Lineare Algebra und Analytische Geometrie, Wintersemester 2019/20, Sommersemester 2020

Birgit Richter, Version: 6. Juli 2020

Fachbereich Mathematik der Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany

Inhaltsverzeichnis

KAPITEL I

Grundlagen

Bevor wir mit dem eigentlichen Stoff der Linearen Algebra beginnen, behandeln wir einige Grundlagen: Was sind Aussagen? Was sind Mengen? Was sind Abbildungen? Welche grundlegen Eigenschaften können Abbildungen haben? Wie beweist man Behauptungen, die für alle natürlichen Zahlen gelten?

I.1. Aussagen

Einen Großteil der Linearen Algebra verbringen wir damit, zu untersuchen, ob Aussagen beweisbar sind oder ob sie falsch sind. Aber was sind Aussagen?

Wir vereinbaren, dass eine *Aussage* ein sprachliches Gebilde ist, das wahr (w) oder falsch (f) ist und das nicht gleichzeitig wahr und falsch sein kann. Im Kontext dieser Vorlesung ist (hoffentlich) immer klar, ob eine Formulierung eine Aussage ergibt, oder nicht.

Beispiele I.1.1.

- Die Aussage „Es gibt mindestens einen Studierenden im Hörsaal H1“ hat Wahrheitswert w.
- Die Ausssage $2 \cdot 3 = 9$ hat Wahrheitswert f. Ebenso ist zwar „57 ist eine Primzahl“ eine Aussage, aber falsch.
- „Ach herrje!“ ist keine Aussage und genausowenig ist „Diese Aussage ist nicht wahr“ eine Aussage.

Definition I.1.2. Es seien A und B zwei gegebene Aussagen. Dann definieren wir die folgenden Verknüpfungen von Aussagen als eine Aussage, deren Wahrheitswert als Funktion der Wahrheitswerte der Aussagen A und B durch die folgenden Wahrheitstafeln gegeben ist:

(a) nicht A, $\neg A$:

A	$\neg A$
w	f
f	w

(b) A und B, $A \wedge B$:

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

(c) A oder B, $A \vee B$:

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Dies ist also kein exklusives „oder“.

(d) A impliziert B, $A \Rightarrow B$:

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Ein klassisches Beispiel zum vierten Fall ist der Satz „Wenn Ptolemäus Recht hat, ist die Erde eine Scheibe“. Dieser Satz ist wahr, auch wenn die Erde keine Scheibe ist.

(e) A ist äquivalent zu B, $A \Leftrightarrow B$:

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

Beispiel I.1.3. Stellen Sie die Wahrheitswerttafel für die Aussage $(\neg A) \vee B$ auf. Sie stimmt mit der für $A \Rightarrow B$ überein. Also gilt

$$((\neg A) \vee B) \Leftrightarrow (A \Rightarrow B).$$

Wenn Sie üben möchten, dann zeigen Sie die folgenden Äquivalenzen:

(a) $\neg(\neg A) \Leftrightarrow A$
(b) $A \wedge B \Leftrightarrow B \wedge A$
(c) $A \vee B \Leftrightarrow B \vee A$
(d) $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$
(e) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$
(f) $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$
(g) $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$
(h) $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$
(i) $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$
(j) $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$

Bemerkung I.1.4. Die Äquivalenz in (h) benutzt man für die Beweismethode des *indirekten Beweises*. Die Äquivalenzen in (f) und (g) heißen auch die *de Morgansche Gesetze* (de Morgan 1806–1871), aber manchmal wird diese Bezeichnung auch für (i) und (j) benutzt.

I.2. Mengen

Zu Mengen gibt es eine berühmte Charakterisierung, die von Georg Cantor (1845-1918) stammt:

> „Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen.“

Das ist natürlich keine Definition. Ich werde Ihnen im Rahmen dieser Vorlesung allerdings auch keine axiomatische Beschreibung von Mengen geben, weil dies zu weit führen würde. Charakteristisch für Mengen ist die Eigenschaft, dass eine Menge eindeutig durch ihre Elemente bestimmt ist. Wenn Sie die Elemente einer Menge kennen, dann kennen Sie die Menge.

Ist X eine Menge und ist x ein Element von X, so schreiben wir $x \in X$. Das Symbol $x \notin X$ besagt, dass x kein Element von X ist.

Manchmal geben wir Mengen direkt an, indem wir auflisten, was die Elemente sind: $X = \{1, 2, 3\}$ ist die Menge, welche die Zahlen 1, 2 und 3 enthält. Diese Menge stimmt mit der Menge $\{2, 1, 3\}$ überein und auch mit der Menge $\{1, 1, 2, 3\}$: Anordnungen oder Doppelungen spielen keine Rolle. Seien Sie vorsichtig mit $\ldots$-Notationen. Auf `https://oeis.org/` finden Sie zum Beispiel weit über tausend Möglichkeiten, was $\{1, 1, 2, 3, 5, \ldots\}$ sein könnte.

Beispiele I.2.1.

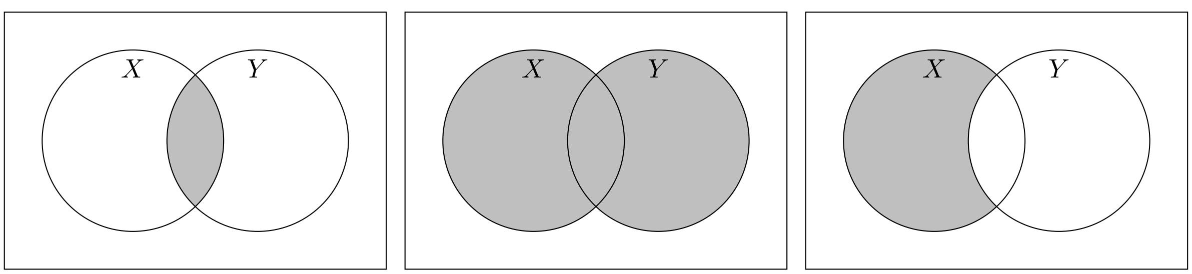
- Die *leere Menge* ist die Menge, die kein Element enthält. Da sie so speziell ist, bekommt sie eine eigene Notation: $\varnothing$. Manchmal schreibt man auch $\varnothing = \{\}$.
- Menge können andere Mengen enthalten: Die Menge $\{1, 2\}$ ist ein Element der Menge
$$X = \{1, \{1\}, \{1, 2\}\}$$
d.h. $\{1, 2\} \in \{1, \{1\}, \{1, 2\}\}$. Aber $2 \notin X$.
- Die Menge $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ der natürlichen Zahlen mit null.

