2 \\
&\quad + \left(\frac{1}{2} \cdot (-1) + (-7) \cdot 1\right) \cdot t + \left(\frac{1}{2} \cdot 1\right) \\
&= -\frac{3}{8} \cdot t^5 + \frac{7}{2} \cdot t^4 - t^3 + \frac{31}{4} \cdot t^2 - \frac{15}{2} \cdot t + \frac{1}{2}.
\end{aligned}$$

Berechnung der Koeffizienten des Produkts $p \cdot q$ mittels Faltungstabelle:

	$\frac{1}{2}$	-7	$\frac{3}{4}$	0
1	$\frac{1}{2}$	-7	$\frac{3}{4}$	0
-1	$-\frac{1}{2}$	7	$-\frac{3}{4}$	0
0	0	0	0	0
$-\frac{1}{2}$	$-\frac{1}{4}$	$\frac{7}{2}$	$-\frac{3}{8}$	0

Die Summation entlang der Diagonalen ergibt wiederum die Koeffizienten

$$c_0 = \frac{1}{2}, \qquad c_1 = -\frac{1}{2} - 7 = -\frac{15}{2}, \qquad c_2 = 0 + 7 + \frac{3}{4} = \frac{31}{4},$$
$$c_3 = -\frac{1}{4} + 0 - \frac{3}{4} + 0 = -1, \qquad c_4 = \frac{7}{2} + 0 + 0 = \frac{7}{2}, \qquad c_5 = -\frac{3}{8} + 0 = -\frac{3}{8}.$$

(*ii*) Für die Polynome

$$p = [1] \mathbin{\tilde{\cdot}} X^3 \mathbin{\tilde{+}} [-3] \mathbin{\tilde{\cdot}} X^2 \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} X \qquad \text{oder auch } p = X^3 \mathbin{\tilde{+}} X^2 \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} X$$
$$q = [-1] \mathbin{\tilde{\cdot}} X \mathbin{\tilde{+}} [7] \qquad \text{oder auch } q = [3] \mathbin{\tilde{\cdot}} X \mathbin{\tilde{+}} [3]$$

über dem Restklassenring $\mathbb{Z}/4\mathbb{Z}$ gilt

$$\begin{aligned}(p \mathbin{\tilde{+}} q)(X) &= [1] \mathbin{\tilde{\cdot}} X^3 \mathbin{\tilde{+}} [1] \mathbin{\tilde{\cdot}} X^2 \mathbin{\tilde{+}} \big([2] \mathbin{\tilde{+}} [-1]\big) \mathbin{\tilde{\cdot}} X \mathbin{\tilde{+}} [7] \\ &= [1] \mathbin{\tilde{\cdot}} X^3 \mathbin{\tilde{+}} [1] \mathbin{\tilde{\cdot}} X^2 \mathbin{\tilde{+}} [1] \mathbin{\tilde{\cdot}} X \mathbin{\tilde{+}} [3],\end{aligned}$$

$$\begin{aligned}(p \mathbin{\tilde{\cdot}} q)(X) &= [1] \mathbin{\tilde{\cdot}} [-1] \mathbin{\tilde{\cdot}} X^4 \mathbin{\tilde{+}} \big([-3] \mathbin{\tilde{\cdot}} [-1] \mathbin{\tilde{+}} [1] \mathbin{\tilde{\cdot}} [7]\big) \mathbin{\tilde{\cdot}} X^3 \\ &\qquad \mathbin{\tilde{+}} \big([-3] \mathbin{\tilde{\cdot}} [7] \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} [-1]\big) \mathbin{\tilde{\cdot}} X^2 \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} [7] \mathbin{\tilde{\cdot}} X \\ &= [3] \mathbin{\tilde{\cdot}} X^4 \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} X^3 \mathbin{\tilde{+}} [1] \mathbin{\tilde{\cdot}} X^2 \mathbin{\tilde{+}} [2] \mathbin{\tilde{\cdot}} X.\end{aligned}$$

△

Zur Vereinfachung der Notation lassen wir in Zukunft das Multiplikationszeichen zwischen Koeffizient und Potenz einer Variable auch oft weg.

Definition 11.8 (Grad eines Polynoms, führender Koeffizient, monisches Polynom)**.**
Es sei p ein Polynom über dem kommutativen Ring R mit den Koeffizienten $a_j \in R$, $j \in \mathbb{N}_0$.

(i) Der **Grad** (englisch: degree) ist definiert als[18]

$$\deg(p) := \begin{cases} -\infty, & \text{falls alle } a_j = 0_R \text{ sind, also } p = 0_R, \\ \max\{j \in \mathbb{N}_0 \mid a_j \neq 0_R\} & \text{sonst.} \end{cases} \tag{11.4}$$

Ein Polynom vom Grad 0 oder $-\infty$ heißt **konstant** (englisch: constant).

(ii) Wenn $p \neq 0_R$ (also nicht das Nullpolynom) ist, dann heißt $\ell(p) := a_{\deg(p)}$ auch der **führende Koeffizient** (englisch: leading coefficient) oder der **Leitkoeffizient** von p. Für $p = 0_R$ definieren wir $\ell(0_R) = 0_R$.

(iii) Ist R ein Ring mit dem Einselement 1_R und gilt $\ell(p) = 1_R$, dann heißt das Polynom $p \neq 0_R$ **normiert** oder **monisch** (englisch: monic). △

Quizfrage 11.3: Was weiß man über das Produkt zweier monischer Polynome?

Beispiel 11.9 (Grad eines Polynoms, führender Koeffizient, monisches Polynom)**.**

(i) Das Polynom $p = \frac{3}{4}\, t^2 - 7\, t + \frac{1}{2} \in \mathbb{Q}[t]$ über dem Körper $\mathbb{Q}$ besitzt $\deg(p) = 2$. Das Polynom p ist nicht monisch, da $\ell(p) = \frac{3}{4} \neq 1 \in \mathbb{Q}$ ist.

(ii) Das Polynom $p = [-7]\, X^3 \mathbin{\tilde{+}} [-3]\, X^2 \mathbin{\tilde{+}} [2]\, X^2 \in (\mathbb{Z}/4\mathbb{Z})[X]$ über dem Restklassenring $\mathbb{Z}/4\mathbb{Z}$ besitzt $\deg(3)$. Das Polynom p ist monisch, da $\ell(p) = [-7] = [1] \in \mathbb{Z}/4\mathbb{Z}$ ist. △

Lemma 11.10 (Grad eines Polynoms)**.**
Es sei R ein kommutativer Ring und $p, q \in R[t]$ zwei Polynome. Dann gilt:

(i) $\deg(p + q) \leqslant \max\{\deg(p),\ \deg(q)\}$.

(ii) $\deg(p \cdot q) \leqslant \deg(p) + \deg(q)$.

(iii) Ist R nullteilerfrei, dann gilt sogar $\deg(p \cdot q) = \deg(p) + \deg(q)$.

Dabei sollen formal für $n \in \mathbb{N}_0$ die Beziehungen $\max\{n,\ (-\infty)\} = \max\{(-\infty), n\} = n$ gelten sowie $\max\{(-\infty),\ (-\infty)\} = -\infty$ und $n + (-\infty) = (-\infty) + n = (-\infty) + (-\infty) = -\infty$.

[18] Manchmal wird der Grad des Nullpolynoms abweichend auch als -1 definiert.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-7.3. □

Folgerung 11.11 (der Polynomring als Integritätsring).
Es sei R ein Integritätsring. Dann ist auch $R[t]$ ein Integritätsring.

Beweis. Nach Definition ist R ein kommutativer, nullteilerfreier Ring mit Eins 1_R, und R ist ungleich dem Nullring. Folglich ist $R[t]$ ein kommutativer Ring mit Eins 1_R (Einspolynom) ungleich dem Nullring, da $1_R \neq 0_R$ (Nullpolynom) gilt. Es bleibt zu zeigen, dass $R[t]$ nullteilerfrei ist. Dazu seien $p, q \in R[t]$ beide nicht das Nullpolynom. Aus Lemma 11.10 Aussage *(iii)* folgt $\deg(p \cdot q) = \deg(p) + \deg(q) \geqslant 0$. Damit ist auch $p \cdot q$ nicht das Nullpolynom. □

§ 11.1 Polynomdivision

Lemma 11.12 (Polynomring ist kein Körper).
Der Polynomring $R[t]$ ist niemals ein Körper.

Beweis. Wenn R der Nullring ist, dann ist auch $R[t]$ der Nullring, besitzt also nur ein Element und ist daher kein Körper (Lemma 10.3). Wir betrachten also im Weiteren nur den Fall, dass R nicht der Nullring ist. Wenn $R[t]$ ein Körper wäre, dann wäre 1_R das neutrale Element bzgl. der Multiplikation in $R[t]$ und damit auch in R. Daraus folgt, dass das Polynom $p = 0_R + 1_R \cdot t$ in $R[t]$ existiert. Dieses besitzt aber kein multiplikatives Inverses, denn: Wäre q das Inverse zu p, gälte also $p \cdot q = 1_R$, dann kann q nicht das Nullpolynom sein. Es müsste also $q = b_n\, t^n + \cdots + b_1\, t + b_0$ gelten für irgendwelche Koeffizienten $b_0, b_1, \ldots, b_n$, $n \in \mathbb{N}_0$. Der 0-te Koeffizient von $p \cdot q$ ist aber $0_R \cdot b_0 = 0_R$, und daher kann $p \cdot q$ nicht das Einspolynom sein. □

Als Ersatz für das Fehlen multiplikativer Inverser führen wir (wie bereits aus $\mathbb{Z}$ bekannt) eine **Division mit Rest** (englisch: division with remainder) von Polynomen ein. Wir arbeiten dabei für den Rest von § 11.1 mit einem Körper K für die Koeffizienten an Stelle eines kommutativen Ringes R.

Definition 11.13 (Teiler eines Polynoms).
Es seien K ein Körper und $p_1, p_2 \in K[t]$ zwei Polynome. p_2 heißt ein **Teiler** (englisch: divisor) von p_1 (kurz: $p_2 \mid p_1$), wenn es ein weiteres Polynom $q \in K[t]$ gibt, sodass gilt:

$$p_1 = q \cdot p_2. \tag{11.5}$$

△

Satz 11.14 (Polynomdivision mit Rest).
Es seien K ein Körper und $p_1, p_2 \in K[t]$ zwei Polynome. Ist $p_2 \neq 0_K$ (Nullpolynom), dann gibt es eindeutig bestimmte Polynome $q, r \in K[t]$, genannt der **Quotient** (englisch: quotient) und der **Rest** (englisch: remainder), sodass gilt:

$$p_1 = q \cdot p_2 + r \quad \text{und } \deg(r) < \deg(p_2). \tag{11.6}$$

Beweis. Wir zeigen zunächst die Existenz der Zerlegung (11.6). Dazu sei $p_2 \in K[t]$ fest und $\deg(p_2) = m \in \mathbb{N}_0$. Wir verwenden vollständige Induktion nach $n := \deg(p_1) \in \mathbb{N}_0 \cup \{-\infty\}$.

Induktionsanfang: Für jedes $n \in \{-\infty\} \cup [\![0, m-1]\!]$ (also wenn $\deg(p_1) < \deg(p_2)$ gilt) setzen wir $q := 0_K$ und $r := p_1$.

Induktionsschritt: Es sei $n \geqslant m$ und die Behauptung für $n-1$ bereits bewiesen. Es sei nun $n = \deg(p_1) \geqslant \deg(p_2)$. Die Darstellungen von p und q seien

$$\begin{aligned} p_1 &= a_n\, t^n \;\; + a_{n-1}\, t^{n-1} \;\; + \cdots + a_1\, t + a_0, \\ p_2 &= b_m\, t^m + b_{m-1}\, t^{m-1} + \cdots + b_1\, t + b_0. \end{aligned}$$

Wir definieren $\widehat{p}_1 := p_1 - a_n\, b_m^{-1}\, t^{n-m} p_2$. Dann ist der Koeffizient von t^n in $\widehat{p}_1$ gleich 0_K. Damit gilt $\deg(\widehat{p}_1) < \deg(p_1)$. Nach Induktionsvoraussetzung existieren also $\widehat{q}, \widehat{r} \in K[t]$ mit der Eigenschaft $\widehat{p}_1 = \widehat{q}_1 \cdot p_2 + \widehat{r}$ und $\deg(\widehat{r}) < \deg(p_2) = m$. Es folgt nun

$$\begin{aligned} p_1 &= \widehat{p}_1 + a_n\, b_m^{-1}\, t^{n-m} p_2 && \text{nach Definition von } \widehat{p}_1 \\ &= \widehat{q}_1 \cdot p_2 + \widehat{r} + a_n\, b_m^{-1}\, t^{n-m} p_2 && \text{gemäß der Zerlegung von } \widehat{p}_1 \\ &= \underbrace{\big(\widehat{q}_1 + a_n\, b_m^{-1}\, t^{n-m}\big)}_{=:q} \cdot p_2 + \underbrace{\widehat{r}}_{=:r} && \text{wegen der Kommutativität und Distributivität im Ring } K[t]. \end{aligned}$$

Dabei gilt $\deg(r) = \deg(\widehat{r}) < \deg(p_2) = m$.

Es bleibt, die Eindeutigkeit der Zerlegung zu bestätigen. Angenommen, es gelte

$$p_1 = q \cdot p_2 + r = \widehat{q} \cdot p_2 + \widehat{r}$$

mit $\deg(r) < \deg(p_2)$ und $\deg(\widehat{r}) < \deg(p_2)$. Dann folgt $(q - \widehat{q}) \cdot p_2 = \widehat{r} - r$ und weiter

$$\begin{aligned} \deg(p_2) &> \deg(r - \widehat{r}) && \text{da } \deg(r - \widehat{r}) \leqslant \max\{\deg(r),\ \deg(-\widehat{r})\} \text{ nach Lemma 11.10} \\ &= \deg((q - \widehat{q}) \cdot p_2) \\ &= \deg(q - \widehat{q}) + \deg(p_2) && \text{nach Lemma 11.10, da } K \text{ als Körper nullteilerfrei ist.} \end{aligned}$$

Deshalb gilt $\deg(q - \widehat{q}) < 0$, woraus $q - \widehat{q} = 0_K$ folgt, also $q = \widehat{q}$. Aus $q \cdot p_2 + r = \widehat{q} \cdot p_2 + \widehat{r}$ folgt dann auch $r = \widehat{r}$. □

Folgerung 11.15 (Polynomdivision mit Rest)**.**
Unter den Voraussetzungen von Definition 11.13 ist q in (11.5) eindeutig bestimmt.

Beispiel 11.16 (Polynomdivision mit Rest)**.**
Die **Polynomdivision** (englisch: polynomial long division) ist ein Verfahren zur Berechnung der Zerlegung (11.6) für zwei gegebene Polynome $p_1, p_2 \in K[t]$. Man sortiert dazu p_1 und p_2 nach absteigenden Potenzen der Variablen und führt dieselben Schritte wie bei einer schriftlichen Division etwa in $\mathbb{Z}$ durch. Sobald der Grad des aktuellen Restes echt kleiner ist als der Grad von p_2, stoppt das Verfahren.

Für $p_1 = 3t^3 + 2t + 1$ und $p_2 = t^2 - 4t$ erhalten wir

$$\begin{array}{rrrl} 3t^3 & & +\,2t + 1 & = (t^2 - 4t)(3t + 12) + 50t + 1 \\ -\,3t^3 & +\,12t^2 & & \\ \hline & 12t^2 & +\,2t & \\ & -\,12t^2 & +\,48t & \\ \hline & & 50t + 1 & \end{array}$$

Also gilt in diesem Beispiel

$$\underbrace{3t^3 + 2t + 1}_{p_1} = \underbrace{(3t + 12)}_{q} \cdot \underbrace{(t^2 - 4t)}_{p_2} + \underbrace{(50t + 1)}_{r}. \qquad \triangle$$

§ 11.2 Polynomfunktionen

Wir gehen nun der Frage nach, welche Objekte man für die Variable t in einem Polynom

$$p = a_n\, t^n + a_{n-1}\, t^{n-1} + \cdots + a_1\, t + a_0 \tag{11.7}$$

über einem kommutativen Ring $(R, +, \cdot)$ sinnvollerweise einsetzen kann. Die naheliegendste Wahl sind sicherlich Elemente aus R selbst, und nur diese lassen wir im Moment zu.

Genauer betrachtet induziert das Polynom p eine Funktion $\widetilde{p}\colon R \to R$, definiert durch

$$\widetilde{p}(r) \coloneqq a_n\, r^n + a_{n-1}\, r^{n-1} + \cdots + a_1\, r + a_0. \tag{11.8}$$

Die Funktion $\widetilde{p}$ heißt die **Polynomfunktion** (englisch: polynomial function) zum Polynom p oder die vom Polynom p induzierte Polynomfunktion.

Bemerkung 11.17 (induzierte Polynomfunktion)**.**
Die Menge R^R der Funktionen $R \to R$, ausgestattet mit den punktweisen Verknüpfungen $+$ und $\cdot$ aus R, bildet einen kommutativen Ring $(R^R, +, \cdot)$. Die Abbildung

$$\Phi\colon (R[t], +, \cdot) \ni p \mapsto \widetilde{p} \in (R^R, +, \cdot) \tag{11.9}$$

ist ein Ringhomomorphismus zwischen zwei kommutativen Ringen. Wenn R das Einselement 1_R besitzt, dann besitzt $R[t]$ das Einselement 1_R (das Einspolynom) und R^R das Einselement 1_R (die Einsabbildung $R \mapsto 1_R$), und es gilt $\Phi(1_R) = 1_R$.

Φ ist i. A. nicht injektiv. Verschiedene Polynome können also dieselbe Polynomfunktion induzieren. Wir betrachten als Beispiel den Körper $K = (\mathbb{Z}_2, +_2, \cdot_2)$ und das Polynom

$$p = t^2 + t.$$

Dann ist die zugehörige Polynomfunktion $K \to K$ gerade $\widetilde{p}(t) = t^2 +_2 t = t \cdot_2 t +_2 t$. Diese erfüllt $p(0) = 0 \cdot_2 0 +_2 0 = 0 +_2 0 = 0$ sowie $p(1) = 1 \cdot_2 1 +_2 1 = 1 +_2 1 = 0$. Es ist also $\widetilde{p}$ die Nullfunktion, obwohl p nicht das Nullpolynom ist. Da das Nullpolynom ebenfalls die Nullfunktion induziert, ist die Zuordnung $p \mapsto \widetilde{p}$ in der Tat nicht injektiv.

Φ ist i. A. auch nicht surjektiv. Bild(Φ) ist der Unterring der Polynomfunktionen des Ringes $(R^R, +, \cdot)$. △

Definition 11.18 (Nullstelle eines Polynoms)**.**
Es sei R ein kommutativer Ring, $p \in R[t]$ ein Polynom und $\widetilde{p}\colon R \to R$ die zugehörige Polynomfunktion. $\lambda \in R$ heißt eine **Nullstelle** (englisch: zero) oder **Wurzel** (englisch: root) von p in R, wenn $\widetilde{p}(\lambda) = 0_R$ gilt. △

Beispiel 11.19 (Nullstelle eines Polynoms)**.**

(i) Das Polynom $p = t^2 + 1 \in \mathbb{R}[t]$ besitzt keine Nullstelle in $\mathbb{R}$, weil für die zugehörige Polynomfunktion $\widetilde{p}\colon \mathbb{R} \to \mathbb{R}$ gilt: $\widetilde{p}(t) = t^2 + 1 \geqslant 1$ für alle $t \in \mathbb{R}$.

(ii) Das Polynom $p = t^2 + 1 \in \mathbb{C}[t]$ besitzt aber zwei Nullstellen in $\mathbb{C}$, und zwar i und $-i$.

(iii) Das Polynom $p = t^2 + 1 \in \mathbb{Z}_5[t]$ besitzt in $\mathbb{Z}_5$ genau die Nullstellen 2 und 3, da für die zugehörige Polynomfunktion $\widetilde{p}\colon \mathbb{Z}_5 \to \mathbb{Z}_5$ gilt: $\widetilde{p}(t) = t \cdot_5 t +_5 1$ und damit

$$\begin{aligned}
\widetilde{p}(0) &= 0 \cdot_5 0 +_5 1 = 1, \\
\widetilde{p}(1) &= 1 \cdot_5 1 +_5 1 = 2, \\
\widetilde{p}(2) &= 2 \cdot_5 2 +_5 1 = 0, \\
\widetilde{p}(3) &= 3 \cdot_5 3 +_5 1 = 0, \\
\widetilde{p}(4) &= 4 \cdot_5 4 +_5 1 = 2.
\end{aligned}$$

(iv) Das Polynom $p = (t-0)\cdot(t-1)\cdot(t-2)\cdot(t-3)\cdot(t-4)+1$, besitzt in $\mathbb{Z}_5$ *keine* Nullstelle, denn es gilt $\widetilde{p}(t) = 1$ für alle $t \in \mathbb{Z}_5$. (Ein solches Polynom gibt es für jeden endlichen Körper.) △

Lemma 11.20 (Nullstellen und Teiler).
Es seien K ein Körper und $p \in K[t]$ ein Polynom. Dann sind äquivalent:

(i) $\lambda \in K$ ist eine Nullstelle von p.

(ii) Das Polynom $t - \lambda \in K[t]$ ist ein Teiler von p.

In diesem Fall gilt für das eindeutig bestimmte Polynom $q \in K[t]$ mit $p = (t-\lambda)\cdot q$ die Eigenschaft $\deg(q) = \deg(p) - 1$.

Beweis. Aussage $(i) \Rightarrow$ Aussage (ii): Es sei $\lambda \in K$ eine Nullstelle von p, also $\widetilde{p}(\lambda) = 0$. Nach Satz 11.14 gibt es (eindeutig bestimmte) Polynome $q, r \in K[t]$, sodass gilt: $p = q\cdot(t-\lambda) + r$ und $\deg(r) < \deg(t-\lambda) = 1$. Also ist r ein konstantes Polynom. Um es zu bestimmen, ist es ausreichend, den Wert der zugehörigen Polynomfunktion $\widetilde{r}$ an einer Stelle zu kennen. An der Stelle λ gilt

$$\widetilde{r}(\lambda) = \big(\widetilde{p} - \widetilde{q}\cdot(t-\lambda)\big)(\lambda) = \widetilde{p}(\lambda) - \widetilde{q}(\lambda)\cdot(\lambda-\lambda) = \widetilde{p}(\lambda) = 0_K$$

gilt. Also ist r das Nullpolynom, und es folgt $p = q\cdot(t-\lambda)$, d. h., $t-\lambda$ ist ein Teiler von p.

Aussage $(ii) \Rightarrow$ Aussage (i): Es sei $t-\lambda$ ein Teiler von p, also existiert ein $q \in K[t]$ mit der Eigenschaft $p = q\cdot(t-\lambda)$. Das Einsetzen von λ liefert $\widetilde{p}(\lambda) = \widetilde{q}(\lambda)\cdot(\lambda-\lambda) = \widetilde{q}(\lambda)\cdot 0_K = 0_K$. Also ist λ eine Nullstelle von p.

Weiter gilt nach Lemma 11.10 (iii) $\deg(p) = \deg(q\cdot(t-\lambda)) = \deg(q) + 1$. □

Satz 11.21 (Zerlegung eines Polynoms).
Es seien K ein Körper und $p \in K[t]$ ein Polynom, $p \neq 0_K$. Dann gilt:

(i) Es existieren $s \in \mathbb{N}_0$, paarweise verschiedene Zahlen $\lambda_1, \ldots, \lambda_s \in K$ sowie Exponenten $n_1, \ldots, n_s \in \mathbb{N}$ und ein Polynom $q \in K[t]$ ohne Nullstelle in K, sodass gilt:

$$p = (t-\lambda_1)^{n_1} \cdots (t-\lambda_s)^{n_s} \cdot q. \tag{11.10}$$

Die Zahl $n_i \in \mathbb{N}$ heißt die **Vielfachheit** (englisch: multiplicity) der Nullstelle λ_i. Jedes Polynom $t - \lambda_i \in K[t]$ heißt ein **Linearfaktor** (englisch: linear factor) von p.

(*ii*) Die Nullstellen von p sind genau die Zahlen $\lambda_1, \dots, \lambda_s \in K$.

Beweis. Aussage (*i*): Wir führen eine Induktion nach $n \coloneqq \deg(p)$ durch. Im Fall $n = 0$ (Induktionsanfang) ist p ein konstantes Polynom ungleich 0_K, das keine Nullstelle besitzt. Daher ist dann $s = 0$ und $q = p$.

Wir nehmen an, die Behauptung sei bereits gezeigt für Polynome vom Grad $n \in \mathbb{N}_0$. Es sei p nun ein Polynom vom Grad $n+1$. Wenn p keine Nullstelle besitzt, so gilt die Behauptung mit $s = 0$ und $q = p$. Andernfalls besitzt p eine Nullstelle $\lambda \in K$. Nach Lemma 11.20 ist also $t - \lambda \in K[T]$ ein Teiler von p, d. h., es gilt $p = (t - \lambda) \cdot \widehat{p}$. Aus Lemma 11.10 folgt $n + 1 = \deg(p) = \deg(t - \lambda) + \deg(\widehat{p}) = 1 + \deg(\widehat{p})$, also gilt $\deg(\widehat{p}) = n$. Nach Induktionsvoraussetzung besitzt also $\widehat{p}$ eine Darstellung wie in (11.10), somit auch $p = (t - \lambda) \cdot \widehat{p}$.

Aussage (*ii*): Ist $\mu \in K$ eine Nullstelle von p, dann folgt

$$0 = \widetilde{p}(\mu) = (\mu - \lambda_1)^{n_1} \cdots (\mu - \lambda_s)^{n_s} \cdot \widetilde{q}(\mu).$$

Da K als Körper nullteilerfrei ist, muss einer der Faktoren gleich 0_K sein. Da aber q nach Voraussetzung keine Nullstelle besitzt, muss es ein $i \in [\![1, s]\!]$ geben mit $\mu = \lambda_i$.

Umgekehrt ist nach Lemma 11.20 jedes λ_i, $i \in [\![1, s]\!]$, eine Nullstelle von p. □

Beachte: Man kann zeigen, dass die Darstellung (11.10) bis auf die Reihenfolge der Faktoren eindeutig ist.

Folgerung 11.22 (Zerlegung eines Polynoms)**.**
Es seien K ein Körper und $p \in K[t]$ ein Polynom, $p \neq 0_K$. Dann gilt:

(*i*) p hat höchstens $\deg(p) \in \mathbb{N}_0$ viele paarweise verschiedene Nullstellen, also $s \leqslant \deg(p)$.

(*ii*) p hat höchstens $\deg(p) \in \mathbb{N}_0$ viele Nullstellen, wenn diese entsprechend ihrer Vielfachheit gezählt werden, also gilt $\sum_{i=1}^s n_i \leqslant \deg(p)$.

Beweis. Aussage (*i*): Es sei (11.10) die (i. W. eindeutige) Zerlegung von p mit $s \in \mathbb{N}_0$ paarweise verschiedenen Nullstellen von p. Dann gilt

$$\begin{aligned} \deg(p) &= \deg(q) + \sum_{i=0}^{s} n_i && \text{nach Lemma 11.10} \\ &\geqslant 0 + \sum_{i=0}^{s} 1 && q \neq 0_K \text{ und die Vielfachheit jeder Nullstelle ist} \geqslant 1 \\ &= s. \end{aligned}$$

Aussage (*ii*): Zählen wir die Nullstellen λ_i in (11.10) gemäß ihrer Vielfachheit n_i, so erhalten wir

$$\begin{aligned} \deg(p) &= \deg(q) + \sum_{i=0}^{s} n_i && \text{nach Lemma 11.10} \\ &\geqslant 0 + \sum_{i=1}^{s} n_i. \end{aligned}$$

□

In Bemerkung 11.17 hatten wir gesehen, dass die Abbildung Polynom $p \mapsto$ Polynomfunktion $\widetilde{p}$ i. A. nicht injektiv ist, weil auch Nicht-Nullpolynome auf die Nullfunktion abgebildet werden können. Das Beispiel dafür war ein Polynom über dem endlichen Körper $\mathbb{Z}_2$. Wie folgendes Ergebnis zeigt, ist die Endlichkeit des Körpers charakteristisch für die Nichtinjektivität von Φ.

Folgerung 11.23.
Es sei K ein unendlicher Körper. Dann ist die Abbildung $\Phi\colon K[t] \to K^K$ aus (11.9) injektiv.

Beweis. Es seien $p_1, p_2 \in K[t]$ Polynome und $\widetilde{p}_1 = \widetilde{p}_2$ die zugehörigen Polynomfunktionen. Wir setzen $q \coloneqq p_1 - p_2$. Wegen $\widetilde{q} = \widetilde{p}_1 - \widetilde{p}_2 = 0_K$ (**Quizfrage 11.4:** Warum gilt diese Gleichheit?) hat q unendlich viele Nullstellen, nämlich alle Elemente aus K. Das widerspricht Folgerung 11.22, es sei denn, q ist das Nullpolynom. Daher gilt $p_1 = p_2$, d. h., die Injektivität von Φ. □

Satz 11.24 (Fundamentalsatz der Algebra[19]).
Jedes Polynom $p \in \mathbb{C}[t]$ mit $\deg(p) > 0$ hat mindestens eine Nullstelle.

Ein Beweis dieses Satz wird i. d. R. in weiterführenden Veranstaltungen über *Funktionentheorie* oder *Algebra* vorgestellt.

Folgerung 11.25 (nicht-konstante Polynome über den komplexen Zahlen $\mathbb{C}$ zerfallen in Linearfaktoren)**.**

Jedes nicht-konstante Polynom $p \in \mathbb{C}[t]$ zerfällt vollständig in Linearfaktoren. In der Darstellung (11.10) gilt also $\deg(q) = 0$.

Beweis. Der Beweis gelingt mit vollständiger Induktion nach $n = \deg(p)$ und Anwendung des Fundamentalsatzes 11.24. □

Ende der Vorlesung 14

Ende der Woche 7

[19] englisch: fundamental theorem of algebra

Kapitel 3 Vektorräume

§ 12 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, sagen wir $(K, +, \cdot)$. In Anlehnung an dessen Verknüpfungen bezeichnen wir die beiden Verknüpfungen im Vektorraum V mit $\oplus$ und $\odot$.

Definition 12.1 (Vektorraum)**.**
Es sei $(K, +, \cdot)$ ein Körper. Ein **Vektorraum** (englisch: vector space) oder **linearer Raum** (englisch: linear space) $(V, \oplus, \odot)$ **über** 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.

(*ii*) Für die Verknüpfung $\odot$, genannt **skalare Multiplikation** (englisch: scalar multiplication) oder **S-Multiplikation**, gelten die folgenden Gesetze für alle $\alpha, \beta \in K$ und $u, v \in V$: die **„gemischten“ Distributivgesetze**[1] (englisch: „mixed“ distributive laws)

$$\alpha \odot (u \oplus v) = (\alpha \odot u) \oplus (\alpha \odot v) \tag{12.1a}$$

$$(\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v) \tag{12.1b}$$

sowie das **„gemischte“ Assoziativgesetz** (englisch: „mixed“ associative law)

$$(\alpha \cdot \beta) \odot v = \alpha \odot (\beta \odot v). \tag{12.1c}$$

Weiterhin ist das neutrale Element 1_K bzgl. $\cdot$ in K auch neutral bzgl. $\odot$:

$$1_K \odot v = v. \tag{12.1d}$$

(*iii*) In einem K-Vektorraum V heißen die Elemente von V auch **Vektoren** (englisch: vectors). Das Nullelement 0_V von $(V, \oplus)$ heißt auch der **Nullvektor** (englisch: zero vector). Die Elemente von K heißen **Skalare** (englisch: scalars), und K selbst heißt der **Skalarkörper** (englisch: scalar field) von V. △

Bemerkung 12.2 (abkürzende Schreibweisen)**.**

(*i*) Das Inverse zu $v \in V$ bzgl. $\oplus$ bezeichnen wir mit $\ominus v$. Die Bezeichnung $u \ominus v$ steht für $u \oplus (\ominus v)$.

[1] Wir bezeichnen die Gesetze hier als „gemischt“, weil in ihnen Skalare $\alpha, \beta \in K$ mit Vektoren $u, v \in V$ gemischt auftauchen.

(*ii*) Wir vereinbaren, dass $\odot$ stärker bindet als $\oplus$, also könnten wir die rechte Seite in (12.1a) auch in der Form $\alpha \odot u \;\oplus\; \alpha \odot v$ schreiben.

(*iii*) 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 $\boxdot\colon V \times K \to V$ definieren, und zwar durch $v \boxdot \alpha \coloneqq \alpha \odot v$. Dann gelten auch dafür die Gesetze (12.1), 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.

(*iv*) 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 12.10.

(*v*) Wir werden im Folgenden statt von einem „Vektorraum $(V, \oplus, \odot)$ über dem Körper $(K, +, \cdot)$" auch einfach von einem „Vektorraum $(V, \oplus, \odot)$" sprechen, wenn der zugrundeliegende Körper aus dem Zusammenhang klar ist oder wenn wir uns nicht explizit auf ihn beziehen müssen. △

Beispiel 12.3 (Vektorraum)**.**

(*i*) Über jedem Körper $(K, +, \cdot)$ gibt es den (bis auf Isomorphie eindeutigen) Vektorraum $(V, \oplus, \odot)$ mit nur einem Element, nämlich $V = \{0_V\}$. Die Verknüpfungen $\oplus$ und $\odot$ sind dann eindeutig festgelegt. Dieser Raum heißt **Nullraum** (englisch: zero vector space) über K.

(*ii*) Jeder Körper $(K, +, \cdot)$, ausgestattet mit den Verknüpfungen $\oplus \coloneqq +$ und $\odot \coloneqq \cdot$, ist ein Vektorraum über sich selbst.

(*iii*) Allgemeiner sei $(K, +, \cdot)$ ein Körper und $(U, +, \cdot)$ ein Unterkörper. Dann ist $(K, +, \cdot)$ ein Vektorraum über U. Beispielsweise ist $\mathbb{R}$ ein $\mathbb{Q}$-Vektorraum, und $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum und ein $\mathbb{Q}$-Vektorraum.

(*iv*) Es sei $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Dann ist die Menge

$$K_n \coloneqq \Big\{(x_1, \ldots, x_n) \,\Big|\, x_i \in K \text{ für } i = 1, \ldots, n\Big\} \tag{12.2}$$

der n-Tupel[2] über K, ausgestattet mit der **komponentenweisen Addition** (englisch: componentwise addition) und der **komponentenweisen skalaren Multiplikation** (englisch: componentwise scalar multiplication)

$$(x_1, \ldots, x_n) \oplus (y_1, \ldots, y_n) \coloneqq (x_1 + y_1, \ldots, x_n + y_n), \tag{12.3a}$$

$$\alpha \odot (x_1, \ldots, x_n) \coloneqq (\alpha \cdot x_1, \ldots, \alpha \cdot y_n), \tag{12.3b}$$

ein Vektorraum über K, genannt der **Vektorraum der Zeilenvektoren** (englisch: row vector space) über K der Dimension n.[3] Der Nullvektor ist der Vektor $(0_K, \ldots, 0_K) \in K_n$.

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 $()$. Es gilt $\alpha \odot () = ()$ für alle $\alpha \in K$.

[2] Gemäß Definition 4.8 müssten wir die n-Tupel eigentlich mit K^n bezeichnen. Es ist aber üblich, die Bezeichnung K^n für den Vektorraum der Spaltenvektoren zu verwenden.

[3] Der Begriff der Dimension für beliebige Vektorräume wird in Definition 13.16 eingeführt. Hier dient er zunächst zur Angabe der Anzahl der Einträge in einem Zeilenvektor.

(*v*) Es sei $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Dann ist die Menge

$$K^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,\middle|\, x_i \in K \text{ für } i = 1, \ldots, n \right\}, \tag{12.4}$$

ausgestattet mit der **komponentenweisen Addition** (englisch: componentwise addition) und der **komponentenweisen skalaren Multiplikation** (englisch: componentwise scalar multiplication)

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \tag{12.5a}$$

$$\alpha \odot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \alpha \cdot x_1 \\ \vdots \\ \alpha \cdot x_n \end{pmatrix}, \tag{12.5b}$$

ein Vektorraum über K, genannt der **Vektorraum der Spaltenvektoren** (englisch: column vector space) oder auch der **Standardvektorraum** (englisch: standard vector space) oder der **Koordinatenraum** (englisch: coordinate space) über K der Dimension n.[4] Der Nullvektor ist der Vektor $\begin{pmatrix} 0_K \\ \vdots \\ 0_K \end{pmatrix}$ 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 $()$. Es gilt $\alpha \odot () = ()$ für alle $\alpha \in K$.

(*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

$$\begin{aligned} (f \oplus g)(x) &:= f(x) + g(x), \\ (\alpha \odot f)(x) &:= \alpha \cdot f(x), \end{aligned}$$

ein Vektorraum über K. Der Nullvektor ist die Nullfunktion.

Insbesondere ist die Menge der Folgen $K^{\mathbb{N}}$ mit Werten in K ein Vektorraum.

(*vii*) Es sei $(K, +, \cdot)$ ein Körper und $K[t]$ der Polynomring (Definition 11.4). In $K[t]$ ist die Addition $+$ gemäß (11.3a) erklärt.[5] Ergänzen wir die skalare Multiplikation[6]

$$\alpha \odot \left(a_n\, t^n + \cdots + a_1\, t + a_0\right) := \alpha \cdot a_n\, t^n + \cdots + \alpha \cdot a_1\, t + \alpha \cdot a_0, \tag{12.6}$$

so wird $K[t]$ zu einem Vektorraum über K, genannt der **Polynomraum** (englisch: vector space of polynomials) über K. △

Lemma 12.4 (Rechenregeln in Vektorräumen, vgl. Lemma 9.3).
Es sei $(K, +, \cdot)$ ein Körper und $(V, \oplus, \odot)$ ein Vektorraum über K. Dann gilt für $\alpha \in K$ und $v \in V$:

[4] Der Begriff **Koordinatenraum** wird später in Satz 19.1 klar werden.
[5] Die zweite Verknüpfung in $K[t]$, also die Multiplikation $\cdot$ von Polynomen gemäß (11.3b), ignorieren wir hier.
[6] Die skalare Multiplikation stimmt überein mit der Multiplikation von Polynomen, wobei einer der Faktoren ein konstantes Polynom ist, das mit einem Element aus K identifiziert werden kann.

(*i*) $0_K \odot v = 0_V = v \odot 0_K$.

(*ii*) $\alpha \odot 0_V = 0_V$.

(*iii*) $\alpha \odot v = 0_V \quad \Rightarrow \quad \alpha = 0_K$ oder $v = 0_V$.

(*iv*) $\alpha \odot (\ominus v) = \ominus (\alpha \odot v) = (-\alpha) \odot v$ und insbesondere $\ominus v = (-1_K) \odot v$.

Beweis. Aussage (*i*):

$$\begin{aligned} 0_V \oplus 0_K \odot v &= 0_K \odot v && \text{da } 0_V \text{ das neutrale Element von } (V,\oplus) \text{ ist} \\ &= (0_K + 0_K) \odot v && \text{da } 0_K \text{ das neutrale Element von } (K,+) \text{ ist} \\ &= 0_K \odot v \oplus 0_K \odot v && \text{nach Distributivgesetz (12.1b).} \end{aligned}$$

Aus der Kürzungsregel (7.10b) in der Gruppe $(V,\oplus)$ folgt nun $0_V = 0_K \odot v$. Der Beweis der zweiten Gleichheit folgt sofort aus $0_K \odot v = v \odot 0_K$.

Aussage (*ii*): Für beliebiges $\alpha \in K$ gilt

$$\begin{aligned} \alpha \odot 0_V \oplus 0_V &= \alpha \odot 0_V && \text{da } 0_V \text{ das neutrale Element von } (V,\oplus) \text{ ist} \\ &= \alpha \odot (0_V \oplus 0_V) && \text{da } 0_V \text{ das neutrale Element von } (V,\oplus) \text{ ist} \\ &= \alpha \odot 0_V \oplus \alpha \odot 0_V && \text{nach Distributivgesetz (12.1a).} \end{aligned}$$

Aus der Kürzungsregel (7.10a) in der Gruppe $(V,\oplus)$ folgt nun $0_V = \alpha \odot 0_V$.

Aussage (*iii*): Es seien $\alpha \in K$ und $v \in V$ sowie $\alpha \odot v = 0_V$. Wir nehmen $\alpha \neq 0_K$ an. Dann gilt

$$\begin{aligned} v &= 1_K \odot v && \text{nach (12.1d)} \\ &= (\alpha \cdot \alpha^{-1}) \odot v && \text{da } \alpha \neq 0_K \text{ in der Gruppe } (K \setminus \{0\}, \cdot) \text{ das Inverse } \alpha^{-1} \text{ besitzt} \\ &= \alpha^{-1} \odot (\alpha \odot v) && \text{nach Assoziativgesetz (12.1c)} \\ &= \alpha^{-1} \odot 0_V && \text{nach Voraussetzung} \\ &= 0_V && \text{nach Aussage } (ii). \end{aligned}$$

Aussage (*iv*): 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:

$$\begin{aligned} \big[(-\alpha) \odot v\big] \oplus \big[\alpha \odot v\big] &= (-\alpha + \alpha) \odot v && \text{nach Distributivgesetz (12.1b)} \\ &= 0_K \odot v && \text{da } \alpha \text{ in der Gruppe } (K,+) \text{ das Inverse } -\alpha \text{ besitzt} \\ &= 0_V && \text{nach Aussage } (i). \end{aligned}$$

Das heißt, es gilt $\ominus (\alpha \odot v) = (-\alpha) \odot v$. Insbesondere für $\alpha = 1_K$ erhalten wir

$$\ominus v = \ominus (1_K \odot v) = (-1_K) \odot v. \tag{$*$}$$

Weiter haben wir

$$\begin{aligned} (-\alpha) \odot v &= \big[(-1_K) \cdot \alpha\big] \odot v && \text{nach Lemma 9.3} \\ &= \big[\alpha \cdot (-1_K)\big] \odot v && \text{da } (K,\cdot) \text{ kommutativ ist} \\ &= \alpha \odot \big((-1_K) \odot v\big) && \text{nach Assoziativgesetz (12.1c)} \\ &= \alpha \odot (\ominus v) && \text{nach } (*). \end{aligned}$$ ☐

Definition 12.5 (Linearkombination)**.**
Es sei $(K, +, \cdot)$ ein Körper und $(V, \oplus, \odot)$ ein Vektorraum über K.

(i) Es sei $E \subseteq V$. Ein Vektor der Form

$$\alpha_1 \odot v_1 \oplus \cdots \oplus \alpha_n \odot v_n \quad \text{oder kurz} \quad \sum_{j=1}^{n} \alpha_j \odot v_j \tag{12.7}$$

mit endlich vielen Skalaren $\alpha_j \in K$ und Vektoren $v_j \in E$ für $j = 1, \ldots, n$ und $n \in \mathbb{N}_0$ heißt eine **Linearkombination** (englisch: linear combination) **der Menge** E oder **Linearkombination von Vektoren der Menge** E. Falls $n = 0$ gilt, so interpretieren wir wie üblich die Addition von null Elementen in (12.7) als den Nullvektor 0_V.

(ii) Es sei $F = (v_i)_{i \in I}$ eine Familie von Vektoren in V. Ein Vektor der Form

$$\alpha_1 \odot v_{i_1} \oplus \cdots \oplus \alpha_n \odot v_{i_n} \quad \text{oder kurz} \quad \sum_{j=1}^{n} \alpha_j \odot v_{i_j} \tag{12.8}$$

mit endlich vielen Indizes $i_j \in I$ und Skalaren $\alpha_j \in K$ für $j = 1, \ldots, n$ und $n \in \mathbb{N}_0$ heißt eine **Linearkombination der Familie** F oder **Linearkombination von Vektoren der Familie** F. Falls $n = 0$ gilt, so interpretieren wir auch hier die Addition von null Elementen in (12.8) als den Nullvektor 0_V.

In beiden Fällen heißen die Skalare α_j die **Koeffizienten** (englisch: coefficients) der Linearkombination. Die Linearkombination heißt **trivial** (englisch: trivial linear combination), wenn alle $\alpha_j = 0$ sind. Falls $n = 0$ gilt, so interpretieren wir wie üblich die Addition von null Elementen als den Nullvektor 0_V. △

Beachte: Linearkombinationen bestehen immer aus endlich vielen Termen!

Beispiel 12.6 (Linearkombination)**.**

(i) Der Vektor $\left(\begin{smallmatrix} 3 \\ -7 \end{smallmatrix}\right)$ ist eine Linearkombination der Menge der Vektoren $\left\{\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\right\}$ im Standardvektorraum $\mathbb{R}^2$, nämlich

$$\begin{pmatrix} 3 \\ -7 \end{pmatrix} = 3 \odot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus (-7) \odot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(ii) Der Vektor $\left(\begin{smallmatrix} 3 \\ -7 \end{smallmatrix}\right)$ ist eine Linearkombination der Menge der Vektoren $\left\{\left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 4 \\ 1 \end{smallmatrix}\right)\right\}$ im Standardvektorraum $\mathbb{R}^2$, nämlich

$$\begin{pmatrix} 3 \\ -7 \end{pmatrix} = \frac{31}{6} \odot \begin{pmatrix} 2 \\ -1 \end{pmatrix} \oplus \frac{-11}{6} \odot \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

(Zur Bestimmung der Koeffizienten wurde ein lineares Gleichungssystem gelöst, siehe Kapitel 4.)

(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 der Funktionen (Vektoren) $\{\sin, \cos\}$ im Vektorraum der Funktionen $\mathbb{R}^{[0,2\pi]}$.

(iv) Das Polynom $3\,t^2 \oplus 5$ im Vektorraum (Polynomraum) $\mathbb{Q}[t]$ ist eine Linearkombination der Menge der Polynome $\{t^2, t, 1\}$, wobei 1 das Einspolynom in $\mathbb{Q}[t]$ bezeichnet. Das Polynom t kommt in der Linearkombination $3\,t^2 \oplus 5$ mit dem Koeffizienten $0 \in \mathbb{Q}$ vor. △

Ende der Vorlesung 15

Definition 12.7 (Unterraum).
Es sei $(K, +, \cdot)$ ein Körper und $(V, \oplus, \odot)$ ein Vektorraum über K.

(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 skalaren Multiplikation $\odot$ mit Elementen in K abgeschlossen ist und wenn $(U, \oplus, \odot)$ selbst wieder ein Vektorraum ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, \oplus)$ eine Untergruppe von $(V, \oplus)$ ist und wenn U bzgl. der skalaren Multiplikation $\odot$ mit Elementen in K abgeschlossen ist, also $K \odot U \subseteq U$.

(ii) Ein Unterraum $(U, \oplus, \odot)$ von $(V, \oplus, \odot)$ heißt **echt** (englisch: proper subspace), wenn $U \subsetneq V$ gilt. △

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 bewerkstelligen:

Satz 12.8 (Unterraumkriterium).
Es sei $(K, +, \cdot)$ ein Körper, $(V, \oplus, \odot)$ ein Vektorraum über K 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 $\alpha \odot u \in U$.
(Mit anderen Worten: $U \neq \emptyset$, und es gilt $U \oplus U \subseteq U$ sowie $K \odot U \subseteq U$.)

(iii) $U \neq \emptyset$, und für alle $u, v \in U$ und $\alpha, \beta \in K$ gilt $\alpha \odot u \;\oplus\; \beta \odot v \in U$.
(Mit anderen Worten: $U \neq \emptyset$, und es gilt $K \odot U \;\oplus\; K \odot U \subseteq U$.)

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $(U, \oplus, \odot)$ ein Unterraum von $(V, \oplus, \odot)$. Dann ist per Definition 12.7 und Definition 12.1 $(U, \oplus)$ 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 (12.1) 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 7.33) 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 12.9 (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 $(\{0_V\}, \oplus, \odot)$ ist der **Nullraum** (englisch: zero vector space) von V.

(ii) Wir betrachten den Standardvektorraum $V = \mathbb{R}^2$ über dem Körper $\mathbb{R}$.

(a) Die Menge
$$U_1 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1 - 2\,x_2 = 0 \right\}$$
ist ein Unterraum von $V = \mathbb{R}^2$, denn die Aussage (ii) des Unterraumkriteriums Satz 12.8 ist erfüllt: Zunächst gilt $\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \in U_1$, also ist $U_1 \neq \emptyset$. Weiter gilt für alle $\alpha \in \mathbb{R}$ und $u = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \in U_1$ sowie $v = \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right) \in U_1$:
$$\begin{aligned} u \oplus v &= \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \in U_1, \quad \text{denn } (x_1 + y_1) - 2\,(x_2 + y_2) = (x_1 - 2\,x_2) + (y_1 - 2\,y_2) = 0 \\ \alpha \odot u &= \begin{pmatrix} \alpha\, x_1 \\ \alpha\, x_2 \end{pmatrix} \in U_1, \quad \text{denn } \alpha\, x_1 - 2\,(\alpha\, x_2) = \alpha\,(x_1 - 2\,x_2) = 0. \end{aligned}$$

(b) Die Menge
$$U_2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1 - 2\,x_2 = 1 \right\}$$
ist kein Unterraum von $V = \mathbb{R}^2$, denn sie enthält nicht den Nullvektor $\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, was aber eine notwendige Bedingung für einen Unterraum darstellt, wie wir im Anschluss an Definition 12.7 bereits gesehen haben.

(c) Die Menge
$$U_3 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1 \geqslant 0,\ x_2 \geqslant 0 \right\}$$
ist kein Unterraum von $V = \mathbb{R}^2$, denn sie erfüllt das Unterraumkriterium nicht. Beispielsweise ist $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \in U_3$, aber $(-1) \odot \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 \\ -1 \end{smallmatrix}\right)$ nicht.

(iii) Es sei $(K, +, \cdot)$ ein Körper und $(K^{\mathbb{N}}, \oplus, \odot)$ der Vektorraum der Folgen mit Werten in K. Dann ist die Menge der Folgen mit endlichem Träger (siehe Bemerkung 11.6) und Werten in K
$$(K^{\mathbb{N}})_{00} := \big\{ (y_i)_{i\in\mathbb{N}} \,\big|\, \{ i \in \mathbb{N} \mid y_i \neq 0_K \} \text{ ist endlich} \big\} \tag{12.9}$$
ein echter Unterraum von $(K^{\mathbb{N}}, \oplus, \odot)$.

(iv) Es sei $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ der Vektorraum der Folgen mit Werten in $\mathbb{R}$. Dann sind die Mengen

$$(\mathbb{R}^{\mathbb{N}})_b := \big\{ (y_i)_{i\in\mathbb{N}} \,\big|\, \exists C \geqslant 0\ \forall i \in \mathbb{N}\ (|y_i| \leqslant C) \big\} \quad \text{der \textbf{beschränkten Folgen}}^{7} \tag{12.10a}$$
$$(\mathbb{R}^{\mathbb{N}})_c := \big\{ (y_i)_{i\in\mathbb{N}} \,\big|\, (y_i)_{i\in\mathbb{N}} \text{ ist konvergent} \big\} \quad \text{der \textbf{konvergenten Folgen}}^{8} \tag{12.10b}$$
$$(\mathbb{R}^{\mathbb{N}})_0 := \big\{ (y_i)_{i\in\mathbb{N}} \,\big|\, (y_i)_{i\in\mathbb{N}} \text{ konvergiert gegen } 0 \big\} \quad \text{der \textbf{Nullfolgen}}^{9} \tag{12.10c}$$
$$(\mathbb{R}^{\mathbb{N}})_{00} := \big\{ (y_i)_{i\in\mathbb{N}} \,\big|\, (y_i)_{i\in\mathbb{N}} \text{ hat endlichen Träger} \big\} \quad \text{der \textbf{Folgen mit endlichem Träger}}^{10} \tag{12.10d}$$

jeweils echte Unterräume von $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$. (**Quizfrage 12.1:** Gibt es auch Unterraumbeziehungen untereinander?)

[7] englisch: bounded sequences
[8] englisch: convergent sequences
[9] englisch: null sequences
[10] englisch: finitely supported sequences

(v) Es sei $(K, +, \cdot)$ ein Körper und $(K[t], +, \cdot)$ der Polynomraum über K; siehe Beispiel 12.3. Dann ist für jedes $n \in \mathbb{N}_0$ die Menge der **Polynome vom Höchstgrad n** (englisch: polynomials of maximum degree n)

$$K_n[t] := \left\{ a_n\, t^n + a_{n-1}\, t^{n-1} + \cdots + a_1\, t + a_0 \,\middle|\, a_i \in K \text{ für } i = 0, \ldots, n \right\} \tag{12.11}$$

ein echter Unterraum von $(K[t], +, \cdot)$. (**Quizfrage 12.2:** Bilden auch die Polynome vom Mindestgrad $n \in \mathbb{N}$ einen Unterraum von $K[t]$?) (**Quizfrage 12.3:** Bilden die Polynome von geradem Grad $n \in 2\,\mathbb{N}_0$ einen Unterraum von $K[t]$?) △

Bemerkung 12.10 (Vereinfachung der Notation)**.**
Zur Vereinfachung der Notation werden wir in Zukunft für die Verknüpfungen $\oplus$ und $\odot$ in Vektorräumen einfach dieselbe Notation verwenden wie für die Verknüpfungen $+$ und $\cdot$ des zugrundeliegenden Körpers. Aus dem Zusammenhang ist klar, ob die jeweilige Verknüpfung mit Skalaren oder mit Vektoren gemeint ist.

Das Zeichen für die Multiplikation $\cdot$ von Vektoren mit Elementen aus dem zugrundeliegenden Körper oder von Körperelementen untereinander wird sogar oft ganz weggelassen, sofern sich dadurch keine Unklarheiten ergeben. Beispielsweise schreiben wir in Zukunft einfach $\sin - \sqrt{2} \cos$ an Stelle von $\sin \ominus \sqrt{2} \odot \cos$.

Wir werden außerdem das Nullelement von K einfach als 0 schreiben, das Einselement von K als 1 und den Nullvektor von V ebenfalls als 0. △

Wie bereits bei Untergruppen in § 7.4 beschäftigt uns nun die Frage, wie man Unterräume erzeugen kann.

Lemma 12.11 (Durchschnitt von Unterräumen)**.**
Es sei $(V, +, \cdot)$ ein Vektorraum und $(U_i, +, \cdot)$ eine Familie von Unterräumen mit der nichtleeren Indexmenge I. Dann ist auch $\bigcap_{i \in I} U_i$ mit $+$ und $\cdot$ ein Unterraum von $(V, +, \cdot)$.

Beweis. Dieser Beweis ist Gegenstand von Hausaufgabe I-8.3. □

Definition 12.12 (erzeugter Unterraum, lineare Hülle, Erzeugendensystem)**.**
Es sei $(V, +, \cdot)$ ein Vektorraum und $E \subseteq V$.

(i) Dann heißt

$$\langle E \rangle := \bigcap \left\{ U \,\middle|\, (U, +, \cdot) \text{ ist Unterraum von } (V, +, \cdot) \text{ und } E \subseteq U \right\} \tag{12.12}$$

der von E **erzeugte Unterraum** (englisch: subspace generated by E) oder die **lineare Hülle** (englisch: linear hull) $\operatorname{Lin}(E)$ von E oder auch der **Spann** (englisch: span) $\operatorname{Span}(E)$ von E in $(V, +, \cdot)$.

Beachte: Bezeichnen wir mit $\mathcal{R}$ die Menge auf rechten Seite von (12.12), über die der Durchschnitt gebildet wird, dann ist $\langle E \rangle$ das Minimum der Menge $\mathcal{R}$ bzgl. der Inklusionshalbordnung und sogar das Minimum der Menge $\mathcal{R}$ bzgl. der Halbordnung „ist Unterraum von".

Ist speziell E die endliche Menge $E = \{v_1, \ldots, v_n\}$ für $v_i \in V$ und $n \in \mathbb{N}$, 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 ein **Erzeugendensystem** (englisch: generating set) von $(V,+,\cdot)$. Falls ein endliches Erzeugendensystem von V existiert, so heißt V **endlich erzeugt** (englisch: finitely generated).

(*iii*) Wenn $F = (v_i)_{i\in I}$ eine Familie von Vektoren in V ist, definieren wir den von F **erzeugten Unterraum** in $(V,+,\cdot)$ bzw. die **lineare Hülle** $\langle F\rangle$ von F bzw. den **Spann** $\operatorname{Span}(F)$ von F in $(V,+,\cdot)$ als

$$\langle F\rangle \coloneqq \bigcap \Big\{ U \,\Big|\, (U,+,\cdot) \text{ ist Unterraum von } (V,+,\cdot) \text{ und } \{v_i \mid i\in I\} \subseteq U \Big\}. \tag{12.13}$$

Wir nennen F ein **Erzeugendensystem** von $(V,+,\cdot)$, wenn $\langle F\rangle = V$ gilt. △

Satz 12.13 (Darstellung des erzeugten Unterraumes)**.**
Es sei $(K,+,\cdot)$ ein Körper, $(V,+,\cdot)$ ein Vektorraum über K und $E\subseteq V$. Dann gilt für den von E erzeugten Unterraum:

$$\langle E\rangle = \Big\{ \sum_{i=1}^{n} \alpha_i\, v_i \,\Big|\, \exists n\in\mathbb{N}_0\ \forall i=1,\ldots,n\ (v_i\in E,\ \alpha_i\in K) \Big\}. \tag{12.14}$$

(Zur Erinnerung: Im Fall $n=0$ interpretieren wir die Linearkombination von null Elementen in der rechten Menge als den Nullvektor 0. Insbesondere im Fall $E=\emptyset$ ist also $\langle E\rangle = \langle\emptyset\rangle = \{0\}$.)

Ist $F = (v_i)_{i\in I}$ eine Familie von Vektoren in V, dann gilt für den von F erzeugten Unterraum:

$$\langle F\rangle = \Big\{ \sum_{j=1}^{n} \alpha_j\, v_{i_j} \,\Big|\, \exists n\in\mathbb{N}_0\ \forall j=1,\ldots,n\ \exists i_j\in I\ (v_{i_j}\in F,\ \alpha_j\in K) \Big\}. \tag{12.15}$$

Beachte: Der von einer Menge E (oder einer Familie F) erzeugte Unterraum stimmt also überein mit der Menge der Linearkombinationen von E (bzw. F).[11] Auf die Reihenfolge der Elemente in F kommt es offenbar nicht an.

Beweis. Zur Abkürzung bezeichnen wir die Menge auf der rechten Seite von (12.14) mit M. Wir führen den Beweis analog zu Satz 7.37 in zwei Schritten.

Schritt 1: $\langle E\rangle \supseteq M$: Es sei U ein beliebiger Unterraum von V, der im Durchschnitt (12.12) 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 (12.12) 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 12.8). Offensichtlich ist $M\neq\emptyset$, denn M enthält mindestens den Nullvektor 0. Sind $\sum_{i=1}^n \alpha_i\, v_i$ und $\sum_{j=1}^m \beta_j\, w_j$ zwei Elemente aus M, so ist auch $\sum_{i=1}^n \alpha_i\, v_i + \sum_{j=1}^m \beta_j\, w_j$ ein Element aus M. Zudem ist für jedes $\alpha\in K$ auch $\alpha\sum_{i=1}^n \alpha_i\, v_i = \sum_{i=1}^n \alpha\,\alpha_i\, v_i$ ein Element aus M. Also ist M ein Unterraum von V. Zusätzlich ist klar, dass $E\subseteq M$ gilt. (**Quizfrage 12.4:** Details?) Das heißt, M ist einer derjenigen Unterräume von V, über die in der Definition von $\langle E\rangle$ der Durchschnitt gebildet wird. Folglich gilt $\langle E\rangle\subseteq M$.

Der Beweis für den Fall (12.15) geht analog. □

Beispiel 12.14 (lineare Hülle)**.**

[11] In manchen Büchern wird daher (12.14) auch als Definition des erzeugten Unterraumes verwendet.

(*i*) Es sei $(K, +, \cdot)$ ein Körper und $(K[t], +, \cdot)$ der Polynomraum über K; siehe Beispiel 12.3. Die Menge aller Monome $E = \{1, t, t^2, \ldots\}$ bildet ein Erzeugendensystem von $K[t]$. Die Menge der Monome $E = \{1, t, \ldots, t^n\}$ bis zum Grad n bildet ein Erzeugendensystem von $K_n[t]$, dem Unterraum der Polynome vom Höchstgrad $n \in \mathbb{N}_0$.

(*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 12.5:** Was passiert im Fall $w \in \langle v \rangle$?) △

Folgerung 12.15 (zu Satz 12.13: lineare Hülle von Vereinigung und Schnitt)**.**
Es sei $(V, +, \cdot)$ ein Vektorraum und $E_1, E_2 \subseteq V$. Dann gilt:

(*i*)
$$\langle E_1 \cup E_2 \rangle = \big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle. \tag{12.16}$$

(*ii*)
$$\langle E_1 \cap E_2 \rangle \subseteq \big\langle \langle E_1 \rangle \cap \langle E_2 \rangle \big\rangle. \tag{12.17}$$

Beweis. Aussage (*i*): Die lineare Hülle ist 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 12.13 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 (12.1c) 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 $\big\langle \langle E_1 \rangle \cup \langle E_2 \rangle \big\rangle \subseteq \langle E_1 \cup E_2 \rangle$.

Aussage (*ii*) Wegen $E_1 \subseteq \langle E_1 \rangle$ und $E_2 \subseteq \langle E_2 \rangle$ gilt $E_1 \cap E_2 \subseteq \langle E_1 \rangle \cap \langle E_2 \rangle$. Durch Bildung der linearen Hülle ergibt sich $\langle E_1 \cap E_2 \rangle \subseteq \big\langle \langle E_1 \rangle \cap \langle E_2 \rangle \big\rangle$. □

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

In diesem Abschnitt beantworten wir die Frage, wie wir einen Vektorraum durch ein möglichst kleines Erzeugendensystem darstellen können und damit Redundanz in der Darstellung vermeiden.

Definition 13.1 (lineare (Un-)abhängigkeit)**.**
Es sei $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K.

(*i*) Eine Menge $E \subseteq V$ heißt **linear unabhängig** (englisch: linearly independent), wenn in jeder Linearkombination 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:

$$v_1, \ldots, v_n \in E, \quad v_i \neq v_j \text{ für } i \neq j \quad \text{und} \quad \sum_{\ell=1}^{n} \alpha_\ell \, v_\ell = 0 \quad \Rightarrow \quad \alpha_\ell = 0 \text{ für alle } \ell = 1, \ldots, n. \tag{13.1}$$

(*ii*) Wenn dagegen eine Linearkombination von $n \in \mathbb{N}$ paarweise verschiedenen Vektoren aus E möglich ist, die den Nullvektor ergibt, wobei nicht alle Koeffizienten gleich 0 sind, dann heißt E **linear abhängig** (englisch: linearly dependent).

(*iii*) Eine Familie F von Vektoren in V heißt **linear unabhängig**, wenn in jeder Linearkombination von $n \in \mathbb{N}$ Vektoren aus F mit paarweise verschiedenen Indizes, die den Nullvektor ergibt, notwendig alle Koeffizienten gleich 0 sind, wenn also für alle $n \in \mathbb{N}$ gilt:

$$i_1, \ldots, i_n \in I, \quad i_j \neq i_k \text{ für } j \neq k \quad \text{und} \quad \sum_{\ell=1}^{n} \alpha_\ell \, v_{i_\ell} = 0 \quad \Rightarrow \quad \alpha_\ell = 0 \text{ für alle } \ell = 1, \ldots, n. \tag{13.2}$$

(*iv*) Wenn dagegen eine Linearkombination von $n \in \mathbb{N}$ Vektoren aus F zu paarweise verschiedenen Indizes möglich ist, die den Nullvektor ergibt, wobei nicht alle Koeffizienten gleich 0 sind, dann heißt F **linear abhängig**. △

Bemerkung 13.2 (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 sagen, die Menge $\{v\} \cup \{v_1, v_2\}$ sei linear unabhängig.

(*ii*) Die Menge $\{v\}$ bestehend aus einem einzigen Vektor ist linear unabhängig, falls $v \neq 0$ ist, ansonsten linear abhängig. (Dasselbe gilt für eine einelementige Familie von Vektoren.)

(*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. △

Die lineare Abhängigkeit einer Menge bzw. einer Familie von Vektoren bedeutet, dass man einen Vektor als Linearkombination der anderen darstellen kann:

Lemma 13.3 (lineare Abhängigkeit ist äquivalent zur Kombinierbarkeit)**.**
Es sei V ein Vektorraum.

(*i*) Es sind äquivalent:

(a) $E \subseteq V$ ist eine linear abhängige Menge.

(b) Es gibt einen Vektor $v \in E$, der als Linearkombination von $E \setminus \{v\}$ darstellbar ist.

(*ii*) Es sind äquivalent:

(a) $F = (v_i)_{i \in I}$ ist eine linear abhängige Familie mit einer Indexmenge I.

(b) Es gibt es einen Vektor v_{i^*} aus der Familie, der als Linearkombination der Familie $(v_i)_{i \in I \setminus \{i^*\}}$ darstellbar ist.

Beweis. Dieser Beweis ist Gegenstand von Hausaufgabe I-9.1. □

Lemma 13.4 (lineare (Un-)abhängigkeit von Teilmengen und Obermengen).
Es sei V ein Vektorraum.

(i) Jede Teilmenge einer Menge linear unabhängiger Vektoren aus V ist ebenfalls linear unabhängig. Jede Teilfamilie einer Familie linear unabhängiger Vektoren aus V ist ebenfalls linear unabhängig.

(ii) Jede Obermenge einer Menge linear abhängiger Vektoren aus V ist ebenfalls linear abhängig. Jede Oberfamilie einer Familie linear abhängiger Vektoren aus V ist ebenfalls linear abhängig.

Beweis. Der Beweis ergibt sich unmittelbar aus der Definition 12.12. □

Beispiel 13.5 (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

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

(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 12.3. Für $y \in X$ definieren wir die **charakteristische Funktion** (englisch: characteristic function) $e_y \colon X \to K$ durch[12]

$$x \mapsto e_y(x) \coloneqq \delta_{xy} \coloneqq \begin{cases} 1, & \text{falls } x = y, \\ 0, & \text{falls } x \neq y. \end{cases} \tag{13.3}$$

Dann ist die Menge $\{e_y \mid y \in X\}$ linear unabhängig.

(iv) Es sei $(K, +, \cdot)$ ein Körper und $(K[t], +, \cdot)$ der Polynomraum über K; siehe Beispiel 12.3. Dann ist die Menge der Monome $E = \{1, t, t^2, \ldots\}$ linear unabhängig. △

Die Bedeutung linear unabhängiger Mengen und Familien von Vektoren stellt folgendes Resultat klar.

Lemma 13.6 (lineare Unabhängigkeit und eindeutige Linearkombination).
Es sei V ein Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren aus V. Dann sind äquivalent:

(i) $(v_i)_{i\in I}$ 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 paarweise verschiedenen Vektoren der Familie $(v_i)_{i\in I}$ linearkombinieren.

[12]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 $x \notin A$.

Ein analoges Resultat gilt für Mengen von Vektoren aus V.

Beweis. 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 12.13 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

$$v = \sum_{i\in I_0} \alpha_i\, v_i. \tag{$*$}$$

Zu zeigen ist noch, dass diese Darstellung eindeutig ist. Wir nehmen dazu an, dass

$$v = \sum_{j\in I_1} \beta_j\, v_j \tag{$**$}$$

eine weitere Darstellung von v als Linearkombination von $(v_i)_{i\in I}$ ist mit endlicher Indexmenge I_1. Wir setzen $N := \{i \in I_0 \mid \alpha_i \neq 0\} \cup \{j \in I_1 \mid \beta_j \neq 0\}$. Dann ist N endlich. Wir haben

$$\begin{aligned}
0 &= v - v \\
&= \sum_{i\in I_0} \alpha_i\, v_i - \sum_{j\in I_1} \beta_j\, v_j \\
&= \sum_{i\in N} \alpha_i\, v_i - \sum_{j\in N} \beta_j\, v_j \quad \text{indem wir } \alpha_i = 0 \text{ für } i \in N \setminus I_0 \text{ und } \beta_j = 0 \text{ für } j \in N \setminus I_1 \text{ ergänzen} \\
&= \sum_{i\in N} (\alpha_i - \beta_i)\, v_i.
\end{aligned}$$

Da nach Voraussetzung $(v_i)_{i\in I}$ linear unabhängig ist, muss $\alpha_i - \beta_i = 0$ für alle $i \in N$ gelten, also $\alpha_i = \beta_i$. Die Darstellungen $(*)$ und $(**)$ unterscheiden sich also nur um eventuell auftauchende Terme mit Nullkoeffizienten.

Aussage (ii) $\Rightarrow$ Aussage (i): Es sei $I_0 \subseteq I$ eine beliebige endliche Teilmenge. Wir müssen zeigen, dass $(v_i)_{i\in I_0}$ linear unabhängig ist. Wir untersuchen also 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.

Der Beweis für Mengen geht analog. □

Lemma 13.7 (lineare Abhängigkeit bedeutet Redundanz)**.**
Es sei V ein Vektorraum und $(v_i)_{i\in I}$ eine nichtleere Familie von Vektoren aus V. Dann sind äquivalent:

(i) $(v_i)_{i\in I}$ ist linear abhängig.

(ii) Es gibt ein $j \in I$, sodass gilt:
$$\langle (v_i)_{i\in I\setminus\{j\}}\rangle = \langle (v_i)_{i\in I}\rangle.$$

Ein analoges Resultat gilt für nichtleere Mengen von Vektoren aus V.

Beachte: Ist ein Erzeugendensystem eines Vektorraumes linear abhängig, so kann man mindestens ein Element aus dem Erzeugendensystem entfernen, ohne den erzeugten Raum zu verkleinern.

Beweis. Aussage $(i) \Rightarrow$ Aussage (ii): Es sei $(v_i)_{i \in I}$ linear abhängig, d. h., es gibt eine endliche Teilmenge $I_0 \subseteq I$ und Koeffizienten $\alpha_i \in K$, $i \in I_0$, die nicht alle gleich Null sind, sodass

$$\sum_{i \in I_0} \alpha_i \, v_i = 0$$

gilt. Es sei $j \in I_0$ ein Index, für den $\alpha_j \neq 0$ gilt. Dann ist

$$\alpha_j \, v_j = -\sum_{i \in I_0 \setminus \{j\}} \alpha_i \, v_i,$$

also auch

$$v_j = -\alpha_j^{-1} \sum_{i \in I_0 \setminus \{j\}} \alpha_i \, v_i = -\sum_{i \in I_0 \setminus \{j\}} \alpha_j^{-1} \alpha_i \, v_i. \tag{$*$}$$

Das zeigt $v_j \in \langle (v_i)_{i \in I \setminus \{j\}} \rangle$.

Es sei nun $v \in \langle (v_i)_{i \in I} \rangle$ beliebig. Dann gibt es eine endliche Teilmenge $I_1 \subseteq I$ und Koeffizienten $\beta_i \in K$, $i \in I_1$, sodass

$$v = \sum_{i \in I_1} \beta_i \, v_i$$

gilt. Wir wollen zeigen, dass wir v auch ohne Verwendung von v_j linearkombinieren können. Falls $j \notin I_1$ liegt (also v_j ohnehin nicht verwendet wird), dann ist $\langle (v_i)_{i \in I \setminus \{j\}} \rangle$ klar. Falls jedoch $j \in I_1$ liegt, dann ersetze v_j durch $(*)$:

$$\begin{aligned} v &= \sum_{i \in I_1} \beta_i \, v_i \\ &= \sum_{i \in I_1 \setminus \{j\}} \beta_i \, v_i + \beta_j \, v_j \\ &= \sum_{i \in I_1 \setminus \{j\}} \beta_i \, v_i - \sum_{i \in I_0 \setminus \{j\}} \beta_j \, \alpha_j^{-1} \alpha_i \, v_i. \end{aligned}$$

Dies ist eine Darstellung von v als Linearkombination aus $(v_i)_{i \in I}$ ohne Verwendung von v_j. Insgesamt gilt also $v \in \langle (v_i)_{i \in I \setminus \{j\}} \rangle$ und damit Aussage (ii).

Aussage $(ii) \Rightarrow$ Aussage (i): Es sei $j \in I$ ein Index, sodass $\langle (v_i)_{i \in I \setminus \{j\}} \rangle = \langle (v_i)_{i \in I} \rangle$ gilt. Dann ist insbesondere $v_j \in \langle (v_i)_{i \in I \setminus \{j\}} \rangle$, d. h., es gibt eine endliche Teilmenge $I_0 \subseteq I \setminus \{j\}$ und Koeffizienten $\alpha_i \in K$, $i \in I_0$, sodass gilt:

$$v_j = \sum_{i \in I_0} \alpha_i \, v_i.$$

Das heißt aber auch

$$\sum_{i \in I_0} \alpha_i \, v_i - 1 \, v_j = 0,$$

d. h., $(v_i)_{i \in I}$ ist linear abhängig.

Der Beweis für Mengen geht analog. □

Im Folgenden interessieren wir uns vor allem für linear unabhängige Erzeugendensysteme, also solche ohne Redundanz.

Definition 13.8 (Basis)**.**
Es sei V ein Vektorraum.

(*i*) Eine Teilmenge $B \subseteq V$ heißt eine **Basis** (englisch: basis) von V, wenn B linear unabhängig ist und $\langle B \rangle = V$ gilt.

(*ii*) Eine Familie B von Vektoren aus V heißt eine **Basis** (englisch: basis) von V, wenn B linear unabhängig ist und $\langle B \rangle = V$ gilt. △

Beachte: Eine Basis ist also ein linear unabhängiges Erzeugendensystem.

Beispiel 13.9 (Basis)**.**

(*i*) $\emptyset$ ist die einzige Basis des Nullraumes $\{0\}$ über jedem Körper K.

(*ii*) Die Menge $E := \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} \right\}$ ist ein Erzeugendensystem des $\mathbb{R}$-Vektorraumes $\mathbb{R}^2$, jedoch keine Basis, da sie linear abhängig ist; siehe Beispiel 13.5. 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 12.3) ist

$$\{e_1, \ldots, e_n\} \quad \text{mit } e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-ter Eintrag} \tag{13.4}$$

eine Basis, genannt die **kanonische Basis** (englisch: canonical basis), **Standardbasis** (englisch: standard basis) oder **Einheitsbasis** (englisch: unit basis) von K^n.

(*iv*) Es sei $(K, +, \cdot)$ ein Körper und $(K[t], +, \cdot)$ der Polynomraum über K; siehe Beispiel 12.3. Dann ist die Menge der Monome $E = \{1, t, t^2, \ldots\}$ eine Basis von $K[t]$. Die Monome $E = \{1, t, \ldots, t^n\}$ bilden eine Basis von $K_n[t]$, dem Unterraum der Polynome vom Höchstgrad $n \in \mathbb{N}_0$. Diese Basen werden die **Monombasis** (englisch: monomial basis) des jeweiligen Vektorraumes genannt. △

Ende der Vorlesung 17

Satz 13.10 (Charakterisierung von Basen)**.**
Es sei V ein Vektorraum und $B \subseteq V$. Dann sind äquivalent:

(*i*) B ist eine Basis von V.

(*ii*) B ist eine maximale linear unabhängige Teilmenge von V.[13] Das heißt: B ist linear unabhängig, und jede echte Obermenge von B ist linear abhängig.

(*iii*) B ist ein minimales Erzeugendensystem von V.[14] Das heißt: B ist ein Erzeugendensystem, und jede echte Teilmenge von B ist kein Erzeugendensystem.

[13] Genauer: B ist ein maximales Element bzgl. der Mengeninklusion (Definition 5.10) in der Menge der linear unabhängigen Teilmengen von V.

[14] Genauer: B ist ein minimales Element bzgl. der Mengeninklusion (Definition 5.10) in der Menge der Erzeugendensysteme von V.

(*iv*) Jeder Vektor $v \in V$ lässt sich auf eindeutige Weise (bis auf Summanden mit Nullkoeffizienten) aus paarweise verschiedenen Elementen von B linearkombinieren.

Ein analoge Charakterisierung der Basiseigenschaft gilt für Familien B in V.

Beweis. 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 Lemma 13.7 („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 $v \in V$ beliebig. Nach Definition von $\langle B \rangle$ ist klar, dass $\langle B \rangle \supseteq B$ gilt. (**Quizfrage 13.1:** Warum?) Wenn also $v \in B$ ist, dann auch $v \in \langle B \rangle$, was zu zeigen war. Wir können also von $v \in V \setminus B$ ausgehen. Nach Voraussetzung ist $B \cup \{v\}$ als echte Obermenge von B linear abhängig. Es existieren also $n \in \mathbb{N}_0$, $v_i \in B$ und $\alpha_i \in K$, $i = 1, \ldots, n$, sowie $\alpha \in K$, sodass gilt:

$$\sum_{i=1}^{n} \alpha_i \, v_i + \alpha \, v = 0.$$

Dabei ist $\alpha \neq 0$, denn sonst wäre bereits B linear abhängig, was der Voraussetzung widerspricht. Wir erhalten also

$$v = -\sum_{i=1}^{n} \alpha^{-1} \alpha_i \, v_i,$$

d. h., 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 Lemma 13.7 („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 Lemma 13.7 („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 13.6 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 paarweise verschiedenen Elementen von B linearkombinieren. Also ist $\langle B \rangle = V$, und aus Lemma 13.6 folgt, dass B linear unabhängig ist, also eine Basis von V. □

Folgerung.

Verschiedene Basen B_1 und B_2 eines Vektorraumes V sind bzgl. der Halbordnung „$\subseteq$" nicht vergleichbar. D. h., es gilt weder $B_1 \subseteq B_2$ noch $B_2 \subseteq B_1$.

Beweis. 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 13.10. □

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 Raum ist nicht alles), und einem Erzeugendensystem, das aber möglicherweise zu groß (linear abhängig) ist, immer eine Basis liegt.

Satz 13.11 (**Basisergänzungssatz**[15]).
Es sei V ein Vektorraum.

(i) Es sei A eine linear unabhängige Menge von Vektoren aus V und E ein Erzeugendensystem von V mit der Eigenschaft $A \subseteq E$. Dann existiert eine Basis B von V mit $A \subseteq B \subseteq E$.

(ii) Es sei $A = (v_i)_{i \in I_A}$ eine linear unabhängige Familie von Vektoren aus V und $E = (v_i)_{i \in I_E}$ ein Erzeugendensystem von V mit der Eigenschaft $I_A \subseteq I_E$. 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, also den Beweis von Aussage (i).

Wir betrachten die Menge aller linear unabhängigen Teilmengen zwischen A und E, also

$$\mathcal{D} := \{D \subseteq V \mid A \subseteq D \subseteq E, \text{ sodass } D \text{ linear unabhängig ist}\} \subseteq \mathcal{P}(E) \subseteq \mathcal{P}(V).$$

$\mathcal{D}$ ist mit der Mengeninklusion eine halbgeordnete Menge. Wegen $A \in \mathcal{D}$ ist $\mathcal{D} \neq \emptyset$. Ziel ist die Anwendung des Lemmas von Zorn 6.35.

Schritt 1: Jede totalgeordnete Teilmenge $\mathcal{C} \subseteq \mathcal{D}$ besitzt eine obere Schranke S in $\mathcal{D}$:

Wir zeigen, dass $S := \bigcup_{C \in \mathcal{C}} C$ eine obere Schranke (Definition 5.10) der Teilmenge $\mathcal{C}$ in $\mathcal{D}$ ist.

Dazu zeigen wir zunächst, dass S überhaupt Element von $\mathcal{D}$ ist:

(a) Für alle $C \in \mathcal{C} \subseteq \mathcal{D}$ gilt $A \subseteq C \subseteq E$, also auch $A \subseteq \bigcup_{C \in \mathcal{C}} C = S \subseteq E$.

(b) 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 13.4 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 6.35 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 Lemma 13.7 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 $\langle B \rangle = V$. □

[15] englisch: basis extension theorem

Es folgen drei unmittelbare Folgerungen als Spezialfälle des Basisergänzungssatzes 13.11, formuliert nur in der Version für Mengen:

Folgerung 13.12 (Basisexistenzsatz: $A = \emptyset$ und $E = V$)**.**
Jeder Vektorraum V besitzt eine Basis.

Folgerung 13.13 (Basisauswahlsatz: $A = \emptyset$)**.**
Aus jedem Erzeugendensystem E eines Vektorraumes V lässt sich eine Basis auswählen.

Folgerung 13.14 (Basisergänzungssatz: $E = V$)**.**
Jede linear unabhängige Menge A eines Vektorraumes V kann zu einer Basis erweitert werden.

Bemerkung 13.15 (zum Basisergänzungssatzes 13.11)**.**
Der Beweis des Basisergänzungssatzes 13.11 und seiner Folgerungen 13.12 bis 13.14 ist nicht konstruktiv, d. h., wir können ihn nicht zur Grundlage eines Verfahrens machen, um eine Basis zu konstruieren. Beispielsweise können wir keine Basis von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum (Beispiel 12.3) angeben.

Wenn der Vektorraum V jedoch endlich erzeugt ist, wenn es also ein endliches Erzeugendensystem gibt, dann lässt sich der Basisergänzungssatz 13.11 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. △

Wir kommen nun zum wichtigen Begriff der Dimension, der in gewissem Sinne die „Größe" eines Vektorraumes beschreibt. Aus Gründen der Übersichtlichkeit formulieren wir die Definition nur für Basen, die Mengen sind, nicht für Familien.

Definition 13.16 (Dimension eines Vektorraumes)**.**
Es sei V ein Vektorraum.

(i) Wenn V eine Basis B der endlichen Kardinalität $n \in \mathbb{N}_0$ besitzt, so sagen wir, V habe **endliche Dimension** (englisch: finite dimension), genauer: die **Dimension** (englisch: dimension) n, in Symbolen: $\dim(V) = n$.

(ii) Wenn V keine Basis endlicher Kardinalität besitzt, so sagen wir, V habe **unendliche Dimension** (englisch: infinite dimension), in Symbolen: $\dim(V) = \infty$.

Zur Verdeutlichung, welcher Körper K verwendet wird, schreiben wir manchmal auch $\dim_K(V)$. △

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 Basis endlicher Kardinalität besitzt, alle seine Basen endliche Kardinalität haben, und zwar alle dieselbe.

Lemma 13.17 (Austauschlemma)**.**
Es sei V ein Vektorraum über dem Körper K und die endliche Menge $B = \{v_1, \ldots, v_n\}$ eine Basis von V mit $n \in \mathbb{N}$. Ist $w = \sum_{i=1}^n \alpha_i \, v_i$ mit Koeffizienten $\alpha_i \in K$ und gilt $\alpha_j \neq 0$ für ein gewisses $j \in [\![1, n]\!]$, dann ist auch $B_0 \coloneqq \{v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n\}$ eine Basis von V.

Ein analoges Resultat gilt, wenn die endliche Familie B eine Basis von V ist.

Beweis. Da es bei einer Basis auf die Reihenfolge der Elemente nicht ankommt, nehmen wir aus Bequemlichkeit und o. B. d. A. an, dass $j = 1$ ist. Aus $w = \sum_{i=1}^n \alpha_i\, v_i$ folgt

$$v_1 = \alpha_1^{-1}\Big(w - \sum_{i=2}^n \alpha_i\, v_i\Big). \tag{$*$}$$

Schritt 1: Wir zeigen: $\langle B_0 \rangle = V$.

Es sei dazu $v \in V$. Da B eine Basis 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 ($*$) folgt

$$\begin{aligned}
v &= \beta_1\, \alpha_1^{-1}\Big(w - \sum_{i=2}^n \alpha_i\, v_i\Big) + \sum_{i=2}^n \beta_i\, v_i \\
&= \beta_1\, \alpha_1^{-1} w - \beta_1\, \alpha_1^{-1} \sum_{i=2}^n \alpha_i\, v_i + \sum_{i=2}^n \beta_i\, v_i \quad && \text{nach Distributivgesetz (10.1a) in } K \\
&= \beta_1\, \alpha_1^{-1} w + \sum_{i=2}^n \big(\beta_i - \beta_1\, \alpha_1^{-1} \alpha_i\big)\, v_i && \text{nach Kommutativ- und Distributivgesetz (10.1b).}
\end{aligned}$$

Damit ist gezeigt, dass in der Tat $\langle B_0 \rangle = \langle w, v_2, \ldots, v_n \rangle = V$ gilt.

Schritt 2: Wir zeigen: B_0 ist linear unabhängig.

Wir betrachten dazu

$$\begin{aligned}
& \beta_1\, w + \sum_{i=2}^n \beta_i\, v_i = 0 \\
\Rightarrow \quad & \beta_1 \sum_{i=1}^n \alpha_i\, v_i + \sum_{i=2}^n \beta_i\, v_i = 0 \\
\Rightarrow \quad & \beta_1\, \alpha_1\, v_1 + \sum_{i=2}^n \big(\beta_1\, \alpha_i + \beta_i\big)\, v_i = 0.
\end{aligned}$$

Da $B = \{v_1, \ldots, v_n\}$ linear unabhängig ist, ist das nur möglich, wenn alle Koeffizienten gleich Null sind, also

$$\beta_1\, \alpha_1 = 0 \quad \text{und} \quad \beta_1\, \alpha_i + \beta_i = 0 \text{ für } i = 2, \ldots, n.$$

Wegen $\alpha_1 \neq 0$ muss $\beta = 0$ sein, woraus dann weiter $\beta_2 = \ldots = \beta_n = 0$ folgt. Das heißt aber, dass $B_0 = \{w, v_2, \ldots, v_n\}$ linear unabhängig ist. □

Satz 13.18 (**Austauschsatz von Steinitz**[16]).
Es sei V ein Vektorraum und die endliche Menge $B = \{v_1, \ldots, v_n\}$ eine Basis von V mit $\#B = n \in \mathbb{N}_0$. Ist $A = \{a_1, \ldots, a_m\}$ eine 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 $B_0 := A \cup D$ ebenfalls eine Basis von V ist. Es gilt $\#B_0 = n$.*

[16] englisch: Steinitz exchange theorem

Beweis. Wir führen einen Induktionsbeweis nach der Mächtigkeit $m = \#A$. 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 $\{a_1, \ldots, a_m\}$ linear unabhängig ist, ist auch $\{a_1, \ldots, a_{m-1}\}$ linear unabhängig. Nach Induktionsannahme gilt $m - 1 \leqslant n$, und es existiert $D = \{v_{i_m}, \ldots, v_{i_n}\} \subseteq B$, sodass $B_1 \coloneqq \{a_1, \ldots, a_{m-1}, v_{i_m}, \ldots v_{i_n}\}$ eine Basis von V ist.

Falls nun $m - 1 = n$ wäre, also $D = \emptyset$, dann wäre bereits $\{a_1, \ldots a_{m-1}\}$ eine Basis von V. Nach dem Satz 13.10 über die Charakterisierung von Basen hieße das aber, dass $\{a_1, \ldots a_m\}$ linear 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, $j = 1, \ldots n$, sodass gilt:

$$a_m = \sum_{j=1}^{m-1} \alpha_j \, a_j + \sum_{j=m}^{n} \alpha_j \, v_{i_j}.$$

Wären alle $\alpha_m = \alpha_{m+1} = \ldots \alpha_n = 0$, so würde das die lineare Abhängigkeit von $A = \{a_1, \ldots, a_m\}$ zeigen, im Widerspruch zur Voraussetzung. Demzufolge gibt es einen Index $j \in [\![m, n]\!]$ mit $\alpha_j \neq 0$. Nach dem Austauschlemma 13.17 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 13.19 (endliche Basen sind gleichmächtig)**.**
Es sei V ein Vektorraum.

(i) Wenn V endlich erzeugt ist, dann ist jede Basis von V endlich, und alle Basen haben dieselbe Mächtigkeit.

(ii) Wenn V nicht endlich erzeugt ist, dann ist jede Basis von V unendlich.

Beweis. Wir führen den Beweis für Mengen.

Aussage (i): Wenn V endlich erzeugt ist, dann gibt es ein endliches Erzeugendensystem und nach Basisergänzungssatz 13.11 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 13.18 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 B auch ein endliches Erzeugendensystem, und V wäre endlich erzeugt, im Widerspruch zur Voraussetzung. □

Bemerkung 13.20 (Dimensionsbegriff für Vektorräume)**.**

(i) Folgerung 13.19 zeigt, dass der Dimensionsbegriff aus Definition 13.16 wohldefiniert ist.

(*ii*) Der Beweis von Folgerung 13.19 verwendet den Basisergänzungssatz 13.11, jedoch nur die Version für endlich-dimensionale (endlich erzeugte) Vektorräume, die ohne das Zornsche Lemma und damit ohne das Auswahlaxiom auskommt.

(*iii*) Sogar für unendlich-dimensionale Vektorräume gilt, dass je zwei Basen dieselbe Mächtigkeit besitzen. (Dieser Beweis verwendet wieder das Auswahlaxiom.) △

Beispiel 13.21 (Dimension eines Vektorraumes)**.**

(*i*) Der „Standardvektorraum K^n der Dimension n" über einem Körper K (Beispiel 12.3) hat tatsächlich die Dimension $n \in \mathbb{N}_0$.

(*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 Basis für $\dim_{\mathbb{R}}(\mathbb{C})$ ist $\{1, i\}$.

(*iv*) Der Nullraum $\{0\}$ ist über jedem Körper der einzige Vektorraum der Dimension 0.

(*v*) Der Polynomraum $(K[t], +, \cdot)$ über einem Körper $(K, +, \cdot)$ hat unendliche Dimension, da $B = \{1, t, t^2, \ldots\}$ eine (abzählbar) unendliche Basis von $K[t]$ ist. Die Monome $B = \{1, t, \ldots, t^n\}$ bilden eine Basis von $K_n[t]$, dem Unterraum der Polynome vom Höchstgrad $n \in \mathbb{N}_0$. Also gilt $\dim(K_n[t]) = n + 1$. △

Folgerung 13.22 (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 Teilmenge, 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.

(*iii*) Für jeden Unterraum U von V gilt: $0 \leqslant \dim(U) \leqslant \dim(V)$.

(*iv*) Für jeden Unterraum U von V gilt $U = V$ genau dann, wenn $\dim(U) = \dim(V)$ ist.

Beweis. Aussage (*i*): Nach Basisergänzungssatz 13.11 existiert eine Basis B von V mit $A \subseteq B$. Nach Folgerung 13.19 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 13.19 gilt $\#A = \dim(V) = n$. Ist umgekehrt A linear unabhängig und gilt $\#A = n$, so gilt für jede Basis $B \supseteq A$ von V einerseits $\#B \geqslant \#A$, andererseite 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 $U = \langle A \rangle = V$. □

Ende der Vorlesung 18

§ 14 Summen von Unterräumen

Literatur: Bosch, 2014, Kapitel 1, Fischer, Springborn, 2020, Kapitel 2.6

§ 14.1 Summen von zwei Unterräumen

Aus Lemma 12.11 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.[17]

Es ist naheliegend, statt $U \cup W$ den kleinsten Unterraum von V zu betrachten, der $U \cup W$ enthält, also $\langle U \cup W \rangle$.

Lemma 14.1 (die lineare Hülle der Vereinigung zweier Unterräume ist die Summe.)**.**
Es sei V ein Vektorraum und U, W zwei Unterräume von V. Dann gilt

$$\langle U \cup W \rangle = U + W. \tag{14.1}$$

Wir bezeichnen den Unterraum $U + W$ als die **Summe der Unterräume** U und W (englisch: sum of two subspaces). (**Quizfrage 14.1:** Wie ist der Ausdruck $U + W$ zu verstehen?)

Beweis. Aus Satz 12.13 wissen wir, dass $\langle U \cup W \rangle$ übereinstimmt mit der Menge M aller Linearkombinationen von $U \cup W$. Wir zeigen nun, dass $M = U + W$ gilt. In der Tat sind die Elemente $u + w$ von $U + W$ auch Elemente von M, also $U + W \subseteq M$. Ist umgekehrt $v \in M$, dann hat v eine Darstellung der Form

$$v = \sum_{i=1}^{n} \alpha_i \, u_i + \sum_{j=1}^{m} \beta_j \, w_j$$

mit $n, m \in \mathbb{N}_0$, $\alpha_i, \beta_j \in K$ und $u_i \in U$ sowie $w_j \in W$. 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 $M \subseteq U + W$, also insgesamt $M = U + W$. □

Bemerkung 14.2 (Ordnung auf der Menge aller Unterräume)**.**
Die Mengeninklusion ist eine Halbordnung auf der Menge aller Unterräume eines gegebenen (möglicherweise unendlich-dimensionalen) Vektorraumes V. Sie stimmt überein mit der Halbordnung „ist Unterraum von". Für je zwei Unterräume U und W ist $U \cap W$ das Infimum von $\{U, W\}$ (Definition 5.10) und $U + W$ das Supremum von $\{U, W\}$. Genau dann, wenn $U \subseteq W$ oder $W \subseteq U$ gilt, ist das Infimum ein Minimum und das Supremum ein Maximum. △

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

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W). \tag{14.2}$$

[17] Tatsächlich gilt: $U \cup W$ ist genau dann ein Unterraum, wenn $U \subseteq W$ oder $W \subseteq U$ gilt (Hausaufgabe I-8.3). Ein analoges Resultat gilt auch für Untergruppen (Hausaufgabe I-5.1), Unterringe und Unterkörper.

Beweis. $U \cap W$ ist ebenfalls ein endlich-dimensionaler Unterraum von V, besitzt also eine endliche Basis $\{v_1, \dots, v_m\}$ mit $m = \dim(U \cap W) \in \mathbb{N}_0$. Nach dem Basisergänzungssatz 13.11 kann diese Basis zu einer Basis von U bzw. einer Basis von W ergänzt werden. Es gibt also $\{u_1, \dots, u_k\} \subseteq U$ und $\{w_1, \dots, w_\ell\} \subseteq W$ mit $k, \ell \in \mathbb{N}_0$, sodass

$$\begin{aligned} B_U &\coloneqq \{v_1, \dots, v_m,\ u_1, \dots, u_k\} && \text{Basis von } U \\ B_W &\coloneqq \{v_1, \dots, v_m,\ w_1, \dots, w_\ell\} && \text{Basis von } W \end{aligned}$$

ist. Wir zeigen nun, dass $B \coloneqq \{v_1, \dots, v_m,\ u_1, \dots, u_k,\ w_1, \dots, w_\ell\}$ eine Basis von $U + W$ ist. Wegen $B = B_U \cup B_W$ ist B ein Erzeugendensystem 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

$$\begin{aligned} \text{einerseits} \quad & u \in U && = \langle v_1, \dots, v_m,\ u_1, \dots, u_k \rangle \\ \text{und andererseits} \quad & u \in U \cap W && = \langle v_1, \dots, v_m \rangle. \end{aligned}$$

Aus der Eindeutigkeit der Darstellung (Lemma 13.6) folgt $\beta_1 = \ldots = \beta_k = 0$. Es folgt also

$$\sum_{i=1}^{m} \alpha_i\, v_i + \sum_{i=1}^{\ell} \gamma_i\, w_i = 0,$$

und da $B_W = \{v_1, \dots, v_m,\ w_1, \dots, w_\ell\}$ 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

$$\underbrace{m + k + \ell}_{\dim(U+W)} = \underbrace{m + k}_{\dim(U)} \quad + \quad \underbrace{m + \ell}_{\dim(W)} \quad - \quad \underbrace{m}_{\dim(U \cap W)}.$$

□

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
$$U = \langle 1, t^2, t^3, t^5 \rangle \quad \text{und} \quad W = \langle t, t^3, t^4 \rangle$$
von $K_5[t]$ über einem beliebigen Körper $(K, +, \cdot)$ gilt
$$U + W = \langle 1, t, t^2, t^3, t^4, t^5 \rangle = K_5[t] \quad \text{und} \quad U \cap W = \langle t^3 \rangle.$$
Die Dimensionsformel (14.2) ergibt
$$\underbrace{\dim(U + W)}_{6} = \underbrace{\dim(U)}_{4} + \underbrace{\dim(W)}_{3} - \underbrace{\dim(U \cap W)}_{1}.$$
△

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 Symbolen: $U \oplus W$. △

Beispiel 14.6 (direkte Summe von zwei Unterräumen)**.**
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*) $\Longrightarrow$ Aussage (*ii*): $V = U \oplus W$ impliziert $V = U + W$. Für gegebenes $v \in V$ gibt es also Vektoren $u \in U$ und $w \in W$, sodass $v = u + w$ gilt. Gilt nun ebenfalls $v = u' + w'$ für Vektoren $u' \in U$ und $w' \in W$, dann gilt
$$u - u' = w' - w \in U \cap W = \{0\}$$
nach Voraussetzung. Daher muss $u = u'$ und $w = w'$ sein.

Aussage (*ii*) $\Longrightarrow$ Aussage (*i*): Aus der Voraussetzung folgt sofort $V = U + W$. Zu zeigen ist $U \cap W = \{0\}$.

Für $v \in U \cap W$ zeigen

$$\begin{aligned} v &= v + 0 \qquad \text{mit } v \in U,\ 0 \in W \\ v &= 0 + v \qquad \text{mit } 0 \in U,\ v \in W \end{aligned}$$

und die 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): Unter den Voraussetzungen zeigt die Dimensionsformel (14.2):

$$\dim(V) = \dim(U) + \dim(W) = \dim(U + W) + \dim(U \cap W) = \dim(U + W).$$

$U + W$ ist also ein Unterraum von V maximaler Dimension und damit identisch zu V. Außerdem zeigt $\dim(U + W) = 0$, dass $U \cap W = \{0\}$ gilt. □

Satz 14.8 (direkte Summe von zwei Unterräumen und Partitionierung einer Basis)**.**
Es sei V ein Vektorraum. Dann gilt:

(i) Ist B eine Basis von V und $\{B_1, B_2\}$ eine Partition[18] von B, dann gilt $V = \langle B_1 \rangle \oplus \langle B_2 \rangle$.

(ii) Sind U_1, U_2 Unterräume von V mit Basen B_1, B_2 und gilt $V = U_1 \oplus U_2$, so ist $B_1 \cup B_2$ eine Basis von V.

Beweis. Aussage (i): Wir zeigen zunächst $V = \langle B_1 \rangle + \langle B_2 \rangle$. Es gilt

$$\begin{aligned} V &= \langle B \rangle && B \text{ ist Basis von } V \\ &= \langle B_1 \cup B_2 \rangle && \text{nach Voraussetzung} \\ &= \big\langle \langle B_1 \rangle \cup \langle B_2 \rangle \big\rangle && \text{nach Folgerung 12.15} \\ &= \langle B_1 \rangle + \langle B_2 \rangle && \text{nach Lemma 14.1 (lineare Hülle der Vereinigung zweier Unterräume)} \\ &\subseteq V. \end{aligned}$$

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 $v \in \langle B_1 \rangle \cap \langle B_2 \rangle$ an, d. h.

$$v = \sum_{i=1}^{n} \alpha_i \, b_i^1 = \sum_{i=1}^{m} \beta_i \, b_i^2$$

für geeignete endliche Teilmengen $\{b_1^1, \ldots, b_n^1\} \subseteq B_1$ und $\{b_1^2, \ldots, b_m^2\} \subseteq B_2$ und Koeffizienten α_i, β_i. Es folgt

$$0 = \sum_{i=1}^{n} \alpha_i \, b_i^1 + \sum_{i=1}^{m} (-\beta_i) \, b_i^2.$$

Da $B = B_1 \cup B_2$ eine Basis ist, ist auch die Teilmenge $\{b_1^1, \ldots, b_n^1,\ b_1^2, \ldots, b_m^2\}$ linear unabhängig. Daraus folgt $\alpha_1 = \ldots = \alpha_n = \beta_1 = \ldots = \beta_m = 0$. Das heißt aber $v = 0$ und damit $\langle B_1 \rangle \cap \langle B_2 \rangle = \{0\}$.

[18] Zur Erinnerung, das heißt $B_1, B_2 \neq \emptyset$, $B = B_1 \cup B_2$ und $B_1 \cap B_2 = \emptyset$, vgl. Definition 5.18. Hier wäre sogar $B_1 = \emptyset$ oder $B_2 = \emptyset$ erlaubt.

Aussage (ii): Es seien nun U_1, U_2 Unterräume von V mit Basen 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

$$\begin{aligned}
\langle B_1 \cup B_2 \rangle &= \big\langle \langle B_1 \rangle \cup \langle B_2 \rangle \big\rangle && \text{nach Folgerung 12.15} \\
&= \langle U_1 \cup U_2 \rangle && \text{da } B_1 \text{ Basis von } U_1 \text{ und } B_2 \text{ Basis von } U_2 \text{ ist} \\
&= U_1 + U_2 && \text{nach Lemma 14.1 (lineare Hülle der Vereinigung zweier Unterräume)} \\
&= V && \text{nach Voraussetzung.}
\end{aligned}$$

Es seien nun $\{b_1^1, \ldots, b_n^1\} \subseteq B_1$ und $\{b_1^2, \ldots, b_m^2\} \subseteq B_2$ endliche Teilmengen mit $n, m \in \mathbb{N}_0$. Wir betrachten die Linearkombination

$$\sum_{i=1}^n \alpha_i \, b_i^1 + \sum_{i=1}^m \beta_i \, b_i^2 = 0 \quad \text{bzw.} \quad \sum_{i=1}^n \alpha_i \, b_i^1 = \sum_{i=1}^m (-\beta_i) \, b_i^2.$$

Die erste Linearkombination gehört zu U_1, die zweite zu U_2. Wegen $U_1 \cap U_2 = \{0\}$ ist jede Linearkombination der Nullvektor. Da B_1 und B_2 Basen sind, gilt $\alpha_1 = \ldots = \alpha_n = \beta_1 = \ldots = \beta_m = 0$. Das heißt, dass $B_1 \cup B_2$ in der Tat linear unabhängig ist. □

Folgerung 14.9 (Existenz eines komplementären Unterraumes[19])**.**
Es seien V ein Vektorraum und U ein Unterraum von V. Dann existiert ein weiterer Unterraum W von V, sodass gilt: $V = U \oplus W$.

Beweis. Es sei B_U eine Basis von U. Aus dem Basisergänzungssatz 13.11 folgt die Existenz einer Basis B von V mit $B_U \subseteq B$. Setzen wir $B_W := B \setminus B_U$ und $W := \langle B_W \rangle$, so ist W nach Satz 14.8 (i) ein Unterraum mit der gesuchten Eigenschaft. □

Definition 14.10 (komplementärer Unterraum)**.**
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 $U = V$ 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.

Quizfrage 14.3: Was gilt für $\operatorname{codim}(U)$ in einem endlich-dimensionalen Vektorraum?

Beispiel 14.11 (komplementärer Unterraum)**.**

(i) Jeder eindimensionale Unterraum W, der nicht identisch mit $U = \big\langle \binom{1}{1} \big\rangle$ ist, ist ein Komplement von U in $\mathbb{R}^2$.

(ii) Die komplementären Unterräume des Unterraumes $(\mathbb{R}^\mathbb{N})_c$ der konvergenten Folgen im Vektorraum $(\mathbb{R}^\mathbb{N}, +, \cdot)$, siehe Beispiel 12.9, haben keine einfache Darstellung. △

[19] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

§ 14.2 Summen von Familien von Unterräumen

Wir beschäftigen uns abschließend noch mit Summen von einer beliebigen Anzahl von Unterräumen eines Vektorraumes.

Definition 14.12 (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

$$\sum_{i\in I} U_i := \Big\langle \bigcup_{i\in I} U_i \Big\rangle \tag{14.3}$$

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 heißt **direkt** (englisch: direct sum), wenn gilt:

$$U_j \cap \sum_{i\in I\setminus\{j\}} U_i = \{0\} \quad \text{für alle } j \in I. \tag{14.4}$$

Wir schreiben dann $\bigoplus_{i\in I} U_i$. Im Fall $I = [\![1,n]\!]$ schreiben wir auch $U_1 \oplus \cdots \oplus U_n$ oder $\bigoplus_{i=1}^n U_i$. △

Beispiel 14.13 (Summe einer Familie von Unterräumen)**.**

(i) Für den Standardvektorraum K^n über einem Körper K mit den Basisvektoren e_i, $i = 1, \dots, n \in \mathbb{N}$, siehe Beispiel 13.9, gilt

$$K^n = \bigoplus_{i=1}^n \langle e_i \rangle.$$

(ii) Für den Polynomraum $(K[t], +, \cdot)$ über einem Körper K mit der Basis $B = \{1, t, t^2, \dots\}$, siehe Beispiel 13.21, gilt

$$K[t] = \bigoplus_{i\in\mathbb{N}_0} \langle t^i \rangle.$$ △

Bemerkung 14.14 (Summe einer Familie von Unterräumen)**.**

(i) Die Elemente der Summe einer Familie $(U_i)_{i\in I}$ von Unterräumen haben die Darstellung

$$\begin{aligned}\sum_{i\in I} U_i &= \bigcup \Big\{ \sum_{i\in I_0} U_i \,\Big|\, I_0 \subseteq I \text{ ist eine endliche Teilfamilie} \Big\} \\ &= \Big\{ \sum_{i\in I_0} u_i \,\Big|\, I_0 \subseteq I \text{ ist eine endliche Teilfamilie und } u_i \in U_i \text{ für alle } i \in I_0 \Big\}.\end{aligned} \tag{14.5}$$

(ii) Im Fall $\#I = 2$ stimmt die Bedingung (14.4), dass die Summe einer Familie von Unterräumen eine direkte Summe ist, überein mit Definition 14.5. Im Fall $\#I > 2$ reicht es jedoch für die Direktheit der Summe nicht aus, dass an Stelle von (14.4) 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 14.15 (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 Teilfamilie $I_0 \subseteq I$ 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 von Hausaufgabe I-9.4. □

Satz 14.16 (direkte Summe von Unterräumen und Partitionierung einer Basis, vgl. Satz 14.8)**.**
Es sei V ein Vektorraum. Dann gilt:

(i) Ist B eine Basis von V und $(B_i)_{i \in I}$ eine Partition von B 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 Basen B_i, $i \in I$, und gilt $V = \bigoplus_{i \in I} U_i$, so ist $\bigcup_{i \in I} B_i$ eine Basis von V.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-9.4. □

Ende der Vorlesung 19

Ende der Woche 9

Kapitel 4 Matrizen und lineare Abbildungen

§ 15 Matrizen

Literatur: Fischer, Springborn, 2020, Kapitel 3.7, Bosch, 2014, Kapitel 3, Deiser, 2022b, Kapitel 3.3, Jänich, 2008, Kapitel 4.2 und 5.1–5.5

Matrizen sind ein universelles Mittel zur Darstellung verschiedener Sachverhalte, beispielsweise in den Wirtschaftswissenschaften, in der Graphentheorie und zur Beschreibung linearer Abbildungen, das sind die Homomorphismen zwischen Vektorräumen. Diese Bedeutung von Matrizen 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 der Dimension** $n \times m$ (englisch: matrix) oder eine $n \times m$**-Matrix** A **über dem Körper** K ist eine endliche, doppelt indizierte Familie von Elementen aus K mit Indexmenge $⟦1, n⟧ \times ⟦1, m⟧$. Wir schreiben sie in der Form[1]

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}. \tag{15.1}$$

Die Menge aller $n \times m$-Matrizen wird mit $K^{n \times m}$ bezeichnet.[2]

(*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.[3]

(*iv*) Eine Matrix heißt **quadratisch** (englisch: square matrix, quadratic matrix), wenn $m = n$ gilt.

(*v*) Die quadratische $n \times n$-Diagonalmatrix

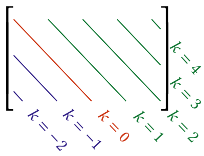
$$I_{n \times n} := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{15.2}$$

[1] Alternativ können auch runde Klammern verwendet werden.
[2] Alternative Bezeichnungen sind $K^{n,m}$ oder $M_{n,m}(K)$.
[3] Man sagt auch, dass bei einer Diagonalmatrix die Nebendiagonalen „nicht besetzt“ sind, d. h., dass dort nur Nullen stehen.

heißt die $n \times n$-**Einheitsmatrix** (englisch: identity matrix).[4] Wir bezeichnen sie auch mit I_n oder einfach I, wenn die Dimension klar ist. △

Beachte: Wir lassen explizit zu, dass eine oder beide Dimensionen gleich 0 sind. Das ist aus vielerlei Gründen praktisch. Eine Matrix der Dimension $n \times 0$ oder $0 \times m$ hat keine Elemente, 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.[5]

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

$$a_{i\bullet} \coloneqq \big(a_{i1}, a_{i2}, \ldots, a_{im}\big) \in K_m, \tag{15.4}$$

vgl. Beispiel 12.3. 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

$$a_{\bullet j} \coloneqq \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \in K^n, \tag{15.5}$$

vgl. nochmals Beispiel 12.3. 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 skalare Multiplikation (S-Multiplikation) wie folgt:

Definition 15.2 (Addition, skalare 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**[6], englisch: addition) und die äußere Verknüpfung $\cdot\colon K \times K^{n \times m} \to K^{n \times m}$ (**skalare Multiplikation**[7], englisch: scalar multiplication) erklärt:

$$(A + B)_{ij} \coloneqq a_{ij} + b_{ij} \tag{15.6a}$$
$$(\alpha \cdot A)_{ij} \coloneqq \alpha \cdot a_{ij} \tag{15.6b}$$

für $A, B \in K^{n \times m}$ und $\alpha \in K$. △

[4] Alternative Bezeichnungen sind I_n oder id_n.
[5] Aus Gründen der Verdeutlichung schreiben wir manchmal auch $a_{i,j}$ an Stelle von a_{ij}.
[6] Die Bezeichnung ist dieselbe wie die der „Addition“ im Körper K.
[7] Auch hier ist die Bezeichnung dieselbe wie die der „Multiplikation“ im Körper K.

Wir werden das Zeichen $\cdot$ für die skalare Multiplikation in der Regel weglassen, vgl. Bemerkung 12.10.

Lemma 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 Vektorraum über K. Dieser besitzt die Dimension $n\,m$, und die Matrizen

$$\{E_{11},\dots,E_{nm}\}\ \text{ mit } E_{ij}\coloneqq \begin{array}{cl} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{array}{l} \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \end{array} \\ \quad\quad\uparrow j\text{-te Spalte} & \end{array} = (\delta_{ik}\,\delta_{j\ell})_{k=1,\dots,n,\ \ell=1,\dots,m} \tag{15.7}$$

bilden eine Basis, genannt die **kanonische Basis** (englisch: canonical basis), **Standardbasis** (englisch: standard basis) oder **Einheitsbasis** (englisch: unit basis) von $K^{n\times m}$.

Beweis. Die Eigenschaften eines Vektorraumes (Definition 12.1) sind leicht nachzurechnen: $(K^{n\times m},+)$ 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 13.10 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 der Basiselemente. □

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 Vektorraum über K. Dieser kann identifiziert werden mit K selbst.

(*iii*) Die Vektorräume der $0\times m$- und der $n\times 0$-Matrizen sind nulldimensionale Vektorräume über K. △

§ 15.1 Matrix-Matrix-Multiplikation

Für Matrizen passender Dimensionen können wir eine Matrix-Matrix-Multiplikation einführen. Diese ist von fundamentaler Bedeutung für den Umgang mit Matrizen.

Definition 15.5 (Matrix-Matrix-Multiplikation)**.**
Es seien $(K, +, \cdot)$ ein Körper und $m, n, \ell \in \mathbb{N}_0$. Für Matrizen ist die **Matrix-Matrix-Multiplikation** (englisch: matrix-matrix multiplication) oder kurz **Matrix-Multiplikation** (englisch: matrix multiplication) wie folgt definiert:

$$\cdot : K^{n\times m} \times K^{m\times \ell} \to K^{n\times \ell}, \tag{15.9a}$$

wobei für $A \in K^{n\times m}$, $B \in K^{m\times \ell}$ und $C := A \cdot B \in K^{n\times \ell}$ gilt:

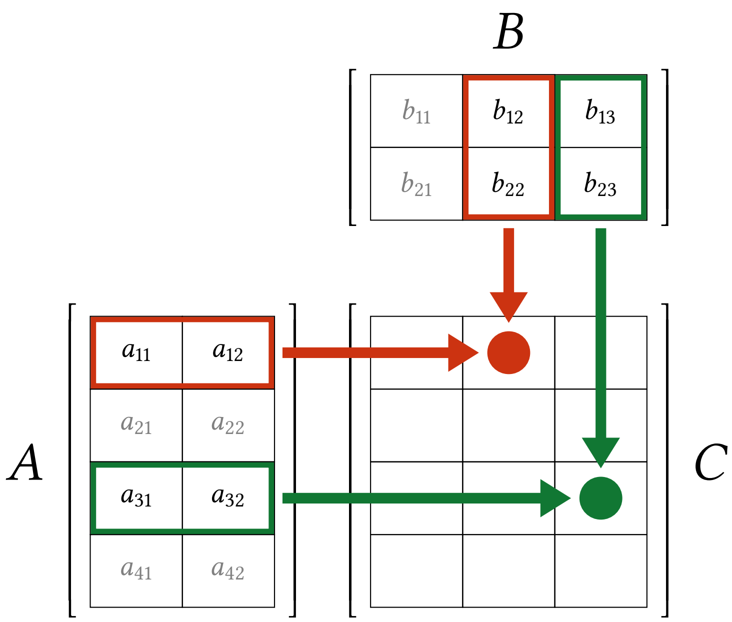
$$c_{ik} := \sum_{j=1}^{m} a_{ij} \cdot b_{jk} \quad \text{für } 1 \leqslant i \leqslant n \text{ und } 1 \leqslant k \leqslant \ell. \tag{15.9b}$$

△

Beachte: Die Summation verwendet die „Addition“ + aus dem Körper K, und in jedem Summanden kommt die „Multiplikation“ · aus K vor.

Beachte: Im Fall $m = 0$ sind die Summen in (15.9b) alle leer. Daher ist das Produkt einer $n \times 0$- und einer $0 \times m$-Matrix die $n \times m$-Nullmatrix.

Den Fall der Matrix-Multiplikation einer 4×2-Matrix A mit einer 2×3-Matrix B können wir wie folgt grafisch darstellen:



$$c_{12} := a_{11} \cdot b_{12} + a_{12} \cdot b_{22}$$

$$c_{33} := a_{31} \cdot b_{13} + a_{32} \cdot b_{23}$$

Beispiel 15.6 (Matrix-Multiplikation).

$$
\begin{aligned}
&\begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 & -4 \\ 0 & -2 & 8 \end{bmatrix} \\
&= \begin{bmatrix} 1 \cdot 5 + (-3) \cdot 0 & 1 \cdot 2 + (-3) \cdot (-2) & 1 \cdot (-4) + (-3) \cdot 8 \\ 2 \cdot 5 + 4 \cdot 0 & 2 \cdot 2 + 4 \cdot (-2) & 2 \cdot (-4) + 4 \cdot 8 \\ 5 \cdot 5 + 0 \cdot 0 & 5 \cdot 2 + 0 \cdot (-2) & 5 \cdot (-4) + 0 \cdot 8 \\ (-3) \cdot 5 + (-6) \cdot 0 & (-3) \cdot 2 + (-6) \cdot (-2) & (-3) \cdot (-4) + (-6) \cdot 8 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 8 & -28 \\ 10 & -4 & 24 \\ 25 & 10 & -20 \\ -15 & 6 & -36 \end{bmatrix}.
\end{aligned}
$$

△

Bemerkung 15.7 (Matrix-Multiplikation).

(i) $A \cdot B$ ist genau dann definiert, wenn die Anzahl der Spalten des linken Faktors A übereinstimmt mit der Anzahl der Zeilen des rechten Faktors B.

(ii) Das Produkt $A \cdot B$ hat soviele Zeilen wie der linke Faktor A und soviele Spalten wie der rechte Faktor B. Die „mittlere Dimension" ist nach der Bildung des Produkts $A \cdot B$ nicht mehr sichtbar.

(iii) 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.

(iv) 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 ergibt sich zwei weitere Sichtweisen auf das Produkt $A \cdot B$:

(1) Die Spalten von $C = A \cdot B$ sind Linearkombinationen der Spalten von 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.

Im Beispiel oben 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 von 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.

Im Beispiel oben 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 $(K, +, \cdot)$ ein Körper und $m, n, p, q \in \mathbb{N}_0$. Für Matrizen $A, A_1, A_2 \in K^{n\times m}$, $B, B_1, B_2 \in K^{m\times p}$, $C \in K^{p\times q}$ und Skalare $\alpha \in K$ gelten die folgenden Eigenschaften:

$$A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2 \quad \textbf{Distributivgesetz}^{8} \tag{15.10a}$$
$$(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B \quad \textbf{Distributivgesetz} \tag{15.10b}$$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \quad \textbf{Assoziativgesetz}^{9} \tag{15.11}$$

$$A \cdot (\alpha \cdot B) = (\alpha \cdot A) \cdot B = \alpha \cdot (A \cdot B) \quad \text{Skalare können überall stehen} \tag{15.12}$$

$$I_n \cdot A = A \cdot I_m = A \quad \text{Neutralität der Einheitsmatrix} \tag{15.13}$$

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

$$\begin{aligned}\big[A \cdot (\alpha \cdot B)\big]_{ik} &= \sum_{j=1}^{m} a_{ij}\,(\alpha \cdot B)_{jk} = \sum_{j=1}^{m} a_{ij}\,(\alpha\, b_{jk}) = \sum_{j=1}^{m} (\alpha\, a_{ij})\, b_{jk} = \big[(\alpha \cdot A) \cdot B\big]_{ik} \\ &= \alpha \sum_{j=1}^{m} a_{ij}\, b_{jk} = \alpha\, \big[A \cdot B\big]_{ik},\end{aligned}$$

[8] englisch: distributive law
[9] englisch: associative law

also (15.12). Schließlich haben wir

$$\begin{aligned}\big[I_n \cdot A\big]_{ik} &= \sum_{j=1}^{m} (I_n)_{ij}\, a_{jk} = \sum_{j=1}^{m} \delta_{ij}\, a_{jk} = a_{ik}\\ \text{und}\quad \big[A \cdot I_m\big]_{ik} &= \sum_{j=1}^{m} a_{ij} (I_m)_{jk} = \sum_{j=1}^{m} a_{ij}\, \delta_{jk} = a_{ik},\end{aligned}$$

also (15.13). □

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 $A \in K^{n\times m}$ und $x \in K^m$ (aufgefasst als $m \times 1$-Matrix) ist. Beispielsweise ist

$$\begin{bmatrix} 1 & -3 \\ 2 & 4 \\ 5 & 0 \\ -3 & -6 \end{bmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 8 \\ -4 \\ 10 \\ 6 \end{pmatrix}.$$

Ein zweiter Spezialfall ist die **Vektor-Matrix-Multiplikation** (englisch: vector-matrix multiplication) $y\,A$, wobei $A \in K^{n\times m}$ und $y \in K_n$ (aufgefasst als $1 \times n$-Matrix) ist. 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 20

§ 15.2 Zeilen- und Spaltenraum

Definition 15.10 (Zeilen- und Spaltenraum, Zeilen- und Spaltenrang)**.**
Es seien $(K, +, \cdot)$ 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

$$\operatorname{SRang}(A) \coloneqq \dim(\operatorname{SR}(A)). \tag{15.15b}$$

△

Beachte: Wir könnten äquivalent den Zeilenrang auch definieren als die maximale Anzahl linear unabhängiger Zeilen von A und den Spaltenrang als die maximale Anzahl linear unabhängiger Spalten von A. (**Quizfrage 15.1:** Warum gilt das?)

Beispiel 15.11 (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

$$\begin{aligned} &\operatorname{ZR}(A) = \Big\langle \overbrace{\begin{pmatrix} 2 & 3 & -1 \end{pmatrix}, \begin{pmatrix} 7 & 4 & 0 \end{pmatrix}}^{\text{linear unabhängig}} \Big\rangle && \text{mit } \operatorname{ZRang}(A) = \dim(\operatorname{ZR}(A)) = 2 \\ \text{und}\quad &\operatorname{SR}(A) = \Big\langle \underbrace{\begin{pmatrix} 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}}_{\text{linear abhängig}} \Big\rangle = \Big\langle \underbrace{\begin{pmatrix} 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}}_{\text{linear unabhängig}} \Big\rangle && \text{mit } \operatorname{SRang}(A) = \dim(\operatorname{SR}(A)) = 2. \end{aligned}$$

△

Die hier beobachtete Übereinstimmung von Zeilen- und Spaltenrang ist kein Zufall:

Satz 15.12 (Zeilenrang gleich Spaltenrang).
Es seien $(K, +, \cdot)$ 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. **Schritt 1:** Wir zeigen $\operatorname{SRang}(A) \leqslant \operatorname{ZRang}(A)$.