- Die Menge $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ der ganzen Zahlen.
- Die Menge $\mathbb{Q} = \{\frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0\}$ der rationalen Zahlen.
- Die Menge der reellen Zahlen, $\mathbb{R}$, lernen Sie in der Analysis genauer kennen. Wir benutzen Sie naiv.
- Das Gebilde
$$\{X, X \text{ ist Menge}\}$$
ist *keine* Menge. Die axiomatische Mengenlehre schließt die Formung solcher Gebilde aus, um Widersprüche zu vermeiden.
- Ist X eine Menge, so kann man allerdings die Menge all ihrer Teilmengen bilden. Dies ist die *Potenzmenge von X*, geschrieben $\mathcal{P}(X)$. Zum Beispiel ist
$$\mathcal{P}(\{1,2\}) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$$
und $\mathcal{P}(\varnothing) = \{\varnothing\}$. Beachten Sie hier die zusätzlichen Mengenklammern: $\{\varnothing\} \neq \varnothing$!

Da Mengen durch ihre Elemente charakterisiert sind, können wir naheliegende Beziehungen zwischen Mengen benennen und können neue Mengen aus alten Mengen konstruieren.

Definition I.2.2. Es seien X und Y Mengen.
(a) Ist jedes Element von X auch ein Element von Y, so schreiben wir $X \subset Y$ und nennen X eine *Teilmenge* von Y.
(b) Die Menge X ist gleich der Menge Y, $X = Y$, falls $X \subset Y$ und $Y \subset X$.
(c) Der *Durchschnitt oder der Schnitt von X und Y* ist die Menge aller Elemente, die sowohl in X als auch in Y liegen. Die Notation dafür ist $X \cap Y$.
(d) Die *Vereinigung von X und Y*, $X \cup Y$, besteht aus allen Elementen, die in X oder Y (oder beiden) liegen.
(e) Die Differenz $X \setminus Y$ besteht aus allen Elementen in X, die nicht Elemente in Y sind.

Oft veranschaulicht man die obigen Konstruktionen mit sogenannten *Venn-Diagrammen* (John Venn, 1834–1923).



Diese Bilder können irreführend sein, weil sie Spezialfälle suggerieren. Zum Beispiel haben wir bei $X \setminus Y$ nicht verlangt, dass $X \cap Y \neq \varnothing$ ist.

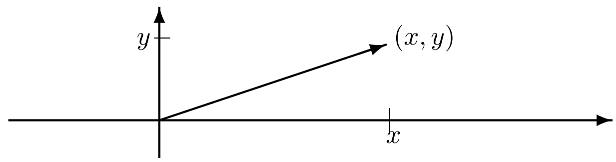
Beispiele I.2.3.
- Es gilt: $\varnothing \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.
- $\{1, 1, 1, 1\} = \{1\}$.
- Ist $x \in X$, so ist $\{x\} \subset X$.
- Es gilt immer, dass $X \cap Y \subset X \cup Y$ und $X \setminus X = \varnothing$.

Eine andere Möglichkeit, neue Mengen aus gegebenen zu formen, ist das Produkt:

Definition I.2.4. Es seien X und Y Mengen.
- Die Menge
$$X \times Y = \{\{\{x\}, \{x, y\}\}, x \in X, y \in Y\}$$
heißt das *kartesische Produkt der Mengen X und Y*.
- Wir schreiben abkürzend (x, y) für ein Element $\{\{x\}, \{x, y\}\}$ von $X \times Y$ und nennen (x, y) das *geordnete Paar mit erstem Eintrag x und zweitem Eintrag y*.

Vorsicht! Die Menge $X \times Y$ stimmt *nicht* mit der Menge $Y \times X$ überein. Für Mengen gilt zwar $\{x, y\} = \{y, x\}$ aber $(x, y) \neq (y, x)$.

Beispiel I.2.5. Ist $X = Y = \mathbb{R}$, so ist $X \times Y = \mathbb{R} \times \mathbb{R}$ und wir schreiben kurz $\mathbb{R}^2$ dafür. Wir veranschaulichen $\mathbb{R}^2$ als Ebene, so dass ein Element $(x, y) \in \mathbb{R}^2$ als erste Koordinate x und als zweite Koordinate y hat.



I.3. Die natürlichen Zahlen

Definition I.3.1. Eine Menge $X \subset \mathbb{R}$ heißt *induktiv*, falls gilt:

(a) $1 \in X$
(b) Ist x ein Element von X, so auch $x + 1$.

Sie kennen induktive Mengen: $\mathbb{N}_0$, $\mathbb{Q}$ und $\mathbb{R}$ sind welche. Wir haben vorher schon die Menge $\mathbb{N}_0$ benutzt. Wir definieren nun:

Definition I.3.2. Die Menge $\mathbb{N}$ ist die kleinste induktive Menge und heißt die *Menge der natürlichen Zahlen.*

Das bedeutet, dass für jede induktive Menge X gilt: $\mathbb{N} \subset X$.

Satz I.3.3. *Ist X eine induktive Menge und gilt $X \subset \mathbb{N}$, so gilt $X = \mathbb{N}$.*

Beweis. Nach Voraussetzung ist $X \subset \mathbb{N}$. Da $\mathbb{N}$ die kleinste induktive Menge ist, gilt $\mathbb{N} \subset X$, also $X = \mathbb{N}$. □

Satz I.3.4. *Es sei für jedes $n \in \mathbb{N}$ eine Aussage A_n gegeben, so dass gilt:*

(a) *A_1 ist wahr.*
(b) *Aus der Annahme, dass A_n für ein beliebig gewähltes $n \in \mathbb{N}$ gilt, folgt immer, dass A_{n+1} wahr ist.*

Dann gilt A_n für alle $n \in \mathbb{N}$.

Beweis. Wir setzen $X := \{n \in \mathbb{N}, A_n \text{ ist wahr}\}$. Damit ist $X \subset \mathbb{N}$ und mit (a) und (b) folgt, dass X induktiv ist, also ist $X = \mathbb{N}$. □

Bemerkung I.3.5. Der Nachweis, dass B_1 gilt, heißt auch *Induktionsanfang*, (IA) . Der Nachweis von (b) heißt *Induktionsschritt*, (IS).

Beispiel I.3.6. Es gibt das berühmte Beispiel von Gauss (Johann Carl Friedrich Gauss, 1777–1855): Was ist die Summe der ersten 100 Zahlen? Wir machen das natürlich allgemeiner. Es sei $n \in \mathbb{N}$. Was ist der Wert von $1 + \ldots + n$?

Die Behauptung ist, dass der Wert gleich $\frac{n(n+1)}{2}$ ist.

(IA) Die Summe von 1 bis 1 ist 1 und das ist $\frac{(1+1)1}{2} = 1$.

(IS) Die Behauptung sei wahr für ein beliebiges $n \in \mathbb{N}$. Dann ist die Summe $1 + \ldots + n + n + 1$ gleich

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

I.4. Relationen

Wir wollen häufig Elemente einer Menge in einer geeigneten Weise als *äquivalent* ansehen.

Definition I.4.1. Es sei X eine Menge.

(a) Eine *Relation auf* X ist eine Teilmenge $R \subset X \times X$. Wir schreiben $x \sim y$, falls $(x, y) \in R$ gilt.

(b) Eine Relation heißt
- *symmetrisch*, falls für alle $x, y \in X$ gilt:
$$x \sim y \Leftrightarrow y \sim x,$$
- *transitiv*, falls für alle $x, y, z \in X$ gilt:
$$(x \sim y) \wedge (y \sim z) \Rightarrow x \sim z$$
- *reflexiv*, falls für alle $x \in X$ gilt: $x \sim x$.

(c) Eine Relation $R \subset X \times X$ heißt *Äquivalenzrelation*, falls R symmetrisch, transitiv und reflexiv ist.

Beispiele I.4.2.
- Es sei X die Menge aller Menschen dieser Erde und R sei die Relation
$$(x, y) \in R \Leftrightarrow x \text{ duzt } y.$$
Reflexivität heißt, dass Sie sich bei Selbstgesprächen duzen. Lehrer*innen duzen Schüler*innen, aber nicht immer umgekehrt, also ist die Relation nicht symmetrisch. Transitiv ist sie auch nicht: Sie duzen Ihre Tante, die duzt ihren Arbeitskollegen, aber deswegen duzen Sie noch lange nicht den Arbeitskollegen Ihrer Tante.
- Eine Relation, die wir in der Mathematik oft benutzen, ist durch Teilbarkeit gegeben. Es sei $X = \mathbb{Z}$ und wir sagen $x \sim y$, falls $x - y$ durch 24 teilbar ist. Es sind also alle Zahlen zu einem $x \in \mathbb{Z}$ äquivalent, die sich nur um ein Vielfaches von 24 von x unterscheiden. Diese Äquivalenzrelation benutzen Sie jeden Tag, wenn Sie über die Uhrzeit reden.
- Die Relation $R \subset \mathbb{R} \times \mathbb{R}$, die durch $(x, y) \in R \Leftrightarrow x < y$ definiert ist, ist transitiv, aber weder reflexiv noch symmetrisch.
- Gleichheit ist eine Äquivalenzrelation: Ist $X \neq \varnothing$, so können wir definieren, dass $x \sim y \Leftrightarrow x = y$.

Definition I.4.3. Es sei X eine Menge und R sei eine Äquivalenzrelation auf X. Eine Teilmenge $Y \subset X$ heißt *Äquivalenzklasse*, falls gilt:
- $Y \neq \varnothing$,
- $x, y \in Y \Rightarrow x \sim y$,
- Ist $y \in Y$ und $x \sim y$, so ist $x \in Y$.

Ist Y eine Äquivalenzklasse der Relation R auf X, so heißt jedes Element y in Y ein *Repräsentant* von Y.

Es gibt meistens keine kanonische Wahl eines Repräsentanten einer Äquivalenzklasse.

In der obigen Äquivalenzrelation $x \sim y \Leftrightarrow 24$ teilt $x - y$ besteht eine Äquivalenzklasse aus denjenigen ganzen Zahlen, die den gleichen Rest bei der Division mit 24 lassen. Was heißt das für die Interpretation als

Uhrzeit? Die Zahl 0 ist Repräsentant der Äquivalenzklasse $Y = \{\ldots, -48, -24, 0, 24, 48, \ldots\}$ aber ebenso 24 oder -24 oder jedes andere $y \in Y$.

Was ist die Äquivalenzklasse von $x \in X$ unter der Gleichheitsrelation?

Lemma I.4.4. *Es sei R eine Äquivalenzrelation auf einer Menge X. Dann gehört jedes $x \in X$ zu genau einer Äquivalenzklasse. Sind Y_1 und Y_2 Äquivalenzklassen bezüglich R, so gilt entweder $Y_1 = Y_2$ oder $Y_1 \cap Y_2 = \varnothing$.*

Beweis. Für ein festes $z \in X$ definieren wir:

$$Y_z := \{y \in X, y \sim z\}.$$

Wir weisen nach, dass Y_z eine Äquivalenzklasse ist:

- Diese Menge enthält wegen der Reflexivität der Relation z und ist somit nicht leer.
- Gilt $x, y \in Y_z$, so gilt nach Definition von Y_z, dass $x \sim z$ und $y \sim z$. Dann gilt aber mit der Symmetrie der Relation, dass auch $z \sim y$. Die Transitivität impliziert dann $x \sim y$.
- Ist $y \in Y_z$ und gilt $x \sim y$, so gilt wiederum nach Definition von Y_z, dass $y \sim z$. Transitivität impliziert, dass $x \sim z$, also ist $x \in Y_z$. Damit erfüllt Y_z alle Bedingungen an eine Äquivalenzklasse.

Wir zeigen nun, dass Äquivalenzklassen entweder gleich sind oder disjunkt. Es seien Y_1 und Y_2 Äquivalenzklassen für die Äquivalenzrelation $\sim$ auf X.

Ist $Y_1 \cap Y_2 \neq \varnothing$, so gibt es ein $z \in Y_1 \cap Y_2$. Wir zeigen, dass gilt: $Y_1 \subset Y_2$. Es sei also $u \in Y_1$, da z und u beide in Y_1 liegen, sind sie äquivalent, also $u \sim z$. Da aber z auch ein Element von Y_2 ist, gilt ebenfalls $u \in Y_2$, also ist jedes Element von Y_1 auch ein Element von Y_2.

Die Inklusion $Y_2 \subset Y_1$ zeigen Sie analog. $\square$

Eine Äquivalenzrelation auf einer Menge X sortiert Elemente in Äquivalenzklassen. Wir fassen diese als Elemente einer neuen Menge auf:

Definition I.4.5. Es sei R eine Äquivalenzrelation auf einer Menge X. Die *Quotientenmenge von X nach R, X/R*, ist die Menge der Äquivalenzklassen.

Beispiele I.4.6.

- Es sei X die Menge aller Menschen und R sei die Relation, die Leute danach sortiert, welchen Nachnamen sie haben. Dann besteht die Äquivalenzklasse *Meier* aus allen Menschen mit diesem Nachnamen. Kurt Mayer ist kein Element dieser Klasse, sondern ein Element der Äquivalenzklasse *Mayer*. Beide Klassen sind Elemente der Menge *Menschen/Nachnamen*.
- Es sei $X = \mathbb{Z}$ und wir betrachten $x \sim y \Leftrightarrow 2$ teilt $x - y$. Dann gibt es die Äquivalenzklasse aller geraden Zahlen und die Äquivalenzklasse aller ungeraden Zahlen und die Quotientenmenge hat zwei Elemente.

I.5. Abbildungen

Definition I.5.1. Es seien X und Y Mengen.

- Eine *zweistellige Relation* zwischen X und Y ist eine Teilmenge $R \subset X \times Y$.
- Eine zweistellige Relation $R_f \subset X \times Y$ heißt *Abbildung*, falls gilt: Ist $(x, y_1) \in R_f$ und $(x, y_2) \in R_f$, so folgt $y_1 = y_2$.
- Der *Definitionsbereich* einer Abbildung, $D(R_f)$, ist die Menge aller $x \in X$, für die es ein $y \in Y$ gibt mit $(x, y) \in R_f$.
- Ist R_f eine Abbildung und $D(R_f) = X$, so nennen wir R_f eine *Abbildung von X nach Y*.

Bemerkung I.5.2. Für eine Abbildung R_f von X nach Y schreiben wir kurz $f\colon X \to Y$. Eine Abbildung ist also eine Vorschrift, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet.