Es sei $r := \operatorname{ZRang}(A) \in \mathbb{N}_0$ und $C \in K^{r\times m}$ eine Matrix, deren linear unabhängige Zeilen $c_{1\bullet}, \ldots, c_{r\bullet} \in K_m$ eine Basis des Zeilenraumes $\operatorname{ZR}(A)$ bilden, also $\langle c_{1\bullet}, \ldots, c_{r\bullet}\rangle = \operatorname{ZR}(A)$. 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

$$a_{i\bullet} = b_{i1}\, c_{1\bullet} + \cdots + b_{ir}\, c_{r\bullet}.$$

Da das Gesagte für jede Zeile von A gilt, erhalten wir die Darstellung

$$A = B\,C = \begin{bmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{ir} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nr} \end{bmatrix} \begin{bmatrix} \text{——} & c_{1\bullet} & \text{——} \\ & \vdots & \\ \text{——} & c_{r\bullet} & \text{——} \end{bmatrix}$$

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

$$\begin{aligned} & \operatorname{SR}(A) \subseteq \operatorname{SR}(B) \\ \Rightarrow\quad & \operatorname{SRang}(A) \leqslant \operatorname{SRang}(B) \leqslant r = \operatorname{ZRang}(A) \quad \text{nach Folgerung 13.22.} \end{aligned}$$

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 $b_{\bullet 1}, \dots b_{\bullet s} \in K^n$ eine Basis des Spaltenraumes $\operatorname{SR}(A)$ bilden, also $\langle b_{\bullet 1}, \dots b_{\bullet s} \rangle = \operatorname{SR}(A)$. 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 von A gilt, erhalten wir die Darstellung

$$A = B\,C = \begin{bmatrix} | & & | \\ b_{\bullet 1} & \cdots & b_{\bullet s} \\ | & & | \end{bmatrix} \begin{bmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1m} \\ \vdots & & \vdots & & \\ c_{s1} & \cdots & c_{sj} & \cdots & c_{sm} \end{bmatrix}$$

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

$$\begin{aligned} & \operatorname{ZR}(A) \subseteq \operatorname{ZR}(C) \\ \Rightarrow \quad & \operatorname{ZRang}(A) \leqslant \operatorname{ZRang}(C) \leqslant s = \operatorname{SRang}(A) \quad \text{nach Folgerung 13.22.} \end{aligned}$$

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 $\operatorname{SRang}(A) \leqslant m$ folgt die Behauptung (15.16). □

Da der Zeilenrang und Spaltenrang einer Matrix übereinstimmen, sprechen wir ab sofort nur noch vom **Rang** (englisch: rank), bezeichnet mit $\operatorname{Rang}(A)$. Als direktes Resultat aus dem Beweis von Satz 15.12 halten wir fest:

Folgerung 15.13 (Rangfaktorisierung)**.**

Es seien $(K, +, \cdot)$ ein Körper, $m, n \in \mathbb{N}_0$ und $A \in K^{n \times m}$. Ist $r = \operatorname{Rang}(A) \in \mathbb{N}_0$, dann existieren Matrizen $B \in K^{n \times r}$ und $C \in K^{r \times m}$, sodass gilt:

$$A = B\,C. \tag{15.17}$$

Die Spalten von B bilden eine Basis von $\operatorname{SR}(A)$. Die Zeilen von C bilden eine Basis von $\operatorname{ZR}(A)$.

Eine solche Faktorisierung der Matrix A, bei der die inneren Dimensionen der Faktoren mit dem $\operatorname{Rang}(A)$ übereinstimmt, heißt eine **Rangfaktorisierung** (englisch: rank factorization) von A.

Satz 15.14 (Rang des Produkts von Matrizen)**.**

Es seien $(K, +, \cdot)$ 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 15.12: 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 $\operatorname{Rang}(A\,B) = \operatorname{ZRang}(A\,B) \leqslant \operatorname{ZRang}(B) = \operatorname{Rang}(B)$.

Die zweite Ungleichung folgt sofort aus $\operatorname{Rang}(A) \leqslant \min\{m, n\}$ und $\operatorname{Rang}(B) \leqslant \min\{\ell, m\}$, siehe Satz 15.12. □

§ 15.3 Zeilenstufenform

Wir geben in diesem Abschnitt eine konstruktive Möglichkeit an, eine Rangfaktorisierung 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 15.15 (elementare Zeilenumformungen, Elementarmatrizen)**.**
Es seien $(K, +, \cdot)$ ein Körper, $m, n \in \mathbb{N}_0$ und

$$A = \begin{bmatrix} \text{---} & a_{1\bullet} & \text{---} \\ & \vdots & \\ \text{---} & a_{i\bullet} & \text{---} \\ & \vdots & \\ \text{---} & a_{n\bullet} & \text{---} \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Unter **elementaren Zeilenumformungen** (englisch: elementary row operations) versteht man die folgenden Operationen, angewendet auf die Matrix A:

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

$$D \coloneqq \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & & \\ & & & \alpha & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} = I + (\alpha - 1)\,E_{ii}. \tag{15.19}$$

Es gilt (nur die Änderungen gegenüber A werden hervorgehoben)

$$D\,A = \begin{bmatrix} & \vdots & \\ \text{---} & \alpha\,a_{i\bullet} & \text{---} \\ & \vdots & \end{bmatrix}.$$

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[10]

$$S := \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & \alpha & 1 & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = I + \alpha\, E_{ij}. \tag{15.20}$$

Es gilt

$$S\,A = \begin{bmatrix} & \vdots & \\ \text{---} & a_{j\bullet} & \text{---} \\ & \vdots & \\ \text{---} & a_{i\bullet} + \alpha\, a_{j\bullet} & \text{---} \\ & \vdots & \end{bmatrix}.$$

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

$$T := \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & 1 & & & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{bmatrix} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}. \tag{15.21}$$

Es gilt

$$T\,A = \begin{bmatrix} & \vdots & \\ \text{---} & a_{j\bullet} & \text{---} \\ & \vdots & \\ \text{---} & a_{i\bullet} & \text{---} \\ & \vdots & \end{bmatrix}.$$

Die Matrizen D, S und T heißen **Elementarmatrizen** (englisch: elementary matrices) vom Typ I, Typ II bzw. Typ III. Die Matrizen T (Typ III) heißen genauer auch **Transpositionsmatrizen** (englisch: transposition matrices) △

Lemma 15.16 (elementare Zeilenumformungen ändern den Zeilenraum und den Zeilenrang nicht).
Es seien $(K, +, \cdot)$ 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 $\operatorname{Rang}(C) = \operatorname{Rang}(A)$.

[10] Die Illustration zeigt den Fall $i > j$.

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 15.17 (Zeilenstufenform)**.**
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 $j_i \coloneqq \min\big\{ j \in [\![1, m]\!] \,\big|\, a_{ij} \neq 0 \big\}$ 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. △

Bemerkung 15.18 (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 weiter nach rechts und können dabei auch Spalten überspringen.

(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 $K \setminus \{0\}$ durchmultiplizieren können. Außerdem können wir zu einer Zeile ein Vielfaches einer weiter unten stehenden Zeile addieren, ohne die Zeilenstufenform zu stören. △

Beispiel 15.19 (Zeilenstufenform)**.**
Hier sind die Besetzungsmuster einiger 3×4-Matrizen in Zeilenstufenform, wobei $\star$ jeweils für einen Eintrag ungleich 0 steht (die Pivot-Elemente) und $?$ für einen beliebigen Eintrag aus dem Körper K.

$$\begin{array}{rc} j_1 = 1 \longrightarrow & \\ j_2 = 2 \longrightarrow & \\ j_3 = 3 \longrightarrow & \end{array}\!\!\! \begin{array}{c} \begin{bmatrix} \star & ? & ? & ? \\ 0 & \star & ? & ? \\ 0 & 0 & \star & ? \end{bmatrix} \\ r = 3 \end{array} \quad \begin{array}{r} j_1 = 1 \longrightarrow \\ j_2 = 3 \longrightarrow \\ \end{array}\!\! \begin{array}{c} \begin{bmatrix} \star & ? & ? & ? \\ 0 & 0 & \star & ? \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ r = 2 \end{array} \quad \begin{array}{r} j_1 = 3 \longrightarrow \\ j_2 = 4 \longrightarrow \\ \end{array}\!\! \begin{array}{c} \begin{bmatrix} 0 & 0 & \star & ? \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ r = 2 \end{array}$$ △

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. Der Spaltenindex sei j_1. Bringe es durch Zeilentausch (elementare Zeilenumformung vom Typ III) an die 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 $j_1 + 1$.

Das vollständige Verfahren kann wie folgt angegeben werden:

Algorithmus 15.20 (Erzeugen der Zeilenstufenform)**.**

Eingabe: Matrix $A \in K^{n \times m}$

Ausgabe: Matrix $C \in K^{n \times m}$ in Zeilenstufenform mit $\mathrm{ZR}(A) = \mathrm{ZR}(C)$ // Die ersten r Zeilen von C bilden eine Basis von $\mathrm{ZR}(A)$.

Ausgabe: $r = \mathrm{Rang}(A) = \mathrm{Rang}(C)$

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 $j \leftarrow \min\big\{ j_r + 1 \leqslant j \leqslant m \,\big|\, (C)_{r+1 \leqslant i \leqslant n,\ j} \big\}$ // erste Nicht-Nullspalte der Restmatrix
  Setze $i \leftarrow \min\big\{ r + 1 \leqslant i \leqslant n \,\big|\, c_{ij} \neq 0 \big\}$ // 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
  Tausche in der Matrix C die Zeilen i und r // Pivot-Element kommt nach oben (Typ III)
  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 unterhalb des Pivot-Elements (r, j) (Typ II)
  end for
end while
return C und r // C ist eine Zeilenstufenform von A und $r = \mathrm{Rang}(A)$

Mit Hilfe von Algorithmus 15.20 können wir folgenden Satz konstruktiv beweisen:

Satz 15.21 (Erreichbarkeit der Zeilenstufenform)**.**
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. Es sei $r \in [\![0, m]\!]$ die Anzahl der Nicht-Nullzeilen in C. Dann bilden die Zeilenvektoren $c_{1\bullet}, \ldots, c_{r\bullet}$ eine Basis von $\mathrm{ZR}(A) = \mathrm{ZR}(C)$. Für den Rang gilt $\mathrm{Rang}(A) = \mathrm{Rang}(C) = r$.

Beispiel 15.22 (Erreichbarkeit der Zeilenstufenform)**.**
Wir betrachten ein Beispiel in $\mathbb{R}^{3 \times 4}$:

$$\begin{bmatrix} 0 & 0 & 3 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 3 & 0 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 3 & 0 & 2 \end{bmatrix} \quad \text{Tauschen der Zeilen 1 und 2}$$

$$\begin{array}{c} \\ \\ -3 \end{array} \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 3 & 0 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -6 & 2 \end{bmatrix} \quad \begin{array}{l} \text{Erzeugen von Nullen unterhalb des Pivot-Elements} \\ \text{Addition des } (-3)\text{-Fachen der Zeile 1 zur Zeile 3} \end{array}$$

$$\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -6 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Erzeugen von Nullen unterhalb des Pivot-Elements
Addition des 2-Fachen der Zeile 2 zur Zeile 3.

Die Zeilenstufenform ist erreicht. Der Zeilenraum von A ist also

$$\operatorname{ZR}(A) = \big\langle \begin{pmatrix} 0 & 1 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 3 & -1 \end{pmatrix} \big\rangle \quad \text{mit } \dim(\operatorname{ZR}(A)) = \operatorname{Rang}(A) = 2. \qquad \triangle$$

Bemerkung 15.23 (Berechnung einer Rangfaktorisierung).
Wir hatten eingangs des Abschnitts angekündigt, eine konstruktive Möglichkeit anzugeben, eine Rangfaktorisierung $A = B\,C$ einer Matrix $A \in K^{n \times m}$ zu bestimmen, sodass also $B \in K^{n \times r}$, $C \in K^{r \times m}$ 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; die resultierende Matrix sei $\overline{C} \in K^{r \times m}$. Wie aber kommen wir an den linken Faktor B?

Die Zeilenstufenform entstand durch Multiplikation mit Elementarmatrizen $E_j \in K^{n \times n}$:

$$E_k \cdots E_2\, E_1\, A = C. \tag{15.22}$$

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[11], 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 (Hausaufgabe I-10.3). Setzen wir nun zur Abkürzung $B \coloneqq E_1'\, E_2' \cdots E_k'$, so erhalten wir die Faktorisierung

$$A = {}_n\underbrace{\left[\begin{array}{c|c} \overset{r}{\overline{B}} & \overset{n-r}{} \end{array}\right]}_{B} \underbrace{\left[\begin{array}{c} \overset{m}{\overline{C}} \\ \hline 0 \end{array}\right]}_{C} \begin{array}{l} {}_r \\ {}_{n-r} \end{array}$$

Da bei der Bildung des Matrix-Produkts $A = B\,C$ die letzten $n - r$ Spalten von B immer nur mit Null-Koeffizienten multipliziert werden, können wir, ohne das Ergebnis zu ändern, den rechten Faktor C durch seine oberen r Zeilen $\overline{C}$ und den linken Faktor B durch seine linken r Spalten $\overline{B}$ ersetzen. So erhalten wir die gewünschte Rangfaktorisierung

$$A = \overline{B}\,\overline{C}$$

mit $\overline{B} \in K^{n \times r}$ und $\overline{C} \in K^{r \times m}$. (**Quizfrage 15.3:** Wie müsste Algorithmus 15.20 ergänzt werden, damit wir als Ergebnis die Faktorisierung $A = B\,C$ erhalten, aus der dann durch Streichen von Spalten bzw. Zeilen die Rangfaktorisierung $A = \overline{B}\,\overline{C}$ folgt?) $\triangle$

§ 15.4 Transposition von Matrizen

Definition 15.24 (Transposition).
Es seien $(K, +, \cdot)$ 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 $\big(A^\intercal\big)_{ij} = (A)_{ji}$ für $i \in [\![1, m]\!]$ und $j \in [\![1, n]\!]$, heißt die zu A **transponierte Matrix** (englisch: transpose matrix) oder kurz die **Transponierte** zu A. $\triangle$

[11] Später (Definition 15.36) werden wir solche Matrizen **invertierbar** nennen.

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 15.25 (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 15.26 (Rechenregeln für Transponierte)**.**
Es seien $(K, +, \cdot)$ 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:

$$(A^\intercal)^\intercal = A \quad \text{Transposition ist } \textbf{involutorisch} \tag{15.24}$$
$$(A + B)^\intercal = A^\intercal + B^\intercal \quad \text{Transposition ist additiv} \tag{15.25a}$$
$$(\alpha\, A)^\intercal = \alpha\, A^\intercal \quad \text{Transposition ist homogen}^{12} \tag{15.25b}$$
$$(A\, C)^\intercal = C^\intercal A^\intercal. \tag{15.26}$$

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 15.27 (Rang der transponierten Matrix)**.**
Es seien $(K, +, \cdot)$ 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 15.13), also gilt $B \in K^{n \times r}$ und $C \in K^{r \times m}$ mit $r = \operatorname{Rang}(A)$. Aufgrund von (15.26) gilt $A^\intercal = C^\intercal B^\intercal$. Aus Satz 15.14 und Satz 15.12 folgt $\operatorname{Rang}(A^\intercal) \leqslant \min\{\operatorname{Rang}(B^\intercal), \operatorname{Rang}(C^\intercal)\} \leqslant \min\{m, n, r\} \leqslant r = \operatorname{Rang}(A)$.

Führen wir dasselbe Argument nochmal mit einer Rangfaktorisierung $A^\intercal = \widehat{B}\,\widehat{C}$ mit $\widehat{B} \in K^{m \times s}$ und $C \in K^{s \times n}$ und $s = \operatorname{Rang}(A^\intercal)$ durch, so ergibt sich unter Beachtung von (15.24) genauso auch $\operatorname{Rang}(A) \leqslant \min\{m, n, s\} \leqslant s = \operatorname{Rang}(A^\intercal)$, zusammen also $\operatorname{Rang}(A) = \operatorname{Rang}(A^\intercal)$. □

Definition 15.28 (symmetrische und antisymmetrische Matrizen)**.**
Es seien $(K, +, \cdot)$ ein Körper, $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$.

[12] Additivität und Homogenität ergibt zusammen Linearität. Die Sprechweise, dass die Transposition eine lineare Abbildung sei, wird in § 17 klar werden.

(*i*) A heißt **symmetrisch** (englisch: symmetric matrix), wenn $A = A^\intercal$ gilt.

(*ii*) A heißt **antisymmetrisch** (englisch: antisymmetric matrix) oder **schiefsymmetrisch** (englisch: skew-symmetric matrix), 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. $K^{n\times n}_{\text{skew}}$. △

Lemma 15.29 (symmetrische und antisymmetrische Anteile quadratischer Matrizen)**.**
Es seien $(K, +, \cdot)$ 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

$$\dim(K^{n\times n}_{\text{sym}}) = \frac{1}{2} n\,(n+1) \tag{15.27a}$$

$$\dim(K^{n\times n}_{\text{skew}}) = \frac{1}{2} n\,(n-1). \tag{15.27b}$$

Es gilt

$$K^{n\times n} = K^{n\times n}_{\text{sym}} \oplus K^{n\times n}_{\text{skew}}. \tag{15.28}$$

Quizfrage 15.4: Was gilt im Körper $\mathbb{Z}_2$ der Charakteristik $\operatorname{char}(K) = 2$?

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-10.4. □

Ende der Vorlesung 21

Ende der Woche 10

§ 15.5 Der Ring quadratischer Matrizen

Die Menge der quadratischen Matrizen $K^{n\times n}$ ist bzgl. der Matrix-Multiplikation abgeschlossen. Zusammen mit der Matrix-Addition ergibt sich eine Ringstruktur $(K^{n\times n}, +, \cdot)$:

Lemma 15.30 (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.

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 (Lemma 15.3). Mit der Matrix-Multiplikation $\cdot$ bildet $(K^{n\times n}, \cdot)$ eine Halbgruppe, da $\cdot$ nach Lemma 15.8 assoziativ ist. Die Distributivgesetze

$$A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2$$
$$(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B$$

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

$$E_{11} \cdot E_{12} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} = E_{12}$$

$$E_{12} \cdot E_{11} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = 0,$$

das für beliebige Körper K gültig ist. □

Der Ring der quadratischen Matrizen $K^{n\times n}$ enthält neben den Diagonalmatrizen noch einige erwähnenswerte Teilmengen. Um Dreiecksmatrizen von strikten Dreiecksmatrizen unterscheiden zu können, schließen wir für die folgenden Resultate die 0×0-Matrizen aus.

Definition 15.31 (obere und untere Dreiecksmatrix)**.**
Es seien $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Eine Matrix $A \in K^{n\times n}$ heißt

(i) eine **obere Dreiecksmatrix** (englisch: upper triangular matrix), wenn $a_{ij} = 0$ für alle $i > j$ gilt.

(ii) eine **strikte obere Dreiecksmatrix** (englisch: strictly upper triangular matrix), wenn $a_{ij} = 0$ für alle $i \geqslant j$ gilt.

(iii) eine **untere Dreiecksmatrix** (englisch: lower triangular matrix), wenn $a_{ij} = 0$ für alle $i < j$ gilt.

(iv) eine **strikte untere Dreiecksmatrix** (englisch: strictly lower triangular matrix), wenn $a_{ij} = 0$ für alle $i \leqslant j$ gilt.

(v) **nilpotent** (englisch: nilpotent), wenn es ein $k \in \mathbb{N}_0$ gibt mit der Eigenschaft $A^k = 0$.

Hier gilt für die Indizes jeweils $1 \leqslant i, j \leqslant n$. △

Wir bezeichnen die Menge der Diagonalmatrizen der Dimension $n \times n$ auch mit $K^{n\times n}_{\diagdown}$ und die Menge der oberen bzw. unteren Dreiecksmatrizen mit $K^{n\times n}_{\text{◹}}$ bzw. $K^{n\times n}_{\text{◺}}$. Es gilt

$$K^{n\times n}_{\diagdown} = K^{n\times n}_{\text{◹}} \cap K^{n\times n}_{\text{◺}}.$$

Beispiel 15.32 (obere und untere Dreiecksmatrix)**.**

(i)
$$\begin{bmatrix} 1 & 3 & -3 \\ 0 & 7 & 4 \\ 0 & 0 & 5 \end{bmatrix}$$
ist eine obere Dreiecksmatrix, aber keine strikte obere Dreiecksmatrix.

(ii)
$$\begin{bmatrix} 0 & 0 & 0 \\ 7 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$
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,[13] denn es gilt

$$\begin{bmatrix} 2 & 2 & -2 \\ 5 & 1 & -3 \\ 1 & 5 & -3 \end{bmatrix}^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$ △

Lemma 15.33 (strikte Dreiecksmatrizen sind nilpotent)**.**
Es seien $(K, +, \cdot)$ 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 von Hausaufgabe I-11.1. □

Beispiel 15.34 (strikte Dreiecksmatrizen sind nilpotent)**.**
Für

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

gilt

$$A^2 = \begin{bmatrix} 0 & 0 & 4 & 17 \\ 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ △

Lemma 15.35 (obere und untere Dreiecksmatrizen bilden Unterringe)**.**
Es seien $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Dann gilt:

(*i*) $K^{n\times n}_{\diagdown}$, $K^{n\times n}_{\text{◹}}$ und $K^{n\times n}_{\text{◺}}$ bilden jeweils einen Unterring mit Eins (Definition 9.10) von $K^{n\times n}$. $K^{n\times n}_{\diagdown}$ ist sogar kommutativ für alle $n \in \mathbb{N}$.

(*ii*) Die strikten oberen und strikten unteren Dreiecksmatrizen bilden einen Unterring (aber keinen Unterring mit Eins) von $K^{n\times n}$. Im Fall $n = 1$ ist das der Nullring.

Beweis. Aussage (*i*): Nach Definition 9.10 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 $K^{n\times n}_{\text{◹}}$ aus. Sind $A, B \in K^{n\times n}_{\text{◹}}$, dann ist auch $-B \in K^{n\times n}_{\text{◹}}$ und damit $A - B \in K^{n\times n}_{\text{◹}}$. Aus dem Untergruppenkriterium (Satz 7.33) folgt, dass $(K^{n\times n}_{\text{◹}}, +)$ eine Untergruppe

[13] Quelle: https://en.wikipedia.org/wiki/Nilpotent_matrix#Example_5, genutzt unter der Lizenz CC-BY-SA 4.0

von $(K^{n\times n},+)$ ist. Da $I_n \in (K^{n\times n}_{\triangledown},+)$ klar ist, bleibt nur zu zeigen, dass das Matrix-Produkt $A \cdot B$ von zwei Matrizen $A, B \in K^{n\times n}_{\triangledown}$ wieder in $K^{n\times n}_{\triangledown}$ liegt. Für Indizes $1 \leqslant k < i \leqslant n$ gilt

$$\big[A \cdot B\big]_{ik} = \sum_{j=1}^{n} \underbrace{a_{ij}}_{=0 \text{ für } j<i} \overbrace{b_{jk}}^{=0 \text{ für } j>k} = 0,$$

da alle Summanden gleich Null sind, also ist tatsächlich $A \cdot B$ wieder eine obere Dreiecksmatrix.

Aussage (ii): Der Beweis der Unterring-Eigenschaft geht genauso wie in Aussage (i). Da I_n keine strikte Dreiecksmatrix ist, handelt es sich nicht um einen Unterring mit Eins[14] 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 7.16) des Monoids $(K^{n\times n},\cdot)$.

Definition 15.36 (invertierbare Matrix)**.**
Es seien $(K,+,\cdot)$ ein Körper und $n \in \mathbb{N}_0$.

(i) Eine Matrix $A \in K^{n\times n}$ heißt **invertierbar** (englisch: invertible matrix) oder **regulär** (englisch: non-singular matrix), wenn sie ein invertierbares Element (Definition 7.11) des Monoids $(K^{n\times n},\cdot)$ ist. Das heißt, es gibt eine Matrix $B \in K^{n\times n}$ mit der Eigenschaft

$$A\,B = I \quad \text{und} \quad B\,A = I. \tag{15.29}$$

In diesem Fall heißt B die zu A **inverse Matrix** (englisch: inverse matrix) oder kurz die **Inverse** (englisch: inverse) von A, in Symbolen: $B = A^{-1}$.

(ii) Andernfalls heißt A **nicht invertierbar** (englisch: non-invertible matrix) oder **singulär** (englisch: singular matrix).

(iii) Die Menge der invertierbaren Matrizen in $K^{n\times n}$, also die Einheitengruppe des Monoids $(K^{n\times n},\cdot)$, heißt die **allgemeine lineare Gruppe** (englisch: general linear group) **vom Grad n über dem Körper K**, in Symbolen

$$\operatorname{GL}(n,K) \coloneqq \big\{A \in K^{n\times n} \,\big|\, A \text{ ist invertierbar}\big\}. \tag{15.30}$$

△

Beachte: B ist Inverse von A genau dann, wenn A Inverse von B ist. Wie in jedem Monoid ist das neutrale Element, also die Einheitsmatrix I, immer invertierbar und selbstinvers.

Beispiel 15.37 (invertierbare Matrix)**.**
Die Matrix

$$A = \begin{bmatrix} 1 & 0 & -7 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

[14] Im Fall $n = 1$ besteht der Unterring nur aus der Nullmatrix, ist also der 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 9.10.

ist invertierbar, denn es gilt

$$\begin{bmatrix} 1 & 0 & -7 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} \frac{1}{17} \begin{bmatrix} 3 & 7 & 0 \\ 1 & -9 & 17 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{und} \quad \frac{1}{17} \begin{bmatrix} 3 & 7 & 0 \\ 1 & -9 & 17 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -7 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \triangle$$

Satz 15.38 (Rechenregeln für Inverse, vgl. Satz 7.17)**.**
Es seien $(K, +, \cdot)$ 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

$$(A^\intercal)^{-1} = (A^{-1})^\intercal. \tag{15.34}$$

Aus diesem Grund können wir statt $(A^\intercal)^{-1}$ auch einfach $A^{-\intercal}$ schreiben.

Beweis. Der Beweis der Eigenschaften Aussagen (i) bis (iii) geht so wie in Satz 7.17.

Aussage (iv): Es sei A invertierbar. Dann gilt

$$A^\intercal\,(A^{-1})^\intercal = (A^{-1}A)^\intercal = \mathrm{id}^\intercal = \mathrm{id},$$
$$(A^{-1})^\intercal\,A^\intercal = (A\,A^{-1})^\intercal = \mathrm{id}^\intercal = \mathrm{id},$$

also ist $(A^{-1})^\intercal$ in der Tat die Inverse zu $A^\intercal$. □

Anhand der Definition 15.36 kann man die Invertierbarkeit oder Nicht-Invertierbarkeit einer Matrix schlecht erkennen. Stattdessen geben wir nun Kriterien für die Invertierbarkeit an.

Lemma 15.39 (Elementarmatrizen sind invertierbar)**.**
Die Elementarmatrizen aus Definition 15.15 sind invertierbar.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-10.3. □

Für Matrizen in Zeilenstufenform ist die Invertierbarkeit leicht zu erkennen:

Satz 15.40 (Invertierbarkeit von Matrizen durch Zeilenstufenform)**.**
Es seien $(K, +, \cdot)$ 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 15.21). 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 maximalen Rang („**vollen Rang**“) besitzt. Das ist genau dann der Fall, wenn die Menge der Zeilenvektoren linear unabhängig ist und 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) $\Leftrightarrow$ Aussage (iii): Da die Elementarmatrizen nach Lemma 15.39 invertierbar sind, ist A genau dann invertierbar, wenn C invertierbar ist. (**Quizfrage 15.5:** Klar?)

Aussage (ii) $\Leftrightarrow$ Aussage (iv): Wie in Lemma 15.16 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 15.19. Dabei sind die Pivot-Elemente $\star \in K \setminus \{0\}$. Durch die Multiplikation $\widehat{C} := D\,C$ mit der Diagonalmatrix D bestehend aus den multiplikativen Inversen der Pivot-Elemente können wir erreichen, dass die Pivot-Elemente in $\widehat{C}$ alle gleich 1 sind.

$\widehat{C}$ ist also nun von der Form $\widehat{C} = I + N$ mit einer strikten oberen Dreiecksmatrix N. Wir rechnen nach, dass die Inverse von $\widehat{C}$ gegeben ist durch

$$\widehat{C}^{-1} = \sum_{k=0}^{n} (-N)^k.$$

Es gilt nämlich

$$\begin{aligned} \widehat{C} \sum_{k=0}^{n} (-N)^k &= \sum_{k=0}^{n} \widehat{C}\,(-N)^k = \sum_{k=0}^{n} (I + N)\,(-N)^k = \sum_{k=0}^{n} (-N)^k - \sum_{k=1}^{n+1} (-N)^k \\ &= (-N)^0 - (-N)^{n+1} = I - 0 = I. \end{aligned}$$

Die letzte Gleichheit gilt aufgrund von Lemma 15.33. Ganz ähnlich können wir auch zeigen:

$$\begin{aligned}\Big(\sum_{k=0}^{n}(-N)^k\Big)\,\widehat{C} = \sum_{k=0}^{n}(-N)^k\,\widehat{C} = \sum_{k=0}^{n}(-N)^k\,(I+N) = \sum_{k=0}^{n}(-N)^k - \sum_{k=1}^{n+1}(-N)^k \\ = (-N)^0 - (-N)^{n+1} = I - 0 = I.\end{aligned}$$

Tatsächlich ist also $\widehat{C}$ invertierbar und besitzt die angegebene Inverse. Damit ist auch $C = D^{-1}\widehat{C}$ 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 15.41 (Multiplikation mit invertierbaren Matrizen ändert den Rang nicht)**.**
Es seien $(K,+,\cdot)$ 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:

$$\operatorname{Rang}(B\,A\,C) = \operatorname{Rang}(A). \tag{15.35}$$

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-11.1. □

Abschließend zeigen wir, dass es für den Nachweis, dass zwei $n \times n$-Matrizen Inverse voneinander sind, ausreichend ist, diese in einer der beiden Reihenfolgen miteinander zu multiplizieren. Mit anderen Worten, jede Rechtsinverse einer quadratischen Matrix ist auch eine Linksinverse (und damit die eindeutige Inverse) und umgekehrt. In einer *Gruppe* kennen wir diese Eigenschaft bereits aus Satz 7.17 Aussage (ii). Allerdings bildet ja $(K^{n\times n},\cdot)$ i. A. lediglich eine (nicht-kommutative) Halbgruppe, und dort gilt diese Eigenschaft i. A. nicht (Hausaufgabe I-4.2). Es ist also bemerkenswert, dass die (für $n \geqslant 2$ nicht-kommutative) Halbgruppe $(K^{n\times n},\cdot)$ die Eigenschaft „Rechtsinverse sind Linksinverse und umgekehrt" besitzt.

Satz 15.42 (Rechtsinverse quadratischer Matrizen sind Linksinverse und umgekehrt)**.**
Es seien $(K,+,\cdot)$ 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 $B\,A = I_n$.

(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

$$\begin{aligned}
n &= \operatorname{Rang}(I_n) && \text{denn } I_n \text{ ist invertierbar und hat daher vollen Rang nach Satz 15.40} \\
&= \operatorname{Rang}(A\,B) && I_n = A\,B \text{ nach Voraussetzung} \\
&\leqslant \min\{\operatorname{Rang}(A),\ \operatorname{Rang}(B)\} && \text{nach Satz 15.14 (Rang des Produkts von Matrizen)} \\
&\leqslant n && \text{nach Satz 15.12 (Rang ist beschränkt durch die Dimensionen).}
\end{aligned}$$

Es muss also überall Gleichheit gelten, insbesondere ist $\operatorname{Rang}(A) = \operatorname{Rang}(B) = n$, und nach Satz 15.40 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 22

---

§ 16 Lineare Gleichungssysteme

Literatur: Fischer, Springborn, 2020, Kapitel 1.4 und 3.3, Bosch, 2014, Kapitel 3.5, Beutelspacher, 2014, Kapitel 4.1, Deiser, 2022b, 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 $(K, +, \cdot)$ 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). 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).

(ii) Die Matrix $[A, b] \in K^{n\times(m+1)}$ heißt die **erweiterte Koeffizientenmatrix** (englisch: augmented coefficient matrix).

(iii) Das lineare Gleichungssystem $A\,x = b$ heißt **homogen** (englisch: homogeneous), wenn $b = 0 \in K^m$ ist, andernfalls **nichthomogen** (englisch: non-homogeneous) oder **inhomogen** (englisch: inhomogeneous).

(iv) Das lineare Gleichungssystem $A\,x = b$ 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):

$$\left.\begin{array}{rcrcrcr} 6\,x_1 & + & 1\,x_2 & + & 1\,x_3 & = & 11 \\ 3\,x_1 & + & 3\,x_2 & + & 1\,x_3 & = & 5 \\ -6\,x_1 & - & 6\,x_2 & - & 3\,x_3 & = & -9 \end{array}\right\} \quad \Leftrightarrow \quad \underbrace{\begin{bmatrix} 6 & 1 & 1 \\ 3 & 3 & 1 \\ -6 & -6 & -3 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 11 \\ 5 \\ -9 \end{pmatrix}}_{b}.$$

(*ii*) Gesucht ist ein Polynom über dem Körper $\mathbb{Z}_5$ vom Höchstgrad 2, also $p = c_0 + c_1\,t + c_2\,t^2 \in (\mathbb{Z}_5)_2[t]$, dessen zugehörige Polynomfunktion $\widetilde{p}$ folgende Bedingungen erfüllt:

$$\widetilde{p}(0) = 3, \quad \widetilde{p}(1) = 2, \quad \widetilde{p}(3) = 4.$$

Das ist ein Beispiel einer **Interpolationsaufgabe**[15] (englisch: interpolation problem). Durch Einsetzen der drei **Interpolationsbedingungen** in die Polynomfunktion erhalten wir das lineare Gleichungssystem

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix}}_{c} = \underbrace{\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}}_{b}$$

für den gesuchten Koeffizientenvektor $c = (c_0, c_1, c_2)^\intercal$. △

Beachte: Jede Spalte der Matrix gehört zu einer der Variablen $x_1, \ldots, x_m$. Jede Zeile der Matrix gehört zu einer Gleichung.

Unser Ziel ist es, *alle* Lösungen von $A\,x = b$ zu bestimmen, also die gesamte **Lösungsmenge** (englisch: solution set)

$$\mathcal{L}(A, b) := \left\{x \in K^m \,\middle|\, A\,x = b\right\}. \tag{16.1}$$

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 $(K, +, \cdot)$ 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**“) (englisch: particular solution) **Lösung** 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) $\operatorname{Rang}(A) = \operatorname{Rang}([A, b])$.

[15] lateinisch: interpolare: neu herrichten, auffrischen

(*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. Aussage (*i*): Das Unterraumkriterium (Satz 12.8) zeigt, dass $\mathcal{L}(A, 0)$ ein Unterraum von K^m ist, 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

$$A\,(\alpha\,x + \beta\,y) = A\,(\alpha\,x) + A\,(\beta\,y) = \alpha\,A\,x + \beta\,A\,y = 0.$$

Wir bestätigen nun $\dim(\mathcal{L}(A, 0)) = m - \operatorname{Rang}(A)$, indem wir eine Basis dieses Unterraumes angeben. 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 $A\,x = 0 \Leftrightarrow \widehat{C}\,x = E_k \cdots E_2\,E_1\,0 = 0$, 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 $\mathcal{L}(A, 0) = \mathcal{L}(\widehat{C}, 0) = \mathcal{L}(C, 0)$.

Wir bezeichnen mit $\mathcal{A} \coloneqq (j_1, \dots, j_r)$ die Familie der Pivot-Spalten und mit $\mathcal{I}$ die komplementäre Familie der Nicht-Pivot-Spalten mit $\#\mathcal{I} = m - r$. Wir können jeden Vektor $x \in K^m$ in zwei Teilvektoren $x_\mathcal{A} \in K^r$ und $x_\mathcal{I} \in K^{m-r}$ partitionieren. Dies geschieht durch Einführung der beiden Matrizen

$$\Pi_\mathcal{A} \coloneqq (e_j^\intercal)_{j\in\mathcal{A}} \quad \text{und} \quad \Pi_\mathcal{I} \coloneqq (e_j^\intercal)_{j\in\mathcal{I}},$$

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 $C\,\Pi_\mathcal{A}^\intercal = [a_{\bullet j_1}, \dots, a_{\bullet j_r}]$ aus den Pivot-Spalten von C besteht und damit eine invertierbare obere Dreiecksmatrix (Satz 15.40) der Dimension $r \times r$ darstellt. Wir können also jede Lösung von $C\,x = 0$ in der Form

$$x_\mathcal{A} = -\big(C\,\Pi_\mathcal{A}^\intercal\big)^{-1} C\,\Pi_\mathcal{I}^\intercal x_\mathcal{I}$$

schreiben, 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 mit

$$x = -\Pi_\mathcal{A}^\intercal\big(C\,\Pi_\mathcal{A}^\intercal\big)^{-1} C\,\Pi_\mathcal{I}^\intercal x_\mathcal{I} + \Pi_\mathcal{I}^\intercal x_\mathcal{I} = \Big[I_m - \Pi_\mathcal{A}^\intercal\big(C\,\Pi_\mathcal{A}^\intercal\big)^{-1} C\Big]\,\Pi_\mathcal{I}^\intercal x_\mathcal{I}$$

mit beliebigem $x_{\mathcal{I}} \in K^{m-r}$ genau die Lösungen von $C\,x = 0$. Da die Spalten der Einheitsmatrix I_{m-r} eine Basis von K^{m-r} bilden, sind die Spalten der Matrix

$$\left[I_m - \Pi_{\mathcal{A}}^\intercal \left(C\,\Pi_{\mathcal{A}}^\intercal\right)^{-1} C\right] \Pi_{\mathcal{I}}^\intercal \in K^{n\times(m-r)}$$

eine Basis von $\mathcal{L}(C, 0) = \mathcal{L}(A, 0)$, und die Kardinalität der Basis ist $m - r$. In der Tat können wir auch

$$C\left[I_m - \Pi_{\mathcal{A}}^\intercal \left(C\,\Pi_{\mathcal{A}}^\intercal\right)^{-1} C\right] \Pi_{\mathcal{I}}^\intercal = \left[C - C\,\Pi_{\mathcal{A}}^\intercal \left(C\,\Pi_{\mathcal{A}}^\intercal\right)^{-1} C\right] \Pi_{\mathcal{I}}^\intercal = 0 \in K^{r\times(m-r)}$$

hier nochmal bestätigen.

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 $b \in \operatorname{SR}(A)$ liegt.

Aussage (b) $\Leftrightarrow$ Aussage (c): Die Hinzunahme des Spaltenvektors b zu einer Matrix A erhöht genau dann den $\operatorname{Rang}(A) = \operatorname{SRang}(A) = \dim(\operatorname{SR}(A))$ nicht, wenn b bereits in $\operatorname{SR}(A)$ 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)* $\operatorname{Rang}(A) = \operatorname{Rang}([A, b])$. 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 $m = \operatorname{Rang}(A)$.

Aussage (b) $\Rightarrow$ Aussage (a): Es gelte $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = m$. Dann ist aufgrund von Aussage *(iii)* $A\,x = b$ lösbar. Aufgrund der Darstellung (16.3) und $\dim(\mathcal{L}(A, 0)) = m - \operatorname{Rang}(A) = 0$ 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 $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = n$ gilt. $\operatorname{Rang}(A) = n$ impliziert nach Satz 15.40 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 15.19). 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 $\operatorname{Rang}(A) = r = 2$
- 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

$$\begin{aligned} x &= \Pi_{\mathcal{A}}^{\mathsf{T}} \, x_{\mathcal{A}} + \Pi_{\mathcal{I}}^{\mathsf{T}} \, x_{\mathcal{I}} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} x_1 \\ 0 \\ x_3 \\ 0 \end{pmatrix}}_{\text{abhängige Variable}} + \underbrace{\begin{pmatrix} 0 \\ x_2 \\ 0 \\ x_4 \end{pmatrix}}_{\text{unabhängige Variable}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \end{aligned}$$

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

$$\left[I_m - \Pi_{\mathcal{A}}^{\mathsf{T}} \left(C \, \Pi_{\mathcal{A}}^{\mathsf{T}} \right)^{-1} C \right] \Pi_{\mathcal{I}}^{\mathsf{T}} \in K^{n \times (m-r)} \tag{16.3}$$

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 $\Pi_{\mathcal{I}}^\intercal$.) Dann rechnen wir die Werte der abhängigen Variablen mit Hilfe der Gleichung

$$(C\,\Pi_{\mathcal{A}}^\intercal)\, x_{\mathcal{A}} = -\underbrace{C\,\Pi_{\mathcal{I}}^\intercal x_{\mathcal{I}}}_{\text{genau eine der } \mathcal{I}\text{-Spalten von } C} \tag{16.4}$$

aus. Weil $C\,\Pi_{\mathcal{A}}^\intercal$ eine obere Dreiecksmatrix ist, können wir die Werte der abhängigen Variablen von hinten nach vorne bestimmen und an der richtigen Stelle in den Lösungsvektor einsortieren ($\Pi_{\mathcal{A}}^\intercal x_{\mathcal{A}}$). Ein Beispiel folgt in Beispiel 16.7.

(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 $m - \operatorname{Rang}(A) \geqslant 1$.[16]
(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? Das Vorgehen ist wie folgt:

(1) Wir bringen die erweiterte Koeffizientenmatrix $[A, b]$ zunächst in Zeilenstufenform. Hier 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.[17]

Dieses Verfahren heißt **Gaußsches Eliminationsverfahren** (englisch: Gaussian elimination method).[18]

Definition 16.5 (reduzierte Zeilenstufenform, vgl. Definition 15.17)**.**
Eine Matrix $A \in K^{n\times m}$ heißt in **reduzierter Zeilenstufenform** (englisch: reduced row echelon form), wenn sie in Zeilenstufenform ist (Definition 15.17) und zusätzlich gilt:

(i) Alle Pivot-Elemente sind gleich 1.

(ii) Elemente, die in einer Spalte oberhalb eines Pivot-Elements stehen, sind gleich 0. △

[16] 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 $\#\mathcal{L}(A, b) = (\#K)^{m-\operatorname{Rang}(A)}$.

[17] Der zusätzliche Aufwand für die Überführung in die reduzierte Zeilenstufenform ist gerechtfertigt, weil dann die $\mathcal{A}$-Spalten $C\,\Pi_{\mathcal{A}}^\intercal$, aus denen sich — siehe (16.4) — die Werte der abhängigen Variablen ergeben, nicht nur eine obere Dreiecksmatrix, sondern die Einheitsmatrix ist, sodass wir die Werte direkt ablesen können.

[18] Das Wort „Elimination" bezieht sich darauf, dass durch Überführung in reduzierte 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 aufgelöst werden kann.

Beachte: Die reduzierte Zeilenstufenform einer Matrix ist eindeutig. Eine Matrix ist genau dann invertierbar, wenn ihre reduzierte Zeilenstufenform die Einheitsmatrix ist.

Beispiel 16.6 (reduzierte Zeilenstufenform, vgl. Beispiel 15.19)**.**

Hier sind die Besetzungsmuster einiger 3×4-Matrizen in reduzierter Zeilenstufenform, wobei ? für einen beliebigen Eintrag aus dem Körper K steht. Die Zahl r gibt wieder den Rang der Matrix an.

$$\begin{array}{r}j_1 = 1 \longrightarrow\\ j_2 = 2 \longrightarrow\\ j_3 = 3 \longrightarrow\end{array}\begin{bmatrix}1 & 0 & 0 & ?\\ 0 & 1 & 0 & ?\\ 0 & 0 & 1 & ?\end{bmatrix} \qquad \begin{array}{r}j_1 = 1 \longrightarrow\\ j_2 = 3 \longrightarrow\\ \end{array}\begin{bmatrix}1 & ? & 0 & ?\\ 0 & 0 & 1 & ?\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad \begin{array}{r}j_1 = 3 \longrightarrow\\ j_2 = 4 \longrightarrow\\ \end{array}\begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$r = 3 \qquad\qquad\qquad r = 2 \qquad\qquad\qquad r = 2$$

△

Wir geben nun anhand von Beispielen die Bestimmung der reduzierten Zeilenstufenform erweiterter 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 16.7 (Lösungsverfahren für lineare Gleichungssysteme (Gaußsches Eliminationsverfahren))**.**

Wir betrachten drei lineare Gleichungssysteme über dem endlichen Körper $(\mathbb{Z}_5, +_5, \cdot_5)$, vgl. Folgerung 10.6. 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

·	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

$$\left[\begin{array}{ccc|c}0 & 0 & 2 & 1\\ 1 & 3 & 0 & 2\\ 3 & 2 & 2 & 2\end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 0 & 2 & 1\\ 3 & 2 & 2 & 2\end{array}\right] \quad \text{Tauschen der Zeilen 1 und 2}$$

$$\begin{array}{c}\\ \\ 2\end{array}\left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 0 & 2 & 1\\ 3 & 2 & 2 & 2\end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 0 & 2 & 1\\ 0 & 3 & 2 & 1\end{array}\right] \quad \begin{array}{l}\text{Erzeugen von Nullen}\\ \text{unterhalb des Pivot-Elements}\end{array}$$

$$\left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 0 & 2 & 1\\ 0 & 3 & 2 & 1\end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 3 & 2 & 1\\ 0 & 0 & 2 & 1\end{array}\right] \quad \text{Tauschen der Zeilen 2 und 3}$$

$$\left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 3 & 2 & 1\\ 0 & 0 & 2 & 1\end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c}1 & 3 & 0 & 2\\ 0 & 3 & 2 & 1\\ 0 & 0 & 2 & 1\end{array}\right] \quad \begin{array}{l}\text{Erzeugen von Nullen}\\ \text{unterhalb des Pivot-Elements}\end{array}$$

An dieser Stelle ist eine Zeilenstufenform hergestellt. Wir erkennen $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = 3 = m$, also ist das System eindeutig lösbar. Wir gehen weiter zur reduzierten Zeilenstufenform:

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \text{Normierung des Pivot-Elements}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{oberhalb des Pivot-Elements}\end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \text{Normierung des Pivot-Elements}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{oberhalb des Pivot-Elements}\end{array}$$

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

$$\left[\begin{array}{ccc|c} 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 2 \\ 3 & 4 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 2 \end{array}\right] \qquad \text{Tauschen der Zeilen 1 und 2}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \qquad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements}\end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements}\end{array}$$

An dieser Stelle ist eine Zeilenstufenform hergestellt. Wir erkennen $\operatorname{Rang}(A) = \operatorname{Rang}([A, b]) = 2 = m - 1$, also ist das System lösbar, und die Lösungsmenge des homogenen Systems hat $\dim(\mathcal{L}(A, 0)) = 1$. Wir gehen weiter zur reduzierten Zeilenstufenform:

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \text{Normierung des Pivot-Elements}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \begin{array}{l}\text{Erzeugen von Nullen} \\ \text{oberhalb des Pivot-Elements}\end{array}$$

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 $x_2 := 0$ 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 Einheitsvektoren 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

$$\left[\begin{array}{ccc|c} 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 2 \\ 3 & 4 & 2 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 3 \end{array}\right] \quad \text{Tauschen der Zeilen 1 und 2}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 3 & 4 & 2 & 3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements} \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements} \end{array}$$

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. △

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 16.8 (Berechnung der inversen Matrix)**.**

Wir führen die Berechnung der inversen Matrix für die erste Matrix aus Beispiel 16.7 vor. Wir erinnern daran, dass wir es hier mit Matrizen über dem Körper $\mathbb{Z}_5$ zu tun haben.

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 \\ 3 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \quad \text{Tauschen der Zeilen 1 und 2}$$

$$\begin{array}{c} \\ \\ 2 \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements} \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \quad \text{Tauschen der Zeilen 2 und 3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{unterhalb des Pivot-Elements} \end{array}$$

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:

$$\begin{array}{c} \\ \\ 3 \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \quad \text{Normierung des Pivot-Elements}$$

$$\begin{array}{c} \\ 3 \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 4 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \quad \begin{array}{l} \text{Erzeugen von Nullen} \\ \text{oberhalb des Pivot-Elements} \end{array}$$

$$\begin{array}{c} \\ 2 \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 4 & 2 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \quad \text{Normierung des Pivot-Elements}$$

$$2\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 4 & 4 \\ 0 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 1 & 3 & 0 & 0 \end{array}\right]$$

Erzeugen von Nullen oberhalb des Pivot-Elements

Die Matrix auf der rechten Seite ist die Inverse der Ausgangsmatrix.

Wir führen die Probe durch:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 4 & 4 \\ 3 & 4 & 2 \\ 3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$ △

Ende der Vorlesung 23

Ende der Woche 11

§ 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, 2022b, 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 8.1 eines Halbgruppenhomomorphismus)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +_1, \cdot_1)$, $(W, +_2, \cdot_2)$ zwei Vektorräume über K.

(*i*) Eine Abbildung $f\colon V \to W$ heißt **strukturverträglich** oder ein **(Vektorraum-)Homomorphismus** (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*) Wie in Definition 8.1 sprechen wir im Fall $(V, +_1, \cdot_1) = (W, +_2, \cdot_2)$ von einem **(Vektorraum-)Endomorphismus** (englisch: vector space endomorphism) oder einem **linearen Endomorphismus** (englisch: linear endomorphism).

(*iii*) Ist $f\colon V \to W$ bijektiv, so heißt f auch **strukturerhaltend** oder ein **(Vektorraum-)Isomorphismus** (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*) Im Fall $(V, +_1, \cdot_1) = (W, +_2, \cdot_2)$ und $f\colon V \to W$ bijektiv sprechen wir auch von einem **(Vektorraum-)Automorphismus** (englisch: vector space automorphism) oder einem **linearen Automorphismus** (englisch: linear automorphism).

(v) Das **Bild** (englisch: image) und der **Kern** (englisch: kernel, null space[19]) eines Vektorraumhomomorphismus $f\colon V \to W$ sind definiert als

$$\operatorname{Bild}(f) := \bigl\{f(u) \in W \bigm| u \in V\bigr\} = f(V), \tag{17.2}$$
$$\operatorname{Kern}(f) := \bigl\{u \in V \bigm| f(u) = 0_W\bigr\} = f^{-1}(\{0_W\}). \tag{17.3}$$

△

Bemerkung 17.2 (zu Definition 17.1)**.**

(i) Die Vektorräume $(V, +_1)$ und $(W, +_2)$ sind insbesondere (abelsche) Gruppen. 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 zu notieren. △

Beispiel 17.3 (Vektorraumhomomorphismus)**.**

(i) Die Vektorräume K_n und K^n über einem Körper K sind zueinander isomorph. Die Abbildung

$$\cdot^\mathsf{T}\colon K_n \ni x \mapsto u^\mathsf{T} \in K^n$$

ist ein linearer Isomorphismus, vgl. (15.25a) und (15.25b). (**Quizfrage 17.1:** Wie sieht der inverse Vektorraumisomorphismus aus?)

(ii) 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

$$\operatorname{vec}\left(\begin{bmatrix} | & & | \\ a_{\bullet 1} & \cdots & a_{\bullet m} \\ | & & | \end{bmatrix}\right) := \begin{pmatrix} a_{\bullet 1} \\ \vdots \\ a_{\bullet m} \end{pmatrix},$$

ist ein linearer Isomorphismus. (**Quizfrage 17.2:** Wie sieht der inverse Vektorraumisomorphismus aus?)

(iii) 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

$$\pi_i\colon K^n \ni x \mapsto x_i \in K \tag{17.4}$$

für $1 \leqslant i \leqslant n \in \mathbb{N}$, eine surjektive lineare Abbildung. (**Quizfrage 17.3:** Ist sie auch injektiv?)

(iv) Die Grenzwert-Abbildung, definiert beispielsweise auf dem Vektorraum der konvergenten $\mathbb{R}$-wertigen Folgen $(\mathbb{R}^\mathbb{N})_c$ (siehe Beispiel 12.9) nach $\mathbb{R}$, also

$$\lim\colon (\mathbb{R}^\mathbb{N})_c \ni (y_i)_{i\in\mathbb{N}_0} \mapsto \lim_{i\to\infty}(y_i) \in \mathbb{R}$$

ist eine surjektive lineare Abbildung. (**Quizfrage 17.4:** Ist sie auch injektiv?)

[19] nicht zu verwechseln mit dem Begriff *zero vector space* (Nullraum) $\{0\}$, siehe Beispiel 12.3

(*v*) 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

$$\cdot' \colon \left\{ f \in \mathbb{R}^{(a,b)} \,\middle|\, f \text{ ist differenzierbar} \right\} \ni f \mapsto f' \in \mathbb{R}^{(a,b)}$$

ist eine lineare Abbildung, die nicht surjektiv und nicht injektiv ist. (**Quizfrage 17.5:** Warum?)

(*vi*) Für eine Matrix $A \in K^{n \times m}$ über einem Körper K ist die Matrix-Vektor-Multiplikation mit A

$$f_A \colon K^m \ni x \mapsto f_A(x) := A\,x \in K^n \tag{17.5}$$

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. △

Lemma 17.4 (Komposition linearer Abbildungen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(U, +, \cdot)$, $(V, +, \cdot)$, $(W, +, \cdot)$ Vektorräume über K. Sind $f \colon V \to W$ und $g \colon U \to V$ lineare Abbildungen, dann ist auch $f \circ g \colon U \to W$ eine lineare Abbildung.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-12.1. □

Als Folgerung ergibt sich (wie auch bereits bei Halbgruppen, Monoiden, Gruppen, Ringen und Körpern), dass Isomorphie eine Äquivalenzrelation auf der Klasse aller Vektorräume über demselben Körper K ist.

Lemma 17.5 (Eigenschaften linearer Abbildungen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Weiter sei $f \colon V \to W$ eine lineare Abbildung.

(*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 $v_i \in V$.

(*iv*) Ist $E \subseteq V$, dann gilt $f\big(\langle E \rangle\big) = \big\langle f(E) \big\rangle$.

(*v*) Ist $F = (v_i)_{i \in I}$ eine Familie von Vektoren in V, dann gilt $f\big(\langle F \rangle\big) = \big\langle f(F) \big\rangle$.

(*vi*) Ist $U \subseteq V$ ein Unterraum, dann ist $f(U) \subseteq W$ ein Unterraum.

(*vii*) Ist $Z \subseteq W$ ein Unterraum, dann ist $f^{-1}(Z) \subseteq V$ 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.

(*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.

Beweis. Aussage (i) und Aussage (ii) folgen sofort aus Lemma 8.5.

Aussage (iii):

$$\begin{aligned} f\Big(\sum_{i=1}^n \alpha_i\, v_i\Big) &= \sum_{i=1}^n f(\alpha_i\, v_i) \quad \text{durch wiederholte Anwendung von (17.1a)} \\ &= \sum_{i=1}^n \alpha_i\, f(v_i) \quad \text{durch Anwendung von (17.1b) auf jeden Summanden} \end{aligned}$$

Aussage (iv): Nach Satz 12.13 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 12.13 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 12.8). 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 12.8). 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 $(f(v_i))_{i \in I}$. □

Lemma 17.6 (Kern und Bild linearer Abbildungen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K sowie $f \colon V \longrightarrow W$ eine lineare Abbildung.

(i) Bild(f) ist ein Unterraum von W.

(ii) f ist surjektiv genau dann, wenn Bild$(f) = W$ gilt.

(iii) Kern(f) ist ein Unterraum von V.

(iv) f ist injektiv genau dann, wenn Kern$(f) = \{0\}$ gilt.

Beweis. Aussage (i): V ist selbst ein Unterraum von V, daher ist Bild$(f) = f(V)$ nach Lemma 17.5 (vi) ein Unterraum von W.

Aussage (ii): Die Aussage folgt sofort aus der Definition von Bild$(f) = f(V)$.

Aussage (iii): $\{0\}$ ist ein Unterraum von W, daher ist Kern$(f) = f^{-1}(\{0\})$ nach Lemma 17.5 (vii) ein Unterraum von V.

Aussage (iv): Weder der Begriff der Injektivität noch die Definition von Kern$(f) := \{u \in V \mid f(u) = 0\}$ ändern sich, wenn wir die lineare Abbildung $f\colon (V,+,\cdot) \to (W,+,\cdot)$ als Gruppenhomomorphismus $f\colon (V,+) \to (W,+)$ auffassen. Die Aussage (iv) folgt daher aus Lemma 8.9. □

§ 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. Zu diesem Zweck bietet es sich an, mit **indizierten Basen** (englisch: indexed basis) zu arbeiten, um eine „Nummerierung" der Basisvektoren zu erhalten. Das heißt, dass wir von nun an bevorzugt mit Basen in Form von *Familien* von Vektoren (Definition 6.29) arbeiten anstatt mit Mengen von Vektoren. Ist die Indexmenge I dabei eine Teilmenge von $\mathbb{N}_0$, so sprechen wir auch von einer **geordneten Basis** (englisch: ordered basis).

Satz 17.7 (Existenz- und Eindeutigkeitssatz für Vektorraumhomomorphismen)**.**
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$, $(W,+,\cdot)$ zwei Vektorräume über K. Weiter sei $(v_i)_{i\in I}$ eine Familie von Vektoren in V und $(w_i)_{i\in I}$ eine Familie von Vektoren in W mit gleicher Indexmenge I.

(i) Ist $(v_i)_{i\in I}$ linear unabhängig, dann gibt es eine lineare Abbildung $f\colon V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i \in I$.[20]

(ii) Ist $B := (v_i)_{i\in I}$ eine Basis von V, dann gibt es genau eine lineare Abbildung $f\colon V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i \in I$.

Diese Abbildung hat außerdem folgende Eigenschaften:

(a) Bild$(f) = \langle (w_i)_{i\in I} \rangle$.

(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.

Beweis. Wir beweisen zunächst die Aussage (ii).

[20] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Schritt 1: Wir konstruieren eine Abbildung $f\colon V \to W$ mit den gesuchten Eigenschaften.

Da $B = (v_i)_{i\in I}$ eine Basis von V ist, kann jedes $v \in V$ nach Satz 13.10 auf (bis auf Nullkoeffizienten) eindeutige Art und Weise als Linearkombination $v = \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}$ geschrieben werden. Daher ist die Setzung

$$f(v) \coloneqq \sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell}$$

wohldefiniert. (**Quizfrage 17.6:** Warum ist die Eindeutigkeit der Darstellung von v bis auf Nullkoeffizienten für die Wohldefiniertheit schon ausreichend?)

Wegen $v_i = 1 \cdot v_i$ gilt

$$f(v_i) = f(1 \cdot v_i) = 1 \, w_i = w_i$$

erfüllt f die Bedingung $f(v_i) = w_i$.

Schritt 2: Wir zeigen: Die so definierte Abbildung $f\colon V \to W$ ist linear.

Sind $v = \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}$ und $u = \sum_{k=1}^m \beta_k \, v_{j_k}$ zwei beliebige Vektoren aus V, dann können wir durch Hinzufügen von Nullkoeffizienten erreichen, dass beide Darstellungen dieselben endlich vielen Indizes aus I verwenden, also $v = \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}$ und $u = \sum_{\ell=1}^n \beta_\ell \, v_{i_\ell}$. Es gilt

$$\begin{aligned}
f(v+u) &= f\Big(\sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell} + \sum_{\ell=1}^n \beta_\ell \, v_{i_\ell}\Big) && \text{nach Darstellung von } v \text{ und } u \\
&= f\Big(\sum_{\ell=1}^n (\alpha_\ell + \beta_\ell) \, v_{i_\ell}\Big) && \text{nach Distributivgesetz (12.1b) und Kommutativität} \\
&= \sum_{\ell=1}^n (\alpha_\ell + \beta_\ell) \, w_{i_\ell} && \text{nach Definition von } f \\
&= \sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell} + \sum_{\ell=1}^n \beta_\ell \, w_{i_\ell} && \text{nach Distributivgesetz (12.1b)} \\
&= f(v) + f(u) && \text{nach Definition von } f
\end{aligned}$$

und außerdem für $\alpha \in K$

$$\begin{aligned}
f(\alpha \, v) &= f\Big(\alpha \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}\Big) && \text{nach Darstellung von } v \\
&= f\Big(\sum_{\ell=1}^n (\alpha \, \alpha_\ell) \, v_{i_\ell}\Big) && \text{nach Distributivgesetz (12.1a)} \\
&= \sum_{\ell=1}^n (\alpha \, \alpha_\ell) \, w_{i_\ell} && \text{nach Definition von } f \\
&= \alpha \sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell} && \text{nach Distributivgesetz (12.1a)} \\
&= \alpha \, f(v) && \text{nach Definition von } f.
\end{aligned}$$

Schritt 3: Wir zeigen die Eindeutigkeit von f.

Dazu sei $g\colon V \to W$ 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 $v = \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}$, dann gilt

$$\begin{aligned} g(v) &= g\Big(\sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}\Big) && \text{wegen der Darstellung von } v \\ &= \sum_{\ell=1}^n \alpha_\ell \, g(v_{i_\ell}) && \text{wegen der Linearität von } g \\ &= \sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell} && \text{wegen der Eigenschaft } g(v_i) = w_i \\ &= f(v) && \text{nach Definition von } f. \end{aligned}$$

Also muss g notwendig mit f übereinstimmen.

Wir kommen zu den weiteren Eigenschaften der Abbildung f.

Aussage (a): Es gilt

$$\begin{aligned} \operatorname{Bild}(f) &= f(V) && \text{nach Definition von } \operatorname{Bild}(f) \\ &= f\big(\langle (v_i)_{i\in I}\rangle\big) && \text{denn } (v_i)_{i\in I} \text{ ist eine Basis von } V \\ &= \big\langle (f(v_i))_{i\in I}\big\rangle && \text{nach Lemma 17.5 } (v) \\ &= \langle (w_i)_{i\in I}\rangle. \end{aligned}$$

Aussage (b): $f\colon V \to W$ 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 $(w_i)_{i\in I}$ linear abhängig, d. h., es gibt ein $n \in \mathbb{N}_0$ und paarweise verschiedene Indizes $i_1, \dots, i_n \in I$ sowie Koeffizienten $\alpha_1, \dots, \alpha_n \in K$, die nicht alle gleich 0 sind, sodass gilt: $\sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell} = 0$ gilt. Dann ist $v \coloneqq \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}$ nicht der Nullvektor in V, da $(v_i)_{i\in I}$ eine Basis von V ist, jedoch gilt

$$f(v) = f\Big(\sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}\Big) = \sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell} = 0.$$

Das bedeutet, dass v und der Nullvektor durch f beide auf den Nullvektor in W abgebildet werden, 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 $v = \sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}$, und es gilt

$$0 = f(v) = f\Big(\sum_{\ell=1}^n \alpha_\ell \, v_{i_\ell}\Big) = \sum_{\ell=1}^n \alpha_\ell \, w_{i_\ell}.$$

Aufgrund der lineare Unabhängigkeit der Familie $(w_i)_{i\in I}$ ist das nur möglich, wenn alle $\alpha_\ell = 0$ sind, also nur dann, wenn $v = 0$ ist. Mit anderen Worten, $\operatorname{Kern}(f) = \{0\}$, und nach Lemma 17.6 (iv) ist f injektiv.

Aussage (d) folgt sofort aus Aussage (b) und Aussage (c).

Nun kommen wir zur Aussage (i). Es sei also $(v_i)_{i\in I}$ linear unabhängig. Wir nutzen den Basisergänzungssatz 13.11 (der im Fall, dass V unendlich-dimensional ist, das Zornsche Lemma und damit das Auswahlaxiom verwendet) und ergänzen die Menge zu einer Basis $(v_i)_{i\in\widehat{I}}$ von V. Wir wählen die fehlenden $(w_i)_{\widehat{I}\setminus I}$ als beliebige Vektoren in W (beispielsweise alle als den Nullvektor). Nach Aussage (i) gibt es dann eine (eindeutige) lineare Abbildung $f\colon V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i \in \widehat{I}$, insbesondere für alle $i \in I$. □

Beispiel 17.8 (Konstruktion linearer Abbildungen).

(i) Es sei K ein Körper und V ein Vektorraum über K. Dann ist eine lineare Abbildung $f\colon K \to V$ durch einen einzigen Funktionswert, etwa $f(1) \in V$, festgelegt.

(ii) Es sei K ein Körper und $(e_1, \ldots, e_m)$ die Standardbasis im Vektorraum K^m. Eine lineare Abbildung $f\colon K^m \to K^n$ 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

$$A := \begin{bmatrix} | & & | \\ f(e_1) & \cdots & f(e_m) \\ | & & | \end{bmatrix}$$

ein, so gilt

$$f(x) = f\Big(\sum_{j=1}^m x_j\, e_j\Big) = \sum_{j=1}^m x_j f(e_j) = A\,x.$$

Eine lineare Abbildung $K^m \to K^n$ kann also durch Matrix-Vektor-Produkte $x \mapsto A\,x$ realisiert werden.

(iii) Im Vektorraum $\mathbb{R}^2$ mit der Standardbasis $(e_1, e_2) = \big(\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\big)$ 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 fest, und zwar eine Drehung um den Winkel 30° im mathematisch positiven Sinn (gegen den Uhrzeigersinn) um den Ursprung. Allgemeiner beschreibt

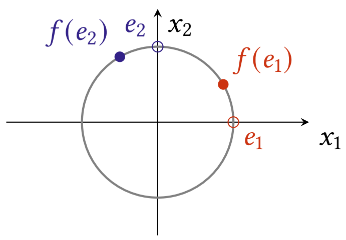
$$f(e_1) = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$$

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 17.7 (ii) bijektiv, also ein Isomorphismus $\mathbb{R}^2 \to \mathbb{R}^2$.

Die Drehabbildung kann 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 fest, und zwar eine Spiegelung an der x_1-Achse. Allgemeiner beschreibt

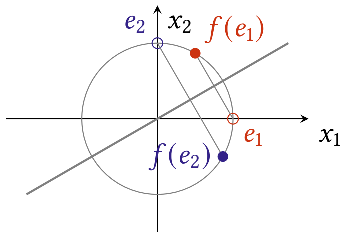
$$f(e_1) = \begin{pmatrix}\cos^2(\alpha) - \sin^2(\alpha)\\ 2\cos(\alpha)\,\sin(\alpha)\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}2\cos(\alpha)\,\sin(\alpha)\\ \sin^2(\alpha) - \cos^2(\alpha)\end{pmatrix}$$

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 17.7 (*ii*) bijektiv, also ein Isomorphismus $\mathbb{R}^2 \to \mathbb{R}^2$.

Die Spiegelungsabbildung kann 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. 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.



(*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

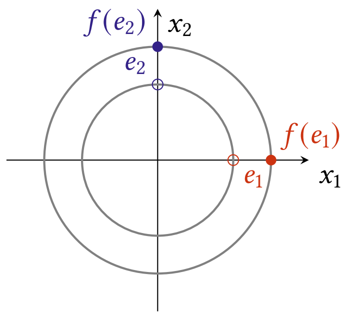
$$f(e_1) = \begin{pmatrix}\alpha\\0\end{pmatrix} \quad \text{und} \quad f(e_2) = \begin{pmatrix}0\\\alpha\end{pmatrix}$$

mit $\alpha \in \mathbb{R}$ eine lineare Abbildung fest, und zwar eine Streckung bzw. (wenn $\alpha < 0$ ist) eine Streckung und Punktspiegelung am Ursprung. Da $(f(e_1), f(e_2))$ für $\alpha \neq 0$ eine Basis von $\mathbb{R}^2$ bildet, ist die Streckungsabbildung nach Satz 17.7 (*ii*) in diesem Fall bijektiv, also ein Isomorphismus $\mathbb{R}^2 \to \mathbb{R}^2$.

Die Streckungsabbildung kann durch Matrix-Vektor-Produkte mit der Matrix

$$\begin{bmatrix}\alpha & 0\\ 0 & \alpha\end{bmatrix} \tag{17.8}$$

realisiert werden.



(*vi*) Es sei K ein Körper und $K[t]$ der Polynomraum über K (Beispiel 12.3). Die **Ableitungsabbildung** $f\colon K[t] \to K[t]$ ist ein linearer Endomorphismus, der durch die Festlegung

$$f(t^n) := \begin{cases} n\,t^{n-1} & \text{für } n \in \mathbb{N} \\ 0 & \text{für } n = 0 \end{cases}$$

einndeutig bestimmt ist, da die Monome $(t^n)_{n\in\mathbb{N}_0}$ eine Basis von $K[t]$ bilden (Beispiel 13.9). Dabei △

Ende der Vorlesung 24

§ 17.2 Die Matrix-Vektor-Multiplikation als lineare Abbildung

Die Matrix-Vektor-Multiplikation als Prototyp einer linearen Abbildung (Beispiel 17.3) ist von so zentraler Bedeutung, dass wir hier ihre Eigenschaften zusammenstellen. Kurz gesagt: 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 17.9 (Matrix-Vektor-Multiplikation als lineare Abbildung)**.**
Es sei K ein Körper und $n, m, k \in \mathbb{N}_0$.

(*i*) Ist $A \in K^{n\times m}$, dann definiert die **von A induzierte lineare Abbildung**

$$f_A\colon K^m \ni x \mapsto f_A(x) := A\,x \in K^n \tag{17.9}$$

eine lineare Abbildung $K^m \to K^n$.

(*ii*) Ist $f\colon K^m \to K^n$ 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

$$f_A \circ f_B = f_{A\,B}. \tag{17.10}$$

(*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

$$(f_A)^{-1} = f_{A^{-1}}. \tag{17.11}$$

Beweis. Der Beweis ist Gegenstand von Hausaufgabe I-12.3. □

§ 17.3 Der Vektorraum der Vektorraumhomomorphismen

Definition 17.10 (Menge der Homomorphismen, Endomorphismen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K.

(*i*) Wir bezeichnen die Menge der linearen Abbildungen (Homomorphismen) $V \to W$ mit

$$\operatorname{Hom}(V, W) := \big\{ f\colon V \to W \,\big|\, f \text{ ist Vektorraumhomomorphismus} \big\}. \tag{17.12}$$

Wollen wir den Skalarkörper betonen, so schreiben wir auch $\operatorname{Hom}_K(V, W)$.

(*ii*) Wir bezeichnen die Menge der Endomorphismen $V \to V$ mit

$$\operatorname{End}(V) := \operatorname{Hom}(V,V) = \big\{ f\colon V \to V \,\big|\, f \text{ ist Vektorraumendomorphismus} \big\}. \tag{17.13}$$

Wollen wir den Skalarkörper betonen, so schreiben wir auch hier $\operatorname{End}_K(V)$. △

Satz 17.11 (die Homomorphismen zwischen Vektorräumen bilden einen Vektorraum)**.**
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$, $(W,+,\cdot)$ zwei Vektorräume über K. $(\operatorname{Hom}(V,W),+,\cdot)$ mit der punktweisen Addition $+$ und der punktweisen S-Multiplikation $\cdot$

$$\begin{aligned}
&+\colon \operatorname{Hom}(V,W) \times \operatorname{Hom}(V,W) \to \operatorname{Hom}(V,W) && \text{mit } f+g \text{ definiert durch } (f+g)(u) := f(u)+g(u) \\
&\cdot\colon K \times \operatorname{Hom}(V,W) \to \operatorname{Hom}(V,W) && \text{mit } \alpha\cdot f \text{ definiert durch } (\alpha\cdot f)(u) := \alpha\, f(u)
\end{aligned}$$

bildet einen Vektorraum über K. Der Nullvektor in $(\operatorname{Hom}(V,W),+,\cdot)$ 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 12.1 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 10.2). Insbesondere ist also $(\operatorname{Hom}(V,W),+)$ eine abelsche Gruppe, da $(W,+)$ eine abelsche Gruppe ist.

Weiter gelten die Distributivgesetze

$$\begin{aligned}
&& \alpha\,(f+g) &= \alpha\, f + \alpha\, g \\
\text{und} && (\alpha+\beta)\, f &= \alpha\, f + \beta\, f
\end{aligned}$$

für alle $\alpha,\beta \in K$ und $f,g \in \operatorname{Hom}(V,W)$, denn wir haben

$$\begin{aligned}
&& [\alpha\,(f+g)](u) &= \alpha\,[(f+g)(u)] = \alpha\,[f(u)+g(u)] = \alpha\, f(u) + \alpha\, g(u) \\
&& &= (\alpha\, f)(u) + (\alpha\, g)(u) = [\alpha\, f + \alpha\, g](u) \\
\text{und} && [(\alpha+\beta)\, f](u) &= (\alpha+\beta)\, f(u) = \alpha\, f(u) + \beta\, f(u) \\
&& &= [\alpha\, f + \beta\, f](u)
\end{aligned}$$

für alle $u \in V$.

Das Assoziativgesetz

$$(\alpha\,\beta)\, f = \alpha\,(\beta\, f)$$

gilt, denn wir haben

$$[(\alpha\,\beta)\, f](u) = (\alpha\,\beta)\, f(u) = \alpha\,(\beta\, f(u)) = \alpha\,([\beta\, f](u)) = [\alpha\,(\beta\, f)](u)$$

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 17.11 ergibt sich, dass insbesondere die Endomorphismen $(\operatorname{End}(V),+,\cdot)$ einen Vektorraum bilden. Ersetzen wir die S-Multiplikation durch die Komposition $\circ$, so erhalten wir einen Ring wie schon beim Endomorphismenring einer abelschen Gruppe (vgl. Beispiel 9.2):

Satz 17.12 (die Endomorphismen eines Vektorraumes bilden einen Ring)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K. $(\operatorname{End}(V), +, \circ)$ mit der punktweisen Addition $+$ und der Komposition $\circ$

$$\begin{aligned} +\colon \operatorname{End}(V) \times \operatorname{End}(V) &\to \operatorname{End}(V) \quad \text{mit } f+g \text{ definiert durch } (f+g)(u) := f(u) + g(u) \\ \circ\colon \operatorname{End}(V) \times \operatorname{End}(V) &\to \operatorname{End}(V) \quad \text{mit } f \circ g \text{ definiert durch } (f \circ g)(u) := f(g(u)) \end{aligned}$$

bildet einen Ring mit dem Einselement id_V, genannt der **Endomorphismenring** (englisch: ring of endomorphisms) des Vektorraumes $(V, +, \cdot)$, vgl. Beispiel 9.2. Er ist i. A. nicht kommutativ.

Beweis. Der Beweis ist derselbe wie für den Endomorphismenring der abelschen Gruppe $(V, +)$ aus Beispiel 9.2, siehe Hausaufgabe I-6.3. Die Vektorraumstruktur von V spielt hier überhaupt keine Rolle. □

§ 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[21] 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 Übergang zur Nebenklasse ergibt dasselbe wie erst Übergang zur Nebenklasse und dann verknüpfen."

Dieser Übergang zur Faktorgruppe $(G / N, \tilde{\star})$ — einer gröberen Version der Gruppe G — war wichtig für das Verständnis der Wirkungsweise von Gruppenhomomorphismen (Homomorphiesatz für Gruppen 8.17). 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 als Normalteiler nicht beliebige Untergruppen verwenden, sondern Unterräume von V.

Aus der Faktormenge wird damit der **Faktorraum** (englisch: factor space) oder **Quotientenraum** (englisch: quotient space) von V nach U. Man sagt auch: „Aus dem Vektorraum $(V, +, \cdot)$ wird der Unterraum U ausfaktorisiert." Jede Nebenklasse $[v] = v + U$ heißt auch ein **affiner Unterraum parallel zu U** (englisch: affine subspace parallel to U). 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) $\dim(v + U) := \dim(U)$ zu.

Satz 17.13 (Faktorraum, vgl. Satz 8.13 über Faktorgruppen)**.**
Es seien $(K, +, \cdot)$ ein Körper, $(V, +, \cdot)$ ein Vektorraum über K und U ein Unterraum von V. Dann gilt:

[21] **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.

(*i*) Auf der Faktormenge

$$V / U = \big\{[v] = v + U \,\big|\, v \in V\big\}$$

sind $\tilde{+}$ und $\tilde{\cdot}$, definiert als

$$[v] \mathbin{\tilde{+}} [w] := [v + w] \quad \text{für } v, w \in V, \tag{17.14a}$$
$$\alpha \mathbin{\tilde{\cdot}} [v] := [\alpha \cdot v] \quad \text{für } \alpha \in K \text{ und } v \in V, \tag{17.14b}$$

eine innere bzw. äußere Verknüpfung, bzgl. der $(V / U, \tilde{+}, \tilde{\cdot})$ einen Vektorraum über K bildet. Das neutrale Element bzgl. $\tilde{+}$ ist $[0] = U$, und für die Inversen gilt $\tilde{-}[v] = [-v]$.

(*ii*) Die Abbildung

$$\pi\colon \begin{cases} V \longrightarrow V / U \\ v \longmapsto [v], \end{cases} \tag{17.15}$$

die jedem Vektor $v \in V$ seine Nebenklasse $[v]$ zuordnet, ist ein surjektiver Vektorraumhomomorphismus. Sie heißt die **kanonische Surjektion** (englisch: canonical surjection) von V auf V / U. Es gilt $\operatorname{Kern}(\pi) = U$.

Beweis. Aussage (*i*): Wir weisen die Bedingungen aus der Definition 12.1 eines Vektorraumes nach.

Nach Satz 8.13 ist $(V / U, \tilde{+})$ eine abelsche Gruppe, da $(U, +)$ insbesondere eine Untergruppe und damit ein Normalteiler der abelschen Gruppe $(V, +)$ ist.

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

$$\begin{aligned}
\alpha \mathbin{\tilde{\cdot}} [v_1] &= [\alpha\, v_1] && \text{nach Definition von } \tilde{\cdot} \\
&= \alpha\, v_1 + U && \text{nach Definition von } [\cdot] \\
&= \alpha\, (v_1 + U) && \text{da } \alpha\, U = U \text{ ist} \\
&= \alpha\, (v_2 + U) && \text{da } [v_1] = [v_2] \text{ vorausgesetzt wurde} \\
&= \alpha\, v_2 + U && \text{da } \alpha\, U = U \text{ ist} \\
&= [\alpha\, v_2] && \text{nach Definition von } [\cdot] \\
&= \alpha \mathbin{\tilde{\cdot}} [v_2] && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Im Fall $\alpha = 0$ gilt

$$\begin{aligned}
\alpha \mathbin{\tilde{\cdot}} [v_1] &= [0\, v_1] && \text{nach Definition von } \tilde{\cdot} \\
&= [0\, v_2] && \text{da } 0\, v_1 = 0 = 0\, v_2 \text{ gilt} \\
&= \alpha \mathbin{\tilde{\cdot}} [v_2] && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Nun weisen wir die Distributivgesetze (12.1a) und (12.1b) in $(V / U, \tilde{+}, \tilde{\cdot})$ nach:

$$\begin{aligned}
\alpha \mathbin{\tilde{\cdot}} ([v_1] \mathbin{\tilde{+}} [v_2]) &= \alpha \mathbin{\tilde{\cdot}} [v_1 + v_2] && \text{nach Definition von } \tilde{+} \\
&= [\alpha\, (v_1 + v_2)] && \text{nach Definition von } \tilde{\cdot} \\
&= [\alpha\, v_1 + \alpha\, v_2] && \text{aufgrund des Distributivgesetzes (12.1a) in } (V, +, \cdot) \\
&= [\alpha\, v_1] \mathbin{\tilde{+}} [\alpha\, v_2] && \text{nach Definition von } \tilde{+} \\
&= \alpha \mathbin{\tilde{\cdot}} [v_1] \mathbin{\tilde{+}} \alpha \mathbin{\tilde{\cdot}} [v_2] && \text{nach Definition von } \tilde{\cdot}
\end{aligned}$$

und

$$\begin{aligned}
(\alpha + \beta) \mathbin{\tilde{\cdot}} [v] &= [(\alpha + \beta)\, v] && \text{nach Definition von } \tilde{\cdot} \\
&= [\alpha\, v + \beta\, v] && \text{aufgrund des Distributivgesetzes (12.1b) in } (V, +, \cdot) \\
&= [\alpha\, v] \mathbin{\tilde{+}} [\beta\, v] && \text{nach Definition von } \tilde{+} \\
&= \alpha \mathbin{\tilde{\cdot}} [v] \mathbin{\tilde{+}} \beta \mathbin{\tilde{\cdot}} [v] && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Es folgt das Assoziativgesetz (12.1c):

$$\begin{aligned}
(\alpha\, \beta) \mathbin{\tilde{\cdot}} [v] &= [(\alpha\, \beta)\, v] && \text{nach Definition von } \tilde{\cdot} \\
&= [\alpha\, (\beta\, v)] && \text{aufgrund des Assoziativgesetzes (12.1c) in } (V, +, \cdot) \\
&= \alpha \mathbin{\tilde{\cdot}} [\beta\, v] && \text{nach Definition von } \tilde{\cdot} \\
&= \alpha \mathbin{\tilde{\cdot}} (\beta \mathbin{\tilde{\cdot}} [v]) && \text{nach Definition von } \tilde{\cdot}.
\end{aligned}$$

Und schließlich zeigen wir, dass $1 \in K$ neutrales Element bzgl. $\tilde{\cdot}$ ist:

$$\begin{aligned}
1 \mathbin{\tilde{\cdot}} [v] &= [1\, v] && \text{nach Definition von } \tilde{\cdot} \\
&= [v] && \text{da 1 neutrales Element bzgl. } \cdot \text{ in } (V, +, \cdot) \text{ ist.}
\end{aligned}$$

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 8.13 über die Eigenschaften der abelschen Gruppe $(V / U, \tilde{+})$.

Aussage (ii): Die Eigenschaft, ein Vektorraumhomomorphismus zu sein, bedeutet

$$\begin{aligned}
\pi(v + w) &= \pi(v) \mathbin{\tilde{+}} \pi(v) \\
\text{und} \quad \pi(\alpha\, v) &= \alpha \mathbin{\tilde{\cdot}} \pi(v)
\end{aligned}$$

für alle $v, w \in V$ und $\alpha \in K$. Nach Definition von π heißt das aber gerade

$$\begin{aligned}
[v + w] &= [v] \mathbin{\tilde{+}} [v] \\
\text{und} \quad [\alpha\, v] &= \alpha \mathbin{\tilde{\cdot}} [v],
\end{aligned}$$

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 $\operatorname{Kern}(\pi) = \pi^{-1}([0]) = U$. □

Bemerkung 17.14 (Faktorraum, vgl. Bemerkung 8.14)**.**
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 17.15 (Faktorraum, vgl. Beispiel 8.15)**.**

(i) Es sei V ein beliebiger Vektorraum und $U = \{0\}$ der Nullraum, einer der beiden trivialen Unterräume von V. Der zugehörige Faktorraum V / U 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 / U ist isomorph zum Nullraum $\{0\}$.

(*iii*) Es sei V der Polynomraum $K[t]$ über einem Körper K. Wählen wir $U = K_0[t]$ als den Unterraum der konstanten Polynome, dann besteht der Faktorraum V/U gerade aus Äquivalenzklassen von Polynomen, wobei zwei Polynome genau dann in derselben Äquivalenzklasse (Nebenklasse) liegen, wenn sie sich nur um eine additive Konstante unterscheiden. △

Bemerkung 17.16 (Unterräume sind genau die Kerne von Vektorraumhomomorphismen, vgl. Bemerkung 8.16).
Es sei $(V, +, \cdot)$ über dem Körper K.

(*i*) Nach Lemma 17.6 ist $\operatorname{Kern}(f)$ für jeden beliebigen Vektorraumhomomorphismus $f\colon V \to W$ in jeden beliebigen Vektorraum $(W, +, \cdot)$ über demselben Körper K immer ein Unterraum von $(V, +, \cdot)$.

(*ii*) Umgekehrt ist jeder Unterraum U von V der Kern eines geeignet gewählten Vektorraumhomomorphismus: Wähle z. B. W als den Faktorraum $(V/U, \tilde{+}, \tilde{\cdot})$ und als Homomorphismus die kanonische Surjektion $V \to V/U$. △

§ 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 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 $(V, +, \cdot)$ und den Unterraum $\operatorname{Kern}(f)$.

Satz 17.17 (**Homomorphiesatz für Vektorräume**[22], vgl. Homomorphiesatz für Gruppen 8.17).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Weiter sei $f\colon V \to W$ ein Homomorphismus. Dann gilt

$$V/\operatorname{Kern}(f) \cong \operatorname{Bild}(f) \tag{17.16a}$$

mit dem Isomorphismus

$$I([v]) := f(v) \quad \text{für } [v] = v + \operatorname{Kern}(f) \in V/\operatorname{Kern}(f). \tag{17.16b}$$

Beweis. Der Vektorraumhomomorphismus ist insbesondere ein Gruppenhomomorphismus $f\colon (V, +) \to (W, +)$, und $\operatorname{Kern}(f) = \{u \in V \mid f(u) = 0\}$ hängt nicht davon ab, ob wir f als Homomorphismus von Gruppen oder von Vektorräumen betrachten. Aus dem Homomorphiesatz für Gruppen 8.17 folgt also sofort, dass $V/\operatorname{Kern}(f)$ und $\operatorname{Bild}(f)$ im Sinne von Gruppen isomorph sind, und zwar durch den Isomorphismus (17.16b).

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:

$$\begin{aligned} I(\alpha \,\tilde{\cdot}\, [v]) &= I([\alpha\, v]) && \text{nach Definition von } \tilde{\cdot} \\ &= f(\alpha\, v) && \text{nach Definition von } I \\ &= \alpha\, f(v) && \text{aufgrund der Homogenität von } f \\ &= \alpha\, I([v]) && \text{nach Definition von } I. \end{aligned}$$
□

[22] englisch: fundamental theorem on vector space homomorphisms

Beispiel 17.18 (Homomorphiesatz für Vektorräume, vgl. Beispiel 8.18)**.**

(i) Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Wir betrachten den Nullhomomorphismus $f \colon V \to W$ mit $f(v) = 0$ für alle $v \in V$. Gemäß Homomorphiesatz für Vektorräume 17.17 ist
$$V \,/\, \operatorname{Kern}(f) = V \,/\, V \;\cong\; \{0\}.$$

(ii) Es seien K ein Körper und $V := K[t]$ der Polynomraum über K (Beispiel 12.3). Wir betrachten die Ableitungsabbildung $f \colon K[t] \to K[t]$ aus Beispiel 17.8, einen linearen Endomorphimus. Diese Abbildung ist surjektiv, aber nicht injektiv. Es gilt $\operatorname{Kern}(f) = K_0[t]$, der Unterraum der konstanten Polynome. Gemäß Homomorphiesatz für Vektorräume 17.17 ist
$$V \,/\, \operatorname{Kern}(f) = K[t] \,/\, K_0[t] \;\cong\; \operatorname{Bild}(f) = K[t].$$

In diesem Fall wird also das Bild trotz Ausfaktorisierung eines eindimensionalen Unterraumes nicht kleiner (geringer in der Dimension) als der Ursprungsraum $V = K[t]$. Das kann nur passieren, wenn $\dim(V) = \infty$ gilt.

Jede Äquivalenzklasse besteht aus denjenigen Polynomen, die sich nur um eine additive Konstante unterscheiden und daher durch die Ableitungsabbildung auf dasselbe Polynom abgebildet werden. Durch $\operatorname{Kern}(f)$ werden die konstante Polynom ausfaktorisiert.

(iii) Es seien $(K, +, \cdot)$ ein Körper und $V := (K^{\mathbb{N}}, +, \cdot)$ der Vektorraum aller K-wertigen Folgen. Wir betrachten den Homomorphismus $f \colon K^{\mathbb{N}} \to K$, der definiert ist durch
$$f((y_i)_{i \in \mathbb{N}}) := y_1 \in K,$$
der also die Folge auf das erste Folgenelement abbildet. Dann ist $\operatorname{Kern}(f) = \{(y_i)_{i \in \mathbb{N}} \mid y_1 = 0\}$, also besteht $K^{\mathbb{N}} \,/\, \operatorname{Kern}(f)$ aus Äquivalenzklassen von Folgen, die jeweils im ersten Folgenglied übereinstimmen. Eine gesamte Äquivalenzklasse wird auf ein- und dasselbe Element von K abgebildet. Gemäß Homomorphiesatz für Vektorräume 17.17 ist
$$V \,/\, \operatorname{Kern}(f) = K^{\mathbb{N}} \,/\, \operatorname{Kern}(f) \;\cong\; \operatorname{Bild}(f) = K. \qquad \triangle$$

Abschließend stellen wir den Homomorphiesatz für Vektorräume 17.17 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.

$$\begin{tikzcd}
W & V \arrow[l, "f"'] \arrow[d, "\pi"] \\
\operatorname{Bild}(f) \arrow[u, "i"] & V \,/\, \operatorname{Kern}(f) \arrow[l, "I"]
\end{tikzcd}$$

Das Diagramm besagt:
$$f = \underbrace{i}_{\text{einbetten}} \circ \overbrace{I}^{\text{isomorph abbilden}} \circ \underbrace{\pi}_{\text{vergröbern}}.$$

Ende der Vorlesung 25

Ende der Woche 12

§ 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 17.17 beteiligten Räume V, $\operatorname{Kern}(f)$ und $\operatorname{Bild}(f)$ untersuchen.

§ 18.1 Zusammenhang von Dimension und Isomorphie

Als Folgerungen aus dem Existenz- und Eindeutigkeitssatz für Vektorraumhomomorphismen 17.7 erhalten wir folgende bemerkenswerte Resultate:

Folgerung 18.1 (isomorphe Vektorräume besitzen dieselbe Dimension[23]).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Sind V und W isomorph, dann gilt $\dim(V) = \dim(W)$.

Beweis. Es sei $f\colon V \to W$ ein Isomorphismus. Nach Folgerung 13.12 existiert eine Basis $(v_i)_{i\in I}$ von V. Setze $w_i := f(v_i)$ für $i \in I$. Dann ist nach Satz 17.7 (ii) (d) $(w_i)_{i\in I}$ eine Basis von W. Beide Basen sind gleichmächtig, also gilt $\dim(V) = \dim(W)$. □

Folgerung 18.2 (Vektorräume gleicher endlicher Dimension sind isomorph).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei endlich-dimensionale Vektorräume über K. Dann sind äquivalent:

(i) V und W sind isomorphe Vektorräume.

(ii) $\dim(V) = \dim(W)$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii) ist gerade die Aussage des Lemmas 18.2.

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 17.7 (ii) gibt es genau eine lineare Abbildung $f\colon V \to W$ mit der Eigenschaft $f(v_i) = w_i$ für alle $i = 1, \ldots, n$. Nach Satz 17.7 (ii) (d) ist dieses f bijektiv, also sind V und W zueinander isomorph. □

Folgerung 18.2 besagt, dass es bis auf Isomorphie nur einen einzigen K-Vektorraum der Dimension $n \in \mathbb{N}_0$ gibt! Alle K-Vektorräume V mit $\dim(V) = n \in \mathbb{N}_0$ sind zueinander isomorph. Das werden wir später in § 19 noch ausnutzen.

Es gilt sogar:

Satz 18.3 (Vektorräume mit gleichmächtigen Basen sind isomorph[24]).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Dann sind äquivalent:

(i) V und W sind isomorphe Vektorräume.

[23] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

[24] Dieses Resultat verwendet das Auswahlaxiom.

(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.

Als Folgerung gibt es also zu jeder Kardinalzahl eine Äquivalenzklasse isomorpher Vektorräume. Wir können die Vektorräume K^n, $n \in \mathbb{N}_0$, als natürliche Repräsentanten der Vektorräume mit Basis der Mächtigkeit n ansehen. Weiter ist der Polynomraum $K[t]$ 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 18.4 (Isomorphiesatz).
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$ ein Vektorraum über K sowie U_1, U_2 zwei Unterräume von V. Dann gilt:

$$(U_1+U_2)/U_1 \cong U_2/(U_1\cap U_2). \tag{18.1}$$

Beweis. Die Beweisidee basiert darauf, 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 17.17 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).

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

$$\begin{aligned} \operatorname{Kern}(f) &= \{u_2\in U_2 \mid f(u_2) = [0]\} && \text{nach Definition des Kerns} \\ & && \text{und des neutralen Elements } [0] \text{ in } (U_1+U_2)/U_1 \\ &= \{u_2\in U_2 \mid [u_2] = [0]\} && \text{nach Definition von } f \\ &= \{u_2\in U_2 \mid u_2 - 0 \in U_1\} && \text{nach Definition von } [\cdot] \\ &= U_1\cap U_2. \end{aligned}$$

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 $[w] = u_1 + u_2 + U_1 = u_2 + U_1 = [u_2] = f(u_2)$. □

Folgerung 18.5 (Isomorphiesatz)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K sowie U_1, U_2 zwei Unterräume von V. Wenn $U_1 \cap U_2 = \{0\}$ gilt, dann ist

$$(U_1 \oplus U_2) / U_1 \cong U_2. \tag{18.2}$$

Beweis. Nach Satz 18.4 gilt

$$(U_1 + U_2) / U_1 \cong U_2 / (U_1 \cap U_2).$$

Die Direktheit der Summe $U_1 \oplus U_2$ bedeutet $U_1 \cap U_2 = \{0\}$, also gilt

$$(U_1 + U_2) / U_1 \cong U_2 / \{0\} = \big\{[u_2] = u_2 + \{0\} \,\big|\, u_2 \in U_2\big\} \cong U_2,$$

vgl. auch Beispiel 17.15. □

Nun können wir die Dimension von Faktorräumen bestimmen:

Satz 18.6 (Dimension des Faktorraumes[25])**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K sowie U ein Unterraum von V. Dann gilt

(i)

$$\dim(V) = \dim(U) + \dim(V / U) \tag{18.3}$$

Wenn $\dim(V) = \infty$ ist, dann hat also mindestens einer der Räume U und V / U ebenfalls unendliche Dimension.

(ii)

$$\dim(V / U) = \operatorname{codim}(U). \tag{18.4}$$

(iii) Ist V endlich-dimensional, dann gilt auch

$$\dim(V / U) = \dim(V) - \dim(U). \tag{18.5}$$

Beweis. Aussage (i) und Aussage (ii): Nach Folgerung 14.9 existiert ein zu U komplementärer Unterraum W, sodass also $V = U \oplus W$ gilt. Dieses Resultat haben wir, ausgehend von einer Basis B_U von U, mit Hilfe des Basisergänzungssatzes 13.11 durch Ergänzung zu einer Basis B von V bewiesen, wobei $B_W \coloneqq B \setminus B_U$ eine Basis von B_W ergibt.

Es gilt also

$$\dim(V) = \dim(U) + \dim(W) \quad \text{und} \quad \dim(W) = \operatorname{codim}(U)$$

nach Definition 14.10 der Kodimension. Diese Gleichung gilt auch im Fall $\dim(V) = \infty$, wobei dann mindestens einer der Räume U und W ebenfalls unendliche Dimension besitzt. Nach Folgerung 18.5 ist

[25] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

$V / U = (U \oplus W) / U \cong W$. Da isomorphe Vektorräume dieselbe Dimension besitzen (Folgerung 18.1), folgen

$$\dim(V) = \dim(U) + \dim(V / U) \quad \text{und} \quad \dim(V / U) = \operatorname{codim}(U),$$

also (18.3) und (18.4).

Aussage *(iii)*: Wenn $\dim(V)$ endlich ist, dann sind auch die Dimensionen der Unterräume $\dim(U)$ und $\dim(W) = \dim(V / U)$ endlich (Folgerung 13.22), und wir können (18.3) auflösen nach $\dim(V / U)$ auflösen, um (18.5) zu erhalten. □

§ 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 17.17 untersuchen:

Folgerung 18.7 (Dimensionen der Räume im Homomorphiesatz für Vektorräume 17.17[26])**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Weiter sei $f \colon V \longrightarrow W$ ein Homomorphismus. Dann gilt

$$\dim(\operatorname{Kern}(f)) + \dim(\operatorname{Bild}(f)) = \dim(V). \tag{18.6}$$

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 17.17 $\operatorname{Bild}(f)$ und $V / \operatorname{Kern} f$ isomorph sind und nach Folgerung 18.1 isomorphe Vektorräume dieselbe Dimension besitzen, ist (18.6) gezeigt. □

Definition 18.8 (Rang und Defekt eines Homomorphismus)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K. Weiter sei $f \colon V \longrightarrow W$ ein Homomorphismus. Dann heißt

$$\operatorname{Rang}(f) \coloneqq \dim(\operatorname{Bild}(f)) \tag{18.7}$$

der **Rang** (englisch: rank) der linearen Abbildung f, und

$$\operatorname{Defekt}(f) \coloneqq \dim(\operatorname{Kern}(f)) \tag{18.8}$$

heißt der **Defekt** (englisch: defect) von f. △

Wir können also die Dimensionsformel (18.6) auch in der Form

$$\operatorname{Defekt}(f) + \operatorname{Rang}(f) = \dim(V) \tag{18.9}$$

schreiben.

Das folgende Resultat erleichtert der Nachweis der Bijektivität (also der Isomorphismus-Eigenschaft) einer linearen Abbildung erheblich.

[26] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Folgerung 18.9 (Charakterisierung der Bijektivität von Homomorphismen endlich-dimensionaler Vektorräume)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei Vektorräume über K sowie $f \colon V \to W$ ein Homomorphismus.

(i) Haben V und W dieselbe endliche Dimension $\dim(V) = \dim(W)$, dann sind äquivalent:

  (a) f ist injektiv.

  (b) $\operatorname{Defekt}(f) = 0$.

  (c) f ist surjektiv.

  (d) $\operatorname{Rang}(f) = \dim(V)$.

  (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 sei 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 von Hausaufgabe I-13.2. □

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 Fälle auftreten:

- Die Ableitungsabbildung $f \colon K[t] \to K[t]$ ist surjektiv, aber nicht injektiv (Beispiel 17.18).
- Die Abbildung $f \colon K[t] \to K[t]$ definiert durch $f(p) = t \cdot p$ ist injektiv, aber nicht surjektiv.
- Die Abbildung $f \colon K[t] \to K[t]$ definiert durch $f(p) = -p$ ist bijektiv.
- Die Nullabbildung $f \colon K[t] \to K[t]$ definiert durch $f(p) = 0$ ist weder injektiv noch surjektiv. △

§ 19 Matrizen zur Darstellung linearer Abbildungen

Literatur: Fischer, Springborn, 2020, Kapitel 3.4–3.5, Beutelspacher, 2014, Kapitel 5.2

Im gesamten § 19 sind alle Vektorräume endlich-dimensional. Wir erinnern an Folgerung 18.2, die besagt, dass es bis auf Isomorphie nur einen einzigen K-Vektorraum der Dimension $n \in \mathbb{N}_0$ gibt! Alle K-Vektorräume V mit $\dim(V) = n \in \mathbb{N}_0$ sind also zueinander isomorph. 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 Die Koordinatendarstellung eines endlich-dimensionalen Vektorraumes

Der folgende Satz gibt an, wie wir mit Hilfe einer Basis von V einen Isomorphismus $K^n \to V$ erhalten können.

Satz 19.1 (Koordinatendarstellung).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei die Familie $B = (v_1, \ldots, v_n)$ eine Basis von V. Dann ist die Abbildung

$$\Phi_B \colon K^n \ni \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^n x_i\, v_i \in V \tag{19.1}$$

ein linearer Isomorphismus $K^n \to V$. 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** (englisch: coordinate vector) $x \in K^n$ bzgl. der Basis B zuordnet. Wir nennen daher Φ_B^{-1} auch die **Koordinatenabbildung** (englisch: coordinate map) bzgl. der Basis B.[27] 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 17.7 *(ii)* (d) bijektiv, also ein linearer Isomorphismus, und die Inverse Φ_B^{-1} ebenfalls. □

Wir werden Koordinatenvektoren typischerweise mit x bezeichnen.

Beispiel 19.2 (Koordinatendarstellung).

(*i*) Das Polynom $7\,t^2 - 3\,t + 5$ hat in der Monombasis $(1,\ t,\ t^2)$ von $\mathbb{R}_2[t]$ den Koeffizientenvektor $\left(\begin{smallmatrix} 5 \\ -3 \\ 7 \end{smallmatrix}\right)$, denn es gilt
$$7\,t^2 - 3\,t + 5 = 5 \cdot 1 + (-3) \cdot t + 7 \cdot t^2.$$

[27] Hier zeigt sich der Grund, warum wir den Standardvektorraum K^n, der in Beispiel 12.3 eingeführt wurde, auch als **Koordinatenraum** bezeichnen. Statt **Koordinaten** kann man auch von den **Komponenten** (englisch: components) bzgl. der Basis B sprechen.

(*ii*) Um dasselbe Polynom $7\,t^2 - 3\,t + 5$ in der Basis $(t^2 - t + 1,\ t^2 + 3,\ t + 1)$ darzustellen, schreiben wir es als Linearkombination der Basisvektoren mit unbekanntem Koeffizientenvektor $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ auf:

$$7\,t^2 - 3\,t + 5 = x_1\,(t^2 - t + 1) + x_2\,(t^2 + 3) + x_3\,(t + 1).$$

Ein Koeffizientenvergleich ergibt das lineare Gleichungssystem

$$\begin{bmatrix} 1 & 3 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ 7 \end{pmatrix}$$

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?)

Probe:

$$\begin{aligned} &13\,(t^2 - t + 1) + (-6)\,(t^2 + 3) + 10\,(t + 1) \\ &\quad = (13 - 6)\,t^2 + (-13 + 10)\,t + (13 - 18 + 10)\,1 \\ &\quad = 7\,t^2 - 3\,t + 5. \end{aligned}$$

△

Beachte: Zur Bestimmung des Koordinatenvektors $\Phi_B^{-1}(v)$ eines Vektors $v \in V$ muss man i. A. ein lineares Gleichungssystem lösen!

Ende der Vorlesung 26

§ 19.2 Darstellung linearer Abbildungen durch Matrizen

Aus Satz 17.7 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 $V \longrightarrow W$ mit Hilfe einer Matrix ist nun,

- alle Vektoren in V durch ihre Koordinatenvektoren bzgl. einer Basis darzustellen,
- ebenso alle Vektoren in W durch ihre Koordinatenvektoren bzgl. einer Basis darzustellen
- und nur noch mit den Koordinatenvektoren zu rechnen, für die die lineare Abbildung notwendigerweise die Form von Matrix-Vektor-Produkten hat.

Als Konsequenz können wir jede lineare Abbildung zwischen beliebigen endlich-dimensionalen Vektorräumen immer in der gleichen, einfachen und konkreten Form von Matrix-Vektor-Produkten darstellen.

Satz 19.3 (Darstellungssatz für lineare Abbildungen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei endlich-dimensionale Vektorräume über K. 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 $f\colon V \longrightarrow W$ 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 auch die **Darstellungsmatrix der linearen Abbildung f bzgl. der Basen B_V und B_W** (englisch: representation matrix), in Symbolen: $A = \mathcal{M}_{B_W}^{B_V}(f)$.

Beweis. f ist durch die Bilder $f(v_j)$ der Basisvektoren v_j, $j = 1, \dots, m$, nach Satz 17.7 eindeutig festgelegt. Für jeden Vektor $f(v_j) \in W$ ist wiederum sein Koordinatenvektor bzgl. der Basis $(w_1, \dots, w_n)$ 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 als Linearkombination in der Basis von W. Die j-te Spalte $\boldsymbol{a}_{\bullet j}$ der Darstellungsmatrix A enthält die Koordinaten von $f(v_j) \in W$ bzgl. der Basis B_W. Unsere Konvention beim Symbol der Darstellungsmatrix ist $\mathcal{M}_{\text{nach}}^{\text{von}}$, d. h., die Basis des Definitionsraumes („von") der darzustellenden Abbildung steht oben, und die Basis des Zielraumes („nach") steht unten.

Beispiel 19.4 (Darstellungsmatrizen von Homomorphismen)**.**

(*i*) Ist $V = K^m$ 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 17.9) gegeben, also $f_A(x) = A\,x$, dann ist A selbst die Darstellungsmatrix $\mathcal{M}_{B_W}^{B_V}(f)$, wenn $B_V = (e_1, \dots, e_m)$ und $B_W = (e_1, \dots, e_n)$ als die Standardbasen in K^m und K^n gewählt werden. Inbesondere hat beispielsweise die Drehabbildung aus Beispiel 17.8 als Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ bzgl. der Standardbasis 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 und die **zyklische Shift-Abbildung** (englisch: cyclic shift map) 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 Basen $B_V = B_W = \big((1,0,0,0,0), \dots, (0,0,0,0,1)\big)$ haben diese Endomorphismen die folgende Darstellung:

$$\mathcal{M}_{B_W}^{B_V}(f) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{bzw.} \quad \mathcal{M}_{B_W}^{B_V}(g) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

(*iii*) Auf dem Vektorraum $V = \mathbb{R}_3[t]$ der Polynome über $\mathbb{R}$ vom Höchstgrad 3 mit der Monombasis $(1,\ t,\ t^2\ t^3)$ betrachten wir die lineare Abbildung (drei Punktauswertungen der zugehörigen Polynomfunktion) mit Werten in $W = \mathbb{R}^3$

$$\mathbb{R}_3[t] \ni p \mapsto \begin{pmatrix} \widetilde{p}(-2) \\ \widetilde{p}(0) \\ \widetilde{p}(2) \end{pmatrix} \in \mathbb{R}^3.$$

Verwenden wir in $\mathbb{R}^3$ die Standardbasis, so besitzt diese Abbildung die Darstellungsmatrix

$$\begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{bmatrix}.$$

(*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 Basis $B_W = (1)$. Die lineare Abbildung $\operatorname{Re}\colon \mathbb{C} \to \mathbb{R}$ (Realteil) hat dann die Darstellungsmatrix

$$\mathcal{M}_{B_W}^{B_V}(\operatorname{Re}) = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

während die lineare Abbildung $\operatorname{Im}\colon \mathbb{C} \to \mathbb{R}$ (Imaginärteil) die Darstellungsmatrix

$$\mathcal{M}_{B_W}^{B_V}(\operatorname{Im}) = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{19.4}$$

besitzt.

(*v*) Es seien $(K, +, \cdot)$ ein Körper und V und W Vektorräume über K. Es sei $\dim(V) = m$, $\dim(W) = n$ und $f\colon V \to W$ ein Homomorphismus mit $r = \operatorname{Rang}(f)$ und daher $\operatorname{Defekt}(f) = \dim(\operatorname{Kern}(f)) = m - r$ nach Dimensionsformel (18.6). Wir wählen die Basen so, dass gilt

$$B_W = (\underbrace{w_1, \ldots, w_r}_{\text{Basis von Bild}(f)}, \underbrace{w_{r+1}, \ldots, w_n}_{\text{Ergänzung zu einer Basis von } W})$$

und

$$B_V = (\underbrace{v_1, \ldots, v_r}_{\text{siehe unten}}, \underbrace{v_{r+1}, \ldots, v_m}_{\text{Basis von Kern}(f)}).$$

Dabei sei $v_j \in f^{-1}(\{w_j\})$, sodass also $f(v_j) = w_j$ gilt für $j = 1, \ldots, r$. Bzgl. dieser angepassten Basen hat f die Darstellungsmatrix

$$\mathcal{M}_{B_W}^{B_V}(f) = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \in K^{n \times m}.$$

(*vi*) 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

$$\mathcal{M}_{B_V}^{B_V}(\operatorname{id}_V) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

also die Einheitsmatrix der passenden Dimension. △

Satz 19.5 (die Zuordnung zur Darstellungsmatrix ist ein Vektorraumisomorphismus)**.**

Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei endlich-dimensionale Vektorräume über K. Weiter seien $B_V = (v_1, \ldots, v_m)$ und $B_W = (w_1, \ldots, w_n)$ Basen von V bzw. von W. Die Zuordnung

$$\mathcal{M}_{B_W}^{B_V}\colon\ \operatorname{Hom}(V, W) \ni f \mapsto \mathcal{M}_{B_W}^{B_V}(f) \in K^{n \times m} \tag{19.5}$$

eines Homomorphismus zu seiner Darstellungsmatrix ist ein Isomorphismus von Vektorräumen.

Beweis. **Schritt** 1: Wir zeigen zunächst die Linearität von $\mathcal{M}_{B_W}^{B_V}$.

Es seien dazu $f, g \in \operatorname{Hom}(V, W)$, $A := \mathcal{M}_{B_W}^{B_V}(f)$ und $B := \mathcal{M}_{B_W}^{B_V}(g)$. 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 $j = 1, \ldots, m$. Das heißt $\mathcal{M}_{B_W}^{B_V}(f+g) = \mathcal{M}_{B_W}^{B_V}(f) + \mathcal{M}_{B_W}^{B_V}(g)$. Weiter gilt für $\alpha \in K$

$$\begin{aligned}(\alpha\, f)(v_j) &= \alpha\, f(v_j) \\ &= \alpha \sum_{i=1}^n a_{ij}\, w_i \\ &= \sum_{i=1}^n (\alpha\, a_{ij})\, w_i\end{aligned} \tag{19.6}$$

für alle $j = 1, \ldots, m$. Das heißt $\mathcal{M}_{B_W}^{B_V}(\alpha\, f) = \alpha\, \mathcal{M}_{B_W}^{B_V}(f)$.

Schritt 2: Wir zeigen: $\mathcal{M}_{B_W}^{B_V}$ ist injektiv.

Es sei dazu $f \in \operatorname{Hom}(V, W)$ so, dass $\mathcal{M}_{B_W}^{B_V}(f) = 0 \in K^{n \times m}$ (die Nullmatrix) ergibt. Das heißt,

$$f(v_j) = \sum_{i=1}^n 0\, w_i = 0$$

für alle $j = 1, \ldots, m$. Damit ist $f\colon V \to W$ der Nullhomomorphismus, also der Nullvektor von $\operatorname{Hom}(V, W)$. Daher gilt $\operatorname{Kern}(\mathcal{M}_{B_W}^{B_V}) = \{0\}$, und nach Lemma 17.6 ist $\mathcal{M}_{B_W}^{B_V}$ injektiv.

Schritt 3: Wir zeigen: $\mathcal{M}_{B_W}^{B_V}$ ist surjektiv.

Es sei dazu $A \in K^{n \times m}$. Nach Satz 17.7 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 19.2: Warum haben wir, um die Bijektivität von $\mathcal{M}_{B_W}^{B_V}$ zu zeigen, nicht die Charakterisierung der Bijektivität von Homomorphismen nach Folgerung 18.9 genutzt?

Folgerung 19.6 (Dimension des Vektorraumes der Homomorphismen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei endlich-dimensionale Vektorräume über K mit $\dim(V) = m$ und $\dim(W) = n$ sowie $m, n \in \mathbb{N}_0$. Dann gilt

$$\dim(\operatorname{Hom}(V, W)) = n\, m. \tag{19.7}$$

Beweis. Das Resultat folgt sofort aus $\dim(K^{n \times m}) = n\, m$ (Lemma 15.3) und der Isomorphie $\operatorname{Hom}(V, W) \cong K^{n \times m}$, aufgrund derer beide Räume dieselbe Dimension haben (Folgerung 18.2). □

Satz 19.7 (Zusammenhang zwischen einem Homomorphismus und seiner Darstellungsmatrix).
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$, $(W,+,\cdot)$ zwei endlich-dimensionale Vektorräume über K. Weiter seien $B_V = (v_1,\dots,v_m)$ und $B_W = (w_1,\dots,w_n)$ Basen von V bzw. von W und $f\colon V \to W$ ein Homomorphismus. Dann gilt für die Darstellungsmatrix $A \coloneqq \mathcal{M}_{B_W}^{B_V}(f)$ und die durch A induzierte lineare Abbildung f_A der Zusammenhang

$$f_A = \underbrace{\Phi_{B_W}^{-1}}_{\text{Koordinaten} \mapsfrom \text{Vektor}} \circ f \circ \underbrace{\Phi_{B_V}}_{\text{Vektor} \mapsfrom \text{Koordinaten}} : K^m \to K^n \tag{19.8a}$$

$$f = \underbrace{\Phi_{B_W}}_{\text{Vektor} \mapsfrom \text{Koordinaten}} \circ f_A \circ \underbrace{\Phi_{B_V}^{-1}}_{\text{Koordinaten} \mapsfrom \text{Vektor}} : V \to W. \tag{19.8b}$$

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 in den Vektor $\Phi_{B_V}(x) \in V$ umgerechnet, dann wirkt f, und schließlich wird das Ergebnis durch $\Phi_{B_W}^{-1}$ als Koordinatenvektor bzgl. der Basis B_W angegeben. (**Quizfrage 19.3:** Wie lässt sich (19.8b) interpretieren?)

Beweis. Wir zeigen, dass die Abbildungen $\Phi_{B_W} \circ f_A \in \operatorname{Hom}(K^m, W)$ und $f \circ \Phi_{B_V} \in \operatorname{Hom}(K^m, W)$ übereinstimmen. Dazu reicht es nach Satz 17.7 aus, zu zeigen, dass ihre Bilder auf einer Basis gleich sind. Wir wählen dazu die Standardbasis $(e_1,\dots,e_m)$ von K^m. Es gilt einerseits

$$\begin{aligned}
(\Phi_{B_W} \circ f_A)(e_j) &= \Phi_{B_W}(f_A(e_j)) && \text{nach Definition 6.13 der Komposition } \circ \\
&= \Phi_{B_W}(A\,e_j) && \text{nach Definition der von } A \text{ induzierten Abbildung } f_A \\
&= \Phi_{B_W}(a_{\bullet j}) && \text{nach Definition des Matrix-Vektor-Produkts} \\
&= \sum_{i=1}^n a_{ij}\,w_i && \text{nach Definition (19.1) von } \Phi_{B_W}
\end{aligned}$$

und andererseits

$$\begin{aligned}
(f \circ \Phi_{B_V})(e_j) &= f(\Phi_{B_V}(e_j)) && \text{nach Definition 6.13 der Komposition } \circ \\
&= f(v_j) && \text{nach Definition (19.1) von } \Phi_{B_V} \\
&= \sum_{i=1}^n a_{ij}\,w_i && \text{nach Definition (19.3) der Darstellungsmatrix } A.
\end{aligned}$$

Aus $\Phi_{B_W} \circ f_A \in \operatorname{Hom}(K^m, W) = f \circ \Phi_{B_V}$ können wir nun (19.8a) und (19.8b) leicht durch Auflösen herleiten, weil Φ_{B_V} und Φ_{B_W} bijektiv sind. □

Wir zeigen nun noch, dass die Komposition linearer Abbildungen durch das Matrix-Matrix-Produkt ihrer Darstellungsmatrizen dargestellt wird.

Satz 19.8 (Darstellungsmatrix der Komposition von Homomorphismen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(U, +, \cdot)$, $(V, +, \cdot)$ und $(W, +, \cdot)$ endlich-dimensionale Vektorräume über K. Weiter seien $B_U = (u_1, \dots, u_k)$, $B_V = (v_1, \dots, v_m)$ und $B_W = (w_1, \dots, w_n)$ 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

$$\mathcal{M}_{B_W}^{B_U}(f \circ g) = \mathcal{M}_{B_W}^{B_V}(f)\, \mathcal{M}_{B_V}^{B_U}(g). \tag{19.9}$$

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:

$$\begin{tikzcd}
W \arrow[ddd, "\Phi_{B_W}^{-1}"'] & & U \arrow[ll, "f \circ g"'] \arrow[dl, "g"] \\
& V \arrow[ul, "f"] \arrow[d, shift left, "\Phi_{B_V}^{-1}"] & \\
& K^m \arrow[u, shift left, "\Phi_{B_V}"] \arrow[dl, "f_A"'] & \\
K^n & & K^k \arrow[ll, "f_{A\,B}"] \arrow[ul, "f_B"'] \arrow[uuu, "\Phi_{B_U}"']
\end{tikzcd}$$

Zur Abkürzung setzen wir $A \coloneqq \mathcal{M}_{B_W}^{B_V}(f)$ und $B \coloneqq \mathcal{M}_{B_V}^{B_U}(g)$. Das rechte und das linke Trapez sind kommutativ nach Satz 19.7, 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 17.9. 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 $\mathcal{M}_{B_W}^{B_U}(f \circ g)$, was zu beweisen war. □

§ 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, W sind Vektorräume über K.

Matrix $A \in K^{n \times m}$		lineare Abbildung $f\colon V \to W$	
Begriff	siehe	Begriff	siehe
Bild(A) $:= \operatorname{SR}(A)$	(15.15a)	$\operatorname{Bild}(f) = \{f(u) \in W \mid u \in V\} = f(V)$	(17.2)
$\operatorname{SRang}(A) = \dim(\operatorname{SR}(A))$	(15.15b)	$\operatorname{Rang}(f) = \dim(\operatorname{Bild}(f))$	(18.7)
Kern(A) $:= \{x \in K^m \mid A\,x = 0\}$		$\operatorname{Kern}(f) = \{u \in V \mid f(u) = 0\} = f^{-1}(\{0\})$	(17.3)
Defekt(A) $:= \dim(\operatorname{Kern}(A))$		$\operatorname{Defekt}(f) = \dim(\operatorname{Kern}(f))$	(18.8)

Beachte: Aus den Definitionen folgt sofort $\operatorname{Bild}(A) = \operatorname{Bild}(f_A)$ und $\operatorname{Kern}(A) = \operatorname{Kern}(f_A)$.

Wir können nun bestätigen, dass die oben genannten Eigenschaften für lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen und ihren Darstellungsmatrizen übereinstimmen.

Satz 19.9 (Eigenschaften linearer Abbildungen und ihrer Darstellungsmatrizen)**.**
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$, $(W,+,\cdot)$ zwei endlich-dimensionale Vektorräume über K mit $\dim(V) = m$ und $\dim(W) = n$. 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 $A := \mathcal{M}_{B_W}^{B_V}(f) \in K^{n \times m}$ die Darstellungsmatrix von f. Dann gilt:

(i) $\operatorname{Bild}(f) = \Phi_{B_W}\big(\operatorname{Bild}(A)\big)$.

(ii) $\operatorname{Rang}(f) = \operatorname{SRang}(A) = \operatorname{Rang}(A)$.

(iii) $\operatorname{Kern}(f) = \Phi_{B_V}\big(\operatorname{Kern}(A)\big)$.

(iv) $\operatorname{Defekt}(f) = \operatorname{Defekt}(A)$.

Beweis. Aussage (i):

$$\begin{aligned}
& w \in \operatorname{Bild}(f) \\
\Leftrightarrow\quad & w \in \operatorname{Bild}(\Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}) && \text{wegen } f = \Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}\text{, siehe (19.8b)} \\
\Leftrightarrow\quad & w \in \Phi_{B_W}\big(\operatorname{Bild}(f_A \circ \Phi_{B_V}^{-1})\big) && \text{nach Definition von Bild} \\
\Leftrightarrow\quad & w \in \Phi_{B_W}\big(\operatorname{Bild}(f_A)\big) && \text{wegen der Bijektivität von } \Phi_{B_V}^{-1} \\
\Leftrightarrow\quad & w \in \Phi_{B_W}\big(\operatorname{Bild}(A)\big) && \text{wegen } f_A(x) = A\,x.
\end{aligned}$$

Aussage (ii): $\operatorname{Bild}(f)$ und $\operatorname{Bild}(A)$ sind nach Aussage (i) zueinander isomorphe Unterräume von W bzw. von K^n. Nach Folgerung 18.9 haben sie also dieselbe Dimension.

Aussage (iii):

$$\begin{aligned}
& v \in \operatorname{Kern}(f) \\
\Leftrightarrow\quad & v \in \operatorname{Kern}(\Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}) && \text{wegen } f = \Phi_{B_W} \circ f_A \circ \Phi_{B_V}^{-1}\text{, siehe (19.8b)} \\
\Leftrightarrow\quad & v \in \Phi_{B_V}\big(\operatorname{Kern}(\Phi_{B_W} \circ f_A)\big) && \text{wegen der Bijektivität von } \Phi_{B_V}^{-1} \\
\Leftrightarrow\quad & v \in \Phi_{B_V}\big(\operatorname{Kern}(f_A)\big) && \text{wegen der Bijektivität von } \Phi_{B_W} \\
\Leftrightarrow\quad & v \in \Phi_{B_V}\big(\operatorname{Kern}(A)\big) && \text{wegen } f_A(x) = A\,x.
\end{aligned}$$

Aussage (iv): $\operatorname{Kern}(f)$ und $\operatorname{Bild}(A)$ sind nach Aussage (iii) zueinander isomorphe Unterräume von V bzw. von K^m. Nach Folgerung 18.9 haben sie also dieselbe Dimension. □

Bemerkung 19.10 (zur Bestimmung von Bild und Kern einer Matrix).
Wie können wir für eine gegebene Matrix $A \in K^{n\times m}$ eine Basis von $\operatorname{Bild}(A) \subseteq K^n$ und eine Basis von $\operatorname{Kern}(A) \subseteq K^m$ bestimmen?

(*i*) Für $\operatorname{Bild}(A)$ haben wir folgende Möglichkeiten:

(1) Wir bestimmen mit Hilfe einer Erweiterung von Algorithmus 15.20, wie in Bemerkung 15.23 beschrieben, eine Rangfaktorisierung $A = B\,C$ mit $B \in K^{n\times r}$ und $C \in K^{r\times m}$ (in Zeilenstufenform) und $r = \operatorname{Rang}(A)$. Dann bilden die Spalten von B 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 15.20) 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. Nur wenn $\dim(\operatorname{Kern}(A)) > 0$ ist, 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 Gleichungsn ausrechnen (Bemerkung 16.4 und Beispiel 16.7).

Ist die Matrix A die Darstellungsmatrix einer linearen Abbildung f, also $A = \mathcal{M}_{B_W}^{B_V}(f)$, dann erhalten wir so auch eine Basis von $\operatorname{Bild}(f)$ und von $\operatorname{Kern}(f)$. Wir müssen lediglich die Basisvektoren von $\operatorname{Bild}(A)$ als Koordinatenvektoren bzgl. der Basis B_W und die Basisvektoren von $\operatorname{Kern}(A)$ als Koordinatenvektoren bzgl. der Basis B_V interpretieren. △

Beispiel 19.11 (zur Bestimmung von Bild und Kern einer Matrix).
Wir betrachten die Punktauswertungsabbildung

$$f\colon \mathbb{R}_3[t] \ni p \mapsto \begin{pmatrix} \widetilde{p}(-2) \\ \widetilde{p}(0) \\ \widetilde{p}(2) \end{pmatrix} \in \mathbb{R}^3$$

aus Beispiel 19.4. Bzgl. der Monombasis $(1,\ t,\ t^2\ t^3)$ in $\mathbb{R}_3[t]$ und der Standardbasis (e_1, e_2, e_3) in $\mathbb{R}^3$ hat dieser Homomorphismus 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

$$\operatorname{Bild}(A) = \left\langle \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \begin{pmatrix} 0\\2\\4 \end{pmatrix}, \begin{pmatrix} 0\\0\\8 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \begin{pmatrix} 0\\1\\2 \end{pmatrix}, \begin{pmatrix} 0\\0\\1 \end{pmatrix} \right\rangle.$$

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

$$\left[\begin{array}{rrrr|r} 1 & -2 & 4 & -8 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 & 0 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{rrrr|r} 1 & -2 & 4 & -8 & 0 \\ 0 & 2 & -4 & 8 & 0 \\ 0 & 0 & 8 & 0 & 0 \end{array}\right],$$

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

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right],$$

an der wir

$$\operatorname{Kern}(A) = \left\langle \begin{pmatrix} 0\\-4\\0\\1 \end{pmatrix} \right\rangle$$

ablesen.[28]

Übersetzen wir $\operatorname{Kern}(A)$ zurück in den Vektorraum $\mathbb{R}_3[t]$, so bedeutet das nach Satz 19.9 Aussage *(iii)*, dass wir $\begin{pmatrix} 0\\-4\\0\\1 \end{pmatrix}$ als Koordinatenvektor bzgl. der gewählten Basis $(1,\ t,\ t^2\ t^3)$ interpretieren müssen. Es sind also genau die Vielfachen des Polynoms

$$p = 0\,1 - 4\,t + 0\,t^2 + 1\,t^3 = -4\,t + t^3,$$

die in $\operatorname{Kern}(f)$ liegen, also die Eigenschaft besitzen, dass alle drei Punktauswertungen Null ergeben. In der Tat gilt

$$\begin{aligned} \widetilde{p}(-2) &= -4 \cdot (-2) + (-2)^3 = 0 \\ \widetilde{p}(0) &= -4 \cdot 0 + 0^3 = 0 \\ \widetilde{p}(2) &= -4 \cdot 2 + 2^3 = 0. \end{aligned}$$

Wir haben also

$$\operatorname{Kern}(f) = \langle -4\,t + t^3 \rangle \subseteq \mathbb{R}_3[t].$$

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 19.4:** Wie äußert sich das in Satz 19.9 Aussage *(i)*?) △

Neben den Zusammenhängen in Satz 19.9 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).

[28] 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.

Satz 19.12 (Invertierbarkeit linearer Abbildungen und ihrer Darstellungsmatrizen).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei endlich-dimensionale Vektorräume über K gleicher Dimension $\dim(V) = \dim(W) = n$ und $n \in \mathbb{N}$. 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 $A \coloneqq \mathcal{M}_{B_W}^{B_V}(f) \in K^{n \times n}$ die Darstellungsmatrix von f. 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 $f^{-1} \colon W \to V$

$$\mathcal{M}_{B_V}^{B_W}(f^{-1}) = A^{-1}. \tag{19.10}$$

Beweis. Die Äquivalenz der Aussagen (i) bis (iii) wurde in Folgerung 18.9 gezeigt. Die Äquivalenz der Aussagen (iv) und (v) wurde in Satz 15.40 gezeigt. Nach Satz 19.9 gilt $\operatorname{Rang}(f) = \operatorname{Rang}(A)$, also sind auch Aussagen (ii) und (v) äquivalent. Ebenfalls nach Satz 19.9 gilt $\operatorname{Defekt}(f) = \operatorname{Defekt}(A)$, also sind auch Aussagen (iii) und (vi) äquivalent.

Ist f bijektiv, dann gilt

$$\begin{aligned}
& f^{-1} \circ f = \mathrm{id}_V \\
\Rightarrow\quad & \mathcal{M}_{B_V}^{B_V}(f^{-1} \circ f) = \mathcal{M}_{B_V}^{B_V}(\mathrm{id}_V) \\
\Rightarrow\quad & \mathcal{M}_{B_V}^{B_W}(f^{-1})\, \mathcal{M}_{B_W}^{B_V}(f) = I_n \qquad \text{nach Satz 19.8 und Beispiel 19.4} \\
\Rightarrow\quad & \mathcal{M}_{B_V}^{B_W}(f^{-1})\, A = I_n.
\end{aligned}$$

Das ist nach Satz 15.42 bereits ausreichend, um

$$A^{-1} = \mathcal{M}_{B_V}^{B_W}(f^{-1}),$$

also (19.10) zu bestätigen. □

§ 19.4 Darstellungsmatrizen von Endomorphismen

Wir betrachten abschließend noch einmal die Endomorphismen eines endlich-dimensionalen Vektorraumes V über einem Körper K. $\operatorname{End}(V, +, \cdot)$ bildet als Spezialfall von $(\operatorname{Hom}(V, V), +, \cdot)$ natürlich einen Vektorraum (Satz 17.11). Endomorphismen können daher wie andere Homomorphismen auch mit Hilfe ihrer Darstellungsmatrix beschrieben werden. Dafür müssen wir eine Basis im Definitionsraum und eine Basis im Zielraum wählen. Obwohl Definitions- und Zielraum bei einem Endomorphismus identisch sind, heißt das nicht, dass wir dieselbe Basis wählen müssten. Es gibt aber einen guten Grund, das zu tun.

Wie wir in Satz 17.12 nachgewiesen haben, bildet $\operatorname{End}(V, +, \circ)$ mit der Komposition $\circ$ einen Ring mit Einselement id_V. Insbesondere ist $\operatorname{End}(V, \circ)$ eine (i. A. nicht-abelsche) Gruppe. Besitzen die Endomorphismen f und g die Darstellungsmatrizen A und B, so hätten wir gerne, dass $f \circ g$ durch die Matrix $A\,B$

dargestellt wird. Das gilt infolge von Satz 19.8 auch, allerdings nur unter der Voraussetzung, dass wir in beiden Kopien des Vektorraumes V (als Definitions- und als Zielraum) dieselbe Basis verwenden!

Unter dieser Voraussetzung gilt folgendes Resultat:

Satz 19.13 (Darstellungssatz für Endomorphismen)**.**
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $B_V = (v_1, \ldots, v_n)$ eine Basis von V. Dann gilt:

(*i*) Die Zuordnung
$$\mathcal{M}_{B_V}^{B_V} \colon (\operatorname{End}(V),+,\circ) \ni f \mapsto \mathcal{M}_{B_V}^{B_V}(f) \in (K^{n\times n},+,\cdot) \tag{19.11}$$
eines Endomorphismus zu seiner Darstellungsmatrix ist ein Isomorphismus von Ringen mit Eins.

(*ii*) Die Abbildung
$$\mathcal{M}_{B_V}^{B_V} \colon (\operatorname{Aut}(V),\circ) \ni f \mapsto \mathcal{M}_{B_V}^{B_V}(f) \in (\operatorname{GL}(n,K),\cdot) \tag{19.12}$$
bildet außerdem die Untergruppe der Automorphismen von V, kurz: $\operatorname{Aut}(V)$, bijektiv auf die Untergruppe der invertierbaren $n \times n$-Matrizen, $\operatorname{GL}(n,K)$, ab.

Beweis. Aussage (*i*): Die additive Verträglichkeit von $\mathcal{M}_{B_V}^{B_V} \colon (\operatorname{End}(V),+,\circ) \to (K^{n\times n},+,\cdot)$, also
$$\mathcal{M}_{B_V}^{B_V}(f+g) = \mathcal{M}_{B_V}^{B_V}(f) + \mathcal{M}_{B_V}^{B_V}(g) \quad \text{für alle } f,g \in \operatorname{End}(V),$$
folgt als Spezialfall von Satz 19.5. Die multiplikative Verträglichkeit, also
$$\mathcal{M}_{B_V}^{B_V}(f \circ g) = \mathcal{M}_{B_V}^{B_V}(f) \cdot \mathcal{M}_{B_V}^{B_V}(g) \quad \text{für alle } f,g \in \operatorname{End}(V),$$
folgt als Spezialfall von Satz 19.8. Schließlich gilt
$$\mathcal{M}_{B_V}^{B_V}(\mathrm{id}_V) = I_n,$$
siehe Beispiel 19.4.

Aussage (*ii*): Nach Satz 19.12 bildet $\mathcal{M}_{B_V}^{B_V}$ tatsächlich $(\operatorname{Aut}(V),\circ) \to (\operatorname{GL}(n,K),\cdot)$ bijektiv ab. □

Ende der Vorlesung 27

Ende der Woche 13

§ 20 Basiswechsel und Normalformen von Darstellungsmatrizen

Literatur: Fischer, Springborn, 2020, Kapitel 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

Die Darstellung eines Vektors v eines endlich-dimensional 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 und insbesondere Endomorphismen 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 die Basen wechseln?

(3) In welcher Basis hat die Darstellungsmatrix besonders einfache Gestalt, die weitergehende Struktureinsichten zulässt?

Die Beantwortung der ersten beiden Fragen gelingt mit Hilfe von **Transformationsmatrizen** für den Basiswechsel.

Definition 20.1 (Transformationsmatrix).
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Es seien $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ zwei Basen von V.[29] Dann heißt

$$\mathcal{T}^{B_V}_{\widehat{B}_V} \coloneqq \mathcal{M}^{B_V}_{\widehat{B}_V}(\mathrm{id}_V) \in K^{n\times n} \tag{20.1}$$

die **Transformationsmatrix des Basiswechsels**, **Übergangsmatrix** oder **Basiswechselmatrix von B_V nach $\widehat{B}_V$** (englisch: transformation matrix, transition matrix, change-of-basis matrix). △

Die Transformationsmatrix $\mathcal{T}^{B_V}_{\widehat{B}_V}$ ist also nichts anderes als die Darstellung der Identitätsabbildung bzgl. der „alten" Basis B_V im Definitionsraum V und der „neuen" Basis $\widehat{B}_V$ im Zielraum V. Die Einträge t_{ij} der Transformationsmatrix $\mathcal{T}^{B_V}_{\widehat{B}_V}$ bestimmen sich daher aus den Bedingungen

$$\underbrace{\mathrm{id}_V(v_j) = v_j}_{\text{Bild des „alten" Basisvektors } v_j \text{ unter der } \mathrm{id}_V\text{-Abbildung}} = \sum_{i=1}^{n} t_{ij}\, \widehat{v}_i \quad \text{für } j = 1, \ldots, n. \tag{20.2}$$

Wir müssen also die „alten" Basisvektoren als Linearkombinationen der „neuen" darstellen.

Beispiel 20.2 (Transformationsmatrix).
Wir betrachten den Vektorraum $V = \mathbb{R}_2[t]$ der Polynome vom Höchstgrad 2 über $\mathbb{R}$ mit der „alten" Basis $B_V = (1,\ t,\ t^2)$ und der „neuen" Basis $\widehat{B}_V = (t^2 - t + 1,\ t^2 + 3,\ t + 1)$. Dann ist die Transformationsmatrix $\mathcal{T}^{B_V}_{\widehat{B}_V} = (t_{ij})$ gegeben durch die Bedingungen

$$\begin{aligned}
1 &= t_{11}\,(t^2 - t + 1) + t_{21}\,(t^2 + 3) + t_{31}\,(t + 1),\\
t &= t_{12}\,(t^2 - t + 1) + t_{22}\,(t^2 + 3) + t_{32}\,(t + 1),\\
t^2 &= t_{13}\,(t^2 - t + 1) + t_{23}\,(t^2 + 3) + t_{33}\,(t + 1),
\end{aligned}$$

was durch Koeffizientenvergleich als lineares Gleichungssystem mit drei rechten Seiten

$$\begin{bmatrix} 1 & 3 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

geschrieben werden kann. Dessen eindeutige Lösung ist

$$\mathcal{T}^{B_V}_{\widehat{B}_V} = \begin{bmatrix} -1 & 1 & 3 \\ 1 & -1 & -2 \\ -1 & 2 & 3 \end{bmatrix}.$$

[29] Wir verwenden die Konvention, dass wir zumindest anfänglich die „alte" Basis in blau und die „neue" Basis in rot kennzeichnen.

Zur Probe rechnen wir Beispiel 19.2 nach, in dem wir das Polynom $7\,t^2 - 3\,t + 5$ mit dem bekannten Koordinatenvektor $x = \left(\begin{smallmatrix} 5 \\ -3 \\ 7 \end{smallmatrix}\right)$ bzgl. der „alten“ Basis in der „neuen“ Basis dargestellt hatten. Für den „neuen“ Koordinatenvektor gilt

$$\widehat{x} = \mathcal{T}_{\widehat{B}_V}^{B_V} x = \begin{bmatrix} -1 & 1 & 3 \\ 1 & -1 & -2 \\ -1 & 2 & 3 \end{bmatrix} \begin{pmatrix} 5 \\ -3 \\ 7 \end{pmatrix} = \begin{pmatrix} 13 \\ -6 \\ 10 \end{pmatrix},$$

vgl. Beispiel 19.2. △

Das folgende Resultat beantwortet insbesondere die Frage (1).

Lemma 20.3 (Eigenschaften von Transformationsmatrizen)**.**

Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Es 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 $x \in K^n$ der Koordinatenvektor eines Vektors $v \in V$ bzgl. der Basis B_V, dann ist $\widehat{x} = \mathcal{T}_{\widehat{B}_V}^{B_V} x$ der Koordinatenvektor von v bzgl. der Basis $\widehat{B}_V$.[30]

(*ii*) Die Transformationsmatrix $\mathcal{T}_{\widehat{B}_V}^{B_V} \in K^{n \times n}$ ist invertierbar.

(*iii*) $\left(\mathcal{T}_{\widehat{B}_V}^{B_V}\right)^{-1} = \mathcal{T}_{B_V}^{\widehat{B}_V}$.

(*iv*) Die von der Matrix $\mathcal{T}_{\widehat{B}_V}^{B_V}$ induzierte lineare Abbildung $K^n \to K^n$ ist $\Phi_{\widehat{B}_V}^{-1} \circ \Phi_{B_V} \in \operatorname{Aut}(K^n)$.

Beweis. Aussage (*i*): Es gilt

$$\begin{aligned} v &= \sum_{j=1}^n x_j\, v_j && \text{nach Voraussetzung} \\ &= \sum_{j=1}^n x_j \sum_{i=1}^n t_{ij}\, \widehat{v}_i && \text{nach (20.2)} \\ &= \sum_{i=1}^n \underbrace{\sum_{j=1}^n (t_{ij}\, x_j)}_{\text{Koeffizient } \widehat{x}_i \text{ bzgl. der Basis } \widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)} \widehat{v}_i && \text{wegen Distributivität und Kommutativität im Körper } K. \end{aligned}$$

Nach Definition des Matrix-Vektor-Produkts (Bemerkung 15.9) gilt also $\widehat{x} = \mathcal{T}_{\widehat{B}_V}^{B_V} x$.

Aussage (*ii*): Die Invertierbarkeit von $\mathcal{T}_{\widehat{B}_V}^{B_V}$ folgt aus Satz 19.12, da $\mathcal{T}_{\widehat{B}_V}^{B_V}$ die Darstellungsmatrix des bijektiven Homomorphismus $\mathrm{id}_V : V \to V$ ist.

[30] Beachte das Muster $\widehat{x} = \mathcal{T}_{\widehat{B}_V}^{B_V} x$ mit alter und neuer Basis und zugehörigen Koordinatenvektoren.

Aussage (iii): Die Eigenschaft $\left(\mathcal{T}_{\widehat{B}_V}^{B_V}\right)^{-1} = \mathcal{T}_{B_V}^{\widehat{B}_V}$ folgt ebenfalls aus Satz 19.12, denn es gilt

$$\begin{aligned}
\mathcal{T}_{B_V}^{\widehat{B}_V} &= \mathcal{M}_{B_V}^{\widehat{B}_V}(\mathrm{id}_V) && \text{nach Definition von Transformationsmatrizen} \\
&= \mathcal{M}_{B_V}^{\widehat{B}_V}((\mathrm{id}_V)^{-1}) && \mathrm{id}_V \text{ ist selbstinvers} \\
&= \left(\mathcal{M}_{\widehat{B}_V}^{B_V}(\mathrm{id}_V)\right)^{-1} && \text{nach Satz 19.12, (19.10)} \\
&= \left(\mathcal{T}_{\widehat{B}_V}^{B_V}\right)^{-1} && \text{nach Definition von Transformationsmatrizen.}
\end{aligned}$$

Aussage (iv): Mit $A \coloneqq \mathcal{T}_{\widehat{B}_V}^{B_V}$ ist folgendes Diagramm kommutativ (vgl. Satz 19.7):

$$\begin{tikzcd}
V \arrow[d, "\Phi_{\widehat{B}_V}^{-1}"'] & V \arrow[l, "\mathrm{id}_V"'] \\
K^n & K^n \arrow[l, "f_A"] \arrow[u, "\Phi_{B_V}"']
\end{tikzcd}$$

Also ist

$$f_A = \Phi_{\widehat{B}_V}^{-1} \circ \mathrm{id}_V \circ \Phi_{B_V} = \underbrace{\Phi_{\widehat{B}_V}^{-1}}_{\text{„neue“ Koordinaten} \hookleftarrow \text{Vektor}} \circ \underbrace{\Phi_{B_V}}_{\text{Vektor} \hookleftarrow \text{„alte“ Koordinaten}}$$

die von A induzierte Abbildung $K^n \to K^n$. □

Folgerung 20.4 (Eigenschaften von Transformationsmatrizen in K^n).
Im Fall $V = K^n$ können wir die Basen $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ als Matrizen

$$B_V = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \quad \text{und} \quad \widehat{B}_V = \begin{bmatrix} | & & | \\ \widehat{v}_1 & \cdots & \widehat{v}_n \\ | & & | \end{bmatrix}$$

in $K^{n \times n}$ auffassen. Dann gilt

$$\mathcal{T}_{\widehat{B}_V}^{B_V} = \left(\widehat{B}_V\right)^{-1} B_V. \tag{20.3}$$

Ist insbesondere eine der Basen die Standardbasis $(e_1, \ldots, e_n)$, so gilt

$$\mathcal{T}_{(e_1,\ldots,e_n)}^{B_V} = B_V \quad \text{bzw.} \quad \mathcal{T}_{\widehat{B}_V}^{(e_1,\ldots,e_n)} = \left(\widehat{B}_V\right)^{-1}. \tag{20.4}$$

Beweis. In diesem Fall gilt $f_{B_V}^{-1} = \Phi_{B_V}$ und $f_{\widehat{B}_V}^{-1} = \Phi_{\widehat{B}_V}$ (**Quizfrage 20.1:** Warum?), also können wir das Diagramm aus dem Beweis von Lemma 20.3 schreiben als

$$\begin{tikzcd}
K^n \arrow[d, "\Phi_{\widehat{B}_V}^{-1} = f_{\widehat{B}_V}^{-1}"'] & K^n \arrow[l, "\mathrm{id}_{K^n}"'] \\
K^n & K^n \arrow[l, "f_A"] \arrow[u, "\Phi_{B_V} = f_{B_V}"']
\end{tikzcd}$$

Wir erhalten

$$\begin{aligned}
f_A &= \Phi_{\widehat{B}_V}^{-1} \circ \Phi_{B_V} && \text{nach Lemma 20.3 } (iv) \\
&= f_{\widehat{B}_V}^{-1} \circ f_{B_V} && \text{denn } f_{B_V} = \Phi_{B_V} \text{ und } f_{\widehat{B}_V} = \Phi_{\widehat{B}_V} \\
&= f_{(\widehat{B}_V)^{-1}} \circ f_{B_V} && \text{nach Lemma 17.9} \\
&= f_{(\widehat{B}_V)^{-1} B_V} && \text{nach Lemma 17.9.}
\end{aligned}$$

Weil f_A die Matrix A eindeutig festlegt (nochmal Lemma 17.9), folgt $\mathcal{T}_{\widehat{B}_V}^{B_V} = A = (\widehat{B}_V)^{-1} B_V$, also (20.3). Die Spezialfälle (20.4) gelten, weil die zur Standardbasis gehörige Matrix $\widehat{B}_V$ bzw. B_V die Einheitsmatrix ist. □

Beispiel 20.5 (Transformationsmatrizen in K^n)**.**

Wir wollen im Raum $V = \mathbb{R}^2$ über dem Körper $\mathbb{R}$ die Transformationsmatrix $\mathcal{T}_{\widehat{B}_V}^{B_V}$ von der Basis $B_V = \left(\binom{-1}{1}, \binom{1}{1}\right)$ zur Basis $\widehat{B}_V = \left(\binom{2}{-1}, \binom{1}{-1}\right)$ finden. Nach (20.3) gilt

$$\mathcal{T}_{\widehat{B}_V}^{B_V} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -1 & -3 \end{bmatrix}.$$ △

§ 20.1 Transformation der Darstellungsmatrizen von Homomorphismen

Wir kommen nun zur Frage (2).

Satz 20.6 (Transformation der Darstellungsmatrix eines Homomorphismus beim Wechsel der Basen)**.**

Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$, $(W, +, \cdot)$ zwei endlich-dimensionale Vektorräume über 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 $f\colon V \to W$:[31]

$$\mathcal{M}_{\widehat{B}_W}^{\widehat{B}_V}(f) = \mathcal{T}_{\widehat{B}_W}^{B_W} \, \mathcal{M}_{B_W}^{B_V}(f) \, \mathcal{T}_{B_V}^{\widehat{B}_V}. \tag{20.5}$$

Beweis. Wir setzen zur Abkürzung

$$\begin{aligned}
T &\coloneqq \mathcal{T}_{\widehat{B}_V}^{B_V} && \text{Übergang von } B_V \text{ nach } \widehat{B}_V, \\
\text{also } T^{-1} &= \mathcal{T}_{B_V}^{\widehat{B}_V} && \text{Übergang von } \widehat{B}_V \text{ nach } B_V, \\
A &\coloneqq \mathcal{M}_{B_W}^{B_V}(f) && \text{„alte" Darstellungsmatrix von } f, \\
S &\coloneqq \mathcal{T}_{\widehat{B}_W}^{B_W} && \text{Übergang von } B_W \text{ nach } \widehat{B}_W, \\
\widehat{A} &\coloneqq \mathcal{M}_{\widehat{B}_W}^{\widehat{B}_V}(f) && \text{„neue" Darstellungsmatrix von } f.
\end{aligned}$$

Wir betrachten das Diagramm

[31]Die Farben deuten wieder den Übergang von den „alten" Basen zu den „neuen" Basen an.

$$\begin{tikzcd}
K^n & & & K^m \arrow[lll, "f_{\widehat{A}}"'] \arrow[dl, "\Phi_{\widehat{B}_V}"'] \arrow[dd, "f_T^{-1}=\Phi^{-1}_{B_V}\circ\Phi_{\widehat{B}_V}"] \\
& W \arrow[ul, "\Phi^{-1}_{\widehat{B}_W}"'] \arrow[dl, shift left, "\Phi^{-1}_{B_W}"] & V \arrow[l, "f"'] \arrow[dr, shift left, "\Phi^{-1}_{B_V}"] & \\
K^n \arrow[uu, "f_S=\Phi^{-1}_{\widehat{B}_W}\circ\Phi_{B_W}"] \arrow[ur, shift left, "\Phi_{B_W}"] & & & K^m \arrow[lll, "f_A"] \arrow[ul, shift left, "\Phi_{B_V}"]
\end{tikzcd}$$

Das obere und das untere Trapez sind kommutativ nach Satz 19.7. Das linke Dreieck und das rechte Dreieck sind kommutativ nach Lemma 20.3. Damit kommutiert auch das äußere Rechteck, d. h., es gilt

$$f_{\widehat{A}} = f_S \circ f_A \circ f_T^{-1} = f_S \circ f_A \circ f_{T^{-1}} = f_{S\,A\,T^{-1}}.$$

Das heißt nach Lemma 17.9 aber gerade

$$\widehat{A} = S\,A\,T^{-1}, \tag{20.6}$$

also die Behauptung (20.5). □

Beispiel 20.7 (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 die Darstellungsmatrix $\mathcal{M}^{(e_1,e_2)}_{(e_1,e_2,e_3)}(f_A)$ bzgl. der Standardbasen in $\mathbb{R}^2$ und $\mathbb{R}^3$. Wir wollen die Darstellungsmatrix 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

$$\widehat{A} = \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_W}(f_A) = S\,A\,T^{-1}$$

bestimmen. Es gilt

$$T^{-1} := \mathcal{T}^{\widehat{B}_V}_{(e_1,e_2)} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

und

$$S := \mathcal{T}^{(e_1,e_2,e_3)}_{\widehat{B}_W} = \left[\mathcal{T}^{\widehat{B}_W}_{(e_1,e_2,e_3)}\right]^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Wir erhalten also die transfomierte Darstellungsmatrix bzgl. der neuen Basen gemäß (20.5)

$$\widehat{A} = S\,A\,T^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 7 \\ 1 & 15 \end{bmatrix}.$$ △

Definition 20.8 (Äquivalenztransformation, äquivalente Matrizen)**.**
Es seien $(K,+,\cdot)$ 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:

$$\widehat{A} = S\,A\,T^{-1}. \tag{20.7}$$

Der Übergang von A zu $S\,A\,T^{-1}$, 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 $K^{n \times m}$.

Satz 20.9 (über äquivalente Matrizen)**.**
Es seien $(K, +, \cdot)$ ein Körper, $n, m \in \mathbb{N}_0$ und $A, \widehat{A} \in K^{n \times m}$. Weiter seien V und W zwei Vektorräume über K mit $\dim(V) = m$ und $\dim(W) = n$ und B_V eine Basis von V, B_W eine Basis von W und $f \colon V \to W$ ein Homomorphismus und $A = \mathcal{M}_{B_W}^{B_V}(f)$.[32] Dann sind äquivalent:

(i) A und $\widehat{A}$ sind äquivalente Matrizen.

(ii) $\widehat{A}$ ist die Darstellungsmatrix von f bzgl. geeigneter Basen $\widehat{B}_V$ und $\widehat{B}_W$.

(iii) Es gilt $\operatorname{Rang}(A) = \operatorname{Rang}(\widehat{A})$.

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 $\widehat{A} = S\,A\,T^{-1}$ gilt. Wir können T^{-1} als Transformationsmatrix eines Basiswechsels $T^{-1} = \mathcal{T}_{B_V}^{\widehat{B}_V}$ auffassen, indem wir die „neue" Basis von V durch

$$\widehat{v}_j = \sum_{i=1}^{m} [T^{-1}]_{ij}\, v_i \quad \text{für } j = 1, \ldots, m$$

definieren, vgl. (20.2). Ebenso können wir S als Transformationsmatrix eines Basiswechsels $S = \mathcal{T}_{\widehat{B}_W}^{B_W}$ auffassen, indem wir die „neue" Basis von W durch

$$\widehat{w}_j = \sum_{i=1}^{n} [S^{-1}]_{ij}\, w_i \quad \text{für } j = 1, \ldots, n$$

definieren. Wir sehen jetzt

$$\begin{aligned}
f_{\widehat{A}} &= f_{S\,A\,T^{-1}} && \text{nach Voraussetzung} \\
&= f_S \circ f_A \circ f_{T^{-1}} && \text{nach Lemma 17.9} \\
&= f_S \circ \Phi_{B_W}^{-1} \circ f \circ \Phi_{B_V} \circ f_{T^{-1}} && \text{nach Satz 19.7} \\
&= \Phi_{\widehat{B}_W}^{-1} \circ \Phi_{B_W} \circ \Phi_{B_W}^{-1} \circ f \circ \Phi_{B_V} \circ \Phi_{B_V}^{-1} \circ \Phi_{\widehat{B}_V} && \text{nach Lemma 20.3} \\
&= \Phi_{\widehat{B}_W}^{-1} \circ f \circ \Phi_{\widehat{B}_V}.
\end{aligned}$$

Das bedeutet wiederum nach Satz 19.7 aber, dass $\widehat{A}$ die Darstellungsmatrix von f bzgl. der „neuen" Basen $\widehat{V}$ und $\widehat{W}$ ist.

Aussage (ii) $\Rightarrow$ Aussage (i): Das folgt sofort aus Satz 20.6.

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 $\widehat{A} = S\,A\,T^{-1}$ gilt. Nach Folgerung 15.41 gilt $\operatorname{Rang}(\widehat{A}) = \operatorname{Rang}(A)$.

[32] Diese Voraussetzung ist nicht einschränkend. Zu einer gegebenen Matrix A können wir immer $V = K^m$, $W = K^n$, B_V und B_W als die Standardbasen und $f = f_A$ wählen.

Aussage (*iii*) $\Longrightarrow$ Aussage (*i*): Wir betrachten $f_A\colon K^m \to K^n$. Nach Satz 19.9 gilt $\operatorname{Rang}(f_A) = \operatorname{Rang}(A)$. Gemäß Beispiel 19.4 können wir eine Basis B von K^m und eine Basis C von K^n finden, sodass

$$
\begin{aligned}
\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] &= \mathcal{M}_B^C(f_A) \\
&= \mathcal{T}_C^{(e_1,\ldots,e_n)}\, \mathcal{M}_{(e_1,\ldots,e_n)}^{(e_1,\ldots,e_m)}(f_A)\, \mathcal{T}_{(e_1,\ldots,e_m)}^B \\
&= \underbrace{\mathcal{T}_C^{(e_1,\ldots,e_n)}}_{=:S}\, A\, \underbrace{\mathcal{T}_{(e_1,\ldots,e_m)}^B}_{=:T^{-1}} \\
&= S\,A\,T^{-1}
\end{aligned}
$$

mit invertierbaren Matrizen $S \in K^{n\times n}$ und $T \in K^{m\times m}$ gilt. Dasselbe Argument können wir für $f_{\widehat{A}}$ verwenden, um zu zeigen, dass es Matrizen $\widehat{S}$ und $\widehat{T}$ gibt, sodass

$$
\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] = \widehat{S}\,\widehat{A}\,\widehat{T}^{-1}
$$

gilt. Zusammen also:

$$
\widehat{A} = (\widehat{S}^{-1}\,S)\,A\,(T^{-1}\,\widehat{T}).
$$

Da die Produkte $\widehat{S}^{-1}S$ und $T^{-1}\widehat{T}$ wieder invertierbare Matrizen sind, folgt die Äquivalenz von A und $\widehat{A}$. □

Bemerkung 20.10 (über äquivalente Matrizen)**.**
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$, $(W,+,\cdot)$ zwei endlich-dimensionale Vektorräume über K mit $\dim(V) = m$ und $\dim(W) = n$. Der Satz 20.9 bedeutet, folgendes Diagramm kommutiert:

$$
\begin{tikzcd}
\operatorname{Hom}(V,W) \arrow[r, "\mathcal{M}_{B_W}^{B_V}", "\text{isomorph}"'] \arrow[d, "\pi"'] & K^{n\times m} \arrow[d, "\pi"] \\
\operatorname{Hom}(V,W)\,/\sim_{\text{Rang}} \arrow[r, "\text{bijektiv}", "\mathcal{M}_{B_W}^{B_V}"'] & K^{n\times m}\,/\sim_{\text{Rang}}
\end{tikzcd}
$$

Dabei ist auf der linken Seite des Diagramms $\sim_{\text{Rang}}$ die Äquivalenzrelation „gleicher Rang" auf $\operatorname{Hom}(V,W)$ mit den Äquivalenzklassen

$$
[f] \coloneqq \big\{ g \in \operatorname{Hom}(V,W) \,\big|\, \operatorname{Rang}(g) = \operatorname{Rang}(f) \big\}.
$$

Weiter ist $\operatorname{Hom}(V,W)\,/\sim_{\text{Rang}}$ die zugehörige Faktormenge (Menge aller Äquivalenzklassen) und π die kanonische Surjektion $f \mapsto [f]$. (**Quizfrage 20.2:** Wieviele Äquivalenzklassen gibt es?)

Auf der rechten Seite des Diagramms ist $\sim_{\text{Rang}}$ die Äquivalenzrelation „gleicher Rang" auf $K^{n,m}$ mit den Äquivalenzklassen

$$
\begin{aligned}
[A] &\coloneqq \big\{ \widehat{A} \in K^{n\times m} \,\big|\, \operatorname{Rang}(\widehat{A}) = \operatorname{Rang}(A) \big\} \\
&= \big\{ \widehat{A} \in K^{n\times m} \,\big|\, \widehat{A} \text{ ist äquivalent zu } A \big\} \\
&= \big\{ S\,A\,T^{-1} \,\big|\, S \in K^{n\times n} \text{ und } T \in K^{m\times m} \text{ sind invertierbar} \big\}.
\end{aligned}
$$

Weiter ist $K^{n\times m} / \sim_{\text{Rang}}$ die zugehörige Faktormenge und π die kanonische Surjektion $A \mapsto [A]$.

Nach Satz 20.9 bildet $\mathcal{M}_{B_W}^{B_V}$ eine gesamte Äquivalenzklasse von $\operatorname{Hom}(V, W)$ auf eine Äquivalenzklasse von $K^{n\times m}$ ab, und verschiedene Äquivalenzklassen von $\operatorname{Hom}(V, W)$ 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 $\operatorname{Hom}(V, W)$ eine sogenannte **glatte Mannigfaltigkeit** (englisch: smooth manifold).[33] Ebenso ist jede Äquivalenzklasse $[A]$ von $K^{n\times m}$ eine glatte Mannigfaltigkeit, und die Abbildung $\mathcal{M}_{B_W}^{B_V}$ ist ein **Diffeomorphismus** zwischen $[f]$ und $[A]$ mit $\operatorname{Rang}(f) = \operatorname{Rang}(A)$. △

Folgerung 20.11 (Normalform der Darstellungsmatrix eines Homomorphismus)**.**
Es seien $(K, +, \cdot)$ 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:

$$S\,A\,T^{-1} = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \in K^{n\times m}. \tag{20.8}$$

Beweis. Dieses Resultat wurde beim Beweis von Satz 20.9 Aussage *(iii)* ⇒ Aussage *(i)* gezeigt; siehe auch Beispiel 19.4. □

Wir nennen die zu A äquivalente Matrix (20.8) die **Rang-Normalform** (englisch: rank normal form) von A. Sie ist eindeutig bestimmt. Eine Matrix in Rang-Normalform kann als natürlicher Repräsentant ihrer Äquivalenzklasse bzgl. der Äquivalenzrelation „gleicher Rang“ angesehen werden. Damit haben wir für allgemeine Homomorphismen auch die Frage (3) beantwortet.

Ende der Vorlesung 28

§ 20.2 Transformation der Darstellungsmatrizen von Endomorphismen

In § 19.4 hatten wir gesehen, dass es sinnvoll ist, bei der Darstellung von Endomorphismen $f : V \to V$ eines endlich-dimensionalen Vektorraumes V durch Darstellungsmatrizen im Definitions- und im Zielraum zweimal dieselbe Basis B_V zu verwenden. (**Quizfrage 20.3:** Was war nochmal der Grund dafür?)

Analog zum Transformationssatz für Darstellungsmatrizen von Homomorphismen 20.6 erhalten wir also als Sonderfall folgende Variante für Endomorphismen:

Satz 20.12 (Transformation der Darstellungsmatrix eines Endomorphismus beim Wechsel der Basis)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein endlich-dimensionaler Vektorraum über K. Weiter seien B_V und $\widehat{B}_V$ Basen von V. Dann gilt für die Darstellungsmatrix eines Endomorphismus $f : V \to V$:[34]

$$\mathcal{M}_{\widehat{B}_V}^{\widehat{B}_V}(f) = \mathcal{T}_{\widehat{B}_V}^{B_V}\,\mathcal{M}_{B_V}^{B_V}(f)\,\mathcal{T}_{B_V}^{\widehat{B}_V}. \tag{20.9}$$

[33] Über solche Strukturen erfahren Sie mehr in Vorlesungen über Differentialgeometrie.
[34] Die Farben deuten wieder den Übergang von der „alten“ Basis zur „neuen“ Basis an.

Das bedeutet, dass wir mit der Notation aus § 20.1

$$\begin{aligned} T &:= \mathcal{T}^{B_V}_{\widehat{B}_V} && \text{Übergang von } B_V \text{ nach } \widehat{B}_V, \\ \text{also } T^{-1} &= \mathcal{T}^{\widehat{B}_V}_{B_V} && \text{Übergang von } \widehat{B}_V \text{ nach } B_V, \\ A &:= \mathcal{M}^{B_V}_{B_V}(f) && \text{„alte“ Darstellungsmatrix von } f, \\ \widehat{A} &:= \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_V}(f) && \text{„neue“ Darstellungsmatrix von } f, \end{aligned}$$

die Beziehung

$$\widehat{A} = T\,A\,T^{-1} \tag{20.10}$$

erhalten.

Definition 20.13 (Ähnlichkeitstransformation, ähnliche Matrizen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}_0$. Zwei Matrizen $A, \widehat{A} \in K^{n \times n}$ heißen **ähnlich** (englisch: similar), wenn es eine invertierbare Matrix $T \in K^{n \times n}$ gibt, sodass gilt:

$$\widehat{A} = T\,A\,T^{-1}. \tag{20.11}$$

Der Übergang von A zu $T\,A\,T^{-1}$, also die Multiplikation von links und von rechts einmal mit einer invertierbaren Matrix und einmal mit ihrer Inversen, heißt auch eine **Ähnlichkeitstransformation** (englisch: similarity transformation) von A. △

Beachte: Die Ähnlichkeit von Matrizen ist eine Äquivalenzrelation auf der Menge $K^{n \times n}$.

Analog zu Satz 20.9 halten wir fest:

Satz 20.14 (über ähnliche Matrizen, vgl. Satz 20.9)**.**
Es seien $(K, +, \cdot)$ ein Körper, $n \in \mathbb{N}_0$ und $A, \widehat{A} \in K^{n \times n}$. Weiter sei V ein Vektorraum über K mit $\dim(V) = n$ und B_V eine Basis von V, $f\colon V \to V$ ein Endomorphismus und $A = \mathcal{M}^{B_V}_{B_V}(f)$.[35] Dann sind äquivalent:

(i) A und $\widehat{A}$ sind ähnliche Matrizen.

(ii) $\widehat{A}$ ist die Darstellungsmatrix von f bzgl. einer geeigneten Basis $\widehat{B}_V$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Der Beweis ist ein Spezialfall des entsprechenden Beweisschrittes von Satz 20.9.

Aussage (ii) $\Rightarrow$ Aussage (i): Das folgt sofort aus Satz 20.12. □

Es stellt sich an dieser Stelle allerdings die Frage, ob wir nicht zusätzlich zu der relativ offensichtlichen Aussage von Satz 20.14 noch eine vergleichbar einfache Charakterisierung der Ähnlichkeit zweier Matrizen finden können, wie es mit der Ranggleichheit bei der Äquivalenz in Satz 20.9 der Fall war. Es stellt sich allerdings heraus, dass es bei der Ähnlichkeit von Matrizen auf sehr viel komplexere Struktur des dargestellten Endomorphismus ankommt, als das beim Rang einer Matrix bzw. dem Rang des dargestellten Homomorphismus der Fall war. Auch die Frage (3) nach einer geeigneten Basis, bzgl. der die Darstellungsmatrix eines Endomorphismus eine möglichst einfache Gestalt hat, also ein Analogon

[35] Diese Voraussetzung ist nicht einschränkend. Zu einer gegebenen Matrix A können wir immer $V = K^n$, B_V als die Standardbasis und $f = f_A$ wählen.

zu Folgerung 20.11, ist für uns noch offen. Da diese Frage eng mit der Charakterisierung der Ähnlichkeit zusammenhängt, ist auch hier die Beantwortung deutlich schwieriger als bei der Rang-Normalform von Homomorphismen. Ein weiterer Hinweis darauf wird durch die Tatsache gegeben, dass wir zur Herstellung einer wie auch immer gearteten einfachen Form der Darstellungsmatrix jetzt nur noch eine Basis, also nur noch eine Transformationsmatrix zur Verfügung haben, um $T\,A\,T^{-1}$ in möglichst einfache Form zu bringen.

Wir werden die gerade gestellten Fragen erst im zweiten Teil der Vorlesung vollständig beantworten können. Wir wollen hier jedoch schon einige Überlegungen dazu anstellen und können zumindest für eine Teilmenge von Matrizen bereits jetzt eine Charakterisierung der Ähnlichkeit erreichen. Dazu führen wir zunächst einen neuen Begriff ein.

Definition 20.15 (invarianter Unterraum eines Endomorphismus bzw. einer Matrix)**.**

(i) Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$ ein Vektorraum über K. Weiter sei $f\colon V \to V$ ein Endomorphismus. Ein Unterraum $U \subseteq V$ heißt ein **f-invarianter Unterraum** (englisch: f-invariant subspace), wenn gilt:

$$f(U) \subseteq U. \tag{20.12}$$

(ii) Es seien $(K,+,\cdot)$ ein Körper und $A \in K^{n\times n}$ eine Matrix mit $n \in \mathbb{N}_0$. Ein Unterraum $U \subseteq K^n$ heißt ein **A-invarianter Unterraum**, wenn gilt:[36]

$$f_A(U) \subseteq U. \tag{20.13}$$

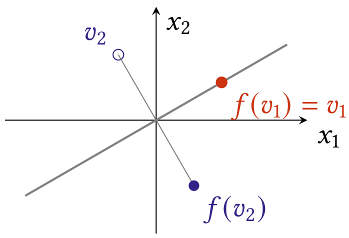
△

Vektoren u in einem f-invarianten Unterraum U haben also die Eigenschaft, dass sie durch f wieder auf einen Vektor $f(u) \in U$ abgebildet werden (auch bei wiederholter Anwendung von f).

Beispiel 20.16 (invarianter Unterraum eines Endomorphismus)**.**

(i) In jedem Vektorraum V sind die trivialen Unterräume $\{0\}$ und V invariante Unterräume für jeden Endomorphismus $f\colon V \to V$.

(ii) Die Spiegelungsabbildung an einer Achse in $\mathbb{R}^2$ aus Beispiel 17.8 hat zwei verschiedene eindimensionale invariante Unterräume, wie wir uns grafisch überlegen können:



Der Vektor $v_1 = \begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}$ wird durch das Matrix-Vektor-Produkt mit der Darstellungsmatrix

$$\begin{aligned}
&\begin{bmatrix}\cos^2(\alpha)-\sin^2(\alpha) & 2\cos(\alpha)\,\sin(\alpha)\\ 2\cos(\alpha)\,\sin(\alpha) & \sin^2(\alpha)-\cos^2(\alpha)\end{bmatrix}\begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}\\
&\quad=\begin{pmatrix}\cos^3(\alpha)-\sin^2(\alpha)\,\cos(\alpha)+2\cos(\alpha)\,\sin^2(\alpha)\\ 2\cos^2(\alpha)\,\sin(\alpha)+\sin^3(\alpha)-\cos^2(\alpha)\,\sin(\alpha)\end{pmatrix}=\begin{pmatrix}\cos(\alpha)\big[\cos^2(\alpha)+\sin^2(\alpha)\big]\\ \sin(\alpha)\big[\cos^2(\alpha)+\sin^2(\alpha)\big]\end{pmatrix}\\
&\quad=\begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}
\end{aligned}$$

[36]Mit anderen Worten: $\{A\,x \mid x \in U\} \subseteq U$ oder kurz: $A \cdot U \subseteq U$.

tatsächlich auf sich selbst abgebildet, während für das Bild des Vektors $v_2 = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$

$$\begin{aligned}
&\begin{bmatrix} \cos^2(\alpha) - \sin^2(\alpha) & 2\cos(\alpha)\ \sin(\alpha) \\ 2\cos(\alpha)\ \sin(\alpha) & \sin^2(\alpha) - \cos^2(\alpha) \end{bmatrix} \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix} \\
&\quad = \begin{pmatrix} -\cos^2(\alpha)\ \sin(\alpha) + \sin^3(\alpha) + 2\cos^2(\alpha)\ \sin(\alpha) \\ -2\cos(\alpha)\ \sin^2(\alpha) + \sin^2(\alpha)\cos(\alpha) - \cos^3(\alpha) \end{pmatrix} = \begin{pmatrix} \sin(\alpha)\big[\cos^2(\alpha) + \sin^2(\alpha)\big] \\ \cos(\alpha)\big[\cos^2(\alpha) + \sin^2(\alpha)\big] \end{pmatrix} \\
&\quad = \begin{pmatrix} \sin(\alpha) \\ \cos(\alpha) \end{pmatrix}
\end{aligned}$$

gilt, d. h., $f(v_2) = -v_2$. Die Unterräume $U_1 = \langle v_1 \rangle$ und $U_2 = \langle v_2 \rangle$ sind also in der Tat zwei verschiedene eindimensionale invariante Unterräume der Spiegelungsabbildung (**Quizfrage 20.4:** klar?). Wegen $U_1 \cap U_2 = \{0\}$ gilt $\mathbb{R}^2 = U_1 \oplus U_2$. In diesem Beispiel kann also der gesamte Vektorraum als direkte Summe invarianter Unterräume des Endomorphismus „Spiegelung“ geschrieben werden.

(*iii*) Für die Ableitungsabbildung als Endomorphismus $f\colon K[t] \to K[t]$ (Beispiel 18.10) sind die invarianten Unterräume genau die Unterräume von der Form $\langle 1, t, t^2, \dots, t^k \rangle = K_k[t]$ für $k \in \mathbb{N}_0$ bzw. der Nullraum. (**Quizfrage 20.5:** Warum?) Wenn wir die Ableitungsabbildung einschränken auf den Polynomraum vom Höchstgrad $n \in \mathbb{N}_0$, also $f\colon K_n[t] \to K_n[t]$, dann sind die invarianten Unterräume noch immer genau die Unterräume der Form $K_k[t]$ für $k \in \mathbb{N}_0$ bzw. der Nullraum. In diesem Beispiel ist also keine Zerlegung des Vektorraumes $K_n[t]$ in eine direkte Summe verschiedener invarianter Unterräume möglich. △

Wenn wir für den Moment annehmen, dass $f\colon V \to V$ ein Endomorphismus auf einem Vektorraum V mit $\dim(V) = n \geqslant 2$ ist und U ein nicht-trivialer f-invarianter Unterraum, d. h., $\dim(U) = k$ mit $1 \leqslant k \leqslant n-1$, dann können wir eine Basis der Form

$$B_V = (\underbrace{v_1, \dots, v_k}_{\text{Basis von } U}, \underbrace{v_{k+1}, \dots, v_n}_{\text{Basis eines zu } U \text{ komplementären Unterraumes } W})$$

verwenden, also $V = U \oplus W$ zerlegen. Bzgl. dieser Basis hat die Darstellungsmatrix von f die Gestalt

$$\mathcal{M}_{B_V}^{B_V}(f) = \begin{array}{cc} \begin{array}{cc} k & n-k \end{array} & \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] & \begin{array}{l} k \\ n-k \end{array} \end{array}$$

mit den angegebenen Dimensionen der Blöcke. An der Struktur der Matrix sieht man sehr schön, dass f auf dem Unterraum U „isoliert“ wirkt, dass also die Einschränkung $f|_U^U$ definiert ist.

Ist zusätzlich auch der zu U komplementäre Unterraum W ein f-invarianter Unterraum, gilt also $f(W) \subseteq W$, dann ist die Darstellungsmatrix von f eine **Blockdiagonalmatrix** (englisch: block diagonal matrix)

$$\mathcal{M}_{B_V}^{B_V}(f) = \begin{array}{cc} \begin{array}{cc} k & n-k \end{array} & \\ \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right] & \begin{array}{l} k \\ n-k \end{array} \end{array}.$$

Hier wirkt also f nun in beiden Unterräumen „isoliert“, sodass auch die Einschränkung $f|_W^W$ definiert ist.

Diese Überlegungen gelten auch für Zerlegungen des Raumes V in direkte Summen von mehr als zwei Unterräumen. Es geht also bei der Aufgabe, eine möglichst einfache Gestalt der Darstellungsmatrix eines Endomorphismus $f\colon V \to V$ zu finden, darum, möglichst niedrig-dimensionale, paarweise verschiedene f-invariante Unterräume von V zu finden, deren direkte Summe den ganzen Raum V ergibt. Der triviale Fall, als einzigen f-invarianten Unterraum V selbst zu nutzen, ist immer möglich, manchmal nicht vermeidbar, bringt aber keine strukturelle Einsicht.

Satz 20.17 (Blockdiagonalgestalt der Darstellungsmatrix eines Endomorphismus)**.**
Es seien $(K,+,\cdot)$ ein Körper und $(V,+,\cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f\colon V \to V$ ein Endomorphismus und $L \in \mathbb{N}_0$. Dann sind äquivalent:

(i) Es existieren f-invariante Unterräume $U_1, \dots, U_L$ der Dimensionen $\dim(U_j) = n_j \in \mathbb{N}_0$, sodass gilt:
$$V = U_1 \oplus \cdots \oplus U_L. \tag{20.14}$$

(ii) Es existiert eine Basis B_V von V, sodass die Darstellungsmatrix von f die Blockdiagonalgestalt
$$\mathcal{M}_{B_V}^{B_V}(f) = \begin{array}{cc} \begin{array}{cccc} n_1 & n_2 & \cdots & n_L \end{array} & \\ \begin{bmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{LL} \end{bmatrix} & \begin{array}{c} n_1 \\ n_2 \\ \vdots \\ n_L \end{array} \end{array} \tag{20.15}$$
besitzt, wobei für die Blöcke $A_{jj} \in K^{n_j \times n_j}$, $j = 1, \dots, L$ gilt.

Beachte: Wie gesagt, der Fall $L = 1$ mit $U_1 = V$ ist immer möglich, bringt aber keine strukturelle Einsicht, da der einzige Block A_{11} in (20.15) die gesamte Matrix ist.

Beweis. □

Beispiel 20.18 (Blockdiagonalgestalt der Darstellungsmatrix eines Endomorphismus)**.**
Das Beispiel der Spiegelungsabbildung aus Beispiel 20.16 in $V = \mathbb{R}^2$ zeigt, dass es Endomorphismen gibt, die die in Satz 20.17 beschriebene Darstellung in Form einer Blockdiagonalmatrix zulassen, die nicht nur aus einem großen Block besteht. Bezüglich der Basis $B_V = \left(\begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}, \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix} \right)$ erhalten wir für die Spiegelungsabbildung die diagonale Darstellungsmatrix (mit Blöcken der minimalen Größe $n_1 = n_2 = 1$)
$$\mathcal{M}_{B_V}^{B_V}(f) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$
△

Wir gehen dem in Beispiel 20.18 beobachteten bestmöglichen Fall, in dem die Darstellung eines Endomorphismus in Form einer Diagonalmatrix möglich ist, weiter nach. Dieser Fall entspricht der Zerlegung von V in die direkte Summe von $L = N$ eindimensionalen Unterräumen, sodass die Blöcke in (20.15) alle die minimale Größe 1×1 haben. Was bedeutet diese Situation für die f-invarianten Unterräume U_j?

Da U_j eindimensional ist, gibt es einen Vektor $v_j \neq 0$, sodass $U_j = \langle v_j \rangle$ gilt. Da $f(U_j) \subseteq U_j$ gilt, muss $f(v_j)$ ein Vielfaches von v_j sein. Das führt uns zu folgender Definition:

Definition 20.19 (Eigenwert, Eigenvektor eines Endomorphismus bzw. einer Matrix)**.**

(i) Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K. Weiter sei $f: V \to V$ ein Endomorphismus. Ein Vektor $v \in V \setminus \{0\}$ heißt ein **Eigenvektor** (englisch: eigenvector) zum **Eigenwert $\lambda \in K$ des Endomorphismus f** (englisch: eigenvalue), wenn gilt:

$$f(v) = \lambda\, v. \tag{20.16}$$

In diesem Fall heißt $(\lambda, v) \in K \times (V \setminus \{0\})$ auch ein **Eigenpaar des Endomorphismus f** (englisch: eigenpair).

(ii) Es seien $(K, +, \cdot)$ ein Körper und $A \in K^{n \times n}$ eine Matrix mit $n \in \mathbb{N}_0$. Ein Vektor $x \in K^n \setminus \{0\}$ heißt ein **Eigenvektor** zum **Eigenwert $\lambda \in K$ der Matrix A**, wenn gilt:

$$A\, x = \lambda\, x. \tag{20.17}$$

In diesem Fall heißt $(\lambda, x) \in K \times (K^n \setminus \{0\})$ auch ein **Eigenpaar der Matrix A**. △

Beachte: Die Eigenpaare einer Matrix $A \in K^{n \times n}$ sind per Definition genau die Eigenpaare des von A induzierten Endomorphismus $f_A: K^n \to K^n$.

Beachte: Die Sprechweise „$\lambda \in K$ ist ein Eigenwert des Endomorphismus $f: V \to V$" bedeutet, dass es ein $v \in V \setminus \{0\}$ gibt, sodass $f(v) = \lambda\, v$ gilt. Analog bedeutet „$\lambda \in K$ ist ein Eigenwert der Matrix $A \in K^{n \times n}$, dass es ein $x \in K^n \setminus \{0\}$ gibt, sodass $A\, x = \lambda\, v$ gilt."

Lemma 20.20 (Eigenpaare von Endomorphismen und ihrer Darstellungsmatrizen)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f: V \to V$ ein Endomorphismus und $A = \mathcal{M}_{B_V}^{B_V}(f) \in K^{n \times n}$ seine Darstellungsmatrix bzgl. einer beliebigen Basis B_V von V. Dann sind äquivalent:

(i) $\lambda \in K$ ist ein Eigenwert von f.

(ii) $\lambda \in K$ ist ein Eigenwert von A.

Weiter gilt:

(iii) Ist $v \in V$ ein Eigenvektor zum Eigenwert λ von f, so ist $x = \Phi_{B_V}^{-1}(v)$ ein Eigenvektor zum Eigenwert λ von A.

(iv) Ist $x \in K^n$ ein Eigenvektor zum Eigenwert λ von A, so ist $v = \Phi_{B_V}(x)$ ein Eigenvektor zum Eigenwert λ von f.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii) und Aussage (iii): Es sei $(\lambda, v) \in K \times (V \setminus \{0\})$ ein Eigenpaar von f und $x = \Phi_{B_V}^{-1}(v)$, also auch $v = \Phi_{B_V}(x)$. Dann gilt

$$\begin{aligned}
A\, x &= f_A(x) && \text{nach Definition von } f_A \\
&= (\Phi_{B_V}^{-1} \circ f \circ \Phi_{B_V})(x) && \text{nach Satz 19.7} \\
&= \Phi_{B_V}^{-1}\big(f(\Phi_{B_V}(x))\big) && \text{nach Definition der Komposition} \\
&= \Phi_{B_V}^{-1}(f(v)) && \text{wegen } v = \Phi_{B_V}(x) \\
&= \Phi_{B_V}^{-1}(\lambda\, v) && \text{wegen } f(v) = \lambda\, v \\
&= \lambda\, \Phi_{B_V}^{-1}(v) && \text{wegen der Linearität von } \Phi_{B_V}^{-1} \\
&= \lambda\, x && \text{wegen } x = \Phi_{B_V}^{-1}(v).
\end{aligned}$$

Das zeigt: $(\lambda, x) \in K \times (K^n \setminus \{0\})$ ist ein Eigenpaar von A.

Aussage (ii) $\Rightarrow$ Aussage (i) und Aussage (iv): Es sei $(\lambda, x) \in K \times (K^n \setminus \{0\})$ ein Eigenpaar von A und $v = \Phi_{B_V}(x)$, also auch $x = \Phi_{B_V}^{-1}(v)$. Dann gilt

$$\begin{aligned}
f(v) &= (\Phi_{B_V} \circ f_A \circ \Phi_{B_V}^{-1})(v) && \text{nach Satz 19.7} \\
&= \Phi_{B_V}\big(f_A(\Phi_{B_V}^{-1}(v))\big) && \text{nach Definition der Komposition} \\
&= \Phi_{B_V}\big(f_A(x)\big) && \text{wegen } x = \Phi_{B_V}^{-1}(v) \\
&= \Phi_{B_V}(A\,x) && \text{nach Definition von } f_A \\
&= \Phi_{B_V}(\lambda\,x) && \text{wegen } A\,x = \lambda\,x \\
&= \lambda\,\Phi_{B_V}(x) && \text{wegen der Linearität von } \Phi_{B_V} \\
&= \lambda\,v && \text{wegen } v = \Phi_{B_V}(x).
\end{aligned}$$

Das zeigt: $(\lambda, v) \in K \times (V \setminus \{0\})$ ist ein Eigenpaar von f. □

Folgerung 20.21 (ähnliche Matrizen haben dieselben Eigenwerte)**.**
Es seien $(K, +, \cdot)$ ein Körper, $n \in \mathbb{N}_0$ und $A, \widehat{A} \in K^{n\times n}$. Wenn A und $\widehat{A}$ ähnlich sind, dann besitzen sie genau dieselben Eigenwerte.

Beweis. Nach Satz 20.14 können A und $\widehat{A}$ als Darstellungsmatrizen desselben Endomorphismus $f\colon V \to V$ in einem Vektorraum V mit $\dim(V) = n$ angesehen werden. Nach Lemma 20.20 sind daher die Eigenwerte von A und auch die Eigenwerte von $\widehat{A}$ identisch mit den Eigenwerten von f. □

Beachte: Die Umkehrung von Folgerung 20.21 gilt nicht! Zwei Matrizen $A, \widehat{A} \in K^{n\times n}$, die dieselben Eigenwerte besitzen, sind also nicht notwendigerweise ähnlich zueinander.

Beispiel 20.22 (nicht-ähnliche Matrizen mit gleichen Eigenwerten)**.**
Die Matrizen

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{und} \quad \widehat{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ in } \mathbb{R}^{2\times 2}$$

besitzen beide genau die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 1$, sie sind aber nicht ähnlich zueinander, denn der Ansatz $\widehat{A} = T\,A\,T^{-1}$ bzw. $\widehat{A}\,T = T\,A$ mit einer Matrix T führt auf

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

oder ausgeschrieben

$$\begin{bmatrix} t_{11} + t_{21} & t_{12} + t_{22} \\ t_{21} & t_{22} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix},$$

was notwendigerweise $t_{21} = t_{22} = 0$ ergibt. Damit hat T aber eine Nullzeile und ist nicht invertierbar. △

Wir kommen zurück zu der in Beispiel 20.18 beobachteten Situation, in der zu einem Endomorphismus (der Spiegelungsabbildung auf $\mathbb{R}^2$) eine diagonale Darstellungsmatrix existierte.

Definition 20.23 (diagonalisierbarer Endomorphismus, diagonalisierbare Matrix)**.**
Es seien $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}_0$.

(*i*) Ein Endomorphismus $f\colon V \to V$ eines Vektorraumes über K mit $\dim(V) = n$ heißt **diagonalisierbar** (englisch: diagonalizable), wenn es f-invariante Unterräume $U_1, \ldots, U_n$ der Dimension 1 gibt, sodass $V = U_1 \oplus \cdots \oplus U_n$ gilt.

(*ii*) Eine Matrix $A \in K^{n \times n}$ heißt **diagonalisierbar** (englisch: diagonalizable), wenn sie zu einer Diagonalmatrix ähnlich ist. △

Den Zusammenhang zwischen der Diagonalisierbarkeit eines Endomorphismus und seinen Darstellungsmatrizen stellt der folgende Satz her:

Satz 20.24 (Diagonalisierbarkeit (der Darstellungsmatrix) eines Endomorphismus)**.**
Es seien $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f\colon V \to V$ ein Endomorphismus und $A = \mathcal{M}_{B_V}^{B_V}(f) \in K^{n \times n}$ seine Darstellungsmatrix bzgl. einer beliebigen Basis B_V von V. Dann sind äquivalent:

(*i*) f ist diagonalisierbar.

(*ii*) V besitzt eine Basis, die nur aus Eigenvektoren von f besteht.

(*iii*) A ist diagonalisierbar.

(*iv*) K^n besitzt eine Basis, die nur aus Eigenvektoren von A besteht.

Beweis. Aussage (*i*) $\Rightarrow$ Aussage (*iii*): Es sei f diagonalisierbar, also gibt es eine Zerlegung $V = U_1 \oplus \cdots \oplus U_n$ mit f-invarianten Unterräumen. Nach Satz 20.17 gibt es also eine Basis $\widehat{B}_V$, sodass die Darstellungsmatrix $\widehat{A} = \mathcal{M}_{\widehat{B}_V}^{\widehat{B}_V}(f)$ eine Diagonalmatrix ist. Da A und $\widehat{A}$ denselben Endomorphismus f darstellen, ist A nach Satz 20.17 ähnlich zur Diagonalmatrix $\widehat{A}$, d. h., A ist diagonalisierbar.

Aussage (*iii*) $\Rightarrow$ Aussage (*iv*): Es sei A diagonalisierbar, also existiert eine invertierbare Matrix $T \in K^{n \times n}$, sodass $\widehat{A} = T\,A\,T^{-1}$ eine Diagonalmatrix ist. Es gilt also

$$\widehat{A} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Wir setzen $X \coloneqq T^{-1}$ und schreiben $A = T^{-1}\,\widehat{A}\,T = X\,\widehat{A}\,X^{-1}$. Wir zeigen, dass die Spalten $x_{\bullet 1}, \ldots, x_{\bullet n}$ von X Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$ von A sind, denn es gilt:

$$\begin{aligned} A\,x_{\bullet j} &= X\,\widehat{A}\,X^{-1} x_{\bullet j} \\ &= X\,\widehat{A}\,e_j && \text{mit dem Standardbasisvektor } e_j \in K^n \\ &= X\,\lambda_j\,e_j && \text{wegen der Diagonalstruktur von } \widehat{A} \\ &= \lambda_j\,X e_j && \text{aufgrund des Assoziativgesetzes (15.12)} \\ &= \lambda_j\,x_{\bullet j} && \text{nach Definition der Matrix-Vektor-Multiplikation.} \end{aligned}$$

Da T invertierbar ist, ist auch $X = T^{-1}$ invertierbar. Die Spalten von X sind also linear unabhängig (Satz 15.40). Das heißt, K^n besitzt die Basis $x_{\bullet 1}, \ldots, x_{\bullet n}$ aus Eigenvektoren von A. (Die Eigenwerte sind gerade die Diagonaleinträge von $\widehat{A}$.)

Aussage (*iv*) $\Rightarrow$ Aussage (*ii*): Es sei $(x_1, \ldots, x_n)$ eine Basis von K^n aus Eigenvektoren der Matrix A zu den Eigenwerten $\lambda_1, \ldots, \lambda_n \in K$. Nach Lemma 20.20 ist dann $v_j = \Phi_{B_V}(x_j)$ ein Eigenwert des

Endomorphismus f zum Eigenwert λ_j für jedes $j = 1, \dots, n$. Da $\Phi_{B_V} : K^n \to V$ ein Isomorphismus ist, bilden die Vektoren $(v_1, \dots, v_n)$ eine Basis von V (Satz 17.7).

Aussage *(ii)* $\Rightarrow$ Aussage *(i)*: Es sei $B_V = (v_1, \dots, v_n)$ eine Basis von V aus lauter Eigenvektoren von f, sagen wir $f(v_j) = \lambda_j v_j$ für den zugehörigen Eigenwert $\lambda_j \in K$. Wir definieren die eindimensionalen Unterräume $U_j \coloneqq \langle v_j \rangle$, $j = 1, \dots, n$. Dann ist jedes U_j f-invariant, denn es gilt

$$\begin{aligned}
f(U_j) &= f(\langle v_j \rangle) && \text{nach Definition von } U_j \\
&= \langle f(v_j) \rangle && \text{nach Lemma 17.5} \\
&= \langle \lambda_j v_j \rangle && \text{wegen } f(v_j) = \lambda_j v_j \\
&= \lambda_j \langle v_j \rangle && \text{wegen Satz 12.13 (lineare Hülle besteht aus Linearkombinationen)} \\
&= \lambda_j U_j && \text{nach Definition von } U_j \\
&\subseteq U_j.
\end{aligned}$$

(**Quizfrage 20.6:** Warum steht hier $\subseteq$ und nicht $=$?) Weiter gilt nach Satz 14.16

$$V = \langle v_1 \rangle \oplus \cdots \langle v_n \rangle = U_1 \oplus \cdots \oplus U_n,$$

d. h., f ist diagonalisierbar. □

Wie der Beweis von Satz 20.24 zeigt, sind es genau die diagonalisierbaren Endomorphismen, die bei geeigneter Wahl der Basis (nämlich einer Basis aus lauter Eigenvektoren) die denkbar einfachste Form einer Darstellungsmatrix besitzen, nämlich eine Diagonalmatrix. Bei dieser sind die Diagonaleinträge gerade die Eigenwerte (die möglicherweise mehrfach auftreten). Zwei *diagonalisierbare* $n \times n$-Matrizen sind also genau dann ähnlich, wenn sie zur selben Diagonalmatrix ähnlich sind.

Wie bereits erwähnt, werden wir

- den allgemeineren Fall von Darstellungsmatrizen in Blockdiagonalgestalt, der eintritt, wenn V als direkte Summe f-invarianter Unterräume darstellbar ist (Satz 20.17),
- sowie den ganz allgemeinen Fall für beliebige Endomorphismen

erst im zweiten Teil der Vorlesung untersuchen können. Wir nehmen jedoch vorweg, dass auch dabei Eigenwerte und Eigenvektoren von f eine wesentliche Rolle spielen.

Ende der Vorlesung 29

Ende der Woche 14

§ 21 Dualräume und adjungierte Homomorphismen

Literatur: Beutelspacher, 2014, Kapitel 5.4, Bosch, 2014, Kapitel 5.1, Fischer, Springborn, 2020, Kapitel 7.1

§ 21.1 Der Dualraum eines Vektorraumes

Definition 21.1 (Dualraum, Linearform)**.**
Es sei V ein Vektorraum über dem Körper K. Der Vektorraum

$$V^* := \operatorname{Hom}(V, K) \tag{21.1}$$

der linearen Abbildungen $V \to K$ heißt der **Dualraum** (englisch: dual space) von V. Die Elemente von V^* heißen auch **lineare Funktionale** (englisch: linear functionals) oder **Linearformen** (englisch: linear forms) auf V oder **Covektoren** (englisch: covectors). △

Linearformen sind also lineare Abbildungen eines Vektorraumes V in den zugrunde liegenden Skalarkörper K, aufgefasst als 1-dimensionaler Vektorraum über sich selbst. Nach Satz 17.11 ist V^* tatsächlich ein Vektorraum. Der Nullvektor in V^* ist die Nullabbildung, die wir auch als die **Nullform** (englisch: zero form) auf V bezeichnen können.

Beispiel 21.2 (Dualraum, Linearform)**.**

(i) Die Projektion auf die i-te Koordinate (Beispiel 17.3)

$$\pi_i \colon K^n \ni x \mapsto x_i \in K$$

ist eine Linearform auf dem Standardvektorraum K^n über dem Körper K.

(ii) Eine Auswertungsabbildung an irgendeinem Punkt $\alpha \in K$, definiert beispielsweise auf dem Vektorraum $K[t]$ der Polynome über einem Körper K, also

$$K[t] \ni p \mapsto \widetilde{p}(\alpha) \in K,$$

ist eine Linearform auf $K[t]$.

(iii) Die Grenzwert-Abbildung, definiert beispielsweise auf dem Vektorraum der konvergenten $\mathbb{R}$-wertigen Folgen $(\mathbb{R}^{\mathbb{N}})_c$ (siehe Beispiel 17.3) nach $\mathbb{R}$, also

$$\lim \colon (\mathbb{R}^{\mathbb{N}})_c \ni (y_i)_{i \in \mathbb{N}_0} \mapsto \lim_{i \to \infty} (y_i) \in \mathbb{R},$$

ist eine Linearform auf $(\mathbb{R}^{\mathbb{N}})_c$.

(iv) Ist $A \in K^{1 \times n}$ eine Matrix, dann ist $f_A \colon K^n \to K$ eine Linearform auf dem Standardvektorraum K^n über dem Körper K.

(v) Ist $B = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$ eine Basis des endlich-dimensionalen Vektorraumes V, so hat eine Linearform $v^* \in V^*$ die Darstellungsmatrix

$$\mathcal{M}^{B_V}_{(1)}(v^*) = \begin{bmatrix} v^*(v_1) & \cdots & v^*(v_n) \end{bmatrix} \in K^{1 \times n}.$$

△

Ab sofort werden wir typische Linearformen auf V mit v^* bezeichnen.

Bemerkung 21.3 (zur Sprechweise)**.**

(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. △

Bemerkung 21.4 (zur Bedeutung von Linearformen)**.**

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 möglichen Einheiten widerspiegelt wie $\mathrm{m\,s^{-1}}$ oder $\mathrm{km\,h^{-1}}$.

Wir können eine Linearform oder einen Covektor interpretieren als eine Art „Maßstab" für Vektoren. Eine Linearform auf dem Vektorraum V trägt gerade die reziproke Dimension der Vektoren in V. Wenn beispielsweise die Elemente von V Geschwindigkeiten repräsentieren mit der Dimension „Länge geteilt durch Zeit", dann haben die Elemente von V^* die Dimension „Zeit geteilt durch Länge". Demzufolge ist das Ergebnis der Auswertung $v^*(v)$ ein dimensionsloser Skalar. △

Bemerkung 21.5 (auch Vektoren „sind" lineare Abbildungen)**.**

Covektoren (Linearformen) sind genauso natürliche Objekte wie Vektoren. Auch Vektoren können als lineare Abbildungen aufgefasst werden. Jedem Vektor $v \in V$ entspricht dabei eine Abbildung

$$K \ni \alpha \mapsto \alpha\, v \in V,$$

also ein Element von $\operatorname{Hom}(K, V)$, genauer: Durch diese Zuordnung ist ein Isomorphismus zwischen V und $\operatorname{Hom}(K, V)$ gegeben. Wir können also Vektoren in V mit Elementen in $\operatorname{Hom}(K, V)$ identifizieren, während Covektoren Elemente von $\operatorname{Hom}(V, K)$ sind. △

Definition 21.6 (kanonische duale Paarung)**.**

Es sei V ein Vektorraum über dem Körper K und V^* sein Dualraum. Die Abbildung

$$\langle \cdot\,, \cdot \rangle \colon V^* \times V \ni (v^*, v) \mapsto \langle v^*\,, v \rangle := v^*(v) \in K \tag{21.2}$$

heißt die **(kanonische) duale Paarung** (englisch: (canonical) dual pairing) der Räume V^* und V.[37] △

Wenn wir die Räume betonen wollen, 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 21.7 (die duale Paarung ist linear in beiden Argumenten)**.**

Es sei V ein Vektorraum über dem Körper K und V^* sein Dualraum. Die duale Paarung (21.2) ist linear in beiden Argumenten, d. h., für $v^*, w^* \in V^*$ und $v, w \in V$ sowie $\alpha, \beta \in K$ gilt

$$\langle \alpha\, v^* + \beta\, w^*\,, v \rangle = \alpha\, \langle v^*\,, v \rangle + \beta\, \langle w^*\,, v \rangle \tag{21.3a}$$

$$\langle v^*\,, \alpha\, v + \beta\, w \rangle = \alpha\, \langle v^*\,, v \rangle + \beta\, \langle v^*\,, w \rangle. \tag{21.3b}$$

[37] In der Physik ist dafür auch die **Bra-Ket-Notation** (englisch: bra-ket notation) üblich. Dort schreibt man $\langle v^* \,|\, v \rangle$ statt $\langle v^*\,, v \rangle$.

Beweis. Es gilt

$$\begin{aligned}\langle \alpha\, v^* + \beta\, w^*, v\rangle &= (\alpha\, v^* + \beta\, w^*)(v) && \text{nach Definition von } \langle\cdot\,,\cdot\rangle \\ &= \alpha\, v^*(v) + \beta\, w^*(v) && \text{nach Definition der Addition und S-Multiplikation} \\ & && \text{von Homomorphismen (Satz 17.11)} \\ &= \alpha\,\langle v^*, v\rangle + \beta\,\langle w^*, v\rangle && \text{nach Definition von } \langle\cdot\,,\cdot\rangle.\end{aligned}$$

Die Gleichung (21.3b) folgt sofort aus der Linearität von $v^* \in V^* = \operatorname{Hom}(V, K)$. ☐

Ist also $B = (v_i)_{i\in I}$ eine Basis von V und ist $(\alpha_i)_{i\in I}$ eine Familie in K, dann gibt es genau eine lineare Abbildung $v^* \in V^*$ mit der Eigenschaft $v^*(v_i) = \langle v^*, v_i\rangle = \alpha_i$ für alle $i \in I$.

Das führt auf folgende Variante von Satz 19.5:

Satz 21.8 (die Zuordnung zur Familie der Bilder ist ein Vektorraumisomorphismus)**.**
Es seien K ein Körper und V ein Vektorraum über K. Weiter sei $B = (v_i)_{i\in I}$ eine Basis von V. Die Zuordnung

$$I_B\colon V^* \ni v^* \mapsto \langle v^*, v_i\rangle_{i\in I} \in K^I \tag{21.4}$$

einer Linearform zur Familie der Bilder ist ein Isomorphismus von Vektorräumen.

Beachte: V^* ist folglich isomorph zu K^I.

Beweis. Der Beweis folgt dem Aufbau von Satz 19.5.

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 (21.3a)

$$\langle \alpha\, v^* + \beta\, w^*, v_i\rangle = \alpha\,\langle v^*, v_i\rangle + \beta\,\langle w^*, v_i\rangle,$$

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 $v^*\colon V \to K$ nur die Nullabbildung sein, also der Nullvektor in V^*. Daher gilt $\operatorname{Kern}(I_B) = \{0\}$, und nach Lemma 17.6 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 17.7 gibt es (genau) einen Homomorphismus $v^*\colon V \to K$, der $\langle v^*, v_i\rangle = \alpha_i$ als Bilder hat, also $I_B(v^*) = (\alpha_i)_{i\in I}$. ☐

Folgerung 21.9 (Dimension des Dualraumes)**.**
Es seien K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Dann gilt $V \cong V^*$ und damit $\dim(V) = \dim(V^*)$.

Beweis. Es sei $B = (v_1, \dots, v_n)$ eine Basis von V. Aus Satz 21.8 folgt, dass

$$I_B \colon V^* \ni v^* \mapsto \big(\langle v^*, v_1\rangle, \dots, \langle v^*, v_n\rangle\big) \in K^{[\![1,n]\!]} \cong K_n$$

ein Isomorphismus ist.[38] Andererseits gilt nach Folgerung 18.2 $K_n \cong V$, also auch $V^* \cong V$, und damit $\dim(V) = \dim(V^*) = n$. □

Beachte: Der hier konstruierte Isomorphismus zwischen V und V^* hängt von der gewählten Basis B in V ab. Es gibt insbesondere keinen kanonischen (natürlichen) Isomorphismus zwischen V und V^*.

Quizfrage 21.1: Warum funktioniert der Beweis der Isomorphie von V und V^* in Folgerung 21.9 nicht auch für unendlich-dimensionale Vektorräume?

Der folgende Satz beantwortet die Frage, wie wir aus einer Basis eines endlich-dimensionalen Vektorraumes V eine Basis von V^* erhalten können.

Satz 21.10 (Basis des Dualraumes).
Es seien K ein Körper und V ein endlich-dimensionaler Vektorraum über K mit Basis $B = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$. Dann bilden die Linearformen $v_i^* \in V^*$ für $i = 1, \dots, n$, definiert durch die Bilder auf der Basis $(v_1, \dots, v_n)$

$$\langle v_i^*, v_j\rangle := \delta_{ij} = \begin{cases} 1 & \text{im Fall } i = j, \\ 0 & \text{im Fall } i \neq j, \end{cases} \tag{21.5}$$

eine Basis von V^*. Für die eindeutige Darstellung von $v^* \in V^*$ als Linearkombination von B^* gilt:

$$v^* = \sum_{i=1}^{n} v^*(v_i)\, v_i^* = \sum_{i=1}^{n} \underbrace{\langle v^*, v_i\rangle}_{\text{Koeffizient}} v_i^*. \tag{21.6}$$

Beweis. Zunächst stellen wir fest, dass die Linearformen $v_1^*, \dots, v_n^*$ eindeutig definiert sind, da wir für jedes v_i^* die Bilder der Basisvektoren $v_1, \dots, v_n$ angegeben haben, die nach Satz 17.7 die lineare Abbildung $v_i^* \in \operatorname{Hom}(V, K) = V^*$ eindeutig festlegen.

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 (21.3a)

$$0 = \langle 0_{V^*}, v_j\rangle = \Big\langle \sum_{i=1}^{n} \alpha_i\, v_i^*, v_j \Big\rangle = \sum_{i=1}^{n} \alpha_i \underbrace{\langle v_i^*, v_j\rangle}_{=\delta_{ij}} = \alpha_j,$$

also sind alle Koeffizienten notwendigerweise gleich Null.

Schritt 2: Die Vektoren $v_1^*, \dots, v_n^*$ erzeugen den ganzen Raum V^*:

Es sei dazu $v^* \in V^*$ beliebig. Wir müssen zeigen, dass v^* als Linearkombination, die dann aufgrund von **Schritt 1** eindeutig ist, von $v_1^*, \dots v_n^*$ darstellbar ist. Dazu überprüfen wir die Darstellung (21.6), also

$$v^* = \sum_{i=1}^{n} \langle v^*, v_i\rangle\, v_i^*.$$

[38] Dass die Menge der Abbildungen $[\![1, n]\!] \to K$ in natürlicher Weise bijektiv auf die Menge der Tupel $(k_1, \dots, k_n) \in K_n$ abgebildet werden kann, hatten wir in Bemerkung 6.33 schon verwenden. Offenbar ist diese Zuordnung auch linear.

Beide Seiten der Gleichung sind Elemente von V^*. Wir zeigen, dass sie auf den Bildern der Basisvektoren $v_1, \dots, v_n$ von V übereinstimmen. In der Tat gilt

$$\Big\langle \sum_{i=1}^n \langle v^*, v_i\rangle \, v_i^*, v_j \Big\rangle = \sum_{i=1}^n \langle v^*, v_i\rangle \, \langle v_i^*, v_j\rangle = \langle v^*, v_j\rangle. \qquad \square$$

Definition 21.11 (duale Basis).
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$. Die durch (21.5) definierte Basis $B^* \coloneqq (v_1^*, \dots, v_n^*)$ von V^* heißt die zur Basis B **duale Basis** (englisch: dual basis). △

Beachte: 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 (21.5) erfüllen müssen.

Beispiel 21.12 (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 21.2), 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 $i, j = 1, \dots, n$. Das heißt also, wir erhalten die i-te Koordinate eines Vektors $x \in K^n$, indem wir das i-te Funktional der zur Standardbasis dualen Basis auf x anwenden.

(*ii*) Allgemeiner seien K ein Körper, V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$ und $B = (v_1, \dots, v_n)$ eine Basis von V. Ist $B^* = (v_1^*, \dots, v_n^*)$ die zu B duale Basis und $x = \Phi_B^{-1}(v)$ der Koordinatenvektor (Satz 19.1) von $v \in V$, dann gilt

$$x_i = \langle v_i^*, v\rangle \quad \text{für } i = 1, \dots, n \tag{21.7}$$

und damit

$$x = \Phi_B^{-1}(v) = \begin{pmatrix} \langle v_1^*, v\rangle \\ \vdots \\ \langle v_n^*, v\rangle \end{pmatrix}. \tag{21.8}$$

Die Linearform v_i^* aus der dualen Basis B^* ermittelt also bei Anwendung auf den Vektor $v \in V$ dessen i-te Koordinate bzgl. der Basis B von V. Wir nennen das Element v_i^* der zu B dualen Basis B^* auch die ***i*-te Koordinatenform** (englisch: *i*-th coordinate form) der Basis B. Es gilt also $v_i^* = \pi_i \circ \Phi_B^{-1}$. △

Bemerkung 21.13 (Rolle der dualen Basis bei Darstellungsmatrizen).
Es seien V und W Vektorräume über demselben Körper K mit Basen $B_V = (v_1, \dots, v_m)$ bzw. $B_W = (w_1, \dots, w_n)$. Die Darstellungsmatrix $A = \mathcal{M}_{B_W}^{B_V}(f)$ eines Homomorphismus $f \colon V \to W$ hatten wir so eingeführt, dass sie spaltenweise die Koordinaten der Bilder der Basisvektoren v_j 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 aber keine Möglichkeit, einzelne Einträge der Darstellungsmatrix A anzugeben. Dies gelingt nun mit (21.7). Ist nämlich $B_W^* = (w_1^*, \dots, w_n^*)$ die zu B_W duale Basis, dann gilt

$$a_{ij} = \langle w_i^*, f(v_j)\rangle. \tag{21.9}$$

△

Bemerkung 21.14 (Dualräume unendlich-dimensionaler Vektorräume).
In dem Fall, dass V unendlich-dimensional ist, erhalten wir mit der Konstruktion (21.5) keine Basis des Dualraumes V^*. Genauer: Es sei $B = (v_i)_{i \in I}$ eine Basis von V mit unendlicher Indexmenge I. Wir definieren wie in (21.5) eine gleichmächtige Familie $B^* := (v_i^*)_{i \in I}$ an Elementen von V^* durch

$$\langle v_i^*, v_j \rangle := \delta_{ij} = \begin{cases} 1 & \text{im Fall } i = j, \\ 0 & \text{im Fall } i \neq j. \end{cases}$$

Dann ist B^* zwar eine linear unabhängige Familie in V^*, aber kein Erzeugendensystem, 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. △

Ende der Vorlesung 1

§ 21.2 Darstellung von Linearformen

In diesem Abschnitt ist V endlich-dimensional mit $\dim(V) = n \in \mathbb{N}_0$. So wie wir Vektoren $v \in V$ durch ihre Koordinatenvektoren $x \in K^n$ bzgl. einer gewählten Basis $B_V = (v_1, \ldots, v_n)$ darstellen, können wir auch Linearformen bzgl. einer Basis im Dualraum darstellen. Wie sich gleich zeigen wird, bietet es sich dazu an, die duale Basis $B_V^* = (v_1^*, \ldots, v_n^*)$ zu verwenden. Sind nämlich

$$v = \sum_{j=1}^n x_j \, v_j \quad \text{und} \quad v^* = \sum_{i=1}^n \xi_i \, v_i^*$$

die Darstellungen von $v \in V$ und $v^* \in V^*$, dann gilt

$$\langle v^*, v \rangle = \Big\langle \sum_{i=1}^n \xi_i \, v_i^*, \sum_{j=1}^n x_j \, v_j \Big\rangle = \sum_{i=1}^n \sum_{j=1}^n \xi_i \, x_j \, \underbrace{\langle v_i^*, v_j \rangle}_{\delta_{ij}} = \sum_{j=1}^n \xi_j \, x_j.$$

Wir notieren die Koordinatenvektoren von v bzw. von v^* als

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n \quad \text{bzw.} \quad \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n.$$

Mit dieser Konvention gilt

$$\langle v^*, v \rangle = \sum_{j=1}^n \xi_j \, x_j = \xi^\intercal x. \tag{21.10}$$

Mit anderen Worten, wir wenden $v^* \in V^*$ auf einen Vektor $v \in V$ an, indem wir das Matrix-Matrix-Produkt $\xi^\intercal x$ zwischen den Koordinatenvektoren $\xi^\intercal \in K_n \cong K^{1 \times n}$ und $x \in K^n \cong K^{n \times 1}$ bestimmen.

Die Einträge des Koordinatenvektoren eines Vektors bzw. einer Linearform erhalten wir übrigens, indem wir die Elemente der jeweils anderen Basis der Reihe nach einsetzen bzw. anwenden:

Lemma 21.15 (Koordinatenvektoren von Vektoren und Linearformen).
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B_V = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$. Weiter sei $B_V^* = (v_1^*, \ldots, v_n^*)$ die zu B_V duale Basis von V^*.

(i) Für den Koordinatenvektor $x = \Phi_{B_V}^{-1}$ eines Vektors $v \in V$ gilt

$$x_i = \langle v_i^*, v \rangle, \quad i = 1, \ldots, n. \tag{21.11}$$

(ii) Für den Koordinatenvektor $\xi = \Phi_{B_V^*}^{-1}$ einer Linearform $v^* \in V^*$ gilt

$$\xi_i = \langle v^*, v_i \rangle, \quad i = 1, \ldots, n. \tag{21.12}$$

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-1.3. □

Beispiel 21.16 (Darstellung von Linearformen).

(i) Im Vektorraum $K_n[t]$ der Polynome vom Höchstgrad $n \in \mathbb{N}_0$ über dem Körper K betrachten wir die Monombasis $B_V = (v_0, v_1, \ldots, v_n) = (1, t, \ldots, t^n)$. Die duale Basis $B_V^* = (v_0^*, v_1^*, \ldots, v_n^*)$ ist dann gegeben durch die Ableitungsoperatoren (vgl. Beispiel 17.8)

$$v_i^* \coloneqq \frac{1}{i!} \frac{\mathrm{d}}{\mathrm{d}t^i}(0)$$

mit Auswertung an der Stelle $t = 0$, genauer:

$$\langle v_i^*, v \rangle \coloneqq \frac{1}{i!} \widetilde{\left(\frac{\mathrm{d}v}{\mathrm{d}t^i} \right)}(0)$$

Wählen wir im Fall $n = 2$ und $K = \mathbb{Q}$ beispielsweise das Polynom $v = 2\,t^2 - 3\,t + 1$, dann hat dieses den Koordinatenvektor $x = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$. Wir betrachten außerdem als Beispiel die durch $\langle v^*, v \rangle = \widetilde{v}(0) - \widetilde{v}(2)$ gegebene Linearform, die also die Differenz der Werte der durch v gegebene Polynomfunktion $\widetilde{v}$ an der Stelle $t = 0$ und $t = 2$ verwendet.

Um die Koordinaten dieser Linearform in der Basis B_V^* zu erhalten, setzen wir gemäß (21.12) der Reihe nach die Elemente der Basis B_V ein. Wir erhalten

$$\begin{aligned} \xi_0 &= \langle v^*, 1 \rangle &&= 1 - 1 = 0 \\ \xi_1 &= \langle v^*, t \rangle &&= 0 - 2 = -2 \\ \xi_2 &= \langle v^*, t^2 \rangle &&= 0 - 4 = -4. \end{aligned}$$

Die gewählte Linearform hat also bzgl. der Basis B_V^* den Koordinatenvektor $\xi = \begin{pmatrix} 0 \\ -2 \\ -4 \end{pmatrix}$.

Die Anwendung von v^* auf v ergibt

$$\langle v^*, v \rangle = \xi^\mathsf{T} x = \begin{pmatrix} 0 \\ -2 \\ -4 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} = -2.$$

(ii) Im Vektorraum K^n stimmen Vektoren mit ihren Koordinatenvektoren x überein, wenn wir die Standardbasis wählen. Eine Linearform in $(K^n)^*$ können wir durch ihren Koordinatenvektor ξ bzgl. der dualen Basis darstellen. Da diese Zuordnung ein Isomorphismus ist (Satz 19.1), können wir Linearformen auf K^n auch direkt mit K^n selbst identifizieren. Die duale Paarung ist dann $\xi^\mathsf{T} x$, siehe (21.10). △

Das folgende Lemma klärt, wie sich die duale Basis bei einem Basiswechsel verhält.

Lemma 21.17 (Zusammenhang der Basiswechsel zwischen Basis und dualer Basis)**.**
Es seien K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Es seien $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ zwei Basen von V. Weiter seien $B_{V^*} = (v_1^*, \ldots, v_n^*)$ und $\widehat{B}_{V^*} = (\widehat{v}_1^*, \ldots, \widehat{v}_n^*)$ die jeweiligen dualen Basen. Ist

$$T := \mathcal{T}_{\widehat{B}_V}^{B_V} = \mathcal{M}_{\widehat{B}_V}^{B_V}(\mathrm{id}_V)$$

die Transformationsmatrix des Basiswechsels von B_V nach $\widehat{B}_V$ (Definition 20.1), dann ist

$$T^{-\mathsf{T}} = \mathcal{T}_{\widehat{B}_{V^*}}^{B_{V^*}} = \mathcal{M}_{\widehat{B}_{V^*}}^{B_{V^*}}(\mathrm{id}_{V^*}) \tag{21.13}$$

die Transformationsmatrix des Basiswechsels von B_{V^*} nach $\widehat{B}_{V^*}$. Ist $\xi \in K^n$ der Koordinatenvektor der Linearform $v^* \in V^*$ bzgl. der Basis B_{V^*}, dann ist $\widehat{\xi} = T^{-\mathsf{T}}\xi$ der Koordinatenvektor von v^* bzgl. der Basis $\widehat{B}_{V^*}$.[39]

Beweis. Mit der Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V}^{B_V}$ des Basiswechsels von B_V nach $\widehat{B}_V$ gilt

$$v_j = \sum_{i=1}^n t_{ij}\, \widehat{v}_i,$$

siehe (20.2). Die gesuchte Transformationsmatrix $S = \mathcal{T}_{\widehat{B}_{V^*}}^{B_{V^*}}$ des Basiswechsels von B_{V^*} nach $\widehat{B}_{V^*}$ erfüllt analog

$$v_j^* = \sum_{i=1}^n s_{ij}\, \widehat{v}_i^*.$$

Wir betrachten nun

$$\begin{aligned}
\langle v_i^*, v_j\rangle &= \Big\langle \sum_{k=1}^n s_{ki}\, \widehat{v}_k^*, \sum_{\ell=1}^n t_{\ell j}\, \widehat{v}_\ell \Big\rangle \\
&= \sum_{k=1}^n \sum_{\ell=1}^n s_{ki}\, t_{\ell j} \langle \widehat{v}_k^*, \widehat{v}_\ell\rangle && \text{wegen der Linearität der dualen Paarung in beiden Argumenten} \\
&= \sum_{k=1}^n s_{ki}\, t_{kj} && \text{da } \widehat{B}_{V^*} \text{ die zu } \widehat{B}_V \text{ duale Basis ist und damit } \langle \widehat{v}_k^*, \widehat{v}_\ell\rangle = \delta_{k\ell} \text{ gilt} \\
&= (S^\mathsf{T} T)_{ij} && \text{nach der Definition (15.9) der Matrix-Matrix-Multiplikation.}
\end{aligned}$$

Andererseits ist B_{V^*} die zu B_V duale Basis, also muss $\langle v_i^*, v_j\rangle = \delta_{ij}$ gelten. Das heißt, es gilt $S^\mathsf{T} T = I$, also $S = T^{-\mathsf{T}}$.

Wie in Lemma 20.3 transfomiert sich der Koordinatenvektor ξ wie folgt:

$$\widehat{\xi} = \mathcal{T}_{\widehat{B}_{V^*}}^{B_{V^*}}\, \xi = T^{-\mathsf{T}}\xi.$$ □

Diese Transformationsvorschrift können wir uns auch aus (21.10) herleiten, denn diese Gleichung ist ja unabhängig davon, welche Basis und zugeörige duale Basis wir wählen. Wir wissen also einerseits

$$\langle v^*, v\rangle = \xi^\mathsf{T} x = \widehat{\xi}^\mathsf{T} \widehat{x}$$

und andererseits $\widehat{x} = T\, x$, also muss $\widehat{\xi} = T^{-\mathsf{T}}\xi$ gelten.

[39] Vergleiche mit der Transformation von Koordinatenvektoren $\widehat{x} = \mathcal{T}_{\widehat{B}_V}^{B_V}\, x = T\, x$ von Vektoren in V, siehe Lemma 20.3.

Bemerkung 21.18 (kovariante und kontravariante Transformationen).
Wir haben gesehen, dass beim Übergang von B_V zu $\widehat{B}_V$ mittels $v_j = \sum_{i=1}^n t_{ij}\, \widehat{v}_i$, also in der Form von alt nach neu (Lemma 20.3)

$$\widehat{v}_j = \sum_{i=1}^n (T^{-1})_{ij}\, v_i,$$

sich die *Koordinatenvektoren* eines Vektors $v \in V$ mit der zu T^{-1} inversen Matrix transformieren, nämlich $\widehat{x} = T\, x$ bzw. in den einzelnen Koordinaten

$$\widehat{x}_i = \sum_{j=1}^n T_{ij}\, x_j.$$

Eine historische, aber immer noch verbreitete Sprechweise für diese Beobachtung ist, dass sich die Koordinaten eines Vektors in V **kontravariant** (englisch: contravariant, lateinisch: contra- (Vorsilbe): gegen, lateinisch: variare: verändern, verwandeln) **transformieren** im Vergleich mit der Transformation der Basisvektoren von V.[40]

Demgegenüber transformieren sich die Koordinaten einer Linearform $v^* \in V^*$ wieder mit T^{-1}, nämlich

$$\widehat{\xi}_j = \sum_{i=1}^n (T^{-1})_{ij}\, \xi_i.$$

Diese Transformationsweise bezeichnet man als **kovariant** (englisch: covariant, lateinisch: co- (Vorsilbe): mit) im Vergleich mit der Transformation der Basisvektoren von V. △

§ 21.3 Annihilatoren

Definition 21.19 ((Prä-)Annihilator).
Es seien K ein Körper und V ein Vektorraum über K.

(*i*) Ist $M \subseteq V$, dann heißt

$$\begin{aligned} M^0 &:= \{v^* \in V^* \mid \langle v^*, v\rangle = 0 \text{ für alle } v \in M\} \\ &= \{v^* \in V^* \mid M \subseteq \operatorname{Kern}(v^*)\} \\ &= \bigcap_{v \in M} \{v\}^0 \subseteq V^* \end{aligned} \tag{21.14}$$

der **Annihilator** von M (englisch: annihilator).[41]

(*ii*) Ist $F \subseteq V^*$, dann heißt

$$\begin{aligned} {}^0F &:= \{v \in V \mid \langle v^*, v\rangle = 0 \text{ für alle } v^* \in F\} \\ &= \bigcap_{v^* \in F} \operatorname{Kern}(v^*) \subseteq V \end{aligned} \tag{21.15}$$

der **Prä-Annihilator** von F (englisch: pre-annihilator).[42] △

[40] 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.
[41] Der Annihilator einer Menge $M \subseteq V$ besteht also aus denjenigen Linearformen auf V, die auf der Menge M verschwinden, also dort den Funktionswert 0 haben.
[42] Der Prä-Annihilator einer Menge $F \subseteq V^*$ besteht also aus denjenigen Vektoren in V, auf denen alle $f \in F$ verschwinden.

Wir geben zunächst die (Prä-)Annihilatoren bestimmter Teilmengen von V und V^* an, nämlich die zu den trivialen Unterräumen.

Lemma 21.20 ((Prä-)Annihilatoren trivialer Unterräume)**.**
Es seien K ein Körper und V ein Vektorraum über K. Dann gilt:

(i) $\{0_V\}^0 = V^*$.

(ii) $V^0 = \{0_{V^*}\}$.

(iii) ${}^0\{0_{V^*}\} = V$.

(iv) ${}^0(V^*) = \{0_V\}$.[43]

Beweis. Aussage (i): Jede Linearform $v^* \in V^*$ bildet 0_V auf $0 \in K$ ab (Lemma 17.5). Das bedeutet

$$\{0_V\}^0 = \{v^* \in V^* \mid \langle v^*, 0_V \rangle = 0\} = V^*.$$

Aussage (ii): 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 $V^0 = \{0_{V^*}\}$.

Aussage (iii): Nach Definition ist

$${}^0\{0_{V^*}\} = \{v \in V \mid \langle v^*, v \rangle = 0 \text{ für } v^* = 0_{V^*}\} = V.$$

Aussage (iv): Nach Definition ist

$${}^0(V^*) = \{v \in V \mid \langle v^*, v \rangle = 0 \text{ für alle } v^* \in V^*\}.$$

Nach Satz 17.7 Aussage (i) existiert für jedes $v \in V \setminus \{0_V\}$ eine Linearform $v^* \in V^*$ mit der Eigenschaft $v^*(v) = 1$. (**Quizfrage 21.2:** Klar?) Das bedeutet, dass jedes $v \in V \setminus \{0_V\}$ 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 der trivialen Unterräume aus Lemma 21.20 sind selbst wieder Unterräume. Das gilt auch allgemein für die Annihilatoren beliebiger Mengen:

Lemma 21.21 ((Prä-)Annihilatoren sind Unterräume)**.**
Es seien K ein Körper und V ein Vektorraum über K. Dann gilt:

(i) Ist $M \subseteq V$, dann ist M^0 ein Unterraum von V^*.

(ii) Ist $F \subseteq V^*$, dann ist ${}^0F$ ein Unterraum von V.

[43] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Beweis. Aussage (i): Wir benutzen das Unterraumkriterium (Satz 12.8). 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.

Aussage (ii): Der Beweis geht analog. □

Beispiel 21.22 ((Prä-)Annihilatoren sind Unterräume).

(i) Wir betrachten den Vektorraum $V := \mathbb{R}^{2\times 2}$ der 2×2-Matrizen über dem Körper $\mathbb{R}$. Weiter sei $U := \mathbb{R}^{2\times 2}_{\text{sym}} \subseteq V$ 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 $v^* \colon V \longrightarrow \mathbb{R}$, die auf U verschwinden. Das sind gerade genau die Vielfachen von der durch $v^*(A) = a_{12} - a_{21}$ definierten Linearform v^*, also gilt

$$U^0 = \{v^* \in V^* \mid v^*(A) = \alpha\, (a_{12} - a_{21}) \text{ für ein } \alpha \in \mathbb{R}\}.$$

(ii) Wir betrachten den Vektorraum $V := \mathbb{R}^{3\times 3}$ der 3×3-Matrizen über dem Körper $\mathbb{R}$. Weiter 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 $\mathbb{R}^{3\times 3}$, sodass diese drei Linearformen (und damit auch alle Linearkombinationen) dort verschwinden, also

$${}^0F = \left\{ A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3} \,\middle|\, \begin{matrix} a_{12} + a_{21} = 0, & a_{11} = 0 \\ a_{13} + a_{31} = 0, & a_{22} = 0 \\ a_{23} + a_{32} = 0, & a_{33} = 0 \end{matrix} \right\},$$

also gilt ${}^0F = \mathbb{R}^{3\times 3}_{\text{skew}}$ (die antisymmetrischen 3×3-Matrizen).

△

Satz 21.23 ((Prä-)Annihilator eines Unterraumes).
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B = (v_1, \ldots, v_n)$ und dualer Basis $B^* = (v_1^*, \ldots, v_n^*)$, $n \in \mathbb{N}_0$.

(i) Ist $(v_1, \ldots, v_k)$ für $k \in \mathbb{N}_0$ eine Basis des Unterraumes $U \subseteq V$, dann ist $(v_{k+1}^*, \ldots, v_n^*)$ eine Basis von U^0. Insbesondere gilt

$$\dim(U^0) = \dim(V) - \dim(U) = \operatorname{codim}(U). \tag{21.16}$$

(ii) Ist $(v_1^*, \ldots, v_k^*)$ für $k \in \mathbb{N}_0$ eine Basis des Unterraumes $F \subseteq V^*$, dann ist $(v_{k+1}, \ldots, v_n)$ eine Basis von ${}^0F$. Insbesondere gilt

$$\dim({}^0F) = \dim(V^*) - \dim(F) = \dim(V) - \dim(F) = \operatorname{codim}(F). \tag{21.17}$$

Beweis. Wir beweisen nur Aussage (*i*). Als Teilfamilie einer Basis ist $(v_{k+1}^*, \dots, v_n^*)$ linear unabhängig (Lemma 13.4). Zu zeigen ist noch, dass $(v_{k+1}^*, \dots, v_n^*)$ ein Erzeugendensystem 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., $v_i^* \in U^0$ 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 $v^* \in U^0 \subseteq V^*$. v^* ist eindeutig durch seine Bilder auf der Basis $B = (v_1, \dots, v_n)$ festgelegt. Es gilt

$$\begin{aligned} v^* &= \sum_{i=1}^n v^*(v_i)\, v_i^* && \text{nach (21.6)} \\ &= \sum_{i=1}^k \underbrace{v^*(v_i)}_{=0}\, v_i^* + \sum_{i=k+1}^n v^*(v_i)\, v_i^* && \text{wegen } v^* \in U^0 \text{ und } v_i \in U \text{ für } i = 1, \dots, k \\ &= \sum_{i=k+1}^n v^*(v_i)\, v_i^*. \end{aligned}$$

Das bedeutet aber $v^* \in \langle v_{k+1}^*, \dots, v_n^* \rangle$. □

Das folgende Lemma zeigt, dass jeder Unterraum der Prä-Annihilator seines Annihilators ist.

Lemma 21.24 (Unterräume sind Prä-Annihilatoren[44]).
Es sei V ein Vektorraum über dem Körper K. Für jeden Unterraum $U \subseteq V$ gilt

$$U = \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*) = {}^0(U^0). \tag{21.18}$$

Beweis. Nach Definition (21.15) gilt

$${}^0(U^0) = \{v \in V \mid \langle v^*, v \rangle_{V^*,V} = 0 \text{ für alle } v^* \in U^0\} = \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*).$$

Es ist also nur noch das erste Gleichheitszeichen in (21.18) zu zeigen.

Wir zeigen zunächst $U \subseteq \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*)$. Es sei dazu $u \in U$. Für beliebiges $v^* \in U^0$ gilt nach Definition (21.14) $\langle v^*, u \rangle = 0$, also $u \in \operatorname{Kern}(v^*)$. Da v^* beliebig war, folgt $U \subseteq \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*)$.

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. Weil u nicht in U liegt, bleibt $\widehat{B}$ nach Hinzufügen von u eine linear unabhängige Menge. Wir ergänzen nun $\widehat{B}$ 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. Wir setzen $\langle v^*, u \rangle := 1$ 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 $u \notin \bigcap_{v^* \in U^0} \operatorname{Kern}(v^*)$. □

[44] Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Abschließend betrachten wir, welche Darstellung der (Prä-)Annihilator einer Teilmenge von K^n bzw. von $(K^n)^*$ hat, wenn wir die Identifikation vonn $(K^n)^*$ mit K^n wie in Beispiel 21.16 vornehmen.

Beispiel 21.25 ((Prä-)Annihilator)**.**

(i) Ist $M \subseteq K^n$ und nehmen wir die Identifikation von $(K^n)^*$ mit K^n wie in Beispiel 21.16 vor, dann gilt

$$M^0 = \{\xi \in K^n \mid \xi^\intercal x = 0 \text{ für alle } x \in M\}.$$

(ii) Wird $F \subseteq K^n$ interpretiert als Teilmenge von $(K^n)^*$, dann gilt

$$^0F = \{x \in K^n \mid \xi^\intercal x = 0 \text{ für alle } \xi \in F\}. \qquad \triangle$$

Das Beispiel zeigt, dass die Konzepte von Annihilator und Prä-Annihilator formal identisch aussehen. Dennoch ist es wichtig zu unterscheiden, ob eine Menge von Vektoren in K^n Objekte in K^n oder in $(K^n)^*$ repräsentiert.

Ende der Vorlesung 2

Ende der Woche 1

§ 21.4 Duale Homomorphismen

Das folgende Diagramm illustriert, wie wir eine lineare Abbildung $f \in \operatorname{Hom}(V, W)$ „dualisieren" können. Das führt zu einer zugehörigen **dualen linearen Abbildung** $f^* \in \operatorname{Hom}(W^*, V^*)$:

$$\begin{tikzcd}
W \arrow[dr, "w^* \in W^*"'] & & V \arrow[ll, "f \in \mathrm{Hom}(V,W)"'] \arrow[dl, "v^* = w^* \circ f \in V^*"] \\
& K &
\end{tikzcd}$$

Definition 21.26 (dualer Homomorphismus)**.**

Es seien V und W Vektorräume über demselben Körper K. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung. Dann heißt

$$f^* \colon W^* \ni w^* \mapsto v^* \coloneqq w^* \circ f \in V^* \tag{21.19}$$

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).[45] $\triangle$

[45] Die Notation für die duale Abbildung ist in der Literatur nicht einheitlich. Man findet auch Bezeichnungen wie f^*, f', $f^\vee$ und weitere. Manche Autoren, vor allem in Texten zur Funktionalanalysis (Analysis in unendlich-dimensionalen Vektorräumen) verwenden auch den Begriff **adjungierter Homomorphismus** oder **adjungierte lineare Abbildung** (englisch: adjoint homomorphism, adjoint linear map). Wir verwenden das Attribut „adjungiert" später in § 34.4 aber für etwas anderes.

Anders formuliert ist die zu f duale Abbildung f^* nichts Anderes als eine Abkürzung für den Ausdruck „$\cdot \circ f$“, in den man an der Stelle $\cdot$ Elemente aus W^* einsetzen kann. Nochmal anders gesagt ist f^* definiert durch die Bedingung

$$\big\langle f^*(w^*)\,,v\big\rangle_{V^*,V} = \big\langle w^*\,,f(v)\big\rangle_{W^*,W} \tag{21.20}$$

für alle $v \in V$ und alle $w^* \in W^*$.

Rücken wir in (21.19) die Linearformen in den Vordergrund, dann stellen wir fest, dass f^* eine Linearform w^* auf W in eine Linearform $v^* = w^* \circ f$ auf V verwandelt. Die lineare Abbildung $f \in \operatorname{Hom}(V, W)$ wird also benutzt, um die Linearform w^* auf W zu einer Linearform v^* auf V „zurückzuziehen“. Man bezeichnet daher die Linearform $v^* = w^* \circ f$ auch als **Pullback** (englisch: pullback) von w^* durch f.

Wir zeigen nun, dass duale Abbildungen wieder linear sind, sodass die Bezeichnungen aus Definition 21.26 gerechtfertigt sind.

Lemma 21.27 (der duale Homomorphismus ist selbst Homomorphismus)**.**
Es seien V und W Vektorräume über demselben Körper K. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung. Dann ist $f^* \in \operatorname{Hom}(W^*, V^*)$.

Beweis. f^* ist additiv, denn es gilt

$$\begin{aligned}
f^*(w_1^* + w_2^*) &= (w_1^* + w_2^*) \circ f && \text{nach Definition von } f^* \\
&= w_1^* \circ f + w_2^* \circ f && \text{nach Definition der Addition von Abbildungen} \\
& && \text{von Homomorphismen (Satz 17.11)} \\
&= f^*(w_1) + f^*(w_2) && \text{nach Definition von } f^*.
\end{aligned}$$

Außerdem ist f^* homogen wegen

$$\begin{aligned}
f^*(\alpha\, w^*) &= (\alpha\, w^*) \circ f && \text{nach Definition von } f^* \\
&= \alpha\, (w^* \circ f) && \text{nach Definition der S-Multiplikation von Abbildungen} \\
& && \text{von Homomorphismen (Satz 17.11)} \\
&= \alpha\, f^*(w^*) && \text{nach Definition von } f^*.
\end{aligned}$$

□

Beispiel 21.28 (der duale Homomorphismus ist selbst Homomorphismus)**.**

(*i*) Es sei V ein Vektorraum über dem Körper K. 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 sei $V = K^n$ über dem Körper K und $\pi_i\colon K^n \to K$ die Projektion auf die i-te Koordinate, siehe (17.4). Dann ist $\pi_i^*\colon K^* \to (K^n)^* \cong K_n$ definiert durch die Bedingungen

$$\big\langle \pi_i^*(y^*)\,,x\big\rangle_{(K^n)^*,K^n} = \big\langle y^*\,,\pi_i(x)\big\rangle_{K^*,K}$$

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 Beispiel 21.16 können wir π_i^* als Abbildung $K^n \to K$ auch darstellen durch

$$\pi_i^*(y) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile}. \\ \\ \\ \\ \end{array}$$

(*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 $f^* \colon W^* \ni w^* \mapsto v^* := f^*(w^*) \in V^*$ 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

$$\big\langle w^*, (z_1, z_2, z_3, z_4, z_5)\big\rangle = z_3 + z_4$$

gegebene Linearform w^* ist $v^* = f^*(w^*)$ gegeben durch

$$\big\langle v^*, (y_1, y_2, y_3, y_4, y_5)\big\rangle = y_2 + y_3. \qquad \triangle$$

Lemma 21.29 (Dualisieren einer Komposition von Homomorphismen)**.**
Es seien K ein Körper und U, V und W Vektorräume über K. Ist $f \in \operatorname{Hom}(V, W)$ und $g \in \operatorname{Hom}(U, V)$, dann gilt für die duale Abbildung der Komposition $f \circ g \in \operatorname{Hom}(U, W)$:

$$(f \circ g)^* = g^* \circ f^*. \tag{21.21}$$

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-2.2. □

Wir zeigen nun, dass verschiedene Homomorphismen auch verschiedene duale Abbildungen haben.

Satz 21.30 (die Zuordnung zur dualen Abbildung ist ein injektiver Vektorraumhomomorphismus)**.**
Es seien V und W Vektorräume über demselben Körper K.

(*i*) Die Zuordnung

$$I\colon \operatorname{Hom}(V, W) \ni f \mapsto f^* \in \operatorname{Hom}(W^*, V^*)$$

einer linearen Abbildung zu ihrer dualen Abbildung ist ein injektiver Homomorphismus von Vektorräumen.[46]

[46] Die Injektivität („Zwei verschiedene Homomorphismen haben auch verschiedene duale Homomorphismen.") hängt im Fall, dass W unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

(*ii*) Sind V und W endlich-dimensional, dann ist I auch surjektiv, also ein Isomorphismus von Vektorräumen. Es gilt dann

$$\dim(\operatorname{Hom}(W^*, V^*)) = \dim(\operatorname{Hom}(V, W)) = \dim(V)\ \dim(W). \tag{21.22}$$

Beweis. Aussage (*i*): I ist additiv, denn für $f, g \in \operatorname{Hom}(V, W)$ und $w^* \in W^*$ gilt

$$\begin{aligned}
(I(f+g))(w^*) &= w^* \circ (f+g) && \text{nach Definition von } I \\
&= w^* \circ f + w^* \circ g && \text{denn } w^* \text{ ist linear} \\
&= (I(f))(w^*) + (I(g))(w^*) && \text{nach Definition von } I \\
&= (I(f) + I(g))(w^*).
\end{aligned}$$

Außerdem ist I homogen, denn es gilt

$$\begin{aligned}
(I(\alpha f))(w^*) &= w^* \circ (\alpha f) && \text{nach Definition von } I \\
&= \alpha\, w^* \circ (f) && \text{denn } w^* \text{ ist linear} \\
&= \alpha\,(I(f)(w^*)) && \text{nach Definition von } I \\
&= ((\alpha I)(f))(w^*).
\end{aligned}$$

Weiterhin ist I injektiv, denn $f \in \operatorname{Kern}(I)$ heißt, dass $f^* = I(f) \in \operatorname{Hom}(W^*, V^*)$ die Nullabbildung ist, also gilt

$$0_{V^*} = f^*(w^*) = w^* \circ f$$

für alle $w^* \in W^*$, oder auch

$$w^*(f(v)) = 0$$

für alle $w^* \in W^*$ und alle $v \in V$. Das bedeutet aber nach Lemma 21.20

$$f(V) \subseteq \bigcap_{w^* \in W^*} \operatorname{Kern}(w^*) = {}^0W^* = \{0_W\}.$$

Mit anderen Worten, f ist die Nullabbildung. Also ist I injektiv.

Aussage (*ii*): Wenn V und W endlich-dimensional sind, dann ist auch $\operatorname{Hom}(V, W)$ endlich-dimensional, und es gilt (Folgerung 19.6) $\dim(\operatorname{Hom}(V, W)) = \dim(V)\ \dim(W)$. Außerdem gilt nach Folgerung 21.9 auch $\dim(V) = \dim(V^*)$ und $\dim(W) = \dim(W^*)$, also

$$\dim(\operatorname{Hom}(V, W)) = \dim(V)\ \dim(W) = \dim(V^*)\ \dim(W^*) = \dim(\operatorname{Hom}(W^*, V^*)).$$

Die lineare Abbildung I bildet also zwischen zwei Vektorräumen derselben endlichen Dimension ab. Da I in Aussage (*i*) bereits als injektiv gezeigt wurde, folgt aus Folgerung 18.9, dass I auch surjektiv ist, also ein Isomorphismus. □

Bemerkung 21.31 (zur Surjektivität von I)**.**
Man kann allgemeiner zeigen, dass I surjektiv ist, sobald W endliche Dimension besitzt. Dieser Beweis benötigt dann wieder das Auswahlaxiom, sofern V unendlich-dimensional ist. Andererseits kann man zeigen, dass I tatsächlich nicht surjektiv ist, wenn W unendlich-dimensional und V nicht der Nullraum ist. △

Schließlich erhalten wir folgende Aussage über den Zusammenhang von Injektivität und Surjektivität von Homomorphismen und ihren Dualen:

Satz 21.32 (Injektivität und Surjektivität des dualen Homomorphismus[47]).
Es seien K ein Körper und V und W zwei Vektorräume über K. Dann gilt:

(i) $f \in \operatorname{Hom}(V, W)$ ist genau dann injektiv, wenn $f^* \in \operatorname{Hom}(W^*, V^*)$ surjektiv ist.

(ii) $f \in \operatorname{Hom}(V, W)$ ist genau dann surjektiv, wenn $f^* \in \operatorname{Hom}(W^*, V^*)$ injektiv ist.

(iii) $f \in \operatorname{Hom}(V, W)$ ist genau dann bijektiv, wenn $f^* \in \operatorname{Hom}(W^*, V^*)$ bijektiv ist.

(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. Aussage (i): 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 $g \in \operatorname{Hom}(U, V)$. Es sei weiter $\widehat{U}$ ein zu U komplementärer Unterraum von W.

Wir wollen zeigen, dass $f^* \in \operatorname{Hom}(W^*, V^*)$ surjektiv ist. Es sei dazu $v^* \in V^*$ beliebig. Wir definieren das Element $w^* \in W^*$ durch

$$w^* : W = U \oplus \widehat{U} \ni (u, \widehat{u}) \mapsto \langle v^*, g(u) \rangle \in K.$$

Damit gilt

$$\begin{aligned} \langle f^*(w^*), v \rangle &= \langle w^*, f(v) \rangle && \text{nach Definition von } f^* \\ &= \langle v^*, g(f(v)) \rangle && \text{nach Definition von } w^* \text{ und wegen } f(v) \in U \\ &= \langle v^*, v \rangle && \text{da } g \text{ die Umkehrabbildung von } f|^U \text{ ist} \end{aligned}$$

für alle $v \in V$. Das heißt aber gerade $f^*(w^*) = v^*$, also ist f^* surjektiv.

Umgekehrt sei nun $f^* \in \operatorname{Hom}(W^*, V^*)$ 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 13.11 können wir leicht zeigen, dass es eine Linearform $v^* \in V^*$ gibt, deren Einschränkung auf $\operatorname{Kern}(f)$ nicht die Nullform ist. (**Quizfrage 21.3:** 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 jeder 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.

Aussage (ii): Es sei zunächst f surjektiv. Es seien $w_1^*, w_2^* \in W^*$ mit der Eigenschaft $f^*(w_1^*) = f^*(w_2^*)$ gegeben. Also gilt $w_1^* \circ f = w_2^* \circ f$. Da f surjektiv ist, existiert zu beliebigem $w \in W$ 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 $w_1^* = w_2^*$, d. h., f^* ist injektiv.

Umgekehrt sei nun $f^* \in \operatorname{Hom}(W^*, V^*)$ 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.

Wie im Beweis von Aussage (i) können mit Hilfe des Basisergänzungssatzes 13.11 eine Linearform $w^* \in W^*$ konstruieren mit der Eigenschaft $w^*|_U = 0$ (Nullform) und $w^*|_{\widehat{U}} \neq 0$. Dann gilt aber

[47] Dieses Resultat hängt im Fall, dass V bzw. W unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

$f^*(w^*) = w^* \circ f = 0_{V^*}$, also liegt w^* in $\operatorname{Kern}(f^*)$, im Widerspruch zur angenommenen Injektivität von f^*. Also muss f surjektiv sein.

Aussage (iii): Das folgt sofort aus Aussage (i) und Aussage (ii).

Aussage (iv): Es gilt

$$\begin{aligned} f^* \circ (f^{-1})^* &= (f^{-1} \circ f)^* && \text{nach Lemma 21.29} \\ &= \mathrm{id}_V^* \\ &= \mathrm{id}_{V^*} && \text{nach Beispiel 21.28} \end{aligned}$$

und außerdem

$$\begin{aligned} (f^{-1})^* \circ f^* &= (f \circ f^{-1})^* && \text{nach Lemma 21.29} \\ &= \mathrm{id}_w^* \\ &= \mathrm{id}_{w^*} && \text{nach Beispiel 21.28.} \end{aligned}$$

Also ist $(f^{-1})^*$ die Inverse von f^*. □

§ 21.5 Darstellungsmatrizen dualer Homomorphismen

Wir untersuchen jetzt, wie die Darstellungsmatrizen einer linearen Abbildung $f \in \operatorname{Hom}(V, W)$ und ihrer dualen linearen Abbildung $f^* \in \operatorname{Hom}(W^*, V^*)$ zusammenhängen. Dazu seien V und W endlichdimensionale Vektorräume über demselben Körper K und $B_V = (v_1, \dots, v_m)$ und $B_W = (w_1, \dots, w_n)$ Basen von V bzw. von W. Wir erinnern uns, dass die Darstellungsmatrix $A = \mathcal{M}_{B_W}^{B_V}(f) = (a_{ij})$ von f bzgl. dieser Basen gegeben ist durch

$$f(v_j) = \sum_{i=1}^n a_{ij}\, w_i \quad \text{für alle } j = 1, \dots, m.$$

Wenden wir auf beide Seiten das Element $w_k^* \in W^*$ der zu B_W dualen Basis an (das ja nach Lemma 21.15 als Koordinatenermittler wirkt), so erhalten wir

$$\underbrace{\langle w_k^*, f(v_j)\rangle}_{=\langle f^*(w_k^*), v_j\rangle} = \Big\langle w_k^*, \sum_{i=1}^n a_{ij}\, w_i \Big\rangle = \sum_{i=1}^n a_{ij}\, \underbrace{\langle w_k^*, w_i\rangle}_{=\delta_{ki}} = a_{kj} \quad \text{für alle } j = 1, \dots, m. \tag{21.23}$$

Wir entscheiden uns dafür, für die Darstellungsmatrix von f^* die dualen Basen zu verwenden. Die gesuchte Matrix

$$B = \mathcal{M}_{B_V^*}^{B_W^*}(f^*) \in K^{m \times n}$$

mit Einträgen b_{ij} erfüllt dann

$$f^*(w_j^*) = \sum_{i=1}^m b_{ij}\, v_i^* \quad \text{für alle } j = 1, \dots, n.$$

Setzen wir in beide Seiten das Element v_k der Basis B_V ein, so erhalten wir

$$\langle f^*(w_j^*), v_k\rangle = \Big\langle \sum_{i=1}^m b_{ij}\, v_i^*, v_k \Big\rangle = \sum_{i=1}^m b_{ij}\, \underbrace{\langle v_i^*, v_k\rangle}_{=\delta_{ik}} = b_{kj} \quad \text{für alle } j = 1, \dots, n. \tag{21.24}$$

Ein Vergleich der linken Seiten von (21.23) und (21.24) zeigt den Zusammenhang $b_{kj} = a_{jk}$, also $B = A^\intercal$!

Wir haben damit folgenden Satz bewiesen:

Satz 21.33 (Darstellungsmatrix des dualen Homomorphismus)**.**
Es seien V und W endlich-dimensionale Vektorräume über demselben Körper K und B_V und B_W Basen von V bzw. von W. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung mit dualer Abbildung $f^* \in \operatorname{Hom}(W^*, V^*)$. Ist $A = \mathcal{M}^{B_V}_{B_W}(f)$ die Darstellungsmatrix von f bzgl. der Basen B_V und B_W, dann ist $A^\intercal = \mathcal{M}^{B_W^*}_{B_V^*}(f^*)$ die Darstellungsmatrix von f^* bzgl. der dualen Basen B_W^* und B_V^*.

Folgerung 21.34 (Rang der dualen Abbildung)**.**
Es seien V und W endlich-dimensionale Vektorräume über demselben Körper K. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung mit dualer Abbildung $f^* \in \operatorname{Hom}(W^*, V^*)$. 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 $A = \mathcal{M}^{B_V}_{B_W}(f)$ und $A^\intercal = \mathcal{M}^{B_W^*}_{B_V^*}(f^*)$ denselben Rang (Lemma 15.27). Nach Satz 19.9 haben wir also $\operatorname{Rang}(f) = \operatorname{Rang}(A) = \operatorname{Rang}(A^\intercal) = \operatorname{Rang}(f^*)$. □

Folgerung 21.35 (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 Beispiel 21.16 gilt für die duale Abbildung $(f_A)^* \in \operatorname{Hom}(K^n, K^m)$ der durch A induzierten Abbildung $f_A \in \operatorname{Hom}(K^m, K^n)$

$$(f_A)^* = f_{A^\intercal}. \tag{21.25}$$

Beweis. Die linke Seite in (21.25) ist nach (21.20) und wegen der Identifikation von K^n und K^m mit ihrem jeweiligen Dualraum definiert durch

$$\big[(f_A)^*(y)\big]^\intercal x = y^\intercal f_A(x) \quad \text{für alle } x \in K^m \text{ und } y \in K^n.$$

Nach Definition von f_A ist die rechte Seite dieser Gleichung gerade $y^\intercal (A\, x)$ oder auch auch $(y^\intercal A)\, x$. Damit muss

$$\big[(f_A)^*(y)\big]^\intercal = y^\intercal A$$

gelten und daher $(f_A)^*(y) = A^\intercal y$. Das ist aber gerade die Behauptung (21.25). □

Abschließend untersuchen wir noch, wie sich die Darstellungsmatrix eines dualen Homomorphismus bei einem Basiswechsel transformiert. Wir kennen aus Lemma 21.17 bereits den Zusammenhang

$$\mathcal{T}^{B_V^*}_{\widehat{B}_V^*} = \big[\mathcal{T}^{\widehat{B}_V}_{B_V}\big]^\intercal = \big[\mathcal{T}^{B_V}_{\widehat{B}_V}\big]^{-\intercal}$$

der Transformationsmatrizen für einen Basiswechsel im Raum V und den zugehörigen Wechsel der dualen Basen im Raum V^*. Entsprechend erhalten wir für die Transformation der Darstellungsmatrizen eines Homomorphismus $f \in \operatorname{Hom}(V, W)$ nach (20.5) und des dualen Homomorphismus $f^* \in \operatorname{Hom}(W^*, V^*)$ die folgenden Beziehungen:

$$\begin{aligned} \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_W}(f) &= \mathcal{T}^{B_W}_{\widehat{B}_W}\, \mathcal{M}^{B_V}_{B_W}(f)\, \mathcal{T}^{\widehat{B}_V}_{B_V} \\ \widehat{A} &= S \quad A \quad T^{-1} \end{aligned} \qquad \text{und} \qquad \begin{aligned} \mathcal{M}^{\widehat{B}_W^*}_{\widehat{B}_V^*}(f^*) &= \mathcal{T}^{B_V^*}_{\widehat{B}_V^*}\, \mathcal{M}^{B_W^*}_{B_V^*}(f^*)\, \mathcal{T}^{\widehat{B}_W^*}_{B_W^*} \\ \widehat{A}^\intercal &= T^{-\intercal} \quad A^\intercal \quad S^\intercal \end{aligned} \tag{21.26}$$

Ende der Vorlesung 3

§ 21.6 Die vier fundamentalen Unterräume zu einer linearen Abbildung

In diesem Abschnitt zeigen wir einen wichtigen Zusammenhang zwischen den Unterräumen Kern und Bild von f und f^*. Man spricht auch von den **vier fundamentalen Unterräumen** (englisch: four fundamental subspaces) zu einer linearen Abbildung f.