Beispiel I.5.3. Es ist durchaus erlaubt, allen $x \in X$ den gleichen Wert y zuzuordnen. Dann erhalten wir die *konstante Abbildung mit Wert y*. Zum Beispiel ist $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = 0$ für alle $x \in \mathbb{R}$ die konstante Abbildung mit Wert 0, oder kurz die Nullabbildung.

Definition I.5.4.

- Sind $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen, so ist die *Verkettung* oder die *Komposition* von g mit f, $g \circ f$, definiert als die Abbildung, die $x \in X$ den Wert $g(f(x))$ zuordnet, also

$$(g \circ f)(x) := g(f(x)), \qquad \begin{tikzcd} X \arrow[rr, "g \circ f"] \arrow[dr, "f"'] & & Z. \\ & Y \arrow[ur, "g"'] & \end{tikzcd}$$

- Ist $f\colon X \to Y$ eine Abbildung und ist $U \subset X$ eine Teilmenge von X, so heißt die Menge
$$f(U) := \{f(u), u \in U\}$$
die *Bildmenge von U unter f* oder kürzer *das Bild von U unter f*.
- Ist $f\colon X \to Y$ eine Abbildung und ist $V \subset Y$ eine Teilmenge von Y, so heißt die Menge
$$f^{-1}(V) := \{x \in X, f(x) \in V\}$$
die *Urbildmenge von V unter f* oder kürzer *das Urbild von V unter f*.

Vorsicht: Die Notation f^{-1} hat in dieser Situation nichts mit Umkehrabbildungen zu tun.

Lemma I.5.5. *Es sei $f\colon X \to Y$ eine Abbildung und seien $U \subset X$ und $V \subset Y$.*
- *Das Bild von U unter f ist eine Teilmenge von Y und das Urbild von V unter f ist eine Teilmenge von X.*
- *Sind $f\colon X \to Y$, $g\colon Y \to Z$ und $h\colon Z \to W$ Abbildungen, dann ist die Komposition assoziativ, d.h. es gilt*
$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Beweis. Dass das Bild und das Urbild Teilmengen ergeben, folgt direkt aus der Definition.
Für die Komposition von Abbildungen rechnen wir für ein beliebiges $x \in X$ nach:
$$h \circ (g \circ f)(x) = h((g \circ f)(x)) = h(g(f(x))) \text{ und } ((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x))).$$
Da beide Werte für ein beliebiges $x \in X$ übereinstimmen und da eine Abbildung durch die Werte auf allen $x \in X$ eindeutig bestimmt ist, folgt die Behauptung. □

Beispiele I.5.6.
- Oft geben wir Abbildungen explizit durch eine Vorschrift an, indem man zum Beispiel kurz schreibt $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$.
- Manchmal benutzen wir auch Fallunterscheidungen, um Abbildungen zu definieren. Zum Beispiel ist die *Heaviside Abbildung* (Oliver Heaviside, 1850–1925) definiert als
$$h\colon \mathbb{R} \to \mathbb{R}, \quad h(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x \geqslant 0. \end{cases}$$
- Sie kennen Abbildungen, die über Wertetabellen definiert sind (Fußballergebnisse, Temperaturtabellen, etc.). Zum Beispiel definiert

x	1	2	3
$f(x)$	2	3	1

eine Abbildung $f\colon \{1,2,3\} \to \{1,2,3\}$.
- Eine unscheinbare, aber wichtige Abbildung ist die *identische Abbildung*. Für jede nichtleere Menge X können wir $\mathrm{id}_X\colon X \to X$ definieren als $\mathrm{id}_X(x) = x$ für alle $x \in X$.

Wir untersuchen oft die Qualität von Abbildungen:

Definition I.5.7. Es sei $f\colon X \to Y$ eine Abbildung.
- f ist *surjektiv*, falls es für alle $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$.
- f ist *injektiv*, falls aus $f(x_1) = f(x_2)$ folgt, dass $x_1 = x_2$ ist.
- f ist *bijektiv*, falls f sowohl injektiv als auch surjektiv ist.

Bemerkung I.5.8. Eine Abbildung $f\colon X \to Y$ ist genau dann surjektiv, wenn das Bild von X unter f ganz Y ist. Die Injektivität von f ist äquivalent dazu, dass verschiedene Elemente aus X auch auf verschiedene Elemente in Y abgebildet werden: Aus $x_1 \neq x_2$ folgt immer, dass $f(x_1) \neq f(x_2)$ ist, falls f injektiv ist.

Satz I.5.9. *Es sei $X \neq \varnothing$ und $f\colon X \to Y$ sei eine Abbildung. Dann gilt:*

(a) *Die Abbildung f ist genau dann injektiv, wenn es eine Abbildung $g\colon Y \to X$ gibt, so dass $g \circ f = \mathrm{id}_X$.*
(b) *Die Abbildung f ist genau dann surjektiv, wenn es eine Abbildung $g\colon Y \to X$ gibt, so dass $f \circ g = \mathrm{id}_Y$.*
(c) *Die Abbildung f ist genau dann bijektiv, wenn es eine Abbildung $g\colon Y \to X$ gibt, so dass $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$ gilt.*

Im ersten Fall heißt g ein *Linksinverses* zu f, im zweiten Fall ein *Rechtsinverses* zu f und im dritten Fall heißt g die *Umkehrabbildung* (oder das *Inverse*) von f.

Beweis. Wir zeigen die Äquivalenz der Behauptungen jeweils, indem wir die Richtungen $\Rightarrow$ und $\Leftarrow$ getrennt beweisen.

(a) Nehmen wir an, dass f injektiv ist. Wir konstruieren ein Linksinverses g zu f: Da $X \neq \varnothing$, gibt es ein $x' \in X$. Wir definieren

$$g(y) = \begin{cases} x', & \text{falls } y \notin f(X), \\ x & \text{falls } f(x) = y. \end{cases}$$

Da f injektiv ist, definiert diese Vorschrift eine Abbildung. Nun ist für jedes $x \in X$ nach Konstruktion

$$g(f(x)) = x,$$

weil $f(x) \in f(X)$ gilt. Also ist g ein Linksinverses zu f.

Wir nehmen nun an, dass f ein Linksinverses g hat und dass $f(x_1) = f(x_2)$ gilt für $x_1, x_2 \in X$. Wir wenden g an auf diese Gleichung und erhalten

$$g(f(x_1)) = g(f(x_2)).$$

Da aber $x_1 = \mathrm{id}_X(x_1) = g(f(x_1))$ und $x_2 = \mathrm{id}_X(x_2) = g(f(x_2))$ ist, folgt dass $x_1 = x_2$ ist, also ist f injektiv.

(b) Es sei f surjektiv. Wir konstruieren ein Rechtsinverses zu f. Da f surjektiv ist, sind die Urbilder $f^{-1}(\{y\})$ für kein $y \in Y$ leer. Wir wählen zu jedem $y \in Y$ ein $x_y \in f^{-1}(\{y\})$ und definieren

$$g(y) := x_y.$$

Damit gilt $f(g(y)) = f(x_y) = y$, weil $x_y \in f^{-1}(\{y\})$. Also ist g ein Rechtsinverses.