Satz 21.36 (vier fundamentale Unterräume zu einer linearen Abbildung[48]).
Es seien V und W Vektorräume über demselben Körper K. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung mit dualer Abbildung $f^* \in \operatorname{Hom}(W^*, V^*)$. Dann gilt:

$$\operatorname{Bild}(f^*) = \operatorname{Kern}(f)^0 \quad \text{in } V^* \tag{21.27a}$$

$$\operatorname{Kern}(f^*) = \operatorname{Bild}(f)^0 \quad \text{in } W^* \tag{21.27b}$$

$$\operatorname{Bild}(f) = {}^0\operatorname{Kern}(f^*) \quad \text{in } W \tag{21.27c}$$

$$\operatorname{Kern}(f) = {}^0\operatorname{Bild}(f^*) \quad \text{in } V \tag{21.27d}$$

Beweis. Zur besseren Übersicht geben wir jeweils die beiden an einer Aussage beteiligten Unterräume nochmals an.

Wir beginnen mit (21.27a), also

$$\operatorname{Bild}(f^*) = \{f^*(w^*) \in V^* \mid w^* \in W^*\}$$

$$\operatorname{Kern}(f)^0 = \{v^* \in V^* \mid \langle v^*, v\rangle = 0 \text{ für alle } v \in \operatorname{Kern}(f)\}.$$

Es sei zunächst $v^* \in \operatorname{Bild}(f^*)$, also existiert $w^* \in W^*$ mit $v^* = f^*(w^*) = w^* \circ f$. Für beliebiges $v \in \operatorname{Kern}(f)$ ist $\langle v^*, v\rangle = \langle w^*, f(v)\rangle = \langle w^*, 0_W\rangle = 0$. Also liegt v^* in $\operatorname{Kern}(f)^0$.

Umgekehrt sei $v^* \in \operatorname{Kern}(f)^0$. Wir definieren ein passendes Element $w^* \in W^*$ mit der Eigenschaft $v^* = f^*(w^*)$. Dazu definieren wir die Werte von w^* zunächst auf dem Unterraum $\operatorname{Bild}(f) \subseteq W$ durch

$$\langle w^*, f(v)\rangle := \langle v^*, v\rangle \quad \text{für alle } v \in V.$$

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, ist $\langle v^*, v_1 - v_2\rangle = 0$, also $\langle v^*, v_1\rangle = \langle v^*, v_2\rangle$.

Wir können nun w^* beliebig auf ganz W fortsetzen. (**Quizfrage 21.4:** Wie geht das genau?) Schließlich gilt per Definition

$$\langle f^*(w^*), v\rangle = \langle w^* \circ f, v\rangle = \langle w^*, f(v)\rangle = \langle v^*, v\rangle$$

für alle $v \in V$, also $f^*(w^*) = w^* \circ f = v^*$. Das zeigt (21.27a).

Nun adressieren wir (21.27b) mit

$$\begin{aligned}\operatorname{Kern}(f^*) &= \{w^* \in W^* \mid f^*(w^*) = 0_{V^*}\}\\ &= \{w^* \in W^* \mid w^* \circ f = 0_{V^*}\}\end{aligned}$$

$$\begin{aligned}\operatorname{Bild}(f)^0 &= \{w^* \in W^* \mid \langle w^*, w\rangle = 0 \text{ für alle } w \in \operatorname{Bild}(f)\}\\ &= \{w^* \in W^* \mid \langle w^*, f(v)\rangle = 0 \text{ für alle } v \in V\}.\end{aligned}$$

[48] Dieses Resultat hängt im Fall, dass V bzw. W unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Es gilt

$$\begin{aligned} & w^* \in \operatorname{Bild}(f)^0 \\ \Leftrightarrow \quad & \langle w^*, f(v)\rangle = 0 \text{ für alle } v \in V \\ \Leftrightarrow \quad & w^* \circ f = 0_{V^*} \\ \Leftrightarrow \quad & w^* \in \operatorname{Kern}(f^*). \end{aligned}$$

Wir beweisen nun (21.27c) mit

$$\operatorname{Bild}(f) = \{f(v) \in W \mid v \in V\}$$

$$\begin{aligned} {}^0\operatorname{Kern}(f^*) &= \{w \in W \mid \langle w^*, w\rangle = 0 \text{ für alle } w^* \in \operatorname{Kern}(f^*)\} \\ &= \{w \in W \mid \langle w^*, w\rangle = 0 \text{ für alle } w^* \in W^* \text{ mit } f^*(w^*) = w^* \circ f = 0_{V^*}\}. \end{aligned}$$

Es sei zunächst $w \in \operatorname{Bild}(f)$, also $w = f(v)$ für ein $v \in V$. Weiter sei $w^* \in W^*$ mit der Eigenschaft $w^* \circ f = 0_{V^*}$. Dann gilt

$$\langle w^*, w\rangle = \langle w^*, f(v)\rangle = \langle w^* \circ f, v\rangle = \langle 0_{V^*}, v\rangle = 0,$$

also $w \in {}^0\operatorname{Kern}(f^*)$.

Für die Umkehrung führen wir einen Beweis durch Kontraposition. Es sei also $w \in W$, aber $w \notin \operatorname{Bild}(f)$. Wir beginnen mit einer Basis $\widehat{B} = (w_i)_{i \in \widehat{I}}$ von $\operatorname{Bild}(f)$. Weil w nicht in $\operatorname{Bild}(f)$ liegt, bleibt $\widehat{B}$ nach Hinzufügen von w eine linear unabhängige Menge. Wir ergänzen nun $\widehat{B}$ plus w zu einer Basis $B = (w_i)_{i \in I}$ von W.

Nun definieren wir eine Linearform $w^* \in W^*$ durch Festlegung der Werte auf der Basis B. Wir setzen $\langle w^*, w\rangle \coloneqq 1$ und $\langle w^*, w_i\rangle = 0$ für alle anderen Basiselemente. Insbesondere ist also w^* die Nullform auf $\operatorname{Bild}(f)$, es gilt also $w^* \circ f = 0_{V^*}$. Dieses w^* zeigt, dass w nicht zu ${}^0\operatorname{Kern}(f^*)$ gehört, denn w^* erfüllt $w^* \circ f = 0_{V^*}$ und $\langle w^*, w\rangle = 1 \neq 0$.

Um (21.27d) zu zeigen, betrachten wir

$$\operatorname{Kern}(f) = \{v \in V \mid f(v) = 0\}$$

$$\begin{aligned} {}^0\operatorname{Bild}(f^*) &= \{v \in V \mid \langle v^*, v\rangle = 0 \text{ für alle } v^* \in \operatorname{Bild}(f^*)\} \\ &= \{v \in V \mid \langle f^*(w^*), v\rangle = 0 \text{ für alle } w^* \in W^*\} \\ &= \{v \in V \mid \langle w^*, f(v)\rangle = 0 \text{ für alle } w^* \in W^*\}. \end{aligned}$$

Es sei $v \in \operatorname{Kern}(f)$, dann gilt

$$\langle w^*, f(v)\rangle = \langle w^*, 0_W\rangle = 0$$

für alle $w^* \in W^*$, also $v \in {}^0\operatorname{Bild}(f^*)$.

Umgekehrt sei $v \in {}^0\operatorname{Bild}(f^*)$. Angenommen $f(v) \neq 0_W$, dann können wir eine Linearform $w^* \in W^*$ finden mit der Eigenschaft $\langle w^*, f(v)\rangle \neq 0$, im Widerspruch zu $v \in {}^0\operatorname{Bild}(f^*)$. Also gilt $v \in \operatorname{Kern}(f)$. □

Abschließend betrachten wir eine Version von Satz 21.36, 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 21.36 mit Hilfe der Matrizen A und $A^\intercal$ ausdrücken.

Folgerung 21.37 (vier fundamentale Unterräume zu einer Matrix)**.**

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 Beispiel 21.16 gilt für die Matrix A und ihre Transponierte $A^\intercal$

$$\begin{aligned} \operatorname{Bild}(A^\intercal) &= \operatorname{Kern}(A)^0 && \text{in } K^m && (21.28a)\\ \operatorname{Kern}(A^\intercal) &= \operatorname{Bild}(A)^0 && \text{in } K^n && (21.28b)\\ \operatorname{Bild}(A) &= {}^0\operatorname{Kern}(A^\intercal) && \text{in } K^n && (21.28c)\\ \operatorname{Kern}(A) &= {}^0\operatorname{Bild}(A^\intercal) && \text{in } K^m. && (21.28d) \end{aligned}$$

Beachte: Wir könnten nach Beispiel 21.25 die Beziehungen (21.28c) und (21.28d) auch als $\operatorname{Bild}(A) = \operatorname{Kern}(A^\intercal)^0$ bzw. als $\operatorname{Kern}(A) = \operatorname{Bild}(A^\intercal)^0$ schreiben. Es ist also (21.28a) $\Leftrightarrow$ (21.28c) und (21.28b) $\Leftrightarrow$ (21.28d).

Beweis. Die Matrix A induziert die Abbildung $f_A \in \operatorname{Hom}(K^m, K^n)$, und nach Folgerung 21.35 induziert $A^\intercal$ die Abbildung $(f_A)^*$. Nach Definition von Bild und Kern für Matrizen (siehe § 19.3) folgt das Resultat sofort aus Satz 21.36. □

Beispiel 21.38 (Beschreibung von Unterräumen durch Basen bzw. Gleichungen)**.**

(i) Gegeben sei der Unterraum

$$U = \left\langle \left\{ \begin{pmatrix} 1\\3\\-2\\1 \end{pmatrix}, \begin{pmatrix} 0\\1\\4\\0 \end{pmatrix} \right\} \right\rangle$$

von $\mathbb{Q}^4$. Gesucht ist eine 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 (21.28c) gilt $\operatorname{Bild}(A) = {}^0\operatorname{Kern}(A^\intercal)$. Wie in § 16 können wir eine Basis von $\operatorname{Kern}(A^\intercal)$ bestimmen, indem wir

$$A^\intercal = \begin{bmatrix} 1 & 3 & -2 & 1\\ 0 & 1 & 4 & 0 \end{bmatrix}$$

auf reduzierte Zeilenstufenform bringen. Es ergibt sich

$$\begin{bmatrix} 1 & 0 & -14 & 1\\ 0 & 1 & 4 & 0 \end{bmatrix},$$

also die Basis

$$\left\{ \begin{pmatrix} 14\\-4\\1\\0 \end{pmatrix}, \begin{pmatrix} -1\\0\\0\\1 \end{pmatrix} \right\}$$

von Kern($A^\intercal$). Wegen (21.28c) gilt nun also

$$U = \operatorname{Bild}(A) = {}^0\operatorname{Kern}(A^\intercal) = \left\{ y \in \mathbb{Q}^4 \,\middle|\, \begin{pmatrix} 14 \\ -4 \\ 1 \\ 0 \end{pmatrix}^\intercal \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = 0 \text{ und } \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}^\intercal \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = 0 \right\}$$
$$= \left\{ y \in \mathbb{Q}^4 \,\middle|\, \begin{aligned} 14\, y_1 - 4\, y_2 + y_3 &= 0 \\ -y_1 + y_4 &= 0 \end{aligned} \right\}.$$

Somit haben wir wie gewünscht U durch Gleichungen beschrieben.

(ii) Es sei umgekehrt der Unterraum eines Vektorraumes durch Gleichungen beschrieben, etwa

$$U = \left\{ x \in \mathbb{R}^5 \,\middle|\, \begin{aligned} x_1 + 3\, x_3 - x_5 &= 0 \\ x_1 + x_2 + x_4 &= 0 \end{aligned} \right\} = \left\{ x \in \mathbb{R}^5 \,\middle|\, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \\ -1 \end{pmatrix}^\intercal \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0 \text{ und } \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}^\intercal \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0 \right\}.$$

Gesucht ist eine Beschreibung mit Hilfe einer Basis. Wir können die gegebene Beschreibung von U auffassen als

$$U = \operatorname{Bild}(A)^0 \text{ mit } A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 3 & 0 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Nach (21.28b) gilt also auch $U = \operatorname{Kern}(A^\intercal)$. Wir bestimmen die reduzierte Zeilenstufenform zu $A^\intercal$ als

$$\begin{bmatrix} 1 & 0 & 3 & 0 & -1 \\ 0 & 1 & -3 & 1 & 1 \end{bmatrix}$$

und können daraus eine Basis von $U = \operatorname{Kern}(A^\intercal)$ 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. △

§ 21.7 Zusammenspiel zwischen Dualräumen und Faktorräumen

In diesem Abschnitt untersuchen wir Dualräume von Faktorräumen und Faktorräume von Dualräumen.

Lemma 21.39 (Dualraum eines Faktorraumes)**.**
Es sei V ein Vektorraum über dem Körper K 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 $\pi^*|^{U^0}$ ist ein Isomorphismus, es gilt also

$$(V/U)^* \cong U^0. \tag{21.29}$$

Beachte: Der Vektorraum der Linearformen auf V, die auf dem Unterraum U verschwinden, ist also isomorph zum Vektorraum der Linearformen, die direkt nur auf dem gröberen Faktorraum V/U wirken.

Beweis. Die kanonische Surjektion $\pi\colon V \to V/U$ ist surjektiv (Satz 17.13). Die duale Abbildung $\pi^*\colon (V/U)^* \to V^*$ ist also injektiv (Satz 21.32); dieses Resultat verwendet nicht das Auswahlaxiom. Wir müssen nun das Bild von π^* identifizieren. Dazu sehen wir uns an, wie π^* auf ein $[v]^* \in (V/U)^*$ wirkt. Es gilt

$$\begin{aligned} \langle \pi^*([v]^*), v\rangle &= \langle [v]^*, \pi(v)\rangle && \text{nach Definition von } \pi^* \\ &= \langle [v]^*, v+U\rangle && \text{nach Definition von } \pi. \end{aligned}$$

Setzen wir für v speziell ein Element aus $u \in U$ ein, so erkennen wir

$$\begin{aligned} \langle \pi^*([v]^*), u\rangle &= \langle [v]^*, U\rangle && \text{denn } u+U=U \\ &= 0 && \text{denn } U=[0] \text{ ist der Nullvektor in } V/U \text{ und } [v]^* \text{ ist linear.} \end{aligned}$$

Das bedeutet aber, dass $\pi^*([v]^*)$ für beliebiges $[v]^* \in (V/U)^*$ zum Unterraum U^0 von V^* gehört.

In der Tat wird jedes Element von U^0 auch angenommen: Zu $v^* \in U^0$ definiere $[v]^* \in (V/U)^*$ durch

$$\langle [v]^*, \underbrace{v+U}_{\text{beliebiges Element von } V/U}\rangle \coloneqq \langle v^*, v\rangle.$$

Diese Vorschrift ist wohldefiniert, also unabhängig von der Wahl des Repräsentanten $v \in v+U = [v]$. (**Quizfrage 21.5:** Warum?) Es gilt nun also

$$\langle \pi^*([v]^*), v\rangle = \langle [v]^*, v+U\rangle = \langle v^*, v\rangle$$

für alle $v \in V$, was wie gewünscht $\pi^*([v]^*) = v^*$ zeigt.

Damit ist die Einschränkung $\pi^*|^{U^0}\colon (V/U)^* \to U^0$ tatsächlich ein Isomorphismus. □

Lemma 21.40 (Faktorraum eines Dualraumes[49])**.**
Es sei V ein Vektorraum über dem Körper K 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

$$V^*/U^0 \cong U^*. \tag{21.30}$$

Beweis. Es sei $i\colon U \to V$ die kanonische Injektion. Nach Satz 21.36 gilt $\operatorname{Kern}(i^*) = \operatorname{Bild}(i)^0 = {}^0U$. Weiter ist nach Satz 21.32 ist $i^*\colon V^* \to U^*$ surjektiv. Der Homomorphiesatz für Vektorräume 17.17, angewendet auf i^*, zeigt nun $V^*/\operatorname{Kern}(i^*) \cong \operatorname{Bild}(i^*)$, also gilt $V^*/U^0 \cong U^0$. □

Ende der Vorlesung 4

Ende der Woche 2

[49]Dieses Resultat hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

§ 21.8 Der Bidualraum

In diesem Abschnitt betrachten wir den Dualraum eines Dualraums.

Definition 21.41 (Bidualraum)**.**
Es sei V ein Vektorraum über dem Körper K 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^{**} besteht also aus Linearformen auf V^*. Wir können uns fragen, wie die Elemente des Bidualraumes 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^* ist. Diese können wir mit $\langle \cdot , v \rangle$ bezeichnen.

Satz 21.42 (kanonische Injektion[50])**.**
Es sei V ein Vektorraum über dem Körper K und V^* sein Dualraum.

(i) Die Abbildung
$$i_V \coloneqq V \ni v \mapsto \langle \cdot , v \rangle \in V^{**} \tag{21.31}$$
ist ein injektiver Homomorphismus, genannt die **kanonische Injektion** (englisch: canonical injection) von V in V^{**}.

(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 $v_1, v_2 \in V^*$ und $v^* \in V^*$ gilt

$$\begin{aligned}
\big(i_V(v_1 + v_2)\big)(v^*) &= \langle v^*, v_1 + v_2 \rangle && \text{nach Definition von } i_V \\
&= \langle v^*, v_1 \rangle + \langle v^*, v_2 \rangle && \text{denn } v^* \text{ ist eine Linearform auf } V \\
&= i_V(v_1)(v^*) + i_V(v_2)(v^*) && \text{nach Definition von } i_V \\
&= \big(i_V(v_1) + i_V(v_2)\big)(v^*) && \text{nach Definition der Addition von Abbildungen,}
\end{aligned}$$

also $i_V(v_1 + v_2) = i_V(v_1) + i_V(v_2)$, d. h., i_V ist additiv. Außerdem ist i_V homogen, denn mit $\alpha \in K$ gilt

$$\begin{aligned}
\big(i_V(\alpha\, v)\big)(v^*) &= \langle v^*, \alpha\, v \rangle && \text{nach Definition von } i_V \\
&= \alpha \, \langle v^*, v \rangle && \text{denn } v^* \text{ ist eine Linearform auf } V \\
&= \alpha \, i_V(v) && \text{nach Definition von } i_V \\
&= (\alpha \, i_V)(v) && \text{nach Definition der S-Multiplikation von Abbildungen.}
\end{aligned}$$

Um die Injektivität zu prüfen, betrachten wir ein $v \in \operatorname{Kern}(i_V)$. Es gilt also

$$i_V(v) = 0_{V^{**}}, \quad \text{d. h., } \big(i_V(v)\big)(v^*) = \langle v^*, v \rangle = 0.$$

Mit anderen Worten, es gilt $v \in \operatorname{Kern}(v^*)$ für alle $v^* \in V^*$. Nach Lemma 21.20 Aussage (iv) bedeutet das aber $v = 0$. Also ist i_V injektiv.

[50] Die Aussage (i) hängt im Fall, dass V unendlich-dimensional ist, vom Zornschen Lemma und damit vom Auswahlaxiom ab.

Aussage (*ii*): Im Fall $\dim(V) = n \in \mathbb{N}_0$ gilt auch $\dim(V^*) = n$, siehe Folgerung 21.9. Wenden wir dieses Resultat erneut an, 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, sondern sogar bijektiv ist, also ein Isomorphismus. □

Beachte: Im Unterschied zu dem in Folgerung 21.9 etablierten Isomorphismus zwischen V und V^* (wenn V endlich-dimensional ist) ist die kanonische Injektion i_V (wie das Attribut „kanonisch" bereits andeutet) nicht von der Wahl einer Basis abhängig.

Folgerung 21.43 (Eigenschaften von Annihilatoren im Dualraum)**.**
Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Dann gilt

(*i*) Für jede Teilmenge $F \subseteq V^*$ gilt $F^0 = i_V({}^0F)$.

(*ii*) Für jeden Unterraum U von V gilt $(U^0)^0 = i_V(U)$.

Beweis. Aussage (*i*): Nach Definition (21.14) gilt

$$\begin{aligned} F^0 &= \{f \in V^{**} \mid \langle f, v^* \rangle_{V^{**},V^*} = 0 \text{ für alle } v^* \in F\} \\ &= \{i_V(v) \in V^{**} \mid \langle v^*, v \rangle_{V^*,V} = 0 \text{ für alle } v^* \in F\}, \end{aligned}$$

wobei die bijektive Ersetzung $f = i_V(v)$ (Satz 21.42) verwendet wurde. Definition (21.15) ergibt

$$\begin{aligned} {}^0F &= \{v \in V \mid \langle v^*, v \rangle_{V^*,V} = 0 \text{ für alle } v^* \in F\}, \\ \text{also} \quad i_V({}^0F) &= \{i_V(v) \in V^{**} \mid \langle v^*, v \rangle_{V^*,V} = 0 \text{ für alle } v^* \in F\} \\ &= F^0. \end{aligned}$$

Aussage (*ii*): Nach Satz 21.42 ist $i_V \in \operatorname{Hom}(V, V^{**})$ bijektiv. Es gilt

$$\begin{aligned} U &= {}^0(U^0) && \text{nach Lemma 21.24} \\ &= i_V^{-1}\Big(i_V\big({}^0(U^0)\big)\Big) && \text{wegen } i_V^{-1} \circ i_V = \mathrm{id}_V \\ &= i_V^{-1}\big((U^0)^0\big) && \text{nach Aussage } (i) \text{ mit } F = U^0. \end{aligned}$$

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 $f \in \operatorname{Hom}(V, W)$ zweimal dualisieren. Das führt zum Begriff des **bidualen Homomorphismus** (englisch: bidual homomorphism) bzw. der **bidualen linearen Abbildung** (englisch: bidual linear map) $f^{**} \in \operatorname{Hom}(V^{**}, W^{**})$. Diese ist definiert durch

$$f^{**} \colon V^{**} \ni v^{**} \mapsto w^{**} := v^{**} \circ f^* \in W^{**}. \tag{21.32}$$

Lemma 21.44 (Zusammenhang zwischen f und f^{**})**.**
Es seien V und W Vektorräume über demselben Körper K sowie $f \in \operatorname{Hom}(V, W)$ ein Homomorphismus. Dann gilt

$$i_W \circ f = f^{**} \circ i_V. \tag{21.33}$$

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 sind per Definition lineare Abbildungen $V \to W^{**}$. Um die Gleichheit zu zeigen, setzen wir ein beliebiges Element $v \in V$ ein und lassen das Ergebnis wirken auf ein beliebiges Element $w^* \in W^*$. Wenn wir zeigen können, dass $\langle (i_W \circ f)(v), w^* \rangle_{W^{**},W^*} = \langle (f^{**} \circ i_V)(v), w^* \rangle_{W^{**},W^*}$ gilt, so ist die Behauptung bewiesen. (**Quizfrage 21.6:** Warum?) Es gilt

$$\begin{aligned}
\left\langle (f^{**} \circ i_V)(v), w^* \right\rangle_{W^{**},W^*} &= \left\langle f^{**}(i_V(v), w^* \right\rangle_{W^{**},W^*} && \text{nach Definition der Komposition} \\
&= \left\langle i_V(v) \circ f^*, w^* \right\rangle_{W^{**},W^*} && \text{nach Definition (21.32) von } f^{**} \\
&= \left\langle i_V(v), f^*(w^*) \right\rangle_{V^{**},V^*} && \text{wegen der Assoziativität der Komposition} \\
&= \left\langle i_V(v), w^* \circ f \right\rangle_{V^{**},V^*} && \text{nach Definition (21.19) von } f^* \\
&= \langle w^* \circ f, v \rangle_{V^*,V} && \text{nach Definition (21.31) von } i_V \\
&= \langle w^*, f(v) \rangle_{W^*,W} && \text{wegen der Assoziativität der Komposition} \\
&= \langle i_W(f(v)), w^* \rangle_{W^{**},W^*} && \text{nach Definition (21.31) von } i_V \\
&= \langle (i_W \circ f)(v), w^* \rangle_{W^{**},W^*} && \text{nach Definition der Komposition.}
\end{aligned}$$

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 (21.33). □

§ 22 Tensorprodukte und multilineare Abbildungen

Literatur: Fischer, Springborn, 2020, Kapitel 7.3–7.4, Hackbusch, 2019, Chapter 3

§ 22.1 Bilineare Abbildungen

Definition 22.1 (bilineare Abbildung, Bilinearform).
Es seien U, V, W Vektorräume über demselben Körper K.

(i) Eine Abbildung
$$f\colon U \times V \to W$$
heißt **bilinear** (englisch: bilinear map), wenn für jedes feste $\overline{u} \in U$ und jedes feste $\overline{v} \in V$ die Abbildungen
$$f(\overline{u}, \cdot)\colon V \ni v \mapsto f(\overline{u}, v) \in W$$
$$f(\cdot, \overline{v})\colon U \ni u \mapsto f(u, \overline{v}) \in W$$
beide linear sind.

(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 $W = K$ nennen wir eine **Bilinearform** (englisch: bilinear form) auf $U \times V$.

(iv) Die Menge aller Bilinearformen $U \times V \to K$ bezeichnen wir mit $\operatorname{Bil}(U, V)$ oder $\operatorname{Bil}(U, V; K)$. △

Beispiel 22.2 (bilineare Abbildung, Bilinearform).

(*i*) Die duale Paarung (21.2), 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 21.7.

(*ii*) Für jede Matrix $A \in K^{n\times m}$ ist die Abbildung

$$K^m \times K^n \ni (x, y) \mapsto x^\mathsf{T} A\, y \in K$$

eine Bilinearform auf $K^m \times K^n$.

(*iii*) Die Multiplikation zweier Polynome über einem Körper K, also

$$K[t] \times K[t] \ni (p, q) \mapsto p \cdot q \in K[t],$$

siehe (11.3c), ist eine bilineare Abbildung in $\operatorname{Bil}(K[t], K[t]; K[t])$.

(*iv*) Auch die Multiplikation zweier Polynome in verschiedenen Variablen über einem Körper K, also

$$K[s] \times K[t] \ni (p, q) \mapsto p \cdot q \in K[s, t]$$

in den Vektorraum $K[s, t]$ der Polynome in beiden Variablen ist eine bilineare Abbildung in $\operatorname{Bil}(K[s], K[t]; K[s, t])$. △

Lemma 22.3 (bilineare Abbildungen bilden einen Vektorraum)**.**

Es seien U, V, W Vektorräume über demselben Körper K. Dann ist $\operatorname{Bil}(U, V; W)$ ein Unterraum des Vektorraumes $W^{U\times V} = \{f\colon U \times V \to W\}$ aller Abbildungen $U \times V \to W$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-3.3. □

Genau wie lineare Abbildungen zwischen Vektorräumen können wir auch bilineare Abbildung dadurch eindeutig bestimmen, dass wir ihre Bilder auf einer geeigneten Menge festlegen.

Satz 22.4 (Existenz- und Eindeutigkeitssatz für bilineare Abbildungen, vgl. Satz 17.7)**.**

Es seien K ein Körper und U, V und W Vektorräume über 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 $f\colon U\times V \to W$ mit der Eigenschaft $f(u_i, v_j) = w_{ij}$ für alle $(i, j) \in I \times J$.

Beweis. Es sei $(u, v) \in U \times V$ beliebig. Wegen der Basiseigenschaft (Satz 13.10) gibt es (bis auf Summanden mit Nullkoeffizienten) eindeutige Darstellungen

$$u = \sum_{i\in I'} \alpha_i\, u_i \quad \text{und} \quad v = \sum_{j\in J'} \beta_j\, v_j$$

mit endlichen Teilfamilien I' von I und J' von J.

Wenn es eine gesuchte bilineare Abbildung $f\colon U \times V \to W$ gibt, so muss diese notwendigerweise

$$\begin{aligned} f(u,v) &= f\Big(\sum_{i\in I'} \alpha_i\, u_i, \sum_{j\in J'} \beta_j\, v_j\Big) \\ &= \sum_{\substack{i\in I'\\ j\in J'}} \alpha_i\, \beta_j\, f(u_i, v_j) && \text{wegen der Bilinearität} \\ &= \sum_{\substack{i\in I'\\ j\in J'}} \alpha_i\, \beta_j\, w_{ij} && \text{wegen der gewünschten Bilder} \end{aligned}$$

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 Vorschrift

$$f(u,v) \coloneqq \sum_{\substack{i\in I'\\ j\in J'}} \alpha_i\, \beta_j\, w_{ij}$$

für $u = \sum_{i\in I'} \alpha_i\, u_i$ und $v = \sum_{j\in J'} \beta_j\, v_j$ tatsächlich eine bilineare Abbildung. Um das zu sehen, halten wir zunächst $v \in V$ fest und betrachten beliebige $u, \overline{u} \in U$. Diese haben die Darstellungen $u = \sum_{i\in I'} \alpha_i\, u_i$ und $\overline{u} = \sum_{i\in I'} \overline{\alpha}_i\, u_i$. (**Quizfrage 22.1:** Warum können wir dieselbe endliche Indexmenge I' verwenden?) Dann ist

$$f(u+\overline{u}, v) = \sum_{\substack{i\in I'\\ j\in J'}} (\alpha_i + \overline{\alpha}_i)\, \beta_j\, w_{ij} = \sum_{\substack{i\in I'\\ j\in J'}} \alpha_i\, \beta_j\, w_{ij} + \sum_{\substack{i\in I'\\ j\in J'}} \overline{\alpha}_i\, \beta_j\, w_{ij} = f(u,v) + f(\overline{u}, v)$$

und für $\alpha \in K$ auch

$$f(\alpha\, u, v) = \sum_{\substack{i\in I'\\ j\in J'}} \alpha\, \alpha_i\, \beta_j\, w_{ij} = \alpha \sum_{\substack{i\in I'\\ j\in J'}} \alpha_i\, \beta_j\, w_{ij} = \alpha\, f(u,v),$$

also ist $U \ni u \mapsto f(u,v) \in W$ für jedes $v \in V$ linear. Analog können wir zeigen, dass auch $V \ni v \mapsto f(u,v) \in W$ für jedes $u \in U$ linear ist. □

Bemerkung 22.5 (zu Satz 22.4)**.**

(i) Ist $(u_i)_{i\in I}$ eine Basis von U und $(v_j)_{j\in J}$ eine Basis von V, so ist die Familie der Vektorpaare $(u_i, v_j)_{(i,j)\in I\times J}$ i. A. keine Basis von $U \times V$! Diese Familie ist i. A. weder linear unabhängig noch ein Erzeugendensystem für $U \times V$.

Betrachten wir zum Beispiel $U = V = \mathbb{R}^3$ jeweils mit der Standardbasis (e_1, e_2, e_3). Dann gilt $\dim(U \times V) = \dim(U) + \dim(V) = 6$, also können die neun Vektoren $(e_1, e_1), \dots, (e_3, e_3)$ in $U \times V$ nicht linear unabhängig sein (Folgerung 13.22). Beispielsweise $(e_1, e_1) - (e_1, e_2) - (e_2, e_1) + (e_2, e_2)$ ist eine nichttriviale Linearkombination zum Nullvektor. Andererseits kann der Vektor $\left(\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}-1\\-1\\-1\end{pmatrix}\right) \in U \times V$ nicht erzeugt werden.

(ii) Der vorherige Punkt zeigt, dass bilineare Abbildungen $U \times V \to W$ nicht durch die Bilder auf einer Basis von $U \times V$ festgelegt werden, sondern durch die Bilder auf dem kartesischen Produkt $(u_i, v_j)_{(i,j)\in I\times J}$ der Basen von U und V. △

Die Bemerkung 22.5 zeigt, dass sich bilineare Abbildungen anders verhalten als lineare Abbildungen bzw. dass Produkträume $U \times V$ möglicherweise nicht die beste Wahl sind, um darauf bilineare Abbildungen zu definieren.

Außerdem gilt für bilineare Abbildungen

$$f(\alpha\, u, v) = \alpha\, f(u, v) = f(u, \alpha\, v)$$

für alle $u \in U$, $v \in V$ und $\alpha \in K$. Es stellt sich die Frage, ob es nicht einen Vektorraum gibt, in dem unter anderem Paare wie $(\alpha\, u, v)$ und $(u, \alpha\, v)$ dieselben Vektoren sind, also identifiziert werden. Eine bilineare Abbildung auf $U \times V$ wäre dann eine lineare Abbildung auf diesem angepassten Vektorraum. Dieser gesuchte Vektorraum nennt sich der **Tensorproduktraum** $U \otimes V$.

Ende der Vorlesung 5

Wir werden jetzt den Tensorproduktraum $U \otimes V$ konstruieren. Wir erheben dabei wegen Satz 22.4 folgende Forderung: Sind $(u_i)_{i\in I}$ bzw. $(v_j)_{j\in J}$ Basen von U bzw. von V, so soll $(u_i, v_j)_{(i,j)\in I\times J}$ gleichmächtig zu einer Basis von $U \otimes V$ sein.

Wir betrachten zur Konstruktion den K-Vektorraum $K^{I\times J} = \{T : I \times J \to K\}$ (**Quizfrage 22.2:** Warum ist das ein Vektorraum?) und dessen Unterraum (**Quizfrage 22.3:** Warum ist das ein Unterraum?) derjenigen $T : I \times J \to K$ mit endlichem Träger, also

$$U \otimes V \coloneqq \big\{T : I \times J \to K \,\big|\, T(i, j) \neq 0 \text{ nur für endlich viele } (i, j) \in I \times J\big\}. \tag{22.1}$$

Sind I und J beide endlich oder ist $I = \emptyset$ oder $J = \emptyset$, dann ist die Bedingung, dass T endlichen Träger hat, natürlich automatisch erfüllt.

Der Ausdruck $u_i \otimes v_j$ steht für dasjenige Element T von $U \otimes V$ mit der Eigenschaft

$$T(k, \ell) \coloneqq \begin{cases} 1 & \text{im Fall } k = i \text{ und } \ell = j \\ 0 & \text{sonst,} \end{cases}$$

also kurz: $u_i \otimes v_j = [(k, \ell) \mapsto \delta_{ik}\, \delta_{j\ell}]$. Wir zeigen nun, dass die Familie

$$B \coloneqq \big(u_i \otimes v_j\big)_{(i,j)\in I\times J} \tag{22.2}$$

eine Basis von $U \otimes V$ bildet. Diese ist wie gewünscht gleichmächtig zu $(u_i, v_j)_{(i,j)\in I\times J}$.

Lemma 22.6 (Basis des Tensorproduktraumes $U \otimes V$)**.**
Es seien U und V Vektorräume über dem Körper K. Sind $(u_i)_{i\in I}$ bzw. $(v_j)_{j\in J}$ Basen von U bzw. von V, dann ist $B \coloneqq \big(u_i \otimes v_j\big)_{(i,j)\in I\times J}$ eine Basis von $U \otimes V$.

Beweis. Zunächst ist B ein Erzeugendensystem, denn für beliebiges $T \in U \otimes V$ haben wir die Darstellung

$$T = \sum_{(i,j)\in E} T(i, j)\, (u_i \otimes v_j),$$

wobei E der Träger von T ist, also diejenige endliche Teilmenge $E \subseteq I \times J$, auf der $T(i, j) \neq 0$ ist.

Zu zeigen ist jetzt noch die lineare Unabhängigkeit von B, vgl. Definition 13.1. Dazu seien

$$T = \sum_{(i,j)\in E} c_{ij}\, (u_i \otimes v_j) = 0 \in U \otimes V$$

die Nullabbildung, $E \subseteq I \times J$ eine beliebige endliche Teilmenge und $c_{ij} \in K$. Es gilt $T(i,j) = c_{ij}$, falls $(i,j) \in E$ liegt, und andernfalls $T(i,j) = 0$. Da die Nullabbildung auf ganz $I \times J$ den Wert 0 hat, müssen die Koeffizienten $c_{ij} = 0$ sein. Das heißt, die Familie B ist linear unabhängig und damit eine Basis von $U \otimes V$. □

Wir fassen zusammen:

Definition 22.7 (Tensorprodukt)**.**
Es seien K ein Körper und U und V Vektorräume über K. Weiter seien $(u_i)_{i \in I}$ bzw. $(v_j)_{j \in J}$ Basen von U bzw. von V.

(i) Der K-Vektorraum

$$U \otimes V \coloneqq \left\{ T \colon I \times J \to K \,\middle|\, T(i,j) \neq 0 \text{ nur für endlich viele } (i,j) \in I \times J \right\} \tag{22.3}$$

heißt ein **Tensorproduktraum** (englisch: tensor product space) $U \otimes V$.[51]

(ii) Elemente von $U \otimes V$ heißen **Tensoren** (englisch: tensors). Der Nullvektor in $U \otimes V$, also die Nullabbildung $T \colon I \times J \to K$ mit $T(i,j) = 0$ für alle $(i,j) \in I \times J$, heißt der **Nulltensor** (englisch: zero tensor).

(iii) Die bilineare Abbildung[52]

$$\otimes \colon U \times V \ni (u,v) \mapsto \otimes(u,v) \eqqcolon u \otimes v \in U \otimes V, \tag{22.4}$$

die durch die Bilder von $(u_i, v_j)_{(i,j) \in I \times J}$ gemäß

$$\otimes(u_i, v_j) \coloneqq u_i \otimes v_j = [(k,\ell) \mapsto \delta_{ik}\, \delta_{j\ell}] \tag{22.5}$$

eindeutig definiert wird, heißt die **universelle bilineare Abbildung** (englisch: universal bilinear map) in den Tensorproduktraum $U \otimes V$.

(iv) Das Paar $(U \otimes V, \otimes)$ heißt ein **Tensorprodukt** (englisch: tensor product) von U und V.

(v) Wir nennen $u \otimes v$ auch das **Tensorprodukt** (englisch: tensor product) der Vektoren $u \in U$ und $v \in V$.

(vi) Elemente von $U \otimes V$ der Form $u \otimes v$ mit $u, v \neq 0$ heißen **Elementartensoren** (englisch: elementary tensors) oder **einfache Tensoren** (englisch: simple tensors). △

Bemerkung 22.8 (zum Tensorprodukt)**.**

(i) Die universelle bilineare Abbildung ist wegen Satz 22.4 durch die Festlegung (22.5) eindeutig bestimmt. Für das Tensorprodukt beliebiger Vektoren $u = \sum_{i \in I'} \alpha_i\, u_i \in U$ und $v = \sum_{j \in J'} \beta_j\, v_j \in V$ mit endlichen Teilmengen $I' \subseteq I$ und $J' \subseteq J$ gilt

$$\otimes(u,v) = \otimes\Big(\sum_{i \in I'} \alpha_i\, u_i, \sum_{j \in J'} \beta_j\, v_j \Big) = \sum_{i \in I'} \sum_{j \in J'} \alpha_i\, \beta_j \;\otimes(u_i, v_j),$$

was wir üblicherweise in der Form

$$u \otimes v = \sum_{i \in I'} \sum_{j \in J'} \alpha_i\, \beta_j\, (u_i \otimes v_j) \tag{22.6}$$

schreiben.

[51] sprich: „U Tensor V“
[52] $u \otimes v$ sprich: „u Tensor v“

(*ii*) „Der" Tensorproduktraum $U \otimes V$ der Vektorräume U und V ist nach Konstruktion abhängig von den gewählten Basen in U und V. Alle so konstruierten Tensorprodukträume von U und V sind jedoch isomorphe Vektorräume, da alle Basen von U und V jeweils gleichmächtig sind (Bemerkung 13.20 und Satz 18.3).[53]

(*iii*) Es gibt noch weitere isomorphe Konstruktionen von Tensorprodukträumen $U \otimes V$, beispielsweise, indem man den Vektorraum aller endlich getragenen linearen Abbildungen $T \colon U \times V \to K$ betrachtet[54] und die Summe der vier Unterräume ausfaktorisiert, die durch die Bedingungen

$$\begin{aligned} T(u+\overline{u}, v) - T(u,v) - T(\overline{u}, v) &= 0 \\ T(u, v+\overline{v}) - T(u,v) - T(u, \overline{v}) &= 0 \\ T(\alpha\, u, v) - \alpha\, T(u,v) &= 0 \\ T(u, \alpha\, v) - \alpha\, T(u,v) &= 0 \end{aligned}$$

für alle $u, \overline{u} \in U$, $v, \overline{v} \in V$ und $\alpha \in K$ gegeben sind, siehe etwa Hackbusch, 2019, Chapter 3.2.

(*iv*) Aufgrund der Vielfalt der (jeweils zueinander isomorphen) Möglichkeiten, Tensorprodukträume $U \otimes V$ zu definieren, ist es wichtig, die universelle bilineare Abbildung $\otimes \colon U \times V \to U \otimes V$, die zu einem konkreten Tensorproduktraum $U \otimes V$ gehört, mit anzugeben.

(*v*) Wir haben $\dim(U \otimes V) = \dim(U) \cdot \dim(V)$, wobei $0 \cdot \infty = \infty \cdot 0 = 0$ gilt sowie $n \cdot \infty = \infty \cdot n = \infty \cdot \infty = \infty$ im Fall $n \in \mathbb{N}$. △

Nach Konstruktion ist das Tensorprodukt $\otimes$ bilinear, d. h., es gelten die folgenden

Lemma 22.9 (Rechenregeln für das Tensorprodukt)**.**
Es seien U und V Vektorräume über dem Körper K und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann gilt:

(*i*) $(u \otimes v) + (\overline{u} \otimes v) = (u + \overline{u}) \otimes v$

(*ii*) $(u \otimes v) + (u \otimes \overline{v}) = u \otimes (v + \overline{v})$

(*iii*) $(\alpha\, u) \otimes v = \alpha\,(u \otimes v) = u \otimes (\alpha\, v)$

für alle $u, \overline{u} \in U$, $v, \overline{v} \in V$ sowie $\alpha \in K$.

Wir vereinbaren, dass das Tensorproduktsymbol $\otimes$ Vorrang vor der Addition bekommt, sodass wir einige der Klammern in Zukunft auch weglassen können.

Jeder Tensor in $T \in U \otimes V$ kann wegen Lemma 22.6 in der Form

$$T = \sum_{(i,j) \in E} T(i,j)\,(u_i \otimes v_j) = \sum_{(i,j) \in E} \underbrace{\big(T(i,j)\, u_i\big)}_{\in U} \otimes \underbrace{v_j}_{\in V}$$

mit einer endlichen Menge $E \subset I \times J$ geschrieben werden, also als Linearkombination von Elementartensoren. Beschränken wir uns dabei nicht auf die Elementartensoren $\big(u_i \otimes v_j\big)_{(i,j) \in I \times J}$ aus der

[53] Dieses Resultat verwendet im Fall unendlich-dimensionaler Vektorräume das Auswahlaxiom.
[54] auch der **freie Vektorraum** (englisch: free vector space) über der Menge $U \times V$ genannt

Basis von $U \otimes V$, sondern lassen Summen von Produkten beliebiger Vektoren zu, so können wir jedes $T \in U \otimes V$ auf i. A. viele Arten in eine Summe von Elementartensoren zerlegen:

$$T = \sum_{k=1}^{n} u_k \otimes v_k, \tag{22.7}$$

wobei die $u_k \in U$ und $v_k \in V$ beliebige Vektoren sind (nicht notwendig aus den Basen) und $n \in \mathbb{N}_0$ gilt.

Definition 22.10 (Rang eines Tensors)**.**
Es seien U und V Vektorräume über dem Körper K 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 $\operatorname{Rang}(T)$, ist die minimale Anzahl von Summanden, mit denen eine Darstellung der Form (22.7) möglich ist. △

Lemma 22.11 (Rang von Elementartensoren und des Nulltensors)**.**
Es seien U und V Vektorräume über dem Körper K und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann gilt:

(i) Der Nulltensor ist der einzige Tensor vom Rang 0.

(ii) Jeder Elementartensor, also $u \otimes v$ mit $u, v \neq 0$, ist vom Rang 1.

Beweis. Aussage (i): Der Nulltensor kann durch die leere Summe in (22.7) 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 Rank 1 haben.

Aussage (ii): Per Definition ist in (22.7) eine Darstellung mit einem Summanden möglich, also gilt $\operatorname{Rang}(u \otimes v) \leqslant 1$. Andererseits sind u und v als Linearkombinationen von Basisvektoren darstellbar, sagen wir $u = \sum_{i \in I'} \alpha_i \, u_i \in U$ und $v = \sum_{j \in J'} \beta_j \, v_j \in V$ mit endlichen Teilmengen $I' \subseteq I$ und $J' \subseteq J$. Nach (22.6) gilt die Darstellung

$$u \otimes v = \sum_{i \in I'} \sum_{j \in J'} \alpha_i \, \beta_j \, (u_i \otimes v_j).$$

Da u und v nicht die Nullvektoren sind, gibt es mindestens ein $\alpha_{i^*} \neq 0$ und mindestens ein $\beta_{j^*} \neq 0$. Damit ist $(u \otimes v)(i^*, j^*) = \alpha_{i^*} \, \beta_{j^*} \neq 0$. Also ist $u \otimes v$ nicht der Nulltensor. □

Wir klären nun, warum die universelle bilineare Abbildung $\otimes \colon U \times V \to U \otimes V$ diesen Namen trägt und wie wir bilineare Abbildungen in $\operatorname{Bil}(U, V; W)$ als lineare Abbildungen in $\operatorname{Hom}(U \otimes V, W)$ umschreiben können.

Satz 22.12 (universelle Eigenschaft der universellen bilinearen Abbildung $\otimes$)**.**
Es seien K ein Körper und U, V und W Vektorräume über K. Weiter seien $(u_i)_{i \in I}$ bzw. $(v_j)_{j \in J}$ Basen von U bzw. von V und $(U \otimes V, \otimes)$ das zugehörige Tensorprodukt. Dann gilt:

(i) Ist $g \in \operatorname{Bil}(U, V; W)$, dann gibt es ein eindeutig bestimmtes $f \in \operatorname{Hom}(U \otimes V; W)$ mit der Eigenschaft $g = f \circ \otimes$, also $g(u, v) = f(u \otimes v)$ für alle $u \in U$ und alle $v \in V$.

(ii) Ist umgekehrt $f \in \operatorname{Hom}(U \otimes V; W)$, dann ist $g \coloneqq f \circ \otimes \in \operatorname{Bil}(U, V; W)$.

Beweis. Aussage (i): Es sei $g \in \operatorname{Bil}(U, V; W)$ gegeben. Wir betrachten die Bilder $w_{ij} := g(u_i, v_j)$ für $(i, j) \in I \times J$. Da die Familie $(u_i \otimes v_j)_{(i,j)\in I\times J}$ eine Basis von $U \otimes V$ ist, gibt es genau eine lineare Abbildung $f \in \operatorname{Hom}(U \otimes V; W)$ mit der Eigenschaft $f(u_i \otimes v_j) = w_{ij}$ (Satz 17.7). Für diese Abbildung gilt

$$g(u_i, v_j) = w_{ij} = f(u_i \otimes v_j) = (f \circ \otimes)(u_i, v_j)$$

für alle $(i, j) \in I \times J$. Damit stimmen die bilinearen Abbildungen g und $f \circ \otimes$ auf der Familie $(u_i, v_j)_{(i,j)\in I\times J}$ überein. Nach Satz 22.4 sind sie damit aber bereits identisch, also gilt $g = f \circ \otimes$.

Aussage (ii): Die Komposition der universellen bilinearen Abbildung $\otimes \colon U \times V \to U \otimes V$ mit $f \in \operatorname{Hom}(U \otimes V; W)$ ergibt eine bilineare Abbildung $U \times V \to W$, denn

$$(f \circ \otimes)(u, v) = f(\otimes(u, v)) = f(u \otimes v)$$

erfüllt

$$\begin{aligned} f((u + \overline{u}) \otimes v) &= f(u \otimes v + \overline{u} \otimes v) = f(u \otimes v) + f(\overline{u} \otimes v) \\ f((\alpha\, u) \otimes v) &= f(\alpha\, (u \otimes v)) = \alpha\, f(u \otimes v) \end{aligned}$$

für $u, \overline{u} \in U$, $v \in V$ und $\alpha \in K$, ist also linear im ersten Argument. Analog kann man die Linearität auch im zweiten Argument zeigen. □

Die universelle bilineare Abbildung $\otimes$ ist also der Vermittler zwischen den bilinearen Abbildungen $U \times V \to W$ und den linearen Abbildungen $U \otimes V \to W$, und zwar unabhängig vom Zielraum W. Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} U \times V \arrow[d, "\otimes"'] \arrow[dr, "g"] & \\ U \otimes V \arrow[r, "f"'] & W \end{tikzcd}$$

Bemerkung 22.13 (abstrakte Definition von Tensorprodukten)**.**
Die in Satz 22.12 erkannte Eigenschaft ist die Grundlage für eine abstrakte Definition „des“ Tensorprodukts $U \otimes V$. Man nennt einen K-Vektorraum $U \otimes V$, zusammen mit einer bilinearen Abbildung $\otimes \colon U \times V \to U \otimes V$, ein **Tensorprodukt** $U \otimes V$ der K-Vektorräume U und V, wenn zu jeder bilinearen Abbildung $g \in \operatorname{Bil}(U, V; W)$ in einen beliebigen K-Vektorraum W stets ein eindeutig bestimmtes $f \in \operatorname{Hom}(U \otimes V; W)$ existiert mit der Eigenschaft $g = f \circ \otimes$.

Die von uns gewählte Konstruktion eines konkreten (basisabhängigen) Tensorproduktraumes $U \otimes V$ in Definition 22.7 war lediglich ein Beispiel. Alle möglichen Tensorprodukträume $U \otimes V$ über den Vektorräumen U und V sind jedoch zueinander isomorph, und die jeweiligen universellen bilinearen Abbildungen lassen sich ineinander überführen.

△

Über die Zuordnung $\operatorname{Bil}(U, V; W) \ni g \mapsto f \in \operatorname{Hom}(U \otimes V; W)$ lässt sich noch mehr sagen:

Satz 22.14 (die Zuordnung $g \mapsto f$ ist ein Vektorraumisomorphismus)**.**
Es seien U, V und W Vektorräume über dem Körper K und $(U \otimes V, \otimes)$ ein Tensorprodukt von U und V. Dann ist die Abbildung

$$\operatorname{Bil}(U, V; W) \ni g \mapsto f \in \operatorname{Hom}(U \otimes V; W) \tag{22.8}$$

(definiert durch die Eigenschaft $g = f \circ \otimes$, siehe Satz 22.12) ein Isomorphismus von Vektorräumen.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-3.4. □

Wir können daher in Zukunft statt mit bilinearen Abbildungen in $\operatorname{Bil}(U, V; W)$ äquivalent auch mit linearen Abbildungen in $\operatorname{Hom}(U \otimes V, W)$ arbeiten.

Beispiel 22.15 (Tensorprodukt)**.**

(*i*) Wir betrachten ein Tensorprodukt von K^n und K^m über einem Körper K. Dessen Dimension ist $m\,n$, wir können die Elemente von $K^n \otimes K^m$ also mit Matrizen in $K^{n \times m}$ identifizieren. Die Identifikation gelingt, indem wir die universelle bilineare Abbildung $\otimes\colon K^n \times K^m \to K^{n \times m}$ als

$$u \otimes v \coloneqq u\,v^\intercal \in K^{n \times m}$$

definieren. Dann gilt

$$e_i \otimes e_j = e_i\,e_j^\intercal = \begin{array}{cl} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{array}{l} \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \end{array} \\ \qquad\quad\ \uparrow j\text{-te Spalte} & \end{array}$$

und

$$u \otimes v = \sum_{i=1}^n \sum_{j=1}^m u_i\,v_j\,(e_i \otimes e_j) = u\,v^\intercal.$$

Für diese als Matrizen dargestellten Elementartensoren gilt offenbar $\operatorname{Rang}(u\,v^\intercal) = 1$ im Sinne des Matrixranges.

Da die $(e_i \otimes e_j)_{i=1,\ldots,n,\ j=1,\ldots,m}$ eine Basis von $K^n \otimes K^m$ darstellen, kann jeder Tensor in $K^n \otimes K^m$ eindeutig in der Form

$$T = \sum_{i=1}^n \sum_{j=1}^m a_{ij}\,(e_i \otimes e_j)$$

dargestellt werden mit einer Koeffizientenmatrix $A \in K^{n \times m}$.

Mit Hilfe einer Rangfaktorisierung (15.17) von Matrizen, also

$$A = B\,C = \sum_{k=1}^r b_{\bullet k}\,c_{k \bullet} = \sum_{k=1}^r b_{\bullet k} \otimes (c_{k \bullet})^\intercal$$

mit $r = \operatorname{Rang}(A)$, kann man zeigen, dass $\operatorname{Rang}(T) = \operatorname{Rang}(A)$ gilt. Der Rang eines Tensors in $K^n \otimes K^m$ nach Definition 22.10 ist also gleich dem Rang der Koeffzientenmatrix.

(*ii*) Es sei K ein Körper und $g\colon K[s] \times K[t] \to K[s,t]$ die bilineare Abbildung „Multiplikation von Polynomen in verschiedenen Variablen“, siehe Beispiel 22.2. Wir wählen die Basen $(s^i)_{i \in \mathbb{N}_0}$ und $(t^j)_{j \in \mathbb{N}_0}$. Dann ist $\big(s^i \otimes t^j\big)_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0}$ eine Basis von $U \otimes V$. Nach Satz 22.12 gibt es genau eine lineare Abbildung $f\colon K[s] \otimes K[t] \to K[s,t]$ mit der Eigenschaft $g = f \circ \otimes$, d. h.,

$$p \cdot q = g(p,q) = f(p \otimes q)$$

für alle $p \in K[s]$ und $q \in K[t]$, insbesondere

$$s^i \cdot t^j = f(s^i \otimes t^j)$$

für alle $i, j \in \mathbb{N}_0$. Da die Familie $(s^i \cdot t^j)_{(i,j)\in\mathbb{N}_0\times\mathbb{N}_0}$ eine Basis von $K[s,t]$ bildet, ist diese Abbildung $f\colon K[s] \otimes K[t] \to K[s,t]$ nach Satz 17.7 sogar ein Isomorphismus. Das bietet eine Möglichkeit, den Polynomring in zwei Veränderlichen[55] aus dem in einer Veränderlichen zu konstruieren.

(*iii*) Es sei W ein beliebiger Vektorraum über $\mathbb{R}$ mit der Basis $(w_j)_{j\in J}$. Weiter sei $V = \mathbb{C}$, aufgefasst als zweidimensionaler Vektorraum über $\mathbb{R}$ mit der Basis $(1, i)$. Dann bildet die Familie $(1 \otimes w_j)_{j\in J}$ zusammen mit $(i \otimes w_j)_{j\in J}$ eine Basis von $\mathbb{C} \otimes W$. Jedes Element von $\mathbb{C} \otimes W$ hat eine (bis auf Nullkoeffizienten) eindeutige Darstellung

$$\widehat{w} = \sum_{j\in J'} \alpha_j\,(1 \otimes w_j) + \sum_{j\in J''} \beta_j\,(i \otimes w_j) = \sum_{j\in J'\cup J''} (\alpha_j + \beta_j\, i)\,(1 \otimes w_j) =: \sum_{j\in J'\cup J''} (\mu_j \otimes w_j)$$

mit endlichen Indexmengen J' und J'' sowie Koeffizienten $\alpha_j, \beta_j \in \mathbb{R}$ bzw. $\mu_j \coloneqq \alpha_j + \beta_j\, i \in \mathbb{C}$. Wir können nun im $\mathbb{R}$-Vektorraum $\mathbb{C} \otimes W$ die skalare Multiplikation von $\mathbb{R}$ auf $\mathbb{C}$ erweitern, indem wir für $\widehat{w}$ mit obiger Darstellung und $\mu \in \mathbb{C}$ definieren:

$$\mu \cdot \widehat{w} \coloneqq \sum_{j\in J'\cup J''} (\mu \cdot \mu_j) \otimes w_j.$$

Wir können mit dieser Definition nun nachprüfen, dass $(\mathbb{C} \otimes W)$ nicht nur ein Vektorraum über $\mathbb{R}$, sondern auch ein Vektorraum über $\mathbb{C}$ ist und als solcher die Basis $(1 \otimes w_j)_{j\in J}$ besitzt. Dieser $\mathbb{C}$-Vektorraum $(\mathbb{C} \otimes W)$ heißt die **Komplexifizierung** (englisch: complexification) des $\mathbb{R}$-Vektorraumes W. Der ursprüngliche Raum W wird in der isomorphen Gestalt $1 \otimes W$ zu einem $\mathbb{R}$-Unterraum von $\mathbb{C} \otimes W$. △

Wir geben noch eine weitere Möglichkeit an, lineare Abbildungen auf einem Tensorproduktraum zu interpretieren:

Lemma 22.16 (Interpretation bilinearer Abbildungen)**.**
Es seien U, V und W endlich-dimensionale Vektorräume über einem Körper K. Dann bestehen Isomorphismen

$$\operatorname{Hom}(U \otimes V; W) \cong \operatorname{Bil}(U, V; W) \cong \operatorname{Hom}(U, \operatorname{Hom}(V, W)) \cong \operatorname{Hom}(V, \operatorname{Hom}(U, W)). \tag{22.9}$$

An Stelle eines Beweises erläutern wir die Isomorphieaussagen: Ist $f\colon U \times V \to W$ eine bilineare Abbildung und setzen wir das erste Argument $u \in U$ ein, so ist die verbleibende Abbildung

$$f_u\colon V \ni v \mapsto f(u,v) \in W$$

linear, also ein Element von $\operatorname{Hom}(V, W)$. Weiterhin ist die Abbildung $U \ni u \mapsto f_u \in \operatorname{Hom}(V, W)$ ebenfalls linear. Das heißt, Elemente von $\operatorname{Bil}(U, V; W)$ können — indem man zuerst das Element von U einsetzt — interpretiert werden als Elemente von $\operatorname{Hom}(U, \operatorname{Hom}(V, W))$. Schließlich ist diese Interpretation sogar ein Vektorraumisomorphismus. Alternativ können wir auch einen Isomorphismus mit $\operatorname{Hom}(V, \operatorname{Hom}(U, W))$ etablieren, indem wir zuerst das Element von V einsetzen.

Folgerung 22.17 (Interpretationen von Bilinearformen)**.**
Insbesondere im Fall $W = K$ ergibt sich aus Lemma 22.16

$$\operatorname{Bil}(U, V; K) \cong \operatorname{Hom}(U \otimes V; K) \cong (U \otimes V)^* \cong \operatorname{Hom}(U, V^*) \cong \operatorname{Hom}(V, U^*).$$

Ende der Vorlesung 6

Ende der Woche 3

[55] allerdings hier nur den Polynomring über einem Körper und nicht über einem kommutativen Ring

§ 22.2 Multilineare Abbildungen

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.

Definition 22.18 (multilineare Abbildung, Multilinearform).
Es seien $V_1, \ldots, V_N, W$ Vektorräume über demselben Körper K für $N \in \mathbb{N}_0$.

(i) Eine Abbildung
$$f \colon V_1 \times \cdots \times V_N \to W$$
heißt **multilinear** (englisch: multilinear map) oder genauer **N-linear** (englisch: N-linear map), wenn für jedes $i \in [\![1, N]\!]$ und alle fest gewählten $\overline{v}_j \in V_j$, $j \in [\![1, N]\!] \setminus \{i\}$, die Abbildung
$$f(\overline{v}_1, \ldots, \overline{v}_{i-1}, \cdot, \overline{v}_{i+1}, \ldots, \overline{v}_N) \colon V_i \ni v_i \mapsto f(\overline{v}_1, \ldots, \overline{v}_{i-1}, v_i, \overline{v}_{i+1}, \ldots, \overline{v}_N) \in W$$
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 $\operatorname{Mult}(V_1, \ldots, V_N)$ oder $\operatorname{Mult}(V_1, \ldots, V_N; K)$. △

Bemerkung 22.19 (multilineare Abbildungen).

(i) Im Fall $N = 0$ ist $V_1 \times \cdots \times V_N$ die Menge, die nur aus dem leeren Tupel () besteht. Der Vektorraum $V_1 \times \cdots \times V_N$ ist also isomorph zum Nullraum. Damit ist jede Abbildung $\{()\} \to W$ multilinear und kann mit dem jeweils einzigen Bildelement in W identifiziert werden.

(ii) Im Fall $N = 1$ sind die multilinearen Abbildungen $V_1 \to W$ gerade die linearen Abbildungen, also $\operatorname{Mult}(V_1; W) = \operatorname{Hom}(V_1, W)$.

(iii) Der Fall $N = 2$ entspricht den bilinearen Abbildungen $V_1 \times V_2 \to W$ aus § 22.1, also $\operatorname{Mult}(V_1, V_2; W) = \operatorname{Bil}(V_1, V_2; W)$. △

Es gelten die mehrdimensionalen Analoga von Lemma 22.3 und Satz 22.4:

Lemma 22.20 (multilineare Abbildungen bilden einen Vektorraum).
Es seien $V_1, \ldots, V_N, W$ Vektorräume über demselben Körper K für $N \in \mathbb{N}_0$. 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 22.21 (Existenz- und Eindeutigkeitssatz für multilineare Abbildungen, vgl. Satz 22.4).
Es seien $V_1, \ldots, V_N, W$ Vektorräume über demselben Körper K für $N \in \mathbb{N}_0$. 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 multilineare Abbildung $f \colon V_1 \times \cdots \times V_N \to W$ mit der Eigenschaft $f(v_{i_1}, \ldots, v_{i_N}) = w_{i_1,\ldots,i_N}$ für alle $(i_1, \ldots, i_N) \in I_1 \times \cdots \times I_N$.

In Erweiterung von Definition 22.7 können wir angeben:

Definition 22.22 (Tensorprodukt).
Es seien K ein Körper und $V_1, \ldots, V_N$ Vektorräume über K und $N \in \mathbb{N}_0$. Weiter seien die Familien $(v_{1,i_1})_{i_1 \in I_1}, \ldots, (v_{N,i_N})_{i_N \in I_N}$ Basen von $V_1, \ldots, V_N$.

(*i*) Der K-Vektorraum

$$V_1 \otimes \cdots \otimes V_N = \bigotimes_{i=1}^{N} V_i :=$$
$$\big\{T \colon I_1 \times \cdots \times I_N \to K \,\big|\, T(i_1, \ldots, i_N) \neq 0 \text{ nur für endlich viele } (i_1, \ldots, i_N) \in I_1 \times \cdots \times I_N\big\} \tag{22.10}$$

heißt ein **Tensorproduktraum** $V_1 \otimes \cdots \otimes V_N$.

(*ii*) Elemente von $V_1 \otimes \cdots \otimes V_N$ heißen **Tensoren** der **Stufe** (englisch: degree) oder **Ordnung** (englisch: order) N. Der Nullvektor in $V_1 \otimes \cdots \otimes V_N$, also die Nullabbildung $T \colon I_1 \times \cdots \times I_N \to K$ mit $T(i_1, \ldots, i_N) = 0$ für alle $(i_1, \ldots, i_N) \in I_1 \times \cdots \times I_N$, heißt der **Nulltensor**.

(*iii*) Die multilineare Abbildung

$$\otimes \colon V_1 \times \cdots \times V_N \ni (v_1, \ldots, v_N) \mapsto \otimes(v_1, \ldots, v_N) =: v_1 \otimes \cdots \otimes v_N \in V_1 \otimes \cdots \otimes V_N, \tag{22.11}$$

die durch die Bilder von $(v_{1,i_1}, \ldots, v_{N,i_N})_{(i_1,\ldots,i_N) \in I_1 \times \cdots \times I_N}$ gemäß

$$\otimes(v_{1,i_1}, \ldots, v_{N,i_N}) := v_{1,i_1} \otimes \cdots \otimes v_{N,i_N} = \big[(k_1, \ldots, k_N) \mapsto \delta_{i_1 k_1} \cdots \delta_{i_N k_N}\big] \tag{22.12}$$

eindeutig definiert wird, heißt die **universelle multilineare Abbildung** (englisch: universal multilinear map) in den Tensorproduktraum $V_1 \otimes \cdots \otimes V_N$. Dabei bezeichnet $v_{1,i_1} \otimes \cdots \otimes v_{N,i_N}$ dasjenige Element $T \in V_1 \otimes \cdots \otimes V_N$ mit der Eigenschaft

$$T(k_1, \ldots, k_N) := \begin{cases} 1 & \text{im Fall } k_j = i_j \text{ für alle } j = 1, \ldots, N \\ 0 & \text{sonst,} \end{cases}$$

also $v_{1,i_1} \otimes \cdots \otimes v_{N,i_N} = \big[(k_1, \ldots, k_N) \mapsto \delta_{i_1 k_1} \cdots \delta_{i_N k_N}\big]$.

(*iv*) Das Paar $(V_1 \otimes \cdots \otimes V_N, \otimes)$ heißt ein **Tensorprodukt** von $V_1, \ldots, V_N$.

(*v*) Wir nennen $v_1 \otimes \cdots \otimes v_N$ auch das **Tensorprodukt** der Vektoren $v_1 \in V_1$ usw. bis $v_N \in V_N$. Statt $v_1 \otimes \cdots \otimes v_N$ schreiben wir manchmal auch $\bigotimes_{j=1}^{N} v_j$.

(*vi*) Elemente von $V_1 \otimes \cdots \otimes V_N$ der Form $v_1 \otimes \cdots \otimes v_N$ mit $v_1, \ldots, v_N \neq 0$ heißen **Elementartensoren** oder **einfache Tensoren** (englisch: simple tensors). △

Auch hier ist wieder die Konstruktion des Tensorproduktraumes $V_1 \otimes \cdots \otimes V_N$ abhängig von der Wahl der Basen, aber alle Instanzen des Tensorproduktraumes sind zueinander isomorph. Es gelten Rechenregeln analog zu Lemma 22.9, d. h., das Tensorprodukt ist nach Konstruktion multilinear. Auch den **Rang** eines Tensors T können wir analog zu Definition 22.10 definieren als die minimale Anzahl von Summanden, mit denen wir T als Summe von Elementartensoren schreiben können.[56] In Analogie zu Lemma 22.11 gilt wieder, dass der Nulltensor der einzige Tensor vom Rang 0 ist und dass Elementartensoren genau diejenigen vom Rang 1 sind.

Bezüglich der universellen bilinearen Abbildung können wir folgende Aussage zeigen:

[56] Im Fall $N \geqslant 3$ ist die Bestimmung des Ranges eines N-stufigen Tensors ein NP-schweres Problem im Sinne der Komplexitätstheorie. Dagegen haben für $N = 0$ und $N = 1$ alle Tensoren bis auf den Nulltensor den Rang 1.

Satz 22.23 (universelle Eigenschaft der universellen multilinearen Abbildung $\otimes$)**.**
Es seien K ein Körper und $V_1, \ldots, V_N$ und W Vektorräume über K und $N \in \mathbb{N}_0$. Weiter seien die Familien $(v_{1,i_1})_{i_1 \in I_1}, \ldots, (v_{N,i_N})_{i_N \in I_N}$ Basen von $V_1, \ldots, V_N$ und $(V_1 \otimes \cdots \otimes V_N, \otimes)$ das zugehörige Tensorprodukt. Dann gilt:

(i) Ist $g \in \operatorname{Mult}(V_1 \times \cdots \times V_N; W)$, dann gibt es ein eindeutig bestimmtes $f \in \operatorname{Hom}(V_1 \otimes \cdots \otimes V_N; W)$ mit der Eigenschaft $g = f \circ \otimes$.

(ii) Ist $f \in \operatorname{Hom}(V_1 \otimes \cdots \otimes V_N; W)$, dann ist $g := f \circ \otimes \in \operatorname{Mult}(V_1 \times \cdots \times V_N; W)$.

Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd}
V_1 \times \cdots \times V_N \arrow[d, "\otimes"'] \arrow[dr, "g"] & \\
V_1 \otimes \cdots \otimes V_N \arrow[r, "f"'] & W
\end{tikzcd}$$

Auch hier gilt analog zu Satz 22.14:

Satz 22.24 (die Zuordnung $g \mapsto f$ ist ein Vektorraumisomorphismus)**.**
Es seien $V_1, \ldots, V_N$ sowie W Vektorräume über dem Körper K, $N \in \mathbb{N}_0$, und $(V_1 \otimes \cdots \otimes V_N, \otimes)$ ein Tensorprodukt von $V_1, \ldots, V_N$. Dann ist die Abbildung

$$\operatorname{Mult}(V_1 \times \cdots \times V_N; W) \ni g \mapsto f \in \operatorname{Hom}(V_1 \otimes \cdots \otimes V_N; W) \tag{22.13}$$

(definiert durch die Eigenschaft $g = f \circ \otimes$, siehe Satz 22.23) ein Isomorphismus von Vektorräumen.

Wir können daher in Zukunft statt mit multilinearen Abbildungen in $\operatorname{Mult}(V_1 \times \cdots \times V_N; W)$ auch mit linearen Abbildungen in $\operatorname{Hom}(V_1 \otimes \cdots \otimes V_N, W)$ arbeiten.

§ 22.3 Tensoren über einem Vektorraum

In diesem Abschnitt betrachten wir eine für die Praxis besonders relevante Klasse von Tensoren noch genauer.

Definition 22.25 (Tensoren über einem Vektorraum)**.**
Es sei V ein Vektorraum über einem Körper K und V^* sein Dualraum. Weiter seien $(r, s) \in \mathbb{N}_0 \times \mathbb{N}_0$. Elemente des Tensorproduktraumes[57]

$$\mathcal{T}_s^r(V) := \underbrace{V \otimes \cdots \otimes V}_{r\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s\text{-mal}} \tag{22.14}$$

heißen **Tensoren vom Typ** (r, s) (englisch: tensors of type (r, s)) über dem Vektorraum V.[58] △

[57]Die Notation für diese Tensorprodukträume ist in der Literatur leider nicht einheitlich. Oft ist die Bedeutung von r und s vertauscht.
[58]Historisch spricht man auch von r-fach **kontravarianten** und s-fach **kovarianten** Tensoren, vgl. Bemerkung 21.18.

Wir stellen uns folgende Fragen:[59]

(1) Wie lassen sich solche Tensoren darstellen?

(2) Wie lassen sich solche Tensoren interpretieren bzw. verwenden?

Bemerkung 22.26 (Positionen von Indizes)**.**
Wir betrachten in diesem Abschnitt überwiegend endlich-dimensionale Vektorräume V mit $\dim(V) = n \in \mathbb{N}_0$. Außerdem gehen wir zu einer für die Tensoralgebra typischen Notation über:

- Die Vektoren einer Basis des „primalen" Raumes V werden mit unteren Indizes nummeriert und heißen typischerweise $e_1, \ldots, e_n$.
- Die Koordinaten von Vektoren in V bzgl. dieser Basis werden mit oberen Indizes notiert, also $v = \sum_{i=1}^n v^i\, e_i$.
- Die Covektoren der dualen Basis von V^* werden mit oberen Indizes nummeriert und heißen typischerweise $\varepsilon^1, \ldots, \varepsilon^n$. Es gilt also $\langle \varepsilon^i, e_j \rangle = \delta_{ij}$.
- Die Koordinaten von Covektoren in V^* bzgl. dieser Basis werden mit unteren Indizes notiert, also $\omega = \sum_{i=1}^n \omega_i\, \varepsilon^i$. △

Wir kommen zur Frage (1). Jeder Tensor in $\mathcal{T}_s^r(V)$ ist eine Linearkombination der elementaren Basistensoren

$$e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_s} \quad \text{mit } 1 \leqslant i_1, \ldots, i_r \leqslant n \text{ und } 1 \leqslant j_1, \ldots, j_s \leqslant n. \tag{22.15}$$

Also kann jeder Tensor in $T \in \mathcal{T}_s^r(V)$ darstellt werden in der Form

$$T = \sum_{i_1=1}^n \cdots \sum_{i_r=1}^n \sum_{j_1=1}^n \cdots \sum_{j_s=1}^n T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}\, e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_s}. \tag{22.16}$$

Die Koeffizienten $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$ werden die **Komponenten** (englisch: components) des Tensors T bzgl. der Basis (22.15) genannt.[60] Sie bilden eine Art „verallgemeinerte Matrix", nämlich ein Zahlenschema mit $r + s$ Dimensionen und daher n^{r+s} Einträgen. In der mathematischen Physik ist es üblich, die Summenzeichen in (22.16) nicht zu notieren, sondern sie sich implizit dazuzudenken. Das nennt sich die **Einsteinsche Summenkonvention** (englisch: Einstein summation convention): „Über jeden Index, der in einer Formel zweifach (und zwar notwendigerweise einmal oben und einmal unten) auftritt, wird summiert."

Die Zuordnung

$$\mathcal{T}_s^r(V) \ni T \mapsto T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} \in K^{n^{r+s}}$$

eines Tensors auf seine Komponenten (bzgl. einer beliebigen, aber festen Basis von V und der zugehörigen dualen Basis) ist ein Vektorraumisomorphismus.

[59] Alle Konzepte in § 22.3 (Darstellung von Tensoren mit Hilfe von Komponenten und Interpretation von Tensoren als multilineare Abbildungen) hätten wir auch bereits für Tensoren in allgemeineren Tensorprodukträumen $V_1 \otimes \cdots \otimes V_N$ einführen können. Wir beschränken uns jedoch aus Gründen der Materialauswahl auf die Tensorprodukträume hier, die nur V und V^* verwenden.

[60] Die Koordinaten von Vektoren v hatten wir stets mit einem anderen Symbol bezeichnet, typischerweise x. Bei Tensoren ist es aber üblich, die Komponenten mit demselben Symbol zu notieren, hier beides Mal T.

Bemerkung 22.27 (Tensoren und ihre Komponenten).
Oft hört man die Aussage, das Zahlenschema der Komponenten $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$ *sei* der eigentliche Tensor T.[61] Dieser Sichtweise schließen wir uns hier nicht an. Jedoch kann man mit einem Tensor „arbeiten" (z. B. die Stufe oder den Rang bestimmen, Vektoren und Covektoren in die durch den Tensor repräsentierte multilineare Abbildung einsetzen etc.), indem man die entsprechenden Aktionen am Schema der Komponenten ausführt. Dazu folgen unten noch Beispiele. △

Wir kommen zu Frage (2). Bisher kennen wir Tensoren nur als Elemente eines Vektorraumes, der als natürlicher Definitionsraum multilinearer Abbildungen angesehen werden kann, sodass diese zu linearen Abbildungen werden (Satz 22.23). Interessanterweise können jedoch auch die Tensoren selbst bereits als multilineare Abbildungen verstanden werden, und zwar sogar auf verschiedene Art und Weise, je nachdem, welche Argumente man einsetzt und welche Argumente man noch zurückhält:

Lemma 22.28 (Interpretation von Tensoren vom Typ (r,s), vgl. Lemma 22.16).
Es sei V ein endlich-dimensionaler Vektorraum über einem Körper K und V^* sein Dualraum. Weiter seien $(r,s) \in \mathbb{N}_0 \times \mathbb{N}_0$ und $\mathcal{T}_s^r(V)$ ein Tensorproduktraum vom Typ (r,s). Dann bestehen folgende Isomorphismen:

$$\text{Typ } (0,0) \quad \mathcal{T}_0^0(V) = K \quad \cancel{\cong \operatorname{Hom}(\{0\},K)} \tag{22.17a}$$
$$\text{Typ } (0,1) \quad \mathcal{T}_1^0(V) = V^* \quad \cong \operatorname{Hom}(V,K) \tag{22.17b}$$
$$\text{Typ } (1,0) \quad \mathcal{T}_0^1(V) = V \quad \cong \operatorname{Hom}(V^*,K) \tag{22.17c}$$
$$\text{Typ } (0,2) \quad \mathcal{T}_2^0(V) = V^* \otimes V^* \cong \operatorname{Hom}(V \otimes V,K) \cong \operatorname{Hom}(V,V^*) \tag{22.17d}$$
$$\text{Typ } (1,1) \quad \mathcal{T}_1^1(V) = V \otimes V^* \cong \operatorname{Hom}(V^* \otimes V,K) \cong \operatorname{Hom}(V,V) \cong \operatorname{Hom}(V^*,V^*) \tag{22.17e}$$
$$\text{Typ } (2,0) \quad \mathcal{T}_0^2(V) = V \otimes V \cong \operatorname{Hom}(V^* \otimes V^*,K) \cong \operatorname{Hom}(V^*,V) \tag{22.17f}$$

usw.

An Stelle eines Beweises erläutern wir die Aussagen. Die Isomorphismen in (22.17a)–(22.17c) sind offensichtlich. Im Fall (22.17d) ist ein Isomorphismus zwischen $\operatorname{Hom}(V \otimes V, K)$ und $\operatorname{Hom}(V, V^*)$ für Basistensoren gegeben durch

$$(\varepsilon^{j_1} \otimes \varepsilon^{j_2})(v,\cdot) \coloneqq \langle \varepsilon^{j_1}, v\rangle\, \varepsilon^{j_2} \quad \text{für } \varepsilon^{j_1}, \varepsilon^{j_2} \in V^* \text{ und } v \in V.$$

Ein weiterer Isomorphismus ist gegeben durch

$$(\varepsilon^{j_1} \otimes \varepsilon^{j_2})(\cdot,v) \coloneqq \langle \varepsilon^{j_2}, v\rangle\, \varepsilon^{j_1} \quad \text{für } \varepsilon^{j_1}, \varepsilon^{j_2} \in V^* \text{ und } v \in V.$$

Im Fall (22.17e) ist ein Isomorphismus zwischen $\operatorname{Hom}(V^* \otimes V, K)$ und $\operatorname{Hom}(V, V)$ für Basistensoren gegeben durch

$$(e_{i_1} \otimes \varepsilon^{j_1})(\cdot,v) \coloneqq \langle \varepsilon^{j_1}, v\rangle\, e_{i_1} \quad \text{für } e_{i_1}, v \in V \text{ und } \varepsilon^{j_1} \in V^*.$$

Ein Isomorphismus zwischen $\operatorname{Hom}(V^* \otimes V, K)$ und $\operatorname{Hom}(V^*, V^*)$ ist für Basistensoren gegeben durch

$$(e_{i_1} \otimes \varepsilon^{j_1})(v^*,\cdot) \coloneqq \langle v^*, e_{i_1}\rangle\, \varepsilon^{j_1} \quad \text{für } e_{i_1} \in V \text{ und } \varepsilon^{j_1}, v^* \in V^*.$$

Im Fall (22.17f) ist ein Isomorphismus zwischen $V \otimes V$ und $\operatorname{Hom}(V^*, V)$ für Basistensoren gegeben durch

$$(e_{i_1} \otimes e_{i_2})(v^*,\cdot) \coloneqq \langle v^*, e_{i_1}\rangle\, e_{i_2} \quad \text{für } e_{i_1}, e_{i_2} \in V \text{ und } v^* \in V^*.$$

[61] Diese Aussage steht auf einer Stufe mit „Der Koordinatenvektor eines Vektors *sei* der Vektor." oder „Die Darstellungsmatrix einer linearen Abbildung *sei* die Abbildung".

Ein weiterer Isomorphismus ist gegeben durch

$$(e_{i_1} \otimes e_{i_2})(\cdot, v^*) := \langle v^*, e_{i_2} \rangle\, e_{i_1} \quad \text{für } e_{i_1}, e_{i_2} \in V \text{ und } v^* \in V^*.$$

Für alle Nicht-Basistensoren des jeweiligen Tensorproduktraumes ergeben sich die Formeln durch Linearkombination der Formeln für Basistensoren.

Daraus erkennen wir: Jeder Faktor in einem Tensorprodukt „konsumiert" ein Element des jeweiligen Dualraumes[62] und liefert — sobald das entsprechende Argument eingesetzt wird — das Ergebnis der dualen Paarung als Element des Skalarkörpers K (eine „Zahl") zurück. Die verbleibenden Faktoren, für die kein Argument eingesetzt wird, bilden einen Tensor niedrigerer Stufe. So würde beispielsweise ein Tensor T aus $\mathcal{T}_1^3(V)$ nach Einsetzen des ersten und dritten Arguments zu einem Tensor S in $\mathcal{T}_1^1(V)$. Am Beispiel eines Basistensors:

$$(e_{i_1} \otimes e_{i_2} \otimes e_{i_3} \otimes \varepsilon^{j_1})(\omega, \cdot, \xi, \cdot) := \langle \omega, e_{i_1} \rangle \langle \xi, e_{i_3} \rangle\, e_{i_2} \otimes \varepsilon^{j_1}.$$

In Komponenten bzgl. irgendeiner Basis von V und der zugehörigen dualen Basis von V^* sieht diese Rechnung für einen beliebigen Tensor $T \in \mathcal{T}_1^3(V)$ wie folgt aus:

$$\sum_{i_1=1}^n \sum_{i_3=1}^n T^{i_1,i_2,i_3}_{j_1}\, \omega_{i_1}\, \xi_{i_3} =: S^{i_2}_{j_1}.$$

Wir werden in Zukunft einen Tensor

$$T \in \mathcal{T}_s^r(V)$$

nach Bedarf mit einem Element in $\operatorname{Hom}(\underbrace{V^* \otimes \cdots \otimes V^*}_{r\text{-mal}} \otimes \underbrace{V \otimes \cdots \otimes V}_{s\text{-mal}}, K)$ bzw. mit der multilinearen Abbildung

$$\underbrace{V^* \times \cdots \times V^*}_{r\text{-mal}} \times \underbrace{V \times \cdots \times V}_{s\text{-mal}} \to K$$

nach Satz 22.23 identifizieren und letztere ebenfalls mit T bezeichnen. Wir schreiben daher einfach auch $T(\omega^1, \ldots, \omega^r, v_1, \ldots, v_s)$ für das Einsetzen aller Argumente in die zum Tensor T gehörige Multilinearform. Hat ω^i die Koordinaten $\begin{pmatrix} \omega_{i,1} \\ \vdots \\ \omega_{i,n} \end{pmatrix}$ bzgl. der Basis $\varepsilon^1, \ldots, \varepsilon^n$ und hat v_j die Koordinaten $\begin{pmatrix} v^{j,1} \\ \vdots \\ v^{j,n} \end{pmatrix}$ bzgl. der Basis $e_1, \ldots, e_n$ und hat T die Komponenten $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$ bzgl. der Basis (22.15), dann gilt

$$T(\omega^1, \ldots, \omega^r, v_1, \ldots, v_s) = \sum_{i_1=1}^n \cdots \sum_{i_r=1}^n \sum_{j_1=1}^n \cdots \sum_{j_s=1}^n T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}\, \omega_{1,i_1} \cdots \omega_{r,i_r}\, v^{1,j_1} \cdots v^{s,j_s}.$$

(**Quizfrage 22.4:** Warum ist das so?)

Ende der Vorlesung 7

[62] wobei der Dualraum V^{**} von V^* mit V identifiziert wird (Satz 21.42)

§ 22.4 Symmetrische und schiefsymmetrische Tensoren

Die Notation in der nachfolgenden Definition wurde nachträglich angepasst, um Doppelindizes zu vermeiden.

Definition 22.29 (symmetrische, schiefsymmetrische und alternierende Tensoren)**.**
Es sei V ein Vektorraum über einem Körper K. Weiter sei $r \in \mathbb{N}_0$.

(i) Ein Tensor $T \in \mathcal{T}_0^r(V) \cong \operatorname{Hom}(V^* \otimes \cdots \otimes V^*, K) \cong \operatorname{Mult}(V^* \times \cdots \times V^*; K)$ heißt **(total) symmetrisch** (englisch: (totally) symmetric tensor), wenn für alle $\omega^1, \ldots, \omega^r \in V^*$ und jede Permutation $\sigma \in S_r$ gilt:

$$T(\omega^1, \ldots, \omega^r) = T(\omega^{\sigma(1)}, \ldots, \omega^{\sigma(r)}). \tag{22.18}$$

Die Menge aller symmetrischen Tensoren in $\mathcal{T}_0^r(V)$ bezeichnen wir mit $\mathcal{T}_0^r(V)_{\text{sym}}$.

(ii) Ein Tensor $T \in \mathcal{T}_0^r(V) \cong \operatorname{Hom}(V^* \otimes \cdots \otimes V^*, K) \cong \operatorname{Mult}(V^* \times \cdots \times V^*; K)$ heißt **(total) schiefsymmetrisch** (englisch: (totally) skew-symmetric tensor), wenn für alle $\omega^1, \ldots, \omega^r \in V^*$ und jede Permutation $\sigma \in S_r$ gilt:

$$T(\omega^1, \ldots, \omega^r) = (\operatorname{sgn} \sigma)\, T(\omega^{\sigma(1)}, \ldots, \omega^{\sigma(r)}). \tag{22.19}$$

Die Menge aller schiefsymmetrischen Tensoren in $\mathcal{T}_0^r(V)$ bezeichnen wir mit $\mathcal{T}_0^r(V)_{\text{skew}}$.

(iii) Ein Tensor $T \in \mathcal{T}_0^r(V) \cong \operatorname{Hom}(V^* \otimes \cdots \otimes V^*, K) \cong \operatorname{Mult}(V^* \times \cdots \times V^*; K)$ heißt **alternierend** (englisch: alternating tensor), wenn für $\omega^1, \ldots, \omega^r \in V^*$ gilt:

$$\omega_i = \omega_k \text{ für ein } k \neq \ell \quad \Rightarrow \quad T(\omega^1, \ldots, \omega^r) = 0. \tag{22.20}$$

Die Menge aller alternierenden Tensoren in $\mathcal{T}_0^r(V)$ bezeichnen wir mit $\mathcal{T}_0^r(V)_{\text{alt}}$. △

Die Symmetrie von $T \in \mathcal{T}_0^r(V)$ bedeutet also gerade, dass die Reihenfolge der r Argumente unerheblich ist. Die Schiefsymmetrie von $T \in \mathcal{T}_0^r(V)$ bedeutet dagegen, dass $T(\varepsilon^{i_1}, \ldots, \varepsilon^{i_r})$ das Vorzeichen wechselt, wenn genau zwei Argumente vertauscht werden. Die alternierende Eigenschaft bedeutet schließlich, dass das Bild von T gleich 0 ist, sobald zwei Argumente identisch sind.

Die Eigenschaften **symmetrisch**, **schiefsymmetrisch** und **alternierend** verwenden wir analog zu Definition 22.29 auch für Multilinearformen in $\operatorname{Mult}(V^* \times \cdots \times V^*; K)$.

Lemma 22.30 (Zusammenhang zwischen schiefsymmetrischen und alternierenden Tensoren)**.**
Es sei V ein Vektorraum über einem Körper K. Weiter sei $r \in \mathbb{N}_0$.

(i) Ist $T \in \mathcal{T}_0^r(V)$ alternierend, dann ist T auch schiefsymmetrisch.

(ii) Im Fall $\operatorname{char}(K) \neq 2$ gilt auch die Umkehrung.

(iii) Im Fall $\operatorname{char}(K) = 2$ gilt $\mathcal{T}_0^r(V)_{\text{sym}} = \mathcal{T}_0^r(V)_{\text{skew}}$.

(iv) Im Fall $r = 0$ und $r = 1$ gilt $\mathcal{T}_0^r(V)_{\text{sym}} = \mathcal{T}_0^r(V)_{\text{skew}} = \mathcal{T}_0^r(V)_{\text{alt}} = \mathcal{T}_0^r(V)$.

Beweis. Aussage (iv): Die einzige Permutation in S_0 ist die leere Permutation, und die einzige Permutation in S_1 ist die Identität. Daher sind (22.18) und (22.19) für jede Permutation und jeden Tensor erfüllt. Außerdem ist im Fall $r = 0$ und $r = 1$ jeder Tensor alternierend, da es nicht möglich ist, zwei verschiedene identische Argumente einzusetzen.

Aussage (i): Im Fall $r = 0$ und $r = 1$ ist wie in Aussage (iv) gezeigt, jeder Tensor alternierend und schiefsymmetrisch.

Es sei nun $r \geqslant 2$. Für jeden Tensor $T \in \mathcal{T}_0^r(V) \cong \operatorname{Mult}(V^* \times \cdots \times V^*; K)$ gilt:

$$\begin{aligned} &T(\omega_1 + \overline{\omega}_1, \omega_2 + \overline{\omega}_2, \ldots, \omega_r) \\ &\qquad = T(\omega_1, \omega_2, \ldots, \omega_r) + T(\overline{\omega}_1, \omega_2, \ldots, \omega_r) + T(\omega_1, \overline{\omega}_2, \ldots, \omega_r) + T(\overline{\omega}_1, \overline{\omega}_2, \ldots, \omega_r). \end{aligned}$$

Ist $T \in \mathcal{T}_0^r(V)$ alternierend und gilt $\omega_1 = \omega_2$ und $\overline{\omega}_1 = \overline{\omega}_2$, also auch $\omega_1 + \overline{\omega}_1 = \omega_2 + \overline{\omega}_2$, so folgt

$$0 = \underbrace{T(\omega_1, \omega_2, \ldots, \omega_r)}_{=0} + T(\overline{\omega}_1, \omega_2, \ldots, \omega_r) + T(\omega_1, \overline{\omega}_2, \ldots, \omega_r) + \underbrace{T(\overline{\omega}_1, \overline{\omega}_2, \ldots, \omega_r)}_{=0},$$

also $T(\overline{\omega}_1, \omega_1, \ldots, \omega_r) = -T(\omega_1, \overline{\omega}_1, \ldots, \omega_r)$. Dasselbe Argument kann nicht nur für die Positionen 1 und 2 angewendet werden, sondern für beliebige zwei Positionen. Das zeigt: T ist schiefsymmetrisch.

Aussage (ii): Im Fall $r = 0$ und $r = 1$ ist wie in Aussage (iv) gezeigt, jeder Tensor alternierend und schiefsymmetrisch.

Es sei daher nun $r \geqslant 2$. Für jeden Tensor $T \in \mathcal{T}_0^r(V) \cong \operatorname{Mult}(V^* \times \cdots \times V^*; K)$ gilt:

$$\begin{aligned} &T(\omega_1 + \overline{\omega}_1, \omega_1 + \overline{\omega}_1, \ldots, \omega_r) \\ &\qquad = T(\omega_1, \omega_1, \ldots, \omega_r) + T(\overline{\omega}_1, \omega_1, \ldots, \omega_r) + T(\omega_1, \overline{\omega}_1, \ldots, \omega_r) + T(\overline{\omega}_1, \overline{\omega}_1, \ldots, \omega_r). \end{aligned}$$

Ist $T \in \mathcal{T}_0^r(V)$ schiefsymmetrisch und gilt $\omega_1 = \overline{\omega}_1$, also auch $\omega_1 + \overline{\omega}_1 = 2\,\omega_1$, dann folgt

$$4\,T(\omega_1, \omega_1, \ldots, \omega_r) = T(\omega_1, \omega_1, \ldots, \omega_r) + \underbrace{T(\overline{\omega}_1, \omega_1, \ldots, \omega_r) + T(\omega_1, \overline{\omega}_1, \ldots, \omega_r)}_{=0} + T(\omega_1, \omega_1, \ldots, \omega_r),$$

also auch

$$2\,T(\omega_1, \omega_1, \ldots, \omega_r) = 0.$$

Falls $\operatorname{char}(K) \neq 2$ gilt, so folgt $T(\omega_1, \omega_1, \ldots, \omega_r) = 0$. Dasselbe Argument kann nicht nur für die Positionen 1 und 2 angewendet werden, sondern für beliebige zwei Positionen. Das zeigt: T ist alternierend.

Aussage (iii): Im Fall $\operatorname{char}(K) = 2$ gilt $-1 = 1$, also ist (22.19) äquivalent zu (22.18). □

Lemma 22.31 (Symmetrie und Schiefsymmetrie von Tensoren und Komponenten)**.**
Es seien V ein endlich-dimensionaler Vektorraum über einem Körper K, $(e_1, \ldots, e_n)$ eine Basis von V und $(\varepsilon^1, \ldots, \varepsilon^n)$ die duale Basis von V^*. Weiter sei $r \in \mathbb{N}_0$. Dann sind äquivalent:

(i) $T \in \mathcal{T}_0^r(V)$ ist symmetrisch.

(ii) Die Komponenten erfüllen $T^{i_1,\ldots,i_r} = T^{i_{\sigma(1)},\ldots,i_{\sigma(r)}}$ für alle $1 \leqslant i_1, \ldots, i_r \leqslant n$ und alle Permutationen $\sigma \in S_r$.

Weiter sind äquivalent:

(i) $T \in \mathcal{T}_0^r(V)$ ist schiefsymmetrisch.

(*ii*) Die Komponenten erfüllen $T^{i_1,\dots,i_r} = (\operatorname{sgn}\sigma)\, T^{i_{\sigma(1)},\dots,i_{\sigma(r)}}$ für alle $1 \leqslant i_1, \dots, i_r \leqslant n$ und alle Permutationen $\sigma \in S_r$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-4.3. □

Lemma 22.32 (symmetrische und schiefsymmetrische Tensoren bilden Unterräume).
Es sei V ein Vektorraum über einem Körper K. Weiter sei $r \in \mathbb{N}_0$. Die Mengen der symmetrischen bzw. der schiefsymmetrischen Tensoren $\mathcal{T}_0^r(V)_{\text{sym}}$ bzw. $\mathcal{T}_0^r(V)_{\text{skew}}$ bilden Unterräume von $\mathcal{T}_0^r(V)$.

Beweis. Der Nulltensor $0 \otimes \cdots \otimes 0$ gehört zu $\mathcal{T}_0^r(V)_{\text{sym}}$, also ist $\mathcal{T}_0^r(V)_{\text{sym}} \neq \emptyset$. Sind $S, T \in \mathcal{T}_0^r(V)_{\text{sym}}$, so gilt

$$\begin{aligned}(S+T)(\varepsilon^{i_1}, \dots, \varepsilon^{i_r}) &= S(\varepsilon^{i_1}, \dots, \varepsilon^{i_r}) + T(\varepsilon^{i_1}, \dots, \varepsilon^{i_r}) \\ &= S(\varepsilon^{i_{\sigma(1)}}, \dots, \varepsilon^{i_{\sigma(r)}}) + T(\varepsilon^{i_{\sigma(1)}}, \dots, \varepsilon^{i_{\sigma(r)}}) \\ &= (S+T)(\varepsilon^{i_{\sigma(1)}}, \dots, \varepsilon^{i_{\sigma(r)}})\end{aligned}$$

und weiterhin für $\alpha \in K$

$$\begin{aligned}(\alpha\, T)(\varepsilon^{i_1}, \dots, \varepsilon^{i_r}) &= \alpha\, T(\varepsilon^{i_1}, \dots, \varepsilon^{i_r}) \\ &= \alpha\, T(\varepsilon^{i_{\sigma(1)}}, \dots, \varepsilon^{i_{\sigma(r)}}) \\ &= (\alpha\, T)(\varepsilon^{i_{\sigma(1)}}, \dots, \varepsilon^{i_{\sigma(r)}}).\end{aligned}$$

Nach dem Unterraumkriterium Satz 12.8 ist $\mathcal{T}_0^r(V)_{\text{sym}}$ damit ein Unterraum. Der Beweis für $\mathcal{T}_0^r(V)_{\text{skew}}$ läuft analog. □

Definition 22.33 (Projektion).
Es sei V ein Vektorraum über dem Körper K. Ein Endomorphismus $P \in \operatorname{End}(V)$ heißt eine **Projektion** (englisch: projection) oder ein **Projektor** (englisch: projector), wenn er idempotent ist, wenn also

$$P \circ P = P \tag{22.21}$$

gilt. Insbesondere sprechen wir von einer **Projektion auf den Unterraum** $\operatorname{Bild}(P)$. △

Wir zeigen jetzt, wie wir einen beliebigen $(r, 0)$-Tensor symmetrisieren bzw. „schiefsymmetrisieren" können. Dazu definieren wir zunächst, wie eine Permutation $\sigma \in S_r$ als lineare Abbildung auf dem Tensorproduktraum $\mathcal{T}_0^r(V)$ wirkt: Für einen Basistensor $e_{i_1} \otimes \cdots \otimes e_{i_r}$ definieren wir

$$\sigma(e_{i_1} \otimes \cdots \otimes e_{i_r}) \coloneqq e_{i_{\sigma(1)}} \otimes \cdots \otimes e_{i_{\sigma(r)}}. \tag{22.22}$$

Dadurch ist die **Permutation** σ durch Satz 22.21 auf ganz $\mathcal{T}_0^r(V)$ eindeutig definiert.

Satz 22.34 (Projektionen auf die Unterräume der symmetrischen und schiefsymmetrischen Tensoren).

Es sei V ein Vektorraum über einem Körper K ~~mit $\dim(V) = n \in \mathbb{N}_0$~~ mit $\operatorname{char}(K) = 0$. Weiter sei $r \in \mathbb{N}_0$.

(*i*) Der Endomorphismus

$$\mathcal{T}_0^r(V) \ni T \mapsto \operatorname{proj}_{\text{sym}}(T) \coloneqq \frac{1}{r!} \sum_{\sigma \in S_r} \sigma(T) \in \mathcal{T}_0^r(V) \tag{22.23}$$

ist eine Projektion auf den Unterraum der symmetrischen Tensoren $\mathcal{T}_0^r(V)_{\text{sym}}$.

(ii) Der Endomorphismus

$$\mathcal{T}_0^r(V) \ni T \mapsto \operatorname{proj}_{\text{skew}}(T) := \frac{1}{r!} \sum_{\sigma \in S_r} (\operatorname{sgn} \sigma)\, \sigma(T) \in \mathcal{T}_0^r(V) \tag{22.24}$$

ist eine Projektion auf den Unterraum der schiefsymmetrischen Tensoren $\mathcal{T}_0^r(V)_{\text{skew}}$.

Quizfrage 22.5: Was bedeutet $\frac{1}{r!}$ in einem beliebigen Körper, und warum fordern wir $\operatorname{char}(K) = 0$?

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-4.4. □

Quizfrage 22.6: Wie sehen die Komponenten eines symmetrisierten Tensors aus? **Quizfrage 22.7:** Wie sehen die Komponenten eines schiefsymmetrisierten Tensors aus?

Lemma 22.35 (Dimension des Unterraumes $\mathcal{T}_0^r(V)_{\text{skew}}$)**.**
Es sei V ein Vektorraum über einem Körper K mit $\operatorname{char}(K) \neq 2$ und $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $r \in \mathbb{N}_0$. Dann gilt[63]

$$\dim(\mathcal{T}_0^r(V)_{\text{skew}}) = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}. \tag{22.25}$$

Insbesondere gilt $\dim(\mathcal{T}_0^n(V)_{\text{skew}}) = 1$ und $\dim(\mathcal{T}_0^r(V)_{\text{skew}}) = 0$ für $r \geqslant n$.

Beweis. Nach Lemma 22.31 liegt die Schiefsymmetrie von $T \in \mathcal{T}_0^r(V)_{\text{skew}}$ genau dann vor, wenn die Komponenten bzgl. einer beliebigen Basis von V schiefsymmetrisch sind. In diesem Fall sind die Komponenten bereits durch die Komponenten mit den Indizes

$$I_{\text{skew}} := \left\{(i_1, \ldots, i_r) \in [\![1, n]\!]^r \,\middle|\, 1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n\right\}$$

eindeutig bestimmt. Für die Indizes nicht in I_{skew} ist entweder ein Index doppelt, woraus folgt, dass dieser Eintrag in den Komponenten gleich 0 ist, oder es gibt eine Permutation σ, sodass $\sigma(i_1, \ldots, i_r) \in I_{\text{skew}}$ liegt. Andererseits können die Einträge mit Indizes in I_{skew} frei gewählt werden, sodass sich $\dim(\mathcal{T}_0^r(V)_{\text{skew}}) = \# I_{\text{skew}}$ ergibt.

Es bleibt zu zeigen, dass $\# I_{\text{skew}} = \binom{n}{r}$ ist. Das ist aber offensichtlich, da $\binom{n}{r}$ die Anzahl der Möglichkeiten angibt, genau r verschiedene Indizes aus dem Vorrat $[\![1, n]\!]$ auszuwählen, die dann aufsteigend sortiert werden. □

Lemma 22.35 zeigt insbesondere, dass es in einem Vektorraum V mit $\dim(V) = n \in \mathbb{N}_0$ bis auf Skalierung nur einen einzigen schiefsymmetrischen Tensor in $\mathcal{T}_0^n(V)_{\text{skew}}$ gibt. Wir können die Resultate aus § 22.4 natürlich auch auf V^* statt V anwenden. Insbesondere gibt es — bis auf Skalierung — nur einen einzigen schiefsymmetrischen Tensor in $\mathcal{T}_n^0(V)_{\text{skew}}$, also nach Satz 22.23 — bis auf Skalierung — auch nur eine einzige schiefsymmetrische Multilinearform $V \times \cdots \times V \to K$. Diese ist Gegenstand von § 23.

Ende der Vorlesung 8

Ende der Woche 4

[63] In (22.25) ist $\binom{n}{r}$ der **Binomialkoeffizient** (englisch: binomial coefficient) „n über r“.

§ 23 Determinanten

Literatur: Fischer, Springborn, 2020, Kapitel 4, Bosch, 2014, Kapitel 4

In diesem Abschnitt ist V stets ein Vektorraum über einem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Die Determinante ist eine Maßzahl für Endomorphismen $f \in \operatorname{End}(V)$, die gewisse Eigenschaften festhält. Unser Vorgehen wird sein, zunächst Determinantenformen auf V^n zu definieren, eine bestimmte davon als *die* Determinante für Matrizen zu erklären und diesen Begriff dann mittels Darstellungsmatrizen auf Endomorphismen zu übertragen.

Definition 23.1 (Determinantenform auf einem Produktraum V^n)**.**
Es sei V ein Vektorraum über einem Körper K und $\dim(V) = n \in \mathbb{N}_0$. Eine Abbildung

$$\Delta \colon V^n \ni (v_1, \ldots, v_n) \mapsto \Delta(v_1, \ldots, v_n) \in K$$

heißt eine **Determinantenform** (englisch: determinant form) auf V^n, wenn folgende Eigenschaften gelten:

(i) Δ ist eine n-lineare Form auf V^n.

(ii) Δ ist alternierend.

(iii) Δ ist nicht die Nullform. $\triangle$

Als n-lineare Form auf V^n wird Δ durch genau einen Tensor $T \in \mathcal{T}_n^0(V)$ repräsentiert. Wegen der alternierenden Eigenschaft ist dieser Tensor auch schiefsymmetrisch (Lemma 22.30). Nach Lemma 22.31 und dem Beweis von Lemma 22.35 ist das Schema der Komponenten bereits durch einen einzigen Index eindeutig festgelegt. O. B. d. A. ist das der Index $(1, 2, \ldots, n)$, also der Funktionswert $\Delta(b_1, \ldots, b_n)$ auf einer beliebigen Basis $(b_1, \ldots, b_n)$ von V. Dass Δ nicht die Nullform ist, bedeutet gerade, dass $\Delta(b_1, \ldots, b_n) \neq 0$ ist. Eine konkrete Determinantenformen auf V^n wird also durch den Funktionswert $\Delta(b_1, \ldots, b_n) \in K \setminus \{0\}$ festgelegt, den man als Normierungswert verstehen kann.

Wie folgendes Resultat zeigt, hat die in Definition 23.1 geforderte alternierende Eigenschaft die Funktion, die lineare Abhängigkeit der eingesetzten Vektoren zu erkennen:

Lemma 23.2 (Bedeutung der alternierenden Eigenschaft)**.**
Es sei V ein Vektorraum über einem Körper K und $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $T \in \mathcal{T}_n^0(V)$. Dann sind äquivalent:

(i) T ist alternierend.

(ii) Wenn die Familie $(v_1, \ldots, v_n)$ linear abhängig ist, dann gilt $T(v_1, \ldots, v_n) = 0$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es sei $(v_1, \ldots, v_n)$ linear abhängig, dann gibt es einen Index $1 \leqslant i \leqslant n$, sodass

$$v_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_j \, v_j$$

gilt (Lemma 13.3). O. B. d. A. gilt $i = 1$. Wir erhalten

$$\begin{aligned} T(v_1, \dots, v_n) &= T\Big(\sum_{j=2}^{n} \alpha_j\, v_j, v_2, \dots, v_n\Big) \\ &= \sum_{j=2}^{n} \alpha_j\, T(v_j, v_2, \dots, v_n) && \text{wegen der Multilinearität} \\ &= \sum_{j=2}^{n} 0 && \text{wegen der alternierenden Eigenschaft} \\ &= 0. \end{aligned}$$

Aussage $(ii) \Rightarrow$ Aussage (i): Wir betrachten $(v_1, \dots, v_n)$, wobei zwei der Vektoren identisch sind. Dann ist $(v_1, \dots, v_n)$ linear abhängig, also $T(v_1, \dots, v_n) = 0$. Damit ist T alternierend. □

Wie sieht nun eine Determinantenform auf V^n konkret aus? Diese Frage beantwortet das folgende Lemma.

Lemma 23.3 (Gestalt von Determinantenformen)**.**
Es sei V ein Vektorraum über einem Körper K und $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $(b_1, \dots, b_n)$ eine Basis von V. Ist Δ eine Determinantenform auf V^n, dann gilt:

$$\Delta(v_1, \dots, v_n) = \Big(\sum_{\sigma \in S_n} \underbrace{(\operatorname{sgn} \sigma)}_{=\pm 1 \in K}\, t_{\sigma(1),1} \cdots t_{\sigma(n),n}\Big)\, \underbrace{\Delta(b_1, \dots, b_n)}_{\text{Normierungswert } \in K \setminus \{0\}} \tag{23.1}$$

für alle $(v_1, \dots, v_n) \in V^n$, wobei T die durch

$$v_j = \sum_{i=1}^{n} t_{ij}\, b_i$$

eindeutig definierte Matrix ist.[64]

Beweis. Aufgrund der Multilinearität der Determinantenform gilt

$$\begin{aligned} \Delta(v_1, \dots, v_n) &= \Delta\Big(\sum_{i=1}^{n} t_{i,1}\, b_i, \dots, \sum_{i=1}^{n} t_{i,n}\, b_i\Big) \\ &= \sum_{i_1=1}^{n} t_{i_1,1}\, \Delta\Big(b_{i_1}, \dots, \sum_{i=1}^{n} t_{i,n}\, b_i\Big) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} t_{i_1,1}\, t_{i_2,2}\, \Delta\Big(b_{i_1}, b_{i_2}, \dots, \sum_{i=1}^{n} t_{i,n}\, b_i\Big) \\ &= \dots \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} t_{i_1,1}\, t_{i_2,2} \cdots t_{i_n,n}\, \Delta(b_{i_1}, b_{i_2}, \dots, b_{i_n}). \end{aligned}$$

[64] In der j-ten Spalte von T 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 T die Transformationsmatrix $T = \mathcal{T}_{(b_1,\dots,b_n)}^{(v_1,\dots,v_n)} \in K^{n \times n}$, siehe Definition 20.1.

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. Mit anderen Worten: Wir beschränken die Summation auf die *Permutationen* der Menge $\{1, \ldots, n\}$, also die Elemente der symmetrischen Gruppe S_n, siehe Definition 7.21. Es ergibt sich:

$$= \sum_{\sigma \in S_n} t_{\sigma(1),1} \cdots t_{\sigma(n),n} \, \Delta(b_{\sigma(1)}, \ldots, b_{\sigma(n)}).$$

Da Δ als alternierende Multilinearform schiefsymmetrisch ist, folgt schließlich

$$= \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \, t_{\sigma(1),1} \cdots t_{\sigma(n),n} \Big) \, \Delta(b_1, \ldots, b_n),$$

was zu beweisen war. □

Folgerung 23.4 (Determinantenformen sind genau auf Basen nicht Null)**.**
Es sei V ein Vektorraum über einem Körper K und $\dim(V) = n \in \mathbb{N}_0$. Ist Δ eine Determinantenform auf V^n, dann sind äquivalent:

(i) $\Delta(v_1, \ldots, v_n) \neq 0$.

(ii) $(v_1, \ldots, v_n)$ ist eine Basis von V.

Beweis. $\neg$ Aussage (ii) $\Rightarrow$ $\neg$ Aussage (i): Ist $(v_1, \ldots, v_n)$ keine Basis von V, dann ist $(v_1, \ldots, v_n)$ eine linear abhängige Familie. Aus Lemma 23.2 folgt $\Delta(v_1, \ldots, v_n) = 0$.

Aussage (ii) $\Rightarrow$ Aussage (i): Δ ist nach Definition 23.1 nicht die Nullform. Also gibt es eine Familie $(b_1, \ldots, b_n)$ mit $\Delta(b_1, \ldots, b_n) \neq 0$. Aus Lemma 23.2 folgt, dass $(b_1, \ldots, b_n)$ linear unabhängig ist, also eine Basis von V.

Wir wenden nun Lemma 23.3 an mit vertauschten Rollen von $(b_1, \ldots, b_n)$ und $(v_1, \ldots, v_n)$, also mit $T := \mathcal{T}_{(v_1,\ldots,v_n)}^{(b_1,\ldots,b_n)}$. Es gilt also der Zusammenhang

$$\Delta(b_1, \ldots, b_n) = \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \, t_{\sigma(1),1} \cdots t_{\sigma(n),n} \Big) \, \Delta(v_1, \ldots, v_n),$$

also $\Delta(b_1, \ldots, b_n) = \alpha \, \Delta(v_1, \ldots, v_n)$ für ein $\alpha \in K$. Wegen $\Delta(b_1, \ldots, b_n) \neq 0$ muss auch $\Delta(v_1, \ldots, v_n) \neq 0$ sein. □

Schließlich müssen wir noch zeigen, dass Determinantenformen auf V^n tatsächlich existieren, denn wir haben bisher nur bewiesen, dass eine Determinantenform *notwendigerweise* die Gestalt (23.1) hat. Die Konstruktion benutzte, dass alternierende Formen notwendigerweise schiefsymmetrisch sind. Wir müssen uns also noch davon überzeugen, dass (23.1) tatsächlich auch die Bedingungen aus Definition 23.1 erfüllt, vor allem, dass eine solche Form alternierend ist, und zwar auch im Fall $\operatorname{char}(K) = 2$.

Satz 23.5 (Existenzsatz für Determinantenformen)**.**

Es sei V ein Vektorraum über einem Körper K und $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $(b_1, \dots, b_n)$ eine Basis von V. Zu jedem $\alpha \in K \setminus \{0\}$ gibt es genau eine Determinantenform auf V^n mit der Eigenschaft $\Delta(b_1, \dots, b_n) = \alpha$. Diese ist durch

$$\Delta(v_1, \dots, v_n) = \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots t_{\sigma(n),n} \Big) \underbrace{\Delta(b_1, \dots, b_n)}_{=\alpha} \tag{23.2}$$

gegeben, wobei T die durch

$$v_j = \sum_{i=1}^n t_{ij}\, b_i$$

eindeutig definierte Matrix ist.

Beweis. Wir wissen bereits aus Lemma 23.3, dass eine Determinantenform auf V^n notwendigerweise (23.2) erfüllt, also durch den Wert von $\alpha \in K \setminus \{0\}$ eindeutig bestimmt ist. Zu zeigen ist, dass die durch (23.2) gegebene Form für $\alpha \in K \setminus \{0\}$ tatsächlich die Bedingungen aus Definition 23.1 erfüllt.

Schritt 1: Δ ist n-linear:

Es gilt $v_j = \sum_{i=1}^n t_{ij}\, b_i$ für $j = 1, \dots, n$. Ersetzen wir also genau einen der Vektoren v_j durch $\beta\, v_j$ mit $\beta \in K$, dann wird die Spalte j der Matrix T mit β skaliert. Es folgt

$$\begin{aligned}
\Delta(v_1, \dots, \beta\, v_j, \dots, v_n) &= \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots (\beta\, t_{\sigma(j),j}) \cdots t_{\sigma(n),n} \Big) \Delta(b_1, \dots, b_n) \\
&= \beta \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots t_{\sigma(n),n} \Big) \Delta(b_1, \dots, b_n) \\
&= \beta\, \Delta(v_1, \dots, v_n).
\end{aligned}$$

Es gelte weiterhin $w_j = \sum_{i=1}^n s_{ij}\, b_i$ für $j = 1, \dots, n$. Ersetzen wir genau einen Vektoren v_j durch $v_j + w_j$, dann wird die j-te Spalte der Matrix T mit der j-ten Spalte der Matrix S addiert. Es folgt

$$\begin{aligned}
\Delta(v_1, \dots, v_j + w_j, \dots, v_n) &= \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots (t_{\sigma(j),j} + s_{\sigma(j),j}) \cdots t_{\sigma(n),n} \Big) \Delta(b_1, \dots, b_n) \\
&= \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots t_{\sigma(j),j} \cdots t_{\sigma(n),n} \Big) \Delta(b_1, \dots, b_n) \\
&+ \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots s_{\sigma(j),j} \cdots t_{\sigma(n),n} \Big) \Delta(b_1, \dots, b_n) \\
&= \Delta(v_1, \dots, v_j, \dots, v_n) + \Delta(v_1, \dots, w_j, \dots, v_n).
\end{aligned}$$

Also ist Δ linear im j-ten Argument.

Schritt 2: Δ ist alternierend:

Wir brauchen nur den Fall $n \geqslant 2$ zu untersuchen. (**Quizfrage 23.1:** Warum?) Es seien o. B. d. A. $v_1 = v_2$. Wir betrachten die Transposition $\tau := \tau(1, 2) \in S_n$. Die symmetrische Gruppe S_n zerfällt in die zwei disjunkten Nebenklassen A_n (die geraden Permutationen) und

$A_n \circ \tau$ (die ungeraden Permutationen), vgl. Beispiel 8.18. Es ist also $S_n = A_n \cup (A_n \circ \tau)$ eine disjunkte Vereinigung und damit

$$\begin{aligned}
&\Delta(v_1, \ldots, v_n) \\
&\quad = \Big(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, t_{\sigma(1),1} \cdots t_{\sigma(n),n} \Big)\, \Delta(b_1, \ldots, b_n) \\
&\quad = \Big(\sum_{\sigma \in A_n} \Big[\underbrace{(\operatorname{sgn} \sigma)}_{=1}\, t_{\sigma(1),1} \cdots t_{\sigma(n),n} + \underbrace{(\operatorname{sgn}(\sigma \circ \tau))}_{=-1}\, t_{(\sigma\circ\tau)(1),1} \cdots t_{(\sigma\circ\tau)(n),n} \Big] \Big)\, \Delta(b_1, \ldots, b_n).
\end{aligned}$$

Für Indizes $j \geqslant 3$ gilt $(\sigma \circ \tau)(j) = \sigma(j)$ und damit $t_{(\sigma\circ\tau)(j),j} = t_{\sigma(j),j}$. Für die ersten zwei Spalten von T gilt $t_{\bullet 1} = t_{\bullet 2}$ und daher

$$t_{(\sigma\circ\tau)(1),1} = t_{\sigma(2),1} = t_{\sigma(2),2} \quad \text{und} \quad t_{(\sigma\circ\tau)(2),2} = t_{\sigma(1),2} = t_{\sigma(1),1}.$$

Das zeigt

$$\begin{aligned}
&\Delta(v_1, \ldots, v_n) \\
&\quad = \Big(\sum_{\sigma \in A_n} \big[t_{\sigma(1),1} \cdots t_{\sigma(n),n} - t_{(\sigma\circ\tau)(1),1} \cdots t_{(\sigma\circ\tau)(n),n} \big] \Big)\, \Delta(b_1, \ldots, b_n) \\
&\quad = \Big(\sum_{\sigma \in A_n} \underbrace{\big[t_{\sigma(1),1} \cdots t_{\sigma(n),n} - t_{\sigma(1),1} \cdots t_{\sigma(n),n} \big]}_{=0} \Big)\, \Delta(b_1, \ldots, b_n) \\
&\quad = 0.
\end{aligned}$$

Das gilt auch im Fall $\operatorname{char}(K) = 2$, also wenn $-1 = 1$ ist. (**Quizfrage 23.2:** Warum?) Also ist Δ alternierend.

Schritt 3: Δ erfüllt $\Delta(b_1, \ldots, b_n) = \alpha$, ist also nicht die Nullform:

Diese Aussage ist klar, weil $\alpha \neq 0$ ist. □

§ 23.1 Die Determinante einer Matrix

Wir betrachten in diesem Abschnitt den Vektorraum $V = K^n$, somit ist $V^n \cong K^{n\times n}$. Wir identifizieren also ein Tupel $(a_1, \ldots, a_n)$ von Vektoren in K^n mit der Matrix $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$. Die Determinantenformen auf $K^{n\times n}$ unterscheiden sich nur im Normierungswert $\Delta(b_1, \ldots, b_n)$ auf einer Basis $(b_1, \ldots, b_n)$ von K^n. Man setzt diesen Wert auf der Standardbasis $(e_1, \ldots, e_n)$ auf 1. Dadurch stimmt die durch

$$a_j = \sum_{i=1}^{n} t_{ij}\, e_i$$

definierte Koeffizientenmatrix T mit A überein. Das führt auf die folgende Definition.

Definition 23.6 (Determinante einer Matrix)**.**
Es sei K ein Körper und $n \in \mathbb{N}_0$. Die **Determinante** (englisch: determinant) einer Matrix $A \in K^{n\times n}$ ist definiert durch die spezielle Determinantenform

$$\det(A) \coloneqq \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}. \tag{23.3}$$

△

Beachte: Mit dieser Definition können wir die Aussage von Satz 23.5 auch formulieren als $\Delta(v_1, \ldots, v_n) = \det(T)\, \Delta(b_1, \ldots, b_n)$.

Die Gleichung (23.3) heißt die **Leibniz-Formel** (englisch: Leibniz formula).

Manche Autoren schreiben statt $\det(A)$ auch $|A|$. Dieser Praxis folgen wir nicht.

Beispiel 23.7 (Determinante einer Matrix)**.**

(i) Im Fall $n = 0$ ist die Determinante der einzigen (leeren) $n \times n$-Matrix $\det(A) = 1$. (**Quizfrage 23.3:** Wie folgt das aus (23.3)?)

(ii) Im Fall $n = 1$ gibt es nur die Permutation $\sigma = \mathrm{id}$. Daher folgt für $A = (a)$ aus (23.3) der Zusammenhang $\det(A) = a$. Die Determinante ist also gerade der einzige Eintrag der Matrix.

(iii) Im Fall $n = 2$ gibt es zwei Permutationen, $\sigma = \mathrm{id}$ und $\sigma = \tau(1, 2)$. Daher folgt aus (23.3):[65]

$$\det(A) = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\, a_{22} - a_{12}\, a_{21}.$$

Grafisch können wir diese Regel wie folgt darstellen:

$$\begin{array}{cc} a_{11} & a_{12} \\ {\scriptstyle +} & \\ a_{21} \, {\scriptstyle -} & a_{22} \end{array}$$

Das Produkt der Hauptdiagonalelemente geht also positiv ein, während das Produkt der Elemente der „Gegendiagonale“ negativ gewichtet wird.

(iv) Im Fall $n = 3$ gibt es sechs Permutionen. Davon sind drei gerade, nämlich $\sigma = \mathrm{id} = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right)$, $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$ 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 (23.3) ergibt

$$\det(A) = \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} \; a_{11}\, a_{22}\, a_{33} + a_{21}\, a_{32}\, a_{13} + a_{31}\, a_{12}\, a_{23} \\ -\; a_{31}\, a_{22}\, a_{13} - a_{21}\, a_{12}\, a_{33} - a_{11}\, a_{32}\, a_{23}. \end{array}$$

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 (23.3) vorgehen. Für größere Matrizen sehen wir später noch effizientere Möglichkeiten.

[65] Das zweite Klammernpaar lässt man bei Ausdrücken wie $\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right)$ häufig weg und schreibt einfach $\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

Lemma 23.8 (Eigenschaften der Determinante von Matrizen).
Es sei 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 Spaltenvektoren $(a_{\bullet 1}, \ldots, a_{\bullet n})$ von A.

(*ii*) $\det(\alpha\, A) = \alpha^n \det(A)$ für alle $\alpha \in K$.

(*iii*) $\det(A) \neq 0 \Leftrightarrow A$ ist regulär $\Leftrightarrow \operatorname{Rang}(A) = n \Leftrightarrow (a_{\bullet 1}, \ldots, a_{\bullet n})$ ist linear unabhängig.

(*iv*) $\det(I) = 1$.

(*v*) $\det(A\,B) = \det(A)\ \det(B)$.[66]

(*vi*) $\det(A^{-1}) = 1/\det(A)$, falls A invertierbar ist.

(*vii*) $\det(A^\intercal) = \det(A)$.

Beweis. Aussage (*i*): siehe Beweis von Satz 23.5.

Aussage (*ii*) folgt sofort aus der Multilinearität der Determinante.