Wir nehmen nun an, dass es ein Rechtsinverses g zu f gibt. Es sei $y \in Y$ ein beliebiges Element. Für $x := g(y)$ gilt dann

$$y = \mathrm{id}_Y(y) = (f \circ g)(y) = f(g(y)) = f(x)$$

und damit liegt y im Bild von X unter f. Da y beliebig war, folgt dass f surjektiv ist.

(c) Da f bijektiv ist, ist es injektiv und surjektiv, also hat f ein Linksinverses $g_1\colon Y \to X$ und ein Rechtsinverses $g_2\colon Y \to X$. Die Assoziativität der Verkettung von Abbildung liefert:

$$g_1 = g_1 \circ \mathrm{id}_Y = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = \mathrm{id}_X \circ g_2 = g_2.$$

Damit ist das Linksinverse gleich dem Rechtsinversen und wir taufen es g. Diese Abbildung ist das gesuchte Inverse zu f.

Es sei g nun ein Inverses zu f, so ist g ein Linksinverses und ein Rechtsinverses von f. Damit folgt mit (a) und (b), dass f sowohl injektiv als auch surjektiv ist. □

Bemerkung I.5.10. Ist f eine bijektive Abbildung, so ist die Umkehrabbildung eindeutig bestimmt. Wir schreiben $f^{-1}\colon Y \to X$ für diese Abbildung. In diesem Fall ist das Urbild von Y unter f genau das Bild von Y unter f^{-1} und somit ist die Notation $f^{-1}(Y)$ nicht mehrdeutig.

Eigenschaften von Abbildungen $f\colon X \to Y$ hängen stark von X und Y ab. So ist zum Beispiel die Abbildung $f\colon \mathbb{Z} \to \mathbb{Z}$, $x \mapsto 2x$ zwar injektiv, aber nicht surjektiv. Dagegen ist $g\colon \mathbb{Q} \to \mathbb{Q}$, $x \mapsto 2x$ sowohl injektiv als auch surjektiv, also bijektiv. Was ist g^{-1}?

Was ist die Umkehrabbildung von $\begin{array}{c|c|c|c} x & 1 & 2 & 3 \\ \hline f(x) & 2 & 3 & 1 \end{array}$?

Zum Abschluss dieses Abschnitts bringen wir Abbildungen in Verbindung mit Äquivalenzrelationen. Es sei $X \neq \varnothing$ und R sei eine Äquivalenzrelation auf X.

Lemma I.5.11. *Es gibt eine surjektive Abbildung $\pi\colon X \to X/R$, die ein $x \in X$ auf seine Äquivalenzklasse $\pi(x)$ abbildet.*

Beweis. Da Äquivalenzklassen entweder disjunkt oder gleich sind, definiert diese Zuordnung eine Abbildung. Ist Y eine Äquivalenzklasse von R, so ist $Y \neq \varnothing$. Also gibt es ein $x \in Y$. Damit ist $\pi(x) = Y$. □

Es sei $f\colon X \to Y$ eine Abbildung. Wir definieren eine Relation R_f auf X durch

$$x_1 \sim_f x_2 \Leftrightarrow f(x_1) = f(x_2).$$

Rechnen Sie bitte nach, dass dies eine Äquivalenzrelation definiert. Die Äquivalenzklassen sind genau die Urbilder einelementiger Untermengen des Bildes, also $f^{-1}(\{y\})$ mit $y \in f(X)$.

Satz I.5.12. *Es gibt genau eine Abbildung $\bar{f}\colon X/R_f \to Y$ mit der Eigenschaft, dass $\bar{f} \circ \pi = f$ ist.*

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "\pi"'] & & Y \\ & X/R_f \arrow[ur, "\bar{f}"'] & \end{tikzcd}$$

Beweis. Wenn es eine solche Abbildung gibt, so muss $\bar{f}(\pi(x)) = f(x)$ für alle $x \in X$ gelten. Dies legt die Abbildung $\bar{f}$ auf X/R_f fest, indem wir definieren, dass für eine Äquivalenzklasse Z $\bar{f}(Z) = f(x)$ ist, falls $x \in Z$.

Diese Definition hängt nicht von der Wahl von x ab. Ist $x_1 \sim_f x_2$ und $x_1 \in Z \in X/R_f$ so ist $f(x_1) = f(x_2)$ und somit gilt

$$\bar{f}(Z) = f(x_1) = f(x_2)$$

unabhängig davon, ob wir x_1 oder x_2 zur Auswertung von $\bar{f}(Z)$ benutzen. □

I.6. Quantoren

Wir haben schon viele Situationen gesehen, in denen wir Ausdrücke wie „Für alle $x \in X \ldots$“ oder „Es gibt ein $x \in X \ldots$“ benutzt haben.

Wir betrachten Aussagen, deren Wahrheitswert von einem Parameter abhängt. Es sei X eine Menge und $x \in X$. Wir benutzen die Notation:

$\forall x \in X : A(x)$ Die Aussage $A(x)$ gilt für alle $x \in X$.
$\exists x \in X : A(x)$ Es gibt ein $x \in X$, für das die Aussage $A(x)$ gilt.
$\exists! x \in X : A(x)$ Es gibt genau ein $x \in X$, für das die Aussage $A(x)$ gilt.

Die Symbole $\forall$ und $\exists$ heißen *Quantoren*.

Seien Sie bitte vorsichtig, wenn Sie Aussagen verneinen möchten, die Quantoren enthalten. Die Verneinung der Aussage „Alle Primzahlen sind ungerade“ ist **nicht** „Alle Primzahlen sind nicht ungerade“ (also „Alle Primzahlen sind gerade“) sondern „Es gibt eine Primzahl, die nicht ungerade ist“. Diese Aussage ist wahr: Die 2 ist gerade aber trotzdem eine Primzahl. Vorsicht ist auch angebracht mit Quantoren und Aussagen, die *und* oder *oder* enthalten. Es gelten unter anderem die folgenden Beziehungen.

(a)
$$(\neg(\forall x \in X : A(x))) \Leftrightarrow (\exists x \in X : \neg A(x)).$$

(b)
$$(\neg(\exists x \in X : A(x))) \Leftrightarrow (\forall x \in X : \neg A(x)).$$

(c)
$$((\forall x \in X : A(x)) \vee (\forall x \in X : B(x))) \Rightarrow (\forall x \in X : A(x) \vee B(x)).$$

(d)
$$(\exists x \in X : A(x) \wedge B(x)) \Rightarrow ((\exists x \in X : A(x)) \wedge (\exists x \in X : B(x))).$$

(e)
$$((\forall x \in X : A(x)) \wedge (\forall x \in X : B(x))) \Leftrightarrow (\forall x \in X : A(x) \wedge B(x)).$$
(f)
$$(\exists x \in X : A(x) \vee B(x)) \Leftrightarrow ((\exists x \in X : A(x)) \vee (\exists x \in X : B(x))).$$

Überlegen Sie sich Beispiele, warum in (c) und (d) nur Implikationen aber keine Äquivalenzen stehen.