Aussage (*iii*):

$$\begin{aligned}
&\det(A) \neq 0 && \\
\Leftrightarrow\ &(a_{\bullet 1}, \ldots, a_{\bullet n}) \text{ ist linear unabhängig} && \text{nach Folgerung 23.4} \\
\Leftrightarrow\ &\operatorname{Rang}(A) = n && \text{nach Satz 15.12} \\
\Leftrightarrow\ &A \text{ ist regulär} && \text{nach Satz 15.40.}
\end{aligned}$$

Aussage (*iv*): Die Einträge der Einheitsmatrix sind δ_{ij} für $i, j = 1, \ldots, n$. Also produziert nur $\sigma = \mathrm{id}$ einen Nicht-Null-Beitrag in der Summe (23.3). Genauer:

$$\det(I) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, \delta_{\sigma(1),1} \cdots \delta_{\sigma(n),n} = (\operatorname{sgn} \mathrm{id})\, \delta_{\mathrm{id}(1),1} \cdots \delta_{\mathrm{id}(n),n} = 1.$$

Aussage (*v*): **Der Beweis dieser Aussage wurde neu formuliert.** Wir schreiben die Matrix B spaltenweise als

$$B = \left[\sum_{i_1=1}^n b_{i_1,1}\, e_{i_1} \quad \cdots \quad \sum_{i_n=1}^n b_{i_n,n}\, e_{i_n}\right].$$

Dann gilt nach Definition des Matrix-Matrix-Produkts

$$\begin{aligned}
A\,B &= \left[\sum_{i_1=1}^n A\, b_{i_1,1}\, e_{i_1} \quad \cdots \quad \sum_{i_n=1}^n A\, b_{i_n,n}\, e_{i_n}\right] \\
&= \left[\sum_{i_1=1}^n b_{i_1,1}\, A\, e_{i_1} \quad \cdots \quad \sum_{i_n=1}^n b_{i_n,n}\, A\, e_{i_n}\right].
\end{aligned}$$

Daher gilt für die Determinante unter Verwendung der Multilinearität

$$\det(A\,B) = \det\left(\sum_{i_1=1}^n b_{i_1,1}\, A\, e_{i_1} \quad \cdots \quad \sum_{i_n=1}^n b_{i_n,n}\, A\, e_{i_n}\right)$$

[66] Die Determinante ist also ein Monoidhomomorphismus vom Monoid $(K^{n\times n}, \cdot)$ in das Monoid $(K, \cdot)$.

$$= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1,1} \cdots b_{i_n,n} \det \begin{pmatrix} A\, e_{i_1} & \cdots & A\, e_{i_n} \end{pmatrix}.$$

Wie im Beweis von Lemma 23.3 können wir die Summation auf diejenigen Indextupel $(i_1, \ldots, i_n)$ beschränken, deren Einträge sämtlich verschieden sind:

$$\begin{aligned} &= \sum_{\sigma \in S_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} \det \begin{pmatrix} A\, e_{i_1} & \cdots & A\, e_{i_n} \end{pmatrix} \\ &= \underbrace{\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, b_{\sigma(1),1} \cdots b_{\sigma(n),n}}_{\det(B)} \underbrace{\det \begin{pmatrix} A\, e_1 & \cdots & A\, e_n \end{pmatrix}}_{\det(A)} \\ &= \det(A)\, \det(B) \end{aligned}$$

Aussage (vi): Es sei A invertierbar, dann folgt aus Aussage (iii) und Aussage (iv)

$$1 = \det(I) = \det(A\, A^{-1}) = \det(A)\, \det(A^{-1})$$

und damit die Behauptung.

Aussage (vii): Es gilt

$$\begin{aligned} \det(A^\mathsf{T}) &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} && \text{nach der Leibniz-Formel (23.3)} \\ &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n} && \text{durch Umsortieren der Faktoren} \\ &= \sum_{\tau \in S_n} (\operatorname{sgn} \tau^{-1})\, a_{\tau(1),1} \cdots a_{\tau(n),n} && \text{durch Ersetzen von } \tau := \sigma^{-1} \\ &= \sum_{\tau \in S_n} (\operatorname{sgn} \tau)\, a_{\tau(1),1} \cdots a_{\tau(n),n} && \text{wegen } \operatorname{sgn} \tau = \operatorname{sgn} \tau^{-1} \text{, siehe Folgerung 7.30} \\ &= \det(A). \end{aligned}$$

(**Quizfrage 23.4:** Warum durfte die Summation über $\sigma \in S_n$ durch die Summation über $\tau \in S_n$ ersetzt werden?) □

Lemma 23.9 (ähnliche Matrizen haben dieselben Determinante)**.**
Es sei K ein Körper und $A, \widehat{A} \in K^{n \times n}$ mit $n \in \mathbb{N}_0$. Sind A und $\widehat{A}$ ähnlich, dann gilt $\det(A) = \det(\widehat{A})$.

Beweis. Sind A und $\widehat{A}$, dann gibt es eine invertierbare Matrix $T \in K^{n \times n}$ mit

$$\widehat{A} = T\, A\, T^{-1}.$$

Mit Lemma 23.8 folgt

$$\begin{aligned} \det(\widehat{A}) &= \det(T)\, \det(A)\, \det(T^{-1}) \\ &= \det(T)\, \det(A)\, \det(T)^{-1} \\ &= \det(A). \end{aligned}$$

□

Ende der Vorlesung 9

Wie kann nun die Determinante einer Matrix effizienter berechnet werden als mit Hilfe der Definition 23.6? Wir geben dazu mehrere Resultate an.

Lemma 23.10 (Determinante einer Dreiecksmatrix).
Es sei 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

$$\det(A) = a_{11} \cdots a_{nn}. \tag{23.4}$$

(ii) Ist A eine obere oder untere Blockdreiecksmatrix, also $A = \left[\begin{smallmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{smallmatrix}\right]$ bzw. $A = \left[\begin{smallmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{smallmatrix}\right]$, dann gilt

$$\det(A) = \det(A_{11}) \det(A_{22}). \tag{23.5}$$

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-5.1. □

Aufgrund des vorherigen Lemmas ist es nützlich, eine Matrix z. B. durch Zeilenumformungen auf Zeilenstufenform (also obere Dreiecksgestalt) zu bringen, um ihre Determinante zu berechnen. Wir betrachten zu diesem Zweck die elementaren Zeilenumformungen aus § 15.3:

Typ I:

$$\text{Für } D := \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \text{ gilt } \det(D) = \alpha.$$

Typ II:

$$\text{Für } S := \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & \alpha & 1 & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \text{ gilt } \det(S) = 1.$$

Typ III:

$$\text{Für } T := \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & 1 & & & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{bmatrix} \text{ gilt } \det(T) = -1.$$

Bei der Berechnung der Zeilenstufenform multiplizieren wir wiederholt die betrachtete Matrix A von links mit Matrizen vom Typ II oder III. (**Quizfrage 23.5:** 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.

Beispiel 23.11 (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 16.7. Wir bringen A dazu auf Zeilenstufenform:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 2 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 2 \end{bmatrix}$$ Tauschen der Zeilen 1 und 2

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 2 \end{bmatrix}$$ Erzeugen von Nullen unterhalb des Pivot-Elements

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$ Tauschen der Zeilen 2 und 3

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$ Erzeugen von Nullen unterhalb des Pivot-Elements

Die Determinante der letzten Matrix ist nach Lemma 23.10 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 $A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 4 & 2 \end{bmatrix}$:

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 0 \\ 3 & 4 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ 3 & 4 & 2 \end{bmatrix}$$ 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 23.8). △

Definition 23.12 (Streichungsmatrix, Minoren, Kofaktoren und Adjunkte)**.**
Es sei 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

$$(A)_{\neq i, \neq j} \coloneqq \begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{bmatrix}. \tag{23.6}$$

(*ii*) Die **Unterdeterminante** (englisch: subdeterminant) oder der **Minor** (englisch: minor) von A bzgl. des Index (i, j) ist die Determinante der zugehörigen Streichungsmatrix, also[67]

$$[A]_{ij} \coloneqq \det\big((A)_{\neq i,\neq j}\big) = \det\begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix} \tag{23.7}$$

(*iii*) Die Größe

$$\widetilde{a}_{ij} \coloneqq (-1)^{i+j}\,[A]_{ij} \tag{23.8}$$

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:

$$\operatorname{adj}(A) \coloneqq \operatorname{cof}(A)^{\mathsf{T}}. \tag{23.9}$$

△

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 23.13 (alternative Definition der Kofaktoren)**.**
Es sei K ein Körper und $A \in K^{n\times n}$, $n \in \mathbb{N}_0$. Für die Kofaktoren von A gilt

$$\widetilde{a}_{ij} = \det \begin{array}{c}\left(\begin{array}{ccccccc} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{array}\right) \\ \uparrow j\text{-te Spalte} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \\ \end{array} \tag{23.10a}$$

$$= \det \begin{array}{c}\left(\begin{array}{ccccccc} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{array}\right) \\ \uparrow j\text{-te Spalte} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \\ \end{array} \tag{23.10b}$$

$$= \det(a_{\bullet 1}, \ldots, a_{\bullet j-1}, e_i, a_{\bullet j+1}, \ldots, a_{\bullet n}). \tag{23.10c}$$

[67]Für Minoren gibt es in der Literatur keine einheitliche Notation.

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 (23.10a) auf die Gestalt

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ \vdots & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{bmatrix}$$

gebracht werden, deren Determinante nach (23.5) gleich $[A]_{ij}$ ist. Durch die Zeilen- und Spaltenvertauschungen kommt der Faktor $(-1)^{i+j}$ hinzu. Also ist die Determinante der Matrix in (23.10a) gleich $(-1)^{i+1}\,[A]_{ij}$, was nach (23.8) die Definition des Kofaktors $\widetilde{a}_{ij}$ bzgl. des Index (i, j) ist.

Die Matrix in (23.10a) geht dadurch aus (23.10b) 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 (23.10b) und (23.10a).

Die Übereinstimmung von (23.10c) und (23.10b) ist klar. □

Beispiel 23.14 (Streichungsmatrix, Minoren, Kofaktoren und Adjunkte)**.**
Die Matrix $A = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$ hat die Kofaktoren

$$\begin{aligned}
\widetilde{a}_{11} &= (-1)^2 [A]_{11} = 4 && \text{oder nach (23.10a)} && \widetilde{a}_{11} = \det\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = 4 \\
\widetilde{a}_{12} &= (-1)^3 [A]_{12} = -3 && \text{oder nach (23.10a)} && \widetilde{a}_{12} = \det\begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} = -3 \\
\widetilde{a}_{21} &= (-1)^3 [A]_{21} = -2 && \text{oder nach (23.10a)} && \widetilde{a}_{21} = \det\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} = -2 \\
\widetilde{a}_{22} &= (-1)^4 [A]_{22} = 1 && \text{oder nach (23.10a)} && \widetilde{a}_{22} = \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.
\end{aligned}$$

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 23.15 (Bedeutung der Adjunkten)**.**
Es sei K ein Körper und $A \in K^{n\times n}$, $n \in \mathbb{N}_0$. Dann gilt

$$\operatorname{adj}(A)\, A = A\, \operatorname{adj}(A) = \det(A)\, I. \tag{23.11}$$

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

$$\begin{aligned}
b_{ik} &= \sum_{j=1}^{n} \widetilde{a}_{ji}\, a_{jk} && \text{nach Definition des Matrix-Matrix-Produkts} \\
&= \sum_{j=1}^{n} a_{jk} \, \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, e_j, a_{\bullet i+1}, \ldots, a_{\bullet n}) && \text{nach (23.10c)} \\
&= \det\Big(a_{\bullet 1}, \ldots, a_{\bullet i-1}, \sum_{j=1}^{n} a_{jk}\, e_j, a_{\bullet i+1}, \ldots, a_{\bullet n}\Big) && \text{wegen der Multilinearität der Determinante} \\
&= \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, a_{\bullet k}, a_{\bullet i+1}, \ldots, a_{\bullet n}).
\end{aligned}$$

Dieser Ausdruck ist gleich 0, wenn $i \neq k$ ist, weil sich dann einer der Vektoren wiederholt und die Determinante alternierend ist. 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 $\operatorname{adj}(A)\,A = \det(A)\,I$ folgt, weil $\frac{1}{\det(A)}\operatorname{adj}(A)$ als Linksinverse von A gleichzeitig auch Rechtsinverse von A ist, siehe Satz 15.42. □

Beispiel 23.16 (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 gegeben durch $A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$, also durch

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \frac{1}{a_{11}\,a_{22} - a_{21}\,a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \qquad \triangle$$

Neben der Transformation auf Zeilenstufenform (Beispiel 23.11) 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 23.17 (**Entwicklungssatz von Laplace**[68])**.**
Es sei K ein Körper und $A \in K^{n \times n}$, $n \in \mathbb{N}_0$. Dann gilt:

Entwicklung nach der i-ten Zeile (englisch: expansion by the i-th row):

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\, \det\big((A)_{\neq i, \neq j}\big) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\, [A]_{ij}. \tag{23.12a}$$

Entwicklung nach der j-ten Spalte (englisch: expansion by the j-th column):

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij}\, \det\big((A)_{\neq i, \neq j}\big) = \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij}\, [A]_{ij}. \tag{23.12b}$$

[68] englisch: Laplace expansion theorem

Beweis. Nach Lemma 23.15 gilt $\det(A)\, I = A \operatorname{adj}(A)$. Für $i \in [\![1, n]\!]$ ist daher

$$\begin{aligned}
\det(A) &= \big(A \operatorname{adj}(A)\big)_{ii} \\
&= \sum_{j=1}^{n} a_{ij} \operatorname{adj}(A)_{ji} && \text{nach Definition der Matrix-Matrix-Multiplikation} \\
&= \sum_{j=1}^{n} a_{ij} \operatorname{cof}(A)_{ij} && \text{nach Definition (23.9) der Adjunkte} \\
&= \sum_{j=1}^{n} a_{ij}\, (-1)^{i+j} \det\big((A)_{\neq i, \neq j}\big) && \text{nach Definition (23.8) des Kofaktors} \\
&= \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\, [A]_{ij} && \text{nach Definition (23.7) des Minors } [A]_{ij}.
\end{aligned}$$

Damit ist (23.12a) gezeigt. (23.12b) folgt analog. □

Beispiel 23.18 (Entwicklungssatz von Laplace).
Wir betrachten die Matrix $A \in \mathbb{Q}^{3\times 3}$ mit

$$A = \begin{pmatrix} -2 & 2 & 3 \\ -1 & 1 & 1 \\ -4 & 0 & 1 \end{pmatrix}.$$

Wir entwickeln die Determinante nach der zweiten Spalte ($j = 2$), weil dort zumindest eine Null auftaucht:[69]

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{3} (-1)^{i+2}\, a_{i2} \det\big((A)_{\neq i, \neq 2}\big) \\
&= -a_{12} \det\big((A)_{\neq 1, \neq 2}\big) + a_{22} \det\big((A)_{\neq 2, \neq 2}\big) - a_{32} \det\big((A)_{\neq 3, \neq 2}\big) \\
&= -2 \det \begin{pmatrix} -1 & 1 \\ -4 & 1 \end{pmatrix} + 1 \det \begin{pmatrix} -2 & 3 \\ -4 & 1 \end{pmatrix} - 0 \det \begin{pmatrix} -2 & 3 \\ -1 & 1 \end{pmatrix} \\
&= -2\,(-1+4) + 1\,(-2+12) + 0 \\
&= -6 + 10 \\
&= 4.
\end{aligned}$$
△

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 23.19 (**Cramersche Regel**[70]).
Es sei 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 $A\,x = b$

$$x_i = \frac{\det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, b, a_{\bullet i+1}, \ldots, a_{\bullet n})}{\det(A)} \quad \text{für } i = 1, \ldots, n. \tag{23.13}$$

[69] Alternative: Entwicklung nach der dritten Zeile ($i = 3$)
[70] englisch: Cramer's rule

Beweis. Es gilt

$$\begin{aligned} x &= A^{-1} b \\ &= \frac{1}{\det(A)} \operatorname{adj}(A)\, b && \text{nach Lemma 23.15,} \\ \text{also} \quad x_i &= \frac{1}{\det(A)} \sum_{j=1}^{n} (\operatorname{adj}(A))_{ij}\, b_j. \end{aligned}$$

Für die Einträge der Adjunkten gilt weiter

$$\begin{aligned} \operatorname{adj}(A)_{ij} &= \widetilde{a}_{ji} && \text{nach Definition der Adjunkten (23.9)} \\ &= \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, e_j, a_{\bullet i+1}, \ldots, a_{\bullet n}) && \text{nach (23.10c).} \end{aligned}$$

Setzen wir diese Beziehung oben ein, so ergibt sich

$$\begin{aligned} x_i &= \frac{1}{\det(A)} \sum_{j=1}^{n} (\operatorname{adj}(A))_{ij}\, b_j \\ &= \frac{1}{\det(A)} \sum_{j=1}^{n} \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, e_j, a_{\bullet i+1}, \ldots, a_{\bullet n})\, b_j \\ &= \frac{1}{\det(A)} \det(a_{\bullet 1}, \ldots, a_{\bullet i-1}, b, a_{\bullet i+1}, \ldots, a_{\bullet n}). \end{aligned}$$

Die letzte Gleichheit gilt wegen der Multilinearität der Determinante. □

Beispiel 23.20 (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.

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 einiger 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 23.17 berechnet werden können. Abseits dieser Fälle findet die Cramersche Regel auch als Mittel in Beweisen weitere Anwendung.

Ende der Vorlesung 10

Ende der Woche 5

§ 23.2 Die Determinante eines Endomorphismus

Wir kommen zurück zu der eingangs von § 23 angeführten Motivation, die Determinante als Maßzahl für Endomorphismen $f \in \operatorname{End}(V)$ einzuführen.

Definition 23.21 (Determinante auf dem Raum $\operatorname{End}(V)$ der Endomorphismen)**.**
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B_V = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$. Die **Determinante** eines Endomorphismus $f \in \operatorname{End}(V)$ ist definiert als

$$\det(f) := \det(A) \tag{23.14}$$

für die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f)$. △

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, denn ist $\widehat{B}_V$ eine weitere Basis von V, dann gilt nach Satz 20.12 für die Darstellungsmatrix $\widehat{A} = \mathcal{M}_{\widehat{B}_V}^{\widehat{B}_V}(f)$

$$\widehat{A} = T\,A\,T^{-1}$$

mit der Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V}^{B_V}$. Da A und $\widehat{A}$ ähnliche Matrizen sind, folgt $\det(A) = \det(\widehat{A})$ aus Lemma 23.9. Daher ist tatsächlich unerheblich, welche Basis für die Darstellungsmatrix in der Definition (23.14) der Determinante eines Endomorphismus verwendet wird.

Lemma 23.22 (Eigenschaften der Determinante für Endormorphismen, vgl. Lemma 23.8)**.**
Es seien K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter seien $f, g \in \operatorname{End}(f)$. Die Determinante $\det\colon \operatorname{End}(f) \to K$ hat folgende Eigenschaften:

- (i) $\det(\alpha\, f) = \alpha^n \det(f)$ für alle $\alpha \in K$.
- (ii) $\det(f) \neq 0 \Leftrightarrow f$ ist invertierbar $\Leftrightarrow \operatorname{Rang}(f) = n \Leftrightarrow$ Jede Darstellungsmatrix von f ist invertierbar.
- (iii) $\det(\operatorname{id}_V) = 1$.
- (iv) $\det(f \circ g) = \det(f)\,\det(g)$.
- (v) $\det(f^{-1}) = 1/\det(f)$, falls f invertierbar ist.
- (vi) $\det(f^*) = \det(f)$ für die zu f duale Abbildung $f^* \in \operatorname{End}(V^*)$.

Beachte: Statt „f ist invertierbar" können wir auch „f ist ein Automorphismus", also „$f \in \operatorname{Aut}(V)$" schreiben.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-6.1. □

§ 23.3 Orientierung eines Vektorraumes

Für den Begriff der Orientierung benötigen wir zunächst den Begriff des geordneten Körpers. Dabei handelt es sich um einen Körper mit einer Totalordnung, die in gewissem Sinne mit den Verknüpfungen der Körpers verträglich ist.

Definition 23.23 (geordneter Körper).
Es sei $(K, +, \cdot)$ ein Körper und $\leqslant$ eine Totalordnung auf K.

(*i*) Der Körper heißt **geordnet** (englisch: ordered field) bzgl. der Totalordnung $\leqslant$, wenn gilt:

$$\alpha \leqslant \beta \quad \Rightarrow \quad \alpha + \gamma \leqslant \beta + \gamma \tag{23.15a}$$

$$\alpha \geqslant 0 \text{ und } \beta \geqslant 0 \quad \Rightarrow \quad \alpha \cdot \beta \geqslant 0 \tag{23.15b}$$

jeweils für $\alpha, \beta, \gamma \in K$.

(*ii*) $\alpha \in K$ heißt **nichtnegativ**, wenn $\alpha \geqslant 0$ ist.

(*iii*) $\alpha \in K$ heißt **positiv**, wenn $\alpha \geqslant 0$ und $\alpha \neq 0$ ist.

(*iv*) $\alpha \in K$ heißt **nichtpositiv**, wenn $\alpha \leqslant 0$ ist.

(*v*) $\alpha \in K$ heißt **negativ**, wenn $\alpha \leqslant 0$ und $\alpha \neq 0$ ist. △

Lemma 23.24 (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*) $\alpha \geqslant 0 \quad \Leftrightarrow \quad -\alpha \leqslant 0$

(*ii*) $\alpha \leqslant \beta$ und $\gamma \leqslant \delta \quad \Rightarrow \quad \alpha + \gamma \leqslant \beta + \delta$

(*iii*) $\alpha \leqslant \beta$ und $\gamma \geqslant 0 \quad \Rightarrow \quad \alpha\,\gamma \leqslant \beta\,\gamma$

(*iv*) $\alpha \leqslant \beta$ und $\gamma \leqslant 0 \quad \Rightarrow \quad \beta\,\gamma \leqslant \alpha\,\gamma$

(*v*) $\alpha^2 \geqslant 0$

(*vi*) $\alpha \neq 0 \quad \Rightarrow \quad \alpha^2 > 0$

(*vii*) $\alpha > 0 \quad \Rightarrow \quad \frac{1}{\alpha} > 0$

(*viii*) $\beta > \alpha > 0 \quad \Rightarrow \quad \frac{1}{\alpha} > \frac{1}{\beta} > 0$

(*ix*) $n\,1 > 0$ für alle $n \in \mathbb{N}$, insbesondere hat K notwendig $\operatorname{char}(K) = 0$.

Beweis. Aussage (*i*):

$$\begin{aligned} \alpha \geqslant 0 \quad &\Rightarrow \quad 0 = \alpha + (-\alpha) \geqslant -\alpha \\ \text{und} \quad -\alpha \leqslant 0 \quad &\Rightarrow \quad 0 = -\alpha + \alpha \leqslant \alpha. \end{aligned}$$

Aussage (*ii*):

$$\begin{aligned} \alpha \leqslant \beta \quad &\Rightarrow \quad \alpha + \gamma \leqslant \beta + \gamma \\ \text{und} \quad \gamma \leqslant \delta \quad &\Rightarrow \quad \beta + \gamma \leqslant \beta + \delta. \end{aligned}$$

Transitivität zeigt, dass dann auch $\alpha + \gamma \leqslant \beta + \delta$ gilt.

Aussage (iii):

$$\begin{aligned} & \alpha \leqslant \beta \\ \Rightarrow \quad & \beta - \alpha \geqslant 0 \\ \Rightarrow \quad & (\beta - \alpha)\,\gamma \geqslant 0 \\ \Rightarrow \quad & \beta\,\gamma - \alpha\,\gamma \geqslant 0 \\ \Rightarrow \quad & \alpha\,\gamma \leqslant \beta\,\gamma. \end{aligned}$$

Aussage (iv): Wegen Aussage (i) ist $-\gamma \geqslant 0$. Aus Aussage (iii) folgt $\alpha\,(-\gamma) \leqslant \beta\,(-\gamma)$, also $\beta\,\gamma \leqslant \alpha\,\gamma$.

Aussage (v): Nehmen wir zunächst $\alpha \geqslant 0$ an. Es folgt $\alpha^2 = \alpha \cdot \alpha \geqslant 0$. Im Fall $\alpha \leqslant 0$ ist $-\alpha \geqslant 0$. Es folgt $\alpha^2 = (-\alpha) \cdot (-\alpha) \geqslant 0$.

Aussage (vi): Die Behauptung folgt aus der Nullteilerfreiheit des Körpers K.

Aussage (vii): Wegen $\alpha > 0$ ist $\alpha \neq 0$, also multiplikativ invertierbar. Das Inverse $\frac{1}{\alpha}$ ist ebenfalls $\neq 0$. Die Annahme $\frac{1}{\alpha} < 0$ würde $1 = \alpha \cdot \frac{1}{\alpha} < 0$ ergeben. Andererseits ist aber $1 = 1 \cdot 1 > 0$ nach Aussage (vi), Widerspruch. Also muss $\frac{1}{\alpha} > 0$ gelten.

Aussage $(viii)$:

$$\begin{aligned} & \beta > \alpha > 0 \\ \Rightarrow \quad & 1 > \frac{\alpha}{\beta} > 0 \quad \text{wegen } \frac{1}{\beta} > 0 \text{ nach Aussage } (vii) \\ \Rightarrow \quad & \frac{1}{\alpha} > \frac{1}{\beta} > 0 \quad \text{wegen } \frac{1}{\alpha} > 0 \text{ nach Aussage } (vii) \end{aligned}$$

Aussage (ix): Aus Aussage (vi) folgt $1 > 0$. Es folgt weiter $1 + 1 > 1 + 0 = 1 > 0$. Induktiv zeigt man $n\,1 > 0$ für alle $n \in \mathbb{N}$. Folglich ist $n\,1 \neq 0$ für all $n \in \mathbb{N}$, d. h., $\operatorname{char}(K) = 0$. □

Beispiel 23.25 (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.

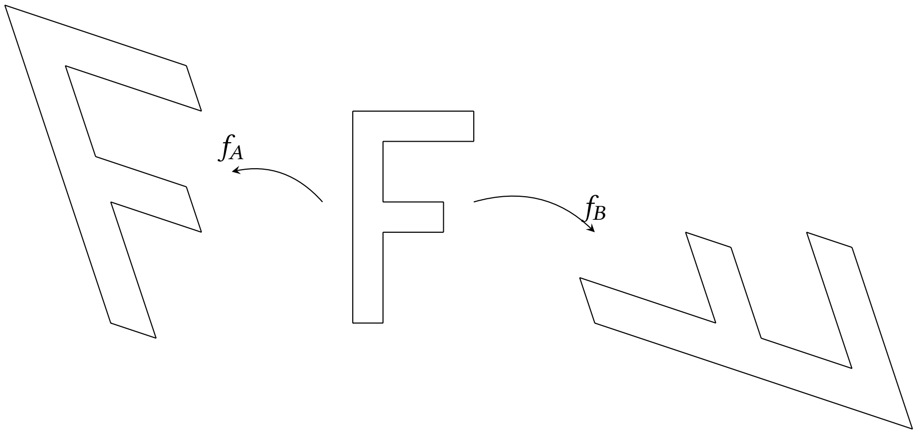
(iii) Die komplexen Zahlen $\mathbb{C}$ sind mit keiner Totalordnung ein geordneter Körper, da $i^2 = -1$ den Aussagen (i) und (vi) aus Lemma 23.24 widerspricht. △

Definition 23.26 (orientierungstreuer Endomorphismus).
Es seien K ein geordneter Körper und V ein endlich-dimensionaler Vektorraum über K. Ein Automorphismus $f \in \operatorname{Aut}(V)$ heißt **orientierungstreu** oder **orientierungserhaltend** (englisch: orientation preserving) im Fall $\det(f) > 0$ und **orientierungsuntreu** oder **orientierungsumkehrend** (englisch: orientation reversing) im Fall $\det(f) < 0$. △

Beispiel 23.27 (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 23.28 (gleichorientierte Basen)**.**
Es seien K ein geordneter Körper und V ein endlich-dimensionaler Vektorraum über K.

(i) Zwei Basen $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ heißen **gleich orientiert** (englisch: equally oriented), wenn die Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V}^{B_V}$ die Bedingung $\det(T) > 0$ erfüllt.

(ii) Zwei Basen $B_V = (v_1, \ldots, v_n)$ und $\widehat{B}_V = (\widehat{v}_1, \ldots, \widehat{v}_n)$ heißen **umgekehrt orientiert** (englisch: differently oriented), wenn die Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V}^{B_V}$ die Bedingung $\det(T) < 0$ erfüllt. △

Lemma 23.29 (Gleichorientierung ist 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 $T = \mathcal{T}_{\widehat{B}_V}^{B_V}$ und $\det(T) > 0$ auch $T^{-1} = \mathcal{T}_{B_V}^{\widehat{B}_V}$ 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 $\operatorname{char}(K) \neq 2$ gilt.

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

Kapitel 5 Normalformen von Endomorphismen

§ 24 Grundlagen zu Eigenwerten und Eigenvektoren

Literatur: Fischer, Springborn, 2020, Kapitel 5, Bosch, 2014, Kapitel 6

Es sei V ein Vektorraum über dem Körper K.

Zur Erinnerung: Ein Vektor $v \in V \setminus \{0\}$ heißt ein **Eigenvektor** zum **Eigenwert** $\lambda \in K$ des Endomorphismus $f \in \operatorname{End}(V)$, wenn $f(v) = \lambda\, v$ gilt. In diesem Fall heißt $(\lambda, v) \in K \times (V \setminus \{0\})$ auch ein **Eigenpaar** von f, siehe Definition 20.19. Ebenso heißt ein Vektor $x \in K^n \setminus \{0\}$ ein **Eigenvektor** zum **Eigenwert** $\lambda \in K$ der Matrix $A \in K^{n \times n}$, wenn $A\,x = \lambda\,x$ gilt. In diesem Fall heißt $(\lambda, v) \in K \times (K^n \setminus \{0\})$ auch ein **Eigenpaar** von A, siehe Definition 20.19.

Wir erinnern auch an die Frage aus § 20, welche Wahl der Basis in einem endlich-dimensionalen Vektorraum V dazu führt, dass die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f)$ eines Endomorphismus $f \in \operatorname{End}(V)$ eine möglichst einfache Struktur hat. Mit anderen Worten, wir suchen nach einer Normalform der Darstellungsmatrizen von Endomorphismen. Wir hatten gesehen, dass dabei f-invariante Unterräume eine wesentliche Rolle spielen, also Unterräume $U \subseteq V$ mit der Eigenschaft $f(U) \subseteq U$. Eindimensionale f-invariante Unterräume sind genau diejenigen, die jeweils durch einen Eigenvektor aufgespannt werden.

Die Diagonalisierbarkeit von f führt auf den einfachsten Fall einer Darstellungsmatrix. Ein Endomorphismus eines Vektorraumes V mit $\dim(V) = n$ heißt **diagonalisierbar** (Definition 20.23), wenn es f-invariante Unterräume $U_1, \dots, U_n$ der Dimension 1 gibt, sodass $V = U_1 \oplus \dots \oplus U_n$ gilt. Es stellt sich die Frage, wann dieser Fall eintritt, also die Frage nach einem Diagonalisierbarkeitskriterium.

Nach Satz 20.24 ist ein Endomorphismus genau dann diagonalisierbar, wenn V eine Basis besitzt, die nur aus Eigenvektoren von f besteht. Das ist wiederum genau dann der Fall, wenn eine (und damit alle) Darstellungsmatrizen A von f ähnlich zu einer Diagonalmatrix (nämlich zur Diagonalmatrix mit den Eigenwerten) ist:

$$X^{-1} A\, X = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} =: D \tag{24.1a}$$

oder äquivalent

$$A\, X = X \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = XD. \tag{24.1b}$$

Spaltenweise gelesen bedeutet die Gleichung (24.1b), dass die j-te Spalte der Matrix

$$X = \begin{bmatrix} x_{\bullet 1} & \cdots & x_{\bullet n} \end{bmatrix} \in K^{n \times n}$$

ein Eigenvektor zum Eigenwert λ_j der Matrix $A \in K^{n\times n}$ ist. Ist $A = \mathcal{M}_{B_V}^{B_V}(f)$ die Darstellungsmatrix von f bzgl. der Basis $B_V = (v_1, \ldots, v_n)$, dann ist nach Lemma 20.20

$$\Phi_{B_V}(x_{\bullet j}) = \sum_{i=1}^{n} x_{ij}\, v_i \tag{24.1c}$$

ein Eigenvektor von f zum Eigenwert λ_j. Die Spalten von X enthalten also die Koordinaten der Eigenvektoren bzgl. der Basis B_V.

Wir untersuchen zunächst die Bedeutung paarweise verschiedener Eigenwerte.

Satz 24.1 (paarweise verschiedene Eigenwerte)**.**
Es sei V ein Vektorraum über dem Körper K. Weiter sei $f \in \operatorname{End}(V)$.

(i) Ist $k \in \mathbb{N}_0$ und sind $(v_1, \ldots, v_k)$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k \in K$, so ist die Familie $(v_1, \ldots, v_k)$ linear unabhängig. Insbesondere gilt $k \leqslant \dim(V)$.

(ii) Gilt $\dim(V) = n \in \mathbb{N}_0$, so hat f höchstens n paarweise verschiedene Eigenwerte.

Beweis. Aussage (i): Wir führen einen Induktionsbeweis. Für $k = 0$ ist nichts zu zeigen. Für $k = 1$ ist die Familie (v_1) linear unabhängig, da der Eigenvektor $v_1 \neq 0$ ist. Wir schließen nun von k auf $k + 1$. Es seien also $\lambda_1, \ldots, \lambda_{k+1}$ paarweise verschiedene Eigenvektoren von f und $v_1, \ldots, v_{k+1}$ zugehörige Eigenvektoren. Wir setzen die Linearkombination

$$\alpha_1\, v_1 + \cdots + \alpha_k\, v_k + \alpha_{k+1}\, v_{k+1} = 0 \tag{24.2}$$

an. Durch Anwendung von f erhalten wir

$$f(\alpha_1\, v_1 + \cdots + \alpha_k\, v_k + \alpha_{k+1}\, v_{k+1}) = \alpha_1\, \lambda_1\, v_1 + \cdots + \alpha_k\, \lambda_k\, v_k + \alpha_{k+1}\, \lambda_{k+1}\, v_{k+1} = 0.$$

Die Subtraktion des λ_{k+1}-Fachen von (24.2) von dieser Gleichung ergibt

$$\alpha_1\, (\lambda_1 - \lambda_{k+1})\, v_1 + \cdots + \alpha_k\, (\lambda_k - \lambda_{k+1})\, v_k + \alpha_{k+1}\, \underbrace{(\lambda_{k+1} - \lambda_{k+1})}_{=0}\, v_{k+1} = 0.$$

Da die k Eigenvektoren $v_1, \ldots, v_k$ aufgrund der Induktionsannahme linear unabhängig sind, folgt

$$\alpha_1\, (\lambda_1 - \lambda_{k+1}) = \ldots = \alpha_k\, (\lambda_k - \lambda_{k+1}) = 0.$$

Da die Eigenwerte paarweise verschieden sind, muss $\alpha_1 = \ldots = \alpha_k = 0$ gelten. Aus der ersten Gleichung folgt nun $\alpha_{k+1}\, v_{k+1} = 0$ und wegen $v_{k+1} \neq 0$ schließlich auch $\alpha_{k+1} = 0$. Das heißt, dass die Familie $(v_1, \ldots, v_{k+1})$ linear unabhängig ist.

Aussage (ii) folgt sofort aus Aussage (i). □

Als Folgerung von Satz 24.1 erhalten wir ein hinreichendes Kriterium für die Diagonalisierbarkeit:

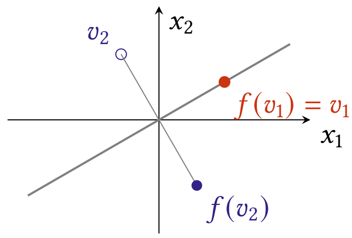
Folgerung 24.2 (hinreichendes Kriterium für Diagonalisierbarkeit)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Besitzt $f \in \operatorname{End}(V)$ n paarweise verschiedene Eigenwerte, dann ist f diagonalisierbar.

Beweis. Es seien $(v_1, \dots, v_n)$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$. Nach Satz 24.1 Aussage (i) ist die Familie $(v_1, \dots, v_n)$ linear unabhängig, also eine Basis von V, die aus lauter Eigenvektoren von f besteht. Nach Satz 20.24 ist f diagonalisierbar. □

Beispiel 24.3 (paarweise verschiedene Eigenwerte)**.**
Wie wir in Beispiel 20.16 gesehen haben, bildet die Spiegelungsabbildung auf $\mathbb{R}^2$ mit der Darstellungsmatrix bzgl. der Standardbasis

$$A = \begin{bmatrix} \cos^2(\alpha) - \sin^2(\alpha) & 2\cos(\alpha)\,\sin(\alpha) \\ 2\cos(\alpha)\,\sin(\alpha) & \sin^2(\alpha) - \cos^2(\alpha) \end{bmatrix}$$

den Vektor $v_1 = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$ (und Vielfache davon) auf v_1 ab und
den Vektor $v_2 = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$ (und Vielfache davon) auf $-v_2$.



Das bedeutet, dass der zugehörige Endomorphismus $f_A \in \operatorname{End}(\mathbb{R}^2)$ die beiden Eigenwerte ± 1 besitzt. Nach Folgerung 24.2 ist f_A also diagonalisierbar. Die Darstellungsmatrix von f_A bzgl. der Basis $B = (v_1, v_2)$ aus Eigenvektoren ist folglich diagonal:

$$\mathcal{M}_B^B(f_A) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \qquad \triangle$$

Wie gleich in Lemma 24.10 am Beispiel von Projektoren klar werden wird, ist das Kriterium von n paarweise verschiedener Eigenwerte nur hinreichend, aber nicht notwendig.

§ 24.1 Berechnung von Eigenvektoren zu gegebenem Eigenwert

Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Bevor wir der Frage nach der Bestimmung von Eigenwerten nachgehen, klären wir, wie zu gegebenem Eigenwert $\lambda \in K$ ein zugehöriger Eigenvektor $v \in V$ des Endomorphismus $f \in \operatorname{End}(V)$ gefunden werden kann. Es gilt

$$\begin{aligned} & f(v) = \lambda\, v \\ \Leftrightarrow\quad & \lambda\, v - f(v) = 0 \\ \Leftrightarrow\quad & (\lambda \operatorname{id}_V - f)(v) = 0 \\ \Leftrightarrow\quad & v \in \operatorname{Kern}(\lambda \operatorname{id}_V - f). \end{aligned} \tag{24.3}$$

Der Vektor $v \in V$ ist also ein Eigenvektor des Endomorphismus f zum Eigenwert $\lambda \in K$ genau dann, wenn $v \neq 0$ ist und $v \in \operatorname{Kern}(\lambda \operatorname{id}_V - f)$ gilt.

Wir können damit insbesondere ein Kriterium formulieren, wann ein Endomorphismus injektiv bzw. bijektiv ist:

Lemma 24.4 (Invertierbarkeitskriterium für einen Endomorphismus mit Eigenwerten)**.**
Es sei V ein Vektorraum über dem Körper K und $f \in \operatorname{End}(V)$. Dann gilt:

(i) f ist injektiv $\Leftrightarrow$ 0 ist kein Eigenwert von f.

(ii) Ist V endlich-dimensional, dann gilt sogar: f ist bijektiv $\Leftrightarrow$ 0 ist kein Eigenwert von f.

Beweis. Aussage (i): Nach Lemma 17.6 ist f genau dann injektiv, wenn $\operatorname{Kern}(f) = \{0\}$ gilt. Wegen $\operatorname{Kern}(f) = \operatorname{Kern}(0\,\mathrm{id}_V - f)$ ist $\operatorname{Kern}(f) = \{0\}$ genau dann der Fall, wenn 0 kein Eigenwert von f ist.

Aussage (ii): Wenn V endliche Dimension hat, dann ist nach Folgerung 18.9 die Injektivität von f bereits hinreichend für die Bijektivität. □

Die Menge aller Eigenwerte eines Endomorphismus bezeichnen wir mit

$$\Lambda(f) \coloneqq \{\lambda \in K \mid \lambda \text{ ist ein Eigenwert von } f\} \subseteq K. \tag{24.4}$$

Wir können analoge Aussagen auch für Matrizen $A \in K^{n\times n}$ formulieren. Es gilt

$$\begin{aligned} & A\,x = \lambda\,x \\ \Leftrightarrow\quad & \lambda\,x - A\,x = 0 \\ \Leftrightarrow\quad & (\lambda\,I - A)(x) = 0 \\ \Leftrightarrow\quad & x \in \operatorname{Kern}(\lambda\,I - A). \end{aligned} \tag{24.5}$$

Der Vektor $x \in K^n \setminus \{0\}$ ist also genau dann ein Eigenvektor der Matrix $A \in K^{n\times n}$ zum Eigenwert $\lambda \in K$, wenn $x \neq 0$ ist und $x \in \operatorname{Kern}(\lambda\,I - A)$ gilt.

Die Bestimmung von Eigenvektoren zu gegebenem Eigenwert $\lambda \in K$ wird damit zu einer Aufgabe, die Lösungsmenge des homogenen linearen Gleichungssystems

$$(\lambda\,I - A)\,x = 0$$

zu bestimmen.

Lemma 24.5 (Invertierbarkeitskriterium für eine Matrix mit Eigenwerten)**.**
Es sei K ein Körper und $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$. Dann sind äquivalent:

(i) A ist regulär.

(ii) 0 ist kein Eigenwert von A.

Beweis. □

Die Menge aller Eigenwerte einer Matrix bezeichnen wir mit

$$\Lambda(A) \coloneqq \{\lambda \in K \mid \lambda \text{ ist ein Eigenwert von } A\} \subseteq K. \tag{24.6}$$

Ist $\lambda \in K$ ein Eigenwert von $f \in \operatorname{End}(V)$, dann ist per Definition $\operatorname{Kern}(\lambda\,\mathrm{id}_V - f)$ ein Unterraum, der mindestens die Dimension 1 besitzt.

Definition 24.6 (Eigenraum, geometrische Vielfachheit)**.**
Es sei V ein Vektorraum über dem Körper K.

(*i*) Ist $f \in \mathrm{End}(V)$ und $\lambda \in K$, dann heißt

$$\mathrm{Eig}(f,\lambda) := \{v \in V \mid f(v) = \lambda\, v\} \tag{24.7a}$$

der **Eigenraum** (englisch: eigenspace) von f zu λ. Die Kardinalzahl

$$\mu^{\mathrm{geo}}(f,\lambda) := \dim(\mathrm{Eig}(f,\lambda)) \tag{24.7b}$$

heißt die **geometrische Vielfachheit** (englisch: geometric multiplicity) von λ.

(*ii*) Ist $A \in K^{n\times n}$ und $\lambda \in K$, dann heißt

$$\mathrm{Eig}(A,\lambda) := \{v \in V \mid A\, v = \lambda\, v\} \tag{24.8a}$$

der **Eigenraum** von A zu λ. Die Kardinalzahl

$$\mu^{\mathrm{geo}}(A,\lambda) := \dim(\mathrm{Eig}(A,\lambda)) \tag{24.8b}$$

heißt die **geometrische Vielfachheit** von λ. △

Beachte: Wir sprechen also auch dann von Eigenräumen, wenn λ kein Eigenwert von f bzw. A ist.

Lemma 24.7 (Eigenschaften von Eigenräumen von Endomorphismen)**.**
Es sei V ein Vektorraum über dem Körper K und $f \in \mathrm{End}(V)$ sowie $\lambda \in K$. Dann gilt:

(*i*) $\mathrm{Eig}(f,\lambda)$ ist ein Unterraum von V.

(*ii*) $\mathrm{Eig}(f,\lambda) \neq \{0\} \Leftrightarrow \lambda$ ist ein Eigenwert von f.

(*iii*) $\mathrm{Eig}(f,\lambda) \setminus \{0\}$ ist die Menge der zu λ gehörenden Eigenvektoren von f.

(*iv*) $\mathrm{Eig}(f,\lambda) = \mathrm{Kern}(\lambda\, \mathrm{id}_V - f)$.

(*v*) Sind $\lambda_1, \lambda_2 \in K$ verschieden, dann gilt $\mathrm{Eig}(f,\lambda_1) \cap \mathrm{Eig}(f,\lambda_2) = \{0\}$.

Beachte: Insbesondere ist $\mathrm{Kern}(f)$ der Eigenraum zu $0 \in K$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-6.3. □

Analog gilt für Matrizen:

Lemma 24.8 (Eigenschaften von Eigenräumen von Matrizen)**.**
Es sei K ein Körper und $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$ sowie $\lambda \in K$. Dann gilt:

(*i*) $\mathrm{Eig}(A,\lambda)$ ist ein Unterraum von V.

(*ii*) $\mathrm{Eig}(A,\lambda) \neq \{0\} \Leftrightarrow \lambda$ ist ein Eigenwert von A.

(*iii*) $\mathrm{Eig}(A,\lambda) \setminus \{0\}$ ist die Menge der zu λ gehörenden Eigenvektoren von A.

(*iv*) $\mathrm{Eig}(A,\lambda) = \mathrm{Kern}(\lambda\, I - A)$.

(*v*) Sind $\lambda_1, \lambda_2 \in K$ verschieden, dann gilt $\mathrm{Eig}(A,\lambda_1) \cap \mathrm{Eig}(A,\lambda_2) = \{0\}$.

Beachte: Insbesondere ist $\mathrm{Kern}(A)$ der Eigenraum zu $0 \in K$.

Beweis. wie Lemma 24.7 □

Das folgende Resultat wurde nachträglich eingefügt. Die Aussage (*v*) von Lemma 24.7 und Lemma 24.8 bedeutet, dass die Summe $\operatorname{Eig}(A, \lambda_1) \oplus \operatorname{Eig}(A, \lambda_2)$ direkt ist. Tatsächlich gilt das sogar auch für mehr als zwei Eigenräume:

Lemma 24.9 (Eigenräume zu verschiedenen Eigenwerten bilden eine direkte Summe)**.**
Es sei K ein Körper.

(*i*) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Sind $\lambda_1, \ldots, \lambda_s$, $s \in \mathbb{N}_0$, paarweise verschiedene Eigenwerte von A, dann ist die Summe der Unterräume $\operatorname{Eig}(A, \lambda_1), \ldots, \operatorname{Eig}(A, \lambda_s)$ direkt.

(*ii*) Es sei V ein Vektorraum über dem Körper K und $f \in \operatorname{End}(V)$. Sind $\lambda_1, \ldots, \lambda_s$, $s \in \mathbb{N}_0$, paarweise verschiedene Eigenwerte von f, dann ist die Summe der Unterräume $\operatorname{Eig}(f, \lambda_1), \ldots, \operatorname{Eig}(f, \lambda_s)$ direkt.

Beweis. Aussage (*i*): Gemäß Definition 14.12 müssen wir zeigen: Für jedes $j \in [\![1, s]\!]$ gilt:

$$\operatorname{Eig}(A, \lambda_j) \cap \sum_{\substack{i=1 \\ i \neq j}}^{s} \operatorname{Eig}(A, \lambda_i) = \{0\}.$$

Es sei also $j \in [\![1, s]\!]$ und $v_j \in \operatorname{Eig}(A, \lambda_j) \cap \sum_{\substack{i=1 \\ i \neq j}}^{s} \operatorname{Eig}(A, \lambda_i)$. Das heißt, es gibt Vektoren $v_i \in \operatorname{Eig}(A, \lambda_i)$ für $i \in [\![1, s]\!] \setminus \{j\}$ mit der Eigenschaft

$$v_j = \sum_{\substack{i=1 \\ i \neq j}}^{s} v_i = \sum_{\substack{i=1 \\ i \neq j \\ v_i \neq 0}}^{s} v_i.$$

Nach Lemma 13.3 bedeutet das aber, dass die Menge $\{v_j\} \cup \{v_i \mid i \in [\![1, s]\!] \setminus \{j\},\; v_i \neq 0\}$ linear abhängig ist. Wenn $v_j \neq 0$ wäre, dann wären die Elemente dieser Menge alles Eigenvektoren zu paarweise verschiedenen Eigenwerten, was nach Satz 24.1 unmöglich ist. Also muss $v_j = 0$ sein.

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

Eine besonders einfache Klasse von Abbildungen, deren Diagonalisierbarkeit wir jetzt schon zeigen können, ist die der Projektoren (Definition 22.33), also der Endomorphismen mit $P \circ P = P$:

Lemma 24.10 (Projektoren sind diagonalisierbar)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $P \in \operatorname{End}(V)$ ein Projektor und $r \coloneqq \operatorname{Rang}(P)$.

(*i*) $V = \operatorname{Kern}(P) \oplus \operatorname{Bild}(P)$.

(*ii*) Für die Eigenwerte von P gilt $\Lambda(P) \subseteq \{0, 1\}$.

(*iii*) Jedes $v \in \operatorname{Bild}(P)$, $v \neq 0$, ist ein Eigenvektor zum Eigenwert 1. Für die geometrische Vielfachheit gilt $\mu^{\text{geo}}(f, 1) = r = \operatorname{Rang}(P) = \dim(\operatorname{Bild}(P))$.

(*iv*) Jedes $v \in \operatorname{Kern}(P)$, $v \neq 0$, ist ein Eigenvektor zum Eigenwert 0. Für die geometrische Vielfachheit gilt $\mu^{\text{geo}}(f, 0) = n - r = \operatorname{Defekt}(P) = \dim(\operatorname{Kern}(P))$.

(*v*) Wählen wir eine Basis von V so, dass

$$B_V = (\underbrace{v_1, \dots, v_r}_{\text{Basis von Bild}(P)}, \underbrace{v_{r+1}, \dots, v_n}_{\text{Basis von Kern}(P)})$$

gilt, dann hat die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(P)$ die Diagonalgestalt

$$A = \begin{array}{cl} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow r\text{-te Zeile} \\ \leftarrow (r+1)\text{-te Zeile} \\ \\ \\ \end{array} \end{array}$$

Beachte: Projektoren haben also höchstens zwei verschiedene Eigenwerte und sind dennoch immer diagonalisierbar. Das zeigt, dass das Diagonalisierbarkeitskriterium aus Folgerung 24.2 nur hinreichend aber nicht notwendig ist.

Beweis. Aussage (*i*): Wir verwenden zum Beweis von $V = \text{Kern}(P) \oplus \text{Bild}(P)$ die Charakterisierung direkter Summen aus Satz 14.7 Aussage (*iii*). Aus dem Dimensionssatz für lineare Abbildungen (Folgerung 18.7) folgt

$$\dim(V) = \dim(\text{Bild}(P)) + \dim(\text{Kern}(P)),$$

es ist also $\dim(\text{Kern}(P)) = n - r$. Es bleibt $\text{Bild}(P) \cap \text{Kern}(P) = \{0\}$ zu zeigen. Es sei dazu $v \in \text{Bild}(P) \cap \text{Kern}(P)$. Dann gibt es ein $u \in V$ mit $v = P(u)$, und außerdem gilt $P(v) = 0$. Es folgt

$$v = P(u) = P(P(u)) = P(v) = 0,$$

also gilt tatsächlich $\text{Bild}(P) \cap \text{Kern}(P) = \{0\}$.

Aussage (*ii*): Es sei $v \neq 0$ ein Eigenvektor von P zum unbekannten Eigenwert $\lambda \in K$. Dann gilt

$$\begin{aligned} \text{einerseits} \quad & P(v) = \lambda\, v \\ \text{und andererseits} \quad & P(v) = P(P(v)) = P(\lambda\, v) = \lambda\, P(v) = \lambda^2\, v. \end{aligned}$$

Also gilt wegen $v \neq 0$ für den Eigenwert $\lambda = \lambda^2$ oder äquivalent $\lambda^2 - \lambda = \lambda\,(\lambda - 1) = 0$. Nach Satz 11.21 gilt für die Nullstellen dieses Polynoms also $\lambda = 0$ oder $\lambda = 1$.

Aussage (*iii*): Es sei $v \in \text{Bild}(P)$, also $v = P(u)$ für ein $u \in V$. Dann gilt

$$P(v) = P(P(u)) = P(u) = v.$$

Also ist v ein Eigenvektor von P zum Eigenwert 1, sofern $v \neq 0$ ist.

Damit gilt $\text{Bild}(P) \subseteq \text{Eig}(P, 1)$, und umgekehrt gilt auch $\text{Eig}(P, 1) \subseteq \text{Bild}(P)$. (**Quizfrage 24.1:** Warum?) Also folgt $\mu^{\text{geo}}(P, 1) = \dim(\text{Bild}(P)) = \text{Rang}(P) = r$.

Aussage (*iv*): Nach Lemma 24.7 ist $\text{Kern}(P) = \text{Eig}(P, 0)$. Also folgt $\mu^{\text{geo}}(P, 0) = \dim(\text{Kern}(P)) = \text{Defekt}(P) = n - r$.

Aussage (*v*) folgt sofort aus der Definition von Darstellungsmatrizen, vgl. (19.3). □

Ende der Vorlesung 12

Ende der Woche 6

§ 24.2 Berechnung von Eigenwerten, charakteristisches Polynom

Wie können wir nun die unbekannten Eigenwerte eines Endomorphismus bzw. einer Matrix bestimmen? Mit anderen Worten, für welche $\lambda \in K$ besteht $\operatorname{Eig}(f, \lambda)$ aus mehr als nur dem Nullvektor?

Lemma 24.11 (Eigenwerte eines Endomorphismus)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f \in \operatorname{End}(V)$ und $\lambda \in K$. Dann sind äquivalent:

(i) λ ist ein Eigenwert von f.

(ii) $\det(\lambda \operatorname{id}_V - f) = 0$.

Beweis. Es gilt:

$$\begin{aligned} & \lambda \in K \text{ ist ein Eigenwert von } f \\ \Leftrightarrow\quad & \operatorname{Eig}(f, \lambda) = \operatorname{Kern}(\lambda \operatorname{id}_V - f) \neq \{0\} && \text{nach Lemma 24.7} \\ \Leftrightarrow\quad & \lambda \operatorname{id}_V - f \text{ ist nicht bijektiv} && \text{nach Folgerung 18.9} \\ \Leftrightarrow\quad & \det(\lambda \operatorname{id}_V - f) = 0 && \text{nach Lemma 23.22.} \end{aligned}$$

□

Analog gilt für die Eigenwerte von Matrizen:

Lemma 24.12 (Eigenwerte einer Matrix)**.**
Es sei K ein Körper, $A \in K^{n \times n}$ und $\lambda \in K$. Dann sind äquivalent:

(i) λ ist ein Eigenwert von A.

(ii) $\det(\lambda I - A) = 0$.

Beweis. wie Lemma 24.11 □

Welche Art Ausdruck verbirgt sich hinter $\det(\lambda I - A)$ mit der Unbekannten λ? Es gilt **In (24.9) wurden einige Vorzeichenfehler korrigiert.**

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & \lambda - a_{nn} \end{pmatrix}. \tag{24.9}$$

Das Ausschreiben dieser Determinante mit Hilfe der Leibniz-Formel (23.3), also

$$\det(\lambda I - A) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \prod_{j=1}^{n} \big(\delta_{\sigma(j),j}\, \lambda - a_{\sigma(j),j}\big) \tag{24.10}$$

zeigt, dass es sich dabei um ein Polynom in der Variablen λ handelt.

Definition 24.13 (charakteristisches Polynom-einer-Matrix).
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Der Ausdruck (24.9) heißt das **charakteristische Polynom** (englisch: characteristic polynomial) der Matrix A, geschrieben χ_A. △

Lemma 24.14 (Eigenwerte von Matrizen sind Nullstellen des charakteristischen Polynoms).
Es sei K ein Körper, $A \in K^{n \times n}$, $n \in \mathbb{N}_0$ und $\lambda \in K$. Dann sind äquivalent:

(i) λ ist ein Eigenwert von A.

(ii) λ ist eine Nullstelle des charakteristischen Polynoms χ_A.

Beweis. Das Resultat folgt sofort aus Lemma 24.12. □

Gemäß Satz 11.21 besitzt das charakteristische Polynom χ_A die Darstellung

$$\chi_A = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s} \cdot q \tag{24.11}$$

mit den $s \in \mathbb{N}_0$ paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_s \in K$, Exponenten $n_1, \ldots, n_s \in \mathbb{N}$ und einem Polynom $q \in K[\lambda]$, das keine Nullstellen in K besitzt.

Definition 24.15 (algebraische Vielfachheit des Eigenwertes einer Matrix).
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$.

(i) Die Zahlen $n_i \in \mathbb{N}$ in der (bis auf die Reihenfolge der Faktoren) eindeutigen Zerlegung (24.11) heißen die **algebraischen Vielfachheiten** (englisch: algebraic multiplicities) der Eigenwerte $\lambda_1, \ldots, \lambda_s$ von A. Wir schreiben $\mu^{\text{alg}}(A, \lambda_i) = n_i$.

(ii) Im Fall $n_i = 1$ nennen wir λ_i einen **einfachen Eigenwert** (englisch: simple eigenvalue) von A. Im Fall $n_i \geqslant 2$ nennen wir λ_i einen **mehrfachen Eigenwert** (englisch: multiple eigenvalue) von A. △

Beachte: Nach Folgerung 11.22 gilt für die Summe der algebraischen Vielfachheiten $\sum_{i=1}^{s} \mu^{\text{alg}}(A, \lambda_i) \leqslant n$.

Wir können die algebraische Vielfachheit eines einzelnen Eigenwertes auch angeben, ohne die gesamte Darstellung (24.11) vorliegen zu haben:

Lemma 24.16 (algebraische Vielfachheit und Teiler des charakteristischen Polynoms).
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Für die algebraische Vielfachheit $\mu^{\text{alg}}(A, \lambda_i)$ eines Eigenwertes λ_i von A gilt

$$\mu^{\text{alg}}(A, \lambda_i) = \max\big\{k \in \mathbb{N} \,\big|\, (\lambda - \lambda_i)^k \mid \chi_A\big\}. \tag{24.12}$$

Beweis. Das Resultat folgt aus der Darstellung (24.11). □

Beispiel 24.17 (charakteristisches Polynom).

(i) Die Darstellungsmatrix der Spiegelungsabbildung aus Beispiel 20.16 bzgl. der Standardbasis in $\mathbb{R}^2$
$$A = \begin{bmatrix} \cos^2(\alpha) - \sin^2(\alpha) & 2\cos(\alpha)\,\sin(\alpha) \\ 2\cos(\alpha)\,\sin(\alpha) & \sin^2(\alpha) - \cos^2(\alpha) \end{bmatrix} = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}$$
besitzt das charakteristische Polynom
$$\begin{aligned} \chi_A &= \det \begin{pmatrix} \lambda - \cos(2\alpha) & -\sin(2\alpha) \\ -\sin(2\alpha) & \lambda + \cos(2\alpha) \end{pmatrix} \\ &= \big(\lambda - \cos(2\alpha)\big)\big(\lambda + \cos(2\alpha)\big) - \sin^2(2\alpha) \\ &= \lambda^2 - \cos^2(2\alpha) - \sin^2(2\alpha) \\ &= \lambda^2 - 1 \\ &= (\lambda - 1)\,(\lambda + 1). \end{aligned}$$
Dessen Nullstellen, also Eigenwerte von A, sind ± 1, was das Ergebnis von Beispiel 24.3 bestätigt. Die algebraische Vielfachheit beider Eigenwerte ist $\mu^{\text{alg}}(A, 1) = \mu^{\text{alg}}(A, -1) = 1$.

(ii) Die Darstellungsmatrix der Drehabbildung aus Beispiel 17.8 bzgl. der Standardbasis in $\mathbb{R}^2$
$$A = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}$$
besitzt das charakteristische Polynom
$$\begin{aligned} \chi_A &= \det \begin{pmatrix} \lambda - \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \lambda - \cos(\alpha) \end{pmatrix} \\ &= \big(\lambda - \cos(\alpha)\big)\big(\lambda - \cos(\alpha)\big) + \sin^2(\alpha) \\ &= \lambda^2 - 2\,\lambda\,\cos(\alpha) + \cos^2(\alpha) + \sin^2(\alpha) \\ &= \lambda^2 - 2\,\lambda\,\cos(\alpha) + 1. \end{aligned}$$
Dieses Polynom hat genau dann eine Nullstelle in $\mathbb{R}$, wenn $\cos^2(\alpha) - 1 \geqslant 0$ gilt, also $\cos^2(\alpha) \geqslant 1$, was wegen $|\cos(\alpha)| \leqslant 1$ gleichbedeutend ist mit $\cos^2(\alpha) = 1$, also $\alpha = 0$ oder $\alpha = \pi$ im Intervall $[0, 2\pi)$.

Im Fall $\alpha = 0$ erhalten wir $\chi_A = \lambda^2 - 2\,\lambda + 1 = (\lambda - 1)^2$, also ist 1 der einzige Eigenwert mit algebraischer Vielfachheit $\mu^{\text{alg}}(A, 1) = 2$. Im Fall $\alpha = \pi$ erhalten wir $\chi_A = \lambda^2 + 2\,\lambda + 1 = (\lambda + 1)^2$, also ist -1 der einzige Eigenwert mit algebraischer Vielfachheit $\mu^{\text{alg}}(A, -1) = 2$. △

Bemerkung 24.18 (charakteristisches Polynom)**.**

(i) Es stellt sich die Frage, ob man ohne Bedenken (lineare) Polynome in die Einträge einer Matrix einsetzen kann. Schließlich haben wir Matrizen an sich sowie auch die Determinante nur für Matrizen definiert, deren Einträgen aus einem Körper K kommen. Nach Lemma 11.12 ist aber der Ring der Polynome $K[\lambda]$ (und auch die Teilmenge der Polynome vom Höchstgrad 1) kein Körper.

Es gibt zwei mögliche Abhilfen: Wir könnten Matrizen allgemeiner über kommutativen Ringen definieren. Es ist aber auch möglich, den Polynomring $K[\lambda]$ zum sogenannten **rationalen**

Funktionenkörper $K(\lambda)$ zu erweitern.[1] Dieser besteht aus den Äquivalenzklassen von

$$\left\{ \frac{p}{q} \,\middle|\, p, q \in K[\lambda],\ q \neq 0 \right\}$$

bzgl. der Äquivalenzrelation $\frac{p_1}{q_1} \sim \frac{p_2}{q_2} \Leftrightarrow p_1 \cdot q_2 = p_2 \cdot q_1$. Der Polynomring $K[\lambda]$ wird durch Identifikation von p mit der Äquivalenzklasse zu $\frac{p}{1}$ zu einem kommutativen Unterring mit Eins von $K(\lambda)$. Der Körper K ist ein Unterkörper von $K(\lambda)$.

(*ii*) Die Matrix $\lambda I - A$ kann also als Matrix im Vektorraum $K(\lambda)^{n \times n}$ über dem Körper $K(\lambda)$ aufgefasst werden. Die Determinante von $\lambda I - A$ ist damit nach Definition 23.6 definiert als Element von $K(\lambda)$. Das heißt, auch $\chi_A = \det(\lambda I - A)$ gehört zu $K(\lambda)$. In Wirklichkeit liegt $\chi_A = \det(\lambda I - A)$ sogar im Unterring der Polynome $K[\lambda] \subsetneq K(\lambda)$. Das liegt daran, dass alle Einträge von $\lambda I - A$ rationale Ausdrücke mit Nenner $1 \in K \subsetneq K(\lambda)$ sind. Damit trifft das auch auf Produkte von Einträgen und schließlich auf Summen von Produkten von Einträgen zu, also auf $\det(\lambda I - A)$.

(*iii*) Oft wird das charakteristische Polynom einer Matrix A als $\det(A - \lambda I)$ definiert, was sich von (24.9) im Vorzeichen unterscheidet. Dieser Unterschied hat aber im Folgenden keine Auswirkungen. △

Definition 24.19 (Spur einer Matrix)**.**
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Die **Spur** (englisch: trace) von A ist definiert als

$$\operatorname{Spur}(A) := \sum_{i=1}^{n} a_{ii}, \tag{24.13}$$

also als die Summe der Hauptdiagonalelemente. △

Satz 24.20 (Eigenschaften des charakteristischen Polynoms)**.**
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Das charakteristische Polynom

$$\chi_A = \sum_{i=0}^{n} \alpha_i \lambda^i \in K_n[\lambda] \tag{24.14}$$

und seine Koeffizienten $\alpha_0, \ldots, \alpha_n$ haben folgende Eigenschaften:

(*i*) χ_A ist ein normiertes Polynom vom Grad $\deg(\chi_A) = n$, also gilt $\alpha_n = 1$.

(*ii*) Im Fall $n \geqslant 1$ gilt $\alpha_{n-1} = -\operatorname{Spur}(A)$.

(*iii*) Es gilt $\alpha_0 = (-1)^n \det(A)$.

Beweis. Wir schreiben wie in (24.10)

$$\chi_A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \prod_{j=1}^{n} \big(\delta_{\sigma(j),j}\, \lambda - a_{\sigma(j),j}\big)$$

[1] Diese Konstruktion ist für alle Integritätsringe R möglich, also für kommutative, nullteilerfreie Ringe mit Eins ungleich dem Nullring (Definition 9.7). Der Polynomring über einem Körper hat diese Eigenschaft, siehe Folgerung 11.11. Allgemein spricht man vom **Quotientenkörper** (englisch: field of fractions) eines Integritätsringes R, das ist der (bis auf Isomorphie eindeutig bestimmte) kleinste Körper, in den R eingebettet werden kann. Ein prominentes Beispiel ist $\mathbb{Q}$, der Quotientenkörper von $\mathbb{Z}$.

$$= \underbrace{\prod_{j=1}^{n} (\lambda - a_{jj})}_{p_1} + \underbrace{\sum_{\sigma \in S_n \setminus \{\mathrm{id}\}} (\operatorname{sgn} \sigma) \prod_{j=1}^{n} \bigl(\delta_{\sigma(j),j}\, \lambda - a_{\sigma(j),j}\bigr)}_{p_2}.$$

In der Summe p_2 treten in jedem Produkt höchstens $n-2$ Diagonalelemente auf. (**Quizfrage 24.2:** Warum?) Damit ist der Grad von p_2 höchstens $n-2$. Die gesuchten Koeffizienten α_n und α_{n-1} von χ_A sind also in p_1 zu finden. Durch Ausmultiplizieren von p_1 erhalten wir

$$\begin{aligned} p_1 &= 1 \cdot \lambda^n - (a_{11} + \cdots + a_{nn})\, \lambda^{n-1} + \ldots \\ &= \lambda^n - \operatorname{Spur}(A)\, \lambda^{n-1} + \ldots \end{aligned}$$

Den Koeffizienten α_0 niedrigster Ordnung erhalten wir durch Einsetzen $\lambda = 0$ in die zu χ_A gehörige Polynomfunktion $\widetilde{\chi_A}$, was $\alpha_0 = \det(-A) = (-1)^n \det(A)$ bestätigt. □

Welcher Zusammenhang besteht zwischen der geometrischen und der algebraischen Vielfachheit eines Eigenwertes?

Satz 24.21 (geometrische und algebraische Vielfachheit)**.**
Es sei K ein Körper, $A \in K^{n \times n}$, $n \in \mathbb{N}_0$ und $\lambda \in K$ ein Eigenwert von A. Dann gilt:

$$1 \leqslant \mu^{\mathrm{geo}}(A, \lambda) \leqslant \mu^{\mathrm{alg}}(A, \lambda). \tag{24.15}$$

Beweis. Es sei $\lambda_i \in K$ ein Eigenwert von A. Nach Definition 24.6 gilt $\mu^{\mathrm{geo}}(A, \lambda_i) \coloneqq \dim(\operatorname{Eig}(A, \lambda_i))$, und nach Lemma 24.8 gilt $1 \leqslant \mu^{\mathrm{geo}}(A, \lambda_i)$.

Wir setzen zur Abkürzung $g \coloneqq \mu^{\mathrm{geo}}(A, \lambda_i)$. Es sei nun $(v_1, \ldots, v_g)$ eine Basis von $\operatorname{Eig}(A, \lambda_i)$. Wir ergänzen diese zu einer Basis $(v_1, \ldots, v_g, v_{g+1}, \ldots, v_n)$ von K^n. Für die Darstellungsmatrix von f_A in dieser Basis gilt dann (siehe die Überlegungen vor Satz 20.17)

$$\widehat{A} = \begin{array}{cc} \begin{array}{cc} g & n-g \end{array} & \\ \left[\begin{array}{c|c} \lambda_i I & * \\ \hline 0 & B \end{array}\right] & \begin{array}{l} g \\ n-g \end{array} \end{array}.$$

Für das charakteristische Polynom dieser zu A ähnlichen Matrix haben wir also (die Gleichheit $\chi_{\widehat{A}} = \chi_A$ wird gleich in Lemma 24.22 bewiesen)

$$\chi_{\widehat{A}} = \chi_A = (\lambda - \lambda_i)^g\, \chi_B.$$

Es ist nun möglich, dass im Restpolynom χ_B nochmals eine Potenz des Linearfaktors $(\lambda - \lambda_i)^g$ auftritt. Damit haben wir $g = \mu^{\mathrm{geo}}(A, \lambda_i) \leqslant \mu^{\mathrm{alg}}(A, \lambda_i)$. □

§ 24.3 Charakteristisches Polynom eines Endomorphismus

Lemma 24.22 (ähnliche Matrizen haben dasselbe charakteristische Polynom)**.**
Es sei K ein Körper und $n \in \mathbb{N}_0$. Sind $A, \widehat{A} \in K^{n \times n}$ ähnlich, dann gilt $\chi_A = \chi_{\widehat{A}}$.

Beweis. Nach Voraussetzung gilt $\widehat{A} = T\,A\,T^{-1}$ mit einer invertierbaren Matrix $T \in K^{n\times n}$. Wir fassen jetzt λ, ein Polynom vom Grad 1, als Element des rationalen Funktionenkörpers $K(\lambda)$ auf. Weiterhin fassen wir I, T, A usw. als Elemente von $K(\lambda)^{n\times n}$ auf, also als $n \times n$-Matrizen über dem Körper $K(t)$. $\lambda\,I$ ist die S-Multiplikation von $\lambda \in K(\lambda)$ mit $I \in K(\lambda)^{n\times n}$. Es gilt

$$\begin{aligned}
\lambda\,I - \widehat{A} &= \lambda\,T\,T^{-1} - T\,A\,T^{-1} \\
&= T\,(\lambda\,T^{-1}) - T\,A\,T^{-1} \qquad \text{wegen (15.12)} \\
&= T\,(\lambda\,I)\,T^{-1} - T\,A\,T^{-1} \\
&= T\,\big(\lambda\,I - A\big)\,T^{-1}.
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
\chi_{\widehat{A}} &= \det(\lambda\,I - \widehat{A}) \\
&= \det\big(T\,\big(\lambda\,I - A\big)\,T^{-1}\big) \\
&= \det(T)\ \det(\lambda\,I - A)\ \det(T^{-1}) \\
&= \det(\lambda\,I - A) \\
&= \chi_A.
\end{aligned}$$ □

Wegen Lemma 24.22 besitzen alle Darstellungsmatrizen eines Endomorphismus dasselbe charakteristische Polynom. Dieses wird damit auch zu einer Eigenschaft des Endomorphismus selbst.

Definition 24.23 (charakteristisches Polynom eines Endomorphismus)**.**
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B_V = (v_1, \dots, v_n)$, $n \in \mathbb{N}_0$. Das **charakteristische Polynom** eines Endomorphismus $f \in \operatorname{End}(V)$ ist definert als

$$\chi_f \coloneqq \chi_A \tag{24.16}$$

für die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f)$. Folglich können wir wie in Definition 24.15 und Lemma 24.16 auch von der **algebraischen Vielfachheit** $\mu^{\mathrm{alg}}(f, \lambda)$ des Eigenwertes λ von f sprechen. △

Wir können nun das Analogon von Lemma 24.14 angeben:

Lemma 24.24 (Eigenwerte von Endomorphismen sind Nullstellen des charakteristischen Polynoms, vgl. Lemma 24.14)**.**
Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Weiter sei $f \in \operatorname{End}(V)$ und $\lambda \in K$. Dann sind äquivalent:

(*i*) λ ist ein Eigenwert von f.

(*ii*) λ ist eine Nullstelle des charakteristischen Polynoms χ_f.

Beweis. Es gilt

$$\begin{aligned}
& \lambda \text{ ist ein Eigenwert von } f \\
\Leftrightarrow\quad & \det(\lambda\,\mathrm{id}_V - f) = 0 && \text{nach Lemma 24.11} \\
\Leftrightarrow\quad & \det(\lambda\,I - A) = 0 && \text{nach Definition 23.21} \\
\Leftrightarrow\quad & \lambda \text{ ist Nullstelle von } \chi_A && \text{nach Definition von } \chi_A \\
\Leftrightarrow\quad & \lambda \text{ ist Nullstelle von } \chi_f && \text{wegen (24.16).}
\end{aligned}$$ □

Beachte: Damit ist auch klar: $\mu^{\text{alg}}(f,\lambda) = \mu^{\text{alg}}(A,\lambda)$ für jede Darstellungsmatrix A von f.

Aus Lemma 24.22 lässt sich außerdem folgendes Resultat gewinnen:

Lemma 24.25 (ähnliche Matrizen haben dieselbe Spur).
Es sei K ein Körper und $n \in \mathbb{N}_0$. Sind $A, \widehat{A} \in K^{n\times n}$ ähnlich, dann gilt $\operatorname{Spur}(A) = \operatorname{Spur}(\widehat{A})$.

Beweis. Nach Satz 24.20 ist die Spur einer Matrix der zweithöchste Koeffizient ihres charakteristischen Polynoms. Die charakteristischen Polynome χ_A und $\chi_{\widehat{A}}$ sind nach Lemma 24.22 identisch, also auch ihre Koeffizienten. (**Quizfrage 24.3:** Wie argumentiert man im Fall $n = 0$?) □

Damit wird auch die Spur zu einer Eigenschaft eines Endomorphismus:

Definition 24.26 (Spur eines Endomorphismus).
Es seien K ein Körper und V ein Vektorraum über K mit Basis $B_V = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$. Die **Spur** eines Endomorphismus $f \in \operatorname{End}(V)$ ist definert als

$$\operatorname{Spur}(f) \coloneqq \operatorname{Spur}(A) \tag{24.17}$$

für die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f)$. △

Wir geben nun neben dem bereits bekannten hinreichenden Kriterium (Folgerung 24.2) auch ein notwendiges Kriterium für die Diagonalisierbarkeit an.

Lemma 24.27 (notwendiges Kriterium für Diagonalisierbarkeit).
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Ist $f \in \operatorname{End}(V)$ diagonalisierbar, dann zerfällt das charakteristische Polynom χ_f in Linearfaktoren, es gilt also

$$\chi_f = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s} \tag{24.18}$$

mit den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_s \in K$ und deren algebraischen Vielfachheiten $n_1, \ldots, n_s \in \mathbb{N}$, für die $n_1 + \cdots + n_s = n$ gilt.

Beweis. Es sei f diagonalisierbar, es gibt also nach Satz 20.24 eine Basis $B_V = (v_1, \ldots, v_n)$ aus Eigenvektoren von f. Die zugehörigen paarweise verschiedenen Eigenwerte seien

$$\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1\text{-mal}}, \ldots, \underbrace{\lambda_s, \ldots, \lambda_s}_{n_s\text{-mal}}.$$

Für die zugehörige Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f)$ gilt dann

$$A = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \lambda_s & \\ & & & & & \ddots \\ & & & & & & \lambda_s \end{bmatrix} \quad \text{und damit} \quad \lambda I - A = \begin{bmatrix} \lambda - \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda - \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \lambda - \lambda_s & \\ & & & & & \ddots \\ & & & & & & \lambda - \lambda_s \end{bmatrix}.$$

Es folgt

$$\chi_f = \chi_A = \det(\lambda I - A) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}. \qquad \square$$

Das folgende Resultat wurde nachträglich eingefügt.

Satz 24.28 (notwendiges und hinreichendes Kriterium für Diagonalisierbarkeit)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Weiter seien $\lambda_1, \dots, \lambda_s$, $s \in \mathbb{N}_0$, die paarweise verschiedenen Eigenwerte von A. Dann sind äquivalent:

(a) A ist diagonalisierbar.

(b) $K^n = \displaystyle\bigoplus_{j=1}^{s} \operatorname{Eig}(A, \lambda_j)$.

(c) $\displaystyle\sum_{j=1}^{s} \mu^{\text{geo}}(A, \lambda_j) = n$.

(d) Das charakteristische Polynom χ_A zerfällt vollständig in Linearfaktoren, und es gilt $\mu^{\text{alg}}(A, \lambda_j) = \mu^{\text{geo}}(A, \lambda_j)$ für alle $j = 1, \dots, s$.

(ii) Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f \in \operatorname{End}(V)$ und $\lambda_1, \dots, \lambda_s$, $s \in \mathbb{N}_0$, die paarweise verschiedenen Eigenwerte von f. Dann sind äquivalent:

(a) f ist diagonalisierbar.

(b) $K^n = \displaystyle\bigoplus_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)$.

(c) $\displaystyle\sum_{j=1}^{s} \mu^{\text{geo}}(f, \lambda_j) = n$.

(d) Das charakteristische Polynom χ_f zerfällt vollständig in Linearfaktoren, und es gilt $\mu^{\text{alg}}(f, \lambda_j) = \mu^{\text{geo}}(f, \lambda_j)$ für alle $j = 1, \dots, s$.

Beweis. Wir beweisen nur Aussage (i).

Aussage (a) $\Rightarrow$ Aussage (b): Es sei A diagonalisierbar, d. h., K^n besitzt eine Basis aus n Eigenvektoren von A (Satz 20.24). Wenn wir diese nach den zugehörigen Eigenwerten λ_j, $j = 1, \dots, s$ gruppieren, so erhalten wir jeweils eine Basis von $\operatorname{Eig}(f, \lambda_j)$, der Mächtigkeit $\dim(\operatorname{Eig}(f, \lambda_j)) = \mu^{\text{geo}}(f, \lambda_j)$. Nach Lemma 24.9 ist $\sum_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)$ eine direkte Summe in K^n. Die Vereinigung der zugehörigen Basen ist nach Satz 14.16 eine Basis von $\bigoplus_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)$. Für die Dimension dieses Unterraumes von K^n gilt daher

$$\dim\Big(\bigoplus_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)\Big) = n,$$

also schließen wir $K^n = \displaystyle\bigoplus_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)$.

Aussage (b) $\Rightarrow$ Aussage (c): Die Dimension $\dim\big(\bigoplus_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)\big)$ ist die Summe der Dimensionen $\dim(\operatorname{Eig}(f, \lambda_j)) = \mu^{\text{geo}}(f, \lambda_j)$, $j = 1, \dots, s$.

Aussage (c) $\Rightarrow$ Aussage (a): Die Vereinigung von Basen der Eigenräume $\operatorname{Eig}(f, \lambda_j)$, $j = 1, \dots, s$, ergibt eine Basis der direkten Summe $\bigoplus_{j=1}^{s} \operatorname{Eig}(f, \lambda_j)$. Nach Voraussetzung ist die Mächtigkeit dieser Basis gleich n, und sie besteht nur aus Eigenvektoren von A. Damit ist A diagonalisierbar.

Aussage (a) $\Rightarrow$ Aussage (d): Ist A diagonalisierbar, dann ist A nach Definition 20.23 ähnlich zu einer Diagonalmatrix. Bezeichnen wir die geometrischen Vielfachheiten mit $n_j \coloneqq \mu^{\text{geo}}(f, \lambda_j)$, so gilt genauer, dass A ähnlich ist zu

$$\begin{array}{cc} \begin{array}{cccc} n_1 & n_2 & \cdots & n_s \end{array} & \\ \begin{bmatrix} \lambda_1 I_{n_1} & & & \\ & \lambda_2 I_{n_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{n_s} \end{bmatrix} & \begin{array}{c} n_1 \\ n_2 \\ \vdots \\ n_s \end{array} \end{array}$$

Da ähnliche Matrizen dasselbe charakteristische Polynom besitzen (Lemma 24.22), gilt

$$\chi_A = \prod_{j=1}^{s} (t - \lambda_j)^{n_j}.$$

Das charakteristische Polynom χ_A zerfällt also in Linearfaktoren, und es gilt $\mu^{\text{alg}}(f, \lambda_j) = \mu^{\text{geo}}(f, \lambda_j)$ für alle $j = 1, \ldots, s$.

Aussage (d) $\Rightarrow$ Aussage (c): Für das charakteristische Polynom von A gelte

$$\chi_A = \prod_{j=1}^{s} (t - \lambda_j)^{n_j}$$

mit Exponenten $n_j \in \mathbb{N}$, $j = 1, \ldots, s$. Dann gilt $n_j = \mu^{\text{alg}}(f, \lambda_j)$, und nach Voraussetzung auch $n_j = \mu^{\text{geo}}(f, \lambda_j)$. Da $\deg(\chi_A) = n$ ist und andererseits nach Lemma 11.10 auch $\sum_{j=1}^{s} n_j = n$, haben wir $\sum_{j=1}^{s} \mu^{\text{geo}}(f, \lambda_j) = n$.

Der Beweis von Aussage (ii) geht analog. □

Wir halten als Ergebnis von Satz 24.28 noch fest, welche Situationen dafür verantwortlich sein können, dass ein Endomorphismus eines n-dimensionalen Vektorraumes bzw. eine $n \times n$-Matrix nicht diagonalisierbar sind:

Beispiel 24.29 (mögliche Ursachen fehlender Diagonalisierbarkeit)**.**

(i) χ_A zerfällt nicht vollständig in Linearfaktoren, es gilt also wie in (24.11)

$$\chi_A = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s} \cdot q$$

mit einem Polynom q mindestens vom Grad 2, das keine Nullstellen in K besitzt. (**Quizfrage 24.4:** Warum gilt $\deg(q) \geqslant 2$?)

In dieser Situation gilt für die Dimension der (direkten) Summe $\bigoplus_{i=1}^{s} \operatorname{Eig}(A, \lambda_i)$ aller Eigenräume

$$\dim\Big(\bigoplus_{i=1}^{s} \operatorname{Eig}(A, \lambda_i)\Big) = \sum_{i=1}^{s} \mu^{\text{geo}}(A, \lambda_i) \leqslant \sum_{i=1}^{s} \mu^{\text{alg}}(A, \lambda_i) = \sum_{i=1}^{s} n_i = n - \deg(q) \leqslant n - 2.$$

Daher können wir mit Eigenvektoren nicht den ganzen Raum K^n aufspannen.

Ein Beispiel für diese Situation ist

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

mit dem charakteristischen Polynom $\chi_A = \lambda^2 + 1$, das über $\mathbb{R}$ keine Nullstellen besitzt.

(*ii*) Für mindestens einen Eigenwert λ_i von f gilt $\mu^{\text{geo}}(A, \lambda_i) < \mu^{\text{alg}}(A, \lambda_i)$.

Ein Beispiel für diese Situation ist

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

mit dem charakteristischen Polynom $\chi_A = (\lambda - 1)^2$ mit $\mu^{\text{alg}}(A, 1) = 2$. Es gilt jedoch

$$\operatorname{Eig}(A, 1) = \operatorname{Kern}(1I - A) = \operatorname{Kern}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \left\langle \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix} \right) \right\rangle$$

mit $\mu^{\text{geo}}(A, 1) = 1$. △

Bemerkung 24.30 (Bestimmung von Nullstellen ist schwierig)**.**
Die Bestimmung von Nullstellen eines Polynoms ist im Allgemeinen schwierig. Man kann mit Hilfe der nach Galois benannten Theorie zeigen, dass für allgemeine Polynome vom Grad $n \geqslant 5$ keine Lösungsformeln mehr existiert, sodass man auf eine numerische Bestimmung von Eigenwerten angewiesen ist.[2] Damit ist es auch schwierig, festzustellen, ob das charakteristische Polynom einer Matrix oder eines Endomorphismus vollständig in Linearfaktoren zerfällt. △

Ende der Vorlesung 13

§ 25 Algebren über Körpern

Wir wollen im weiteren Verlauf von Kapitel 5 auch für den nicht-diagonalisierbaren Fall klären, bzgl. welcher Basen die Darstellungsmatrix eines Endomorphismus eine möglichst einfache Gestalt hat und wie diese aussieht. Dazu müssen wir etwas weiter ausholen.

Wir hatten zu Beginn von § 11.2 die Frage gestellt, welche Elemente man für die Variable t in einem Polynom

$$p = \alpha_n\, t^n + \alpha_{n-1}\, t^{n-1} + \cdots + \alpha_1\, t + \alpha_0$$

sinnvollerweise (in die zugehörige Polynomfunktion) einsetzen kann. Dort hatten wir zunächst Elemente des kommutativen Ringes R eingesetzt, aus dem auch die Koeffizienten $\alpha_0, \ldots, \alpha_n$ stammen. Dabei nutzen wir die „Multiplikation“ im Ring, um für $a \in R$ den Ausdruck

$$\alpha_n \cdot a^n = \alpha_n \cdot \underbrace{a \cdots a}_{n\text{-mal}} \in R$$

zu bilden. Weiterhin nutzen wir die „Addition“ im Ring, um mehrere solcher Terme zu „addieren“.

Es ist jedoch auch möglich, allgemeinere Objekte als Elemente des Koeffizientenringes R einzusetzen. Stammen diese Elemente aus einer Struktur A, dann benötigen wir offenbar

$$\begin{aligned}
\text{neben einer inneren Verknüpfung } \star\colon A \times A \to A \quad & \text{zur Bildung von } a^n = a \star a \cdots \star a \text{ für } a \in A \\
\text{auch die innere Verknüpfung } +\colon A \times A \to A \quad & \text{zur „Addition“ } \alpha_0 + \alpha_1 \cdot a + \alpha_2 \cdot a^2 + \cdots \\
\text{sowie eine äußere Verknüpfung } \cdot\colon R \times A \to A \quad & \text{zur Bildung von } \alpha_n \cdot a^n \text{ für } \alpha_n \in R \text{ und } a \in A.
\end{aligned}$$

Eine solche Struktur, deren Elemente sich u. a. für das Einsetzen in Polynome eignen, heißt eine **Algebra** über dem (kommutativen) Ring R. Wir benötigen in dieser Vorlesung aber nur Algebren über Körpern.

[2] Mehr dazu in Vorlesungen zur Numerik, insbesondere zur numerischen linearen Algebra.

Definition 25.1 (Algebra über einem Körper).
Es sei $(K, +, \cdot)$ ein Körper. Eine **(assoziative) Algebra** (englisch: (associative) algebra) $(A, +, \cdot, \star)$ **über** K (kurz: eine K-Algebra) ist eine Menge A mit zwei inneren Verknüpfung $+\colon A \times A \to A$ und $\star\colon A \times A \to A$ sowie einer äußeren Verknüpfung $\cdot\colon K \times A \to A$, die die folgenden Bedingungen erfüllen:

(i) $(A, +, \cdot)$ ist ein K-Vektorraum.

(ii) $(A, +, \star)$ ist ein Ring.

(iii) Die Verknüpfung $\star$, genannt die **Multiplikation** (englisch: multiplication) ist verträglich mit der S-Multiplikation, d. h., es gilt ein **„gemischtes" Assoziativgesetz**

$$(\alpha \cdot a) \star b = \alpha \cdot (a \star b) = a \star (\alpha \cdot b) \tag{25.1}$$

für alle $\alpha \in K$ und $a, b \in A$.[3]

Eine Algebra A heißt **kommutativ** (englisch: commutative), wenn $\star$ kommutativ ist.

Eine Algebra A heißt eine **Algebra mit Eins** (englisch: algebra with unity) oder eine **unitäre Algebra** (englisch: unitary algebra), wenn es in A ein neutrales Element bzgl. $\star$ gibt.[4] Existiert dann zu $a \in A$ bzgl. $\star$ ein inverses Element, so bezeichnen wir dieses mit a^{-1}. △

Wir vereinbaren, dass $\star$ stärker bindet als $+$ und $-$, sodass wir beispielsweise statt $(a \star b) + (a \star c)$ auch $a \star b + a \star c$ schreiben können. Zwischen $\cdot$ und $\star$ müssen wir wegen (25.2) keine Bindungsstärke festlegen.

Lemma 25.2 (Multiplikation in einer Algebra ist bilinear).
Es sei $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K. Dann ist die Multiplikation $\star$ bilinear, d. h., es gilt

$$(\alpha \cdot a + \beta \cdot b) \star c = \alpha \cdot (a \star c) + \beta \cdot (b \star c) \tag{25.2a}$$
$$a \star (\beta \cdot b + \gamma \cdot c) = \beta \cdot (a \star b) + \gamma \cdot (a \star c) \tag{25.2b}$$

für alle $a, b, c \in$ und alle $\alpha, \beta, \gamma \in K$.

Beweis. Aufgrund der Ringeigenschaft von $(A, +, \star)$ gelten die Distributivgesetze (vgl. (9.1))

$$\begin{aligned}(a + b) \star c &= (a \star c) + (b \star c)\\ a \star (b + c) &= (a \star b) + (a \star c)\end{aligned}$$

für alle $a, b, c \in A$. Wir haben also

$$\begin{aligned}(\alpha \cdot a + \beta \cdot b) \star c &= (\alpha \cdot a) \star c + (\beta \cdot b) \star c && \text{nach Distributivgesetz}\\ &= \alpha \cdot (a \star c) + \beta \cdot (b \star c) && \text{wegen (25.2)}\end{aligned}$$

Analog gilt auch

$$\begin{aligned}a \star (\beta \cdot b + \gamma \cdot c) &= a \star (\beta \cdot b) + a \star (\gamma \cdot c) && \text{nach Distributivgesetz}\\ &= \beta \cdot (a \star b) + \gamma \cdot (a \star c) && \text{wegen (25.2).}\end{aligned}$$
□

[3] Zur Verdeutlichung haben wir in (25.2) die S-Multiplikation vorübergehend wieder mit $\cdot$ gekennzeichnet. Wie in Vektorräumen üblich, werden wir das aber ab sofort wieder unterlassen.

[4] Vergleiche die Bezeichnungsweise bei Ringen (Definition 9.1).

Beispiel 25.3 (Algebra über einem Körper).

(i) Über jedem Körper ist die **Nullalgebra** (englisch: zero algebra) die (bis auf Isomorphie) eindeutig bestimmte Algebra mit $A = \{0\}$ und die dadurch eindeutig bestimmten Verknüpfungen $0 + 0 = 0$, $\alpha \cdot 0 = 0$ für $\alpha \in K$ und $0 \star 0 = 0$. Wie beim Nullring (Beispiel 9.2) ist auch die Nullalgebra eine Algebra mit Einselement $1 = 0$. Sie ist die einzige Algebra, in der das Nullelement und das Einselement identisch sind, siehe Lemma 9.3.

(ii) Für jeden Körper $(K, +, \cdot)$ ist $(K, +, \cdot, \cdot)$ eine kommutative Algebra mit Einselement 1 über sich selbst.

(iii) Allgemeiner sei $(K, +, \cdot)$ ein Körper und $(U, +, \cdot)$ ein Unterkörper. Dann ist $(K, +, \cdot_U, \cdot)$ eine kommutative Algebra mit Eins über U. (Dabei ist $\cdot_U : U \times K \to K$ die als S-Multiplikation dienende Einschränkung der Multiplikation $\cdot : K \times K \to K$ im Körper K.) Beispielsweise ist $\mathbb{R}$ eine $\mathbb{Q}$-Algebra, und $\mathbb{C}$ ist eine $\mathbb{R}$-Algebra und eine $\mathbb{Q}$-Algebra.

(iv) Es sei $(K, +, \cdot)$ ein Körper. Für $n \in \mathbb{N}_0$ ist die Menge der $n \times n$-Matrizen $(K^{n \times n}, +, \cdot, \cdot)$ mit der Matrix-Matrix-Multiplikation $\cdot$ ist eine (im Fall $n \geqslant 2$ nicht-kommutative) Algebra über K mit Einselement I_n.

Die Verknüpfungssymbole $\cdot$ für die S-Multiplikation und $\cdot$ für die Matrix-Matrix-Multiplikation sind identisch, aber wir können sie anhand der Objekte, die verknüpft werden unterscheiden, etwa $\alpha \cdot A$ und $A \cdot B$. Die Verknüpfungssymbole in dieser **Matrixalgebra** (englisch: matrix algebra) werden häufig auch gar nicht geschrieben, also $\alpha\, A\, B$ statt $\alpha \cdot A \cdot B$.

(v) Ist $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K, dann bildet $(\operatorname{End}(V), +, \cdot, \circ)$ mit der Komposition $\circ$ als „Multiplikation“ eine i. A. nicht kommutative Algebra über K mit Einselement id_V, genannt die **Endomorphismenalgebra von V** (englisch: algebra of endomorphisms).

(vi) Ist $(K, +, \cdot)$ ein Körper, dann bilden die Polynome $(K[t], +, \cdot, \cdot)$ mit der Polynommultiplikation $\cdot$ (siehe (11.3)) eine kommutative Algebra über K mit Einselement 1 (Einspolynom), genannt die **Polynomialgebra** (englisch: algebra of polynomials). (**Quizfrage 25.1:** Gilt das auch für die Polynome $K_n[t]$ vom Höchstgrad $n \in \mathbb{N}_0$?)

(vii) Ist X eine beliebige Menge, $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K, dann bildet die Menge der Funktionen

$$A^X = \{f : X \to A\}$$

mit den punktweisen Addition, S-Multiplikation und inneren Multiplikation eine Algebra über K. Diese erbt die möglicherweise vorhandene Kommutativität und das neutrale Element bzgl. $\star$ von A.

($viii$) Insbesondere bildet also die Menge der Funktionen $K^X = \{f : X \to K\}$ eine kommutative Algebra mit Eins über K. △

Definition 25.4 (Unteralgebra, vgl. Definition 7.31 einer Untergruppe).
Es sei $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K.

(i) Eine Teilmenge $U \subseteq A$ heißt ein **Unteralgebra** (englisch: subalgebra) von $(A, +, \cdot, \star)$, wenn U bzgl. $+$, $\cdot$ und $\star$ abgeschlossen ist und wenn $(U, +, \cdot, \star)$ selbst wieder ein Algebra ist.

Beachte: Das ist genau dann erfüllt, wenn $(U, +, \cdot)$ ein Unterraum von $(A, +, \cdot)$ und wenn $(U, +, \star)$ in Unterring von $(A, +, \star)$ ist.

(*ii*) Ist $(A, +, \cdot, \star)$ eine Algebra mit Einselement e, dann fordern wir für eine **Unteralgebra mit Eins** (englisch: subalgebra with unity) $(U, +, \cdot, \star)$ zusätzlich zu Eigenschaft (*i*), dass $e \in U$ liegt.[5]

Beachte: Es reicht nicht aus, zu fordern, dass $(U, \cdot)$ *irgendein* neutrales Element besitzt.

(*iii*) Eine Unteralgebra $(U, +, \cdot, \star)$ von $(A, +, \cdot, \star)$ heißt **echt** (englisch: proper subalgebra), wenn $U \subsetneq A$ gilt. △

Definition 25.5 (Algebrahomomorphismus, vgl. Definition 8.1 eines Halbgruppenhomomorphismus)**.**
Es seien $(A_1, +_1, \cdot_1, \star)$ und $(A_2, +_2, \cdot_2, \square)$ zwei Algebren über demselben Körper K.

(*i*) Eine Abbildung $f\colon A_1 \to A_2$ heißt **strukturverträglich** oder ein **(Algebra-)Homomorphismus** (englisch: algebra homomorphism) von $(A_1, +_1, \cdot_1, \star)$ in $(A_2, +_2, \cdot_2, \square)$, wenn gilt:

$$f(a +_1 b) = f(a) +_2 f(b) \quad \text{für alle } a, b \in A_1, \tag{25.3a}$$
$$f(\alpha \cdot_1 a) = \alpha \cdot_2 f(a) \quad \text{für alle } \alpha \in K \text{ und } a \in A_1. \tag{25.3b}$$
$$f(a \star b) = f(a) \square f(b) \quad \text{für alle } a, b \in A_1. \tag{25.3c}$$

Besitzen beide Algebren ein Einselement e_{A_1} bzw. e_{A_2} und fordern wir zusätzlich

$$f(e_{A_1}) = e_{A_2} \tag{25.3d}$$

dann nennen wir f genauer einen **Homomorphismus von Algebren mit Eins** (englisch: homomorphism of algebras with unity).

(*ii*) Wie in Definition 8.1 sprechen wir im Fall $(A_1, +_1, \cdot_1, \star) = (A_2, +_2, \cdot_2, \square)$ von einem **(Algebra-) Endomorphismus** (englisch: algebra endomorphism) bzw. von einem **Endomorphismus einer Algebra mit Eins** (englisch: endomorphism of an algebra with unity).

(*iii*) Ist $f\colon A_1 \to A_2$ bijektiv, so heißt f auch **strukturerhaltend** oder ein **(Algebra-)Isomorphismus** (englisch: algebra isomorphism) bzw. ein **Isomorphismus von Algebren mit Eins** (englisch: endomorphism of an algebra with unity). In diesem Fall nennen wir $(A_1, +_1, \cdot_1, \star)$ und $(A_2, +_2, \cdot_2, \square)$ auch zueinander **isomorphe Algebren** (englisch: isomorphic algebras) bzw. zueinander **isomorphe Algebren mit Eins** (englisch: isomorphic algebras with unity) und schreiben

$$(A_1, +_1, \cdot_1, \star) \cong (A_2, +_2, \cdot_2, \square).$$

(*iv*) Im Fall $(A_1, +_1, \cdot_1, \star) = (A_2, +_2, \cdot_2, \square)$ und $f\colon A_1 \to A_2$ bijektiv sprechen wir auch von einem **(Algebra-)Automorphismus** (englisch: algebra automorphism) bzw. von einem **Automorphismus einer Algebra mit Eins** (englisch: automorphism of an algebra with unity).

(*v*) Das **Bild** (englisch: image) und der **Kern** (englisch: kernel) eines Algebrahomomorphismus $f\colon A_1 \to A_2$ sind definiert als

$$\begin{aligned} \operatorname{Bild}(f) &:= \left\{ f(x) \in A_2 \,\middle|\, x \in A_1 \right\} = f(A_1), \\ \operatorname{Kern}(f) &:= \left\{ x \in A_1 \,\middle|\, f(x) = 0_{A_2} \right\} = f^{-1}(\{0_{A_2}\}). \end{aligned} \tag{25.4}$$

△

Beispiel 25.6 (Algebrahomomorphismus)**.**
Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Dann sind die Algebra der Endomorphismen von V $(\operatorname{End}(V), +, \cdot, \circ)$ und die Algebra der Matrizen $(K^{n \times n}, +, \cdot, \cdot)$ isomorph als Algebren mit Eins. Für jede Basis B_V von V ist $f \mapsto \mathcal{M}_{B_V}^{B_V}(f)$ ein Isomorphismus. △

[5] Dadurch ist die Unteralgebra $(U, +, \cdot, \star)$ dann natürlich selbst wieder ein Algebra mit dem Einselement e.

Wir kommen zurück zu unserer Motivation vom Anfang des § 25, nämlich Elemente einer K-Algebra $(A, +, \cdot, \star)$ in ein Polynom aus $K[t]$ einzusetzen. Es sei dazu

$$p = \sum_{i=0}^{n} \alpha_i \, t^i$$

ein Polynom mit Koeffizienten in K. Wir setzen voraus, dass A ein Einselement $e \in A$ besitzt und definieren (vgl. Bemerkung 7.13) für $a \in A$

$$\begin{aligned} a^0 &:= e \\ a^1 &:= a \\ a^2 &:= a \star a \\ a^3 &:= a \star a \star a \end{aligned}$$

und damit allgemein a^n für $n \in \mathbb{N}_0$. (**Quizfrage 25.2:** Welche Eigenschaft der Algebra A ist für die Wohldefiniertheit dieser abkürzenden Schreibweise wichtig?)

Damit können wir nun die durch p induzierte **Polynomfunktion** (vgl. (11.8)) $\widetilde{p}_A \colon A \to A$ definieren als

$$\widetilde{p}_A(a) := \sum_{i=0}^{n} \alpha_i \, a^n = \alpha_0 \, e + \alpha_1 \, a + \cdots + \alpha_{n-1} \, a^{n-1} + \alpha_n \, a^n \in A. \tag{25.5}$$

Definition 25.7 (Einsetzungshomomorphismus)**.**

Es sei $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K mit Eins. Für $a \in A$ heißt die Abbildung

$$\varphi_a \colon K[t] \ni p \mapsto \widetilde{p}_A(a) \in A \tag{25.6}$$

der **Einsetzungshomomorphismus** (englisch: substitution homomorphism) oder der **Auswertungshomomorphismus** (englisch: evaluation homomorphism) **zu** a. △

Lemma 25.8 (Einsetzungshomomorphismus ist Homomorphismus von Algebren mit Eins)**.**

Es sei $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K mit Eins. Für jedes $a \in A$ ist der Einsetzungshomomorphismus $\varphi_a \colon K[t] \to A$ ein Homomorphismus von Algebren mit Eins.

Beweis. $(K[t], +, \cdot, \cdot)$ ist eine Algebra über dem Körper K mit Eins (Beispiel 25.3). Dasselbe gilt nach Voraussetzung für $(A, +, \cdot, \star)$. Es sei $a \in A$. Weiter seien $p, q \in K[t]$ und $\alpha \in K$. Wir füllen p oder q ggf. mit Nullkoeffizienten auf, sodass wir p und q in der Form

$$p = \sum_{i=0}^{n} \alpha_i \, t^i \quad \text{und} \quad q = \sum_{i=0}^{n} \beta_i \, t^i$$

schreiben können, wobei die Koeffizienten $\alpha_i, \beta_i \in K$ liegen.

Wir müssen die drei Bedingungen auf (25.3) nachweisen. Es gilt

$$\begin{aligned}
\varphi_a(p+q) &= \varphi_a\Big(\sum_{i=0}^n \alpha_i\, t^i + \sum_{i=0}^n \beta_i\, t^i\Big) && \text{wegen der Form von } p \text{ und } q \\
&= \varphi_a\Big(\sum_{i=0}^n (\alpha_i + \beta_i)\, t^i\Big) && \text{nach Definition der Addition von Polynomen (11.3a)} \\
&= \sum_{i=0}^n (\alpha_i + \beta_i)\, a^i && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6)} \\
&= \sum_{i=0}^n \alpha_i\, a^i + \sum_{i=0}^n \beta_i\, a^i && \text{weil } (A, +, \cdot) \text{ ein Vektorraum ist} \\
&= \varphi_a(p) + \varphi_a(q) && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6).}
\end{aligned}$$

Weiter gilt

$$\begin{aligned}
\varphi_a(\alpha\, p) &= \varphi_a\Big(\alpha \sum_{i=0}^n \alpha_i\, t^i\Big) && \text{wegen der Form von } p \text{ und } q \\
&= \varphi_a\Big(\sum_{i=0}^n (\alpha\, \alpha_i)\, t^i\Big) && \text{nach Definition der S-Multiplikation von Polynomen (12.6)} \\
&= \sum_{i=0}^n (\alpha\, \alpha_i)\, a^i && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6)} \\
&= \alpha \sum_{i=0}^n \alpha_i\, a^i && \text{weil } (A, +, \cdot) \text{ ein Vektorraum ist} \\
&= \alpha\, \varphi_a(p) && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6).}
\end{aligned}$$

Schließlich gilt

$$\begin{aligned}
\varphi_a(p \cdot q) &= \varphi_a\Big(\Big(\sum_{i=0}^n \alpha_i\, t^i\Big) \cdot \Big(\sum_{i=0}^n \beta_i\, t^i\Big)\Big) && \text{wegen der Form von } p \text{ und } q \\
&= \varphi_a\Big(\sum_{k=0}^{2n} \sum_{i=0}^k \alpha_i\, \beta_{k-i}\, t^k\Big) && \text{nach Def. der Multipl. von Polynomen (11.3b) und (11.3c)} \\
&= \sum_{k=0}^{2n} \sum_{i=0}^k \alpha_i\, \beta_{k-i}\, a^k && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6)} \\
&= \Big(\sum_{i=0}^n \alpha_i\, a^i\Big) \cdot \Big(\sum_{i=0}^n \beta_i\, a^i\Big) && \text{weil } (A, +, \cdot) \text{ ein Ring ist} \\
&= \varphi_a(p) \cdot \varphi_a(q) && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6)}
\end{aligned}$$

Schließlich wird durch φ_a das Einspolynom in $K[t]$ auf das Einselement in $e \in A$ abgebildet:

$$\begin{aligned}
\varphi_a(1\, t^0) &= 1\, a^0 && \text{nach Definition von } \varphi_a\text{, siehe (25.5) und (25.6)} \\
&= 1\, e && \text{nach Definition von } a^0 \\
&= e && \text{weil } (A, +, \cdot) \text{ ein Vektorraum ist.}
\end{aligned}$$

□

Quizfrage 25.3: Warum wird die Kommutativität der Algebra A in Lemma 25.8 nicht benötigt?

Bemerkung 25.9 (Festlegung des Einsetzungshomomorphismus).
Es sei $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K mit Einselement e. Der Einsetzungshomomorphismus zu $a \in A$, $\varphi_a \colon K[t] \to A$, ist durch

$$\varphi_a(1) = e \quad \text{(das Einspolynom } 1 \in K[t] \text{ wird auf das Einselement } e \in A \text{ abgebildet)} \tag{25.7a}$$
$$\varphi_a(t) = a \quad \text{(das Polynom } t \in K[t] \text{ wird auf das Element } a \in A \text{ abgebildet)} \tag{25.7b}$$

eindeutig festgelegt. △

Bemerkung 25.10 (induzierte Polynomfunktion, vgl. Bemerkung 11.17).
Es sei $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine Algebra über K mit Eins. Die Menge A^A der Funktionen $A \to A$, ausgestattet mit den punktweisen Verknüpfungen $+$, $\cdot$ und $\star$ aus A, bildet eine K-Algebra mit Eins $(A^A, +, \cdot, \star)$ (Beispiel 25.3). Die Abbildung

$$\Phi \colon (K[t], +, \cdot, \cdot) \ni p \mapsto \widetilde{p}_A \in (A^A, +, \cdot, \star) \tag{25.8}$$

ist ein Homomorphismus von zwei Algebren mit Eins.

Wie wir in Bemerkung 11.17 bereits gesehen haben, ist Φ i. A. nicht injektiv. Verschiedene Polynome können also dieselbe Polynomfunktion induzieren. Es gilt jedoch das Analogon von Folgerung 11.23: Φ ist injektiv, falls K unendlich und A nicht die Nullalgebra ist.

Φ ist i. A. auch nicht surjektiv. $\operatorname{Bild}(\Phi)$ ist die Unteralgebra mit Eins der Polynomfunktionen der Algebra mit Eins $(A^A, +, \cdot, \star)$. △

Beispiel 25.11 (Einsetzen von Algebraelementen in ein Polynom).

(*i*) Wir betrachten einen Körper $(K, +, \cdot)$ und die K-Algebra mit Eins der Matrizen $(K^{n \times n}, +, \cdot, \cdot)$. Das Einsetzen einer Matrix $A \in K^{n \times n}$ in das Polynom $p = t^2 - 1$ ergibt

$$\widetilde{p}(A) = A^2 - A^0 = A \cdot A - I.$$

(*ii*) Wir betrachten einen Körper $(K, +, \cdot)$, einen K-Vektorraum $(V, +, \cdot)$ und die K-Algebra mit Eins der Endomorphismen $(\operatorname{End}(V), +, \cdot, \circ)$. Das Einsetzen eines Endomorphismus $f \in \operatorname{End}(V)$ in das Polynom $p = t^2 - 1$ ergibt

$$\widetilde{p}(f) = f^2 - f^0 = f \circ f - \mathrm{id}_V.$$ △

Ende der Vorlesung 14

Ende der Woche 7

Wir schließen den Abschnitt mit der Beobachtung, dass Algebrahomomorphismen mit Polynomen kommutieren, genauer:

Satz 25.12 (Algebrahomomorphismen kommutieren mit Polynomen).
Es sei K ein Körper und A_1, A_2 zwei Algebren mit Eins über K. Weiter sei $f \colon A_1 \to A_2$ ein Homormophismus von Algebren mit Eins und $a \in A_1$. Dann gilt für die Einsetzungshomomorphismen φ_a von A_1 und $\psi_{f(a)}$ von A_2:

$$f \circ \varphi_a = \psi_{f(a)}. \tag{25.9}$$

Das bedeutet, für jedes $a \in A_1$ und jedes Polynom $p \in K[t]$ gilt:

$$f(\widetilde{p}(a)) = \widetilde{p}(f(a)). \tag{25.10}$$

Mit anderen Worten, das folgende Diagramm kommutiert:

$$\begin{tikzcd} & K[t] \arrow[dl, "\varphi_a"'] \arrow[dr, "\varphi_{f(a)}"] & \\ A_1 \arrow[rr, "f"'] & & A_2 \end{tikzcd}$$

Beweis. Wir zeigen zunächst, dass (25.9) und (25.10) dieselben Aussagen sind. Nach Definition 25.7 bildet $\varphi_a \colon K[t] \to A_1$ ab, und zwar durch

$$\varphi_a(p) = \widetilde{p}(a).$$

Nach Definition der Komposition bildet also $f \circ \varphi_a \colon K[t] \to A_2$ ab, und zwar durch

$$(f \circ \varphi_a)(p) = f(\varphi_a(p)) = f(\widetilde{p}(a)).$$

Für $b \in A_2$ bildet wiederum nach Definition 25.7 $\psi_b \colon K[t] \to A_2$ ab, und zwar durch

$$\psi_b(p) = \widetilde{p}(b),$$

speziell für $b = f(a)$ erhalten wir also

$$\psi_{f(a)}(p) = \widetilde{p}(f(a)).$$

Die Aussage, dass die Abbildungen in (25.9) für beliebige Elemente $p \in K[t]$ übereinstimmen, ist also gerade (25.10).

Wir zeigen nun (25.10). Dazu sei $p = \sum_{i=0}^n \alpha_i\, t^i \in K[t]$ beliebig. Dann gilt

$$\begin{aligned} f(\widetilde{p}(a)) &= f\Big(\sum_{i=0}^n \alpha_i\, a^i\Big) \\ &= \sum_{i=0}^n f(\alpha_i\, a^i) && \text{da } f \text{ additiv ist, siehe (25.3a)} \\ &= \sum_{i=0}^n \alpha_i\, f(a^i) && \text{da } f \text{ homogen ist, siehe (25.3b)} \\ &= \sum_{i=0}^n \alpha_i\, f(a)^i && \text{da } f \text{ mit der Multiplikation verträglich ist, siehe (25.3c)} \\ &= \widetilde{p}(f(a)), \end{aligned}$$

was zu zeigen war. □

§ 26 Der Satz von Cayley-Hamilton

Satz 26.1 (Satz von Cayley-Hamilton[6]).
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Dann gilt $\widetilde{\chi_A}(A) = 0 \in K^{n \times n}$.

[6] englisch: Cayley-Hamilton theorem

(*ii*) Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f \in \operatorname{End}(V)$. Dann gilt $\widetilde{\chi_f}(f) = 0 \in \operatorname{End}(V)$.

Die Aussage (*i*) besagt also, dass eine beliebige Matrix A, eingesetzt in ihr eigenes charakteristisches Polynom, immer die Nullmatrix ergibt. Aussage (*ii*) bedeutet, dass ein beliebiger Endomorphismus f eines endlich-dimensionalen Vektorraumes, eingesetzt in sein eigenes charakteristisches Polynom, immer den Nullendomorphismus ergibt. Man sagt auch, das charakteristische Polynom **annulliere** seine Matrix bzw. seinen Endomorphismus.

Beweis. Wir beweisen die Aussage hier unter der Voraussetzung Fall $n \geqslant 1$. (**Quizfrage 26.1:** Warum gilt sie auch im Fall $n = 0$?) Aussage (*i*): Wir fassen $C := \lambda I - A$ als Matrix in $K(\lambda)^{n \times n}$ über dem Körper $K(\lambda)$ auf. Wir betrachten die Adjunkte $\operatorname{adj}(C) \in K(\lambda)^{n \times n}$. Nach Lemma 23.15 gilt dann

$$\operatorname{adj}(C)\, C = \operatorname{adj}(C)\, (\lambda I - A) = \det(C)\, I = \chi_A\, I \in K(\lambda)^{n \times n}. \tag{26.1}$$

Die Terme $C = \lambda I - A$ sowie I in dieser Gleichung sind Elemente von $K(\lambda)^{n \times n}$, und χ_A ist Element des Körpers $K(\lambda)$.

Die Einträge der Adjunkten $\operatorname{adj}(C)$ sind nach Definition 23.12 vorzeichenbehaftete Unterdeterminanten von C. Die Einträge von $\operatorname{adj}(C)$ liegen daher in $K(\lambda)$ und $\operatorname{adj}(C)$ insgesamt in $K(\lambda)^{n \times n}$. Genauer liegen die Einträge $(\operatorname{adj}(C))_{ij} = \widetilde{c}_{ji} = \det\big((C)_{\neq i, \neq j}\big)$ sogar im Unterring $K[\lambda] \subsetneq K(\lambda)$, sind also rationale Ausdrücke mit Nenner $1 \in K \subsetneq K(\lambda)$, weil die Einträge von C im Unterring $K[\lambda]$ liegen und dieser multiplikativ und additiv abgeschlossen ist. Als Unterdeterminanten von C sind die Einträge in $\operatorname{adj}(C)$ schließlich Polynome vom Höchstgrad $n - 1$.

Wir können daher $\operatorname{adj} C \in K(\lambda)^{n \times n}$ als Linearkombination

$$\operatorname{adj}(C) = \sum_{j=0}^{n-1} \underbrace{\lambda^j}_{\in K(\lambda)} \overbrace{B_j}^{\in K^{n \times n} \subsetneq K(\lambda)^{n \times n}}$$

im Vektorraum $K(t)^{n \times n}$ schreiben. Alle weiteren Rechnungen finden in der Algebra $K(\lambda)^{n \times n}$ statt. Es gilt

$$\begin{aligned}
\chi_A\, I &= \operatorname{adj}(C)\, C && \text{wegen (26.1)} \\
&= \Big(\sum_{j=0}^{n-1} \lambda^j B_j\Big) C \\
&= \sum_{j=0}^{n-1} \lambda^j B_j\, (\lambda I - A) && \text{nach Definition von } C \text{ und Distributivgesetz} \\
&= \sum_{j=0}^{n-1} \lambda^{j+1} B_j - \sum_{j=0}^{n-1} \lambda^j B_j\, A && \text{nach Distributivgesetz und (25.2)} \\
&= \sum_{j=0}^{n} \lambda^j\, (B_{j-1} - B_j\, A) && \text{mit den Definitionen } B_{-1} = B_n := 0 \in K^{n \times n}.
\end{aligned} \tag{26.2}$$

Das charakteristische Polynom χ_A hat eine Darstellung

$$\chi_A = \sum_{j=0}^{n} c_j\, \lambda^j$$

mit Koeffizienten $c_j \in K$, insbesondere $c_n = 1$. Es gilt

$$\begin{aligned}
\chi_A\, I &= \Big(\sum_{j=0}^{n} c_j\, \lambda^j\Big)\, I \\
&= \sum_{j=0}^{n} c_j\, \lambda^j\, I && \text{nach Distributivgesetz} \\
&= \sum_{j=0}^{n} \lambda^j\, (c_j\, I) && \text{nach (25.2).}
\end{aligned} \tag{26.3}$$

Wir vergleichen jetzt die beiden rechten Seiten in (26.2) und (26.3). Beides sind Matrizen in $K(\lambda)^{n\times n}$. Gleichheit liegt genau dann vor, wenn alle Einträge übereinstimmen. Jeder Eintrag dieser Matrizen ist ein Element von $K(\lambda)$, genauer sogar ein Polynom in $K_n[\lambda]$. Diese stimmen genau dann überein, wenn alle Koeffizienten übereinstimmen. Führen wir den Koeffizientenvergleich für $j = n, n-1, \ldots, 1, 0$ durch, so erhalten wir

$$\begin{aligned}
I = c_n\, I &= B_{n-1} - B_n\, A = B_{n-1} && \text{im Fall } j = n \\
c_j\, I &= B_{j-1} - B_j\, A && \text{im Fall } j = n-1, \ldots, 1 \\
c_i\, I &= B_{-1} - B_0\, A = -B_0\, A && \text{im Fall } j = 0.
\end{aligned}$$

Wir setzen nun A in die Polynomfunktion zu χ_A ein:

$$\begin{aligned}
\widetilde{\chi_A}(A) &= \sum_{j=0}^{n} c_j\, A^j = \sum_{j=0}^{n} c_j\, I\, A^j \\
&= c_n\, I\, A^n + \Big(\sum_{j=1}^{n-1} c_j\, I\, A^j\Big) + c_0\, I\, A^0 \\
&= B_{n-1}\, A^n + \Big(\sum_{j=1}^{n-1} (B_{j-1} - B_j\, A)\, A^j\Big) - B_0\, A && \text{Einsetzen der Koeffizienten } c_j \\
&= B_{n-1}\, A^n + \Big(\sum_{j=1}^{n-1} B_{j-1}\, A^j\Big) - \Big(\sum_{j=1}^{n-1} B_j\, A^{j+1}\Big) - B_0\, A \\
&= B_{n-1}\, A^n + \Big(\sum_{j=0}^{n} B_j\, A^{j+1}\Big) - \Big(\sum_{j=1}^{n-1} B_j\, A^{j+1}\Big) - B_0\, A \\
&= 0.
\end{aligned}$$

Der Beweis von Aussage *(ii)* ist Gegenstand von Hausaufgabe II-8.1. □

Bemerkung 26.2 (zu Satz 26.1)**.**

Folgende Argumentation zum Beweis des Satzes von Cayley-Hamilton 26.1 wäre falsch: Das Einsetzen von A an Stelle von λ in das charakteristische Polynom $\chi_A = \det(\lambda\, I - A)$ ergibt

$$\widetilde{\chi_A}(A) = \det(A\, I - A) = \det(A - A) = \det(0) = 0.$$

Quizfrage 26.2: Wo liegt der Fehler? △

Beispiel 26.3 (Satz von Cayley-Hamilton).
Die Matrix $A = \left[\begin{smallmatrix} 2 & 1 \\ -1 & -3 \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$ besitzt das charakteristische Polynom

$$\chi_A = \det \begin{pmatrix} \lambda - 2 & -1 \\ 1 & \lambda - 3 \end{pmatrix} = (\lambda - 2)\,(\lambda + 3) + 1 = \lambda^2 + \lambda - 5.$$

Das Einsetzen von A ergibt

$$\begin{aligned}
\widetilde{\chi_A}(A) &= A^2 + A - 5\,I \\
&= \begin{bmatrix} 2 & 1 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -3 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ -1 & -3 \end{bmatrix} - 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 3 & -1 \\ 1 & 8 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ -1 & -3 \end{bmatrix} + \begin{bmatrix} -5 & 0 \\ 0 & -5 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

wie erwartet. △

Folgerung 26.4 (A^{-1} ist ein Polynom in A).
Es sei K ein Körper und $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$. Wenn A invertierbar ist, dann gibt es ein Polynom $p \in K_{n-1}[t]$ mit der Eigenschaft[7]

$$A^{-1} = \widetilde{p}(A). \tag{26.4}$$

Beweis. Das charakteristische Polynom hat nach Satz 24.20 die Form

$$\chi_A = \sum_{i=0}^{n} \alpha_i\,\lambda^i \in K_n[\lambda].$$

Für den niedrigsten Koeffizienten gilt $\alpha_0 = (-1)^n \det(A) \neq 0$, da A invertierbar ist. Setzen wir die Matrix A ein, so ergibt sich nach dem Satz von Cayley-Hamilton 26.1:

$$0 = \widetilde{\chi_A}(A) = \alpha_n\,A^n + \cdots + \alpha_1\,A + \alpha_0\,I,$$

also

$$\alpha_0\,I = -\sum_{j=1}^{n} \alpha_j\,A_j = A\Big(-\sum_{j=1}^{n} \alpha_j\,A_{j-1}\Big).$$

Die Multiplikation mit α_0^{-1} zeigt die Behauptung. □

Beispiel 26.5 (A^{-1} ist ein Polynom in A).
Wir betrachten die Matrix $A = \left[\begin{smallmatrix} 2 & 1 \\ -1 & -3 \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$, die das charakteristische Polynom

$$\chi_A = \lambda^2 + \lambda - 5$$

besitzt, siehe Beispiel 26.3. Nach dem Satz von Cayley-Hamilton 26.1 gilt also $A^2 + A - 5\,I = 0$ und damit

$$I = \frac{1}{5}A^2 + \frac{1}{5}A = A\Big(\frac{1}{5}A + \frac{1}{5}I\Big).$$

Das bedeutet:

$$A^{-1} = \frac{1}{5}A + \frac{1}{5}I. \qquad △$$

[7]Im Fall $n = 0$ ist mit $p \in K_{-1}[t]$ das Nullpolynom gemeint.

§ 27 Ideale in Ringen

Ideale in Ringen sind spezielle Unterringe, die dieselbe Funktion einnehmen wie Normalteiler in Gruppen. Mit ihrer Hilfe können wir beispielsweise Faktorringe definieren und einen Homomorphiesatz für Ringe erhalten.

Definition 27.1 (Ideal, vgl. Definition 8.10)**.**
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
$$R\,J \subseteq J \quad \text{und} \quad J\,R \subseteq J \tag{27.1}$$
gilt.[8] Manchmal notiert man die Eigenschaft, dass $(J, +, \cdot)$ ein Ideal des Ringes $(R, +, \cdot)$ ist, als $(J, +, \cdot) \trianglelefteq (R, +, \cdot)$.

Beachte: Das ist genau dann erfüllt, wenn $(J, +)$ eine Untergruppe von $(R, +)$ ist und wenn (27.1) gilt, denn (27.1) impliziert bereits, dass J bzgl. $\cdot$ abgeschlossen ist.

(ii) Ein Ideal $(J, +, \cdot)$ von $(R, +, \cdot)$ heißt **echt** (englisch: proper ideal), wenn $J \subsetneq R$ gilt. △

Lemma 27.2 (Kerne von Ringhomomorphismen sind Ideale)**.**
Es seien $(R_1, +, \cdot)$ und $(R_2, +, \cdot)$ Ringe und $f \colon R_1 \to R_2$ ein Homomorphismus. Dann gilt:
$$\operatorname{Kern}(f) = \{a \in R_1 \mid f(a) = 0\}$$
ist ein Ideal von R_1.

Beweis. Nach Lemma 8.7 ist $\operatorname{Kern}(f)$ eine Untergruppe von $(R_1, +)$. Es sei nun $a \in R$ und $j \in \operatorname{Kern}(f)$. Dann gilt
$$f(a \cdot j) = f(a) \cdot f(j) = f(a) \cdot 0 = 0$$
nach Lemma 9.3 und analog
$$f(j \cdot a) = f(j) \cdot f(a) = 0 \cdot f(a) = 0.$$
Also gehören $a \cdot j$ und $j \cdot a$ wieder zu J. □

Beispiel 27.3 (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) Für $m \in \mathbb{N}$ ist $(m\,\mathbb{Z}, +, \cdot)$ ein Ideal von $(\mathbb{Z}, +, \cdot)$.

(iii) Es sei $(R, +, \cdot)$ ein kommutativer Ring. Für $m \in \mathbb{N}_0$ ist $\left\{p \in R[t] \,\middle|\, p = \alpha_m\, t^m + \alpha_{m+1}\, t^{m+1} + \cdots\right\} = t^m\, R[t]$ ein Ideal von $(R[t], +, \cdot)$.

△

Lemma 27.4 (Durchschnitt von Idealen)**.**
Es sei $(R, +, \cdot)$ ein Ring und $(J_i, +, \cdot)_{i \in I}$ eine Familie von Idealen mit der nichtleeren Indexmenge I. Dann ist auch $\bigcap_{i \in I} J_i$ ein Ideal in J.

[8] Ausgeschrieben heißt das also: $a \in R$ und $j \in J$ impliziert $a\,j \in J$ und $j\,a \in J$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-8.2. □

Definition 27.5 (erzeugtes Ideal, Hauptideal)**.**
Es sei $(R, +, \cdot)$ ein Ring und $E \subseteq R$.

(*i*) Dann heißt

$$(E) := \bigcap \big\{ J \,\big|\, (J, +, \cdot) \text{ ist Ideal von } (R, +, \cdot) \text{ und } E \subseteq J \big\} \tag{27.2}$$

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 gilt: $(a) = J$. △

Satz 27.6 (Darstellung des erzeugten Ideals)**.**

(*i*) Es sei $(R, +, \cdot)$ ein Ring, $E \subseteq R$ und $a \in R$. Dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in E \;\cup\; -E \;\cup\; R\,E \;\cup\; E\,R \;\cup\; R\,E\,R\big) \Big\}, \tag{27.3a}$$

$$(a) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in \{\pm a\} \;\cup\; R\,a \;\cup\; a\,R \;\cup\; R\,a\,R\big) \Big\}. \tag{27.3b}$$

(*ii*) Es sei $(R, +, \cdot)$ ein Ring mit Eins, $E \subseteq R$ und $a \in R$. Dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in R\,E\,R\big) \Big\}, \tag{27.4a}$$

$$(a) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in R\,a\,R\big) \Big\}. \tag{27.4b}$$

Insbesondere ist $(1) = R$.

(*iii*) Es sei $(R, +, \cdot)$ ein kommutativer Ring, $E \subseteq R$ und $a \in R$. Dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in E \;\cup\; -E \;\cup\; R\,E\big) \Big\}, \tag{27.5a}$$

$$(a) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in \{\pm a\} \;\cup\; R\,a\big) \Big\}. \tag{27.5b}$$

(*iv*) Es sei $(R, +, \cdot)$ ein kommutativer Ring mit Eins, $E \subseteq R$ und $a \in R$. Dann gilt für das von E bzw. von a erzeugte Ideal:

$$(E) = \Big\{ \sum_{i=1}^n a_i \,\Big|\, \exists n \in \mathbb{N}_0 \; \forall i = 1, \ldots, n \; \big(a_i \in R\,E\big) \Big\}, \tag{27.6a}$$

$$(a) = R\,a. \tag{27.6b}$$

Insbesondere ist $(1) = R$.

In jedem Fall gilt $(\emptyset) = (0) = \{0\}$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-8.2. □

Es gibt Ringe, in denen jedes Ideal ein Hauptideal ist, also von einem einzigen Ringelement erzeugt werden kann. Integritätsringe (kommutative, nullteilerfreie Ringe mit Eins ungleich dem Nullring) mit dieser Eigenschaft heißen **Hauptidealringe** (englisch: principal ideal domains).

Satz 27.7 (der Polynomring über einem Körper ist ein Hauptidealring)**.**
Es sei K in Körper. Dann gilt:

(i) Zu jedem Ideal J in $K[t]$ existiert ein $p \in K[t]$ mit der Eigenschaft $J = (p)$.

(ii) Das Polynom p aus Aussage (i) ist bis auf einen Faktor $\alpha \in K$ eindeutig bestimmt.

Es gilt also: Sind $p_1, p_2 \in K[t]$ zwei Polynome mit der Eigenschaft $J = (p_1) = (p_2)$, dann gilt $p_2 = \alpha\, p_1$ mit einem $\alpha \in K \setminus \{0\}$.

(iii) Ist $J \neq \{0\}$, dann ist p aus Aussage (i) eines der Polynome minimalen Grades in $J \setminus \{0\}$.

(iv) Ist $J = \{0\}$, dann gilt $p = 0$.

Beweis. Aussage (i): Im Fall $J = \{0\}$ wählen wir $p = 0$. Nach (27.3b) gilt $J = (p)$.

Nun sei $J \neq \{0\}$, dann besitzt die Menge $\big\{\deg(p) \,\big|\, p \in J \setminus \{0\}\big\}$ ein Minimum $m \in \mathbb{N}_0$. Es sei $p \in J \setminus \{0\}$ ein Polynom mit $\deg(p) = m$.

Schritt 1: $(p) \subseteq J$:

Wegen $p \in J$ haben wir $(p) \subseteq (J)$, und da J bereits in Ideal ist, gilt $(J) = J$.

Schritt 2: $J \subseteq (p)$:

Es sei $\bar{p} \in J$ beliebig. Nach Satz 11.14 gibt es (eindeutig bestimmte) Polynome $q, r \in K[t]$ mit $\deg(r) < \deg(p)$ und

$$\bar{p} = q\, p + r.$$

Daraus folgt

$$r = \bar{p} - q\, p.$$

Weil J ein Ideal ist und $\bar{p}, p \in J$ liegen, ist auch $r \in J$. (**Quizfrage 27.1:** Klar?) Da aber p ein Element minimalen Grades von $J \setminus \{0\}$ ist, muss $r = 0$ gelten, d. h., wir haben $\bar{p} = q\, p \in K[t]\, J \subseteq J$.

Aussage (ii): Es gelte $J = (p_1) = (p_2)$. Da $K[t]$ ein kommutativer Ring mit Eins ist, gilt nach Satz 27.6 $(p_1) = K[t]\, p_1$ und $(p_2) = K[t]\, p_2$. Aus $p_1 \in (p_2)$ folgt also die Existenz von $q_1 \in K[t]$ mit

$$p_1 = q_1\, p_2,$$

und analog folgt aus $p_2 \in (p_1)$ die Existenz von $q_2 \in K[t]$ mit

$$p_2 = q_2\, p_1.$$

Zusammen haben wir also

$$p_1 = q_1\, q_2\, p_1 \quad \text{und damit} \quad p_1\,(q_1\, q_2 - 1) = 0.$$

$K[t]$ ist nullteilerfrei (Folgerung 11.11). Es muss also $p_1 = 0$ oder $q_1\, q_2 = 1$ gelten. Im ersten Fall folgt $J = \{0\}$ und damit auch $p_2 = 0$. Im zweiten Fall folgt aus Lemma 11.10 $\deg(q_1\, q_2) = \deg(q_1) + \deg(q_2) = \deg(1) = 0$. Daraus folgt $\deg(q_1) = \deg(q_2) = 0$. Das heißt aber $q_2 = \alpha \in K$ mit $\alpha \neq 0$ und $q_1 = \alpha^{-1}$. Damit folgt wie behauptet $p_2 = \alpha\, p_1$.

Aussage (iii): Im Beweis von Aussage (i) wurde p als ein Polynom in $J \setminus \{0\}$ minimalen Grades gewählt. Aussage (ii) zeigt, dass jedes andere erzeugende Polynom von J ein Vielfaches von p ist (mit Faktor ungleich 0), also ebenfalls ein Polynom in $J \setminus \{0\}$ minimalen Grades.

Aussage (iv): Diese Aussage ist klar. □

Ende der Vorlesung 15

§ 28 Das Minimalpolynom

Wir betrachten zu gegebener Matrix $A \in K^{n \times n}$ die Menge derjenigen Polynome, die A beim Einsetzen auf die Nullmatrix abbilden (die A annullieren). Das charakteristische Polynom ist nach dem Satz von Cayley-Hamilton 26.1 eines davon, aber wir interessieren uns für das Polynom kleinsten Grades. Wir untersuchen dazu zunächst die Struktur dieser Menge von Polynomen mit der genannten Eigenschaft:

Lemma 28.1 (die annullierenden Polynome bilden ein Ideal)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Die Menge

$$J_A := \left\{ p \in K[\lambda] \,\middle|\, \widetilde{p}(A) = 0 \right\} \tag{28.1}$$

ist ein Ideal in $K[\lambda]$ ungleich dem Nullideal.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Die Menge

$$J_f := \left\{ p \in K[\lambda] \,\middle|\, \widetilde{p}(f) = 0 \right\} \tag{28.2}$$

ist ein Ideal in $K[\lambda]$ ungleich dem Nullideal.

Beachte: Das Nullpolynom und das charakteristische Polynom χ_A sind Elemente von J_A. Ebenso sind das Nullpolynom und das charakteristische Polynom χ_f Elemente von I_f.

Beweis. Aussage (i): Wir betrachten den Einsetzungshomomorphismus (Definition 25.7) zu A in der Matrixalgebra $K^{n \times n}$, also

$$\varphi_A \colon K[t] \ni p \mapsto \widetilde{p}(A) \in K^{n \times n}.$$

Insbesondere ist φ_A ein Homomorphismus von Ringen mit Eins, und es gilt

$$J_A = \operatorname{Kern}(\varphi_A).$$

Nach Lemma 27.2 ist J_A also ein Ideal. Das charakteristische Polynom χ_A gehört nach dem Satz von Cayley-Hamilton 26.1 zu J_A, und es ist ungleich dem Nullpolynom. Also ist J_A nicht das Nullideal.

Aussage (ii) kann analog gezeigt werden. □

Nach Satz 27.7 ist jedes gradminimale Polynom in J_A ein Erzeuger des Ideals J_A. Durch Normierung ist dieses Polynom sogar eindeutig bestimmt. Dasselbe gilt für J_f und führt zu folgender Definition.

Definition 28.2 (Minimalpolynom).
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Das eindeutig bestimmte normierte Polynom $\mu_A \neq 0$ geringsten Grades mit der Eigenschaft $\mu_A \in J_A$, also $\widetilde{\mu_A}(A) = 0$, heißt das **Minimalpolynom** (englisch: minimal polynomial) **von** A.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Weiter sei $f \in \operatorname{End}(V)$. Das eindeutig bestimmte normierte Polynom $\mu_f \neq 0$ geringsten Grades mit der Eigenschaft $\mu_f \in J_f$, also $\widetilde{\mu_f}(f) = 0$, heißt das **Minimalpolynom von** f. △

Wir zeigen zunächst einen wichtigen Zusammenhang zwischen dem Minimalpolynom eines Endomorphismus und dem einer beliebigen Darstellungsmatrix:

Satz 28.3 (Minimalpolynome eines Endomorphismus und seiner Darstellungsmatrizen).
Es sei K ein Körper und V ein endlich-dimensionaler Vektorraum über K mit der Basis B_V. Weiter sei $f \in \operatorname{End}(V)$ und $A = \mathcal{M}_{B_V}^{B_V}(f)$. Dann gilt

$$\mu_f = \mu_A. \tag{28.3}$$

Beachte: Die Gleichheit haben wir auch für die charakteristischen Polynome, allerdings aufgrund der Definition 24.23.

Beweis. Wir betrachten den Algebraisomorphismus $\Phi\colon (\operatorname{End}(V), +, \cdot, \circ) \to (K^{n \times n}, +, \cdot, \cdot)$, der durch $\Phi(f) = \mathcal{M}_{B_V}^{B_V}(f)$ gegeben ist, siehe Beispiel 25.6. Nach Voraussetzung gilt $A = \Phi(f)$.

Nach Satz 25.12 gilt für jedes $p \in K[t]$

$$\Phi(\widetilde{p}(f)) = \widetilde{p}(\Phi(f)) = \widetilde{p}(A),$$

mit anderen Worten: Die Darstellungsmatrix des Endomorphismus $\widetilde{p}(f)$ ist gerade $\widetilde{p}(A)$. Das heißt aber auch:

$$\widetilde{p}(f) = 0 \quad \Rightarrow \quad \Phi(\widetilde{p}(f)) = \widetilde{p}(A) = 0.$$

Da Φ also Isomorphismus injektiv ist, gilt auch ~~Mit Hilfe des inversen Isomorphismus Φ^{-1} zeigen wir analog~~

$$\widetilde{p}(A) = 0 \quad \Rightarrow \quad \widetilde{p}(f) = 0.$$

Damit sind die Ideale J_A und J_f aus (28.1) und (28.2) identisch, also auch die Minimalpolynome μ_f und μ_A. □

Folgerung 28.4 (ähnliche Matrizen haben dasselbe Minimalpolynom).
Es sei K ein Körper und $n \in \mathbb{N}_0$. Sind $A, \widehat{A} \in K^{n \times n}$ ähnlich, dann sind ihre Minimalpolynome identisch, also $\mu_A = \mu_{\widehat{A}}$.

Beweis. Das Resultat folgt sofort aus Satz 28.3 und der Tatsache, dass zwei Matrizen genau dann ähnlich sind, wenn sie Darstellungsmatrizen ein- und desselben Endomorphismus sind (Satz 20.14). □

§ 28.1 Bestimmung des Minimalpolynoms

Setzen wir in ein Polynom $p \in K[t]$ mit der Darstellung

$$p = \sum_{i=0}^{m} \alpha_i \, t^i$$

die Matrix A ein, so ergibt sich

$$\widetilde{p}(A) = \sum_{i=0}^{m} \alpha_i \, A^i.$$

Gilt $\widetilde{p}(A) = 0$, so bedeutet die äquivalente Bedingung

$$\sum_{i=0}^{m} \alpha_i \, A^i = 0, \tag{28.4}$$

dass die Familie der Matrizen $F_m = (A^0, A^1, \ldots, A^m)$ linear abhängig sind, sofern nicht alle Koeffizienten gleich Null sind. Wir können daher bei der Suche nach dem Minimalpolynom die Familien F_m, angefangen bei $m = 0$, auf lineare Unabhängigkeit prüfen. Dazu „vektorisieren" wir die Matrizen z. B. mit Hilfe des Vektorisierungsisomorphismus aus Beispiel 17.3 $\operatorname{vec}\colon K^{n \times n} \ni A \mapsto \operatorname{vec}(A) \in K^{n^2}$. Bezeichnen wir $v_i := \operatorname{vec}(A^i)$, dann ist (28.4) äquivalent zu

$$\sum_{i=0}^{m} \alpha_i \, v_i = 0, \tag{28.5}$$

also zu einem homogenen linearen Gleichungssystem. Wir prüfen durch Überführung auf Zeilenstufenform, ob dessen Koeffizientenmatrix $(v_0, v_1, \ldots, v_m)$ vollen Rang $m + 1$ hat. Falls nicht, so haben wir mit m den Grad des Minimalpolynoms gefunden. Wir bestimmen dann die Lösungsmenge des Gleichungssystems (28.5). Diese wird eindimensional sein. (**Quizfrage 28.1:** Woher wissen wir das?) In ihr finden wir genau einen Koeffizientenvektor $\begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_m \end{pmatrix}$ mit der Eigenschaft $\alpha_m = 1$. (**Quizfrage 28.2:** Warum ist das so?)

Beispiel 28.5 (Bestimmung des Minimalpolynoms)**.**

Wir bestimmen das Minimalpolynom der Matrix

$$A = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \in \mathbb{R}^{3 \times 3}.$$

Es gilt

$$A^2 = \begin{bmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{bmatrix} \quad \text{und} \quad A^3 = \begin{bmatrix} -18 & 9 & 9 \\ 9 & -18 & 9 \\ 9 & 9 & -18 \end{bmatrix}.$$

Wir betrachten jetzt die Matrix

$$\begin{bmatrix}\operatorname{vec}(A^0) & \operatorname{vec}(A^1) & \operatorname{vec}(A^2) & \operatorname{vec}(A^3)\end{bmatrix} = \begin{bmatrix} 1 & -2 & 6 & -18 \\ 0 & 1 & -3 & 9 \\ 0 & 1 & -3 & 9 \\ 0 & 1 & -3 & 9 \\ 1 & -2 & 6 & -18 \\ 0 & 1 & -3 & 9 \\ 0 & 1 & -3 & 9 \\ 0 & 1 & -3 & 9 \\ 1 & -2 & 6 & -18 \end{bmatrix}$$

Wir können bei A^3 aufhören, da ja spätestens das charakteristische Polynom $\chi_A = \lambda^3 + 6\,\lambda^2 + 9\,\lambda$ die Bedingung $\widetilde{\chi_A}(A) = 0$ erfüllen wird.

Die Spalten $\{0, 1\}$ zusammen sind noch linear unabhängig, die Spalten $\{0, 1, 2\}$ aber nicht mehr. Es gilt daher $m = 2$, und der Kern von $\begin{bmatrix}\operatorname{vec} A^0 & \operatorname{vec} A^1 & \operatorname{vec} A^2\end{bmatrix}$ besteht offenbar aus den Vielfachen von $\left(\begin{smallmatrix}0\\3\\1\end{smallmatrix}\right)$. Das Minimalpolynom von A ist also

$$\mu_A = \lambda^2 + 3\,\lambda. \qquad \triangle$$

§ 28.2 Weitere Eigenschaften des Minimalpolynoms

Lemma 28.6 (das Minimalpolynom teilt jedes Polynom, das seine Matrix annulliert)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$. Dann teilt das Minimalpolynom μ_A jedes Polynom in $p \in J_A$, also jedes $p \in K[t]$ mit der Eigenschaft $\widetilde{p}(A) = 0$.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Dann teilt das Minimalpolynom μ_f jedes Polynom in $p \in J_f$, also jedes $p \in K[t]$ mit der Eigenschaft $\widetilde{p}(f) = 0$.

Beweis. Aussage (i): Das Minimalpolynom μ_A ist nach Satz 27.7 ein Erzeuger des Ideals J_A von $K[t]$, kurz: $J_A = (\mu_A)$. Da $K[t]$ kommutativer Ring mit Eins ist, gilt nach (27.6b)

$$J_A = (\mu_A) = K[t]\,\mu_A.$$

Mit anderen Worten, für jedes $p \in J_A$ gibt es ein $q \in K[t]$ mit der Eigenschaft $p = q\,\mu_A$.

Aussage (ii) kann analog gezeigt werden. □

Folgerung 28.7 (das Minimalpolynom teilt das charakteristische Polynom)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$. Dann teilt das Minimalpolynom μ_A das charakteristische Polynom χ_A.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Dann teilt das Minimalpolynom μ_f das charakteristische Polynom χ_f.

Beweis. Die Aussage folgt sofort aus dem Satz von Cayley-Hamilton 26.1 und Lemma 28.6. □

Das folgende Beispiel stellt eine Alternative dar zur Bestimmung des Minimalpolynoms nach Beispiel 28.5. Es nutzt bereits die Aussage von Lemma 28.9.

Beispiel 28.8 (das Minimalpolynom teilt das charakteristische Polynom)**.**

(i) Die Matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{Q}^{2\times 2}$$

besitzt das charakteristische Polynom $\chi_A = (\lambda - 1)^2$. Da das Minimalpolynom ein Teiler von χ_A mit denselben Nullstellen ist, gilt entweder $\mu_A = (\lambda - 1)$ oder $\mu_A = (\lambda - 1)^2$. Durch Einsetzen finden wir, dass $\mu_A = (\lambda - 1)$ die richtige Wahl ist, da $A - I = 0$ gilt.

(ii) Die Matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{Q}^{2\times 2}$$

besitzt das charakteristische Polynom $\chi_A = (\lambda - 1)\,(\lambda + 1)$. Da das Minimalpolynom ein Teiler von χ_A mit denselben Nullstellen ist, bleibt nur die Möglichkeit $\mu_A = \chi_A$.

(iii) Die Matrix

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -8 & 1 & 4 \\ 2 & -1 & -1 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

besitzt das charakteristische Polynom $\chi_A = (\lambda+1)^2\,(\lambda-3)$. Da das Minimalpolynom ein Teiler von χ_A mit denselben Nullstellen ist, gilt entweder $\mu_A = (\lambda+1)\,(\lambda-3)$ oder $\mu_A = \chi_A = (\lambda+1)^2\,(\lambda-3)$. Da

$$(A+I)\,(A-3\,I) = \begin{bmatrix} 2 & -1 & 0 \\ -8 & 2 & 4 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ -8 & -2 & 4 \\ 2 & -1 & -4 \end{bmatrix} = \begin{bmatrix} 4 & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \neq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

gilt, muss $\mu_A = \chi_A = (\lambda + 1)^2\,(\lambda - 3)$ die richtige Wahl sein.

(iv) Die Matrix

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -6 & 1 & 2 \\ 3 & -1 & -2 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

besitzt dasselbe charakteristische Polynom $\chi_A = (\lambda + 1)^2\,(\lambda - 3)$ wie die Matrix oben. Da das Minimalpolynom ein Teiler von χ_A mit denselben Nullstellen ist, gilt wiederum entweder $\mu_A = (\lambda + 1)\,(\lambda - 3)$ oder $\mu_A = \chi_A = (\lambda + 1)^2\,(\lambda - 3)$. Da

$$(A+I)\,(A-3\,I) = \begin{bmatrix} 3 & -1 & -1 \\ -6 & 2 & 2 \\ 3 & -1 & -1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ -6 & -2 & 2 \\ 3 & -1 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

gilt, muss $\mu_A = \chi_A = (\lambda + 1)^{\cancel{2}}\,(\lambda - 3)$ die richtige Wahl sein. △

Wir zeigen nun einen weiteren Zusammenhang zwischen dem Minimalpolynom und dem charakteristischen Polynom:

Lemma 28.9 (Minimalpolynom und charakteristisches Polynom besitzen dieselben Nullstellen)**.**
Es sei K ein Körper.

(*i*) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Dann sind äquivalent:

(a) $\lambda \in K$ ist eine Nullstelle des Minimalpolynoms μ_A.

(b) $\lambda \in K$ ist eine Nullstelle des charakteristischen Polynoms χ_A.

(*ii*) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Dann sind äquivalent:

(a) $\lambda \in K$ ist eine Nullstelle des Minimalpolynoms μ_f.

(b) $\lambda \in K$ ist eine Nullstelle des charakteristischen Polynoms χ_f.

Beweis. Aussage (*i*):

Aussage (a) $\Rightarrow$ Aussage (b): Nach Folgerung 28.7 teilt das Minimalpolynom das charakteristische Polynom, d. h., es existiert ein $q \in K[t]$ mit der Eigenschaft $\chi_A = q\,\mu_A$. Daraus folgt

$$\widetilde{\chi_A}(\lambda) = \widetilde{q}(\lambda)\,\widetilde{\mu_A}(\lambda) = \widetilde{q}(\lambda)\,0 = 0.$$

Aussage (b) $\Rightarrow$ Aussage (a): Nach Lemma 24.14 ist λ ein Eigenwert von A. Es sei v ein zugehöriger Eigenvektor.

Das Minimalpolynom habe die Darstellung $\mu_A = \sum_{i=0}^m \alpha_i\, t^i$ mit $m \in \mathbb{N}_0$ und Koeffizienten $\alpha_i \in K$. Wir erhalten

$$\begin{aligned}
0 &= 0\,v && \text{„Der Nullvektor ist die Nullmatrix mal Vektor } v\text{“} \\
&= \widetilde{\mu_A}(A)\,v && \text{nach dem Satz von Cayley-Hamilton 26.1} \\
&= \Big(\sum_{i=0}^m \alpha_i\, A^i\Big)\,v \\
&= \sum_{i=0}^m \alpha_i\, A^i\, v && \text{nach Distributivgesetz der Matrix-Vektor-Multiplikation} \\
&= \sum_{i=0}^m \alpha_i\, \lambda^i\, v && \text{wegen } A\,v = \lambda\,v \\
&= \Big(\sum_{i=0}^m \alpha_i\, \lambda^i\Big)\,v && \text{nach Distributivgesetz im Vektorraum } K^n \\
&= \widetilde{\mu_A}(\lambda)\,v.
\end{aligned}$$

Weil aber $v \neq 0$ ist, muss $\widetilde{\mu_A}(\lambda) = 0$ gelten.

Aussage (*ii*) kann analog gezeigt werden. □

Mit Lemma 24.14 und Lemma 24.24 folgt aus Lemma 28.9 unmittelbar:

Folgerung 28.10 (Eigenwerte von Matrizen und Endomorphismen sind Nullstellen des Minimalpolynoms)**.**
Es sei K ein Körper.

(*i*) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Dann sind äquivalent:

(a) λ ist ein Eigenwert von A.

(b) λ ist eine Nullstelle des Minimalpolynoms μ_A.

(*ii*) Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Weiter sei $f \in \operatorname{End}(V)$ und $\lambda \in K$. Dann sind äquivalent:

(a) λ ist ein Eigenwert von f.

(b) λ ist eine Nullstelle des Minimalpolynoms μ_f.

Ende der Vorlesung 16

Ende der Woche 8

Lemma 28.11 (Minimalpolynom einer Diagonalmatrix)**.**
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$ eine Diagonalmatrix

$$A = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}.$$

Ist $\{b_1, \ldots, b_s\} = \{a_1, \ldots, a_n\}$ und sind $b_1, \ldots, b_s$ paarweise verschieden[9], dann gilt

$$\mu_A = (\lambda - b_1) \cdots (\lambda - b_s). \tag{28.6}$$

Beweis. Für jedes Polynom $p \in K[t]$ gilt

$$\widetilde{p}(A) = \begin{bmatrix} \widetilde{p}(a_1) & & \\ & \ddots & \\ & & \widetilde{p}(a_n) \end{bmatrix}.$$

(**Quizfrage 28.3:** Klar?) Ein Polynom $p \in K[t]$ erfüllt also $\widetilde{p}(a_j) = 0$ genau dann, wenn p den Linearfaktor $\lambda - a_j$ enthält (Lemma 11.20). Daraus folgt: $\widetilde{p}(A) = 0$ gilt genau dann, wenn p alle Linearfaktoren $\lambda - b_1, \ldots, \lambda - b_s$ enthält.

Per Definition ist das Minimalpolynom μ_A das normierte Polynom minimalen Grades mit dieser Eigenschaft $\widetilde{p}(A) = 0$. Offenbar ist das gerade $\mu_A = (\lambda - b_1) \cdots (\lambda - b_s)$. □

Schließlich können mit Hilfe des Minimalpolynomials ein weiteres Kriterium für die Diagonalisierbarkeit einer Matrix bzw. eines Endomorphismus angeben. Dieses verwendet im Unterschied zu Satz 24.28 das Minimalpolynom:

Satz 28.12 (notwendiges und hinreichendes Kriterium für Diagonalisierbarkeit)**.**
Es sei K ein Körper.

(*i*) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Dann sind äquivalent:

(a) A ist diagonalisierbar.

(b) Das Minimalpolynom μ_A zerfällt in Linearfaktoren und besitzt nur einfache Nullstellen.

(*ii*) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Dann sind äquivalent:

(a) f ist diagonalisierbar.

[9] Die Menge $\{b_1, \ldots, b_s\}$ besteht also aus den verschiedenen Elementen von $\{a_1, \ldots, a_n\}$.

(b) Das Minimalpolynom μ_f zerfällt in Linearfaktoren und besitzt nur einfache Nullstellen.

Beweis. Wir beweisen nur Aussage (ii). (**Quizfrage 28.4:** Wie kann man von Aussage (ii) auf Aussage (i) schließen?)

Aussage (a) $\Rightarrow$ Aussage (b): Es sei f diagonalisierbar. Das heißt, es gibt eine Basis B_V von V, sodass die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f)$ eine Diagonalmatrix ist, siehe Satz 20.24 oder die Einleitung von § 24. Auf der Diagonale von A stehen die paarweise verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_s$ entsprechend ihrer geometrischen Vielfachheit. Lemma 28.11 zeigt, dass für das Minimalpolynom gilt:

$$\mu_f = \mu_A = (\lambda - \lambda_1) \cdots (\lambda - \lambda_s).$$

Dessen Nullstellen $\lambda_1, \ldots, \lambda_s$ sind alle einfach.

Aussage (b) $\Rightarrow$ Aussage (a): Es gelte $\mu_f = \prod_{j=1}^s (\lambda - \lambda_j)$, wobei $\lambda_1, \ldots, \lambda_s \in K$ paarweise verschieden sind. Wir führen einen Induktionsbeweis nach $n = \dim(V) \in \mathbb{N}_0$. Im Fall $n = 0$ und $n = 1$ ist jeder Endormorphismus von V diagonalisierbar.

Die Behauptung sei bewiesen für $\dim(V) = n \in \mathbb{N}_0$. Bevor wir den Induktionsschritt durchführen, folgen mehrere Überlegungen:

Schritt 1: Wir zerlegen

$$\mu_f = (\lambda - \lambda_1) \cdots (\lambda - \lambda_s) =: (\lambda - \lambda_1)\, p.$$

Dabei ist $p = (\lambda - \lambda_2) \cdots (\lambda - \lambda_s) \in K_{s-1}[t]$ ein Polynom, das ebenfalls in Linearfaktoren zerfällt und lauter einfache Nullstellen hat. Außerdem gilt $\widetilde{p}(\lambda_1) \neq 0$.

Nach Satz 11.14 existieren (eindeutig bestimmte) Polynome $q, r \in K[t]$ mit

$$p = q \cdot (\lambda - \lambda_1) + r, \tag{28.7}$$

wobei $\deg(r) < \deg(\lambda - \lambda_1) = 1$ ist. Also ist $r = \alpha \in K$ ein konstantes Polynom. Wegen

$$\begin{aligned}\widetilde{r}(\lambda_1) &= \widetilde{p}(\lambda_1) - \widetilde{q}(\lambda_1) \cdot (\lambda_1 - \lambda_1) \\ &= \widetilde{p}(\lambda_1) \\ &= (\lambda - \lambda_2) \cdots (\lambda - \lambda_s) \\ &\neq 0\end{aligned}$$

gilt $\alpha = r = \widetilde{r}(\lambda_1) \neq 0$.

Schritt 2: Wir zeigen: $V = U \oplus W$ mit $U := \operatorname{Kern}(f - \lambda_1 \operatorname{id}_V) = \operatorname{Eig}(f, \lambda_1)$ und $W := \operatorname{Bild}(f - \lambda_1 \operatorname{id}_V)$.

Dazu setzen wir f in (28.7) ein:

$$\widetilde{p}(f) = \widetilde{q}(f) \circ (f - \lambda_1 \operatorname{id}_V) + \alpha \operatorname{id}_V.$$

Folglich gilt für alle $v \in V$:

$$\big(\widetilde{p}(f)\big)(v) = \big(\widetilde{q}(f) \circ (f - \lambda_1 \operatorname{id}_V)\big)(v) + \alpha\, v$$

und wegen $\alpha \neq 0$

$$\begin{aligned}v &= \alpha^{-1}\big(\widetilde{p}(f)\big)(v) - \alpha^{-1}\big(\widetilde{q}(f) \circ (f - \lambda_1 \operatorname{id}_V)\big)(v) \\ &= \underbrace{\alpha^{-1}\big((f - \lambda_2 \operatorname{id}_V) \circ \cdots \circ (f - \lambda_s \operatorname{id}_V)\big)(v)}_{=:u} - \underbrace{\alpha^{-1}\big(\widetilde{q}(f) \circ (f - \lambda_1 \operatorname{id}_V)\big)(v)}_{=:w}.\end{aligned}$$

Für u gilt

$$\begin{aligned}(f - \lambda_1 \operatorname{id}_V)(u) &= \alpha^{-1}\big((f - \lambda_1 \operatorname{id}_V) \circ (f - \lambda_2 \operatorname{id}_V) \circ \dots \circ (f - \lambda_s \operatorname{id}_V)\big)(u) \\ &= \alpha^{-1}(\mu_f(f))(v) \\ &= 0 \quad \text{wegen } \mu_f(f) = 0.\end{aligned}$$

Also liegt $u \in U = \operatorname{Kern}(f - \lambda_1 \operatorname{id}_V)$.

Weiter gilt

$$\begin{aligned}w &= \alpha^{-1}\big(\widetilde{q}(f) \circ (f - \lambda_1 \operatorname{id}_V)\big)(v) \\ &= \alpha^{-1}\big((f - \lambda_1 \operatorname{id}_V) \circ \widetilde{q}(f)\big)(v)\end{aligned}$$

(**Quizfrage 28.5:** Warum gilt die Gleichheit?) und somit $w \in W = \operatorname{Bild}(f - \lambda_1 \operatorname{id}_V)$.

Wir haben also gezeigt, dass jedes $v \in V$ in der Form $v = u + w$ mit $u \in U$ und $w \in W$ dargestellt werden kann, also gilt $V = U + W$.

Weiter gilt nach der Dimensionsformel für lineare Abbildungen (Folgerung 18.7)

$$\dim(U) + \dim(W) = \dim(\operatorname{Kern}(f - \lambda_1 \operatorname{id}_V)) + \dim(\operatorname{Bild}(f - \lambda_1 \operatorname{id}_V)) = \dim(V)$$

und weiter nach (14.2)

$$\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W) = \dim(V) - \dim(V) = 0.$$

Es gilt also tatsächlich $V = U \oplus W$.

Schritt 3: Wir zeigen: $W = \operatorname{Bild}(f - \lambda_1 \operatorname{id}_V)$ ist f-invarianter Unterraum mit $0 \leqslant \dim(W) < \dim(V)$.

Es sei dazu $w \in W$, d. h., $w = (f - \lambda_1 \operatorname{id}_V)(v)$ für ein $v \in V$. Wir haben

$$\begin{aligned}f(w) &= \big(f \circ (f - \lambda_1 \operatorname{id}_V)\big)(v) \\ &= \big((f - \lambda_1 \operatorname{id}_V) \circ f\big)(v) \quad \text{wegen } f \circ (f - \lambda_1 \operatorname{id}_V) = (f - \lambda_1 \operatorname{id}_V) \circ f \\ &\in \operatorname{Bild}(f - \lambda_1 \operatorname{id}_V) = W.\end{aligned}$$

Wegen $\dim(U) = \dim(\operatorname{Eig}(f, \lambda_1)) \geqslant 1$ gilt tatsächlich

$$\dim(W) = \dim(V) - \dim(U) \leqslant \dim(V) - 1.$$

Wir können nun den Induktionsschritt von $\dim(V) = n$ auf $n + 1$ durchführen. Das Minimalpolynom $\mu_f = \prod_{j=1}^{s} (\lambda - \lambda_j)$ habe die Koeffizientendarstellung $\mu_f = \sum_{j=1}^{s} \alpha_j \lambda^j$.

Wir setzen $g := f|_W^W \in \operatorname{End}(W)$. Für $w \in W$ gilt dann

$$\begin{aligned}\big(\widetilde{\mu_f}(g)\big)(w) &= \Big(\sum_{j=1}^{s} \alpha_j g^j\Big)(w) \\ &= \sum_{j=1}^{s} \alpha_j g^j(w) \\ &= \sum_{j=1}^{s} \alpha_j f^j(w) \\ &= \Big(\sum_{j=1}^{s} \alpha_j f^j\Big)(w) \\ &= \big(\widetilde{\mu_f}(f)\big)(w) \\ &= 0 \quad \text{wegen } \mu_f(f) = 0.\end{aligned}$$

Da $w \in W$ beliebig war, bedeutet das: $\widetilde{\mu_f}(g) = 0 \in \operatorname{End}(W)$. Damit gehört μ_f zum Ideal $J_g \subseteq K[t]$. Nach Lemma 28.6 gilt also

$$\mu_g \mid \mu_f = (\lambda - \lambda_1) \cdots (\lambda - \lambda_s)$$

Nach Voraussetzung sind die $\lambda_1, \ldots, \lambda_s$ paarweise verschieden. Das heißt aber, auch μ_g zerfällt in Linearfaktoren und hat nur einfache Nullstellen.

Aus der Induktionsvoraussetzung können wir schließen, dass $g \in \operatorname{End}(W)$ diagonalisierbar ist, denn es gilt $\dim(W) \leqslant \dim(V) - 1 \leqslant n$. Es existiert also eine Basis von W, die aus lauter Eigenvektoren zu den Eigenwerten von $g = f|_W^W$ besteht. Diese Eigenpaare sind auch Eigenpaare von f. (**Quizfrage 28.6:** Warum?) Ergänzen wir diese Basis von W aus Eigenvektoren von f durch eine Basis von $U = \operatorname{Eig}(f, \lambda_1)$, dann erhalten wir nach Satz 14.8 eine Basis von $V = U \oplus W$, die wiederum aus lauter Eigenvektoren von f besteht. Also ist f diagonalisierbar. □

Beispiel 28.13 (notwendiges und hinreichendes Kriterium für Diagonalisierbarkeit, vgl. Beispiel 28.8)**.**

(*i*) Die Matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{Q}^{2 \times 2}$$

besitzt das Minimalpolynom $\mu_A = (\lambda - 1)\,(\lambda + 1)$, ist also diagonalisierbar.

(*ii*) Die Matrix

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -8 & 1 & 4 \\ 2 & -1 & -1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$

besitzt das Minimalpolynom $\mu_A = (\lambda + 1)^2\,(\lambda - 3)$, ist also nicht diagonalisierbar.

(*iii*) Die Matrix

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -6 & 1 & 2 \\ 3 & -1 & -2 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$

besitzt das Minimalpolynom $\mu_A = (\lambda + 1)\,(\lambda - 3)$, ist also diagonalisierbar. △

§ 28.3 Die Begleitmatrix eines Polynoms

Definition 28.14 (Begleitmatrix eines Polynoms)**.**
Es sei K ein Körper und

$$p = \lambda^n + \sum_{j=0}^{n-1} \alpha_j \, \lambda^j \in K[\lambda] \tag{28.8a}$$

ein normiertes Polynom vom Grad $n \in \mathbb{N}_0$ (bzw. das Einspolynom im Fall $n = 0$). Die Matrix

$$C_p \coloneqq \begin{bmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & \vdots & -\alpha_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -\alpha_{n-1} \end{bmatrix} \in K^{n \times n} \tag{28.8b}$$

heißt die **Begleitmatrix** (englisch: companion matrix) von p. △

Beachte: Die Einsen laufen direkt unterhalb der Hauptdiagonale. Der führende Koeffizient von p ist Eins und wird in der Begleitmatrix nicht notiert. Im Fall $n = 0$ ist $p = 1$ und C_p die leere Matrix. Im Fall $n = 1$ mit $p = \alpha_0$ gilt $C_p = -\alpha_0$.

Lemma 28.15 (charakteristisches und Minimalpolynom von Begleitmatrizen)**.**
Es sei K ein Körper und $p \in K[\lambda]$ ein normiertes Polynom. Dann gilt:

$$\chi_{C_p} = \mu_{C_p} = p. \tag{28.9}$$

Beachte: In diesem Sinne ist die Begleitmatrix C_p ein besonders einfacher Repräsentant der Äquivalenzklasse aller zueinander ähnlichen Matrizen, deren charakteristisches und Minimalpolynom dem vorgegebenen Polynom p entsprechen.

Beweis. Die Behauptung $\chi_{C_p} = p$ ist Gegenstand von Hausaufgabe II-7.1. In der dazugehörigen Lösung wurde die Möglichkeit durch Zeilentransformationen zu einer einfachen Matrixstruktur und einer Darstellung von p nach dem Horner-Schema zu kommen vorgestellt. Alternativ lässt sich der Nachweis per Induktion nach dem Grad $\deg(p) = n \in \mathbb{N}_0$ zeigen: Für $n = 0$ ist $p = 1$ und $C_p = ()$ mit $\det(\lambda I_0 - ()) = 1$. Wir betrachten nun ein normiertes Polynom $p = \lambda^{n+1} + \sum_{j=0}^n \alpha_j \lambda^j$ vom Grad $\deg(p) = n+1$, wobei die Behauptung für den Grad n bereits bewiesen sei. Wir haben

$$C_p = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & \vdots & -\alpha_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -\alpha_n \end{bmatrix} \text{ und entwickeln } \det(\lambda I - C_p) = \det \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 & \alpha_0 \\ -1 & \lambda & 0 & \cdots & 0 & \alpha_1 \\ 0 & -1 & \ddots & \ddots & \vdots & \alpha_2 \\ \vdots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & -1 & \lambda & \alpha_{n-1} \\ 0 & \cdots & \cdots & 0 & -1 & \lambda + \alpha_n \end{pmatrix}$$

nach der ersten Zeile mit Hilfe des Laplaceschen Entwicklungssatzes 23.17. Wir erhalten

$$\det(\lambda I - C_p) = \lambda \det \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 & \alpha_1 \\ -1 & \lambda & 0 & \cdots & 0 & \alpha_2 \\ 0 & -1 & \ddots & \ddots & \vdots & \alpha_3 \\ \vdots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & -1 & \lambda & \alpha_{n-1} \\ 0 & \cdots & \cdots & 0 & -1 & \lambda + \alpha_n \end{pmatrix} + (-1)^{n+2} \alpha_0 \det \begin{pmatrix} -1 & \lambda & 0 & \cdots & 0 \\ 0 & -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & \lambda \\ 0 & \cdots & \cdots & 0 & -1 \end{pmatrix}.$$

Die erste Matrix ist gerade $\lambda I - C_q$ mit der Begleitmatrix C_q des Polynoms $q = \alpha_1 + \alpha_2 \lambda + \cdots + \alpha_n \lambda^{n-1} + \lambda^n$ vom Grad $\deg(q) = n$. Nach Induktionsvoraussetzung ist also $\det(\lambda I - C_q) = q$. Die zweite Matrix ist eine obere Dreiecksmatrix, und nach Lemma 23.10 ist ihre Determinante gleich $(-1)^n$. Wir erhalten also

$$\begin{aligned} \det(\lambda I - C_p) &= \lambda q + \alpha_0 \\ &= \lambda (\alpha_1 + \alpha_2 \lambda + \cdots + \alpha_n \lambda^{n-1} + \lambda^n) + \alpha_0 \\ &= p, \end{aligned}$$

also $\chi_{C_p} = p$, was zu zeigen war.

Der Beweis der Behauptung $\mu_{C_p} = p$ ist Gegenstand von Hausaufgabe II-8.3. □

§ 29 Die Frobenius-Normalform

Literatur: Liesen, Mehrmann, 2024, Kapitel 16

Normalformen von Endomorphismen endlich-dimensionaler Vektorräume sind dadurch gekennzeichnet, dass die Darstellungsmatrizen Blockdiagonalstruktur

$$\begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{bmatrix}$$

mit „möglichst einfachen“ Blöcken besitzt. Jeder Block entspricht nach Satz 20.17 einem f-invarianten Unterraum.

Für die Wahl dieser Blöcke bzw. invarianten Unterräume gibt es verschiedene Möglichkeiten und daher auch verschiedene Normalformen, darunter die **Frobenius-Normalform**, die **Weierstraß-Normalform** und die **Jordan-Normalform**. Bei der Frobenius-Normalform, die wir in diesem Abschnitt betrachten, sind die Blöcke gerade die Begleitmatrizen bestimmter Polynome.

In Definition 28.2 hatten wir das Minimalpolynom einer Matrix $A \in K^{n \times n}$ als das eindeutige normierte Polynom $p \in K[t]$ geringsten Grades definiert, das die Eigenschaft $\widetilde{p}(A) = 0$ besitzt. Nun betrachten wir eine verfeinerte Version davon, indem wir nur fordern, dass $\widetilde{p}(A)\, x = 0$ für einen gegebenen Vektor $x \in K^n$ gilt. Man sagt auch, das Polynom **annulliere** die Matrix A **lokal**. „Lokal“ bedeutet hier: bei Anwendung auf einen gegebenen Vektor.

Lemma 29.1 (die lokal annullierenden Polynome bilden ein Ideal)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Für $x \in K^n$ ist die Menge

$$J_{A,x} \coloneqq \big\{ p \in K[\lambda] \,\big|\, \widetilde{p}(A)\, x = 0 \big\} \tag{29.1}$$

ein Ideal in $K[\lambda]$ ungleich dem Nullideal.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Für $v \in V$ ist die Menge

$$J_{f,v} \coloneqq \big\{ p \in K[\lambda] \,\big|\, \widetilde{p}(f)(v) = 0 \big\} \tag{29.2}$$

ein Ideal in $K[\lambda]$ ungleich dem Nullideal.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-9.1. □

Ganz analog zu Definition 28.2 haben wir mit Hilfe von Satz 27.7:

Definition 29.2 (lokales Minimalpolynom, vgl. Definition 28.2)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$ und $x \in K^n$. Das eindeutig bestimmte normierte Polynom $\mu_{A,x} \neq 0$ geringsten Grades mit der Eigenschaft $\mu_{A,x} \in J_{A,x}$, also $\widetilde{\mu_{A,x}}(A)\, x = 0$, heißt das **(lokale) Minimalpolynom** (englisch: local minimal polynomial) **von A bzgl. x**.

(*ii*) Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K. Weiter sei $f \in \operatorname{End}(V)$ und $v \in V$. Das eindeutig bestimmte normierte Polynom $\mu_{f,v} \neq 0$ geringsten Grades mit der Eigenschaft $\mu_{f,v} \in J_{f,v}$, also $\widetilde{\mu_{f,v}}(f)(v) = 0$, heißt das **(lokale) Minimalpolynom von f bzgl. v**. △

Satz 29.3 (lokale Minimalpolynome eines Endomorphismus und seiner Darstellungsmatrizen, vgl. Satz 28.3)**.**
Es sei K ein Körper und V ein endlich-dimensionaler Vektorraum über K mit der Basis $B_V = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$. Weiter sei $f \in \operatorname{End}(V)$ und $A = \mathcal{M}_{B_V}^{B_V}(f)$. Dann gilt für $v \in V$ und seinen Koordinatenvektor $x = \Phi_{B_V}^{-1}(v)$:

$$\mu_{f,v} = \mu_{A,x}. \tag{29.3}$$

Beweis. □

Zur Bestimmung des lokalen Minimalpolynoms einer Matrix A bzgl. eines Vektors x können wir ähnlich wie in Beispiel 28.5 vorgehen. Statt die Matrizen $A^0, A^1, \ldots$ etc. auf das erstmalige Auftreten linearer Abhängigkeit zu prüfen, testen wir nun die Vektoren $A^0\,x, A^1\,x, \ldots$

Beispiel 29.4 (Bestimmung des lokalen Minimalpolynoms, vgl. Beispiel 28.5)**.**

(*i*) Wir bestimmen das lokale Minimalpolynom der Matrix

$$A = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \in \mathbb{R}^{3\times 3} \text{ bzgl. des Vektors } x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Es gilt

$$A\,x = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad A^2\,x = \begin{bmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{bmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ -3 \end{pmatrix}.$$

Wir betrachten jetzt die Matrix

$$\begin{bmatrix} A^0\,x & A^1\,x & A^2\,x \end{bmatrix} = \begin{bmatrix} 1 & -2 & 6 \\ 0 & 1 & -3 \\ 0 & 1 & -3 \end{bmatrix}$$

Wie in Beispiel 28.5 sind auch hier die Spalten $\{0, 1\}$ zusammen noch linear unabhängig, die Spalten $\{0, 1, 2\}$ aber nicht mehr. Es gilt daher $\deg(\mu_{A,x}) = 2$, und der Kern von $\begin{bmatrix} A^0\,x & A^1\,x & A^2\,x \end{bmatrix}$ besteht aus den Vielfachen von $\left(\begin{smallmatrix} 0 \\ 3 \\ 1 \end{smallmatrix}\right)$. Das lokale Minimalpolynom von A bzgl. x ist also

$$\mu_{A,x} = \lambda^2 + 3\,\lambda = \mu_A.$$

(*ii*) Verwenden wir dieselbe Matrix A, aber den Vektor $y = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, so erhalten wir

$$A\,y = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{und erst recht} \quad A^2\,y = \begin{bmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{bmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Nun gilt daher $\deg(\mu_{A,y}) = 1$, und der Kern von $\begin{bmatrix} A^0\, y & A^1\, y \end{bmatrix}$ besteht aus den Vielfachen von $\binom{0}{1}$. Das lokale Minimalpolynom von A bzgl. y ist also

$$\mu_{A,y} = \lambda. \qquad \triangle$$

Analog zu Lemma 28.6 gilt:

Lemma 29.5 (das lokale Minimalpolynom teilt jedes Polynom, das seine Matrix lokal annulliert)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$ und $x \in K^n$. Dann teilt das Minimalpolynom $\mu_{A,x}$ bzgl. x jedes Polynom in $p \in J_{A,x}$, also jedes $p \in K[t]$ mit der Eigenschaft $\widetilde{p}(A)\, x = 0$.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K, $f \in \operatorname{End}(V)$ und $v \in V$. Dann teilt das Minimalpolynom $\mu_{f,v}$ bzgl. v jedes Polynom in $p \in J_{f,v}$, also jedes $p \in K[t]$ mit der Eigenschaft $\widetilde{p}(f)(v) = 0$.

Insbesondere gilt $\mu_{A,x} \mid \mu_A$ für jedes $x \in K^n$ und $\mu_{f,v} \mid \mu_f$ für jedes $v \in V$.

Das folgende Resultat zeigt, dass jedes lokale Minimalpolynom von A einen A-invarianten Unterraum von K^n liefert:

Lemma 29.6 (A-invariante Unterräume durch lokale Minimalpolynome)**.**
Es sei K ein Körper und $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$ eine Matrix. Weiter sei $x \in K^n$ und

$$\mu_{A,x} = \sum_{j=0}^{d-1} \alpha_j\, \lambda^j + \lambda^d$$

das lokale Minimalpolynom bzgl. x. Dann ist der Unterraum

$$U_x \coloneqq \langle x,\ A\,x,\ A^2\,x, \dots, A^{d-1}\,x \rangle \tag{29.4a}$$

ein A-invarianter Unterraum mit der Basis

$$B_x = (x_1, \dots, x_d) \text{ mit } x_j \coloneqq A^{j-1}\,x \text{ für } j = 1, \dots, d. \tag{29.4b}$$

Beweis. Die Familie der Vektoren $B_x \coloneqq (x,\ A\,x,\ A^2\,x, \dots, A^{d-1}\,x)$ ist linear unabhängig, jedoch $(x,\ A\,x,\ A^2\,x, \dots, A^{d-1}\,x,\ A^d\,x)$ ist linear abhängig, vgl. Beispiel 29.4, denn es gilt

$$\widetilde{\mu_{A,x}}(A)\, x = \sum_{j=0}^{d-1} \alpha_j\, A^j\, x + A^d\, x = 0.$$

Um zu zeigen, dass der Unterraum $U_x \coloneqq \langle x,\ A\,x,\ A^2\,x, \dots, A^{d-1}\,x \rangle$ A-invariant ist, betrachten wir eine Linearkombination $u = \sum_{j=0}^{d-1} \beta_j\, A^j\, x \in U$. Es folgt

$$A\,u = \sum_{j=0}^{d-1} \beta_j\, A^{j+1}\, x = \sum_{j=1}^{d} \beta_{j-1}\, A^j\, x = \beta_{d-1}\, A^d\, x + \underbrace{\sum_{j=1}^{d-1} \beta_{j-1}\, A^j\, x}_{\in U},$$

und wir können das lokale Minimalpolynom verwenden, um $A^d\, y = -\sum_{j=0}^{d-1} \alpha_j\, A^j\, y \in U$ zu ersetzen.

Das zeigt $A\,u \in U$. □

Ende der Vorlesung 17

Wir zeigen nun, dass diejenigen Vektoren, deren lokale Minimalpolynom $\mu_{A,x}$ maximal möglichen Grad haben, besondere Bedeutung besitzen.[10] Mit ihrer Hilfe können wir nämlich den Vektorraum K^n in zwei komplementäre, A-invariante Unterräume zerlegen:

Lemma 29.7 (Zerlegung in komplementäre, A-invariante Unterräume)**.**

Es sei K ein Körper und $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$ eine Matrix. Weiter sei $x \in K^n$ so gewählt, dass das lokale Minimalpolynom bzgl. x

$$\mu_{A,x} = \sum_{j=0}^{d-1} \alpha_j\, \lambda^j + \lambda^d$$

den maximal möglichen Grad über alle lokalen Minimalpolynome von A besitzt[11], also

$$d = \max\{\deg(\mu_{A,y}) \mid y \in K^n\}.$$

Dann gilt: Ergänzen wir die Basis B_x des Unterraumes U_x aus (29.4) zu irgendeiner Basis

$$B = (x_1, \dots, x_d, x_{d+1}, \dots, x_n)$$

von K^n und definieren die Linearform ξ durch die Bilder auf der Basis B wie folgt: **Im Folgenden wurde $\xi^\intercal\cdot$ korrigiert zur dualen Paarung $\langle \xi\,, \cdot \rangle$.**

$$\langle \xi\,, x_j \rangle = \delta_{jd} = \begin{cases} 1 & \text{im Fall } j = d, \\ 0 & \text{sonst,} \end{cases}$$

dann ist der Unterraum

$$W := \left\{ w \in K^n \,\middle|\, \langle \xi\,, A^j\, w \rangle = 0 \text{ für alle } j = 0, \dots, d-1 \right\} \tag{29.5a}$$

A-invariant, besitzt Dimension $\dim(W) = n - d$ und erfüllt

$$K^n = U_x{}^{\mathrm{GM}} \oplus W. \tag{29.5b}$$

Beachte: Der Unterraum W ist gerade der Prä-Annihilator von $\left\{\xi,\; A^\intercal \xi,\; (A^\intercal)^2\, \xi, \dots, (A^\intercal)^{d-1}\, \xi\right\}$.

Beweis. Wir ergänzen B_x zu irgendeiner Basis $B = (x_1, \dots, x_d, x_{d+1}, \dots, x_n)$ von K^n und definieren die Linearform $\xi \in (K^n)^*$ wie angegeben.

---

[10] Ein solcher Vektor heißt ein **maximaler Vektor** (englisch: maximal vector) **bzgl. A bzw. f**.

[11] Es ist natürlich immer möglich, ein solches $x \in K^n$ zu wählen, da die Menge der möglichen Grade der lokalen Minimalpolynome nach oben durch den $\deg(\mu_A$ beschränkt ist.

Schritt 1: Wir zeigen: W ist A-invariant.

Es sei dazu $w \in W$ beliebig. Nach Definition gehört $A\,w$ genau dann zu W, wenn die Bedingungen

$$\langle \xi\,, A^j\,A\,w\rangle = \langle \xi\,, A^{j+1}\,w\rangle = 0 \quad \text{für } j = 0, \dots, d-1$$

gelten. Für die Indizes $j = 0, \dots, d-2$ ist das wegen $w \in W$ klar nach (29.5b).

Für die Aussage im Fall $j = d-1$ betrachten wir $d' \coloneqq \deg(\mu_{A,w})$. Wegen der Maximalgradeigenschaft des lokalen Minimalpolynoms zu x gilt $d' \leqslant d$. Der Vektor $A^d\,w$ liegt daher in $U_w \coloneqq \big\langle w,\ A\,w,\ A^2\,w, \dots, A^{d'-1}\,w\big\rangle$. Es ist also möglich, $A^d\,w$ als Linearkombination der Vektoren $A^j\,w$ für $j = 0, \dots, d'-1 \leqslant d-1$ darzustellen. Daher gilt auch $\langle \xi\,, A^d\,w\rangle = 0$ und damit $A\,w \in W$.

Schritt 2: Wir zeigen: $\dim(W) \geqslant n-d$.

Per Definition (29.5b) gilt mit $Y \coloneqq \big\langle \xi,\ A^\intercal\xi,\ (A^\intercal)^2\,\xi, \dots, (A^\intercal)^{d-1}\,\xi\big\rangle$ die Beziehung

$$W = {}^0Y = \big\{w \in K^n \,\big|\, \langle \xi\,, w\rangle = 0 \text{ für alle } \xi \in Y\big\}.$$

Da $\dim(Y) \leqslant d$ gilt, haben wir nach Satz 21.23 $\dim(W) = n - \dim(Y) \geqslant n-d$.

Schritt 3: Wir zeigen: $U_x \cap W = \{0\}$.

Es sei dazu $w \in U_x \cap W$, also $w = \sum_{i=1}^d \beta_i\,x_i$ mit Koeffizienten $\beta_i \in K$. Wegen $w \in W$ gilt weiter wegen (29.5b) $\langle \xi\,, A^j\,w\rangle = 0$ für $j = 0, \dots, d-1$. Für $j = 0$ heißt das zunächst

$$0 = \Big\langle \xi\,, \sum_{i=1}^d \beta_i\,x_i\Big\rangle = \sum_{i=1}^d \beta_i\,\langle \xi\,, x_i\rangle = \beta_d,$$

wobei die letzte Gleichheit aus der Definition von ξ folgt. Nun ist also $w = \sum_{i=1}^{d-1} \beta_i\,x_i$, und wir werten die Bedingung (29.5b) für $j = 1$ aus:

$$0 = \Big\langle \xi\,, \sum_{i=1}^{d-1} \beta_i\,A\,x_i\Big\rangle = \sum_{i=1}^{d-1} \beta_i\,\langle \xi\,, A\,x_i\rangle = \sum_{i=1}^{d-1} \beta_i\,\langle \xi\,, x\rangle_{i+1} = \beta_{d-1}.$$

So verfahren wir weiter, bis sich schließlich für alle Koeffizienten $\beta_d = \beta_{d-1} = \cdots = \beta_1 = 0$ ergeben hat, also $x = 0$.

Schritt 4: Wir zeigen: $\dim(W) = n-d$ und $K^n = U \oplus W$.

Nach Satz 14.3 und — wie gerade gezeigt — $\dim(U \cap W) = 0$ gilt

$$\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W) = \dim(U) + \dim(W) \geqslant d + (n-d) = n.$$

Die Dimension kann aber nicht größer sein als n, also haben wir in der Tat $\dim(U+W) = n$ und damit $U + W = K^n$. Außerdem folgt $\dim(W) = n-d$, und wegen $U \cap W = \{0\}$ ist die Summe direkt. □

Noch ist nicht klar, ob lokale Minimalpolynome maximalen Grades mit dem Minimalpolynom übereinstimmen. In Beispiel 29.4 hatten wir zumindest ein Beispiel gesehen, bei dem ein lokales Minimalpolynom $\mu_{A,x}$ mit dem globalen Minimalpolynom μ_A übereinstimmte. Tatsächlich ist das immer möglich:

Lemma 29.8 (unter den lokalen Minimalpolynomen ist das Minimalpolynom).
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Für jedes $x \in K^n$, für das $\mu_{A,x}$ maximalen Grad besitzt, gilt $\mu_{A,x} = \mu_A$.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K, $f \in \operatorname{End}(V)$. Für jedes $v \in V$, für das $\mu_{f,v}$ maximalen Grad besitzt, gilt $\mu_{f,v} = \mu_f$.

Beweis. Aussage (i): Es sei $x \in K^n$ ein maximaler Vektor bzgl. A. Wir zeigen: $\widetilde{\mu_{A,x}}(A)\, y = 0$ für alle $y \in K^n$.

Es sei dazu $y \in K^n$ beliebig. Nach (29.5) können wir $y = u + w$ schreiben für (eindeutig bestimmte) $u \in U_x$ und $w \in W$.

Die Eigenschaft $\widetilde{\mu_{A,x}}(A)\, u = 0$ ist klar. Um $\widetilde{\mu_{A,x}}(A)\, w = 0$ zu zeigen, betrachten wir $p \coloneqq \mu_{A,x+w}$. Dann gilt

$$0 = \widetilde{p}(A)\,(x + w) = \widetilde{p}(A)\, x + \widetilde{p}(A)\, w,$$

also $\widetilde{p}(A)\, w = -\widetilde{p}(A)\, x$. Nach Aussage von Lemma 29.7 ist W A-invariant, also gilt auch $\widetilde{p}(A)\, w \in W$. Damit liegt $\widetilde{p}(A)\, w = -\widetilde{p}(A)\, x$ sowohl in W als auch in U_x, also muss $\widetilde{p}(A)\, w = 0$ und auch $\widetilde{p}(A)\, x = 0$ gelten. Daraus folgt $\mu_{A,x} \mid p$.

Da x ein maximaler Vektor war, gilt aber auch

$$\deg(\mu_{A,x}) \geqslant \deg(\mu_{A,x+w}) = \deg(p)$$

und somit $p = \mu_{A,x}$. Damit folgt $\widetilde{\mu_{A,x}}(A)\, w = \widetilde{p}(A)\, w = 0$.

Wir haben also

$$\widetilde{\mu_{A,x}}(A)\, y = \widetilde{\mu_{A,x}}(A)\, u + \widetilde{\mu_{A,x}}(A)\, w = 0 + 0 = 0$$

gezeigt, wobei $y \in K^n$ beliebig war. Das heißt aber $\widetilde{\mu_{A,x}}(A) = 0$ und damit $\mu_A \mid \mu_{A,x}$. Weil aber auch $\mu_{A,x} \mid \mu_A$ gilt, schließen wir wie behauptet $\mu_{A,x} = \mu_A$.

Aussage (ii) kann analog gezeigt werden. □

Wir können nun die Existenz der Frobenius-Normalform einer Matrix bzw. eines Endomorphismus beweisen:

Satz 29.9 (Frobenius-Normalform).
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Dann existieren normierte, für $n \geqslant 1$ nicht-konstante Polynome $p_1, \ldots, p_k$, $k \in \mathbb{N}$, die folgende Eigenschaften haben:

(a) $p_1 = \mu_A$ und $p_{j+1} \mid p_j$ für $j = 1, \ldots, k-1$.

(b) A ist ähnlich zu der Blockdiagonalmatrix

$$\begin{bmatrix} C_{p_1} & & & \\ & C_{p_2} & & \\ & & \ddots & \\ & & & C_{p_k} \end{bmatrix} \tag{29.6}$$

mit den Begleitmatrizen C_{p_j}, $j = 1, \ldots, k$.

(c) Sowohl die Polynome p_j also auch die Matrix in (29.6) sind eindeutig.

(ii) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Dann existieren normierte, nicht-konstante Polynome $p_1, \dots, p_k$, $k \in \mathbb{N}$, die folgende Eigenschaften haben:

(a) $p_1 = \mu_f$ und $p_{j+1} \mid p_j$ für $j = 1, \dots, k-1$.

(b) Es gibt eine Basis B_V von V, sodass die Darstellungsmatrix von f die Blockdiagonalgestalt

$$\mathcal{M}_{B_V}^{B_V}(f) = \begin{bmatrix} C_{p_1} & & & \\ & C_{p_2} & & \\ & & \ddots & \\ & & & C_{p_k} \end{bmatrix} \tag{29.7}$$

mit den Begleitmatrizen C_{p_j}, $j = 1, \dots, k$ besitzt.

(c) Sowohl die Polynome p_j also auch die Matrix in (29.7) sind eindeutig.

Die durch A bzw. durch f eindeutig bestimmte Matrix in (29.6) bzw. (29.7) heißt die **Frobenius-Normalform** (englisch: Frobenius normal form) oder die **rationale Normalform** (englisch: rational normal form) von A bzw. von f.

Beweis. Aussage (i): Im Fall $n = 0$ gilt $\mu_A = 1$ und $k = 1$, d. h., und es ist nichts zu zeigen.

Schritt 1: Wir spalten einen A-invarianten Unterraum U_1 der Bauart $\langle x_1,\ A\,x_1,\ A^2\,x_1, \dots, A^{m-1}\,x_1 \rangle$ maximaler Dimension ab.

Es sei dazu x_1 ein maximaler Vektor bzgl. A. Es gilt also $\mu_{A,x_1} = \mu_A$ nach Lemma 29.8. Der Unterraum

$$U_1 \coloneqq \langle x_1,\ A\,x_1,\ A^2\,x_1, \dots, A^{m-1}\,x_1 \rangle$$

hat Dimension m und ist A-invariant. Aufgrund der A-Invarianz von U_1 können wir f_A auf U_1 einschränken. Für die Darstellungsmatrix bzgl. der Basis

$$B_1 \coloneqq (x_1,\ A\,x_1,\ A^2\,x_1, \dots, A^{m-1}\,x_1)$$

erhalten wir die zu A ähnliche Matrix

$$\mathcal{M}_{B_1}^{B_1}(f_A|_{U_1}^{U_1}) = C_{p_1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & \vdots & -\alpha_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -\alpha_{m-1} \end{bmatrix} \in K^{m \times m}$$

für $p_1 \coloneqq \mu_A$.

Falls $m = n$ gilt, also $U_1 = K^n$, so ist der Beweis hier erbracht.

Schritt 2: Wir konstruieren einen zu U_1 komplementären, ebenfalls A-invarianten Unterraum W_1, also $K^n = U_1 \oplus W_1$ mit $A \cdot W_1 \subseteq W_1$.

Einen solchen Unterraum erhalten wir Lemma 29.7. Es gilt $\dim(U_1) = m = \deg(\mu_A)$ und damit $\dim(W_1) = n - m \leqslant n - 1$.

Ergänzen wir die Basis B_1 von U_1 durch eine Basis $\widetilde{B}_1$ von W_1 zur Basis B, dann besitzt die zu A ähnliche Darstellungsmatrix bzgl. der Basis B die Blockdiagonalstruktur

$$\mathcal{M}_B^B(f_A) = \begin{bmatrix} C_{p_1} & \\ & \mathcal{M}_{\widetilde{B}_1}^{\widetilde{B}_1}(f_A|_{W_1}^{W_1}) \end{bmatrix}.$$

Für das Minimalpolynom gilt $\mu_A = \mu_{\mathcal{M}_B^B(f_A)}$ (Folgerung 28.4), also auch

$$0 = \widetilde{\mu_A}\big(\mathcal{M}_B^B(f_A)\big) = \begin{bmatrix} \widetilde{\mu_A}(C_{p_1}) & \\ & \widetilde{\mu_A}\big(\mathcal{M}_{\widetilde{B}_1}^{\widetilde{B}_1}(f_A|_{W_1}^{W_1})\big) \end{bmatrix}.$$

Insbesondere ist also $\widetilde{\mu_A}\big(\mathcal{M}_{\widetilde{B}_1}^{\widetilde{B}_1}(f_A|_{W_1}^{W_1})\big) = 0$. Das heißt, das Minimalpolynom p_2 des eingeschränkten Endomorphismus $f_A|_{W_1}^{W_1}$ ist ein Teiler des Minimalpolynoms $p_1 = \mu_A$ von A.

Schritt 3: Wir wiederholen die Konstruktion:

Wir können nun die Konstruktion von **Schritt 1** an für den eingeschränkten Endomorphismus $f_A|_{W_1}^{W_1}$ wiederholen. Dadurch erhalten induktiv wir weitere Polynome $p_2, \dots, p_k$ (die Minimalpolynome der jeweiligen „Restabbildungen") mit der behaupteten Eigenschaft $p_{j+1} \mid p_j$ sowie die Darstellung (29.6). Das Verfahren stoppt, sobald $V = U_1 \oplus \cdots \oplus U_k$ und somit $W_k = \{0\}$ ist.

Noch zu zeigen ist die Eindeutigkeit. Es seien dazu

$$A_1 = \begin{bmatrix} C_{p_1} & & \\ & \ddots & \\ & & C_{p_k} \end{bmatrix} \quad \text{und} \quad A_2 = \begin{bmatrix} C_{q_1} & & \\ & \ddots & \\ & & C_{q_\ell} \end{bmatrix}$$

zwei zu A ähnliche Matrizen in Frobenius-Normalform mit normierten, nicht-konstanten Polynomen $p_1, \dots, p_k$ und $q_1, \dots, q_\ell$ und $k, \ell \in \mathbb{N}$, die die Eigenschaften $p_1 = q_1 = \mu_A = \mu_{A_1} = \mu_{A_2}$ sowie $p_{j+1} \mid p_j$ und $q_{j+1} \mid q_j$ erfüllen für $j = 1, \dots, k-1$ bzw. $j = 1, \dots, \ell - 1$.

Im Fall $k = 1$ sind wegen $p_1 = q_1$ die Dimensionen der führenden Blöcke C_{p_1} und C_{q_1} beide gleich n, also ist auch $\ell = 1$ und $A_1 = A_2$.

Wir betrachten nun den Fall $k > 1$, also auch $\ell > 1$. Da die Matrizen A_1 und A_2 beide zu A ähnlich sind, sie auch zueinander ähnlich, d. h., es gibt eine reguläre Matrix $T \in K^{n \times n}$ mit $A_1 = T\,A_2\,T^{-1}$. Da Polynome beim Einsetzen von Blockdiagonalmatrizen blockweise wirken, haben wir

$$p_2(A_1) = \begin{bmatrix} p_2(C_{p_1}) & & & \\ & p_2(C_{p_2}) & & \\ & & \ddots & \\ & & & p_2(C_{p_k}) \end{bmatrix}.$$

Dabei ist $p_j(C_{p_j}) = 0$ für alle $j = 1, \ldots, k$, denn nach Lemma 28.15 ist p_j das Minimalpolynom von C_{p_j}. Nach Voraussetzung ist p_2 ein Vielfaches von $p_2, \ldots, p_k$, also sind auch sind auch $p_2(C_{p_3})$ usw. bis $p_2(C_{p_k})$ Nullmatrizen. Also haben wir

$$\begin{aligned}
\begin{bmatrix} p_2(C_{p_1}) & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} &= p_2(A_1) = p_2(T\, A_2\, T^{-1}) = T\, p_2(A_2)\, T^{-1} \\
&= T \begin{bmatrix} p_2(C_{q_1}) & & & \\ & p_2(C_{q_2}) & & \\ & & \ddots & \\ & & & p_2(C_{q_\ell}) \end{bmatrix} T^{-1} \\
&= T \begin{bmatrix} p_2(C_{p_1}) & & & \\ & p_2(C_{q_2}) & & \\ & & \ddots & \\ & & & p_2(C_{q_\ell}) \end{bmatrix} T^{-1}.
\end{aligned}$$

Insbesondere sind $p_2(A_1)$ und $p_2(A_2)$ ähnliche Matrizen und haben als solche nach Folgerung 15.41 denselben Rang. Da die Ränge der beiden Matrizen links und rechts durch die Summe der Ränge der Diagonalblöcke gegeben sind, muss insbesondere $p_2(C_{q_2}) = 0$ gelten. Also folgt $q_2 \mid p_2$. Durch Vertauschen der Rollen haben wir ebenfalls $p_2 \mid q_2$, und weil beide Polynome normiert sind, muss $p_2 = q_2$ gelten. Also haben auch die zweiten Diagonalblöcke identische Größe. Induktiv zeigt man dasselbe weiter, bis sich schließlich $k = \ell$ und $p_k = q_k$ ergibt. Damit ist die Eindeutigkeit der Frobenius-Normalform gezeigt.

Aussage (ii) kann analog gezeigt werden. □

Bemerkung 29.10 (zur Frobenius-Normalform)**.**

(i) Die in der Frobenius-Normalform auftretenden normierten, nicht-konstanten Polynome $p_1, \ldots, p_k$ heißen die **Invariantenteiler** der Matrix A bzw. des Endomorphismus f.

(ii) Jede Begleitmatrix der Dimension $m \in \mathbb{N}$ besitzt maximal $2\,m - 1$ Einträge ungleich Null. Damit hat die Frobenius-Normalform mit $k \in \mathbb{N}$ Blöcken maximal $2\,n - k$ Einträge ungleich Null. △

Beispiel 29.11 (Frobenius-Normalform)**.**

(i) Die Matrix

$$A = \begin{bmatrix} 0 & 1 & 3 \\ 3 & 1 & -4 \\ -2 & 1 & 5 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

besitzt das Minimalpolynom $\mu_A = (\lambda - 2)^3 = \lambda^3 - 6\,\lambda^2 + 12\,\lambda - 8$. A ist daher ähnlich zu seiner Frobenius-Normalform, die nur aus einem Block C_{p_1} zu $p_1 = \mu_A$ besteht:

$$\widehat{A} = \begin{bmatrix} 0 & 0 & 8 \\ 1 & 0 & -12 \\ 0 & 1 & 6 \end{bmatrix}.$$

(*ii*) Die Matrix

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ -1 & 1 & 4 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

besitzt das Minimalpolynom $\mu_A = (\lambda - 2)^2 = \lambda^2 - 4\,\lambda + 4$. Die Frobenius-Normalform beginnt mit dem 2×2-Block C_{p_1} zu $p_1 = \mu_A$. Da nur noch Dimension 1 fehlt, muss das normierte, nicht-konstante Folgepolynom p_2 vom Grad 1 sein und außerdem $p_2 \mid p_1$ erfüllen. Es muss also $p_2 = \lambda - 2$ gelten. A ist daher ähnlich zu seiner Frobenius-Normalform, die aus zwei Blöcken C_{p_1} und C_{p_2} besteht:

$$\widehat{A} = \left[\begin{array}{cc|c} 0 & -4 & \\ 1 & 4 & \\ \hline & & 2 \end{array}\right].$$

(*iii*) Die Matrix

$$A = \begin{bmatrix} 4 & -1 & -2 & 3 \\ -1 & 5 & 2 & -4 \\ 0 & 1 & 3 & -1 \\ -1 & 2 & 2 & -1 \end{bmatrix} \in \mathbb{Q}^{4\times 4}$$

besitzt das Minimalpolynom $\mu_A = (\lambda - 3)^2(\lambda - 2) = \lambda^3 - 8\,\lambda^2 + 21\,\lambda - 18$. Die Frobenius-Normalform beginnt mit dem 3×3-Block C_{p_1} zu $p_1 = \mu_A$. Da nur noch Dimension 1 fehlt, kommen als normiertes, nicht-konstantes Folgepolynom p_2 vom Grad 1 wegen $p_2 \mid p_1$ beide Möglichkeiten $p_2 = \lambda - 3$ oder $p_2 = \lambda - 2$ in Frage.

Hier ist $p_2 = \lambda - 3$ die richtige Wahl, denn ähnliche Matrizen besitzen das gleiche charakteristische Polynom, was in diesem Fall gerade $\chi_A = (\lambda - 3)\,\mu_A$ ist. Die Frobenius-Normalform von A, zu der A ja ähnlich ist, muss also gerade dieses Polynom als charakteristisches Polynom haben. Da das charakteristische Polynom der Frobeniusnormalform von A blockweise berechnet werden kann ergibt sich gerade $\lambda - 3$ als der verbleibende Faktor. A ist daher ähnlich zu seiner Frobenius-Normalform, die aus zwei Blöcken C_{p_1} und C_{p_2} besteht:

$$\widehat{A} = \left[\begin{array}{ccc|c} 0 & 0 & 18 & \\ 1 & 0 & -21 & \\ 0 & 1 & 8 & \\ \hline & & & 3 \end{array}\right].$$

(*iv*) **Das folgende Beispiel wurde nachträglich eingefügt.** Die Matrix

$$A = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

besitzt das Minimalpolynom $\mu_A = \lambda - 1$ Die weiteren Polynome sind wegen der Eigenschaft $p_2 \mid p_1$ und $p_3 \mid p_2$ gleich $p_3 = p_2 = p_1 = \lambda - 1$. A ist daher ähnlich zu seiner Frobenius-Normalform, die aus drei Blöcken der Dimension 1 besteht:

$$\widehat{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$ △

Die folgende Bemerkung wurde nachträglich hinzugefügt.

Bemerkung 29.12 (Feststellen der Ähnlichkeit zweier Matrizen)**.**

(i) Die Frobenius-Normalform ist ein ausgezeicheter Repräsentant der Äquivalenzklasse ähnlicher Matrizen in $K^{n\times n}$.

Da die Frobenius-Normalform eindeutig ist, gilt: Zwei Matrizen sind genau dann ähnlich, wenn sie dieselbe Frobenius-Normalform haben. Das kann man zum Test auf Ähnlichkeit verwenden.

(ii) Die Äquivalenzklassen der Ähnlichkeits-Relation sind nicht alle gleich groß. Beispielsweise ist die Einheitsmatrix I_n der einzig mögliche Repräsentant ihrer Äquivalenzklasse. (**Quizfrage 29.1:** Klar?) Außerdem ist jede Begleitmatrix der einzig mögliche Repräsentant ihrer Äquivalenzklasse. (**Quizfrage 29.2:** Klar?) △

Ende der Vorlesung 18

Ende der Woche 9

§ 30 Die Jordan-Normalform

Literatur: Liesen, Mehrmann, 2024, Kapitel 16

Die Frobenius-Normalform aus § 29 verwendete als Grundlage der Zerlegung des Raumes in eine direkte Summe von Unterräumen die A-invarianten Unterräume $\langle x_1,\ A\,x_1,\ A^2\,x_1,\dots,A^{m-1}\,x_1\rangle$. Die Jordan-Normalform hingegen benutzt eine andere Zerlegung, wobei die Betonung auf Eigenpaaren liegt.

Da im Allgemeinen die geometrische Vielfachheit eines Eigenvektors unterhalb der algebraischen Vielfachheit liegen kann, reichen Eigenvektoren möglicherweise nicht aus, um eine Basis des gesamten Raumes zusammenzustellen, selbst wenn wir annehmen, dass das charakteristische Polynom vollständig in Linearfaktoren zerfällt (Beispiel 24.29). Es gibt aber eine Möglichkeit, den Begriff des Eigenvektors zu verallgemeinern, um Zugriff auf weitere Basisvektoren zu bekommen.

Definition 30.1 (verallgemeinerter Eigenvektor, verallgemeinerter Eigenraum)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n\times n}$ für $n \in \mathbb{N}_0$. Ein Vektor $x \in K^n \setminus \{0\}$ heißt ein **verallgemeinerter Eigenvektor** oder **Hauptvektor** (englisch: generalized eigenvector) zum **Eigenwert $\lambda \in K$ der Matrix A**, wenn

$$(\lambda\, I - A)^k\, x = 0, \quad \text{also } x \in \operatorname{Kern}\big((\lambda\, I - A)^k\big) \tag{30.1}$$

für ein $k \in \mathbb{N}$ gilt. Wir sprechen genauer von einem verallgemeinerten Eigenvektor der **Stufe** (englisch: rank) $k \in \mathbb{N}$, wenn $x \in \operatorname{Kern}\big((\lambda\, I - A)^k\big) \setminus \operatorname{Kern}\big((\lambda\, I - A)^{k-1}\big)$ gilt.

(ii) Die Menge

$$\begin{aligned}\operatorname{GEig}(A,\lambda) &:= \big\{x \in K^n \,\big|\, (\lambda\, I - A)^k\, x = 0 \text{ für ein } k \in \mathbb{N}\big\} \\ &= \bigcup_{k\in\mathbb{N}} \operatorname{Kern}\big((\lambda\, I - A)^k\big)\end{aligned} \tag{30.2}$$

heißt der **verallgemeinerte Eigenraum** (englisch: generalized eigenspace) zu $\lambda \in K$.[12]

[12]Wie bereits in Definition 24.6 ist es auch hier praktisch, für den Begriff des verallgemeinerten Eigenraumes auch Parameter $\lambda \in K$ zuzulassen, die keine Eigenwerte sind.

(*iii*) Es sei V ein Vektorraum über K und $f \in \operatorname{End}(V)$. Ein Vektor $v \in V \setminus \{0\}$ heißt ein **verallgemeinerter Eigenvektor** oder **Hauptvektor** zum **Eigenwert** $\lambda \in K$ **des Endomorphismus** f, wenn

$$(\lambda \operatorname{id}_V - f)^k(v) = 0, \quad \text{also } v \in \operatorname{Kern}\big((\lambda \operatorname{id}_V - f)^k\big) \tag{30.3}$$

für ein $k \in \mathbb{N}$ gilt. Wir sprechen genauer von einem verallgemeinerten Eigenvektor der **Stufe** $k \in \mathbb{N}$, wenn $x \in \operatorname{Kern}\big((\lambda \operatorname{id}_V - f)^k\big) \setminus \operatorname{Kern}\big((\lambda \operatorname{id}_V - f)^{k-1}\big)$ gilt.

(*iv*) Die Menge

$$\begin{aligned} \operatorname{GEig}(f, \lambda) &:= \big\{v \in V \,\big|\, (\lambda \operatorname{id}_V - f)^k(v) = 0 \text{ für ein } k \in \mathbb{N}\big\} \\ &= \bigcup_{k \in \mathbb{N}} \operatorname{Kern}\big((\lambda \operatorname{id}_V - f)^k\big) \end{aligned} \tag{30.4}$$

heißt der **verallgemeinerte Eigenraum** zu $\lambda \in K$. △

Die verallgemeinerten Eigenräume sind A- bzw. f-invariant, denn es gilt

$$(\lambda I - A)\, A = \lambda A - A^2 = A\, (\lambda I - A)$$

und analog $(\lambda \operatorname{id}_V - f) \circ f = f \circ (\lambda \operatorname{id}_V - f)$.

Für jede quadratische Matrix A (bzw. analog für jeden Endomorphismus $f \in \operatorname{End}(V)$) gilt

$$\{0\} = \operatorname{Kern}(A^0) \subseteq \operatorname{Kern}(A^1) \subseteq \cdots \subseteq \operatorname{Kern}(A^k) \subseteq \operatorname{Kern}(A^{k+1}) \subseteq \cdots .$$

Durch das Potenzieren werden die Kerne größer, und wir können neben den Eigenvektoren aus $\operatorname{Kern}\big((\lambda I - A)^k\big)$ für $k = 1$ im Allgemeinen noch zusätzliche verallgemeinerte Eigenvektoren hinzugewinnen für $k \geqslant 2$. Es ist dabei ausreichend, den Exponenten gleich der Dimension des betrachteten Vektorraumes zu wählen, denn für höhere Potenzen werden die Kerne nicht mehr größer:

Lemma 30.2 (Kerne von Potenzen von Endormorphismen)**.**
Es sei K ein Körper.

(*i*) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Dann gilt

$$\operatorname{Kern}(A^n) = \operatorname{Kern}(A^{n+j}) \quad \text{für alle } j \in \mathbb{N}_0. \tag{30.5}$$

(*ii*) Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$ und $f \in \operatorname{End}(V)$. Dann gilt

$$\operatorname{Kern}(f^n) = \operatorname{Kern}(f^{n+j}) \quad \text{für alle } j \in \mathbb{N}_0. \tag{30.6}$$

Wegen Lemma 30.2 können wir also in der Darstellung des verallgemeinerten Eigenraumes den Exponenten auch einfach auf die Raumdimension festlegen:

$$\operatorname{GEig}(A, \lambda) = \operatorname{Kern}\big((\lambda I - A)^n\big) \quad \text{bzw.} \quad \operatorname{GEig}(f, \lambda) = \operatorname{Kern}\big((\lambda \operatorname{id}_V - f)^n\big),$$

wenn $n \in \mathbb{N}_0$ die Dimension der Matrix A bzw. des Raumes V ist.

Der folgende Satz besagt, dass wir mit Hilfe verallgemeinerter Eigenvektoren tatsächlich eine Basis des zugrundeliegenden Vektorraumes finden, sofern das charakteristische und damit auch das Minimalpolynom vollständig in Linearfaktoren zerfallen, also

$$\chi_A = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s} \tag{30.7a}$$

$$\text{bzw.} \quad \chi_f = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s} \tag{30.7b}$$

mit den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_s \in K$ und deren algebraischen Vielfachheiten $n_1, \ldots, n_s \in \mathbb{N}$ gilt. Für die algebraischen Vielfachheiten gilt dann $n_j = \dim(\mathrm{GEig}(A, \lambda_j))$ bzw. $n_j = \dim(\mathrm{GEig}(f, \lambda_j))$.

Satz 30.3 (Zerlegung in verallgemeinerte Eigenräume)**.**
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Wenn das charakteristische Polynom χ_A wie in (30.7a) vollständig in Linearfaktoren zerfällt, dann gilt:

$$K^n = \bigoplus_{j=1}^{s} \mathrm{GEig}(A, \lambda_j). \tag{30.8}$$

(ii) Es sei V ein Vektorraum über dem Körper K mit $\dim(V) = n \in \mathbb{N}_0$. Wenn das charakteristische Polynom χ_f wie in (30.7b) vollständig in Linearfaktoren zerfällt, dann gilt:

$$V = \bigoplus_{j=1}^{s} \mathrm{GEig}(f, \lambda_j). \tag{30.9}$$

Bemerkung 30.4 (zerfallendes charakteristische Polynom)**.**
Körper K, in denen jedes Polynom in $K[t]$ vom Grad $\geqslant 1$ mindestens eine Nullstelle in K besitzt, heißen **algebraisch abgeschlossen** (englisch: algebraically closed). Dazu äquivalent ist, dass jedes Polynom vom Grad $\geqslant 1$ vollständig in Linearfaktoren zerfällt. Nach dem Fundamentalsatz 11.24 ist $\mathbb{C}$ ein solcher Körper. △

Wir gehen nun auf die innere Struktur der verallgemeinerten Eigenräume einer Matrix A ein. Jeder verallgemeinerte Eigenraum $\mathrm{Kern}\big((\lambda\, I - A)^n\big)$ ist A-invariant, zerfällt aber i. A. nochmal in kleinere A-invariante Unterräume. Deren Anzahl ist gleich $\mu^{\mathrm{geo}}(A, \lambda)$. Jeder dieser Unterräume besitzt eine Basis $(x_1, \ldots, x_k)$, $k \in \mathbb{N}$, die mit einem Eigenvektor x_1 beginnt, also einem Hauptvektor der Stufe 1. Dieser wird ggf. ergänzt durch einen Hauptvektor der jeweils folgenden Stufe. Diese können wir aus den Beziehungen

$$(\lambda\, I - A)\, x_j = -x_{j-1} \text{ für } j = 2, \ldots, r \tag{30.10}$$

bestimmen. Wegen $x_1 \in \mathrm{Kern}(\lambda\, I - A)$ und dieser Konstruktion gilt in der Tat $x_j \in \mathrm{Kern}\big((\lambda\, I - A)^j\big)$. Der finale Index $r \in \mathbb{N}$ ist der erste, für den (30.10) nicht lösbar ist.

Auf dem Unterraum $\langle x_1, \ldots, x_r \rangle$ hat die durch A induzierte Abbildung bezüglich der Basis $(x_1, \ldots, x_r)$ wegen der Beziehungen

$$\begin{aligned} A\, x_1 &= \lambda\, x_1 \\ A\, x_2 &= \lambda\, x_2 + x_1 \\ \vdots\; &= \;\vdots \\ A\, x_r &= \lambda\, x_r + x_{r-1} \end{aligned}$$

eine Darstellungsmatrix wie in der folgenden Definition:

Definition 30.5 (Jordan-Block, Jordan-Matrix)**.**
Es sei K ein Körper.

(i) Es sei $\lambda \in K$ und $r \in \mathbb{N}$. Dann heißt die Matrix

$$J_r(\lambda) \coloneqq \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix} \in K^{r \times r} \tag{30.11}$$

der **Jordan-Block** (englisch: Jordan block) der Größe $r \times r$ mit Diagonaleintrag λ.[13]

(ii) Eine Matrix $A \in K^{n \times n}$, $n \in \mathbb{N}_0$, heißt eine **Jordan-Matrix** (englisch: Jordan matrix), wenn sie eine Blockdiagonalmatrix aus lauter Jordan-Blöcken ist. △

Bemerkung 30.6 (Eigenschaften von Jordan-Blöcken).

(i) Für $r = 1$ ergibt sich $J_1(\lambda) = \big(\lambda\big) \in K^{1 \times 1}$.

(ii) Der Jordan-Block $J_r(\lambda)$ ist ähnlich zur Begleitmatrix C_p des Polynoms $p = (t - \lambda)^r \in K[t]$. Daher gilt (Lemma 28.15)

$$\chi_{J_r(\lambda)} = \mu_{J_r(\lambda)} = (t - \lambda)^r \in K[t].$$

(iii) $\lambda \in K$ ist der einzige Eigenwert von $J_r(\lambda)$. Dieser weist die größtmögliche Differenz zwischen seiner algebraischen und geometrischen Vielfachheit auf, denn es gilt

$$\mu^{\text{alg}}(J_r(\lambda), \lambda) = r \quad \text{und} \quad \mu^{\text{geo}}(J_r(\lambda), \lambda) = 1.$$ △

Wir können nun das Hauptresultat dieses Abschnitts angeben:

Satz 30.7 (Jordan-Normalform).
Es sei K ein Körper.

(i) Es sei $A \in K^{n \times n}$ für $n \in \mathbb{N}_0$. Wenn das charakteristische Polynom χ_A wie in (30.7a) vollständig in Linearfaktoren zerfällt, dann gilt:

(a) A ist ähnlich zu einer Jordan-Matrix

$$\begin{bmatrix} J_{r_1}(\lambda_1) & & & \\ & J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{r_k}(\lambda_k) \end{bmatrix}. \tag{30.12}$$

Dabei sind $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von A entsprechend ihrer geometrischen Vielfachheit, also nicht notwendig alle verschieden. Wir bezeichnen (30.12) als eine **Jordan-Normalform** (englisch: Jordan normal form) von A.

(b) Die Jordan-Normalformen von A unterscheiden sich nur in der Reihenfolge ihrer Jordan-Blöcke.

[13] Manche Autoren benutzen auch den Begriff **Jordan-Kästchen**. Manchmal stehen die Einsen auch unterhalb der Hauptdiagonalen, was einer umgekehrten Nummerierung der Vektoren $x_1, \ldots, x_k$ entspricht.

(*ii*) Es sei V ein endlich-dimensionaler Vektorraum über K und $f \in \operatorname{End}(V)$. Wenn das charakteristische Polynom χ_f wie in (30.7b) vollständig in Linearfaktoren zerfällt, dann gilt:

(a) f besitzt eine Darstellungsmatrix der Gestalt (30.12). Dabei sind $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von f entsprechend ihrer geometrischen Vielfachheit, also nicht notwendig alle verschieden. Wir bezeichnen (30.12) als eine **Jordan-Normalform** von f.

(b) Die Jordan-Normalformen von f unterscheiden sich nur in der Reihenfolge ihrer Jordan-Blöcke.

Bemerkung 30.8 (Eigenschaften der Jordan-Normalform)**.**
Eine Jordan-Normalform (30.12) gibt Aufschluss über die algebraischen und geometrischen Vielfachheiten der Eigenwerte von A sowie über das charakteristische und das Minimalpolynom. Es gilt

$$\chi_A = \prod_{j=1}^{k} (t - \lambda_j)^{r_j} \qquad \text{charakteristisches Polynom} \tag{30.13a}$$

$$\text{und} \quad \mu_A = \prod_{\lambda \in \Lambda(A)} (t - \lambda)^{\max\{r_j \,|\, \lambda_j = \lambda\}} \qquad \text{Minimalpolynom.} \tag{30.13b}$$

Das bedeutet, für die algebraischen und geometrischen Vielfachheiten eines Eigenwertes $\lambda \in \Lambda(A)$ gilt

$$\mu^{\text{alg}}(A, \lambda) = \sum_{\lambda_j = \lambda} r_j \qquad \text{algebraische Vielfachheit} \tag{30.13c}$$

$$\mu^{\text{geo}}(A, \lambda) = \sum_{\lambda_j = \lambda} 1 \qquad \text{geometrische Vielfachheit.} \tag{30.13d}$$

Die algebraische Vielfachheit eines Eigenwertes λ entspricht also der Summe der Dimensionen aller Jordan-Blöcke mit dem Diagonaleintrag λ, und die geometrische Vielfachheit entspricht der Anzahl der Jordan-Blöcke mit diesem Diagonaleintrag. Die Vielfachheit der Nullstelle λ im Minimalpolynom entspricht der Größe des größten Jordan-Blocks mit dem Diagonaleintrag λ.

Zu jedem Jordan-Block mit Diagonaleintrag (Eigenwert) λ gehört genau ein Eigenvektor. Diese Eigenvektoren spannen gemeinsam den Eigenraum $\operatorname{Eig}(A, \lambda)$ auf. △

Beispiel 30.9 (Jordan-Normalform, vgl. Beispiel 29.11)**.**

(*i*) Die Matrix

$$A = \begin{bmatrix} 0 & 1 & 3 \\ 3 & 1 & -4 \\ -2 & 1 & 5 \end{bmatrix} \in \mathbb{Q}^{3 \times 3}$$

besitzt das charakteristische Polynom und Minimalpolynom $\chi_A = \mu_A = (\lambda - 2)^3$, also den einzigen Eigenwert 2 mit algebraischer Vielfachheit $\mu^{\text{alg}}(A, 2) = 3$. Aus (30.13) erfahren wir, dass wir nur Jordan-Blöcke zum Eigenwert 2 bekommen werden und dass der größte Jordan-Block die Größe 3×3 haben wird. Daher kann es nur einen einzigen Jordan-Block geben, also gilt $\mu^{\text{geo}}(A, 2) = 3$. Wir können also an dieser Stelle bereits eine Jordan-Normalform von A angeben:[14]

$$\widehat{A} = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

[14] Der besseren Lesbarkeit wegen lassen wir die strukturellen Nullen in diesem Beispiel weg.

Wollen wir auch eine Transformationsmatrix, also eine zugehörige Basis von $\mathbb{Q}^3$ bestimmen, so können wir wie folgt vorgehen: Aus dem linearen Gleichungssystem $(2 \cdot I - A)\,x_1 = 0$ erhalten wir einen ersten Eigenvektor:

$$\begin{bmatrix} 2 & -1 & -3 \\ -3 & 1 & 4 \\ 2 & -1 & -3 \end{bmatrix} x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

etwa $x_1 = \left(\begin{smallmatrix} 1 \\ -1 \\ 1 \end{smallmatrix}\right)$. Aus dem linearen Gleichungssystem

$$\begin{bmatrix} 2 & -1 & -3 \\ -3 & 1 & 4 \\ 2 & -1 & -3 \end{bmatrix} x_2 = -x_1 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$$

erhalten wir den verallgemeinerten Eigenvektor 2. Stufe $x_2 = \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right)$ und anschließend aus

$$\begin{bmatrix} 2 & -1 & -3 \\ -3 & 1 & 4 \\ 2 & -1 & -3 \end{bmatrix} x_3 = -x_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

den verallgemeinerten Eigenvektor 3. Stufe $x_3 = \left(\begin{smallmatrix} 1 \\ 2 \\ 0 \end{smallmatrix}\right)$.

Wir können jetzt mit Hilfe der verallgemeinerten Eigenvektoren die Spalten der Transformationsmatrix angeben, um die neue Basis $\widehat{B}_V$ in der alten Basis (e_1, e_2, e_3) darzustellen:

$$T^{-1} := \mathcal{T}_{\widehat{B}_V}^{(e_1,e_2,e_3)} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{und} \quad T = \begin{bmatrix} 0 & 0 & 1 \\ -2 & 1 & 3 \\ 1 & 0 & -1 \end{bmatrix}$$

Es gilt also

$$\underbrace{\begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}}_{\widehat{A}} = \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ -2 & 1 & 3 \\ 1 & 0 & -1 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} 0 & 1 & 3 \\ 3 & 1 & -4 \\ -2 & 1 & 5 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}}_{T^{-1}}.$$

(ii) Die Matrix

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ -1 & 1 & 4 \end{bmatrix} \in \mathbb{Q}^{3\times 3}$$

besitzt das charakteristische Polynom $\chi_A = (\lambda - 2)^3$ und das Minimalpolynom $\mu_A = (\lambda - 2)^2$, also den einzigen Eigenwert 2 mit algebraischer Vielfachheit $\mu^{\text{alg}}(A, 2) = 3$. Aus (30.13) erfahren wir, dass wir nur Jordan-Blöcke zum Eigenwert 2 bekommen und dass der größte Jordan-Block die Größe 2×2 haben wird. Daher muss es genau einen weiteren Jordan-Block geben, mit der Größe 1×1, also gilt $\mu^{\text{geo}}(A, 2) = 2$. Wir können also an dieser Stelle bereits eine Jordan-Normalform von A angeben:

$$\widehat{A} = \left[\begin{array}{cc|c} 2 & 1 & \\ & 2 & \\ \hline & & 2 \end{array}\right].$$

Wollen wir auch eine Transformationsmatrix, also eine zugehörige Basis von $\mathbb{Q}^3$ bestimmen, so können wir wie folgt vorgehen: Aus dem linearen Gleichungssystem $(2 \cdot I - A)\, x_1 = 0$ erhalten wir wie erwartet zwei linear unabhängige Eigenvektoren:

$$\begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ 1 & -1 & -2 \end{bmatrix} x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

nämlich $\left\langle \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 2 \\ 0 \\ 1 \end{smallmatrix}\right) \right\rangle$. Wir wissen, dass es einen Vektor x_1 in diesem Eigenraum geben muss, zu dem ein verallgemeinerter Eigenvektor der 2. Stufe existiert, also eine Lösung von

$$\begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ 1 & -1 & -2 \end{bmatrix} x_2 = -x_1 = \alpha \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ 0 \\ -1 \end{pmatrix}.$$

Die Transformation auf Zeilenstufenform zeigt, dass das mit $\alpha = 1$ und $\beta = -1$ möglich ist und wir dann $x_1 = \left(\begin{smallmatrix} -1 \\ 1 \\ -1 \end{smallmatrix}\right)$ und beispielsweise $x_2 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$ erhalten. Ein verallgemeinerter Eigenvektor 3. Stufe wird nicht existieren, da wir dann einen Jordan-Block der Größe 3×3 bekämen. In der Tat ist

$$\begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ 1 & -1 & -2 \end{bmatrix} x_3 = -x_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

nicht lösbar.[15]

Nachdem wir nun den zwei-dimensionalen A-invarianten Unterraum $\left\langle \left(\begin{smallmatrix} -1 \\ 1 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \right\rangle$ von $\operatorname{GEig}(A, 2)$ gefunden haben, benötigen wir noch einen weiteren Eigenvektor, der in einem komplementären Unterraum liegt. Dazu wählen wir beispielsweise den Vektor $\left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)$ aus $\operatorname{Eig}(A, 2)$.

Wie in Teil (i) erhalten wir mit Hilfe der verallgemeinerten Eigenvektoren die Transformationsmatrix

$$T^{-1} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \quad \text{und} \quad T = \begin{bmatrix} 0 & 0 & -1 \\ 1 & -1 & -2 \\ 0 & 1 & 1 \end{bmatrix},$$

und es gilt

$$\underbrace{\left[\begin{array}{cc|c} 2 & 1 & \\ & 2 & \\ \hline & & 2 \end{array}\right]}_{\widehat{A}} = \underbrace{\begin{bmatrix} 0 & 0 & -1 \\ 1 & -1 & -2 \\ 0 & 1 & 1 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ -1 & 1 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}}_{T^{-1}}.$$

(iii) Die Matrix

$$A = \begin{bmatrix} 4 & -1 & -2 & 3 \\ -1 & 5 & 2 & -4 \\ 0 & 1 & 3 & -1 \\ -1 & 2 & 2 & -1 \end{bmatrix} \in \mathbb{Q}^{4 \times 4}$$

besitzt das charakteristische Polynom $\chi_A = (\lambda - 3)^3 (\lambda - 2)$ und das Minimalpolynom $\mu_A = (\lambda - 3)^2 (\lambda - 2)$. Aus (30.13) erfahren wir, dass wir Jordan-Blöcke zu den Eigenwerten 3 und 2

[15] Eine alternative Vorgehensweise wäre gewesen, zuerst einen Vektor in $x_2 \in \operatorname{Kern}\big((2 \cdot I - A)^2\big) \setminus \operatorname{Kern}\big((2 \cdot I - A)^1\big)$ zu berechnen und dann $x_1 = (2 \cdot I - A)\, x_2$ zu verwenden.

bekommen werden. Der größte Jordan-Block zum Eigenwert 3 wird die Größe 3×3 haben und der größte Jordan-Block zum Eigenwert 2 die Größe 1×1. Also gibt es nur jeweils diesen einen Jordan-Block, und es gilt $\mu^{\text{geo}}(A, 3) = 1$ und $\mu^{\text{geo}}(A, 2) = 1$. Wir können also an dieser Stelle bereits eine Jordan-Normalform von A angeben:

$$\widehat{A} = \left[\begin{array}{cc|c|c} 3 & 1 & & \\ & 3 & & \\ \hline & & 3 & \\ \hline & & & 2 \end{array}\right].$$

Die Bestimmung einer Transformationsmatrix kann wie in Teil (i) erfolgen. Da hier zu jedem Eigenwert nur ein Jordan-Block vorliegt, können wir mit irgendeinem Eigenvektor starten und im Fall des Eigenwertes 3 dazu einen verallgemeinerten Eigenvektor 2. und 3. Stufe bestimmen.

(iv) Die Matrix

$$A = I = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \in \mathbb{Q}^{3 \times 3}$$

besitzt das charakteristische Polynom $\chi_A = (\lambda - 1)^3$ und das Minimalpolynom $\mu_A = \lambda - 1$. Sie liegt bereits in Jordan-Normalform vor.

Unabhängig davon erfahren wir aus (30.13), dass wir nur Jordan-Blöcke zum Eigenwert 1 bekommen und dass der größte Jordan-Block die Größe 1×1 haben wird. Daher muss es genau drei solcher Jordan-Blöcke geben, also gilt $\mu^{\text{geo}}(A, 1) = 3$. In diesem Beispiel ist jeder Vektor (außer dem Nullvektor) ein Eigenvektor von A, sodass wir jede reguläre Matrix T als Transformationsmatrix wählen können, was durch

$$I = T\, I\, T^{-1}$$

auch bestätigt wird.

(v) Die Ableitungsabbildung $f\colon K_n[t] \to K_n[t]$ auf dem Unterraum der Polynome vom Höchstgrad $n \in \mathbb{N}_0$ über einem Körper K ist durch die

$$f(t^n) := \begin{cases} n\, t^{n-1} & \text{für } n \in \mathbb{N} \\ 0 & \text{für } n = 0 \end{cases}$$

gegeben, siehe Beispiel 17.8. Ihre Darstellungsmatrix bzgl. der Monombasis $(t^0, t^1, \dots, t^n)$ ist

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & 2 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \ddots & n \\ 0 & \cdots & & \cdots & 0 \end{pmatrix} \in K^{(n+1) \times (n+1)}$$

Diese liegt schon beinahe in Jordan-Normalform vor. Durch Verwendung der Basis

$$\left(t^0, t^1, \frac{1}{2!} t^2, \frac{1}{3!} t^3, \dots \frac{1}{n!} t^n\right)$$

können wir sie leicht in Jordan-Normalform bringen. Es ergibt sich ein einzelner großer Jordan-Block $J_{n+1}(0)$ zum einzigen Eigenwert 0. △

Ende der Vorlesung 19

Kapitel 6 Innenprodukte

In diesem Kapitel kommen wir zum Begriff des Innenprodukts in Vektorräumen. Mit diesem werden wir den Vektorraum mit zusätzlichen Möglichkeiten ausstatten, u. a., den Konzepten „Länge eines Vektors“, „Winkel zwischen Vektoren“ und „Winkel zwischen Unterräumen“.

§ 31 Bilinearformen

Literatur: Fischer, Springborn, 2020, Kapitel 6

Wir beginnen mit einer Rekapitulation von Bilinearformen. Dazu sei V ein Vektorraum über dem Körper K. Eine Abbildung $\gamma\colon V \times V \to K$ heißt eine **Bilinearform** auf $V \times V$ (oder einfach: auf V), wenn für jedes feste $\overline{u} \in V$ und jedes feste $\overline{v} \in V$ die Abbildungen

$$\begin{aligned} \gamma(\overline{u}, \cdot)\colon V \ni v &\mapsto \gamma(\overline{u}, v) \in K \\ \gamma(\cdot, \overline{v})\colon V \ni u &\mapsto \gamma(u, \overline{v}) \in K \end{aligned}$$

beide linear sind. Die Menge $\operatorname{Bil}(V, V)$ aller Bilinearformen $V \times V \to K$ ist ein K-Vektorraum (Lemma 22.3).

Wir nennen eine Bilinearform wie in Definition 22.29

$$\textbf{symmetrisch}\text{, wenn} \quad \gamma(u, v) = \gamma(v, u) \quad \text{für alle } u, v \in V \tag{31.1a}$$

$$\textbf{schiefsymmetrisch}\text{, wenn} \quad \gamma(u, v) = -\gamma(v, u) \quad \text{für alle } u, v \in V \tag{31.1b}$$

$$\textbf{alternierend}\text{, wenn} \quad u = v \quad \Rightarrow \quad \gamma(u, v) = 0 \tag{31.1c}$$

gilt. Wie in Lemma 22.30 gezeigt, ist jede alternierende Bilinearform schiefsymmetrisch. Im Fall $\operatorname{char}(K) \neq 2$ ist auch jede schiefsymmetrische Bilinearform alternierend. Wir bezeichnen den Unterraum der symmetrischen Bilinearformen mit $\operatorname{Bil}_{\text{sym}}(V, V)$ und den der schiefsymmetrischen mit $\operatorname{Bil}_{\text{skew}}(V, V)$.

Beispiel 31.1 (Bilinearformen auf einem Vektorraum, vgl. Beispiel 22.2)**.**

(*i*) Für jede Matrix $A \in K^{n \times n}$ ist die Abbildung

$$\gamma_A\colon K^n \times K^n \ni (x, y) \mapsto x^\mathsf{T} A\, y \in K$$

eine Bilinearform auf K^n. Wir nennen sie die **von A induzierte Bilinearform** auf K^n und bezeichnen sie auch mit γ_A. Die Symmetrieeigenschaften von γ_A sind genau dieselben wie die von A. A ist die Darstellungsmatrix von γ_A bzgl. der Standardbasis $(e_1, \ldots, e_n)$ von K^n.

(*ii*) Insbesondere ist ($A = I$) die Abbildung

$$\gamma\colon K^n \times K^n \ni (x, y) \mapsto x^\mathsf{T} y = \sum_{i=1}^n x_i\, y_i \in K$$

eine symmetrische Bilinearform auf K^n.

(*iii*) Ein bestimmtes Integral über das Produkt zweier Polynome über $\mathbb{R}$, z. B.

$$\gamma\colon \mathbb{R}[t] \times \mathbb{R}[t] \ni (p,q) \mapsto \int_0^1 \widetilde{p}(t) \cdot \widetilde{q}(t) \,\mathrm{d}t \in \mathbb{R}$$

ist eine symmetrische Bilinearform auf $\mathbb{R}[t]$.

(*iv*) **Das folgende Beispiel wurde nachträglich hinzugefügt.** Das Produkt zweier Punktauswertungen an verschiedenen Punkten, also

$$\gamma\colon K[t] \times K[t] \ni (p,q) \mapsto \widetilde{p}(0) \cdot \widetilde{q}(1) \in K$$

ist eine nicht-symmetrische Bilinearform auf $K[t]$. Für $p = t+1$ und $q = t$ gilt nämlich $\widetilde{p}(0) = 1$ und $\widetilde{q}(1) = 1$, also $\gamma(p,q) = 1 \cdot 1 = 1$. Hingegen gilt $\widetilde{q}(0) = 0$, also auch $\gamma(q,p) = 0 \neq 1$. △

Definition 31.2 (Darstellungsmatrix einer Bilinearform)**.**
Es sei V ein endlich-dimensionaler Vektorraum über dem Körper K und $\gamma \in \operatorname{Bil}(V,V)$. Weiter sei $B_V = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$, eine Basis von V. Die Matrix

$$\mathcal{M}^{B_V}_{B_V^*}(\gamma) = \big(\gamma(v_i, v_j)\big)_{i,j=1}^n \in K^{n \times n} \tag{31.2}$$

heißt die **Darstellungsmatrix** der Bilinearform γ bzgl. der Basis B_V. △

Symmetrische Bilinearformen sind genau diejenigen mit symmetrischen Darstellungsmatrizen. Ebenso sind schiefsymmetrische Bilinearformen genau diejenigen mit schiefsymmetrischen (antisymmetrischen) Darstellungsmatrizen.

Sind $u = \sum_{i=1}^n \alpha_i \, v_i$ und $v = \sum_{j=1}^n \beta_j \, v_j$ zwei beliebige Vektoren in V, so gilt

$$\gamma(u,v) = \gamma\Big(\sum_{i=1}^n \alpha_i \, v_i, \sum_{j=1}^n \beta_j \, v_j\Big) = \sum_{i=1}^n \alpha_i \sum_{j=1}^n \beta_j \, \gamma(u_i, v_j) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}^{\mathsf{T}} \mathcal{M}^{B_V}_{B_V^*}(\gamma) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. \tag{31.3}$$

Der Wert einer Bilinearform beim Einsetzen von zwei Vektoren ergibt sich also dadurch, dass wir deren Koeffizientenvektoren von links bzw. von rechts an die Darstellungsmatrix heranmultiplizieren. Insbesondere gilt

$$\mathcal{M}^{(e_1,\ldots,e_n)}_{(e_1^*,\ldots,e_n^*)}(\gamma_A) = A.$$

Bemerkung 31.3 (zur Darstellungsmatrix von Bilinearformen)**.**

(*i*) Nach Lemma 22.28 kann eine Bilinearform auf V identifiziert werden mit einem Element $\widehat{\gamma}$ von $\operatorname{Hom}(V, V^*)$, indem man das zweite Argument von $\gamma(\cdot,\cdot)$ einsetzt und das erste freilässt, also $\widehat{\gamma}(v) = \gamma(\cdot, v)$. Die Gleichheit in (31.2) lautet eigentlich

$$\mathcal{M}^{B_V}_{B_V^*}(\widehat{\gamma}) = \big(\gamma(v_i, v_j)\big)_{i,j=1}^n.$$

(*ii*) Nach Lemma 22.28 kann eine Bilinearform auf V außerdem mit einem Tensor vom Typ $(0,2)$ über V, also einem Element von $\mathcal{T}_2^{\,0}(V)$, identifiziert werden. Wählen wir die zu $B_V = (v_1, \ldots, v_n)$ duale Basis $B_V^* = (v_1^*, \ldots, v_n^*)$ und als Basis von $\mathcal{T}_2^{\,0}(V)$ die elementaren Tensoren $v_i^* \otimes v_j^*$ (vgl. Bemerkung 22.26), so stimmt die Darstellungsmatrix $\mathcal{M}^{B_V}_{B_V^*}(\gamma)$ gerade mit der Matrix der Komponenten des Tensors überein. △

In einem endlich-dimensionalen Vektorraum V ist die zu $\gamma \in \operatorname{Hom}(V, V^*)$ **duale Bilinearform** (englisch: dual bilinear form) $\gamma^*\colon V^{**} \cong V \to V^*$ nach der allgemeinen Definition 21.26 für duale Homomorphismen definiert durch

$$\gamma^*(u, v) = \gamma(v, u) \tag{31.4}$$

für $u, v \in V$. Insofern könnten wir symmetrische Bilinearformen auch als **selbstduale Bilinearformen** (englisch: self-dual bilinear form) bezeichnen. (**Quizfrage 31.1:** In welcher Beziehung stehen die Darstellungsmatrizen von γ und γ^* zueinander?)

Für jede Basis B_V eines endlich-dimensionalen Vektorraumes V ist die Zuordnung

$$\mathcal{M}_{B_V^*}^{B_V}\colon \underbrace{\operatorname{Hom}(V, V^*)}_{\cong \operatorname{Bil}(V,V)} \ni \gamma \mapsto \mathcal{M}_{B_V^*}^{B_V}(\gamma) \in K^{n \times n} \tag{31.5}$$

einer Bilinearform zu ihrer Darstellungsmatrix ein Isomorphismus von Vektorräumen.

Aus dem Transformationssatz für Homomorphismen 20.6 folgt sofort:

Satz 31.4 (Transformation der Darstellungsmatrix einer Bilinearform beim Wechsel der Basis)**.**
Es seien K ein Körper und V ein endlich-dimensionaler Vektorraum über K. Weiter seien B_V und $\widehat{B}_V$ Basen von V. Dann gilt für die Darstellungsmatrix einer Bilinearform $\gamma\colon V \times V \to K$:

$$\mathcal{M}_{\widehat{B}_V^*}^{\widehat{B}_V}(\gamma) = \mathcal{T}_{\widehat{B}_V^*}^{B_V^*}\, \mathcal{M}_{B_V^*}^{B_V}(\gamma)\, \mathcal{T}_{B_V}^{\widehat{B}_V}. \tag{31.6}$$

Bezeichnen wir dabei wie üblich mit $T := \mathcal{T}_{\widehat{B}_V}^{B_V}$ die Transformationsmatrix und mit $A := \mathcal{M}_{B_V^*}^{B_V}(\gamma)$ und $\widehat{A} := \mathcal{M}_{\widehat{B}_V^*}^{\widehat{B}_V}(\gamma)$ die Darstellungsmatrizen, so können wir (31.6) schreiben als

$$\widehat{A} = T^{-\intercal} A\, T^{-1}.$$

Definition 31.5 (Kongruenztransformation, kongruente Matrizen)**.**
Es seien K ein Körper und $n \in \mathbb{N}_0$. Zwei Matrizen $A, \widehat{A} \in K^{n \times n}$ heißen **kongruent** (englisch: congruent), wenn es eine invertierbare Matrix $T \in K^{n \times n}$ gibt, sodass gilt:

$$\widehat{A} = T^{-\intercal} A\, T^{-1}. \tag{31.7}$$

Der Übergang von A zu $T^{-\intercal} A\, T^{-1}$ heißt auch eine **Kongruenztransformation** (englisch: congruence transformation) von A. △

Beachte: Die Kongruenz von Matrizen ist eine Äquivalenzrelation auf der Menge $K^{n \times n}$. A ist genau dann symmetrisch, wenn $\widehat{A}$ symmetrisch ist. **Quizfrage 31.2:** Klar?

Analog zu Satz 20.9 (Äquivalenztransformationen) und Satz 20.14 (Ähnlichkeitstransformationen) können wir zeigen:

Satz 31.6 (über kongruente Matrizen, vgl. Satz 20.9 und Satz 20.14)**.**
Es seien K ein Körper, $n \in \mathbb{N}_0$ und $A, \widehat{A} \in K^{n \times n}$. Weiter sei V ein Vektorraum über K mit $\dim(V) = n$ und B_V eine Basis von V mit dualer Basis B_V^*, $\gamma\colon V \times V \to K$ eine Bilinearform und $A = \mathcal{M}_{B_V^*}^{B_V}(\gamma)$.[1]
Dann sind äquivalent:

[1]Diese Voraussetzung ist nicht einschränkend. Zu einer gegebenen Matrix A können wir immer $V = K^n$, B_V als die Standardbasis und $\gamma = \gamma_A$ wählen.

(*i*) A und $\widehat{A}$ sind kongruente Matrizen.

(*ii*) $\widehat{A}$ ist die Darstellungsmatrix von γ bzgl. einer geeigneten Basis $\widehat{B}_V$ und zugehöriger dualer Basis $\widehat{B}_V^*$.

Beweis. Aussage $(i) \Rightarrow$ Aussage (ii): Wie im Beweis von Satz 20.9 können wir T^{-1} als Transformationsmatrix eines Basiswechsels $T^{-1} = \mathcal{T}_{B_V}^{\widehat{B}_V}$ auffassen. Nach (21.13) gilt dann

$$T^{-\intercal} = \mathcal{T}_{\widehat{B}_{V^*}}^{B_{V^*}}.$$

Gemäß dem Transformationssatz 20.6 für Darstellungsmatrizen haben wir

$$\mathcal{M}_{\widehat{B}_V^*}^{\widehat{B}_V}(\gamma) = \mathcal{T}_{\widehat{B}_V^*}^{B_V^*}\,\mathcal{M}_{B_V^*}^{B_V}(\gamma)\,\mathcal{T}_{B_V}^{\widehat{B}_V} = T^{-\intercal} A\, T^{-1} = \widehat{A}.$$

Aussage $(ii) \Rightarrow$ Aussage (i): Definieren wir $T^{-1} = \mathcal{T}_{B_V}^{\widehat{B}_V}$, so folgt die Behauptung sofort aus dem Transformationssatz 20.6. □

Bemerkung 31.7 (zum Transformationsverhalten von Darstellungsmatrizen)**.**
Wir haben jetzt bereits drei verschiedene Transformationsverhalten für Darstellungsmatrizen bei einem Basiswechsel kennengelernt:

Name	Transformation	Darstellungsmatrizen von
Äquivalenztransformation	$S\,A\,T^{-1}$	$\operatorname{Hom}(V, W)$
Ähnlichkeitstransformation	$T\,A\,T^{-1}$	$\operatorname{End}(V)$
Kongruenztransformation	$T^{-\intercal} A\, T^{-1}$	$\operatorname{Bil}(V, V) \cong \operatorname{Hom}(V, V^*)$

Ähnlichkeits- und Kongruenztransformationen sind dabei Spezialfälle von Äquivalenztransformationen.

Einer Matrix kann man nicht ansehen, welche Art von Abbildung sie repräsentiert! Für die Auswahl der richtigen Transformation ist das aber wesentlich, aber auch für viele andere Konzepte. Beispielsweise ist die Frage nach Eigenwerten einer Matrix oder nach ihrer Jordan-Normalform nur dann sinnvoll, wenn diese einen Endomorphismus repräsentiert!

Die Tatsache, dass eine Eigenschaft einer Matrix unter einer Klasse von Transformationen erhalten bleibt, zeigt uns ebenfalls, welche Art von Abbildung die Matrix repräsentiert. Beispielsweise bleibt die Symmetrie einer Matrix nur durch Kongruenztransformationen erhalten, nicht aber unter Äquivalenz- und Ähnlichkeitstransformationen. Insofern ist die Symmetrie einer Matrix eine strukturelle Eigenschaft der dahinterliegenden linearen Abbildung genau dann, wenn diese Abbildung ein Element von $\operatorname{Hom}(V, V^*)$ ist, also eine Bilinearform! △

Der Rang einer Matrix bleibt unter einer Äquivalenztransformation erhalten (Folgerung 15.41), also auch unter einer Kongruenztransformation. Alle Darstellungsmatrizen einer Bilinearform besitzen daher denselben Rang. Der Rang wird damit zu einer Eigenschaft der Bilinearform selbst:

Definition 31.8 (Rang einer Bilinearform).
Es seien K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $\gamma\colon V \times V \to K$ eine Bilinearform und $A = \mathcal{M}_{B_V^*}^{B_V}(\gamma) \in K^{n \times n}$ ihre Darstellungsmatrix bzgl. einer beliebigen Basis B_V von V. Wir definieren den **Rang** (englisch: rank) der Bilinearform γ durch

$$\operatorname{Rang}(\gamma) := \operatorname{Rang}(A). \tag{31.8}$$

△

Beachte: Da die duale Bilinearform durch die Abbildungsmatrix $A^\intercal$ dargestellt wird, gilt $\operatorname{Rang}(\gamma) = \operatorname{Rang}(\gamma^*)$.

Definition 31.9 (nicht-ausgeartete und perfekte Bilinearformen).
Es seien K ein Körper und V ein Vektorraum über K. Weiter sei $\gamma\colon V \times V \to K$ eine Bilinearform.

(*i*) γ heißt **nicht ausgeartet** (englisch: non-degenerate), wenn die linearen Abbildungen

$$V \ni u \mapsto \gamma(u, \cdot) \in V^* \tag{31.9a}$$
$$V \ni v \mapsto \gamma(\cdot, v) \in V^* \tag{31.9b}$$

beide injektiv sind. Andernfalls heißt γ **ausgeartet** (englisch: degenerate).

(*ii*) γ heißt **perfekt**, wenn die linearen Abbildungen

$$V \ni u \mapsto \gamma(u, \cdot) \in V^* \tag{31.10a}$$
$$V \ni v \mapsto \gamma(\cdot, v) \in V^* \tag{31.10b}$$

beide bijektiv sind. △

Das folgende Lemma stellt klar, wie wir prüfen können, ob eine Bilinearform ausgeartet ist oder nicht:

Lemma 31.10 (nicht-ausgeartete Bilinearformen).
Es seien K ein Körper und V ein Vektorraum über K. Weiter sei $\gamma\colon V \times V \to K$ eine Bilinearform. Dann sind äquivalent:

(*i*) γ ist nicht ausgeartet.

(*ii*) Es gelten

$$\gamma(u, v) = 0 \text{ für alle } v \in V \quad \Rightarrow \quad u = 0 \tag{31.11a}$$
$$\gamma(u, v) = 0 \text{ für alle } u \in V \quad \Rightarrow \quad v = 0 \tag{31.11b}$$

(*iii*) Es gelten

$$\text{für alle } u \neq 0 \text{ existiert } v \in V \text{ mit } \gamma(u, v) \neq 0 \tag{31.12a}$$
$$\text{für alle } v \neq 0 \text{ existiert } u \in V \text{ mit } \gamma(u, v) \neq 0 \tag{31.12b}$$

Beweis. Aussage (i) $\Leftrightarrow$ Aussage (ii): Es gilt:

$$\begin{aligned}
& u \mapsto \gamma(u,\cdot) \text{ ist injektiv} \\
\Leftrightarrow\quad & \operatorname{Kern}\big(u \mapsto \gamma(u,\cdot)\big) = \{0\} \quad \text{nach Lemma 17.6} \\
\Leftrightarrow\quad & \text{Wenn } \gamma(u,\cdot) \text{ die Nullabbildung in } V^* \text{ ist, dann muss } u = 0 \text{ gelten.} \\
\Leftrightarrow\quad & (31.11a)
\end{aligned}$$

Analog folgt auch die Äquivalenz der Injektivität von $v \mapsto \gamma(\cdot,v)$ und (31.11b).

Aussage (ii) $\Leftrightarrow$ Aussage (iii): (31.11a) ist äquivalent zu (31.12a), und (31.11b) ist äquivalent zu (31.12b). □

Im Fall endlich-dimensionaler Vektorräume fallen die Eigenschaften „nicht ausgeartet" und „perfekt" zusammen und lassen sich zudem anhand einer Darstellungsmatrix prüfen:

Lemma 31.11 (nicht-ausgeartete Bilinearformen in endlicher Dimension)**.**
Es seien K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $\gamma\colon V \times V \to K$ eine Bilinearform und $A = \mathcal{M}^{B_V}_{B_V^*}(\gamma) \in K^{n \times n}$ ihre Darstellungsmatrix bzgl. einer beliebigen Basis B_V von V. Dann sind äquivalent:

(i) $u \mapsto \gamma(u,\cdot)$ ist injektiv.

(ii) $v \mapsto \gamma(\cdot,v)$ ist injektiv.

(iii) γ ist nicht ausgeartet.

(iv) γ ist perfekt.

(v) A ist regulär.