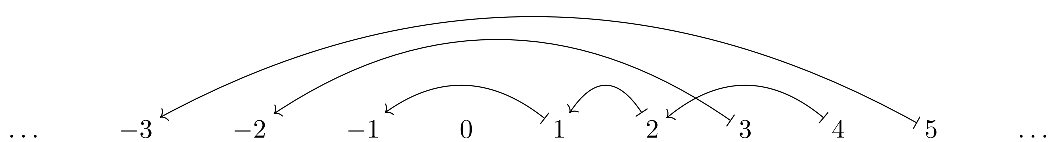
Hängt eine Aussage von zwei Parametern ab, darf man Quantoren nicht einfach vertauschen. Ist $A(x, y)$ eine Aussage mit $x \in X$ und $y \in Y$, so gilt nur
$$(\exists x \in X : \forall y \in Y : A(x, y)) \Rightarrow \forall y \in Y : \exists x \in X : A(x, y).$$
Ein Beispiel dazu ist $X = Y = \mathbb{Z}$ und $A(x, y)$ sei die Aussage, dass $x < y$ ist. Dann ist es falsch, dass es ein $x \in \mathbb{Z}$ gibt, welches kleiner ist als alle $y \in \mathbb{Z}$, aber es ist wahr, dass es für jedes $y \in \mathbb{Z}$ ein $x \in \mathbb{Z}$ gibt mit $x < y$.

I.7. Mächtigkeit von Mengen

Wir verwenden die folgenden Referenzmengen: Für $n \in \mathbb{N}_0$ sei
$$\mathbf{n} := \begin{cases} \varnothing, & \text{falls } n = 0, \\ \{1, \ldots, n\}, & \text{sonst.} \end{cases}$$

Definition I.7.1.
(a) Eine Menge X ist *endlich*, falls es ein $n \in \mathbb{N}_0$ und eine Bijektion $f\colon \mathbf{n} \to X$ gibt. Die Zahl n heißt dann die *Mächtigkeit von* X. Wir benutzen die Notation $|X| = n$ oder auch $\#X = n$.
(b) Zwei Mengen X und Y heißen *gleichmächtig*, falls es eine bijektive Abbildung $f\colon X \to Y$ gibt.
(c) Ist X eine Menge, die gleichmächtig zu $\mathbb{N}_0$ ist, so heißt X *abzählbar*.

Beispiele I.7.2.
- Die Mengen $\mathbb{N}_0$ und $\mathbb{Z}$ sind gleichmächtig. Die Abbildung
$$f\colon \mathbb{N}_0 \to \mathbb{Z}, \quad f(n) = \begin{cases} \frac{n}{2}, & \text{falls } n \text{ gerade ist,} \\ -\frac{n+1}{2}, & \text{falls } n \text{ ungerade ist,} \end{cases}$$
ist eine Bijektion.

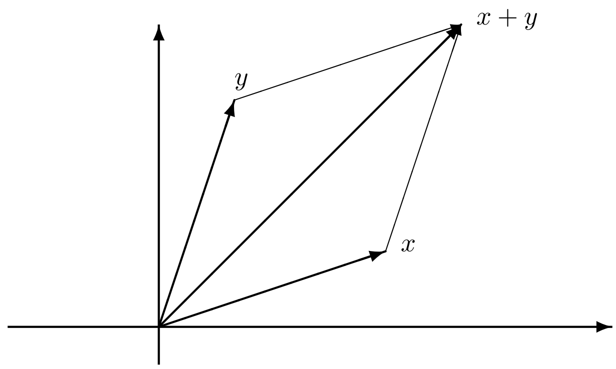


- Die Mengen $\mathbb{N}_0$ und $\mathbb{Q}$ sind gleichmächtig, aber $\mathbb{R}$ ist nicht abzählbar.

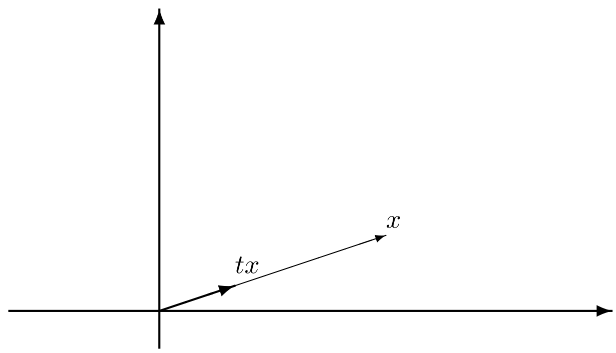
I.8. Geraden und Geometrie in der Ebene

Elemente in $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ schreiben wir oft als Vektoren, also
$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, x_1, x_2 \in \mathbb{R} \right\}.$$

Definition I.8.1.
(a) Die *Summe* von $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ist
$$x + y := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}.$$



(b) Die *Streckung von* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *um den Wert* $t \in \mathbb{R}$ ist $tx := \begin{pmatrix} tx_1 \\ tx_2 \end{pmatrix}$.



Bemerkung I.8.2. Für alle $x, y, z \in \mathbb{R}^2$ und alle $s, t \in \mathbb{R}$ gilt:

(a) $(x+y)+z = x+(y+z)$.
(b) Mit $0 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ist $0 + x = x = x + 0$.
(c) Für alle $x \in \mathbb{R}^2$ gibt es ein $-x \in \mathbb{R}^2$ mit $x + (-x) = 0 = (-x) + x$, nämlich
$$-x = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}.$$
(d) $x + y = y + x$.
(e) $s(tx) = (st)x$.
(f) $1x = x$.
(g) $t(x+y) = tx + ty$.
(h) $(s+t)x = sx + tx$.

Für den $\mathbb{R}^2$ sind diese Rechenregeln einfach nachzurechnen. Tun Sie das bitte. Analoge Aussagen gelten für den $\mathbb{R}^n$ für alle $n \geqslant 1$. Wir werden später genau diese Rechenregeln benutzen, um allgemein zu sagen, was ein Vektorraum sein soll.

Definition I.8.3. Es sei $n \in \mathbb{N}$ fest.

Für $p, v \in \mathbb{R}^n$ mit $v \neq 0$ heißt
$$G_{p,v} := \{p + tv, t \in \mathbb{R}\}$$
eine *affine Gerade mit Fußpunkt* p *und Richtungsvektor* v.

Oft schreibt man kurz $G_{p,v} = p + \mathbb{R}v$. Die obige Darstellung heißt *Parameterdarstellung der Geraden.* Vorsicht! Diese Darstellung ist nicht eindeutig:

Satz I.8.4. *Für $v, w \in \mathbb{R}^n \setminus \{0\}$ und $p, q \in \mathbb{R}^n$ gilt:*

$$G_{p,v} = G_{q,w} \Leftrightarrow (q \in G_{p,v} \text{ und } \exists s \in \mathbb{R} \setminus \{0\} : w = sv).$$

Beweis. Wir nehmen an, dass $G_{p,v} = G_{q,w}$ gilt. Da $q \in G_{q,w}$ ist also auch $q \in G_{p,v}$. Damit gibt es ein $\mu \in \mathbb{R}$ mit

$$q = p + \mu v.$$

Ebenso ist $q + w \in G_{p,v}$, also gibt es ein $\lambda \in \mathbb{R}$ mit

$$q + w = p + \lambda v.$$

Da $q + w \neq q$ ist, folgt $\mu \neq \lambda$ und

$$\begin{aligned} w =& q + w - q \\ =& p + \lambda v - (p + \mu v) \\ =& (\lambda - \mu)v. \end{aligned}$$

Wir setzen $s := \lambda - \mu$ und erhalten $w = sv$.