(vi) $\operatorname{Rang}(\gamma) = \operatorname{Rang}(A) = n$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (v): Es gilt $\dim(V) = \dim(V^*)$. Die Injektivität der Abbildung $u \mapsto \gamma(u,\cdot) \in \operatorname{Hom}(V,V^*)$ impliziert also bereits die Bijektivität (Folgerung 18.9). Das heißt wiederum (Satz 19.12), dass die Darstellungsmatrix $A^\intercal$ regulär ist, also auch A.

Aussage (ii) $\Rightarrow$ Aussage (v) folgt analog.

Aussage (iii) $\Rightarrow$ Aussage (i): klar nach Definition.

Aussage (iii) $\Rightarrow$ Aussage (ii): klar nach Definition.

Aussage (iii) $\Rightarrow$ Aussage (iv): Es gilt $\dim(V) = \dim(V^*)$. Die Injektivität der Abbildung $u \mapsto \gamma(u,\cdot) \in \operatorname{Hom}(V,V^*)$ impliziert also bereits die Bijektivität (Folgerung 18.9). Weiter impliziert die Injektivität der Abbildung $v \mapsto \gamma(\cdot,v) \in \operatorname{Hom}(V,V^*)$ bereits die Bijektivität. Also ist γ perfekt.

Aussage (iv) $\Rightarrow$ Aussage (iii): klar nach Definition.

Aussage (v) $\Rightarrow$ Aussage (iii): Die Darstellungsmatrix A ist regulär, also ist die Abbildung $v \mapsto \gamma(\cdot,v)$ bijektiv (Satz 19.12). Da auch $A^\intercal$ regulär ist, ist auch die Abbildung $u \mapsto \gamma(u,\cdot)$ bijektiv.

Aussage (v) $\Leftrightarrow$ Aussage (vi): klar nach Satz 15.40. □

Ende der Vorlesung 20

Ende der Woche 10

§ 32 Quadratische Formen

Literatur: Fischer, Springborn, 2020, Kapitel 6

Definition 32.1 (quadratische Form)**.**
Es seien K ein Körper und V ein Vektorraum über K.

(i) Eine Abbildung $q\colon V \to K$ heißt eine **quadratische Form** (englisch: quadratic form) auf V, wenn gilt:

$$q(\alpha\, u) = \alpha^2 q(u) \quad \text{für alle } \alpha \in K \text{ und alle } v \in V \tag{32.1a}$$

$$\text{Die Abbildung } \Gamma\colon V \times V \ni (u,v) \mapsto q(u+v) - q(u) - q(v) \in K \text{ ist eine Bilinearform.} \tag{32.1b}$$

(ii) Die Menge aller quadratischen Formen auf V bezeichnen wir mit $\operatorname{QF}(V)$. △

Beachte: Die Abbildung in (32.1b) ist immer symmetrisch, d. h., $\Gamma(u,v) = \Gamma(v,u)$.

Die Bedingung (32.1b) lautet ausgeschrieben:

$$\begin{aligned}
&\Gamma(u_1 + u_2, v) = \Gamma(u_1, v) + \Gamma(u_2, v), \text{ also}\\
&q(u_1 + u_2 + v) - q(u_1 + u_2) - q(v) = q(u_1 + v) - q(u_1) - q(v) + q(u_2 + v) - q(u_2) - q(v)
\end{aligned} \tag{32.2a}$$

$$\begin{aligned}
&\Gamma(\alpha\, u, v) = \alpha\, \Gamma(u, v), \text{ also}\\
&q(\alpha\, u + v) - q(\alpha\, u) - q(v) = \alpha\, q(u + v) - \alpha\, q(u) - \alpha\, q(v)
\end{aligned} \tag{32.2b}$$

für alle $u, u_1, u_2, v, v_1, v_2 \in V$ und $\alpha \in K$. (**Quizfrage 32.1:** Warum müssen wir nicht auch noch $\Gamma(u, v_1 + v_2) = \Gamma(u, v_1) + \Gamma(u, v_2)$ sowie $\Gamma(u, \alpha\, v) = \alpha\, \Gamma(u, v)$ fordern?)

Lemma 32.2 (quadratische Formen bilden einen Vektorraum)**.**
Es seien K ein Körper und V ein Vektorraum über K. Dann ist $\operatorname{QF}(V)$ ein Unterraum des Vektorraumes $K^V = \{f\colon V \to K\}$ aller Abbildungen $V \to K$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-11.1. □

Das folgende Lemma zeigt, dass die **Diagonale** jeder Bilinearform eine quadratische Form ergibt:

Lemma 32.3 (Bilinearformen induzieren quadratische Formen)**.**
Es seien K ein Körper und V ein Vektorraum über K. Ist $\gamma\colon V \times V \to K$ eine Bilinearform auf V, dann ist

$$q_\gamma(u) \coloneqq \gamma(u,u) \tag{32.3}$$

eine quadratische Form auf V mit zugehörigem $\Gamma = \gamma + \gamma^*$.

Beachte: Bis auf den Faktor 2 ist Γ der symmetrische Anteil von γ, vgl. Lemma 15.29.

Beweis. Für $\alpha \in K$ gilt

$$\begin{aligned} q_\gamma(\alpha\, u) &= \gamma(\alpha\, u, \alpha\, u) && \text{nach Definition von } q_\gamma \\ &= \alpha^2 \gamma(u,u) && \text{wegen der Bilinearität von } \gamma \\ &= \alpha^2 q_\gamma(u) && \text{nach Definition von } q_\gamma \end{aligned}$$

Das heißt, (32.1a) ist für q_γ erfüllt.

Für die Abbildung Γ gilt außerdem

$$\begin{aligned} \Gamma(u,v) &= q_\gamma(u+v) - q_\gamma(u) - q_\gamma(v) && \text{nach Definition von } \gamma_{(q_\gamma)}\text{, siehe (32.1b)} \\ &= \gamma(u+v, u+v) - \gamma(u,u) - \gamma(v,v) && \text{nach Definition von } q_\gamma \\ &= \pm\gamma(u,u) \pm \gamma(v,v) + \gamma(u,v) + \gamma(v,u) && \text{wegen der Bilinearität von } \gamma \\ &= \gamma(u,v) + \gamma^*(u,v) && \text{nach Definition (31.4) der dualen Form,} \end{aligned}$$

also $\Gamma = \gamma + \gamma^*$. Sowohl γ als auch γ^* sind Bilinearformen auf V, also auch die Summe (Lemma 22.3). Damit ist auch (32.1b) für Γ erfüllt. □

Umgekehrt ergibt sich aus jeder quadratischen Form q durch (32.1b) eine symmetrische Bilinearform Γ. Der Zusammenhang wird durch folgenden Satz klargestellt:

Satz 32.4 (Zusammenhang zwischen quadratischen Formen und symmetrischen Bilinearformen)**.** Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und V ein Vektorraum über K. Dann ist

$$\begin{aligned} q_\bullet\colon\ & \operatorname{Bil_{sym}}(V,V) \ni \gamma \mapsto q_\gamma \in \operatorname{QF}(V) \\ & q_\gamma(u) \coloneqq \gamma(u,u) \end{aligned} \tag{32.4}$$

ein Isomorphismus von Vektorräumen. Dabei heißt q_γ die durch γ **induzierte quadratische Form** (englisch: induced quadratic form). Die zu $q_\bullet$ inverse Abbildung ist

$$\begin{aligned} \gamma_\bullet\colon\ & \operatorname{QF}(V) \ni q \mapsto \gamma_q \in \operatorname{Bil_{sym}}(V,V) \\ & \gamma_q(u,v) \coloneqq \frac{1}{2}\big(q(u+v) - q(u) - q(v)\big), \end{aligned} \tag{32.5}$$

genannt die durch q **induzierte (symmetrische) Bilinearform** (englisch: induced symmetric bilinear form). Die Gleichung (32.5) heißt **Polarisierungsformel** (englisch: polarizing identity).

Beachte: Hier und im folgenden Beweis benutzen wir die Abkürzung $2 \coloneqq 1+1 \in K$ sowie $4 = 2 \cdot 2 \in K$. Wegen $\operatorname{char}(K) \neq 2$ ist $2 \neq 0$

Beweis. In Lemma 32.3 hatten wir gezeigt, dass $q_\bullet$ tatsächlich in $\operatorname{QF}(V)$ abbildet (sogar für beliebige und nicht nur für symmetrische Bilinearformen). Nach Definition 32.1 ist $(u,v) \mapsto q(u+v) - q(u) - q(v)$ eine Bilinearform, also auch $\frac{1}{2}\big(q(u+v) - q(u) - q(v)\big)$. Also bildet $\gamma_\bullet$ tatsächlich in $\operatorname{Bil_{sym}}(V,V)$ ab.

Wir zeigen die Linearität von $q_\bullet$:

$$q_{\alpha\,\gamma_1 + \beta\,\gamma_2}(u) = (\alpha\,\gamma_1 + \beta\,\gamma_2)(u,u) = \alpha\,\gamma_1(u,u) + \beta\,\gamma_2(u,u) = \alpha\, q_{\gamma_1}(u) + \beta\, q_{\gamma_2}(u).$$

Es bleibt zu bestätigen, dass $q_\bullet$ und $\gamma_\bullet$ tatsächlich Inverse voneinander sind. Für $q \in \mathrm{QF}(V)$ gilt

$$\begin{aligned}
&\big((q_\bullet \circ \gamma_\bullet)(q)\big)(u) \\
&\quad = \big(q_\bullet(\gamma_q)\big)(u) && \text{nach Definition der Komposition von Funktionen} \\
&\quad = \gamma_q(u,u) && \text{nach Definition von } q_\bullet \\
&\quad = \frac{1}{2}\big(q(u+u) - q(u) - q(u)\big) && \text{nach Definition von } \gamma_q \\
&\quad = \frac{1}{2} q(2\,u) - q(u) \\
&\quad = \frac{1}{2} \cdot 4\,q(u) - q(u) && \text{wegen (32.1a)} \\
&\quad = q(u).
\end{aligned}$$

Damit ist $q_\bullet \circ \gamma_\bullet = \mathrm{id}_{\mathrm{QF}(V)}$ bestätigt.

Nun betrachten wir $\gamma_\bullet \circ q_\bullet$. Für $\gamma \in \mathrm{Bil}_{\mathrm{sym}}(V,V)$ gilt

$$\begin{aligned}
&\big((\gamma_\bullet \circ q_\bullet)(\gamma)\big)(u,v) \\
&\quad = \big(\gamma_\bullet(q_\gamma)\big)(u,v) && \text{nach Definition der Komposition von Funktionen} \\
&\quad = \frac{1}{2}\big(q_\gamma(u+v) - q_\gamma(u) - q_\gamma(v)\big) && \text{nach Definition von } \gamma_\bullet \\
&\quad = \frac{1}{2}\big(\gamma(u+v,u+v) - \gamma(u,u) - \gamma(v,v)\big) && \text{nach Definition von } q_\gamma \\
&\quad = \frac{1}{2}\big(\gamma(u,v) + \gamma^*(u,v)\big) && \text{wie im Beweis von Lemma 32.3} \\
&\quad = \gamma(u,v) && \text{wegen der Symmetrie von } \gamma.
\end{aligned}$$

Damit ist auch $\gamma_\bullet \circ q_\bullet = \mathrm{id}_{\mathrm{Bil}_{\mathrm{sym}}(V,V)}$ bestätigt. □

Bemerkung 32.5 (Zusammenhang zwischen quadratischen Formen und symmetrischen Bilinearformen)**.**

(*i*) In Vektorräumen über Körpern K mit $\mathrm{char}(K) \neq 2$ können aufgrund von Satz 32.4 quadratische Formen auf V mit symmetrischen Bilinearformen auf V identifiziert werden.

Damit sind auch die Dimensionen beider Vektorräume gleich, es gilt also

$$\dim\big(\mathrm{QF}(V)\big) = \dim\big(\mathrm{Bil}_{\mathrm{sym}}(V,V)\big) = \frac{1}{2} n\,(n+1)$$

im Fall $\dim(V) = n \in \mathbb{N}_0$, siehe Lemma 15.29.

(*ii*) Im Fall $\mathrm{char}(K) = 2$ ist $q_\bullet\colon \mathrm{Bil}_{\mathrm{sym}}(V,V) \to \mathrm{QF}(V)$ noch wohldefiniert, aber im Allgemeinen weder surjektiv noch injektiv.

(*iii*) In Vektorräumen über Körpern K mit $\mathrm{char}(K) \neq 2$ ist eine symmetrische Bilinearform auf V durch die Polarisierungsformel (32.5) bereits durch ihre Werte entlang der Diagonalen eindeutig festgelegt, denn es gilt (siehe Beweis von Satz 32.4)

$$\gamma(u,v) = \frac{1}{2}\big(\gamma(u+v,u+v) - \gamma(u,u) - \gamma(v,v)\big).$$ △

Beispiel 32.6 (Zusammenhang zwischen quadratischen Formen und symmetrischen Bilinearformen)**.**

(*i*) Die Matrizen

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \in \mathbb{Q}^{2\times 2} \quad \text{und} \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \in \mathbb{Q}^{2\times 2}$$

induzieren verschiedene Bilinearformen, nämlich

$$\begin{aligned} \gamma_A(x,y) &= x^\mathsf{T} A\, y = x_1\, y_1 + 2\, x_1\, y_2 + 3\, x_2\, y_2 && \text{(nicht symmetrisch)} \\ \gamma_B(x,y) &= x^\mathsf{T} B\, y = x_1\, y_1 + x_1\, y_2 + x_2\, y_1 + 3\, x_2\, y_2 && \text{(symmetrisch).} \end{aligned}$$

Diese jedoch induzieren dieselbe quadratische Form, nämlich

$$q(x) = \gamma_A(x,x) = x_1^2 + 2\, x_1\, x_2 + 3\, x_2^2 = \gamma_B(x,x).$$

(*ii*) Aus der quadratischen Form

$$q(x) = -x_1^2 - 8\, x_1\, x_2 + 5\, x_2^2$$

über $\mathbb{Q}^2$ können wir durch die Polarisierungsformel (32.5) die zugehörige Bilinearform bestimmen. Wir können sie allerdings hier auch direkt aus den Koeffizienten ablesen. Es gilt

$$\gamma_q(x,y) = x^\mathsf{T} \begin{bmatrix} -1 & -4 \\ -4 & 5 \end{bmatrix} y.$$

(*iii*) Die Abbildung

$$f\colon \mathbb{Q}^2 \times \mathbb{Q}^2 \ni (x,y) \mapsto 3\, x_1\, y_1 + y_1\, y_2 \in \mathbb{Q}$$

ist keine Bilinearform, denn

$$\begin{aligned} f(2\,x, y) &= 6\, x_1\, y_1 + y_1\, y_2 \\ 2\, f(x,y) &= 6\, x_1\, y_1 + 2\, y_1\, y_2 \end{aligned}$$

sind verschiedene Abbildungen.

(*iv*) Die Abbildung

$$K[t] \ni p \mapsto \int_0^1 \widetilde{p}(t)\, \widetilde{p}'(t)\, \mathrm{d}t \in K$$

ist eine quadratische Form. △

§ 33 Quadratische Räume und Orthogonalität

Definition 33.1 (quadratischer Raum, Orthogonalität, orthogonales Komplement)**.**
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und V ein Vektorraum über K.

(*i*) Ist $\gamma \in \operatorname{Bil}_{\text{sym}}(V,V)$ eine symmetrische Bilinearform auf V und q die zugehörige quadratische Form, dann heißt (V,γ) ein **quadratischer Raum** (englisch: quadratic space) **über** V.[2]

(*ii*) Zwei Vektoren $u, v \in V$ heißen **orthogonal bzgl. der symmetrischen Bilinearform** γ, wenn $\gamma(u,v) = 0$ gilt. Wir schreiben dafür auch $u \perp_\gamma v$ oder $u \perp v$.

(*iii*) Wir führen auch die Mengenschreibweise $E_1 \perp E_2$ für $E_1, E_2 \subseteq V$ ein. Diese bedeutet $u_1 \perp u_2$ für alle $u_1 \in E_1$ und alle $u_2 \in E_2$. Vereinfachend schreiben wir $u \perp E$ an Stelle von $\{u\} \perp E$.

[2] Sachgerechter wäre es, (V,q) als quadratischen Raum zu bezeichnen, allerdings benutzen wir im Folgenden die Bilinearform γ viel häufiger als die zugehörige quadratische Form q.

(*iv*) Eine Menge $E \subseteq V$ von Vektoren in V heißt **orthogonal** (englisch: orthogonal set) **bzgl.** γ, wenn ihre Elemente paarweise orthogonal sind, d. h., wenn gilt: $u \perp v$ für alle $u, v \in E$ mit $u \neq v$.

(*v*) Eine Familie $(v_i)_{i \in I}$ von Vektoren in V heißt **orthogonal** (englisch: orthogonal family) **bzgl.** γ, wenn ihre Mitglieder paarweise orthogonal sind, d. h., wenn gilt: $v_i \perp v_j$ für alle $i, j \in I$ mit $i \neq j$.

(*vi*) Der Unterraum

$$E^\perp \coloneqq \left\{ v \in V \,\middle|\, \gamma(u, v) = 0 \text{ für alle } u \in E \right\} \tag{33.1}$$

heißt das **orthogonale Komplement** (englisch: orthogonal complement) **bzgl.** γ der Menge $E \subseteq V$. Vereinfachend schreiben wir $u^\perp$ an Stelle von $\{u\}^\perp$. △

Statt „**orthogonal bzgl.** γ“ sagen wir kurz auch „γ**-orthogonal**“.

Beispiel 33.2 (quadratischer Raum, Orthogonalität)**.**

(*i*) Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und V ein Vektorraum über K. Weiter sei γ die Nullform auf $V \times V$. Dann ist (V, γ) ein quadratischer Raum, in dem zwei beliebige Vektoren stets γ-orthogonal sind.

(*ii*) In $V = \mathbb{R}^n$ erzeugt die symmetrische Bilinearform $\gamma(x, y) \coloneqq x^\intercal y$ den bekannten Begriff von Orthogonalität. △

Satz 33.3 (**Satz des Pythagoras**[3])**.**
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ) ein quadratischer Raum über K. Ist die Familie $(v_1, \ldots, v_k)$ γ-orthogonal, dann gilt

$$\gamma\Big(\sum_{i=1}^k v_i, \sum_{j=1}^k v_j\Big) = \sum_{i=1}^k \gamma(v_i, v_i). \tag{33.2}$$

Beweis.

$$\gamma\Big(\sum_{i=1}^k v_i, \sum_{j=1}^k v_j\Big) = \sum_{i=1}^k \sum_{j=1}^k \gamma(v_i, v_j) = \sum_{i=1}^k \gamma(v_i, v_i).$$

□

Definition 33.4 (Orthogonalbasis, orthogonale direkte Summe)**.**
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ) ein quadratischer Raum über K.

(*i*) Eine Familie $B \coloneqq (v_i)_{i \in I}$ von Vektoren in V heißt eine **Orthogonalbasis** (englisch: orthogonal basis) von (V, γ), wenn B eine Basis von V und außerdem orthogonal bzgl. γ ist.

(*ii*) Es sei $(U_i)_{i \in I}$ eine nichtleere Familie von Unterräumen von V. Die Summe $\sum_{i \in I} U_i = \left\langle \bigcup_{i \in I} U_i \right\rangle$ dieser Familie heißt eine **orthogonale direkte Summe** (englisch: direct sum), wenn die Summe direkt ist (Definition 14.12) und $U_i \perp U_j$ gilt für alle $i, j \in I$ mit $i \neq j$.[4]

Wir schreiben für eine orthogonale direkte Summe auch $\bigoplus_{i \in I} U_i$. △

Beachte: Beispielsweise im Fall $V = U_1 \oplus U_2$ sind U_1 und U_2 gegenseitig orthogonale Komplemente.

[3] englisch: Pythagorean Theorem

[4] Aus $U_i \perp U_j$ folgt automatisch auch $U_i \perp \sum_{j \in I \setminus \{i\}} U_j$. (**Quizfrage 33.1:** Klar?)

Satz 33.5 (orthogonale direkte Summe von Unterräumen und Partitionierung einer Orthogonalbasis, vgl. Satz 14.16).
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ) ein quadratischer Raum über K. Dann gilt:

(i) Ist B eine Orthogonalbasis von V und $(B_i)_{i \in I}$ eine Partition von B 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 Orthogonalbasen B_i, $i \in I$, und gilt $V = \bigoplus_{i \in I} U_i$, so ist $\bigcup_{i \in I} B_i$ eine Orthogonalbasis von V.

Beweis. Aussage (i): Aus Satz 14.16 folgt, dass $V = \bigoplus_{i \in I} \langle B_i \rangle$ eine direkte Summe ist. Es bleibt $\langle B_i \rangle \perp \langle B_j \rangle$ zu zeigen für $i, j \in I$, $i \neq j$. Dazu sei $u \in \langle B_i \rangle$ und $v \in \langle B_j \rangle$. Diese Vektoren haben eine Darstellung $u = \sum_{k=1}^m \alpha_k \, v_k$ bzw. $v = \sum_{\ell=1}^n \beta_\ell \, v_\ell$ mit Basisvektoren $v_k \in B_i$ und $v_\ell \in B_j$. Es gilt

$$\gamma(u, v) = \gamma\Big(\sum_{k=1}^m \alpha_k \, v_k, \sum_{\ell=1}^n \beta_\ell \, v_\ell\Big) = \sum_{k=1}^m \sum_{\ell=1}^n \alpha_k \, \beta_\ell \, \gamma(v_k, v_\ell) = 0,$$

denn B_i und B_j sind disjunkte Teilmengen einer Orthogonalbasis, und daher gilt $\gamma(v_k, v_\ell) = 0$.

Aussage (ii): Aus Satz 14.16 folgt zunächst, dass $B \coloneqq \bigcup_{i \in I} B_i$ eine Basis von V ist. Zwei Vektoren aus B sind entweder aus derselben Teilmenge B_i oder aus zwei verschiedenen Teilmengen B_i und B_j mit $i \neq j$. Im ersten Fall sind sie orthogonal, da B_i eine Orthogonalbasis ist. Im zweiten Fall sind sie orthogonal, da $B_i \subseteq U_i$ und $B_j \subseteq U_j$ gilt und die Orthogonalität von U_i und U_j angenommen wurde. □

Folgerung 33.6 (orthogonale direkte Summe von Unterräumen und **Satz des Pythagoras**).
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ) ein quadratischer Raum über K. Weiter sei $(U_i)_{i=1}^m$, $m \in \mathbb{N}$, eine Familie von Unterräumen von V und $V = \bigoplus_{i=1}^m U_i$. Ist $v = \sum_{i=1}^m v_i$ die eindeutige Zerlegung von $v \in V$ in $v_i \in U_i$, dann gilt

$$\gamma(v, v) = \sum_{i=1}^m \gamma(v_i, v_i). \tag{33.3}$$

Beweis. Zunächst halten wir fest, dass nach Satz 14.15 jedes $v \in V$ eine eindeutige Zerlegung der Form $v = \sum_{i=1}^m v_i$ mit $v_i \in U_i$ besitzt. Da die Familie $(v_1, \ldots, v_m)$ orthogonal ist, folgt die Behauptung aus Satz 33.3. □

Lemma 33.7 (Orthogonalbasen und diagonale Darstellungsmatrizen).
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ) ein endlich-dimensionaler quadratischer Raum über K. Weiter sei B_V eine Basis von V. Dann sind äquivalent:

(i) B_V ist eine Orthogonalbasis von V.

(ii) Die Darstellungsmatrix $\mathcal{M}_{B_V^*}^{B_V}(\gamma)$ ist diagonal.

Beweis. Es sei $B_V = (v_1, \ldots, v_n)$ mit $n = \dim(V) \in \mathbb{N}_0$. Die Orthogonalität der Familie ist nach Definition äquivalent zu $\gamma(v_i, v_j) = 0$ für $i \neq j$. Da die Darstellungsmatrix $\mathcal{M}_{B_V^*}^{B_V}(\gamma)$ die Einträge $\gamma(v_i, v_j)$ hat, ist die Orthogonalität äquivalent zur Diagonalität der Darstellungsmatsrix. □

Definition 33.8 (Homomorphismus quadratischer Räume, vgl. Definition 17.1 eines Vektorraumhomomorphismus).
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ_1) sowie (W, γ_2) quadratische Räume über K.

(*i*) Eine Abbildung $f\colon V \to W$ heißt **strukturverträglich** oder ein **Homomorphismus quadratischer Räume** (englisch: homomorphism of quadratic spaces) von (V, γ_1) in (W, γ_2), wenn gilt:

$$f \text{ ist linear} \tag{33.4a}$$

$$\gamma_2(f(u), f(v)) = \gamma_1(u, v) \quad \text{für alle } u, v \in V. \tag{33.4b}$$

(*ii*) Ist $f\colon V \to W$ bijektiv, so heißt f auch **strukturerhaltend** oder ein **Isomorphismus quadratischer Räume** (englisch: isomorphism of quadratic spaces). In diesem Fall nennen wir (V, γ_1) und (W, γ_2) auch zueinander **isomorphe quadratische Räume** (englisch: isomorphic quadratic spaces) und schreiben

$$(V, \gamma_1) \cong (W, \gamma_2).$$ △

Lemma 33.9 (Inverse eines Isomorphismus quadratischer Räume)**.**
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ_1) sowie (W, γ_2) quadratische Räume über K. Ist $f\colon V \to W$ ein Isomorphismus quadratischer Räume, dann auch f^{-1}. Wenn V endlich-dimensional ist, dann gilt außerdem $\operatorname{Rang}(\gamma_1) = \operatorname{Rang}(\gamma_2)$.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-11.1. □

§ 33.1 Normalformen symmetrischer Bilinearformen

Wir klären nun die Frage nach Normalformen von quadratischen Formen bzw. symmetrischen Bilinearformen.

Satz 33.10 (jeder endlich-dimensionale quadratische Raum besitzt eine Orthogonalbasis)**.**
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und (V, γ) ein endlich-dimensionaler quadratischer Raum über K. Dann gilt:

(*i*) V besitzt eine Orthogonalbasis B_V. In dieser ist die Darstellungsmatrix $\mathcal{M}_{B_V^*}^{B_V}(\gamma)$ diagonal.

(*ii*) Die Basis B_V kann so gewählt werden, dass die Darstellungsmatrix die Gestalt

$$\mathcal{M}_{B_V^*}^{B_V}(\gamma) = \begin{bmatrix} \alpha_1 & & & & & \\ & \ddots & & & & \\ & & \alpha_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \tag{33.5}$$

hat, wobei $r = \operatorname{Rang}(\gamma)$ gilt und die $\alpha_1, \ldots, \alpha_r \in K \setminus \{0\}$ sind.

Beweis. Aussage (*i*): Wir können annehmen, dass γ nicht die Nullform ist, denn ansonsten ist jede beliebige Basis B_V eine Orthogonalbasis und die zugehörige Darstellungsmatrix die Nullmatrix, also diagonal.

Wir führen einen Induktionsbeweis nach $n = \dim(V) \in \mathbb{N}_0$. Im Fall $n = 0$ ist die einzig mögliche Basis die leere Familie $B_V = ()$, also eine Orthogonalbasis. Unter der Annahme, die Behauptung sei für

$n \in \mathbb{N}_0$ bereits gezeigt, schließen wir nun auf $n+1$. Da γ nicht die Nullform ist, gibt es ein $v_1 \in V$ mit $\gamma(v_1, v_1) \neq 0$. Wir identifizieren γ mit der Abbildung $v \mapsto \gamma(\cdot, v)$. Wir betrachten nun

$$\operatorname{Kern}(\gamma(\cdot, v_1)) = \operatorname{Kern}(v \mapsto \gamma(v, v_1)) = \{v \in V \mid \gamma(v, v_1) = 0\}.$$

Aus dem Dimensionssatz für lineare Abbildungen (Folgerung 18.7) folgt

$$\begin{aligned} n+1 &= \dim(V) \\ &= \dim \operatorname{Kern}(\gamma(\cdot, v_1)) + \dim \operatorname{Bild}(\gamma(\cdot, v_1)). \end{aligned}$$

Da $\gamma(\cdot, v_1)$ nach K abbildet, gilt $\dim \operatorname{Bild}(\gamma(\cdot, v_1)) \in \{0, 1\}$. Wegen $\gamma(v_1, v_1) \neq 0$ ist nur $\dim \operatorname{Bild}(\gamma(\cdot, v_1)) = 1$ möglich, also $\dim \operatorname{Kern}(\gamma(\cdot, v_1)) = n$. Weiterhin ist $v_1 \notin \operatorname{Kern}(\gamma(\cdot, v_1))$.

Zur Abkürzung setzen wir $H \coloneqq \operatorname{Kern}(\gamma(\cdot, v_1))$. Aus Satz 14.15 folgt nun, dass $V = \langle v_1 \rangle \oplus H$ eine direkte Summe ist. Wir haben sogar $V = \langle v_1 \rangle \ominus H$, da $\gamma(v, v_1) = 0$ ist für alle $v \in H = \operatorname{Kern}(\gamma(\cdot, v_1))$.

Nun ist $(H, \gamma|_{H \times H})$ ein quadratischer Raum mit $\dim(H) = n$. Nach Induktionsvoraussetzung existiert also eine Orthogonalbasis $(v_2, \ldots, v_n)$ von $(H, \gamma|_{H \times H})$. Aus Satz 33.5 folgt, dass $(v_1, v_2, \ldots, v_n)$ eine Orthogonalbasis von V ist.

Aussage (ii): Durch Umsortieren der Basisvektoren können wir die diagonale Darstellungsmatrix $\mathcal{M}_{B_V^*}^{B_V}(\gamma)$ in die Gestalt (33.5) bringen. Dieses Umsortieren entspricht einer weiteren Kongruenztransformation. (**Quizfrage 33.2:** Klar?) Da alle Darstellungsmatrizen von γ denselben Rang besitzen (vgl. Definition 31.8), muss auch (33.5) den Rang r besitzen. Das heißt, $\alpha_1, \ldots, \alpha_r \in K \setminus \{0\}$. □

Bemerkung 33.11 (zu Satz 33.10).

(i) Die Werte $\alpha_1, \ldots, \alpha_r$ sind nicht eindeutig bestimmt. (**Quizfrage 33.3:** Was passiert zum Beispiel, wenn wir den ersten Basisvektor in B_V mit einem $\alpha \notin \{0, 1\}$ multiplizieren?)

(ii) Bezeichnen wir die Matrix in (33.5) mit A, so können wir die Aussage von Satz 33.10 auch wie folgt formulieren: Der quadratische Raum (V, γ) ist isomorph zu (K^n, γ_A). △

Folgerung 33.12 (jede symmetrische Matrix ist kongruent zu einer Diagonalmatrix).
Es seien K ein Körper mit $\operatorname{char}(K) \neq 2$ und $A \in K^{n \times n}$ symmetrisch, $n \in \mathbb{N}_0$. Dann gibt es eine reguläre Matrix $T \in K^{n \times n}$, sodass $T^{-\mathsf{T}} A \, T^{-1}$ eine Diagonalmatrix der Gestalt (33.5) ist.

Beweis. Wir betrachten die durch A induzierte symmetrische Bilinearform

$$\gamma_A \colon K^n \times K^n \ni (x, y) \mapsto x^\mathsf{T} A \, y \in K$$

auf K^n, vgl. Beispiel 31.1. A ist die Darstellungsmatrix von γ_A bzgl. der Standardbasis $B_V = (e_1, \ldots, e_n)$.

Der quadratische Raum (K^n, γ_A) besitzt nach Satz 33.10 eine Orthogonalbasis $\widehat{B}_V$, bzgl. der die Darstellungsmatrix $\widehat{A} \coloneqq \mathcal{M}_{\widehat{B}_V^*}^{\widehat{B}_V}(\gamma_A)$ eine Diagonalmatrix der Gestalt (33.5) ist. Nach Satz 31.4 gilt für den Zusammenhang der beiden Darstellungsmatrizen $\widehat{A} = T^{-\mathsf{T}} A \, T^{-1}$. □

Ende der Vorlesung 21

§ 33.2 Normalformen symmetrischer Bilinearformen in reellen Vektorräumen

Über dem Körper $K = \mathbb{R}$ der reellen Zahlen erhalten wir folgende Version von Satz 33.10:

Satz 33.13 (jeder endlich-dimensionale reelle quadratische Raum besitzt eine Orthogonalbasis, vgl. Satz 33.10)**.**
Es sei (V, γ) ein endlich-dimensionaler quadratischer Raum über $\mathbb{R}$. Dann gilt: V besitzt eine Orthogonalbasis B_V, bzgl. der die Darstellungsmatrix die Gestalt

$$\mathcal{M}_{B_V^*}^{B_V}(\gamma) = \begin{bmatrix} I_{n_+} & & \\ & -I_{n_-} & \\ & & 0_{n_0} \end{bmatrix} \tag{33.6}$$

hat. Dabei ist die **Signatur** (n_+, n_-, n_0) eindeutig bestimmt, und wir bezeichnen sie mit $\operatorname{signature}(\gamma) = (n_+, n_-, n_0) \in \mathbb{N}_0^3$. Es gilt $\operatorname{Rang}(\gamma) = n_+ + n_-$.

Beweis. Wir beginnen mit einer Darstellungsmatrix der Gestalt (33.5), wobei $\alpha_1, \dots, \alpha_r \in \mathbb{R} \setminus \{0\}$ sind mit $r = \operatorname{Rang}(\gamma)$.

Schritt 1: Wir zeigen, dass die Form (33.6) erreichbar ist.

Wir nutzen dazu die Eigenschaft, dass $\mathbb{R}$ mit der üblichen Ordnung ein geordneter Körper ist (Definition 23.23).[5] Wir betrachten zunächst die Diagonaleinträge mit $\alpha_i > 0$. O. B. d. A. sind dies die führenden $n_+ \in \mathbb{N}_0$ Diagonaleinträge. Es sei nun $\beta_i > 0$ die (positive) Wurzel von α_i, sodass also $\alpha_i = \beta_i^2$ gilt. Durch Skalierung der bisherigen Basisvektoren v_i mit $\frac{1}{\beta_i} v_i$ für $i = 1, \dots, n_+$ gilt für die neuen Diagonaleinträge der Darstellungsmatrix $\gamma(v_i, v_i) = +1$ für $i = 1, \dots, n_+$.

Die folgenden $n_- \in \mathbb{N}_0$ Diagonaleinträge sind diejenigen mit $\alpha_i < 0$. Wir setzen dann $\beta_i > 0$ die (positive) Wurzel von $-\alpha_i > 0$, sodass also $-\alpha_i = \beta_i^2$ gilt. Wiederum durch Skalierung der bisherigen Basisvektoren v_i mit $\frac{1}{\beta_i} v_i$ für $i = n_+ + 1, \dots, n_+ + n_-$ gilt für die neuen Diagonaleinträge der Darstellungsmatrix $\gamma(v_i, v_i) = -1$.

Schritt 2: Es bleibt zu zeigen, dass die Signatur eindeutig ist.

Dazu bemerken wir zunächst, dass die Darstellung (33.6) uns eine orthogonale Zerlegung $V = V_+ \oplus V_- \oplus V_0$ gibt, wobei mit $q(v) = \gamma(v, v)$

$$\begin{aligned} q(v) &> 0 \quad \text{für alle } v \in V_+ \setminus \{0\} \\ q(v) &< 0 \quad \text{für alle } v \in V_- \setminus \{0\} \\ q(v) &= 0 \quad \text{für alle } v \in V_0 \end{aligned}$$

gilt.

Schritt 3: Wir zeigen: Ist $U \subseteq V$ ein Unterraum mit der Eigenschaft $q(u) > 0$ für alle $u \in U \setminus \{0\}$, dann folgt $\dim(U) \leqslant \dim(V_+)$.

Dazu nehmen wir im Gegenteil an, es gelte $\dim(U) > \dim(V_+) = \dim(V) - \dim(V_- \oplus V_0)$, also auch

$$\dim(U) + \dim(V_- \oplus V_0) > \dim(V).$$

[5] Diese ist auch die einzige, mit der $\mathbb{R}$ zu einem geordneten Körper wird.

Das heißt aber nach Satz 14.7, dass die Summe von U und $V_- \oplus V_0$ nicht direkt sein kann. Es gibt also einen Vektor $u \in U \cap (V_- \oplus V_0)$, $u \neq 0$. Dieser Vektor hat eine eindeutige Zerlegung $u = u_- + u_0$ mit $u_- \in V_-$ und $u_0 \in V_0$. Daher gilt

$$q(u) = \gamma(u_- + u_0, u_- + u_0) = \gamma(u_-, u_-) + \gamma(u_0, u_0) = \gamma(u_-, u_-) \leqslant 0.$$

Das ist im Widerspruch zur Annahme $q(u) > 0$ für alle $u \in U \setminus \{0\}$.

Schritt 4: Wir betrachten nun zwei Zerlegungen $V = V_+ \oplus V_- \oplus V_0$ und $V = \widehat{V}_+ \oplus \widehat{V}_- \oplus \widehat{V}_0$, wobei

$$\begin{aligned} q(v) &> 0 \quad \text{für alle } v \in V_+ \setminus \{0\} \text{ und alle } v \in \widehat{V}_+ \setminus \{0\} \\ q(v) &< 0 \quad \text{für alle } v \in V_- \setminus \{0\} \text{ und alle } v \in \widehat{V}_- \setminus \{0\} \\ q(v) &= 0 \quad \text{für alle } v \in V_0 \text{ und alle } v \in \widehat{V}_0 \end{aligned}$$

gilt. Dann ist nach Definition $V_0 = \widehat{V}_0$. Wegen des Ergebnisses von **Schritt** 3 gilt $\dim(\widehat{V}_+) \leqslant \dim(V_+)$, mit vertauschten Rollen aber auch $\dim(V_+) \leqslant \dim(\widehat{V}_+)$, also $\dim(V_+) = \dim(\widehat{V}_+)$.

Damit sind die Dimensionen n_0 und n_+ eindeutig bestimmt, also auch n_-. □

Kurz gesagt bedeutet Satz 33.13: Ein endlich-dimensionaler quadratischer Raum (V, γ) über $\mathbb{R}$ ist durch signature(γ) bereits bis auf Isomorphie eindeutig bestimmt. Aus signature(γ) können wir auch $\dim(V) = n_+ + n_- + n_0$ sowie $\operatorname{Rang}(\gamma) = n_+ + n_-$ ablesen.

Wir formulieren das Ergebnis von Satz 33.13 auch noch einmal in einer Form ähnlich zu Folgerung 33.12:

Satz 33.14 (**Trägheitssatz von Sylvester**[6]).

Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, $n \in \mathbb{N}_0$. Dann gibt es eine reguläre Matrix $T \in \mathbb{R}^{n \times n}$, sodass $T^{-\intercal} A T^{-1}$ eine Diagonalmatrix der Gestalt (33.6) ist. Die Anzahl der Diagonaleinträge, die gleich +1, −1 oder 0 sind, also die Signatur der durch A induzierten Bilinearform γ_A, ist durch A eindeutig festgelegt.

Bemerkung 33.15 (zur Formulierung des Trägheitssatzes von Sylvester).

In der Literatur wird der Satz 33.14 häufig so formuliert, dass Kongruenztransformationen das Vorzeichen der Eigenwerte symmetrischer Matrizen über $\mathbb{R}$ erhalten. Dieser Sichtweise schließen wir uns nicht an, weil Eigenwerte für Darstellungsmatrizen von Bilinearformen keine Funktion haben, vgl. Bemerkung 31.7. △

Abschließend bemerken wir, dass wir die Ergebnisse für den Fall $K = \mathbb{R}$ nicht auf die rationalen Zahlen $K = \mathbb{Q}$ übertragen können, da das Instrument des Wurzelziehens nicht zur Verfügung steht.

[6] englisch: Sylvester's law of inertia

§ 33.3 Normalformen symmetrischer Bilinearformen in komplexen Vektorräumen

Auch über dem Körper $K = \mathbb{C}$ der komplexen Zahlen können wir den Satz 33.10 verschärfen:

Satz 33.16 (jeder endlich-dimensionale quadratische Raum über $K = \mathbb{C}$ besitzt eine Orthogonalbasis, vgl. Satz 33.10)**.**
Es sei (V, γ) ein endlich-dimensionaler quadratischer Raum über $\mathbb{C}$. Dann gilt: V besitzt eine Orthogonalbasis B_V, bzgl. der die Darstellungsmatrix die Gestalt

$$\mathcal{M}_{B_V^*}^{B_V}(\gamma) = \begin{bmatrix} I_r & \\ & 0 \end{bmatrix} \tag{33.7}$$

hat, wobei gleich $r = \operatorname{Rang}(\gamma)$ ist.

Beweis. Wir beginnen mit einer Darstellungsmatrix der Gestalt (33.5), wobei $\alpha_1, \ldots, \alpha_r \in \mathbb{C} \setminus \{0\}$ sind mit $r = \operatorname{Rang}(\gamma)$. Es sei nun $\beta_i \in \mathbb{C} \setminus \{0\}$ eine der beiden komplexen Wurzeln von α_i, sodass also $\alpha_i = \beta_i^2$ gilt. Indem wir die bisherigen Basisvektoren $v_1, \ldots, v_r$ durch $\frac{1}{\beta_1} v_1, \ldots, \frac{1}{\beta_r} v_r$ ersetzen, gilt für die neuen Diagonaleinträge der Darstellungsmatrix

$$\gamma\Big(\frac{1}{\beta_i} v_i, \frac{1}{\beta_i} v_i\Big) = \frac{1}{\beta_i^2}\, \gamma(v_i, v_i) = \frac{1}{\beta_i^2}\, \alpha_i = 1. \qquad \square$$

Kurz gesagt bedeutet Satz 33.16: Ein quadratischer Raum (V, γ) über $\mathbb{C}$ ist durch $\dim(V)$ und $\operatorname{Rang}(\gamma)$ bereits bis auf Isomorphie eindeutig bestimmt.

§ 34 Innenprodukte in reellen Vektorräumen

In diesem Abschnitt betrachten wir **reelle Vektorräume** (englisch: real vector spaces), also solche über dem Körper $K = \mathbb{R}$. Wir nutzen dabei u. a. aus, dass $\mathbb{R}$ mit der üblichen Totalordnung ein geordneter Körper ist (Definition 23.23), um weitere Eigenschaften von Bilinearformen fordern und untersuchen zu können.

Definition 34.1 (Definitheit und Indefinitheit von Bilinearformen in reellen Vektorräumen)**.**
Es sei V ein Vektorraum über $\mathbb{R}$. Eine (nicht notwendig symmetrische) Bilinearform γ auf V heißt

(*i*) **positiv definit** (englisch: positive definite), wenn $\gamma(v, v) > 0$ gilt für alle $v \in V \setminus \{0\}$.

(*ii*) **positiv semidefinit** (englisch: positive semi-definite), wenn $\gamma(v, v) \geqslant 0$ gilt für alle $v \in V$.

(*iii*) **negativ definit** (englisch: negative definite), wenn $\gamma(v, v) < 0$ gilt für alle $v \in V \setminus \{0\}$.

(*iv*) **negativ semidefinit** (englisch: negative semi-definite), wenn $\gamma(v, v) \leqslant 0$ gilt für alle $v \in V$.

(*v*) **indefinit** (englisch: indefinite), wenn γ weder positiv semidefinit noch negativ semidefinit ist, wenn es also Vektoren $v_1, v_2 \in V$ gibt mit $\gamma(v_1, v_1) > 0$ und $\gamma(v_2, v_2) < 0$. △

Ist V endlich-dimensional, B_V eine beliebige Basis von V und ist $A = \mathcal{M}_{B_V^*}^{B_V}(\gamma)$ die Darstellungsmatrix von γ bzgl. dieser Basis, dann können wir die Eigenschaften aus der Definition 34.1 auch mit Hilfe von A ausdrücken. Beispielsweise heißt die Darstellungsmatrix $A \in \mathbb{R}^{n \times n}$ **positiv definit**, wenn $x^\intercal A\, x > 0$ gilt für alle $x \in \mathbb{R}^n \setminus \{0\}$. Das ist wegen des bijektiven Zusammenhangs $x = \Phi_{B_V}^{-1}(v)$ genau dann der Fall, wenn γ positiv definit ist. Eine analoge Aussage gilt für **positiv semidefinite**, **negativ definite**, **negativ semidefinite** und **indefinite** Matrizen.

Bemerkung 34.2 (Definitheit einer Bilinearform und Diagonaleinträge von Darstellungsmatrizen)**.**
Die positive bzw. negative Definitheit einer Bilinearform γ auf einem endlich-dimensionalen Vektorraum über $\mathbb{R}$ ist äquivalent dazu, dass *jede* Darstellungsmatrix nur positive bzw. nur negative Diagonaleinträge besitzt. Ist γ weder positiv noch negativ definit, dann gibt es eine Basis, bzgl. der die Darstellungsmatrix eine Null auf der Diagonale hat, selbst wenn γ nicht ausgeartet ist. △

Definition 34.3 (Innenprodukt in einem reellen Vektorraum, Euklidischer Raum)**.**
Es sei (V, γ) ein quadratischer Raum über $\mathbb{R}$. Die symmetrische Bilinearform γ heißt ein **Innenprodukt** (englisch: inner product) **auf** V, wenn γ positiv definit ist.[7] In diesem Fall heißt (V, γ) auch ein **reeller Innenproduktraum** (englisch: real inner product space) oder **Euklidischer Raum** (englisch: Euclidean space). △

Beachte: Innenprodukte sind genau diejenigen symmetrischen Bilinearformen mit Signatur $(n, 0, 0)$.

Bemerkung 34.4 (Bezeichnung von Innenprodukten)**.**
Manche Autoren sprechen statt von Innenprodukten auch von **Skalarprodukten** (englisch: scalar products). Wir vermeiden diesen Begriff, um Verwechslungen mit der S-Multiplikation (Definition 12.1) auszuschließen. △

Beispiel 34.5 (Innenprodukt in einem reellen Vektorraum)**.**

(i) Die durch die Einheitsmatrix $I \in \mathbb{R}^{n \times n}$ induzierte Bilinearform

$$\gamma_I \colon \mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto x^\intercal y = \sum_{i=1}^n x_i\, y_i \in \mathbb{R}$$

heißt das **Standardinnenprodukt** (englisch: standard inner product, dot product) auf $\mathbb{R}^n$.

(ii) **Beachte:** Um die positive Definitheit einer symmetrischen Bilinearform nachzuweisen, genügt es nicht, nur die Werte auf Basisvektoren anzusehen: Betrachten wir beispielsweise

$$A = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

und die durch A induzierte Bilinearform γ_A, dann gilt $\gamma(e_1, e_1) = \gamma(e_2, e_2) = 1$, aber

$$\gamma\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}^\intercal \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -2 < 0.$$ △

Das folgende Resultat wurde nachträglich hinzugefügt.

Lemma 34.6 (positiv definite Bilinearformen und Matrizen sind invertierbar)**.**

[7] $\gamma \colon V \times V \to \mathbb{R}$ ist also bilinear, erfüllt $\gamma(u, v) = \gamma(v, u)$ und $\gamma(v, v) > 0$ für alle $v \in V \setminus \{0\}$.

(*i*) Es sei V ein endlich-dimensionaler Vektorraum über $\mathbb{R}$ und γ eine (nicht notwendig symmetrische) positiv definite Bilinearform auf V. Dann ist die Abbildung

$$\Gamma\colon V \ni u \mapsto \gamma(\cdot, u) \in V^* \tag{34.1}$$

bijektiv. Für jedes $\xi \in V^*$ existiert also eine eindeutige Lösung des „linearen Gleichungssystems" $\Gamma(u) = \xi$, also

$$\gamma(v, u) = \langle \xi \,, v \rangle \quad \text{für alle } v \in V.$$

Die **inverse Bilinearform** (englisch: inverse bilinear form) $\gamma^{-1} \in \operatorname{Bil}(V^*, V^*)$, definiert durch $\gamma^{-1}(\xi, \eta) := \langle \eta \,, \Gamma^{-1}\xi \rangle$, ist wiederum positiv definit.

(*ii*) Es sei $A \in \mathbb{R}^{n \times n}$ für $n \in \mathbb{N}_0$ eine (nicht notwendig symmetrische) positiv definite Matrix. Dann ist A invertierbar. Die inverse Matrix ist wiederum positiv definit.

Beweis. Aussage (*i*): Wir nehmen an, die lineare Abbildung Γ sei nicht injektiv. Dann existiert $u \in V$, $u \neq 0$, mit $\gamma(\cdot, u) = 0$ (die Nullform auf V). Es folgt $\gamma(u, u) = 0$, also ist γ nicht positiv definit. Γ muss also injektiv sein, und wegen der endlichen Dimension auch bijektiv (Folgerung 18.9).

Wir jetzt $\Gamma^{-1} \in \operatorname{Hom}(V^*, V)$ und $\xi \in V^*$. Ist $\xi \neq 0$, dann ist auch $u := \Gamma^{-1}\xi \in V \neq 0$. Unter dieser Voraussetzung gilt

$$\gamma^{-1}(\xi, \xi) = \langle \xi \,, \Gamma^{-1}\xi \rangle = \langle \xi \,, u \rangle = \gamma(u, u) > 0.$$

Also ist auch γ^{-1} positiv definit.

Aussage (*ii*): Wir nehmen an, A ist nicht invertierbar. Dann exististiert $x \in \mathbb{R}^n$, $x \neq 0$, mit $A\,x = 0$. Es folgt $x^\intercal A\,x = 0$, also ist A nicht positiv definit.

Wir betrachten jetzt $\xi \in (\mathbb{R}^n)^* \cong \mathbb{R}^n$. Ist $\xi \neq 0$, dann ist auch $x := A^{-1}\xi \neq 0$. Unter dieser Voraussetzung gilt

$$\xi^\intercal A^{-1}\xi = x^\intercal A^\intercal A^{-1} A\,x = x^\intercal A^\intercal x = x^\intercal A x > 0.$$

Also ist auch A^{-1} positiv definit. □

Für allgemeine symmetrische Bilinearformen γ sind die Konzepte der linearen Unabhängigkeit von Vektoren und deren Orthogonalität noch voneinander unabhängig, selbst wenn γ nicht ausgeartet ist. Für Innenprodukte jedoch besteht folgendes Resultat:

Lemma 34.7 (Orthogonalität bzgl. eines Innenprodukts impliziert lineare Unabhängigkeit)**.**
Es sei (V, γ) ein Euklidischer Raum. Ist $(v_1, \ldots, v_k)$ eine orthogonale Familie von Vektoren in $V \setminus \{0\}$, $k \in \mathbb{N}_0$, dann ist $(v_1, \ldots, v_k)$ linear unabhängig.

Beweis. Wir machen den Ansatz $\sum_{j=1}^k \alpha_j\, v_j = 0$. Dann gilt auch

$$0 = \gamma\Big(v_i, \sum_{j=1}^k \alpha_j\, v_j\Big) = \sum_{j=1}^k \alpha_j\, \gamma(v_i, v_j) = \alpha_i\, \gamma(v_i, v_i)$$

für alle $i = 1, \ldots, k$. Da γ ein Innenprodukt und $v_i \neq 0$ ist, gilt $\gamma(v_i, v_i) > 0$, daher folgt $\alpha_i = 0$. □

§ 34.1 Innenprodukte induzieren Normen

Eine weitere besondere Eigenschaft jedes Innenprodukts auf einem reellen Vektorraum ist, dass es ein Norm induziert. Um das zu beweisen, benötigen wir zunächst das folgende wichtige Resultat:

Satz 34.8 (Cauchy-Schwarz-Ungleichung[8]).
Es sei (V, γ) ein Euklidischer Raum. Dann gilt:

$$\gamma(u,v)^2 \leqslant \gamma(u,u)\,\gamma(v,v) \quad \text{oder äquivalent} \quad |\gamma(u,v)| \leqslant \sqrt{\gamma(u,u)}\,\sqrt{\gamma(v,v)} \tag{34.2}$$

für alle $u, v \in V$. Gleichheit gilt in (34.2) genau dann, wenn (u, v) eine linear abhängige Familie ist.

Beweis. Für $v = 0$ ist (34.2) mit Gleichheit erfüllt, und (u, v) ist linear abhängig.

Für den Rest des Beweises sei nun also $v \neq 0$. Für $\beta \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leqslant \gamma(u - \beta\,v, u - \beta\,v) && \text{wegen der positiven Definitheit von } \gamma \\ &= \gamma(u,u) - 2\,\beta\,\gamma(u,v) + \beta^2\,\gamma(v,v) && \text{wegen der Bilinearität und Symmetrie von } \gamma. \end{aligned}$$

Hierbei ist $\gamma(v, v) > 0$ wegen der positiven Definitheit, und wir setzen $\beta := \frac{\gamma(u,v)}{\gamma(v,v)}$. Es folgt

$$\begin{aligned} 0 &\leqslant \gamma(u,u) - 2\,\frac{\gamma(u,v)}{\gamma(v,v)}\,\gamma(u,v) + \frac{\gamma(u,v)^2}{\gamma(v,v)} \\ &= \gamma(u,u) - \frac{\gamma(u,v)^2}{\gamma(v,v)}. \end{aligned}$$

Die Multiplikation mit $\gamma(v, v) > 0$ ergibt (34.2).

Wir müssen noch untersuchen, wann die Ungleichung (34.2) mit Gleichheit erfüllt ist und können dabei weiterhin von $v \neq 0$ ausgehen. Ist (u, v) linear abhängig, so haben wir also $u = \delta\,v$ für ein $\delta \in \mathbb{R}$. Aus der Bilinearität folgt $\gamma(u, u) = \delta^2\,\gamma(v, v)$ sowie

$$\gamma(u,v)^2 = \big(\delta\,\gamma(v,v)\big)^2 = \delta^2\,\gamma(v,v)^2 = \gamma(u,u)\,\gamma(v,v),$$

also Gleichheit in (34.2).

Umgekehrt gelte $\gamma(u, v)^2 = \gamma(u, u)\,\gamma(v, v)$. Dann haben wir mit $\beta := \frac{\gamma(u,v)}{\gamma(v,v)}$ und denselben Umformungen wie oben

$$\begin{aligned} 0 &= \gamma(u,u) - \frac{\gamma(u,v)^2}{\gamma(v,v)} \\ &= \gamma(u,u) - 2\,\beta\,\gamma(u,v) + \beta^2\,\gamma(v,v) \\ &= \gamma(u - \beta\,v, u - \beta\,v). \end{aligned}$$

Aufgrund der positiven Definitheit folgt $u - \beta\,v = 0$, also ist (u, v) linear abhängig. □

Bemerkung 34.9 (Cauchy-Schwarz-Ungleichung für semidefinite symmetrische Bilinearformen)**.**
Der Beweis zeigt, dass (34.2) auch dann richtig bleibt, wenn γ nur als positiv semidefinit angenommen wird. Wir müssen lediglich, nachdem $0 \leqslant \gamma(u, u) - 2\,\beta\,\gamma(u, v) + \beta^2\,\gamma(v, v)$ gezeigt wurde, den Fall $\gamma(v, v) = 0$ gesondert behandeln. Daraus folgt nämlich auch $\gamma(u, v) = 0$, denn andernfalls liefern die Wahlen $\beta := \frac{1}{2}\frac{\gamma(u,u\pm v)}{\gamma(u,v)}$ den Widerspruch $0 \leqslant \mp\gamma(u, v)$. Mit $\gamma(v, v) = \gamma(u, v) = 0$ sehen wir aber, dass (34.2) ebenfalls gilt. △

[8] englisch: Cauchy-Schwarz inequality

Definition 34.10 (Norm auf einem reellen Vektorraum).
Es sei V ein Vektorraum über $\mathbb{R}$.

(i) Eine Abbildung $\|\cdot\|\colon V \to \mathbb{R}$ heißt eine **Norm** (englisch: norm) **auf** V, wenn gilt:

$$\|u\| \geqslant 0 \quad \text{und} \quad \|u\| = 0 \;\Rightarrow\; u = 0 \qquad \textbf{positive Definitheit}^{9} \tag{34.3a}$$

$$\|\alpha\, u\| = |\alpha|\, \|u\| \qquad \textbf{absolute Homogenität}^{10} \tag{34.3b}$$

$$\|u + v\| \leqslant \|u\| + \|v\| \qquad \textbf{Dreiecksungleichung} \text{ oder } \textbf{Subadditivität}^{11} \tag{34.3c}$$

für alle $u, v \in V$ und alle $\alpha \in \mathbb{R}$.

(ii) Das Paar $(V, \|\cdot\|)$ heißt ein **normierter reeller Vektorraum** (englisch: normed real vector space). △

Das folgende Resultat nutzt das Wurzelziehen in $\mathbb{R}$, also die Existenz einer eindeutigen Lösung $x = \sqrt{\alpha} \geqslant 0$ quadratischer Gleichungen der Form $x^2 = \alpha$ mit $\alpha \geqslant 0$.

Satz 34.11 (jedes Innenprodukt induziert eine Norm).
Es sei (V, γ) ein Euklidischer Raum. Dann definiert

$$\|\cdot\|_\gamma \colon V \ni u \mapsto \|u\|_\gamma \coloneqq \sqrt{\gamma(u,u)} \in \mathbb{R}_{\geqslant 0} \tag{34.4}$$

eine Norm auf V.

Beweis. Die Behauptung $\|u\|_\gamma \geqslant 0$ folgt aus der Eigenschaft der Wurzelfunktion $\sqrt{\cdot}\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, wie auch die Behauptung $\|u\|_\gamma = 0 \Rightarrow u = 0$. Die positive Homogenität ist eine Konsequenz aus der Bedingung (32.1a) an die quadratische Form $u \mapsto \gamma(u,u)$:

$$\|\alpha\, u\|_\gamma^2 = \gamma(\alpha\, u, \alpha\, u) = \alpha^2\, \gamma(u,u).$$

Durch Wurzelziehen folgt

$$\|\alpha\, u\|_\gamma = \sqrt{\alpha^2}\, \sqrt{\gamma(u,u)} = |\alpha|\, \|u\|_\gamma.$$

Schließlich zeigen wir die Dreiecksungleichung. Es ist

$$\begin{aligned}
\|u+v\|_\gamma^2 &= \gamma(u+v, u+v) && \text{nach Definition der Norm } \|\cdot\|_\gamma \text{ nach (34.4)}\\
&= \gamma(u,u) + 2\,\gamma(u,v) + \gamma(v,v) && \text{wegen der Bilinearität und Symmetrie von } \gamma\\
&= \|u\|_\gamma^2 + 2\,\gamma(u,v) + \|v\|_\gamma^2 && \text{nach Definition der Norm } \|\cdot\|_\gamma \text{ nach (34.4)}\\
&\leqslant \|u\|_\gamma^2 + 2\,\|u\|_\gamma \|v\|_\gamma + \|v\|_\gamma^2 && \text{wegen der Cauchy-Schwarz-Ungleichung (34.2)}\\
&= \big(\|u\|_\gamma + \|v\|_\gamma\big)^2.
\end{aligned}$$

Durch Wurzelziehen folgt die Behauptung $\|u+v\|_\gamma \leqslant \|u\|_\gamma + \|v\|_\gamma$. □

Die Norm $\|u\|$ eines Vektors u können wir als ein Konzept von „Länge“ interpretieren. Die Norm $\|u - v\|$ einer Differenz von Vektoren können wir als den „Abstand“ zwischen u und v verstehen. Wir betrachten in diesem Abschnitt nur Normen, die von Innenprodukten induziert werden. Wenn das Innenprodukt klar ist, schreiben wir auch einfach $\|\cdot\|$ statt $\|\cdot\|_\gamma$.

[9] englisch: positive definiteness
[10] englisch: absolute homogeneity
[11] englisch: triangle inequality, subadditivity

Satz 34.12 (**Satz des Pythagoras**).
Es sei (V, γ) ein Euklidischer Raum. Ist die Familie $(v_1, \ldots, v_k)$ orthogonal, dann gilt

$$\Big\|\sum_{i=1}^k v_i\Big\|^2 = \sum_{i=1}^k \|v_i\|^2. \tag{34.5}$$

Beweis. Wir nutzen den Satz des Pythagoras 33.3. Es gilt

$$\Big\|\sum_{i=1}^k v_i\Big\|^2 = \gamma\Big(\sum_{i=1}^k v_i, \sum_{j=1}^k v_j\Big) = \sum_{i=1}^k \gamma(v_i, v_i) = \sum_{i=1}^k \|v_i\|^2. \qquad \square$$

Definition 34.13 (normierter Vektor, orthonormale Familie, Orthonormalbasis).
Es sei (V, γ) ein Euklidischer Raum mit induzierter Norm $\|\cdot\|$.

(*i*) Ein Vektor $u \in V$ heißt ein **normierter Vektor** oder **Einheitsvektor** (englisch: unit vector), wenn $\|u\| = 1$ gilt.

(*ii*) Eine Menge $E \subseteq V$ von Vektoren in V heißt **orthonormal** (englisch: orthonormal set) **bzgl.** γ, wenn die Menge orthogonal ist und alle Elemente normiert sind.

(*iii*) Eine Familie $(v_i)_{i \in I}$ von Vektoren in V heißt **orthonormal** (englisch: orthonormal family) **bzgl.** γ, wenn die Familie orthogonal ist und alle ihre Mitglieder normiert sind.

(*iv*) Eine Familie $B := (v_i)_{i \in I}$ von Vektoren in V heißt eine **Orthonormalbasis** (englisch: orthonormal basis) von (V, γ), wenn B eine Basis von V und außerdem orthonormal bzgl. γ ist. △

Beachte: Jeder Vektor $u \in V \setminus \{0\}$ wird durch Übergang zu $\frac{u}{\|u\|} = \frac{1}{\|u\|}\, u$ zu einem normierten Vektor. Wir können also aus einer Orthogonalbasis leicht eine Orthonormalbasis machen, indem wir jeden Basisvektor normieren.

Damit folgt:

Satz 34.14 (jeder endlich-dimensionale Euklidische Raum besitzt eine Orthonormalbasis, vgl. Satz 33.13).

Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Dann gilt: V besitzt eine Orthonormalbasis B_V. Die Darstellungsmatrix erfüllt daher $\mathcal{M}_{B_V^*}^{B_V}(\gamma) = I$.

§ 34.2 Orthogonale Projektion und Gram-Schmidt-Verfahren

Es stellt sich die Frage, wie wir eine Orthogonalbasis bzw. Orthonormalbasis algorithmisch bestimmen können. Wir geben dazu in diesem Abschnitt das **Gram-Schmidt-Verfahren** an, mit dem wir eine gegebene Familie linear unabhängiger Vektoren in eine orthogonale bzw. orthonormale Familie umwandeln können. Grundlage des Verfahrens ist die orthogonale Projektion:

Definition 34.15 (orthogonale Projektion auf einen Vektor).
Es sei (V, γ) ein Euklidischer Raum und $u \in V$, $u \neq 0$. Dann ist durch

$$\operatorname{proj}_u^\gamma \colon V \ni v \mapsto \operatorname{proj}_u^\gamma(v) := \frac{\gamma(v, u)}{\|u\|^2}\, u \in V \tag{34.6}$$

die **orthogonale Projektion** (englisch: orthogonal projection) **bzgl.** γ **auf** $\langle u \rangle$ definiert.[12] △

[12]Etwas ungenau spricht man auch von der „Projektion auf u" statt von der Projektion auf den Unterraum $\langle u \rangle$.

Lemma 34.16 (Eigenschaften der orthogonalen Projektion auf einen Vektor).
Es sei (V, γ) ein Euklidischer Raum und $u \in V$, $u \neq 0$. Dann gilt:

(i) Gleichung (34.6) definiert eine Projektion (Definition 22.33).

(ii) Es gilt $V = \operatorname{Bild}(\operatorname{proj}_u^\gamma) \oplus \operatorname{Kern}(\operatorname{proj}_u^\gamma)$. Jedes $v \in V$ besitzt also eine eindeutige Zerlegung $v = u_\parallel + u_\perp$ mit $u_\parallel \in \langle u \rangle$ und $u_\perp \in \langle u \rangle^\perp$.

Beweis. Aussage (i): Zunächst stellen wir fest, dass $\operatorname{proj}_u^\gamma$ eine lineare Abbildung, also ein Endomorphismus von V ist. Die Idempotenz rechnen wir wie folgt nach:

$$\begin{aligned}
\operatorname{proj}_u^\gamma(\operatorname{proj}_u^\gamma(v)) &= \operatorname{proj}_u^\gamma\Big(\frac{\gamma(v,u)}{\|u\|^2}\, u\Big) \\
&= \frac{\gamma(v,u)}{\|u\|^2} \operatorname{proj}_u^\gamma(u) \\
&= \frac{\gamma(v,u)}{\|u\|^2} \frac{\|u\|^2}{\|u\|^2}\, u \\
&= \frac{\gamma(v,u)}{\|u\|^2}\, u \\
&= \operatorname{proj}_u^\gamma(v).
\end{aligned}$$

Aussage (ii): Da $\operatorname{proj}_u^\gamma(v)$ per Definition ein Vielfaches von u ist, gilt $\operatorname{Bild}(\operatorname{proj}_u^\gamma) \subseteq \langle u \rangle$. Wegen $\operatorname{proj}_u^\gamma(u) = u$ und der Linearität gilt auch $\operatorname{proj}_u^\gamma(\alpha\, u) = \alpha\, u$. Also folgt $\operatorname{Bild}(\operatorname{proj}_u^\gamma) = \langle u \rangle$. Aus Lemma 24.10 ergibt sich $V = \operatorname{Bild}(\operatorname{proj}_u^\gamma) \oplus \operatorname{Kern}(\operatorname{proj}_u^\gamma)$. Es bleibt zu zeigen, dass $\operatorname{Kern}(\operatorname{proj}_u^\gamma) = \operatorname{Bild}(\operatorname{proj}_u^\gamma)^\perp$ ist. Aus $v \in \operatorname{Kern}(\operatorname{proj}_u^\gamma)$ folgt aber sofort $\gamma(u, v) = 0$, also $v \perp u$ und damit $v \perp \langle u \rangle = \operatorname{Bild}(\operatorname{proj}_u^\gamma)$. □

Anschaulich bestimmt die γ-orthogonale Projektion $\operatorname{proj}_u^\gamma(v)$ den Anteil des Vektors v im Unterraum $\langle u \rangle$. (Dieser hängt von der Bilinearform γ ab!)

Die Idee des **Gram-Schmidt-Verfahrens** (englisch: Gram-Schmidt process) besteht darin, aus den Vektoren einer linear unabhängigen Familie $(v_1, \ldots, v_k)$ sukzessive die Anteile vorangegangener Vektoren „herauszurechnen". Dadurch wird die gegebene Familie in eine orthogonale Familie $(u_1, \ldots, u_k)$ umgewandelt, und es gilt $\langle v_1, \ldots, v_k \rangle = \langle u_1, \ldots, u_j \rangle$ für alle $j = 1, \ldots, k$. Das Verfahren wird auch **Gram-Schmidt-Orthogonalisierung** (englisch: Gram-Schmidt orthogonalization) genannt.

Algorithmus 34.17 (Gram-Schmidt-Orthogonalisierung).
Eingabe: V ein Euklidischer Vektorraum
Eingabe: γ ein Innenprodukt auf V
Eingabe: linear unabhängige Familie $(v_1, \ldots, v_k)$ in V
Ausgabe: orthogonale Familie $(u_1, \ldots, u_k)$ in V mit $\langle u_1, \ldots, u_k \rangle = \langle v_1, \ldots, v_k \rangle$

for $j = 1, \ldots, k$ **do**
  Setze $u_j \leftarrow v_j - \sum_{i=1}^{j-1} \frac{\gamma(v_j, u_i)}{\alpha_i} u_i$ // $u_j = v_j - \sum_{i=1}^{j-1} \operatorname{proj}_{u_i}^\gamma(v_j)$
  Bestimme $\alpha_j := \|u_j\|^2$ // $\alpha_j > 0$
end for
return $(u_1, \ldots, u_k)$

Durch eine zusätzliche Normierung der entstehenden Vektoren erhalten wir sogar eine orthonormale Familie:

Algorithmus 34.18 (Gram-Schmidt-Orthonormalisierung).

Eingabe: V ein Euklidischer Vektorraum
Eingabe: γ ein Innenprodukt auf V
Eingabe: linear unabhängige Familie $(v_1, \ldots, v_k)$ in V
Ausgabe: orthonormale Familie $(u_1, \ldots, u_k)$ in V mit $\langle u_1, \ldots, u_k \rangle = \langle v_1, \ldots, v_k \rangle$

for $j = 1, \ldots, k$ **do**
  Setze $u_j \leftarrow v_j - \sum_{i=1}^{j-1} \gamma(v_j, u_i)\, u_i$ // $u_j = v_j - \sum_{i=1}^{j-1} \operatorname{proj}_{u_i}^{\gamma}(v_j)$
  Bestimme $\alpha_j := \|u_j\|^2$ // $\alpha_j > 0$
  Setze $u_j \leftarrow \dfrac{u_j}{\sqrt{\alpha_j}}$ // $u_1, \ldots, u_j$ erfüllen jetzt $\|u_i\| = 1$
end for
return $(u_1, \ldots, u_k)$

Satz 34.19 (Eigenschaften des Gram-Schmidt-Verfahrens).
Es sei (V, γ) ein Euklidischer Raum. Ist $(u_1, \ldots, u_k)$ eine linear unabhängige Familie in V, dann liefert Algorithmus 34.17 eine orthogonale Familie $(v_1, \ldots, v_k)$ von V mit der Eigenschaft $\langle v_1, \ldots, v_j \rangle = \langle u_1, \ldots, u_j \rangle$ für alle $j = 1, \ldots, k$. Dasselbe gilt für Algorithmus 34.18, wobei die Familie $(v_1, \ldots, v_k)$ sogar orthonormal ist.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-11.2. □

Ende der Vorlesung 22

Ende der Woche 11

§ 34.3 Orthogonale Abbildungen und die orthogonale Gruppe

In diesem Abschnitt betrachten wir die Homomorphismen Euklidischer Räume, also (siehe Definition 33.8) lineare Abbildungen zwischen zwei Euklidischen Vektorräumen, die mit den Innenprodukten verträglich sind. So wie die Homomorphismen von Vektorräumen auch als „lineare Abbildungen" bezeichnet werden, so heißen die Homomorphismen Euklidischer Räume **orthogonale Abbildungen**.

Definition 34.20 (Homomorphismus Euklidischer Räume, vgl. Definition 33.8).
Es seien (V, γ_1) und (W, γ_2) zwei Euklidische Räume. Eine Abbildung $f \colon V \to W$ heißt **orthogonal** (englisch: orthogonal map) oder eine **(lineare) Isometrie** (englisch: linear isometry, altgriechisch: *ἴσος*: gleich, altgriechisch: *μετρικός*: bezogen auf die Messung) **bzgl.** (γ_1, γ_2), wenn f ein Homormophismus der quadratischen Räume $(V, \gamma_1) \to (W, \gamma_2)$ ist, wenn also gilt:

$$f \text{ ist linear} \tag{34.7a}$$
$$\gamma_2(f(u), f(v)) = \gamma_1(u, v) \quad \text{für alle } u, v \in V. \tag{34.7b}$$

Statt **orthogonal bzgl.** (γ_1, γ_2) sagen wir auch kurz (γ_1, γ_2)-**orthogonal**. △

Die Orthogonalität einer linearen Abbildung zwischen Euklidischen Vektorräumen ist eine starke Eigenschaft, die auf viele verschiedene Arten charakterisiert werden kann:

Satz 34.21 (Charakterisierung orthogonaler Abbildungen)**.**
Es seien (V, γ_1) und (W, γ_2) zwei Euklidische Räume. Weiter sei $f \in \operatorname{Hom}(V, W)$. Dann sind äquivalent:

(i) f ist (γ_1, γ_2)-orthogonal.

(ii) $\|f(v)\|_{\gamma_2} = \|v\|_{\gamma_1}$ für alle $v \in V$, d. h., f ist **längentreu** (englisch: length preserving).

(iii) $\|f(v_1) - f(v_2)\|_{\gamma_2} = \|v_1 - v_2\|_{\gamma_1}$ für alle $v_1, v_2 \in V$, d. h., f ist **abstandstreu**[13] (englisch: distance preserving).

(iv) Ist $(v_i)_{i \in I}$ eine orthonormale Familie in (V, γ_1), dann ist auch $(f(v_i))_{i \in I}$ eine orthonormale Familie in (W, γ_2).

(v) Ist v ein Einheitsvektor in (V, γ_1), dann ist $f(v)$ ein Einheitsvektor in (W, γ_2).

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Es gilt

$$\|f(v)\|_{\gamma_2}^2 = \gamma_2(f(v), f(v)) = \gamma_1(v, v) = \|v\|_{\gamma_1}^2.$$

Durch Wurzelziehen folgt die Behauptung.

Aussage (ii) $\Rightarrow$ Aussage (iii):

$$\|f(v_1) - f(v_2)\|_{\gamma_2} = \|f(v_1 - v_2)\|_{\gamma_2} = \|v_1 - v_2\|_{\gamma_1}.$$

Aussage (iii) $\Rightarrow$ Aussage (i): Aufgrund der Polarisierungsformel (32.5) ergibt sich

$$\begin{aligned} \gamma_1\big(v_1, -v_2\big) &= \frac{1}{2}\big(\|v_1 - v_2\|_{\gamma_1}^2 - \|v_1\|_{\gamma_1}^2 - \|-v_2\|_{\gamma_1}^2\big) \\ &= \frac{1}{2}\big(\|v_1 - v_2\|_{\gamma_1}^2 - \|v_1\|_{\gamma_1}^2 - \|v_2\|_{\gamma_1}^2\big) \end{aligned}$$

und ebenso

$$\begin{aligned} \gamma_2\big(f(v_1), f(-v_2)\big) &= \gamma_2\big(f(v_1), -f(v_2)\big) \\ &= \frac{1}{2}\big(\|f(v_1) - f(v_2)\|_{\gamma_2}^2 - \|f(v_1)\|_{\gamma_2}^2 - \|-f(v_2)\|_{\gamma_2}^2\big) \\ &= \frac{1}{2}\big(\|f(v_1) - f(v_2)\|_{\gamma_2}^2 - \|f(v_1) - f(0)\|_{\gamma_2}^2 - \|f(v_2) - f(0)\|_{\gamma_2}^2\big) \\ &= \frac{1}{2}\big(\|v_1 - v_2\|_{\gamma_1}^2 - \|v_1\|_{\gamma_1}^2 - \|v_2\|_{\gamma_1}^2\big). \end{aligned}$$

Das zeigt $\gamma_1\big(v_1, -v_2\big) = \gamma_2\big(f(v_1), f(-v_2)\big)$ für alle $v_1, v_2 \in V$, also ist f (γ_1, γ_2)-orthogonal.

Aussage (i) und Aussage (ii) $\Rightarrow$ Aussage (iv): Wir betrachten $v_1, v_2 \in V$. Es gilt $\gamma_1(v_1, v_2) = \gamma_2(f(v_1), f(v_2))$. $v_1 \perp v_2$ impliziert daher $f(v_1) \perp f(v_2)$. Ist also $(v_i)_{i \in I}$ eine orthogonale Familie, dann ist auch $(f(v_i))_{i \in I}$ eine orthogonale Familie. Sind zudem die Mitglieder von $(v_i)_{i \in I}$ normiert, dann sind wegen der Längentreue auch die Mitglieder von $(f(v_i))_{i \in I}$ normiert.

Aussage (iv) $\Rightarrow$ Aussage (v): Betrachte die Familie (v) mit einem normierten Mitglied.

[13]Diese Eigenschaft erklärt die Bezeichnung **Isometrie**.

Aussage (v) $\Longrightarrow$ Aussage (ii): Für $v = 0$ gilt $\|f(v)\|_{\gamma_2} = \|f(0)\|_{\gamma_2} = \|0\|_{\gamma_2} = 0$ und auch $\|v\|_{\gamma_1} = \|0\|_{\gamma_1} = 0$. Es sei nun also $v \neq 0$, dann gilt

$$\begin{aligned}
\|v\|_{\gamma_1} &= \left\| \|v\|_{\gamma_1} \frac{v}{\|v\|_{\gamma_1}} \right\|_{\gamma_1} \\
&= \|v\|_{\gamma_1} \Bigg\| \underbrace{\frac{v}{\|v\|_{\gamma_1}}}_{\text{normiert}} \Bigg\|_{\gamma_1} && \text{denn die Norm ist absolut homogen} \\
&= \|v\|_{\gamma_1} \left\| f\Big(\frac{v}{\|v\|_{\gamma_1}}\Big) \right\|_{\gamma_2} && \text{nach Voraussetzung} \\
&= \|v\|_{\gamma_1} \left\| \frac{1}{\|v\|_{\gamma_1}} f(v) \right\|_{\gamma_2} && \text{denn } f \text{ ist linear} \\
&= \left\| \frac{\|v\|_{\gamma_1}}{\|v\|_{\gamma_1}} f(v) \right\|_{\gamma_2} && \text{denn die Norm ist absolut homogen} \\
&= \|f(v)\|_{\gamma_2}.
\end{aligned}$$

□

Lemma 34.22 (orthogonale Abbildungen sind injektiv).
Es seien (V, γ_1) und (W, γ_2) zwei Euklidische Räume.

(i) Ist $f \in \operatorname{Hom}(V, W)$ eine (γ_1, γ_2)-orthogonale Abbildung, dann ist f injektiv.

(ii) Ist $f \in \operatorname{Hom}(V, W)$ eine (γ_1, γ_2)-orthogonale Abbildung und gilt zusätzlich $\dim(V) = \dim(W) = n \in \mathbb{N}_0$, dann ist f bijektiv und f^{-1} ebenfalls eine bijektive orthogonale Abbildung.

Beachte: Ein orthogonales $f \in \operatorname{Hom}(V, W)$ kann für endlich-dimensionales W also nur dann existieren, wenn $\dim(V) \leqslant \dim(W)$ gilt.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-12.1. □

Lemma 34.23 (Komposition orthogonaler Abbildungen).
Es seien (U, γ_1), (V, γ_2) und (W, γ_3) Euklidische Räume. Ist $f\colon V \to W$ eine (γ_2, γ_3)-orthogonale Abbildung und $g\colon U \to V$ eine (γ_1, γ_2)-orthogonale Abbildung, dann ist $f \circ g\colon U \to W$ eine (γ_1, γ_3)-orthogonale Abbildung.

Beweis. Die Komposition $f \circ g\colon U \to W$ ist linear nach Lemma 17.4. Weiter gilt für $u, v \in U$

$$\gamma_3\big((f \circ g)(u), (f \circ g)(v)\big) = \gamma_3\big(f(g(u)), f(g(v))\big) = \gamma_2\big(g(u), g(v)\big) = \gamma_1(u, v),$$

also ist $f \circ g$ (γ_1, γ_3)-orthogonal. □

Lemma 34.24 (bijektive Isometrie in den Koordinatenraum).
Es sei (V, γ) ein Euklidischer Raum mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei B_V eine Basis von V und $M \coloneqq \mathcal{M}_{B_V^*}^{B_V}(\gamma)$ die Darstellungsmatrix von γ. Dann ist die Abbildung

$$\Phi_{B_V} : (\mathbb{R}^n, \gamma_M) \to (V, \gamma)$$

eine bijektive Isometrie. Das heißt, für alle $u, v \in V$ gilt

$$\gamma(u, v) = x^\intercal M\, y \quad \text{mit den Koordinatenvektoren } x = \Phi_{B_V}^{-1}(u) \text{ und } y = \Phi_{B_V}^{-1}(v). \tag{34.8}$$

Beachte: Ist insbesondere B_V eine Orthonormalbasis von V, also $M = I$, dann ist Φ_{B_V} eine bijektive Isometrie zwischen $\mathbb{R}^n$ mit dem Standardinnenprodukt und (V, γ_I).

Beweis. Es sei $M = \mathcal{M}_{B_V^*}^{B_V}(\gamma)$ die Darstellungsmatrix von γ. Für Vektoren $u, v \in V$ und ihre Koordinatenvektoren $x = \Phi_{B_V}^{-1}(u)$ bzw. $y = \Phi_{B_V}^{-1}(v)$ gilt also

$$\gamma(u, v) = x^\intercal M\, y,$$

vgl. (31.3). □

Das folgende Resultat erklärt, wie wir die Orthogonalität eines Homomorphismus f zwischen Euklidischen Räumen anhand einer Darstellungsmatrix erkennen kann:

Lemma 34.25 (Orthogonalität in Darstellungsmatrizen von Homomorphismen)**.**
Es seien (V, γ_1) und (W, γ_2) zwei endlich-dimensionale Euklidische Räume mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$.

(*i*) Es seien B_V und B_W Basen von V bzw. von W und $M_1 := \mathcal{M}_{B_V^*}^{B_V}(\gamma_1) \in \mathbb{R}^{m \times m}$ sowie $M_2 := \mathcal{M}_{B_W^*}^{B_W}(\gamma_2) \in \mathbb{R}^{n \times n}$ die Darstellungsmatrizen von γ_1 bzw. von γ_2. Für $f \in \operatorname{Hom}(V, W)$ sind äquivalent:

(a) f ist (γ_1, γ_2)-orthogonal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_W}^{B_V}(f) \in \mathbb{R}^{n \times m}$ erfüllt

$$A^\intercal M_2\, A = M_1. \tag{34.9}$$

(*ii*) Sind B_V und B_W sogar Orthonormalbasen von V bzw. von W, dann sind für $f \in \operatorname{Hom}(V, W)$ äquivalent:

(a) f ist (γ_1, γ_2)-orthogonal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_W}^{B_V}(f) \in \mathbb{R}^{n \times m}$ erfüllt

$$A^\intercal I_n\, A = I_m. \tag{34.10}$$

Beachte: Wir schreiben $A^\intercal I_n\, A$ statt $A^\intercal A$, weil letztere Schreibweise von den Abbildungseigenschaften der dargestellten Objekte $A \mathrel{\hat{=}} f \in \operatorname{Hom}(V, W)$ und $A^\intercal \mathrel{\hat{=}} f^* \in \operatorname{Hom}(W^*, V^*)$ nicht passt. Es fehlt der Übergang durch I_n, das die Abbildung $w \mapsto \gamma_2(w, \cdot) \in \operatorname{Hom}(W, W^*)$ darstellt, die wir später (Definition 34.35) als Riesz-Isomorphismus kennenlernen werden.

Beweis. Aussage (*i*): Wir betrachten als Hilfestellung das kommutative Diagramm

$$\begin{tikzcd}
(W, \gamma_2) & (V, \gamma_1) \arrow[l, "f"'] \arrow[d, "\Phi_{B_V}^{-1}"] \\
(\mathbb{R}^n, \gamma_{M_2}) \arrow[u, "\Phi_{B_W}"] & (\mathbb{R}^m, \gamma_{M_1}) \arrow[l, "f_A"]
\end{tikzcd}$$

wie in Satz 19.7. Es sei $A = \mathcal{M}_{B_W}^{B_V}(f)$ die Darstellungsmatrix von f. Für Vektoren $u, v \in V$ mit Koordinatenvektoren $x = \Phi_{B_V}^{-1}(u)$ bzw. $y = \Phi_{B_V}^{-1}(v)$ gilt also

$$\begin{aligned} A\,x &= \Phi_{B_W}^{-1}(f(u)) \\ A\,y &= \Phi_{B_W}^{-1}(f(v)), \end{aligned}$$

vgl. (19.8a). Für die Innenprodukte gilt

$$\gamma_1(u,v) = [\Phi_{B_V}^{-1}(u)]^\intercal M_1\,[\Phi_{B_V}^{-1}(v)] = x^\intercal M_1\,y$$

und

$$\gamma_2(f(u), f(v)) = [\Phi_{B_W}^{-1}(f(u))]^\intercal M_2\,[\Phi_{B_W}^{-1}(f(v))] = [A\,x]^\intercal M_2\,[A\,y] = x^\intercal A^\intercal M_2\,A\,y.$$

Also ist f orthogonal genau dann, wenn $A^\intercal M_2\,A = M_1$ gilt.

Aussage (ii) ist ein spezieller Fall von Aussage (i), weil für Orthonormalbasen $M_1 = I_m$ und $M_2 = I_n$ gilt. □

Folgerung 34.26 (Orthogonalität in Darstellungsmatrizen von Endomorphismen)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum mit $\dim(V) = n \in \mathbb{N}_0$.

(i) Es sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\gamma) \in \mathbb{R}^{n\times n}$ die Darstellungsmatrix von γ. Für $f \in \operatorname{End}(V)$ sind äquivalent:

(a) f ist γ-orthogonal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{R}^{n\times n}$ erfüllt

$$A^\intercal M\,A = M. \tag{34.11}$$

(ii) Ist B_V sogar eine Orthonormalbasis von V, dann sind für $f \in \operatorname{End}(V)$ äquivalent:

(a) f ist γ-orthogonal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{R}^{n\times n}$ erfüllt

$$A^\intercal I_n\,A = I_n. \tag{34.12}$$

Beweis. Das Resultat folgt sofort aus Lemma 34.25. □

Die Eigenschaft (34.9) einer Darstellungsmatrix eines Homomorphismus ist zwar bzgl. der basisabhängigen Darstellungsmatrizen $M_1 := \mathcal{M}_{B_V^*}^{B_V}(\gamma_1)$ sowie $M_2 := \mathcal{M}_{B_W^*}^{B_W}(\gamma_2)$ definiert, muss aber tatsächlich unabhängig von der Wahl der Basen gelten. (**Quizfrage 34.1:** Warum ist das der Fall?)

In der Tat erhalten wir beim Übergang zu neuen Basen $\widehat{B}_V$ und $\widehat{B}_W$ mit den Transformationsmatrizen $T := \mathcal{T}_{\widehat{B}_V}^{B_V}$ und $S := \mathcal{T}_{\widehat{B}_W}^{B_W}$ die Beziehungen

$$\widehat{M}_1 = T^{-\intercal} M_1\,T^{-1}, \quad \widehat{M}_2 = S^{-\intercal} M_2\,S^{-1} \quad \text{und} \quad \widehat{A} = S\,A\,T^{-1} \quad \text{sowie} \quad \widehat{A}^\intercal = T^{-\intercal} A^\intercal S^\intercal.$$

Daraus ergibt sich.

$$\begin{aligned}
& \text{(34.9) gilt für die Matrizen bzgl. der neuen Basen } \widehat{B_V} \text{ und } \widehat{B_W} \\
\Leftrightarrow\quad & \widehat{A}^\intercal \widehat{M_2}\, \widehat{A} = \widehat{M_1} \\
\Leftrightarrow\quad & (T^{-\intercal} A^\intercal S^\intercal)\,(S^{-\intercal} M_2\, S^{-1})\, S\, A\, T^{-1} = T^{-\intercal} M_1\, T^{-1} \\
\Leftrightarrow\quad & T^{-\intercal} A^\intercal\, M_2\, A\, T^{-1} = T^{-\intercal} M_1\, T^{-1} \\
\Leftrightarrow\quad & A^\intercal M_2\, A = M_1 \\
\Leftrightarrow\quad & \text{(34.9) gilt für die Matrizen bzgl. der alten Basen } B_V \text{ und } B_W.
\end{aligned}$$

Ein analoges Resultat gilt natürlich für (34.11), also für den Sonderfall von Endomorphismen. Die Orthogonalitätseigenschaft eines Homomorphismus (oder Endomorphismus) zwischen endlich-dimensionalen Euklidischen Räumen kann also an einer beliebigen Darstellungsmatrix überprüft werden.

Definition 34.27 (orthogonale Matrix)**.**

(i) Es seien $m, n \in \mathbb{N}_0$, $M_1 \in \mathbb{R}^{m\times m}$ und $M_2 \in \mathbb{R}^{n\times n}$ beide symmetrisch und positiv definit.[14] Eine Matrix $A \in \mathbb{R}^{n\times m}$ heißt (M_1, M_2)**-orthogonal** (englisch: (M_1, M_2)-orthogonal matrix), wenn (34.9) gilt, also

$$A^\intercal M_2\, A = M_1.$$

(ii) Es sei $n \in \mathbb{N}_0$ und $M \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit. Eine Matrix $A \in \mathbb{R}^n$ heißt M**-orthogonal** (englisch: M-orthogonal matrix), wenn (34.11) gilt, also

$$A^\intercal M\, A = M. \qquad \triangle$$

Bemerkung 34.28 (zum Begriff der Orthogonalität von Matrizen)**.**

(i) (M_1, M_2)-Orthogonalität ist eine Eigenschaft, die für Darstellungsmatrizen von Homomorphismen zwischen Euklidischen Räumen erklärt ist.[15]

(ii) M-Orthogonalität ist eine Eigenschaft, die für Darstellungsmatrizen von Endomorphismen eines Euklidischen Raumes erklärt ist.[16] △

Lemma 34.29 (Eigenwerte orthogonaler Endomorphismen)**.**
Es sei (V, γ) ein Euklidischer Raum. Ist $f \in \operatorname{End}(V)$ γ-orthogonal, dann gilt $\Lambda(f) \subseteq \{\pm 1\}$.

Beachte: Hat der Endomorphismus f des reellen Vektorraumes V einen Eigenwert ungleich ± 1, so ist er also in keinem Innenprodukt γ von V orthogonal.

Beweis. Es sei $\lambda \in \mathbb{R}$ ein Eigenwert von f und $v \in V$ ein zugehöriger Eigenvektor. Dann gilt

$$\|v\| = \|f(v)\| = \|\lambda\, v\| = |\lambda|\, \|v\|.$$

Da $v \neq 0$ und damit $\|v\| \neq 0$ ist, muss $|\lambda| = 1$ gelten. □

[14] M_1 induziert also ein Innenprodukt auf $\mathbb{R}^m$, und M_2 induziert eines auf $\mathbb{R}^n$.

[15] Die (M_1, M_2)-Orthogonalität von A bedeutet, dass für jede Wahl von Basen B_V und B_W von Vektorräumen V und W passender Dimension der durch A dargestellte Homomorphismus orthogonal bzgl. der durch M_1 bzw. M_2 auf V bzw. W induzierten Innenprodukte ist. Gemeint ist: $\gamma_1(u, v) = [\Phi_{B_V}^{-1}(u)]^\intercal M_1\, [\Phi_{B_V}^{-1}(v)]$.

[16] Die M-Orthogonalität von A bedeutet, dass für jede Wahl einer Basis B_V eines Vektorraumes V passender Dimension der durch A dargestellte Endomorphismus orthogonal bzgl. des durch M auf V induzierten Innenprodukts ist.

Lemma 34.30 (orthogonale Endomorphismen bilden eine Gruppe).
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Dann gilt:

(i) Die γ-orthogonalen Endomorphismen von (V, γ) bilden eine Gruppe bzgl. der Komposition, genannt die **orthogonale Gruppe** (englisch: orthogonal group) des Euklidischen Raumes (V, γ):

$$\mathrm{O}(V, \gamma) := \{f \in \mathrm{End}(V) \mid f \text{ ist } \gamma\text{-orthogonal}\}. \tag{34.13}$$

(ii) Die γ-orthogonalen Endomorphismen $f \in \mathrm{O}(V, \gamma)$ mit $\det(f) = 1$ bilden einen Normalteiler von $\mathrm{O}(V, \gamma)$, genannt die **spezielle orthogonale Gruppe** (englisch: special orthogonal group) des Euklidischen Raumes (V, γ):

$$\mathrm{SO}(V, \gamma) := \{f \in \mathrm{O}(V, \gamma) \mid \det(f) = 1\}. \tag{34.14}$$

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-12.2. □

Die Aussage von Lemma 34.30 übersetzt sich wie folgt in Matrizen:

Folgerung 34.31 (Darstellungsmatrizen orthogonaler Endomorphismen bilden eine Gruppe).
Es sei $n \in \mathbb{N}_0$ und $M \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Dann gilt:

(i) Die M-orthogonalen Matrizen in $\mathbb{R}^{n \times n}$ bilden eine Gruppe bzgl. der Matrix-Multiplikation, genannt die **orthogonale Gruppe** des Euklidischen Raumes $(\mathbb{R}^n, \gamma_M)$:

$$\mathrm{O}(\mathbb{R}^n, \gamma_M) := \{A \in \mathbb{R}^{n \times n} \mid A \text{ ist } M\text{-orthogonal}\}. \tag{34.15}$$

(ii) Die M-orthogonalen Matrizen $A \in \mathbb{R}^{n \times n}$ mit $\det(A) = 1$ bilden einen Normalteiler von $\mathrm{O}(\mathbb{R}^n, \gamma_M)$, genannt die **spezielle orthogonale Gruppe** des Euklidischen Raumes $(\mathbb{R}^n, \gamma_M)$:

$$\mathrm{SO}(\mathbb{R}^n, \gamma_M) := \{A \in \mathrm{O}(\mathbb{R}^n, \gamma_M) \mid \det(A) = 1\}. \tag{34.16}$$

Bemerkung 34.32 (zum Begriff der Orthogonalität).
In Lehrbüchern wird der Begriff der Orthogonalität für Matrizen in der Regel nur für den Fall $M = I$ beschrieben, also für das Standardinnenprodukt auf $\mathbb{R}^n$. Das heißt, der Begriff der **orthogonalen Matrix** wird dort ausschließlich im Sinne von $A^\intercal I A = A^\intercal A = I$ verwendet. Dieser Fall bedeutet, dass für B_V stets eine Orthonormalbasis gewählt wird.

Die orthogonale Gruppe von $\mathrm{O}(\mathbb{R}^n, \gamma_I)$ wird dann oft einfach mit $\mathrm{SO}(n)$ bezeichnet. Wir folgen dieser Praxis nicht. Unsere Notation $\mathrm{O}(\mathbb{R}^n, \gamma_M)$ stellt klar, dass es um die orthogonale Gruppe des Euklidischen Raumes $(\mathbb{R}^n, \gamma_M)$ geht. △

Beispiel 34.33 (orthogonale Endomorphismen und Matrizen).

(i) Die identische Abbildung id_V ist in jedem reellen Vektorraum γ-orthogonal bzgl. jedes beliebigen Innenprodukts γ. Sie gehört sogar zu $\mathrm{SO}(V, \gamma)$ für alle γ.

(ii) Die Einheitsmatrix I_n ist in jedem Vektorraum $\mathbb{R}^n$ M-orthogonal bzgl. jeder symmetrischen, positiv definiten Matrix M. Sie gehört sogar zu $\mathrm{SO}(\mathbb{R}^n, \gamma_M)$ für alle symmetrischen, positiv definiten Matrizen M.

(*iii*) Die Matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

ist I-orthogonal, denn es gilt

$$A^\intercal I\, A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Die Matrix A ist aber nicht M-orthogonal bzgl. der symmetrischen, positiv definiten Matrix

$$M = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

denn es gilt

$$A^\intercal M\, A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \neq M.$$

(*iv*) Die Drehabbildung, ein Endomorphismus von $\mathbb{R}^2$ mit der Darstellungsmatrix

$$A = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

bzgl. der Standardbasis, ist orthogonal bzgl. des Standardinnenprodukts, denn es gilt

$$A^\intercal\, I\, A = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Wegen $\det(A) = 1$ gehört A sogar zu $\mathrm{SO}(\mathbb{R}^2, \gamma_I)$.

Man kann zeigen, dass Drehabbildungen die einzigen orthogonalen Endomorphismen von $\mathbb{R}^2$ sind.

(*v*) Die Spiegelungsabbildung, ein Endomorphismus von $\mathbb{R}^2$ mit der Darstellungsmatrix

$$A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}$$

bzgl. der Standardbasis, ist orthogonal bzgl. des Standardinnenprodukts, denn es gilt

$$A^\intercal\, I\, A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Wegen $\det(A) = -1$ gehört A aber nicht zu $\mathrm{SO}(\mathbb{R}^2, \gamma_I)$.

Man kann zeigen, dass Spiegelungsabbildungen die einzigen Endomorphismen von $\mathbb{R}^2$ sind, die zu $\mathrm{O}(\mathbb{R}^2, \gamma_I) \setminus \mathrm{SO}(\mathbb{R}^2, \gamma_I)$ gehören. △

Ende der Vorlesung 23

§ 34.4 Die Riesz-Abbildung und adjungierte Homomorphismen

Zwischen einem Vektorraum V und seinem Dualraum V^* gibt es auch bei endlicher Dimension keinen *kanonischen* Isomorphismus. Wenn wir jedoch ein Innenprodukt in V gewählt haben, dann induziert dieses eine bevorzugte Wahl eines solchen Isomorphimus in $\Gamma \in \operatorname{Hom}(V, V^*)$, die mit den Innenprodukten γ in V und γ^{-1} in V^* kompatibel ist, also eine lineare Isometrie. Dieser Isomorphismus wird der **Riesz-Isomorphismus** des Euklidischen Raumes (V, γ) genannt.

Satz 34.34 (**Darstellungssatz von Riesz**[17]**).**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Dann ist die Abbildung

$$\Gamma \colon V \ni v \mapsto \gamma(\cdot, v) \in V^* \tag{34.17}$$

ein Isomorphismus. Wird V^* mit dem Innenprodukt γ^{-1} ausgestattet, dann ist $\Gamma \colon (V, \gamma) \to (V^*, \gamma^{-1})$ eine bijektive Isometrie. Es gilt

$$\gamma(u, v) = \langle \Gamma(v) \,, u \rangle = \langle \Gamma(u) \,, v \rangle = \gamma^{-1}\big(\Gamma(u), \Gamma(v)\big) \tag{34.18}$$

für alle $u, v \in V$.

Beweis. Nach Lemma 34.6 ist Γ bijektiv, und die inverse Bilinearform $\gamma^{-1}(\xi, \eta) \coloneqq \langle \eta \,, \Gamma^{-1}\xi \rangle$ ist wieder positiv definit. Sie erbt außerdem die Symmetrie von γ, ist also tatsächlich ein Innenprodukt auf V^*. Folglich ist (V^*, γ^{-1}) ein Euklidischer Raum.

Es bleibt die Orthogonalität, also die Isometrieeigenschaft von Γ zu zeigen. Es gilt aber nach Definition von Γ

$$\gamma(u, v) = \langle \Gamma(v) \,, u \rangle$$

und nach Definition von γ^{-1}

$$\gamma^{-1}\big(\Gamma(u), \Gamma(v)\big) = \big\langle \Gamma(v) \,, \Gamma^{-1}(\Gamma(u)) \big\rangle = \langle \Gamma(v) \,, u \rangle.$$

Die noch fehlende Gleichheit $\langle \Gamma(v) \,, u \rangle = \langle \Gamma(u) \,, v \rangle$ folgt aus der Symmetrie von γ. □

Definition 34.35 (Riesz-Isomorphismus)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Die durch (34.17) definierte Abbildung, also

$$\Gamma_{V \to V^*} \colon V \ni v \mapsto \gamma(\cdot, v) \in V^*$$

heißt der **Riesz-Isomorphismus** (englisch: Riesz isomorphism) **von V nach V^***. △

Beachte: Wir können $\Gamma_{V \to V^*}$ als *kanonischen* Isomorphismus der Euklidischen Räume (V, γ) und (V^*, γ^{-1}) bezeichnen. Manchmal wird der Riesz-Isomorphismus auch als die inverse Abbildung von V^* nach V definiert, die natürlich ebenfalls eine lineare Isometrie ist. Die Bijektivität von (34.7) bedeutet, dass jede Linearform auf V von der Form $\gamma(\cdot, v)$ für ein $v \in V$ ist. Darin begründet sich die Bezeichnung „Darstellungssatz“.

Wir drücken die Aussage des Rieszschen Darstellungssatzes 34.34 noch einmal für den Fall $V = \mathbb{R}^n$ mit dem Innenprodukt $\gamma = \gamma_M$ aus, induziert durch die symmetrische, positiv definite Matrix M:

[17] Der Satz wird auch **Darstellungssatz von Fréchet-Riesz** genannt, englisch: (Fréchet-)Riesz representation theorem.

Folgerung 34.36 (Rieszscher Darstellungssatz für $(\mathbb{R}^n, \gamma_M)$)**.**
Es sei $n \in \mathbb{N}_0$ und $M \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Dann ist die Abbildung

$$\Gamma\colon \mathbb{R}^n \ni x \mapsto M\,x \in (\mathbb{R}^n)^* \tag{34.19}$$

eine bijektive Isometrie der Euklidischen Räume $(\mathbb{R}^n, \gamma_M)$ und $((\mathbb{R}^n)^*, \gamma_{M^{-1}})$. Es gilt

$$x^\mathsf{T} M\, y = \underbrace{(x^\mathsf{T} M)\, y}_{=\langle M\,x, y\rangle} = \underbrace{(y^\mathsf{T} M)\, x}_{=\langle M\,y, x\rangle} = (x^\mathsf{T} M)\, M^{-1}(M\, y) \tag{34.20}$$

für alle $x, y \in \mathbb{R}^n$.

Satz 34.37 (Orthogonalität dualer Homomorphismen)**.**
Es seien (V, γ_1) und (W, γ_2) zwei endlich-dimensionale Euklidische Räume derselben Dimension. Weiter sei $f \in \operatorname{Hom}(V, W)$. Dann sind äquivalent:

(i) f ist (γ_1, γ_2)-orthogonal.

(ii) f^* ist $(\gamma_2^{-1}, \gamma_1^{-1})$-orthogonal.

Beweis. Wir beweisen das Resultat mit Hilfe von Darstellungsmatrizen. Es sei dazu $\dim(V) = n \in \mathbb{N}_0$ und B_V und B_W irgendwelche Basen von V bzw. von W und B_V^* sowie B_W^* die zugehörigen dualen Basen. Wir setzen

$$M_1 \coloneqq \mathcal{M}_{B_V^*}^{B_V}(\gamma_1), \quad M_2 \coloneqq \mathcal{M}_{B_W^*}^{B_W}(\gamma_2), \quad A \coloneqq \mathcal{M}_{B_W}^{B_V}(f), \quad \text{also gilt } A^\mathsf{T} = \mathcal{M}_{B_V^*}^{B_W^*}(f^*).$$

Nach Lemma 34.25 ist die (γ_1, γ_2)-Orthogonalität von f äquivalent zu

$$A^\mathsf{T} M_2\, A = M_1. \tag{34.21}$$

Da γ_1^{-1} bzgl. der dualen Basis B_V^* durch M_1^{-1} dargestellt wird und analog γ_2^{-1} durch M_2^{-1}, ist die $(\gamma_2^{-1}, \gamma_1^{-1})$-Orthogonalität von f^* nach Lemma 34.25 äquivalent zu

$$A\, M_1^{-1} A^\mathsf{T} = M_2^{-1}. \tag{34.22}$$

Nimmt man Aussage (i) an, dann ist nach Lemma 34.22 f injektiv und wegen übereinstimmender Dimensionen von V und W nach Folgerung 18.9 sogar bijektiv. Also ist A und damit auch A^T invertierbar. Dieselbe Schlussfolgerung erhält man aus der Annahme von Aussage (ii). Es gilt daher

$$\begin{aligned}
& A^\mathsf{T} M_2\, A = M_1 \\
\Leftrightarrow \quad & M_2\, A = A^{-\mathsf{T}} M_1 \\
\Leftrightarrow \quad & A = M_2^{-1} A^{-\mathsf{T}} M_1 \\
\Leftrightarrow \quad & A\, M_1^{-1} = M_2^{-1} A^{-\mathsf{T}} \\
\Leftrightarrow \quad & A\, M_1^{-1} A^\mathsf{T} = M_2^{-1}.
\end{aligned}$$

Daraus folgt die Behauptung. □

Quizfrage 34.2: Ist die Voraussetzung übereinstimmender Dimensionen von V und W in Satz 34.37 wesentlich?

Als Spezialfall von Satz 34.37 ergibt sich sofort:

Folgerung 34.38 (Orthogonalität dualer Endomorphismen)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Weiter sei $f \in \operatorname{End}(V)$. Dann sind äquivalent:

(i) f ist γ-orthogonal.

(ii) f^* ist γ^{-1}-orthogonal.

Dabei werden die γ-Orthogonalität von f und die γ^{-1}-Orthogonalität von f^* durch die äquivalenten Beziehungen

$$A^\intercal M A = M \tag{34.23}$$

und

$$A M^{-1} A^\intercal = M^{-1} \tag{34.24}$$

ausgedrückt.

Wir definieren nun die **Adjungierte** f° einer linearen Abbildung f zwischen den Euklidischen Räumen (V, γ_1) und (W, γ_2).[18] Die Definition ergibt sich aus der Forderung, dass das folgende Diagramm kommutiert:

$$\begin{tikzcd}
(V, \gamma_1) \arrow[d, "\Gamma_{V \to V^*}"'] & (W, \gamma_2) \arrow[l, "f^\circ"'] \arrow[d, "\Gamma_{W \to W^*}"] \\
(V^*, \gamma_1^{-1}) & (W^*, \gamma_2^{-1}) \arrow[l, "f^*"']
\end{tikzcd}$$

Definition 34.39 (adjungierter Homomorphismus)**.**
Es seien (V, γ_1) und (W, γ_2) zwei endlich-dimensionale Euklidische Räume. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung. Dann heißt

$$f^\circ \coloneqq \Gamma_{V \to V^*}^{-1} \circ f^* \circ \Gamma_{W \to W^*} \colon W \longrightarrow V \tag{34.25}$$

der zu f **(γ_1, γ_2)-adjungierte Homomorphismus** (englisch: adjoint homomorphism) bzw. die zu f **(γ_1, γ_2)-adjungierte Abbildung** (englisch: adjoint linear map).[19] △

Die Definition (34.25) lautet ausgeschrieben

$$\gamma_2\big(w, f(v)\big) = \gamma_1\big(f^\circ(w), v\big) \tag{34.26}$$

für alle $v \in V$ und alle $w \in W$.

Bemerkung 34.40 (zur (γ_1, γ_2)-adjungierten Abbildung)**.**
Die adjungierte Abbildung hängt — im Unterschied zur dualen Abbildung f^*! — von den beiden Innenprodukten (γ_1, γ_2) ab, was sich in der Notation f° aber nicht niederschlägt. Häufig wird in der Literatur auch nur von „der adjungierten Abbildung" gesprochen, wobei die Innenprodukte aus dem Zusammenhang erschlossen werden müssen. △

[18] Die Notation für die adjungierte Abbildung ist in der Literatur nicht einheitlich. Man findet auch Bezeichnungen wie $f^\sharp$, f^{ad} und weitere.
[19] Zur Klarstellung: $\Gamma_{V \to V^*}^{-1}$ bedeutet $(\Gamma_{V \to V^*})^{-1}$.

Satz 34.41 (Darstellungsmatrix des (γ_1, γ_2)-adjungierten Homomorphismus, vgl. Satz 21.33)**.**

Es seien (V, γ_1) und (W, γ_2) zwei endlich-dimensionale Euklidische Räume. Es seien B_V und B_W Basen von V bzw. von W und $M_1 \coloneqq \mathcal{M}_{B_V^*}^{B_V}(\gamma_1) \in \mathbb{R}^{m \times m}$ sowie $M_2 \coloneqq \mathcal{M}_{B_W^*}^{B_W}(\gamma_2) \in \mathbb{R}^{n \times n}$ die Darstellungsmatrizen von γ_1 bzw. von γ_2. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung. Ist $A = \mathcal{M}_{B_W}^{B_V}(f)$ die Darstellungsmatrix von f bzgl. der Basen B_V und B_W, dann ist

$$A^\circ \coloneqq \mathcal{M}_{B_V}^{B_W}(f^\circ) = M_1^{-1} A^\intercal M_2 \tag{34.27}$$

die Darstellungsmatrix von $f^\circ \in \operatorname{Hom}(W, V)$.

Beweis. Die Darstellungsmatrizen der Abbildungen in (34.25) sind

$$\begin{aligned}
\mathcal{M}_{B_W^*}^{B_W}(\Gamma_{W \to W^*}) &= M_2 \\
\mathcal{M}_{B_V^*}^{B_W^*}(f^*) &= A^\intercal \qquad \text{nach Satz 21.33} \\
\mathcal{M}_{B_V^*}^{B_V}(\Gamma_{V \to V^*}) &= M_1 \\
\text{und daher} \quad \mathcal{M}_{B_V}^{B_V^*}(\Gamma_{V \to V^*}^{-1}) &= M_1^{-1}
\end{aligned}$$

Nach Satz 19.8 wird die Komposition der Abbildungen durch das Produkt der Matrizen dargestellt, also

$$\mathcal{M}_{B_V}^{B_W}(f^\circ) = \mathcal{M}_{B_V}^{B_V^*}(\Gamma_{V \to V^*}^{-1})\, \mathcal{M}_{B_V^*}^{B_W^*}(f^*)\, \mathcal{M}_{B_W^*}^{B_W}(\Gamma_{W \to W^*}) = M_1^{-1} A^\intercal M_2. \qquad \square$$

Lemma 34.42 ((γ_1, γ_2)-biadjungierte Abbildung)**.**

Es seien (V, γ_1) und (W, γ_2) zwei endlich-dimensionale Euklidische Räume. Dann ist die **(γ_1, γ_2)-biadjungierte Abbildung** (englisch: biadjoint homomorphism, biadjoint map) $f^{\circ\circ} \coloneqq (f^\circ)^\circ$ identisch zu f, unabhängig von den Innenprodukten γ_1 und γ_2.

Beweis. Wir beweisen das Resultat unter Zuhilfenahme der Darstellungsmatrizen und Satz 34.41. Die Darstellungsmatrix der (γ_1, γ_2)-adjungierten Abbildung $f^\circ \in \operatorname{Hom}(W, V)$ ist nach (34.27)

$$A^\circ = M_1^{-1} A^\intercal M_2.$$

Die Darstellungsmatrix der biadjungierten $f^{\circ\circ} \in \operatorname{Hom}(V, W)$ ist daher

$$\begin{aligned}
A^{\circ\circ} &= M_2^{-1} (A^\circ)^\intercal M_1 && \text{nach (34.27)} \\
&= M_2^{-1} (M_1^{-1} A^\intercal M_2)^\intercal M_1 \\
&= M_2^{-1} (M_2^\intercal\, A\, M_1^{-\intercal})\, M_1 \\
&= M_2^{-1} M_2\, A\, M_1^{-1}\, M_1 && \text{denn } M_1 = M_1^\intercal \text{ und } M_2 = M_2^\intercal \\
&= A.
\end{aligned}$$

Damit haben $A^{\circ\circ}$ und A dieselben Darstellungsmatrizen bzgl. identischer Basen (B_V, B_W) und sind daher dieselben Abbildungen. $\square$

§ 34.5 Nochmal vier fundamentale Unterräume zu einer linearen Abbildung

In § 21.6 hatten wir die vier fundamentalen Unterräume, die zu einer linearen Abbildung $f \in \operatorname{Hom}(V, W)$ gehören, untersucht. Dort hatten wir in Satz 21.36 die folgenden Beziehungen gezeigt:

$$\begin{aligned} \operatorname{Bild}(f^*) &= \operatorname{Kern}(f)^0 && \text{in } V^* \\ \operatorname{Kern}(f^*) &= \operatorname{Bild}(f)^0 && \text{in } W^* \\ \operatorname{Bild}(f) &= {}^0\operatorname{Kern}(f^*) && \text{in } W \\ \operatorname{Kern}(f) &= {}^0\operatorname{Bild}(f^*) && \text{in } V. \end{aligned}$$

Wir wollen diese Beziehungen nun noch einmal anders formulieren, und zwar unter Verwendung einer zu f adjungierten Abbildung f° und orthogonalen Komplementen an Stelle von (Prä-)Annihilatoren. Wir beschränken uns hier auf den Fall endlicher Dimensionen, sodass das Resultat nicht vom Auswahlaxiom abhängt.

Satz 34.43 (vier fundamentale Unterräume zu einer linearen Abbildung)**.**
Es seien (V, γ_1) und (W, γ_2) zwei endlich-dimensionale Euklidische Räume. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung mit (γ_1, γ_2)-adjungierter Abbildung $f^\circ \in \operatorname{Hom}(W, V)$. Dann gilt:

$$\operatorname{Bild}(f^\circ) = \operatorname{Kern}(f)^\perp \quad \text{in } V \tag{34.28a}$$

$$\operatorname{Kern}(f^\circ) = \operatorname{Bild}(f)^\perp \quad \text{in } W \tag{34.28b}$$

$$\operatorname{Bild}(f) = \operatorname{Kern}(f^\circ)^\perp \quad \text{in } W \tag{34.28c}$$

$$\operatorname{Kern}(f) = \operatorname{Bild}(f^\circ)^\perp \quad \text{in } V. \tag{34.28d}$$

Die orthogonalen Komplemente in V sind bzgl. γ_1 und in W bzgl. γ_2 zu verstehen.

Beweis. Wir beginnen mit (34.28b). Es gilt

$$\begin{aligned} & w \in \operatorname{Bild}(f)^\perp \\ \Leftrightarrow\quad & \gamma_2(f(v), w) = 0 && \text{für alle } v \in V \\ \Leftrightarrow\quad & \gamma_1(v, f^\circ(w)) = 0 && \text{für alle } v \in V \\ \Leftrightarrow\quad & f^\circ(w) = 0 && \text{denn } \gamma_1 \text{ ist als Innenprodukt nicht ausgeartet} \\ \Leftrightarrow\quad & w \in \operatorname{Kern}(f^\circ). \end{aligned}$$

Nun adressieren wir (34.28d). Es gilt

$$\begin{aligned} & v \in \operatorname{Kern}(f) \\ \Leftrightarrow\quad & v \in \operatorname{Kern}(f^{\circ\circ}) && \text{denn } f^{\circ\circ} = f \text{ nach Lemma 34.42} \\ \Leftrightarrow\quad & v \in \operatorname{Bild}(f^\circ)^\perp && \text{wegen (34.28b).} \end{aligned}$$

(34.28a) folgt nun aus (34.28d), denn es gilt

$$\begin{aligned} & \operatorname{Kern}(f) = \operatorname{Bild}(f^\circ)^\perp \\ \Leftrightarrow\quad & \operatorname{Kern}(f)^\perp = \operatorname{Bild}(f^\circ)^{\perp\perp} = \operatorname{Bild}(f^\circ). \end{aligned}$$

(34.28c) folgt ähnlich aus (34.28b):

$$\begin{aligned} & \operatorname{Kern}(f^\circ) = \operatorname{Bild}(f)^\perp \\ \Leftrightarrow\quad & \operatorname{Kern}(f^\circ)^\perp = \operatorname{Bild}(f)^{\perp\perp} = \operatorname{Bild}(f). \end{aligned}$$ □

Folgerung 34.44 (Endomorphismen induzieren orthogonale direkte Zerlegungen)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Weiter sei $f \in \operatorname{End}(V)$ eine lineare Abbildung mit γ-adjungierter Abbildung $f^\circ \in \operatorname{End}(V)$. Dann gilt:

$$V = \operatorname{Kern}(f) \ominus \operatorname{Bild}(f^\circ) \tag{34.29a}$$

$$V = \operatorname{Kern}(f^\circ) \ominus \operatorname{Bild}(f). \tag{34.29b}$$

Beweis. Für jeden Unterraum $U \subseteq V$ gilt $V = U \ominus U^\perp$, insbesondere also

$$V = \operatorname{Kern}(f) \ominus \operatorname{Kern}(f)^\perp = \operatorname{Kern}(f) \ominus \operatorname{Bild}(f^\circ)$$

wegen (34.28a) und analog

$$V = \operatorname{Bild}(f) \ominus \operatorname{Bild}(f)^\perp = \operatorname{Bild}(f) \ominus \operatorname{Kern}(f^\circ)$$

wegen (34.28b). □

Ende der Vorlesung 24

Ende der Woche 12

§ 34.6 Selbstadjungierte, orthogonale und normale Endomorphismen

Wir untersuchen in diesem Abschnitt die Begriffe Selbstadjungiertheit und Normalität, in welcher Beziehung sie zueinander stehen und wie sich die Selbstadjungiertheit und die Normalität auf die Diagonalisierbarkeit von Endomorphismen in Euklidischen Räumen auswirken.

Definition 34.45 (selbstadjungierte und normale Endomorphismen)**.**
Es sei (V, γ) ein Euklidischer Raum.

(*i*) $f \in \operatorname{End}(V)$ heißt **γ-selbstadjungiert** (englisch: γ-self adjoint), wenn $f = f^\circ$ gilt.

(*ii*) $f \in \operatorname{End}(V)$ heißt **γ-normal** (englisch: γ-normal), wenn $f \circ f^\circ = f^\circ \circ f$ gilt.[20] △

Wir untersuchen nun, wie wir die Selbstadjungiertheit bzw. die Normalität eines Endomorphismus eines Euklidischen Raumes anhand einer Darstellungsmatrix erkennen können:

Lemma 34.46 (Selbstadjungiertheit in Darstellungsmatrizen von Endomorphismen)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum.

(*i*) Es sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\gamma) \in \mathbb{R}^{n \times n}$ die Darstellungsmatrix von γ. Für $f \in \operatorname{End}(V)$ sind äquivalent:

   (a) f ist γ-selbstadjungiert.

   (b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{R}^{n \times n}$ erfüllt $A^\circ = A$, also

$$M^{-1} A^\mathsf{T} M = A. \tag{34.30}$$

[20] Das heißt, f kommutiert mit seiner γ-adjungierten Abbildung. Anders gesagt, f ist γ-normal genau dann, wenn $f \circ f^\circ$ γ-selbstadjungiert ist.

(ii) Ist B_V sogar eine Orthonormalbasis von V, dann sind für $f \in \operatorname{End}(V)$ äquivalent:

(a) f ist γ-selbstadjungiert.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{R}^{n \times n}$ erfüllt

$$I_n^{-1} A^\intercal I_n = A. \tag{34.31}$$

Beachte: Wir schreiben in (34.31) $I_n^{-1} A^\intercal I_n$ statt $A^\intercal$, weil letztere Schreibweise von den Abbildungseigenschaften der dargestellten Objekte $A \mathrel{\widehat{=}} f \in \operatorname{End}(V)$ und $A^\intercal \mathrel{\widehat{=}} f^* \in \operatorname{End}(V^*)$ nicht passt. Es fehlt der Übergang durch I_n und I_n^{-1}, das den Riesz-Isomorphismus $(V, \gamma) \to (V^*, \gamma^{-1})$ bzw. den inversen Riesz-Isomorphismus darstellt.

Beweis. Aussage (i): Aus Satz 34.41 folgt die Darstellung $A^\circ = M^{-1} A^\intercal M$. $f^\circ = f$ ist äquivalent dazu, dass die Darstellungsmatrizen von f° und f übereinstimmen, siehe Satz 19.5.

Aussage (ii) ist ein Spezialfall von Aussage (i). □

Lemma 34.47 (Normalität in Darstellungsmatrizen von Endomorphismen)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum mit $\dim(V) = n \in \mathbb{N}_0$.

(i) Es sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\gamma) \in \mathbb{R}^{n \times n}$ die Darstellungsmatrix von γ. Für $f \in \operatorname{End}(V)$ sind äquivalent:

(a) f ist γ-normal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{R}^{n \times n}$ erfüllt $A\,A^\circ = A^\circ\,A$, also

$$A\,M^{-1} A^\intercal M = M^{-1} A^\intercal M\,A. \tag{34.32}$$

(ii) Ist B_V sogar eine Orthonormalbasis von V, dann sind für $f \in \operatorname{End}(V)$ äquivalent:

(a) f ist γ-normal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{R}^{n \times n}$ erfüllt

$$A\,I_n^{-1} A^\intercal I_n = I_n^{-1} A^\intercal I_n\,A. \tag{34.33}$$

Beachte: Wir schreiben auch hier in (34.33) wieder I_n und I_n^{-1} explizit, um die Abbildungseigenschaften der dargestellten Objekte erkennbar wiederzugeben.

Beweis. Der Beweis ist ähnlich zum Beweis von Lemma 34.46. □

Lemma 34.48 (Zusammenhang zwischen Orthogonalität, Selbstadjungiertheit und Normalität)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum.

(i) Ist $f \in \operatorname{End}(V)$ γ-selbstadjungiert, dann ist f auch γ-normal.

(ii) Ist $f \in \operatorname{End}(V)$ γ-orthogonal, dann ist f auch γ-normal.

Beweis. Aussage (i): $f^\circ = f$ impliziert $f \circ f^\circ = f^\circ \circ f$.

Aussage (ii): Es sei $\dim(V) = n \in \mathbb{N}_0$. Wir beweisen das Resultat mit Hilfe von Darstellungsmatrizen bzgl. irgendeiner Basis B_V. Die γ-Orthogonalität von f ist nach Folgerung 34.26 gerade $A^\intercal M A = M$. Daraus folgt einerseits

$$M^{-1}(A^\intercal M A) = M^{-1} M = I.$$

Zudem ist $A \in \mathbb{R}^{n\times n}$ invertierbar, da die γ-Orthogonalität die Injektivität und damit die Bijektivität impliziert. Damit folgt andererseits

$$A\, M^{-1}(A^\intercal M\, A)\, A^{-1} = A\, M^{-1} M\, A^{-1} = I.$$

Das heißt nach (34.32), dass f γ-normal ist. □

Beispiel 34.49 (Zusammenhang zwischen Orthogonalität, Selbstadjungiertheit und Normalität).

(i) Die Drehabbildung, ein Endomorphismus von $\mathbb{R}^2$ mit der Darstellungsmatrix

$$A = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

bzgl. der Standardbasis, ist orthogonal bzgl. des Standardinnenprodukts, siehe Beispiel 34.33. Also ist A auch I-normal.

Im Allgemeinen ist A aber nicht I-selbstadjungiert, denn die Prüfung von (34.31) ergibt

$$A^\circ = I_n^{-1} A^\intercal I_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}^\intercal \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \neq A,$$

außer im Fall $\sin(\alpha) = 0$, was im Intervall $[0, 2\pi)$ genau für $\alpha = 0$ und $\alpha = \pi$ der Fall ist.

Die I-adjungierte Abbildung mit Darstellungsmatrix A° kann gedeutet werden als Drehung um den Winkel $-\alpha$.

(ii) Die Spiegelungsabbildung, ein Endomorphismus von $\mathbb{R}^2$ mit der Darstellungsmatrix

$$A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}$$

bzgl. der Standardbasis, ist orthogonal bzgl. des Standardinnenprodukts, siehe Beispiel 34.33. Also ist A auch I-normal.

Darüber hinaus ist A auch I-selbstadjungiert, denn die Prüfung von (34.31) ergibt

$$A^\circ = I_n^{-1} A^\intercal I_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}^\intercal \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix} = A. \quad \triangle$$

Folgerung 34.50 (normale Endomorphismen induzieren orthogonale direkte Zerlegungen, vgl. Folgerung 34.44).
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Weiter sei $f \in \operatorname{End}(V)$ γ-normal. Dann gilt:

$$\operatorname{Kern}(f^\circ) = \operatorname{Kern}(f) \tag{34.34a}$$

$$\operatorname{Bild}(f^\circ) = \operatorname{Bild}(f) \tag{34.34b}$$

$$V = \operatorname{Kern}(f) \oplus \operatorname{Bild}(f) \tag{34.34c}$$

$$V = \operatorname{Kern}(f^\circ) \oplus \operatorname{Bild}(f^\circ). \tag{34.34d}$$

Beweis. Es sei $v \in \operatorname{Kern}(f)$. Dann gilt

$$0 = \gamma\big(f(v), f(v)\big) = \gamma\big(v, (f^\circ \circ f)(v)\big) = \gamma\big(v, f \circ (f^\circ)(v)\big) = \gamma\big(f \circ (f^\circ)(v), v\big) = \gamma\big(f^\circ(v), f^\circ(v)\big).$$

Also ist $f^\circ(v) = 0$, d. h., $v \in \operatorname{Kern}(f^\circ)$ und damit $\operatorname{Kern}(f) \subseteq \operatorname{Kern}(f^\circ)$. Wegen $f^{\circ\circ} = f$ folgt auch $\operatorname{Kern}(f^\circ) \subseteq \operatorname{Kern}(f)$. Das zeigt (34.34a).

Nach Satz 34.43 gilt

$$\operatorname{Bild}(f^\circ) = \operatorname{Kern}(f)^\perp = \operatorname{Kern}(f^\circ)^\perp = \operatorname{Bild}(f)^{\perp\perp} = \operatorname{Bild}(f).$$

Damit ist (34.34b) gezeigt.

Schließlich folgt aus Folgerung 34.44 und dem bereits Gezeigten:

$$V = \operatorname{Kern}(f) \oplus \operatorname{Bild}(f^\circ) = \operatorname{Kern}(f) \oplus \operatorname{Bild}(f) = \operatorname{Kern}(f^\circ) \oplus \operatorname{Bild}(f^\circ)$$

Damit sind auch (34.34c) und (34.34d) bestätigt. □

Lemma 34.51 (Eigenvektoren selbstadjungierter Endomorphismen)**.**
Es sei (V, γ) ein Euklidischer Raum. Ist $f \in \operatorname{End}(V)$ γ-selbstadjungiert und sind $\lambda_1, \lambda_2 \in \mathbb{R}$ verschiedene Eigenwerte von f mit Eigenvektoren v_1 bzw. v_2, dann sind (v_1, v_2) γ-orthogonal.

Beweis. Es gilt

$$\begin{aligned}
\lambda_1\, \gamma(v_1, v_2) &= \gamma(\lambda_1\, v_1, v_2) && \text{wegen der Linearität im ersten Argument} \\
&= \gamma(f(v_1), v_2) && \text{weil } (\lambda_1, v_1) \text{ ein Eigenpaar ist} \\
&= \gamma(v_1, f(v_2)) && \text{aufgrund der Selbstadjungiertheit} \\
&= \gamma(v_1, \lambda_2\, v_2) && \text{weil } (\lambda_2, v_2) \text{ ein Eigenpaar ist} \\
&= \lambda_2\, \gamma(v_1, v_2) && \text{wegen der Linearität im zweiten Argument.}
\end{aligned}$$

Aus $\lambda_1 \neq \lambda_2$ folgt $\gamma(v_1, v_2) = 0$. □

Bei der Herleitung von Normalformen allgemeiner Endomorphismen (wie in § 29 für die Frobenius-Normalform und in § 30 für die Jordan-Normalform) ging es immer darum, möglichst kleine invariante Unterräume U zu finden. Die Hürde dabei war, sicherzustellen, dass ein dazu komplementärer Unterraum existiert, der wiederum invariant ist, sodass das Argument wiederholt werden kann. Bei selbstadjungierten Endomorphismen gibt es diese Hürde nicht, denn für jeden invarianten Unterraum ist sein γ-orthogonales Komplement ebenfalls ein invarianter Unterraum:

Lemma 34.52 (γ-orthogonale Komplemente invarianter Unterräume γ-selbstadjungierter Endomorphismen sind invariant)**.**

(i) Es sei (V, γ) ein Euklidischer Raum. Weiter sei $f \in \operatorname{End}(V)$ γ-selbstadjungiert. Ist $U \subseteq V$ ein f-invarianter Unterraum, dann ist auch $U^\perp$ ein f-invarianter Unterraum.

(ii) Es sei $(\mathbb{R}^n, \gamma_M)$ ein Euklidischer Raum, also $M \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Weiter sei $A \in \mathbb{R}^{n \times n}$, $n \in \mathbb{N}_0$ eine γ_M-selbstadjungierte Matrix. Ist $U \subseteq \mathbb{R}^n$ ein A-invarianter Unterraum, dann ist auch $U^\perp$ ein A-invarianter Unterraum.

Beweis. Der Beweis ist Gegenstand von Hausaufgabe II-13.1. □

Definition 34.53 (Einheitssphäre, Rayleigh-Quotient)**.**
Es sei (V, γ) ein ~~endlich-dimensionaler~~ Euklidischer Raum.

(*i*) Die Menge
$$S(V, \gamma) \coloneqq \{v \in V \mid \|v\|_\gamma = 1\} \tag{34.35}$$
heißt die **Einheitssphäre** (englisch: unit sphere) des Raumes (V, γ).[21]

(*ii*) Für $f \in \operatorname{End}(V)$ γ-selbstadjungiert heißt die Funktion
$$R_{f,\gamma} \colon V \setminus \{0\} \ni v \mapsto \frac{\gamma\big(v, f(v)\big)}{\|v\|_\gamma^2} \in \mathbb{R} \tag{34.36}$$
der **Rayleigh-Quotient** (englisch: Rayleigh quotient) von f. △

Der folgende Satz stellt einen Zusammenhang zwischen den Werten des Rayleigh-Quotienten und den Eigenwerten von f her. Für den Beweis benötigen wir Resultate aus der Analysis über dem Körper $\mathbb{R}$ der reellen Zahlen.

Satz 34.54 (Maximum des Rayleigh-Quotienten)**.**
Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum mit $\dim(V) = n \in \mathbb{N}$. Weiter sei $f \in \operatorname{End}(V)$ γ-selbstadjungiert. Dann gilt:

(*i*) Der Rayleigh-Quotient ist auf $V \setminus \{0\}$ nach oben beschränkt und nimmt sein Supremum als Maximum an.

(*ii*) Das Maximum $m \coloneqq \max\big\{R_{f,\gamma}(v) \,\big|\, v \in V \setminus \{0\}\big\}$ des Rayleigh-Quotienten ist der größte Eigenwert von f. Die Vektoren $v \in V \setminus \{0\}$, für die $R_{f,\gamma}(v) = m$ gilt, sind genau die Eigenvektoren zum Eigenwert m, also
$$\{v \in V \setminus \{0\} \mid R_{f,\gamma} = m\} = \operatorname{Eig}(f, m) \setminus \{0\}.$$

Beweis. Aussage (*i*): Für alle $\alpha \in \mathbb{R}$, $\alpha \in 0$ gilt
$$R_{f,\gamma}(\alpha\, v) = \frac{\gamma\big(\alpha\, v, f(\alpha\, v)\big)}{\|\alpha\, v\|_\gamma^2} = \frac{\alpha^2\, \gamma\big(v, f(v)\big)}{|\alpha|^2\, \|v\|_\gamma^2} = \frac{\gamma\big(v, f(v)\big)}{\|v\|_\gamma^2} = R_{f,\gamma}(v).$$
wegen der Linearität von f, der Bilinearität von γ und der absoluten Homogenität der Norm.[22] Das Bild von $R_{f,\gamma}$ auf $V \setminus \{0\}$ ist daher das gleiche wie das Bild von $R_{f,\gamma}$ auf $S(V, \gamma)$.

Wir benötigen im Folgenden mehrere Resultate sowie den Begriff der Differenzierbarkeit einer Funktion $\mathbb{R} \to \mathbb{R}$ aus der reellen Analysis:

(1) Die Einheitssphäre $S(V, \gamma)$ in endlich-dimensionalen reellen normierten Vektorräumen ist kompakt (d. h. beschränkt und abgeschlossen).

(2) Jede stetige reellwertige Funktion nimmt über jeder nichtleeren, kompakten Teilmenge eines reellen normierten Vektorraumes ihr Maximum an. Dieses Resultat wird auch der Satz von Weierstraß genannt.

(3) Die Ableitung $\dot{r}$ einer differenzierbaren Funktion $r \colon \mathbb{R} \to \mathbb{R}$ ist an einem lokalen Maximum gleich 0.

[21]Im Unterschied dazu heißt $\{v \in V \mid \|v\|_\gamma \leqslant 1\}$ die **Einheitskugel** (englisch: unit ball) von (V, γ). Im Fall $n = 0$ ist $S(V, \gamma) = \emptyset$.
[22]Der Rayleigh-Quotient ist somit eine 0-homogene Funktion.

Der Rayleigh-Quotient ist stetig auf seiner Definitionsmenge $V \setminus \{0\}$, also auch auf der Einheitssphäre $S(V,\gamma)$. (**Quizfrage 34.3:** Klar?) Die Menge der Funktionswerte des Rayleigh-Quotienten

$$\big\{R_{f,\gamma}(v) \,\big|\, v \in V \setminus \{0\}\big\} = \big\{R_{f,\gamma}(v) \,\big|\, v \in S(V,\gamma)\big\},$$

nimmt nach dem Satz von Weierstraß auf der nichtleeren, kompakten Menge $S(V,\gamma)$ ihr Supremum M an, das damit zum Maximum wird (Definition 5.10).

Aussage *(ii)*: Es sei $v \in S(V,\gamma)$ einer der Maximierer von $R_{f,\gamma}$.

Schritt 1: Wir zeigen: Ist $d \in v^\perp$, dann gilt auch $\gamma\big(f(d), v\big) = 0$, also $f(d) \in v^\perp$.

Die Aussage ist klar für $d = 0$. Für $d \neq 0$ können wir o. B. d. A. im Beweis $\|d\|_\gamma = 1$ annehmen. (**Quizfrage 34.4:** Warum?)

Wir betrachten dazu den Punkt $u(t) = \cos(t)\, v + \sin(t)\, d$ für $t \in \mathbb{R}$. Für diesen gilt

$$\begin{aligned}
\|u(t)\|_\gamma^2 &= \gamma(\cos(t)\, v + \sin(t)\, d, \cos(t)\, v + \sin(t)\, d) \\
&= \cos^2(t)\, \gamma(v,v) + 2\, \cos(t)\, \sin(t)\, \gamma(v,d) + \sin^2(t)\, \gamma(d,d) \\
&= \cos^2(t)\, \underbrace{\|v\|_\gamma^2}_{=1} + \sin^2(t)\, \underbrace{\|d\|_\gamma^2}_{=1} \\
&= 1.
\end{aligned}$$

Also verläuft die Kurve $t \mapsto u(t)$ auf der Sphäre $S(V,\gamma)$ und bei $t = 0$ durch den Punkt v.

Wir betrachten nun die Funktionswerte des Rayleigh-Quotienten entlang der Kurve $u(t)$, also die Funktion $r(t) \coloneqq R_{f,\gamma}(u(t))$. Diese ist differenzierbar, und ihre Ableitung ist gegeben durch

$$\begin{aligned}
\dot r(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \gamma\big(u(t), f(u(t))\big) \\
&= \gamma\big(\dot u(t), f(u(t))\big) + \gamma\big(u(t), \dot f(u(t))\big) && \text{nach Produktregel} \\
&= \gamma\big(\dot u(t), f(u(t))\big) + \gamma\big(u(t), f(\dot u(t))\big) && \text{wegen der Linearität von } f \\
&= 2\, \gamma\big(f(\dot u(t)), u(t)\big) && \text{wegen der } \gamma\text{-Selbstadjungiertheit von } f.
\end{aligned}$$

Es gilt

$$\dot u(t) = -\sin(t)\, v + \cos(t)\, d,$$

also $\dot u(0) = d$.

Nach Voraussetzung hat r an der Stelle $t = 0$ ein globales und damit auch ein lokales Maximum. Daher gilt $\dot r(0) = 0$. Wir erhalten also wie gewünscht

$$\dot r(0) = 2\, \gamma\big(f(\dot u(0)), u(0)\big) = 2\, \gamma\big(f(d), v\big) = 0.$$

Schritt 2: Wir zeigen: v ist ein Eigenvektor zum Eigenwert $m \coloneqq \max\big\{R_{f,\gamma}(v) \,\big|\, v \in V \setminus \{0\}\big\}$.

In **Schritt** 1 haben wir gezeigt: $d \in v^\perp \;\Rightarrow\; f(d) \in v^\perp$. Das heißt aber: $v^\perp = \langle v \rangle^\perp$ ist ein f-invarianter Unterraum. Aus Lemma 34.52 folgt, dass auch $\langle v \rangle^{\perp\perp} = \langle v \rangle$ ein f-invarianter Unterraum sein muss. Das bedeutet aber, dass v ein Eigenvektor von f sein muss. Also gilt $f(v) = \lambda\, v$ für irgendein $\lambda \in \mathbb{R}$. Das impliziert schließlich

$$m = \gamma\big(v, f(v)\big) = \gamma(v, \lambda\, v) = \lambda\, \|v\|_\gamma^2 = \lambda. \qquad \square$$

Der folgende **Spektralsatz für γ-selbstadjungierte Endomorphismen** (englisch: spectral theorem for γ-selfadjoint endomorphisms), hier für Euklidische Räume, ist das Hauptresultat dieses Abschnitts.[23]

Satz 34.55 (selbstadjungierte Endomorphismen Euklidischer Räume sind orthogonal diagonalisierbar)**.**

Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Für $f \in \operatorname{End}(V)$ sind äquivalent:

(i) f ist γ-selbstadjungiert.

(ii) f ist diagonalisierbar, und es gibt es Basis aus Eigenvektoren von f, die γ-orthonormal ist. (Wir nennen f dann **γ-orthogonal diagonalisierbar** (englisch: γ-orthogonally diagonalizable).)

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Wir führen einen Induktionsbeweis nach $\dim(V) = n \in \mathbb{N}_0$. Im Fall $n = 0$ ist die Behauptung wahr, denn f ist die Nullabbildung, und die einzige Basis von V ist die leere Menge. Wir machen nun den Induktionsschritt von n auf $n + 1$. Wegen $n + 1 \geqslant 1$ ist Satz 34.54 anwendbar. Es folgt, dass $m := \max\{R_{f,\gamma}(v) \,|\, v \in V \setminus \{0\}\}$ ein Eigenwert von f ist (und zwar der größte). Der zugehörige Eigenraum ist f-invariant, also gilt

$$V = \operatorname{Eig}(f, m) \oplus \operatorname{Eig}(f, m)^\perp.$$

Dabei ist $U := \operatorname{Eig}(f, m)^\perp$ nach Lemma 34.52 wieder f-invariant und von echt kleinerer Dimension als V, da $\dim(\operatorname{Eig}(f, m)) \geqslant 1$ ist. Wir können also f zu $f|_U^U$ einschränken. Nach Induktionsvoraussetzung besitzt U eine γ-Orthonormalbasis aus Eigenvektoren von $f|_U^U$. Diese sind auch Eigenvektoren von f. Insgesamt erhalten wir also eine orthogonale direkte Summe

$$V = \operatorname{Eig}(f, \lambda_1) \oplus \cdots \oplus \operatorname{Eig}(f, \lambda_s)$$

mit den paarweise verschiedenen Eigenwerten λ_i von f. Innerhalb jedes Eigenraumes kann eine Basis γ-orthonormal gewählt werden. Die Orthogonalität der Basisvektoren über die verschiedenen Eigenräume hinweg ergibt sich aus Lemma 34.51.

Aussage (ii) $\Rightarrow$ Aussage (i): Es sei B_V eine γ-orthonormale Basis aus Eigenvektoren von f. Es ist also $D := \mathcal{M}_{B_V}^{B_V}(f)$ eine Diagonalmatrix und $M := \mathcal{M}_{B_V^*}^{B_V}(\gamma) = I$. Die Darstellungsmatrix des γ-adjungierten Endomorphismus ist nach (34.27) $D^\circ = M^{-1} D^\intercal M = I^{-1} D^\intercal I = D$. Damit ist f γ-selbstadjungiert. □

Bemerkung 34.56 (zur orthogonalen Diagonalisierbarkeit)**.**

(i) Es sei (V, γ) ein endlich-dimensionaler Euklidischer Raum. Die γ-orthogonale Diagonalisierbarkeit eines Endomorphismus auf V (oder einer Darstellungsmatrix) ist i. A. eine stärkere Eigenschaft als Diagonalisierbarkeit. Die Diagonalisierbarkeit alleine bedeutet, dass eine Basis aus Eigenvektoren existiert. Innerhalb eines Eigenraumes können dabei die Basisvektoren immer orthonormalisiert werden. Die γ-orthogonalen Diagonalisierbarkeit bedeutet, dass *alle* Eigenvektoren auch über die Eigenräume hinweg paarweise orthogonal gewählt werden können.

[23] Im Beweis dieses Satzes kommen über Satz 34.54 unweigerlich Argumente der Analysis ins Spielt. Ein alternativer Beweis von Satz 34.55 nutzt den Fundamentalsatz der Algebra 11.24, dessen Beweis ebenfalls auf analytischen Argumenten beruht.

(*ii*) Es sei V ein endlich-dimensionaler reeller Vektorraum. Ist ein Endomorphismus (oder eine Darstellungsmatrix) diagonalisierbar und ist $B_V = (v_1, \ldots, v_n)$ eine Basis von V aus lauter Eigenvektoren, dann können wir ein Innenprodukt γ auf V so wählen, dass f γ-orthogonal diagonalisierbar ist: Definieren wir γ durch

$$\gamma(v_i, v_j) := \delta_{ij},$$

also mit $\mathcal{M}_{B_V^*}^{B_V}(\gamma) = I$, dann ist (V, γ) ein Euklidischer Raum, und f ist γ-orthogonal diagonalisierbar. △

Die Übersetzung von Satz 34.55 in Matrizen lautet wie folgt:

Folgerung 34.57 (M-selbstadjungierte Matrizen sind M-orthogonal diagonalisierbar)**.**
Es sei $(\mathbb{R}^n, \gamma_M)$ ein Euklidischer Raum, also $M \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit. Weiter sei $A \in \mathbb{R}^{n\times n}$, $n \in \mathbb{N}_0$. Dann sind äquivalent:

(*i*) A ist M-selbstadjungiert, erfüllt also $M^{-1}A^\intercal M = A$.

(*ii*) A ist diagonalisierbar, und es gibt eine Basis $(v_1, \ldots, v_n)$ von $\mathbb{R}^n$ aus Eigenvektoren von A, die M-orthonormal ist. Es gibt also eine invertierbare Matrix $V \in \mathbb{R}^{n\times n}$, sodass $V^\intercal M\, V = I$ gilt und

$$AV = V\Lambda \tag{34.37}$$

mit der Diagonalmatrix $\Lambda \in \mathbb{R}^{n\times n}$ der Eigenwerte von A entsprechend ihrer Vielfachheit und zugehörigen M-orthonormalen Eigenvektoren $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$.

Wegen der M-Orthonormalität von V, also $V^\intercal M\, V = I$, gilt $V^{-1} = V^\intercal M$. Wir können daher (34.37) auch schreiben als

$$A = V\Lambda V^{-1} = V\Lambda V^\intercal M. \tag{34.38}$$

Ende der Vorlesung 25

§ 35 Innenprodukte in Vektorräumen über $K = \mathbb{C}$

In diesem Abschnitt betrachten wir **komplexe Vektorräume** (englisch: complex vector spaces), also solche über dem Körper $K = \mathbb{C}$. Im Unterschied zu den reellen Zahlen $\mathbb{R}$ ist $\mathbb{C}$ kein geordneter Körper. Die Definition 34.3 eines Innenprodukts muss also für komplexe Vektorräume angepasst werden. Der Körper der komplexen Zahlen enthält jedoch den geordneten Unterkörper $\mathbb{R} \subsetneq \mathbb{C}$, sodass man von einem Innenprodukt fordern kann, dass $\gamma(x, x) \in \mathbb{R}$ und $\gamma(x, x) \geqslant 0$ gilt. Mit einer Bilinearform γ ist das wegen

$$\gamma(i\, x, i\, x) = i^2\, \gamma(x, x) = -\gamma(x, x)$$

aber unmöglich. Für die Definition eines Innenprodukts in komplexen Vektorräumen verwendet man daher statt Bilinearformen sogenannte **Sesquilinearformen**. Für deren Definition benötigen wir zunächst den Begriff der **antilinearen Abbildung**.

Die Argumente im gesamten § 35 sind denen in § 34 so ähnlich, dass wir Beweise nur angeben, wo nennenswerte Abweichungen bestehen.

Definition 35.1 (antilineare Abbildung)**.**

Es seien V, W komplexe Vektorräume.

(i) Eine Abbildung $f\colon V \to W$ heißt **antilinear** (englisch: antilinear, lateinisch: anti: gegen) oder **konjugiert linear** (englisch: conjugate linear), wenn gilt:

$$f(u+v) = f(u) + f(v) \quad \text{für alle } u, v \in V, \tag{35.1a}$$

$$f(\alpha\, u) = \overline{\alpha}\, f(u) \quad \text{für alle } u \in V \text{ und alle } \alpha \in \mathbb{C}. \tag{35.1b}$$

Dabei bezeichnet $\overline{\alpha}$ die zu $\alpha \in \mathbb{C}$ komplex konjugierte Zahl, vgl. Anhang A.

Die Eigenschaft (35.1a) ist bereits als die **Additivität** (englisch: additivity) von f bekannt. Die Eigenschaft (35.1b) heißt die **Antihomogenität** (englisch: anti-homogeneity) oder **konjugierte Homogenität** (englisch: conjugate homogeneity) von f.

(ii) Antilineare Abbildungen werden auch als **antilineare Homomorphismen** (englisch: antilinear homomorphism) bezeichnet. Entsprechend gibt es auch **antilineare Endomorphismen**, **antilineare Isomorphismen** und **antilineare Automorphismen**. Diese bilden jeweils wieder Vektorräume, die wir mit $\operatorname{antiHom}(V, W)$, $\operatorname{antiEnd}(V)$, $\operatorname{antiIso}(V, W)$ bzw. $\operatorname{antiAut}(V)$ bezeichnen. △

Beachte: Den Begriff der linearen Abbildung gibt es zu jedem Paar von Vektorräumen über demselben Körper K. Der Begriff der **antilinearen Abbildung** ist jedoch nur für Vektorräume über dem Körper $K = \mathbb{C}$ vorgesehen.[24]

Beispiel 35.2 (antilineare Abbildung)**.**

(i) Die komponentenweise komplexe Konjugation in $\mathbb{C}^n$, $n \in \mathbb{N}_0$, also

$$\mathbb{C}^n \ni \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto \begin{pmatrix} \overline{v_1} \\ \vdots \\ \overline{v_n} \end{pmatrix} \in \mathbb{C}^n,$$

ist eine antilineare Abbildung. Wir bezeichnen diese auch mit $\overline{\cdot}\colon \mathbb{C}^n \to \mathbb{C}^n$. Insbesondere die komplexe Konjugation $\mathbb{C} \to \mathbb{C}$ ist eine antilineare Abbildung.

(ii) Für $A \in \mathbb{C}^{n \times m}$ ist die Matrix-Vektor-Multiplikation mit dem komplex konjugierten Vektor

$$\mathbb{C}^m \ni z \mapsto A\, \overline{z} \in \mathbb{C}^n$$

eine antilineare Abbildung. △

Wie eine lineare Abbildung ist auch eine antilineare Abbildung $V \to W$ ist durch die Bilder auf einer Basis von V bereits eindeutig festgelegt:

[24] Genauer: Man benötigt Vektorräume über einem Körper, der mit einer Involution ausgestattet ist, also einer bijektiven Abbildung $\iota\colon K \to K$ mit der Eigenschaft $\iota \circ \iota = \mathrm{id}_K$.

Satz 35.3 (Darstellungssatz für antilineare Abbildungen)**.**
Es seien V, W zwei endlich-dimensionale komplexe Vektorräume. 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 antilinearen Abbildung $f : V \to W$ eine eindeutig definierte Matrix $A \in \mathbb{C}^{n \times m}$ mit der Eigenschaft

$$f(v_j) = \sum_{i=1}^{n} a_{ij}\, w_i \quad \text{für alle } j = 1, \ldots, m. \tag{35.2}$$

Diese Matrix A heißt die **Darstellungsmatrix der antilinearen Abbildung f bzgl. der Basen B_V und B_W**, in Symbolen: $A = \mathcal{M}_{B_W}^{B_V}(f)$.

Die Darstellungsmatrix einer antilinearen Abbildung ist also von der einer linearen Abbildung (Satz 19.3) nicht zu unterscheiden. Ihre Wirkungen auf Koordinatenvektoren unterscheiden sich jedoch: Ist $v = \sum_{j=1}^{m} \alpha_j\, v_j$ die eindeutige Darstellung eines Vektors $v \in V$ und $A = \mathcal{M}_{B_W}^{B_V}(f)$, so gilt

$$f(v) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, \alpha_j\, w_i, \quad \text{wenn } f \text{ linear ist} \tag{35.3a}$$

$$f(v) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, \overline{\alpha_j}\, w_i, \quad \text{wenn } f \text{ antilinear ist.} \tag{35.3b}$$

Wir definieren nun den Typ von Abbildungen $V \times V \to \mathbb{C}$, der die Bilinearformen bei der Bildung von Innenprodukten auf komplexen Vektorräumen ersetzen wird.

Definition 35.4 (Sesquilinearform, hermitesche und schiefhermitesche Sesquilinearform, vgl. Definition 22.1)**.**
Es sei V ein komplexer Vektorraum.

(i) Eine Abbildung $\theta : V \times V \to \mathbb{C}$ heißt eine **Sesquilinearform** (englisch: sesquilinear form, lateinisch: sesqui: anderthalb) **auf $V \times V$** (oder einfach: **auf V**), wenn θ im ersten Argument antilinear und im zweiten Argument linear ist, wenn also gilt:

$$\theta(\alpha\, u + \beta\, v, w) = \overline{\alpha}\, \theta(u, w) + \overline{\beta}\, \theta(v, w) \tag{35.4a}$$

$$\theta(u, \alpha\, v + \beta\, w) = \alpha\, \theta(u, v) + \beta\, \theta(u, w) \tag{35.4b}$$

für alle u, v, w und alle $\alpha, \beta \in \mathbb{C}$.

Beachte: In der Literatur findet man auch Definition, bei denen θ im ersten Argument linear und im zweiten Argument antilinear ist.

(ii) Die Menge aller Sesquilinearformen $V \times V \to \mathbb{C}$ bezeichnen wir mit $\operatorname{Sesq}(V, V)$.

(iii) Wir nennen eine Sesquilinearform

$$\textbf{hermitesch}, \text{ wenn} \quad \theta(u, v) = \overline{\theta(v, u)} \quad \text{für alle } u, v \in V \tag{35.5a}$$

$$\textbf{schiefhermitesch}, \text{ wenn} \quad \theta(u, v) = -\overline{\theta(v, u)} \quad \text{für alle } u, v \in V \tag{35.5b}$$

gilt.

(iv) Die Menge aller hermiteschen Sesquilinearformen $V \times V \to \mathbb{C}$ bezeichnen wir mit $\operatorname{Sesq}_{\text{herm}}(V, V)$ und die der schiefhermiteschen Sesquilinearformen mit $\operatorname{Sesq}_{\text{skew}}(V, V)$.

(*v*) Zwei Vektoren $u, v \in V$ heißen **orthogonal bzgl. der hermiteschen Sesquilinearform** θ, wenn $\theta(u, v) = 0$ gilt. Wir schreiben dafür auch $u \perp_\theta v$ oder $u \perp v$. (**Quizfrage 35.1:** Ist Orthogonalität eine symmetrische Relation?) △

Die Menge $\operatorname{Sesq}(V, V)$ aller Sesquilinearformen $V \times V \to \mathbb{C}$ ist ein $\mathbb{C}$-Vektorraum, und $\operatorname{Sesq}_{\text{herm}}$ und $\operatorname{Sesq}_{\text{skew}}$ sind Unterräume von $\operatorname{Sesq}(V, V)$.

Das folgende Lemma besagt, dass für die Definition positiver definiter Sesquilinearformen aus komplexen Vektorräumen nur solche Sesquilinearformen in Frage kommen, die hermitesch sind. (Bei positiv definiten Bilinearformen auf reellen Vektorräumen war die positive Definitheit auch ohne Symmetrie möglich.)

Lemma 35.5 (genau die hermiteschen Sesquilinearformen sind auf der Diagonale reell)**.**
Es sei V ein komplexer Vektorraum. Für $\theta \in \operatorname{Sesq}(V, V)$ sind äquivalent:

(i) θ ist hermitesch.

(ii) $\theta(v, v) \in \mathbb{R}$ für alle $v \in V$.

Beweis. Aussage (i) $\Rightarrow$ Aussage (ii): Wenn θ hermitesch ist, erhalten wir $\theta(v, v) = \overline{\theta(v, v)}$. Also ist $\theta(v, v)$ reell.

Aussage (ii) $\Rightarrow$ Aussage (i): Umgekehrt sei $\theta(v, v)$ reell für alle $v \in V$. Wir fixieren Vektoren $u, v \in V$. Dann ist

$$\begin{aligned} \theta(u+v, u+v) &= \overbrace{\theta(u,u)}^{\in\mathbb{R}} + \theta(u,v) + \theta(v,u) + \overbrace{\theta(v,v)}^{\in\mathbb{R}} \\ \text{und} \quad \theta(u+i\,v, u+i\,v) &= \theta(u,u) + \theta(u,i\,v) + \theta(i\,v,u) + \theta(i\,v,i\,v) \\ &= \theta(u,u) + i\,\theta(u,v) - i\,\theta(v,u) - i^2\theta(v,v) \\ &= \underbrace{\theta(u,u)}_{\in\mathbb{R}} + i\,\theta(u,v) - i\,\theta(v,u) + \underbrace{\theta(v,v)}_{\in\mathbb{R}}. \end{aligned}$$

Setzen wir $\theta(u, v) = a + b\,i$ und $\theta(v, u) = c + d\,i$ mit $a, b, c, d \in \mathbb{R}$, so folgt aus $\theta(u + v, u + v) \in \mathbb{R}$, dass

$$\theta(u,v) + \theta(v,u) = (a+c) + (b+d)\,i \in \mathbb{R}$$

gilt, also $b = -d$. Weiter folgt aus $\theta(u + i\,v, u + i\,v) \in \mathbb{R}$, dass

$$i\,\theta(u,v) - i\,\theta(v,u) = (d-b) + (a-c)\,i \in \mathbb{R}$$

gilt, also $a = c$. Das heißt aber gerade $\theta(u, v) = \overline{\theta(v, u)}$. □

Definition 35.6 (Definitheit und Indefinitheit von Sesquilinearformen in komplexen Vektorräumen, vgl. Definition 34.1)**.**
Es sei V ein komplexer Vektorraum. Eine (notwendigerweise hermitesche) Sesquilinearform θ auf V heißt

(i) **positiv definit**, wenn $\theta(v, v) > 0$ gilt für alle $v \in V \setminus \{0\}$.

(ii) **positiv semidefinit**, wenn $\theta(v, v) \geqslant 0$ gilt für alle $v \in V$.

(iii) **negativ definit**, wenn $\theta(v, v) < 0$ gilt für alle $v \in V \setminus \{0\}$.

(iv) **negativ semidefinit**, wenn $\theta(v, v) \leqslant 0$ gilt für alle $v \in V$.

(*v*) **indefinit**, wenn θ weder positiv semidefinit noch negativ semidefinit ist, wenn es also Vektoren $v_1, v_2 \in V$ gibt mit $\theta(v_1, v_1) > 0$ und $\theta(v_2, v_2) < 0$. △

Definition 35.7 (Innenprodukt in einem komplexen Vektorraum, unitärer Raum, vgl. Definition 34.1)**.**
Es sei V ein komplexer Vektorraum. Die hermitesche Sesquilinearform $\theta \in \operatorname{Sesq}(V, V)$ heißt ein **Innenprodukt auf** V, wenn θ positiv definit ist.[25] In diesem Fall heißt (V, θ) auch ein **komplexer Innenproduktraum** (englisch: complex inner product space) oder ein **unitärer Raum** (englisch: unitary space). △

Beispiel 35.8 (Innenprodukt in einem komplexen Vektorraum)**.**

(*i*) Die durch die Einheitsmatrix $I \in \mathbb{C}^{n \times n}$ induzierte Sesquilinearform

$$\theta_I \colon \mathbb{C}^n \times \mathbb{C}^n \ni (x, y) \mapsto x^{\mathrm{H}} y = \sum_{i=1}^{n} \overline{x_i}\, y_i \in \mathbb{C}$$

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

(*ii*) Im Vektorraum $\mathbb{C}^{n \times m}$ ist

$$\theta(A, B) \coloneqq \operatorname{trace}(A^{\mathrm{H}} B)$$

ein Innenprodukt, genannt das **Frobenius-Innenprodukt** (englisch: Frobenius inner product auf $C^{n \times m}$). △

Definition 35.9 (Darstellungsmatrix einer Sesquilinearform, vgl. Definition 31.2)**.**
Es sei V ein endlich-dimensionaler komplexer Vektorraum und $\theta \in \operatorname{Sesq}(V, V)$. Weiter sei $B_V = (v_1, \ldots, v_n)$, $n \in \mathbb{N}_0$, eine Basis von V. Die Matrix

$$\mathcal{M}^{B_V}_{B_V^*}(\theta) = \big(\theta(v_i, v_j)\big)_{i,j=1}^{n} \in \mathbb{C}^{n \times n} \tag{35.6}$$

heißt die **Darstellungsmatrix** der Sesquilinearform θ bzgl. der Basis B_V. △

Eine Sesquilinearform $\theta \colon V \times V \to \mathbb{C}$ induziert durch

$$V \ni v \mapsto \theta(\cdot, v) \in V^*$$

eine antilineare Abbildung in $\operatorname{antiHom}(V, V^*)$. Deren Darstellungsmatrix ist in (35.6) gemeint.

Sind $u = \sum_{i=1}^{n} \alpha_i\, v_i$ und $v = \sum_{j=1}^{n} \beta_j\, v_j$ zwei beliebige Vektoren in V, so gilt

$$\theta(u, v) = \theta\Big(\sum_{i=1}^{n} \alpha_i\, v_i, \sum_{j=1}^{n} \beta_j\, v_j\Big) = \sum_{i=1}^{n} \overline{\alpha_i} \sum_{j=1}^{n} \beta_j\, \theta(u_i, v_j) = \begin{pmatrix} \overline{\alpha_1} \\ \vdots \\ \overline{\alpha_n} \end{pmatrix}^{\mathsf{T}} \mathcal{M}^{B_V}_{B_V^*}(\theta) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. \tag{35.7}$$

Definition 35.10 (konjugiert transponierte Matrix, hermitesch transponierte Matrix, vgl. Definition 15.24)**.**
Es sei $A \in \mathbb{C}^{n \times m}$, $n, m \in \mathbb{N}_0$. Die **konjugiert transponierte Matrix** (englisch: conjugate transpose matrix) oder **hermitesch transponierte Matrix** (englisch: Hermitian transpose matrix) **zu** $A = (a_{ij})$ ist die Matrix mit den Einträgen $\overline{a_{ji}}$. Wir bezeichnen sie mit $\overline{A}^{\mathsf{T}}$ oder kurz: $A^{\mathrm{H}} \in \mathbb{C}^{m \times n}$.[26] △

[25] $\theta \colon V \times V \to \mathbb{C}$ ist also sesquilinear, erfüllt $\theta(u, v) = \overline{\theta(v, u)}$ und $\theta(v, v) > 0$ für alle $v \in V \setminus \{0\}$.

[26] Die Notation für die konjugiert transponierte Matrix ist in der Literatur nicht einheitlich. Man findet auch Bezeichnungen wie A^* und $A^\dagger$.

Insbesondere können wir auch zu Zeilenvektoren in $\mathbb{C}_m$ und Spaltenvektoren in $\mathbb{C}^n$ die hermitsche „Matrix“ bilden. Die Beziehung (35.7) können wir also auch in der Form

$$\theta(u,v) = \alpha^{\mathrm{H}} \mathcal{M}_{B_V^*}^{B_V}(\theta)\,\beta$$

schreiben.

Lemma 35.11 (Rechenregeln für die konjugierte Transposition, vgl. Lemma 15.26)**.**
Für $A, B \in \mathbb{C}^{n\times m}$ und $C \in \mathbb{C}^{m\times \ell}$, $n, m, \ell \in \mathbb{N}_0$, und $\alpha \in \mathbb{C}$ gelten die folgenden Eigenschaften:

$$(A^{\mathrm{H}})^{\mathrm{H}} = A \qquad \text{konjugierte Transposition ist } \textbf{involutorisch} \tag{35.8}$$
$$(A+B)^{\mathrm{H}} = A^{\mathrm{H}} + B^{\mathrm{H}} \qquad \text{konjugierte Transposition ist additiv} \tag{35.9a}$$
$$(\alpha\, A)^{\mathrm{H}} = \overline{\alpha}\, A^{\mathrm{H}} \qquad \text{konjugierte Transposition ist antihomogen} \tag{35.9b}$$
$$(A\, C)^{\mathrm{H}} = C^{\mathrm{H}} A^{\mathrm{H}}. \tag{35.10}$$

Beachte: (35.9a) und (35.9b) zusammen bedeuten, dass die konjugierte Transposition eine antilineare Abbildung ist.

Definition 35.12 (hermitesche, schiefhermitesche Matrix, vgl. Definition 15.28)**.**
Es $A \in \mathbb{C}^{n\times n}$, $n \in \mathbb{N}_0$.

(*i*) A heißt **hermitesch** (englisch: Hermitian matrix), wenn $A = A^{\mathrm{H}}$ gilt.[27]

(*ii*) A heißt **antihermitesch** (englisch: anti-Hermitian matrix) oder **schiefhermitesch** (englisch: skew-Hermitian matrix), wenn $A = -A^{\mathrm{H}}$ gilt.

Die Menge der hermiteschen bzw. schiefhermiteschen $n \times n$-Matrizen bezeichnen wir mit $\mathbb{C}^{n\times n}_{\mathrm{herm}}$ bzw. $\mathbb{C}^{n\times n}_{\mathrm{skew}}$. △

Beachte: So wie symmetrische Matrizen in $K^{n\times n}$ die Darstellungsmatrizen symmetrischer Bilinearformen (auf Vektorräumen über beliebigen Körpern K) sind, so sind hermitesche Matrizen in $\mathbb{C}^{n\times n}$ die Darstellungsmatrizen hermitescher Sesquilinearformen auf komplexen Vektorräumen.

Man kann zeigen, dass (notwendigerweise hermitesche) Sesquilinearformen bzw. Matrizen invertierbar sind, vgl. Lemma 34.6.

Beispiel 35.13 (Sesquilinearform)**.**
Für jede Matrix $A \in \mathbb{C}^{n\times n}$ ist die Abbildung

$$\theta_A \colon \mathbb{C}^n \times \mathbb{C}^n \ni (x,y) \mapsto x^{\mathrm{H}} A\, y \in \mathbb{C}$$

eine Sesquilinearform auf $\mathbb{C}^n$. Wir nennen sie die **von A induzierte Sesquilinearform** auf $\mathbb{C}^n$ und bezeichnen sie auch mit θ_A. Die „Hermite-Eigenschaften“ von θ_A sind genau dieselben wie die von A. A ist die Darstellungsmatrix von θ_A bzgl. der Standardbasis $(e_1, \ldots, e_n)$ von $\mathbb{C}^n$. △

Da eine Sesquilinearform auf V ein Element von $\operatorname{antiHom}(V, V^*)$ darstellt, ergibt sich für die Darstellungsmatrizen eine bisher nicht verwendete Transformationsvorschrift.

[27]Mit anderen Worten: A ist ihre eigene hermitesch transponierte Matrix.

Satz 35.14 (Transformation der Darstellungsmatrix einer Sesquilinearform beim Wechsel der Basis, vgl. Satz 31.4)**.**
Es sei V ein endlich-dimensionaler komplexer Vektorraum. Weiter seien B_V und $\widehat{B}_V$ Basen von V. Dann gilt für die Darstellungsmatrix einer Sesquilinearform $\theta\colon V \times V \to K$:

$$\mathcal{M}^{\widehat{B}_V}_{\widehat{B}_V^*}(\theta) = \overline{\mathcal{T}^{B_V^*}_{\widehat{B}_V^*}}\, \mathcal{M}^{B_V}_{B_V^*}(\theta)\, \mathcal{T}^{\widehat{B}_V}_{B_V}. \tag{35.11}$$

Bezeichnen wir dabei wie üblich mit $T \coloneqq \mathcal{T}^{B_V}_{\widehat{B}_V}$ die Transformationsmatrix und mit $A \coloneqq \mathcal{M}^{B_V}_{B_V^*}(\theta)$ und $\widehat{A} \coloneqq \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_V^*}(\theta)$ die Darstellungsmatrizen, so können wir (35.11) schreiben als

$$\widehat{A} = T^{-\mathrm{H}} A\, T^{-1}.$$

Quizfrage 35.2: Ist klar, dass Transponieren, elementweises Konjugieren und Invertieren kommutieren, sodass wir einfach $T^{-\mathrm{H}}$ schreiben dürfen? Es gilt also $T^{-\mathrm{H}} = \overline{(T^{-1})^\intercal} = (\overline{T^{-1}})^\intercal = \overline{(T^\intercal)^{-1}} = (\overline{T^\intercal})^{-1} = (\overline{T}^\intercal)^{-1} = (\overline{T}^{\intercal})^{-1}$.

Definition 35.15 (hermitesche Kongruenztransformation, hermitesch kongruente Matrizen, vgl. Definition 31.5)**.**
Zwei Matrizen $A, \widehat{A} \in \mathbb{C}^{n\times n}$, $n \in \mathbb{N}_0$, heißen **hermitesch kongruent** (englisch: Hermitian congruent), wenn es eine invertierbare Matrix $T \in \mathbb{C}^{n\times n}$ gibt, sodass gilt:

$$\widehat{A} = T^{-\mathrm{H}} A\, T^{-1}. \tag{35.12}$$

Der Übergang von A zu $T^{-\mathrm{H}} A\, T^{-1}$ heißt auch eine **hermitesche Kongruenztransformation** (englisch: Hermitian congruence transformation) von A. △

Lemma 35.16 (Orthogonalität bzgl. eines Innenprodukts impliziert lineare Unabhängigkeit, vgl. Lemma 34.7)**.**
Es sei (V, θ) ein unitärer Raum. Ist $(v_1, \ldots, v_k)$ eine θ-orthogonale Familie von Vektoren in $V \setminus \{0\}$, $k \in \mathbb{N}_0$, dann ist $(v_1, \ldots, v_k)$ linear unabhängig.

§ 35.1 Innenprodukte induzieren Normen

Auch in komplexen Innenprodukträumen (unitären Räumen) induziert jedes Innenprodukt eine Norm. Auch hier gilt die Cauchy-Schwarz-Ungleichung:

Satz 35.17 (**Cauchy-Schwarz-Ungleichung**, vgl. Satz 34.8)**.**
Es sei (V, θ) ein unitärer Raum. Dann gilt:

$$\theta(u,v)\,\overline{\theta(u,v)} = |\theta(u,v)|^2 \leqslant \theta(u,u)\,\theta(v,v) \quad \text{oder äquivalent} \quad |\theta(u,v)| \leqslant \sqrt{\theta(u,u)}\,\sqrt{\theta(v,v)} \tag{35.13}$$

für alle $u, v \in V$. Gleichheit gilt in (35.13) genau dann, wenn (u, v) eine linear abhängige Familie ist.

Definition 35.18 (Norm auf einem komplexen Vektorraum, vgl. Definition 34.10)**.**
Es sei V ein Vektorraum über $\mathbb{C}$.

(*i*) Eine Abbildung $\|\cdot\| \colon V \to \mathbb{R}$ heißt eine **Norm** (englisch: norm) **auf** V, wenn gilt:

$$\|u\| \geqslant 0 \quad \text{und} \quad \|u\| = 0 \;\Rightarrow\; u = 0 \qquad \textbf{positive Definitheit} \tag{35.14a}$$

$$\|\alpha\, u\| = |\alpha|\, \|u\| \qquad \textbf{absolute Homogenität} \tag{35.14b}$$

$$\|u + v\| \leqslant \|u\| + \|v\| \qquad \textbf{Dreiecksungleichung} \text{ oder } \textbf{Subadditivität} \tag{35.14c}$$

für alle $u, v \in V$ und alle $\alpha \in \mathbb{C}$.

(*ii*) Das Paar $(V, \|\cdot\|)$ heißt ein **normierter komplexer Vektorraum** (englisch: normed complex vector space). △

Satz 35.19 (jedes Innenprodukt induziert eine Norm, vgl. Satz 34.11)**.**
Es sei (V, θ) ein unitärer Raum. Dann definiert

$$\|\cdot\|_\theta \colon V \ni u \mapsto \|u\|_\theta \coloneqq \sqrt{\theta(u, u)} \in \mathbb{R}_{\geqslant 0} \tag{35.15}$$

eine Norm auf V.

Beispiel 35.20 (jedes Innenprodukt induziert eine Norm)**.**

(*i*) Das Standardinnenprodukt auf $\mathbb{C}^n$ (Beispiel 35.8)

$$\theta_I \colon \mathbb{C}^n \times \mathbb{C}^n \ni (x, y) \mapsto x^{\mathrm{H}} y = \sum_{i=1}^{n} \overline{x_i}\, y_i \in \mathbb{C}$$

induziert die Norm (hier quadradiert)

$$\|x\|^2 \coloneqq \sum_{i=1}^{n} \overline{x_i}\, x_i = \sum_{i=1}^{n} |x_i|^2,$$

die manchmal die **Standardnorm** (englisch: standard norm) auf $\mathbb{C}^n$ genannt wird.

(*ii*) Die vom Frobenius-Innenprodukt auf $\mathbb{C}^{n \times m}$

$$\theta(A, B) \coloneqq \operatorname{trace}(A^{\mathrm{H}} B)$$

induzierte Norm ist (hier quadradiert)

$$\|A\|_F^2 \coloneqq \operatorname{trace}(A^{\mathrm{H}} A) = \sum_{i=1}^{n} \sum_{j=1}^{m} \overline{a_{ij}}\, a_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2,$$

genannt die **Frobenius-Norm** (englisch: Frobenius norm) auf $\mathbb{C}^{n \times m}$. △

Satz 35.21 (**Satz des Pythagoras**, vgl. Satz 34.12)**.**
Es sei (V, θ) ein unitärer Raum. Ist die Familie $(v_1, \ldots, v_k)$ orthogonal, dann gilt

$$\Big\| \sum_{i=1}^{k} v_i \Big\|^2 = \sum_{i=1}^{k} \|v_i\|^2. \tag{35.16}$$

Definition 35.22 (normierter Vektor, orthonormale Familie, Orthonormalbasis, vgl. Definition 34.13)**.**
Es sei (V, θ) ein unitärer Raum mit induzierter Norm $\|\cdot\|$.

(i) Ein Vektor $u \in V$ heißt ein **normierter Vektor** oder **Einheitsvektor**, wenn $\|u\| = 1$ gilt.

(ii) Eine Menge $E \subseteq V$ von Vektoren in V heißt **orthonormal bzgl. θ**, wenn die Menge orthogonal ist und alle Elemente normiert sind.

(iii) Eine Familie $(v_i)_{i \in I}$ von Vektoren in V heißt **orthonormal bzgl. θ**, wenn die Familie orthogonal ist und alle ihre Mitglieder normiert sind.

(iv) Eine Familie $B := (v_i)_{i \in I}$ von Vektoren in V heißt eine **Orthonormalbasis** von (V, θ), wenn B eine Basis von V und außerdem orthonormal bzgl. θ ist. △

Satz 35.23 (jeder endlich-dimensionale unitäre Raum besitzt eine Orthonormalbasis, vgl. Satz 34.14)**.**
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Dann gilt: V besitzt eine Orthonormalbasis B_V. Die Darstellungsmatrix erfüllt daher $\mathcal{M}_{B_V^*}^{B_V}(\theta) = I$.

§ 35.2 Orthogonale Projektion und Gram-Schmidt-Verfahren

Wie in Definition 34.15 können wir auch in unitären Räumen die orthogonale Projektion auf einen Vektor definieren:

Definition 35.24 (orthogonale Projektion auf einen Vektor)**.**
Es sei (V, θ) ein unitärer Raum und $u \in V$, $u \neq 0$. Dann ist durch

$$\operatorname{proj}_u^\theta \colon V \ni v \mapsto \operatorname{proj}_u^\theta(v) := \frac{\theta(v, u)}{\|u\|^2}\, u \in V \tag{35.17}$$

die **orthogonale Projektion** (englisch: orthogonal projection) **bzgl. θ auf $\langle u \rangle$** definiert.[28] △

Lemma 35.25 (Eigenschaften der orthogonalen Projektion auf einen Vektor)**.**
Es sei (V, θ) ein unitärer Raum und $u \in V$, $u \neq 0$. Dann gilt:

(i) Gleichung (35.17) definiert eine Projektion (Definition 22.33).

(ii) Es gilt $V = \operatorname{Bild}(\operatorname{proj}_u^\theta) \oplus \operatorname{Kern}(\operatorname{proj}_u^\theta)$. Jedes $v \in V$ besitzt also eine eindeutige Zerlegung $v = u_\parallel + u_\perp$ mit $u_\parallel \in \langle u \rangle$ und $u_\perp \in \langle u \rangle^\perp$.

Auch in unitären Räumen können wir das **Gram-Schmidt-Verfahren** mit und ohne Normiernug definieren, um eine gegebene linear unabhängige Familie von Vektoren in eine orthogonale bzw. orthonormale Familie $(u_1, \ldots, u_k)$ umzuwandeln.

Algorithmus 35.26 (Gram-Schmidt-Orthogonalisierung)**.**
Eingabe: V ein unitärer Vektorraum
Eingabe: θ ein Innenprodukt auf V
Eingabe: linear unabhängige Familie $(v_1, \ldots, v_k)$ in V
Ausgabe: orthogonale Familie $(u_1, \ldots, u_k)$ in V mit $\langle u_1, \ldots, u_k \rangle = \langle v_1, \ldots, v_k \rangle$
1: **for** $j = 1, \ldots, k$ **do**
2: Setze $u_j \leftarrow v_j - \sum_{i=1}^{j-1} \frac{\theta(v_j, u_i)}{\alpha_i} u_i$ // $u_j = v_j - \sum_{i=1}^{j-1} \operatorname{proj}_{u_i}^\theta(v_j)$

[28] Etwas ungenau spricht man auch von der „Projektion auf u“ statt von der Projektion auf den Unterraum $\langle u \rangle$.

Bestimme $\alpha_j := \|u_j\|^2$ // $\alpha_j > 0$
end for
return $(u_1, \dots, u_k)$

Durch eine zusätzliche Normierung der entstehenden Vektoren erhalten wir wieder eine orthonormale Familie:

Algorithmus 35.27 (Gram-Schmidt-Orthonormalisierung).
Eingabe: V ein unitärer Vektorraum
Eingabe: θ ein Innenprodukt auf V
Eingabe: linear unabhängige Familie $(v_1, \dots, v_k)$ in V
Ausgabe: orthonormale Familie $(u_1, \dots, u_k)$ in V mit $\langle u_1, \dots, u_k \rangle = \langle v_1, \dots, v_k \rangle$
for $j = 1, \dots, k$ **do**
Setze $u_j \leftarrow v_j - \sum_{i=1}^{j-1} \theta(v_j, u_i)\, u_i$ // $u_j = v_j - \sum_{i=1}^{j-1} \operatorname{proj}_{u_i}^{\theta}(v_j)$
Bestimme $\alpha_j := \|u_j\|^2$ // $\alpha_j > 0$
Setze $u_j \leftarrow \dfrac{u_j}{\sqrt{\alpha_j}}$ // $u_1, \dots, u_j$ erfüllen jetzt $\|u_i\| = 1$
end for
return $(u_1, \dots, u_k)$

Satz 35.28 (Eigenschaften des Gram-Schmidt-Verfahrens).
Es sei (V, θ) ein unitärer Raum. Ist $(u_1, \dots, u_k)$ eine linear unabhängige Familie in V, dann liefert Algorithmus 35.26 eine orthogonale Familie $(v_1, \dots, v_k)$ von V mit der Eigenschaft $\langle v_1, \dots, v_j \rangle = \langle u_1, \dots, u_j \rangle$ für alle $j = 1, \dots, k$. Dasselbe gilt für Algorithmus 35.27, wobei die Familie $(v_1, \dots, v_k)$ sogar orthonormal ist.

§ 35.3 Unitäre Abbildungen und die unitäre Gruppe

In diesem Abschnitt betrachten wir die Homomorphismen unitärer Räume, also lineare Abbildungen zwischen zwei unitären Vektorräumen, die mit den Innenprodukten verträglich sind. Während die Homomorphismen Euklidischer Räume (Definition 34.20) als **orthogonale Abbildungen** bezeichnet werden, heißen die Homomorphismen unitärer Räume **unitäre Abbildungen**.

Definition 35.29 (Homomorphismus unitärer Räume, vgl. Definition 34.20).
Es seien (V, θ_1) und (W, θ_2) zwei unitäre Räume. Eine Abbildung $f\colon V \to W$ heißt **unitär** (englisch: unitary map) oder eine **(lineare) Isometrie bzgl.** (θ_1, θ_2), wenn f ein Homormophismus der unitären Räume $(V, \theta_1) \to (W, \theta_2)$ ist, wenn also gilt:

$$f \text{ ist linear} \tag{35.18a}$$

$$\theta_2(f(u), f(v)) = \theta_1(u, v) \quad \text{für alle } u, v \in V. \tag{35.18b}$$

Statt **unitär bzgl.** (θ_1, θ_2) sagen wir auch kurz (θ_1, θ_2)**-unitär**. △

Satz 35.30 (Charakterisierung unitärer Abbildungen, vgl. Satz 34.21).
Es seien (V, θ_1) und (W, θ_2) zwei unitäre Räume. Weiter sei $f \in \operatorname{Hom}(V, W)$. Dann sind äquivalent:

(*i*) f ist (θ_1, θ_2)-unitär.

(ii) $\|f(v)\|_{\theta_2} = \|v\|_{\theta_1}$ für alle $v \in V$, d. h., f ist **längentreu**.

(iii) $\|f(v_1) - f(v_2)\|_{\theta_2} = \|v_1 - v_2\|_{\theta_1}$ für alle $v_1, v_2 \in V$, d. h., f ist **abstandstreu**.

(iv) Ist $(v_i)_{i\in I}$ eine orthonormale Familie in (V, θ_1), dann ist auch $(f(v_i))_{i\in I}$ eine orthonormale Familie in (W, θ_2).

(v) Ist v ein Einheitsvektor in (V, θ_1), dann ist $f(v)$ ein Einheitsvektor in (W, θ_2).

Lemma 35.31 (unitäre Abbildungen sind injektiv, vgl. Lemma 34.22)**.**
Es seien (V, θ_1) und (W, θ_2) zwei unitäre Räume.

(i) Ist $f \in \operatorname{Hom}(V, W)$ eine (θ_1, θ_2)-unitäre Abbildung, dann ist f injektiv.

(ii) Ist $f \in \operatorname{Hom}(V, W)$ eine (θ_1, θ_2)-unitäre Abbildung und gilt zusätzlich $\dim(V) = \dim(W) = n \in \mathbb{N}_0$, dann ist f bijektiv und f^{-1} ebenfalls eine bijektive unitäre Abbildung.

Lemma 35.32 (Komposition unitärer Abbildungen, vgl. Lemma 34.23)**.**
Es seien (U, θ_1), (V, θ_2) und (W, θ_3) unitäre Räume. Ist $f\colon V \to W$ eine (θ_2, θ_3)-unitäre Abbildung und $g\colon U \to V$ eine (θ_1, θ_2)-unitäre Abbildung, dann ist $f \circ g\colon U \to W$ eine (θ_1, θ_3)-unitäre Abbildung.

Lemma 35.33 (bijektive Isometrie in den Koordinatenraum, vgl. Lemma 34.24)**.**
Es sei (V, θ) ein unitärer Raum mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\theta)$ die Darstellungsmatrix von θ. Dann ist die Abbildung

$$\Phi_{B_V}\colon (\mathbb{C}^n, \theta_M) \to (V, \theta)$$

eine bijektive Isometrie. Das heißt, für alle $u, v \in V$ gilt

$$\theta(u, v) = x^{\mathrm{H}} M\, y \quad \text{mit den Koordinatenvektoren } x = \Phi_{B_V}^{-1}(u) \text{ und } y = \Phi_{B_V}^{-1}(v). \tag{35.19}$$

Beachte: Ist insbesondere B_V eine Orthonormalbasis von V, also $M = I$, dann ist Φ_{B_V} eine bijektive Isometrie zwischen $\mathbb{C}^n$ mit dem Standardinnenprodukt und (V, θ_I).

Wir untersuchen jetzt, wie wir die Unitarität eines Homomorphismus f zwischen unitären Räumen anhand einer Darstellungsmatrix erkennen können:

Lemma 35.34 (Unitarität in Darstellungsmatrizen von Homomorphismen, vgl. Lemma 34.25)**.**
Es seien (V, θ_1) und (W, θ_2) zwei endlich-dimensionale unitäre Räume mit $\dim(V) = m \in \mathbb{N}_0$ und $\dim(W) = n \in \mathbb{N}_0$.

(i) Es seien B_V und B_W Basen von V bzw. von W und $M_1 := \mathcal{M}_{B_V^*}^{B_V}(\theta_1) \in \mathbb{C}^{m\times m}$ sowie $M_2 := \mathcal{M}_{B_W^*}^{B_W}(\theta_2) \in \mathbb{C}^{n\times n}$ die Darstellungsmatrizen von θ_1 bzw. von θ_2. Für $f \in \operatorname{Hom}(V, W)$ sind äquivalent:

   (a) f ist (θ_1, θ_2)-unitär.

   (b) Die Darstellungsmatrix $A = \mathcal{M}_{B_W}^{B_V}(f) \in \mathbb{C}^{n\times m}$ erfüllt

$$A^{\mathrm{H}} M_2\, A = M_1. \tag{35.20}$$

(ii) Sind B_V und B_W sogar Orthonormalbasen von V bzw. von W, dann sind für $f \in \operatorname{Hom}(V, W)$ äquivalent:

(a) f ist (θ_1, θ_2)-unitär.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_W}^{B_V}(f) \in \mathbb{C}^{n \times m}$ erfüllt

$$A^{\mathrm{H}} I_n \, A = I_m. \tag{35.21}$$

Beachte: Wir schreiben in (35.21) $A^{\mathrm{H}} I_n \, A$ statt $A^{\mathrm{H}} A$, weil letztere Schreibweise von den Abbildungseigenschaften der dargestellten Objekte $A \mathrel{\widehat{=}} f \in \mathrm{Hom}(V, W)$ und $A^{\mathsf{T}} \mathrel{\widehat{=}} f^* \in \mathrm{Hom}(W^*, V^*)$ nicht passt. Es fehlt der Übergang durch I_n, das den Riesz-Isomorphismus $W \to W^*$ darstellt. In (35.21) taucht A^{H} statt A^{T} auf, weil die Antilinearität im „von links kommenden" Argument benötigt wird.

Folgerung 35.35 (Unitarität in Darstellungsmatrizen von Endomorphismen, vgl. Folgerung 34.26)**.** Es sei (V, θ) ein endlich-dimensionaler unitärer Raum mit $\dim(V) = n \in \mathbb{N}_0$.

(i) Es sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\theta) \in \mathbb{C}^{n \times n}$ die Darstellungsmatrix von θ. Für $f \in \mathrm{End}(V)$ sind äquivalent:

(a) f ist θ-unitär.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{C}^{n \times n}$ erfüllt

$$A^{\mathrm{H}} M \, A = M. \tag{35.22}$$

(ii) Ist B_V sogar eine Orthonormalbasis von V, dann sind für $f \in \mathrm{End}(V)$ äquivalent:

(a) f ist θ-unitär.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{C}^{n \times n}$ erfüllt

$$A^{\mathrm{H}} I_n \, A = I_n. \tag{35.23}$$

Definition 35.36 (unitäre Matrix, vgl. Definition 34.27)**.**

(i) Es seien $m, n \in \mathbb{N}_0$, $M_1 \in \mathbb{C}^{m \times m}$ und $M_2 \in \mathbb{C}^{n \times n}$ beide hermitesch und positiv definit.[29] Eine Matrix $A \in \mathbb{C}^{n \times m}$ heißt **(M_1, M_2)-unitär** (englisch: (M_1, M_2)-unitär matrix), wenn (35.20) gilt, also

$$A^{\mathrm{H}} M_2 \, A = M_1.$$

(ii) Es sei $n \in \mathbb{N}_0$ und $M \in \mathbb{C}^{n \times n}$ hermitesch und positiv definit. Eine Matrix $A \in \mathbb{C}^n$ heißt **M-unitär** (englisch: M-unitär matrix), wenn (35.20) gilt, also

$$A^{\mathrm{H}} M \, A = M. \qquad \triangle$$

Bemerkung 35.37 (zum Begriff der Unitarität von Matrizen, Bemerkung 34.28)**.**

(i) (M_1, M_2)-Unitarität ist eine Eigenschaft, die für Darstellungsmatrizen von Homomorphismen zwischen unitären Räumen erklärt ist.[30]

[29] M_1 induziert also ein Innenprodukt auf $\mathbb{C}^m$, und M_2 induziert eines auf $\mathbb{C}^n$.

[30] Die (M_1, M_2)-Unitarität von A bedeutet, dass für jede Wahl von Basen B_V und B_W von Vektorräumen V und W passender Dimension der durch A dargestellte Homomorphismus unitär bzgl. der durch M_1 bzw. M_2 auf V bzw. W induzierten Innenprodukte ist. Gemeint ist: $\theta_1(u, v) = [\Phi_{B_V}^{-1}(u)]^{\mathrm{H}} M_1 \, [\Phi_{B_V}^{-1}(v)]$.

(*ii*) M-Unitarität ist eine Eigenschaft, die für Darstellungsmatrizen von Endomorphismen eines unitären Raumes erklärt ist.[31] △

Lemma 35.38 (Eigenwerte unitärer Endomorphismen, vgl. Lemma 34.29).
Es sei (V, θ) ein unitärer Raum. Ist $f \in \operatorname{End}(V)$ θ-unitär, dann gilt $|\lambda| = 1$ für alle $\lambda \in \Lambda(f)$.

Beachte: Hat der Endomorphismus f des komplexen Vektorraumes V einen Eigenwert vom Betrag ungleich 1, so ist er also in keinem Innenprodukt θ von V unitär.

Beweis. Es sei $\lambda \in \mathbb{C}$ ein Eigenwert von f und $v \in V$ ein zugehöriger Eigenvektor. Dann gilt

$$\|v\| = \|f(v)\| = \|\lambda\, v\| = |\lambda|\, \|v\|.$$

Da $v \neq 0$ und damit $\|v\| \neq 0$ ist, muss $|\lambda| = 1$ gelten. □

Lemma 35.39 (unitäre Endomorphismen bilden eine Gruppe, vgl. Lemma 34.30).
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Dann gilt:

(*i*) Die θ-unitären Endomorphismen von (V, θ) bilden eine Gruppe bzgl. der Komposition, genannt die **unitäre Gruppe** (englisch: unitary group) des unitären Raumes (V, θ):

$$\operatorname{U}(V, \theta) \coloneqq \{f \in \operatorname{End}(V) \mid f \text{ ist } \theta\text{-unitär}\}. \tag{35.24}$$

(*ii*) Die θ-unitären Endomorphismen $f \in \operatorname{U}(V, \theta)$ mit $\det(f) = 1$ bilden einen Normalteiler von $\operatorname{U}(V, \theta)$, genannt die **spezielle unitäre Gruppe** (englisch: special unitary group) des unitären Raumes (V, θ):

$$\operatorname{SU}(V, \theta) \coloneqq \{f \in \operatorname{U}(V, \theta) \mid \det(f) = 1\}. \tag{35.25}$$

Die Aussage von Lemma 35.39 übersetzt sich wie folgt in Matrizen:

Folgerung 35.40 (Darstellungsmatrizen unitärer Endomorphismen bilden eine Gruppe, vgl. Folgerung 34.31).
Es sei $n \in \mathbb{N}_0$ und $M \in \mathbb{C}^{n \times n}$ hermitesch und positiv definit. Dann gilt:

(*i*) Die M-unitären Matrizen in $\mathbb{C}^{n \times n}$ bilden eine Gruppe bzgl. der Matrix-Multiplikation, genannt die **unitäre Gruppe** des unitären Raumes $(\mathbb{C}^n, \theta_M)$:

$$\operatorname{U}(\mathbb{C}^n, \theta_M) \coloneqq \{A \in \mathbb{C}^{n \times n} \mid A \text{ ist } M\text{-unitär}\}. \tag{35.26}$$

(*ii*) Die M-unitären Matrizen $A \in \mathbb{C}^{n \times n}$ mit $\det(A) = 1$ bilden einen Normalteiler von $\operatorname{U}(\mathbb{C}^n, \theta_M)$, genannt die **spezielle unitäre Gruppe** des unitären Raumes $(\mathbb{C}^n, \theta_M)$:

$$\operatorname{SU}(\mathbb{C}^n, \theta_M) \coloneqq \{A \in \operatorname{U}(\mathbb{C}^n, \theta_M) \mid \det(A) = 1\}. \tag{35.27}$$

[31] Die M-Unitarität von A bedeutet, dass für jede Wahl einer Basis B_V eines Vektorraumes V passender Dimension der durch A dargestellte Endomorphismus unitär bzgl. des durch M auf V induzierten Innenprodukts ist.

Bemerkung 35.41 (zum Begriff der Unitarität, vgl. Bemerkung 34.32).
In Lehrbüchern wird der Begriff der Unitarität für Matrizen in der Regel nur für den Fall $M = I$ beschrieben, also für das Standardinnenprodukt auf $\mathbb{C}^n$. Das heißt, der Begriff der **orthogonalen Matrix** wird dort ausschließlich im Sinne von $A^{\mathrm{H}} I A = A^{\mathrm{H}} A = I$ verwendet. Dieser Fall bedeutet, dass für B_V stets eine Orthonormalbasis gewählt wird.

Die unitäre Gruppe von $\mathrm{U}(\mathbb{C}^n, \kappa_I)$ wird dann oft einfach mit $\mathrm{SO}(n)$ bezeichnet. Wir folgen dieser Praxis nicht. Unsere Notation $\mathrm{U}(\mathbb{C}^n, \kappa_M)$ stellt klar, dass es um die unitäre Gruppe des unitären Raumes $(\mathbb{C}^n, \kappa_M)$ geht. △

§ 35.4 Die Riesz-Abbildung und adjungierte Homomorphismen

Auch in unitären Vektorräumen (V, θ) induziert das Innenprodukt θ ein Innenprodukt θ^{-1} im Dualraum V^*, sodass der antilineare Vektorraum-Isomorphismus $V \ni u \mapsto \theta(u, \cdot) \in V^*$ unitär wird:

Satz 35.42 (**Darstellungssatz von Riesz**, vgl. Satz 34.34).
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Dann ist die Abbildung

$$\Theta \colon V \ni u \mapsto \theta(u, \cdot) \in V^* \tag{35.28}$$

ein antilinearer Isomorphismus. Wird V^* mit dem Innenprodukt θ^{-1} ausgestattet, dann ist $\Theta \colon (V, \theta) \to (V^*, \theta^{-1})$ eine bijektive Isometrie. Es gilt

$$\theta(u, v) = \langle \Theta(u)\, , v \rangle = \overline{\langle \Theta(v)\, , u \rangle} = \theta^{-1}\big(\Theta(u), \Theta(v)\big) \tag{35.29}$$

für alle $u, v \in V$.

Definition 35.43 (Riesz-Isomorphismus, vgl. Definition 34.35).
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Die durch (35.28) definierte Abbildung, also

$$\Theta_{V \to V^*} \colon V \ni u \mapsto \theta(u, \cdot) \in V^*$$

heißt der **Riesz-Isomorphismus von V nach V^***. △

Folgerung 35.44 (Rieszscher Darstellungssatz für $(\mathbb{C}^n, \theta_M)$, vgl. Folgerung 34.36).
Es sei $n \in \mathbb{N}_0$ und $M \in \mathbb{C}^{n \times n}$ hermitesch und positiv definit. Dann ist die Abbildung

$$\Theta \colon \mathbb{C}^n \ni x \mapsto (x^{\mathrm{H}} M)^{\mathsf{T}} = M^{\mathsf{T}}\, \overline{x} \in (\mathbb{C}^n)^* \tag{35.30}$$

eine bijektive antilineare Isometrie der unitären Räume $(\mathbb{C}^n, \theta_M)$ und $\big((\mathbb{C}^n)^*, \theta_{M^{-1}}\big)$. Es gilt

$$x^{\mathrm{H}} M\, y = \underbrace{(x^{\mathrm{H}} M)\, y}_{= \langle M^{\mathsf{T}} \overline{x}\, , y \rangle} = \underbrace{\overline{(y^{\mathrm{H}} M)\, x}}_{= \overline{\langle M^{\mathsf{T}} \overline{y}\, , x \rangle}} = (x^{\mathrm{H}} M)\, M^{-1} (M\, y) \tag{35.31}$$

für alle $x, y \in \mathbb{C}^n$.

Satz 35.45 (Unitarität dualer Homomorphismen, vgl. Satz 34.37).
Es seien (V, θ_1) und (W, θ_2) zwei endlich-dimensionale unitäre Räume derselben Dimension. Weiter sei $f \in \mathrm{Hom}(V, W)$. Dann sind äquivalent:

(*i*) f ist (θ_1, θ_2)-unitär.

(*ii*) f^* ist $(\theta_2^{-1}, \theta_1^{-1})$-unitär.

Als Spezialfall von Satz 35.45 ergibt sich sofort:

Folgerung 35.46 (Unitarität dualer Endomorphismen, vgl. Folgerung 34.38)**.**
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Weiter sei $f \in \operatorname{End}(V)$. Dann sind äquivalent:

(*i*) f ist θ-unitär.

(*ii*) f^* ist θ^{-1}-unitär.

Wir definieren nun die **Adjungierte** f° einer linearen Abbildung f zwischen den unitären Räumen (V, θ_1) und (W, θ_2). Die Definition ergibt sich aus der Forderung, dass das folgende Diagramm kommutiert:

$$\begin{tikzcd}
(V, \theta_1) \arrow[d, "\Theta_{V \to V^*}"'] & (W, \theta_2) \arrow[l, "f^\circ"'] \arrow[d, "\Theta_{W \to W^*}"] \\
(V^*, \theta_1^{-1}) & (W^*, \theta_2^{-1}) \arrow[l, "f^*"']
\end{tikzcd}$$

Beachte: Die Riesz-Isomorphismen $\Theta_{W \to W^*}$ und $\Theta_{V \to V^*}^{-1}$ sind beide antilinear, die Komposition $\Theta_{V \to V^*}^{-1} \circ f^* \circ \Theta_{W \to W^*}$ ist aber wieder linear.

Definition 35.47 (adjungierter Homomorphismus, vgl. Definition 34.39)**.**
Es seien (V, θ_1) und (W, θ_2) zwei endlich-dimensionale unitäre Räume. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung. Dann heißt

$$f^\circ := \Theta_{V \to V^*}^{-1} \circ f^* \circ \Theta_{W \to W^*} \colon W \to V \tag{35.32}$$

der zu f **(θ_1, θ_2)-adjungierte Homomorphismus** bzw. die zu f **(θ_1, θ_2)-adjungierte Abbildung**. △

Die Definition (35.32) lautet ausgeschrieben

$$\theta_2\big(w, f(v)\big) = \theta_1\big(f^\circ(w), v\big) \tag{35.33}$$

für alle $v \in V$ und alle $w \in W$.

Bemerkung 35.48 (Abbildungseigenschaften rund um die adjungierte Abbildung)**.**

(*i*) Für $f \in \operatorname{Hom}(V, W)$ ist auch f° wieder linear, also $f^\circ \in \operatorname{Hom}(W, V)$.

(*ii*) Die Zuordnung
$$\operatorname{Hom}(V, W) \ni f \mapsto f^* \in \operatorname{Hom}(W^*, V^*)$$
ist ebenfalls linear.

(*iii*) Die Zuordnung
$$\operatorname{Hom}(V, W) \ni f \mapsto f^\circ \in \operatorname{Hom}(W, V)$$
ist jedoch antilinear. △

Satz 35.49 (Darstellungsmatrix des (θ_1, θ_2)-adjungierten Homomorphismus, vgl. Satz 34.41).
Es seien (V, θ_1) und (W, θ_2) zwei endlich-dimensionale unitäre Räume. Es seien B_V und B_W Basen von V bzw. von W und $M_1 := \mathcal{M}_{B_V^*}^{B_V}(\theta_1) \in \mathbb{R}^{m \times m}$ sowie $M_2 := \mathcal{M}_{B_W^*}^{B_W}(\theta_2) \in \mathbb{R}^{n \times n}$ die Darstellungsmatrizen von θ_1 bzw. von θ_2. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung. Ist $A = \mathcal{M}_{B_W}^{B_V}(f)$ die Darstellungsmatrix von f bzgl. der Basen B_V und B_W, dann ist

$$A^\circ := \mathcal{M}_{B_V}^{B_W}(f^\circ) = M_1^{-1} A^{\mathrm{H}} M_2 \tag{35.34}$$

die Darstellungsmatrix von $f^\circ \in \operatorname{Hom}(W, V)$.

Lemma 35.50 ((θ_1, θ_2)-biadjungierte Abbildung, vgl. Lemma 34.42).
Es seien (V, θ_1) und (W, θ_2) zwei endlich-dimensionale unitäre Räume. Dann ist die **(θ_1, θ_2)-biadjungierte Abbildung** $f^{\circ\circ} := (f^\circ)^\circ$ identisch zu f, unabhängig von den Innenprodukten θ_1 und θ_2.

§ 35.5 Nochmal vier fundamentale Unterräume zu einer linearen Abbildung

Satz 35.51 (vier fundamentale Unterräume zu einer linearen Abbildung, vgl. Satz 34.43).
Es seien (V, θ_1) und (W, θ_2) zwei endlich-dimensionale unitäre Räume. Weiter sei $f \in \operatorname{Hom}(V, W)$ eine lineare Abbildung mit (θ_1, θ_2)-adjungierter Abbildung $f^\circ \in \operatorname{Hom}(W, V)$. Dann gilt:

$$\operatorname{Bild}(f^\circ) = \operatorname{Kern}(f)^\perp \quad \text{in } V \tag{35.35a}$$
$$\operatorname{Kern}(f^\circ) = \operatorname{Bild}(f)^\perp \quad \text{in } W \tag{35.35b}$$
$$\operatorname{Bild}(f) = \operatorname{Kern}(f^\circ)^\perp \quad \text{in } W \tag{35.35c}$$
$$\operatorname{Kern}(f) = \operatorname{Bild}(f^\circ)^\perp \quad \text{in } V. \tag{35.35d}$$

Die orthogonalen Komplemente in V sind bzgl. θ_1 und in W bzgl. θ_2 zu verstehen.

Folgerung 35.52 (Endomorphismen induzieren orthogonale direkte Zerlegungen, vgl. Folgerung 34.44).

Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Weiter sei $f \in \operatorname{End}(V)$ eine lineare Abbildung mit θ-adjungierter Abbildung $f^\circ \in \operatorname{End}(V)$. Dann gilt:

$$V = \operatorname{Kern}(f) \ominus \operatorname{Bild}(f^\circ) \tag{35.36a}$$
$$V = \operatorname{Kern}(f^\circ) \ominus \operatorname{Bild}(f). \tag{35.36b}$$

§ 35.6 Selbstadjungierte, unitäre und normale Endomorphismen

In § 34.6 hatten wir gesehen, dass γ-selbstadjungierte Endomorphismen Euklidischer Räume (Satz 34.55) diagonalisierbar sind, wobei die Basis aus Eigenvektoren sogar γ-orthonormal gewählt werden können. Wir werden in diesem Abschnitt zeigen, dass in unitären Räumen diese Rolle von normalen Endomorphismen übernommen wird, also von einer größeren Klasse von Abbildungen.

Definition 35.53 (selbstadjungierte und normale Endomorphismen, vgl. Definition 34.45).
Es sei (V, θ) ein unitärer Raum.

(*i*) $f \in \operatorname{End}(V)$ heißt **θ-selbstadjungiert**, wenn $f = f^\circ$ gilt.

(*ii*) $f \in \operatorname{End}(V)$ heißt **θ-normal**, wenn $f \circ f^\circ = f^\circ \circ f$ gilt. △

Selbstadjungiertheit bzw. Normalität eines Endomorphismus eines unitären Raumes zeigt sich anhand einer Darstellungsmatrix wie folgt:

Lemma 35.54 (Selbstadjungiertheit in Darstellungsmatrizen von Endomorphismen, vgl. Lemma 34.46)**.**

Es sei (V, θ) ein endlich-dimensionaler unitärer Raum mit $\dim(V) = n \in \mathbb{N}_0$.

(i) Es sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\theta) \in \mathbb{C}^{n\times n}$ die Darstellungsmatrix von θ. Für $f \in \operatorname{End}(V)$ sind äquivalent:

(a) f ist θ-selbstadjungiert.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{C}^{n\times n}$ erfüllt $A^\circ = A$, also

$$M^{-1}A^{\mathrm{H}}M = A. \tag{35.37}$$

(ii) Ist B_V sogar eine Orthonormalbasis von V, dann sind für $f \in \operatorname{End}(V)$ äquivalent:

(a) f ist θ-selbstadjungiert.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{C}^{n\times n}$ erfüllt

$$I_n^{-1}A^{\mathrm{H}}I_n = A. \tag{35.38}$$

Beachte: Wir schreiben in (35.38) $I_n^{-1}A^{\mathrm{H}}I_n$ statt A^{H}, weil letztere Schreibweise von den Abbildungseigenschaften der dargestellten Objekte $A \mathrel{\widehat{=}} f \in \operatorname{End}(V)$ und $A^\intercal \mathrel{\widehat{=}} f^* \in \operatorname{End}(V^*)$ nicht passt. Es fehlt der Übergang durch I_n und I_n^{-1}, das den Riesz-Isomorphismus $(V, \theta) \to (V^*, \theta^{-1})$ bzw. den inversen Riesz-Isomorphismus darstellt. In (35.38) taucht A^{H} statt $A^\intercal$ auf, weil die Antilinearität im „von links kommenden" Argument benötigt wird.

Lemma 35.55 (Normalität in Darstellungsmatrizen von Endomorphismen, vgl. Lemma 34.47)**.**
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum mit $\dim(V) = n \in \mathbb{N}_0$.

(i) Es sei B_V eine Basis von V und $M := \mathcal{M}_{B_V^*}^{B_V}(\theta) \in \mathbb{C}^{n\times n}$ die Darstellungsmatrix von θ. Für $f \in \operatorname{End}(V)$ sind äquivalent:

(a) f ist θ-normal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{C}^{n\times n}$ erfüllt $A\,A^\circ = A^\circ\,A$, also

$$A\,M^{-1}A^{\mathrm{H}}M = M^{-1}A^{\mathrm{H}}M\,A. \tag{35.39}$$

(ii) Ist B_V sogar eine Orthonormalbasis von V, dann sind für $f \in \operatorname{End}(V)$ äquivalent:

(a) f ist θ-normal.

(b) Die Darstellungsmatrix $A = \mathcal{M}_{B_V}^{B_V}(f) \in \mathbb{C}^{n\times n}$ erfüllt

$$A\,I_n^{-1}A^{\mathrm{H}}I_n = I_n^{-1}A^{\mathrm{H}}I_n\,A. \tag{35.40}$$

Beachte: Wir schreiben auch hier in (35.40) wieder I_n und I_n^{-1} explizit, um die Abbildungseigenschaften der dargestellten Objekte erkennbar wiederzugeben.

Lemma 35.56 (Zusammenhang zwischen Unitarität, Selbstadjungiertheit und Normalität, vgl. Lemma 34.48)**.**
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum.

(i) Ist $f \in \operatorname{End}(V)$ θ-selbstadjungiert, dann ist f auch θ-normal.

(ii) Ist $f \in \operatorname{End}(V)$ θ-orthogonal, dann ist f auch θ-normal.

Folgerung 35.57 (normale Endomorphismen induzieren orthogonale direkte Zerlegungen, vgl. Folgerung 34.50)**.**
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Weiter sei $f \in \operatorname{End}(V)$ θ-normal. Dann gilt:

$$\operatorname{Kern}(f^\circ) = \operatorname{Kern}(f) \tag{35.41a}$$
$$\operatorname{Bild}(f^\circ) = \operatorname{Bild}(f) \tag{35.41b}$$
$$V = \operatorname{Kern}(f) \ominus \operatorname{Bild}(f). \tag{35.41c}$$
$$V = \operatorname{Kern}(f^\circ) \ominus \operatorname{Bild}(f^\circ). \tag{35.41d}$$

In Erweiterung von Lemma 34.51 gilt für θ-selbstadjungierte Endomorphismen:

Lemma 35.58 (Eigenwerte und Eigenvektoren selbstadjungierter Endomorphismen, vgl. Lemma 34.51)**.**

Es sei (V, θ) ein endlich-dimensionaler unitärer Raum mit $\dim(V) = n \in \mathbb{N}_0$. Weiter sei $f \in \operatorname{End}(V)$ θ-selbstadjungiert.

(i) Die Eigenwerte von f sind reell.

(ii) Sind $\lambda_1, \lambda_2 \in \mathbb{R}$ verschiedene Eigenwerte von f mit Eigenvektoren v_1 bzw. v_2, dann sind (v_1, v_2) θ-orthogonal.

Beweis. Aussage (i): Es sei (λ, v) ein Eigenpaar von f. Dann gilt

$$\begin{aligned}
\lambda\, \theta(v, v) &= \theta(v, \lambda\, v) && \text{wegen der Linearität im zweiten Argument} \\
&= \theta(v, f(v)) && \text{denn } (\lambda, v) \text{ ist ein Einpaar von } f \\
&= \theta(f(v), v) && \text{wegen der } \theta\text{-Selbstadjungiertheit von } f \\
&= \theta(\lambda\, v, v) && \text{denn } (\lambda, v) \text{ ist ein Einpaar von } f \\
&= \overline{\lambda}\, \theta(v, v) && \text{wegen der Antilinearität im ersten Argument.}
\end{aligned}$$

Da $\theta(v, v) = \|v\|^2 \neq 0$ ist, muss $\lambda = \overline{\lambda}$ gelten, also folgt $\lambda \in \mathbb{R}$.

Aussage (ii): Der Beweis ist identisch zum Beweis von Lemma 34.51. □

Wir untersuchen nun den Zusammenhang zwischen der θ-Normalität eines Endomorphismus f eines unitären Vektorraumes und den Eigenwerten von f und f°.

Lemma 35.59 (Eigenwerte normaler Endomorphismen)**.**
Es sei (V, θ) ein endlich-dimensionaler unitärer Raum. Weiter sei $f \in \operatorname{End}(V)$ θ-normal sowie $\lambda \in \mathbb{C}$. Dann gilt:

(i) $\lambda \operatorname{id}_V - f$ ist ebenfalls θ-normal.

(*ii*) $\operatorname{Eig}(f,\lambda) = \operatorname{Eig}(f^\circ,\overline{\lambda})$. Insbesondere ist $\lambda \in \mathbb{C}$ genau dann ein Eigenwert von f, wenn $\overline{\lambda}$ ein Eigenwert von f° ist.

Beweis. Aussage (*i*): Wir setzen zur Abkürzung $g := \lambda \operatorname{id}_V - f$. Wir bestimmen zunächst g°. Nach (35.33) ist g° definiert durch $\theta\big(w, g(v)\big) = \theta\big(g^\circ(w), v\big)$ für alle $v, w \in V$. Es ist

$$\theta\big(w, (\lambda \operatorname{id}_V - f)(v)\big) = \theta\big(w, \lambda\, v\big) - \theta\big(w, f(v)\big) = \theta\big(\overline{\lambda}\, w, v\big) - \theta\big(f^\circ(w), v\big) = \theta\big((\overline{\lambda} \operatorname{id}_V - f^\circ)(w), v\big),$$

also folgt $g^\circ = \overline{\lambda} \operatorname{id}_V - f^\circ$, vgl. auch Bemerkung 35.48.

Wir können jetzt die Normalität von g nachweisen:

$$\begin{aligned}
g \circ g^\circ &= (\lambda \operatorname{id}_V - f) \circ (\overline{\lambda} \operatorname{id}_V - f^\circ) \\
&= \lambda \overline{\lambda} \operatorname{id}_V - f \circ (\overline{\lambda} \operatorname{id}_V) - (\lambda \operatorname{id}_V) \circ f^\circ + f \circ f^\circ \\
&= \overline{\lambda} \lambda \operatorname{id}_V - (\overline{\lambda} \operatorname{id}_V) \circ f - f^\circ \circ (\lambda \operatorname{id}_V) + f \circ f^\circ && \text{denn } f \text{ und } f^\circ \text{ sind linear} \\
&= \overline{\lambda} \lambda \operatorname{id}_V - (\overline{\lambda} \operatorname{id}_V) \circ f - f^\circ \circ (\lambda \operatorname{id}_V) + f^\circ \circ f && \text{denn } f \text{ ist } \theta\text{-normal} \\
&= (\overline{\lambda} \operatorname{id}_V - f^\circ) \circ (\lambda \operatorname{id}_V - f) \\
&= g^\circ \circ g.
\end{aligned}$$

Aussage (*ii*): Es gilt

$$\begin{aligned}
\operatorname{Eig}(f,\lambda) &= \operatorname{Kern}(\lambda \operatorname{id}_V - f) && \text{nach Definition 24.6} \\
&= \operatorname{Kern}(g) && \text{nach Definition von } g \\
&= \operatorname{Kern}(g^\circ) && \text{nach Folgerung 35.57} \\
&= \operatorname{Kern}(\overline{\lambda} \operatorname{id}_V - f^\circ) && \text{nach dem Beweis von Aussage } (i) \\
&= \operatorname{Eig}(f,\overline{\lambda}) && \text{nach Definition 24.6.}
\end{aligned}$$

□

Wir kommen nun zum Hauptresultat dieses Abschnitts, dem **Spektralsatz für θ-normale Endomorphismen** (englisch: spectral theorem for θ-normal endomorphisms) in unitären Räumen.

Satz 35.60 (normale Endomorphismen unitärer Räume sind orthogonal diagonalisierbar, vgl. Satz 34.55)**.**

Es sei (V,θ) ein endlich-dimensionaler unitärer Raum. Für $f \in \operatorname{End}(V)$ sind äquivalent:

(*i*) f ist θ-normal.

(*ii*) f ist diagonalisierbar, und es gibt es Basis aus Eigenvektoren von f, die θ-orthonormal ist. (Wir nennen f dann **θ-orthogonal diagonalisierbar**.)

Beweis. Aussage (*i*) ⟹ Aussage (*ii*): Wir führen einen Induktionsbeweis nach $\dim(V) = n \in \mathbb{N}_0$. Im Fall $n = 0$ ist die Behauptung wahr, denn f und f° sind die Nullabbildungen, und die einzige Basis von V ist die leere Menge.

Wir machen nun den Induktionsschritt von n auf $n+1$. Die Strategie ist, den Vektorraum V in eine orthogonale direkte Summe von Räumen geringerer Dimension zu zerlegen, die f-invariant sowie f°-invariant sind und auf denen die Einschränkung von f wiederum θ-normal ist.

Schritt 1: Wir zerlegen $V = U \oplus W$ mit einem Eigenraum U.

Nach dem Fundamentalsatz der Algebra 11.24 existiert mindestens eine Nullstelle $\lambda \in \mathbb{C}$ des charakteristischen Polynoms χ_f von f, also ein Eigenwert. Wir setzen $g := \lambda \operatorname{id}_V - f$ und $U := \operatorname{Kern}(g) = \operatorname{Kern}(\lambda \operatorname{id}_V - f) = \operatorname{Eig}(f, \lambda)$. Nach Lemma 35.59 ist g wiederum θ-normal. Aus Folgerung 35.57 folgt daher

$$V = \operatorname{Kern}(g) \oplus \operatorname{Bild}(g) = U \oplus W$$

mit $W := \operatorname{Bild}(g) = \operatorname{Bild}(\lambda \operatorname{id}_V - f)$.

Schritt 2: Wir zeigen: W ist f-invariant.

Es sei $w \in W = \operatorname{Bild}(\lambda \operatorname{id}_V - f)$, also $w = \lambda v - f(v)$ für ein $v \in V$. Dann gilt

$$f(w) = f(\lambda v - f(v)) = \lambda f(v) - f(f(v)) = (\lambda \operatorname{id}_V - f)(f(v)),$$

also ist $f(w) \in \operatorname{Bild}(\lambda \operatorname{id}_V - f) = W$. Damit ist die Einschränkung $f|_W^W$ definiert.

Schritt 3: Wir zeigen: W ist auch f°-invariant.

Es gilt

$$\begin{aligned}
f^\circ(W) &= f^\circ(\operatorname{Bild}(\lambda \operatorname{id}_V - f)) && \text{nach Definition von } W \\
&= f^\circ\big((\lambda \operatorname{id}_V - f)(V)\big) && \text{nach Definition von Bild} \\
&= \big[f^\circ \circ (\lambda \operatorname{id}_V - f)\big](V) && \text{nach Definition der Komposition} \\
&= \big[(\lambda \operatorname{id}_V - f) \circ f^\circ\big](V) && \text{denn } f^\circ \text{ ist linear und } f \text{ ist } \theta\text{-normal} \\
&= (\lambda \operatorname{id}_V - f)(f^\circ(V)) && \text{nach Definition der Komposition} \\
&\subseteq (\lambda \operatorname{id}_V - f)(V) && \text{denn } f^\circ \in \operatorname{End}(V) \\
&= W && \text{nach Definition von } W.
\end{aligned}$$

Damit ist auch die Einschränkung $f^\circ|_W^W$ definiert.

Schritt 4: Wir zeigen: $f|_W^W$ ist θ-normal.

Es gilt $(f|_W^W)^\circ = f^\circ|_W^W$, denn adjungierte Abbildungen sind eindeutig, und $f^\circ|_W^W$ erfüllt die Bedingung, die adjungierte Abbildung zu $f|_W^W$ zu sein, nämlich

$$\theta(w_1, f(w_2)) = \theta(f^\circ(w_1), w_2)$$

für alle $w_1, w_2 \in W$. Die Normalität von $f|_W^W$ folgt aus

$$\begin{aligned}
f|_W^W \circ f^\circ|_W^W &= f \circ f^\circ|_W^W \\
&= f^\circ \circ f|_W^W \\
&= f^\circ|_W^W \circ f|_W^W.
\end{aligned}$$

Der Unterraum $W \subsetneq V$ hat echt kleinere Dimension als V, also gilt $\dim(W) \leqslant n$. Da $f|_W^W$ wie gezeigt θ-normal ist, ist $f|_W^W$ nach Induktionsvoraussetzung diagonalisierbar, und W besitzt eine θ-orthonormale Basis aus Eigenvektoren. Diese sind auch Eigenvektoren von f. Wir ergänzen diese Basis von W durch eine θ-orthonormale Basis von U, womit die Behauptung gezeigt ist.

Aussage (ii) $\Rightarrow$ Aussage (i): Es sei B_V eine θ-orthonormale Basis aus Eigenvektoren von f. Es ist also $D := \mathcal{M}_{B_V}^{B_V}(f)$ eine Diagonalmatrix und $M := \mathcal{M}_{B_V^*}^{B_V}(\theta) = I$. Die Darstellungsmatrix des θ-adjungierten Endomorphismus ist nach (35.34) $D^\circ = M^{-1} D^{\mathrm{H}} M = I^{-1} D^{\mathrm{H}} I = D^{\mathrm{H}}$. Es folgt also $D\, D^{\mathrm{H}} = D^{\mathrm{H}} D$. Damit ist f θ-normal. □

Quizfrage 35.3: Warum folgt aus $D^\circ = D^{\mathrm{H}}$ i. A. nicht, dass f sogar θ-selbstadjungiert ist, wie es im Beweis von Satz 34.55 der Fall war?

Analog zu Bemerkung 34.56 gilt auch hier wieder:

Bemerkung 35.61 (zur orthogonalen Diagonalisierbarkeit, vgl. Bemerkung 34.56)**.**
Es sei V ein endlich-dimensionaler komplexer Vektorraum. Ist ein Endomorphismus (oder eine Darstellungsmatrix) diagonalisierbar und ist $B_V = (v_1, \dots, v_n)$ eine Basis von V aus lauter Eigenvektoren, dann können wir ein Innenprodukt θ auf V so wählen, dass f θ-orthogonal diagonalisierbar ist: Definieren wir θ durch

$$\theta(v_i, v_j) \coloneqq \delta_{ij},$$

also mit $\mathcal{M}_{B_V^*}^{B_V}(\theta) = I$, dann ist (V, θ) ein unitärer Raum, und f ist θ-orthogonal diagonalisierbar. △

Die Übersetzung von Satz 35.60 in Matrizen lautet wie folgt:

Folgerung 35.62 (M-normale Matrizen sind M-orthogonal diagonalisierbar, vgl. Folgerung 34.57)**.**
Es sei $(\mathbb{C}^n, \theta_M)$ ein unitärer Raum, also $M \in \mathbb{C}^{n \times n}$ hermitesch und positiv definit. Weiter sei $A \in \mathbb{C}^{n \times n}$, $n \in \mathbb{N}_0$. Dann sind äquivalent:

(i) A ist M-normal, erfüllt also $M^{-1} A^{\mathrm{H}} M = A$.

(ii) A ist diagonalisierbar, und es gibt eine Basis $(v_1, \dots, v_n)$ von $\mathbb{C}^n$ aus Eigenvektoren von A, die M-orthonormal ist. Es gibt also eine invertierbare Matrix $V \in \mathbb{C}^{n \times n}$, sodass $V^{\mathrm{H}} M V = I$ gilt und

$$A V = V \Lambda \tag{35.42}$$

mit der Diagonalmatrix $\Lambda \in \mathbb{C}^{n \times n}$ der Eigenwerte von A entsprechend ihrer Vielfachheit und zugehörigen M-orthonormalen Eigenvektoren $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$.

Wegen der M-Orthonormalität von V, also $V^{\mathrm{H}} M V = I$, gilt $V^{-1} = V^{\mathrm{H}} M$. Wir können daher (35.42) auch schreiben als

$$A = V \Lambda V^{-1} = V \Lambda V^{\mathrm{H}} M. \tag{35.43}$$

Ende der Vorlesung 26

Ende der Woche 13

§ 36 Die Singulärwertzerlegung

Kapitel A Die komplexen Zahlen

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 $\mathbb{C}$.

Die **komplexen Zahlen** lassen sich aus den reellen Zahlen $\mathbb{R}$ aufbauen. Die Menge $\mathbb{C}$ der komplexen Zahlen besteht aus allen Ausdrücken der Form

$$a + b\,i,$$

wobei $a, b \in \mathbb{R}$ sind und i ein Symbol ist, das nicht in $\mathbb{R}$ enthalten ist.

Die **Addition** (englisch: addition) mit dem Symbol $+$ ist definiert durch

$$(a + b\,i) + (c + d\,i) \coloneqq (a + c) + (b + d)\,i.$$

Aufgrund der Kommutativität von $+$ in $\mathbb{R}$ bildet $(\mathbb{C}, +)$ eine kommutative Halbgruppe. Wir prüfen leicht nach, dass $0 + 0\,i$ neutrales Element in $(\mathbb{C}, +)$ ist und dass $a + b\,i$ und $-a - b\,i \coloneqq (-a) + (-b)\,i$ additive Inverse zueinander sind. Wir schreiben daher auch $-(a + b\,i)$ für $(-a) + (-b)\,i$. Wir haben damit gezeigt: $(\mathbb{C}, +)$ ist eine abelsche Gruppe.

Wir definieren weiter die **Multiplikation** (englisch: multiplication) mit dem Symbol $\cdot$ durch[1]

$$(a + b\,i) \cdot (c + d\,i) \coloneqq (a \cdot c - b \cdot d) + (a \cdot d + b \cdot c)\,i.$$

Motiviert ist diese Definition durch das formale Distributivgesetz $(a + b\,i) \cdot (c + d\,i) = a \cdot c + (b\,i) \cdot c + a \cdot (d\,i) + (b\,i) \cdot (d\,i)$, wobei die Kommutativitäts-Regeln $(b\,i) \cdot c = (b \cdot c)\,i$ und $a \cdot (d\,i) = (a \cdot d)\,i$ und $(b\,i) \cdot (d\,i) = (b \cdot d)\,i^2$ gelten sollen sowie $i^2 = -1$. Aufgrund der Kommutativität von $\cdot$ in $\mathbb{R}$ bildet $(\mathbb{C}, \cdot)$ eine kommutative Halbgruppe. Wir prüfen leicht nach, dass $1 + 0\,i$ neutrales Element in $(\mathbb{C}, \cdot)$ ist und dass $a + b\,i$ und $a/(a^2 + b^2) - b/(a^2 + b^2)\,i$ multiplikative Inverse zueinander sind, sofern $a + b\,i \neq 0 + 0\,i$ gilt. Wir schreiben daher auch $(a + b\,i)^{-1}$ für $a/(a^2 + b^2) - b/(a^2 + b^2)\,i$. Wir haben damit gezeigt: $(\mathbb{C} \setminus \{0 + 0\,i\}, \cdot)$ eine abelsche Gruppe.

Weiterhin können wir die (wegen der Kommutativität von $\cdot$ zusammenfallenden) Distributivgesetze

$$\begin{aligned}
(a + b\,i) \cdot \big((c + d\,i) + (e + f\,i)\big) &= (a + b\,i) \cdot (c + d\,i) + (a + b\,i) \cdot (e + f\,i) \\
\big((a + b\,i) + (c + d\,i)\big) \cdot (e + f\,i) &= (a + b\,i) \cdot (e + f\,i) + (c + d\,i) \cdot (e + f\,i)
\end{aligned}$$

nachprüfen. Nach Definition 10.1 ist damit $(\mathbb{C}, +, \cdot)$ ein Körper.

[1]Die Einführung von $\mathbb{C}$ als formale Ausdrücke der Form $a + b\,i$ ist der Einführung von Polynomen in Definition 11.1 nicht unähnlich. Die Besonderheit hier ist aber die Festlegung $i^2 = -1$, die es ermöglicht, jede komplexe Zahl als „Polynom maximal ersten Grades“ in i zu schreiben.

Die Abbildung

$$\mathbb{C} \ni z = a + b\,i \mapsto \overline{z} = a - b\,i \in \mathbb{C}$$

ist ein Körperautomorphismus $(\mathbb{C}, +, \cdot) \to (\mathbb{C}, +, \cdot)$. 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).

Das Produkt $z \cdot \overline{z}$ einer komplexen Zahl $z = a + b\,i$ mit ihrer konjugiert komplexen Zahl ergibt

$$z \cdot \overline{z} = (a + b\,i) \cdot (a - b\,i) = a^2 + b^2.$$

Mit Hilfe der komplexen Konjugation können wir die oben eingeführte Darstellung der multiplikativen Inversen motivieren, denn es gilt

$$\frac{1}{a + b\,i} = \frac{1}{a + b\,i} \cdot \frac{a - b\,i}{a - b\,i} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}\,i.$$

In der Darstellung einer komplexen Zahl $z = a + b\,i$ heißt $a =: \operatorname{Re}(z) \in \mathbb{R}$ der **Realteil** (englisch: real part) und $b =: \operatorname{Im}(z) \in \mathbb{R}$ der **Imaginärteil** (englisch: imaginary part). Es gilt

$$\begin{aligned} \frac{z + \overline{z}}{2} &= \frac{a + b\,i + a - b\,i}{2} = a = \operatorname{Re}(z) \\ \frac{z - \overline{z}}{2\,i} &= \frac{a + b\,i - a + b\,i}{2\,i} = b = \operatorname{Im}(z). \end{aligned}$$

Beachte: $\operatorname{Re}\colon (\mathbb{C}, +, \cdot) \to (\mathbb{C}, +, \cdot)$ und $\operatorname{Im}\colon (\mathbb{C}, +, \cdot) \to (\mathbb{C}, +, \cdot)$ sind keine Körperhomomorphismen! (**Beachte:** Warum nicht?)

Schließlich halten wir fest, dass $\mathbb{R}$ mittels der Identifikation $a = a + 0\,i$ zu einem Teilkörper von $\mathbb{C}$ wird. Eine komplexe Zahl $z \in \mathbb{C}$ ist genau dann reell, wenn $\operatorname{Im}(z) = 0$ gilt, also genau dann, wenn $z = \overline{z}$ gilt. Eine komplexe Zahl $z \in \mathbb{C}$ ist genau dann imaginär (also $z \in i\,\mathbb{R}$), wenn $\operatorname{Re}(z) = 0$ gilt, also genau dann, wenn $z = -\overline{z}$ gilt.

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 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 und 7.20
$(\mathbb{N}_0, +)$		✓	0	Beispiele 7.2, 7.4, 7.8 und 7.20
$(\mathbb{N}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8, 7.14 und 7.20
$(\mathbb{N}_0, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8 und 7.20
$(\mathbb{Z}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8 und 7.14
				Beispiele 7.16 und 7.20
$(\mathbb{Q}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8 und 7.20
$(\mathbb{R}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8, 7.14 und 7.20
$(\mathbb{C}, \cdot)$		✓	1	Beispiele 7.2, 7.4, 7.8 und 7.20
$(\mathbb{Z}_m, \cdot_m)$	multiplikatives Monoid $\mathbb{Z}$ modulo m	✓	1	Beispiele 7.2, 7.8, 7.14 und 7.16
$(H^X, +)$	Funktionen $X \to H$ in die Halbgruppe $(H, +)$	wie in $(H, +)$		Beispiel 10.2
$(\mathbb{N}^X, +)$		✓	—	
$(\mathbb{N}_0^X, +)$		✓	$x \mapsto 0$	
$(H^X, \cdot)$	Funktionen $X \to H$ in die Halbgruppe $(H, \cdot)$	wie in $(H, \cdot)$		Beispiel 10.2
$(\mathbb{N}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{N}_0^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{Z}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{Q}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{R}^X, \cdot)$		✓	$x \mapsto 1$	Beispiele 7.2, 7.4 und 7.16
$(\mathbb{C}^X, \cdot)$		✓	$x \mapsto 1$	
$(X^X, \circ)$		—	id_X	Beispiele 7.2, 7.4, 7.14 und 7.16
$(\mathcal{P}(X), \cap)$		✓	X	Beispiele 7.4 und 7.8
$(\mathcal{P}(X), \cup)$		✓	$\emptyset$	Beispiele 7.4 und 7.8
$(\mathcal{P}(X), \triangle)$		✓	$\emptyset$	Beispiele 7.4 und 7.8
$(\Sigma^*, \circ)$		—	$()$	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 und 7.14 Beispiele 7.16, 7.20 und 7.38
$(\mathbb{Q}, +)$		✓	0	Beispiele 7.2, 7.4, 7.8 und 7.14 Beispiele 7.16 und 7.20
$(\mathbb{R}, +)$		✓	0	Beispiele 7.2, 7.4, 7.8 und 7.14 Beispiele 7.16 und 7.20
$(\mathbb{C}, +)$		✓	0	Beispiele 7.2, 7.4, 7.8 und 7.14 Beispiele 7.16 und 7.20
$(\mathbb{Q}_{\neq 0}, \cdot)$		✓	1	Beispiele 7.14 und 7.16
$(\mathbb{R}_{\neq 0}, \cdot)$		✓	1	Beispiele 7.14 und 7.16
$(\mathbb{C}_{\neq 0}, \cdot)$		✓	1	Beispiele 7.14 und 7.16
$(m\,\mathbb{Z}, +)$	ganzzahlige Vielfache von m	✓	1	Beispiele 7.34 und 7.38
$(\mathbb{Z}_m, +_m)$	additive Gruppe $\mathbb{Z}$ modulo m	✓	0	Beispiele 7.2, 7.8, 7.14 und 7.16
$(\mathbb{Z} / m\,\mathbb{Z}, \tilde{+})$	Faktorgruppe, isomorph zu $(\mathbb{Z}_m, +_m)$	✓	$[1]$	Beispiel 8.15
$(G^X, +)$	Funktionen $X \to G$ in die Gruppe $(G, +)$	wie in $(G, +)$		Beispiel 10.2
$(\mathbb{Z}^X, +)$		✓	$x \mapsto 0$	
$(\mathbb{Q}^X, +)$		✓	$x \mapsto 0$	
$(\mathbb{R}^X, +)$		✓	$x \mapsto 0$	Beispiele 7.2, 7.4 und 7.16
$(\mathbb{C}^X, +)$		✓	$x \mapsto 0$	
$(G^X, \cdot)$	Funktionen $X \to G$ in die Gruppe $(G, \cdot)$	wie in $(G, \cdot)$		Beispiel 10.2
$(\mathbb{Q}_{\neq 0}^X, \cdot)$		✓	$x \mapsto 1$	
$(\mathbb{R}_{\neq 0}^X, \cdot)$		✓	$x \mapsto 1$	Beispiel 7.16
$(\mathbb{C}_{\neq 0}^X, \cdot)$		✓	$x \mapsto 1$	
$(S_n, \circ)$	symmetrische Gruppe auf $[\![1, n]\!]$	—	$\mathrm{id}_{[\![1,n]\!]}$	Definition 7.21
$(A_n, \circ)$	alternierende Gruppe auf $[\![1, n]\!]$	—	$\mathrm{id}_{[\![1,n]\!]}$	Beispiel 7.34

Symbol	Beschreibung	komm.	neut. E.	Referenz
Ringe	$(m \in \mathbb{N}, n \in \mathbb{N}_0)$			
$(\{0_R\}, +, \cdot)$	Nullring	✓	0_R	Beispiel 9.2
$(\mathbb{Z}, +, \cdot)$		✓	1	Beispiel 9.2
$(m\,\mathbb{Z}, +, \cdot)$	ganzzahlige Vielfache von m	✓	—	Beispiel 9.2
$(\mathbb{Z}_m, +_m, \cdot_m)$	$\mathbb{Z}$ modulo m	✓	1	Beispiel 9.2
$(\mathbb{Z} / m\,\mathbb{Z}, \tilde{+}, \tilde{\cdot})$	Restklassenring modulo m, isomorph zu $(\mathbb{Z}_m, +_m, \cdot_m)$	✓	$[1]$	Beispiel 9.2
$(R^X, +, \cdot)$	Funktionen $X \to R$ in den Ring $(R, +, \cdot)$	wie in $(R, +, \cdot)$		Beispiele 9.8 und 10.2
$(\mathbb{Z}^X, +, \cdot)$		✓	1	
$(\mathbb{Q}^X, +, \cdot)$		✓	1	
$(\mathbb{R}^X, +, \cdot)$		✓	1	
$(\mathbb{C}^X, +, \cdot)$		✓	1	
$(R[t], +, \cdot)$	Polynomring über dem kommutativen Ring $(R, +, \cdot)$	✓	1	Definition 11.4
$(R^R, +, \cdot)$	Funktionen $R \to R$ in den Ring $(R, +, \cdot)$	wie in $(R, +, \cdot)$		Bemerkung 11.17
$(\operatorname{End}(G), +, \circ)$	Endomorphismenring der abelschen Gruppe G	—	id_G	Beispiel 9.2
$(K^{n \times n}, +, \cdot)$	quadratische $n \times n$-Matrizen über einem Körper K	—	I_n	Lemma 15.30
$(\operatorname{End}(V), +, \circ)$	Endomorphismen eines Vektorraumes $(V, +, \cdot)$ über einem Körper $(K, +, \cdot)$	—	id_V	Satz 17.12

Körper	$(m \in \mathbb{N})$			
$(\mathbb{Q}, +, \cdot)$		✓	1	Beispiele 9.2 und 10.2
$(\mathbb{R}, +, \cdot)$		✓	1	Beispiele 9.2 und 10.2
$(\mathbb{C}, +, \cdot)$		✓	1	Beispiele 9.2 und 10.2
$(\mathbb{Z}_m, +_m, \cdot_m)$	Körper von $\mathbb{Z}$ modulo m für Primzahlen m	✓	1	Beispiel 9.2
$(\mathbb{Z} / m\,\mathbb{Z}, \tilde{+}, \tilde{\cdot})$	Restklassenkörper mod. m für Primz. m, isomorph zu $(\mathbb{Z}_m, +_m, \cdot_m)$	✓	$[1]$	Beispiel 9.2
$(K(t), +, \cdot)$	rationaler Funktionenkörper	✓	1	Bemerkung 24.18

R. Herzog

Symbol	Beschreibung	Referenz
Vektorräume	$(n, m \in \mathbb{N}_0)$	
$(K_n, +, \cdot)$	Zeilenvektoren über einem Körper $(K, +, \cdot)$	Beispiel 12.3
$(K^n, +, \cdot)$	Spaltenvektoren über einem Körper $(K, +, \cdot)$	Beispiel 12.3
$(K^X, +, \cdot)$	Funktionen $X \to K$ in den Körper $(K, +, \cdot)$	Beispiel 12.3
$(V^X, +, \cdot)$	Funktionen $X \to V$ in den Vektorraum $(V, +, \cdot)$ über dem Körper $(K, +, \cdot)$	
$(K[t], +, \cdot)$	Polynome über dem Körper $(K, +, \cdot)$	Beispiel 12.3
$(K_n[t], +, \cdot)$	Polynome vom Höchstgrad $n \in \mathbb{N}_0$ über dem Körper $(K, +, \cdot)$	Beispiel 12.9
$(K^{\mathbb{N}}, +, \cdot)$	Folgen mit Werten im Körper $(K, +, \cdot)$	Beispiel 12.9
$(K^{\mathbb{N}}, +, \cdot)_{00}$	Folgen mit endlichem Träger und Werten im Körper $(K, +, \cdot)$	Beispiel 12.9
$(\mathbb{Q}^{\mathbb{N}}, +, \cdot)$	Folgen mit Werten im Körper $(\mathbb{Q}, +, \cdot)$	
$(\mathbb{Q}^{\mathbb{N}}, +, \cdot)_b$	beschränkte Folgen mit Werten im Körper $(\mathbb{Q}, +, \cdot)$	
$(\mathbb{Q}^{\mathbb{N}}, +, \cdot)_c$	konvergente Folgen mit Werten im Körper $(\mathbb{Q}, +, \cdot)$	
$(\mathbb{Q}^{\mathbb{N}}, +, \cdot)_0$	Nullfolgen mit Werten im Körper $(\mathbb{Q}, +, \cdot)$	
$(\mathbb{Q}^{\mathbb{N}}, +, \cdot)_{00}$	Folgen mit endlichem Träger mit Werten im Körper $(\mathbb{Q}, +, \cdot)$	
$(\mathbb{R}^{\mathbb{N}}, +, \cdot)$		Beispiel 12.9
$(\mathbb{R}^{\mathbb{N}}, +, \cdot)_b$		Beispiel 12.9
$(\mathbb{R}^{\mathbb{N}}, +, \cdot)_c$		Beispiel 12.9
$(\mathbb{R}^{\mathbb{N}}, +, \cdot)_0$		Beispiel 12.9
$(\mathbb{R}^{\mathbb{N}}, +, \cdot)_{00}$		Beispiel 12.9
$(\mathbb{C}^{\mathbb{N}}, +, \cdot)$		
$(\mathbb{C}^{\mathbb{N}}, +, \cdot)_b$		
$(\mathbb{C}^{\mathbb{N}}, +, \cdot)_c$		
$(\mathbb{C}^{\mathbb{N}}, +, \cdot)_0$		
$(\mathbb{C}^{\mathbb{N}}, +, \cdot)_{00}$		
$(K^{n \times m}, +, \cdot)$	$n \times m$-Matrizen über einem Körper $(K, +, \cdot)$	Definition 15.2
$(\mathrm{Hom}(V, W), +, \cdot)$	Homomorphismen $V \to W$ mit VR $(V, +, \cdot)$ und $(W, +, \cdot)$ über demselben Körper $(K, +, \cdot)$	Satz 17.11

Symbol	Beschreibung	assoz.	komm.	neut. E.	Referenz
Algebren	$(n \in \mathbb{N}_0)$				
$(A^X, +, \cdot, \star)$	Funktionen $X \longrightarrow A$ in die Algebra $(A, +, \cdot, \star)$ über dem Körper $(K, +, \cdot)$	wie in $(A, +, \cdot, \star)$			Beispiel 25.3
$(K[t], +, \cdot, \cdot)$	Polynome über einem Körper $(K, +, \cdot)$	✓	✓	1	Beispiel 25.3
$(K^{n \times n}, +, \cdot, \cdot)$	$n \times n$-Matrizen über einem Körper $(K, +, \cdot)$	✓		I_n	Beispiel 25.3
$(\operatorname{End}(V), +, \cdot, \circ)$	Endomorphismen eines Vektorraumes V über einem Körper $(K, +, \cdot)$	✓		id_V	Beispiel 25.3

Symbol	Beschreibung	Referenz
Euklidische Räume	$(n \in \mathbb{N}_0)$	
$(\mathbb{R}^n, \gamma_M)$	$M \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit	Beispiel 34.5

Symbol	Beschreibung	Referenz
Unitäre Räume	$(n \in \mathbb{N}_0)$	
$(\mathbb{C}^n, \theta_M)$	$M \in \mathbb{R}^{n \times n}$ hermitesch und positiv definit	Beispiel 35.8

R. Herzog

Kapitel C Transformationen und Normalformen

Im Folgenden seien V und W endlich-dimensionale Vektorräume über demselben Körper K. Es seien

- B_V und $\widehat{B}_V$ Basen von V mit Transformationsmatrix $T = \mathcal{T}_{\widehat{B}_V}^{B_V}$, siehe (20.1),
- B_V^* und $\widehat{B}_V^*$ die zugehörigen dualen Basen von V^* mit Transformationsmatrix $T^{-\intercal} = \mathcal{T}_{\widehat{B}_{V^*}}^{B_{V^*}}$, siehe (21.13),
- B_W und $\widehat{B}_W$ Basen von W mit Basiswechselmatrix $S = \mathcal{T}_{\widehat{B}_W}^{B_W}$.

Name der Transformation	Voraussetzung	Darstellungen	Transformation	Referenz	Normalform	Stichwort	Referenz
Äquivalenztransformation	$f \in \mathrm{Hom}(V, W)$	$A = \mathcal{M}^{B_V}_{B_W}(f)$ $\widehat{A} = \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_W}(f)$	$\widehat{A} = S\,A\,T^{-1}$	Satz 20.9	$\begin{bmatrix} I_r & \\ & 0 \end{bmatrix}$	Rang	Folgerung 20.11
Ähnlichkeitstransformation	$f \in \mathrm{End}(V)$	$A = \mathcal{M}^{B_V}_{B_V}(f)$ $\widehat{A} = \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_V}(f)$	$\widehat{A} = T\,A\,T^{-1}$	Satz 20.14	$\begin{bmatrix} C_{p_1} & & \\ & \ddots & \\ & & C_{p_k} \end{bmatrix}$	Frobenius	Satz 29.9
					$\begin{bmatrix} J_{r_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{r_k}(\lambda_k) \end{bmatrix}$	Jordan	Satz 30.7
Kongruenztransformation	$f \in \mathrm{Hom}(V, V^*)$	$A = \mathcal{M}^{B_V}_{B_V^*}(f)$ $\widehat{A} = \mathcal{M}^{\widehat{B}_V}_{\widehat{B}_V^*}(f)$	$\widehat{A} = T^{-\intercal} A\,T^{-1}$	Satz 31.4	$\begin{bmatrix} \alpha_1 & & & & \\ & \ddots & & & \\ & & \alpha_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}$	$A = A^\intercal$	Satz 33.10
					$\begin{bmatrix} I_r & \\ & 0 \end{bmatrix}$ $K = \mathbb{C}$	$A = A^\intercal$	Satz 33.16
					$\begin{bmatrix} I_{n_+} & & \\ & -I_{n_-} & \\ & & 0_{n_0} \end{bmatrix}$ $K = \mathbb{R}$	$A = A^\intercal$	Satz 33.13

Kapitel D Eigenschaften linearer Abbildungen und Matrizen

Lineare Abbildungen bzw. ihre Abbildungsmatrizen können bestimmte Eigenschaften haben. Wir stellen hier nochmals zusammen, welche Eigenschaften zu Abbildungen bzw. ihren Matrizen welchen Typs gehören und was sie bedeuten.

Kapitel E Exkurs: Lie-Algebren

In der obigen Definition 25.1 und in Beispiel 25.3 sind alle Algebren assoziativ. Eine wichtige Klasse i. A. nicht assoziativer Algebren sind die **Lie-Algebren**.

Definition 0.1 (Lie-Algebra über einem Körper)**.**
Es sei $(K, +, \cdot)$ ein Körper. Eine **Lie-Algebra** (englisch: Lie algebra) $(A, +, \cdot, [\cdot, \cdot])$ **über** K (kurz: eine K-Lie-Algebra) ist eine Menge A mit zwei inneren Verknüpfung $+ \colon A \times A \to A$ und $[\cdot, \cdot] \colon A \times A \to A$ sowie einer äußeren Verknüpfung $\cdot \colon K \times A \to A$, die die folgenden Bedingungen erfüllen:

(i) $(A, +, \cdot)$ ist ein K-Vektorraum.

(ii) Die Verknüpfung $[\cdot, \cdot]$, genannt die **Lie-Klammer** (englisch: Lie bracket), ist bilinear, erfüllt also

$$[\alpha\, a + \beta\, b, c] = \alpha\, [a, c] + \beta\, [b, c] \tag{0.1a}$$
$$[a, \beta\, b + \gamma\, c] = \beta\, [a, b] + \gamma\, [a, c] \tag{0.1b}$$

für alle $a, b, c \in$ und alle $\alpha, \beta, \gamma \in K$.

(iii) Die Lie-Klammer ist alternierend, erfüllt also

$$[a, a] = 0 \quad \text{für alle } a \in A. \tag{0.2}$$

(iv) Die Lie-Klammer erfüllt außerdem die **Jacobi-Identität** (englisch: Jacobi identity)

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0 \quad \text{für alle } a, b, c \in A. \tag{0.3}$$

△

Bemerkung 0.2 (zu Lie-Algebren)**.**
Im Gegensatz zu Definition 25.1 wird also die Assoziativität der Multiplikation (hier $[\cdot, \cdot]$ genannt) bei einer Lie-Algebra nicht vorausgesetzt. Daher ist die Bilinearität keine Folge der Definition (Lemma 25.2), sondern wird gefordert. Die Bilinearität (0.1) beinhaltet dabei die Distributivgesetze

$$[a, b + c] = [a, b] + [a, c] \tag{0.4a}$$
$$[a + b, c] = [a, c] + [b, c] \tag{0.4b}$$

sowie die Verträglichkeit mit der S-Multiplikation

$$[\alpha\, a, b] = \alpha[a, b] = [a, \alpha\, b] \tag{0.5}$$

als Spezialfälle, die Bestandteile der Definition 25.1 sind. △

Beispiel 0.3 (Lie-Algebra über einem Körper)**.**

(*i*) $(\mathbb{R}^3, +, \cdot, \times)$ mit dem Kreuzprodukt

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \coloneqq \begin{pmatrix} x_2\, y_3 - x_3\, y_2 \\ x_3\, y_1 - x_1\, y_3 \\ x_1\, y_2 - x_2\, y_1 \end{pmatrix} \tag{0.6}$$

ist eine Lie-Algebra über $\mathbb{R}$.

(*ii*) Ist $(K, +, \cdot)$ ein Körper und $(A, +, \cdot, \star)$ eine assoziative Algebra über K, dann bildet $(A, +, \cdot, [\cdot, \cdot])$ mit dem **Kommutator**

$$[a, b] \coloneqq a \star b - b \star a \tag{0.7}$$

als „Multiplikation“ eine Lie-Algebra über K.

(*iii*) Ist $(K, +, \cdot)$ ein Körper, dann bildet $(K^{n \times n}, +, \cdot, [\cdot, \cdot])$ mit dem **Kommutator**

$$[A, B] \coloneqq A\, B - B\, A \tag{0.8}$$

als „Multiplikation“ eine Lie-Algebra über K.

(*iv*) Ist $(K, +, \cdot)$ ein Körper und $(V, +, \cdot)$ ein Vektorraum über K, dann bildet $(\operatorname{End}(V), +, \cdot, [\cdot, \cdot])$ mit dem **Kommutator**

$$[f, g] \coloneqq f \circ g - g \circ f \tag{0.9}$$

als „Multiplikation“ eine Lie-Algebra über K. △

Kapitel F 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 G 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

Index

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. (2022a). *Einführung in die Mengenlehre*. URL: https://www.aleph1.info/?call=Puc&permalink=mengenlehre1.

– (2022b). *Grundbegriffe der Mathematik*. 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.

Hackbusch, W. (2019). *Tensor Spaces and Numerical Tensor Calculus*. Springer International Publishing. DOI: 10.1007/978-3-030-35554-8.

Jänich, K. (2008). *Lineare Algebra*. 11. Aufl. Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-75502-9.

Liesen, J.; V. Mehrmann (2024). *Lineare Algebra. Ein Lehrbuch über die Theorie mit Blick auf die Praxis*. Springer Studium Mathematik. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-67944-9.

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.