Für die Rückrichtung sei also $q \in G_{p,v}$, also gibt es ein $\lambda \in \mathbb{R}$ mit $q = p + \lambda v$. Da $w = sv$ ist, können wir jedes $x \in G_{q,w}$ schreiben als

$$x = q + tsv.$$

Damit gilt aber auch

$$\begin{aligned} x =& q + tsv \\ =& p + \lambda v + tsv \\ =& p + (\lambda + ts)v \end{aligned}$$

und damit ist x ein Element von $G_{p,v}$. Dies zeigt $G_{q,w} \subset G_{p,v}$. Ist $y \in G_{p,v}$, so schreiben wir wieder $y = p + \mu v$, benutzen wiederum $q = p + \lambda v$ und rechnen nach:

$$\begin{aligned} y =& q - \lambda v + \mu v \\ =& q + (\mu - \lambda)v. \end{aligned}$$

Da nach Voraussetzung $w = sv$ ist mit $s \in \mathbb{R} \setminus \{0\}$, können wir schließen, dass

$$y = q + (\mu - \lambda)v = q + ((\mu - \lambda)s^{-1})w$$

und damit gilt auch $G_{p,v} \subset G_{q,w}$. □

Korollar I.8.5. *Ist $G \subset \mathbb{R}^n$ eine Gerade und sind $x, y \in G$ mit $x \neq y$, so hat G die Parameterform*

$$G = G_{x,y-x}.$$

Insbesondere ist jede Gerade im $\mathbb{R}^n$ durch zwei Punkte festgelegt.

Beweis. Da $x \in G$, können wir x als Fußpunkt nehmen. Also ist $G = G_{x,v}$ für ein $v \in \mathbb{R}^n \setminus \{0\}$. Da $y \in G = G_{x,v}$, gibt es ein $\lambda \in \mathbb{R}$ mit $y = x + \lambda v$. Damit ist

$$\lambda v = y - x,$$

aber da $x \neq y$, ist $\lambda \neq 0_K$ und wir können v schreiben als

$$v = \lambda^{-1}(y - x).$$

Aus Satz I.8.4 folgt die Behauptung. □

Wie sieht die Paramterdarstellung aus, wenn Sie y als Fußpunkt wählen?

Im Folgenden betrachten wir der Einfachheit halber wieder den $\mathbb{R}^2$. Die Konzepte und Ergebnisse lassen sich aber auf beliebige $\mathbb{R}^n$ für $n \geqslant 2$ verallgemeinern.

Definition I.8.6. Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ sei das *Skalarprodukt von x und y*

$$\langle x, y \rangle := x_1 y_1 + x_2 y_2 \in \mathbb{R}.$$

Satz I.8.7. *Die Abbildung*

$$\langle -,-\rangle \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \langle x,y\rangle$$

hat die folgenden Eigenschaften:

(a) $\forall x, x', y \in \mathbb{R}^2 : \langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$.
(b) $\forall x, y \in \mathbb{R}^2, \forall t \in \mathbb{R} : \langle tx, y\rangle = t\langle x, y\rangle$.
(c) $\forall x, y \in \mathbb{R}^2 : \langle x, y\rangle = \langle y, x\rangle$.
(d) *Für alle* $x \in \mathbb{R}^2$ *gilt, dass* $\langle x, x\rangle \geqslant 0$*;* $\langle x, x\rangle = 0$ *gilt genau dann, wenn* $x = 0$.

Beweis. (a) bis (c) rechnen Sie direkt nach. Zu (d):

$$\langle x, x\rangle = x_1^2 + x_2^2$$

ist die Summe zweier reeller Quadrate und $x_i^2 \geqslant 0$ für $x_i \in \mathbb{R}$. Die Gleichheit $x_1^2 + x_2^2 = 0$ gilt nur, wenn $x_1 = x_2 = 0$, aber dass ist äquivalent dazu, dass $x = 0$. □

Das Skalarprodukt hat vielfältige Anwendungen, zum Beispiel zur Längen- und Abstandsmessung und zur Identifizierung von Winkeln zwischen Vektoren.

Definition I.8.8. Es seien $x, y \in \mathbb{R}^2$.

(a) Die *Norm (oder auch Länge) von* x sei

$$||x|| := \sqrt{\langle x, x\rangle}.$$

(b) Wir nennen $d(x, y) := ||x - y||$ den *euklidischen Abstand* von x und y.

Aus den Eigenschaften des Skalarprodukts folgt sofort, dass $||x|| \geqslant 0$ ist und dass der Wert 0 nur für $x = 0$ angenommen wird. Für ein $t \in \mathbb{R}$ gilt für die Streckung von x durch t:

$$||tx|| = \sqrt{\langle tx, tx\rangle} = \sqrt{t^2\langle x, x\rangle} = |t|||x||.$$

Satz I.8.9. *Es seien* $x, y, z \in \mathbb{R}^2$.

(a) *Es gilt die* Cauchy-Schwarzsche Ungleichung*:*

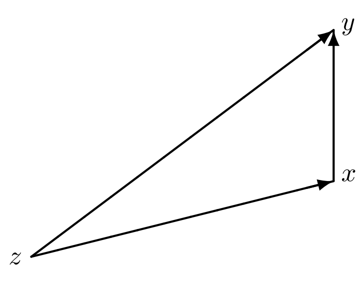
$$|\langle x, y\rangle| \leqslant ||x||||y||.$$

(b)

$$||x + y|| \leqslant ||x|| + ||y||.$$

(c)

$$d(x, z) \leqslant d(x, y) + d(y, z).$$



Die Ungleichung (c) heißt die *Dreiecksungleichung*.

Beweis. Für (a) rechnen wir aus, dass gilt:

$$\begin{aligned}(||x||||y||)^2 - \langle x, y\rangle^2 =&\langle x, x\rangle\langle y, y\rangle - \langle x, y\rangle^2 \\ =&(x_1^2 + x_2^2)(y_1^2 + y_2^2) - (x_1y_1 + x_2y_2)^2 \\ =&x_1^2y_2^2 + x_2^2y_1^2 - 2x_1y_1x_2y_2 \\ =&(x_1y_2 - x_2y_1)^2 \geqslant 0.\end{aligned}$$

Wir erhalten also, dass

$$(||x||||y||)^2 \geqslant \langle x, y\rangle^2$$

und aus der Monotonie der Wurzelfunktion folgt mit $\sqrt{r^2} = |r|$ für alle $r \in \mathbb{R}$, dass

$$||x||||y|| \geqslant |\langle x, y\rangle|.$$

Für (b) rechnen wir aus, dass

$$\begin{aligned} ||x + y||^2 =& \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ =& ||x||^2 + 2\langle x, y\rangle + ||y||^2 \end{aligned}$$

gilt. Mit der Cauchy-Schwarzschen Ungleichung können wir dies abschätzen:

$$\begin{aligned} ||x||^2 + 2\langle x, y\rangle + ||y||^2 \leqslant& ||x||^2 + 2||x||||y|| + ||y||^2 \\ =& (||x|| + ||y||)^2. \end{aligned}$$

Die Dreiecksungleichung in (c) folgt direkt, weil nach (b) gilt:

$$||x - z|| = ||x - y + y - z|| \leqslant ||x - y|| + ||y - z||.$$

□

Bemerkung I.8.10. Die Cauchy-Schwarzsche Ungleichung stellt sicher, dass für $x, y \in \mathbb{R}^2 \setminus \{0\}$ gilt

$$\left|\frac{\langle x, y\rangle}{||x||||y||}\right| = \frac{|\langle x, y\rangle|}{||x||||y||} \leqslant 1.$$

Wir wissen also, dass $\frac{\langle x,y\rangle}{||x||||y||}$ Werte im Intervall $[-1, 1] := \{t \in \mathbb{R}, -1 \leqslant t \leqslant 1\}$ annimmt. Die Kosinusabbildung

$$\cos\colon [0, \pi] \to [-1, 1]$$

ist bijektiv und arccos bezeichne die Umkehrabbildung

$$\arccos\colon [-1, 1] \to [0, \pi],$$

die also einem $t \in [-1, 1]$ einen Winkel in $[0, \pi]$ zuordnet.

Definition I.8.11. Für $x, y \in \mathbb{R}^2 \setminus \{0\}$ heißt

$$\measuredangle(x, y) := \arccos\left(\frac{\langle x, y\rangle}{||x||||y||}\right)$$

der *Innenwinkel von x und y.*

Beispiele I.8.12.

- Ist $y = -x$, so erhalten wir

  $$\langle x, y\rangle = -\langle x, x\rangle$$

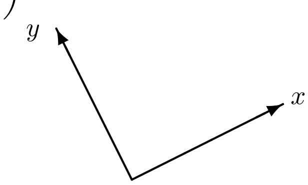
  und

  $$\measuredangle(x, y) = \arccos\left(\frac{\langle x, y\rangle}{||x||||y||}\right) = \left(\frac{-\langle x, x\rangle}{||x||||x||}\right) = \arccos(-1) = \pi.$$

- Ist $x = y$, so gilt

  $$\measuredangle(x, y) = \measuredangle(x, x) = \arccos(1) = 0.$$

- Ist $\langle x, y\rangle = 0$ für $x \neq 0 \neq y$, so ist $\arccos(0) = \frac{\pi}{2}$ entsprechend 90 Grad. Also stehen x und y senkrecht aufeinander. Wir sagen auch, dass x und y *orthogonal* zueinander sind. Zum Beispiel ist für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ immer $y = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ orthogonal.



Wir benutzen im Folgenden das Skalarprodukt, um Geraden im $\mathbb{R}^2$ zu beschreiben.

Es seien $p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \in \mathbb{R}^2$ und $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2 \setminus \{0\}$. Ist $G_{p,v} \subset \mathbb{R}^2$ eine Gerade, dann gilt für ein $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$:

$$x \in G_{p,v} \Leftrightarrow \exists \lambda \in \mathbb{R} : x_1 = p_1 + \lambda v_1 \text{ und } x_2 = p_2 + \lambda v_2.$$

Reskalieren wir diese Gleichungen und subtrahieren sie, so erhalten wir

$$v_2 x_1 - v_1 x_2 = v_2 p_1 - v_1 p_2.$$

Also gilt:

$$G_{p,v} \subset \left\{ x \in \mathbb{R}^2, \langle x, \begin{pmatrix} v_2 \\ -v_1 \end{pmatrix} \rangle = \langle p, \begin{pmatrix} v_2 \\ -v_1 \end{pmatrix} \rangle \right\}.$$

Definition I.8.13. Es sei $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ und $\mu \in \mathbb{R}$.

(a) Es sei $v^\perp := \begin{pmatrix} v_2 \\ -v_1 \end{pmatrix}$.

(b) Ist $v \neq 0$, so sei

$$H_{v,\mu} := \{x \in \mathbb{R}^2, \langle x, v \rangle = \mu\} = \{x \in \mathbb{R}^2, x_1 v_1 + x_2 v_2 = \mu\}.$$

Durch Nachrechnen sehen Sie direkt:

Lemma I.8.14. *Für alle $v, w \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$ gilt:*

- $(v + w)^\perp = v^\perp + w^\perp$,
- $(\lambda v)^\perp = \lambda(v^\perp)$,
- $||v^\perp|| = ||v||$,
- $(v^\perp)^\perp = -v$,
- $\langle v, w^\perp \rangle = -\langle v^\perp, w \rangle$. *Insbesondere ist $\langle v, v^\perp \rangle = -\langle v^\perp, v \rangle = -\langle v, v^\perp \rangle$ und da die einzige reelle Zahl, die gleich ihrem Negativen ist, die 0 ist, steht $v^\perp$ also immer senkrecht auf v.*

Satz I.8.15. *Es sei $0 \neq v \in \mathbb{R}^2$ und $\mu \in \mathbb{R}$. Dann ist $H_{v,\mu}$ eine Gerade im $\mathbb{R}^2$ und zwar*

$$H_{v,\mu} = G_{\frac{\mu}{||v||^2} v, v^\perp}.$$

Beweis. Es sei $x \in G_{\frac{\mu}{||v||^2} v, v^\perp}$, also existiert ein $\lambda \in \mathbb{R}$ mit

$$x = \frac{\mu}{||v||^2} v + \lambda v^\perp.$$

Wir wenden die Abbildung $\langle -, v \rangle$ auf diese Gleichung an und erhalten:

$$\begin{aligned} \langle x, v \rangle &= \langle \frac{\mu}{||v||^2} v, v \rangle + \langle \lambda v^\perp, v \rangle \\ &= \frac{\mu}{||v||^2} \langle v, v \rangle + \lambda \langle v^\perp, v \rangle \\ &= \mu \frac{||v||^2}{\langle v, v \rangle} + \lambda 0 \\ &= \mu + 0 = \mu \end{aligned}$$

und damit ist $x \in H_{v,\mu}$.

Ist umgekehrt $x \in H_{v,\mu}$ so gilt

$$\langle x - \frac{\mu}{||v||^2} v, v \rangle = \langle x, v \rangle - \frac{\mu}{||v||^2} \langle v, v \rangle = \mu - \mu = 0.$$

Wir setzen $y := x - \frac{\mu}{||v||^2} v$ und erhalten in Koordinaten

$$0 = \langle y, v \rangle = y_1 v_1 + y_2 v_2.$$

Wir machen eine Fallunterscheidung:

Fall 1: Ist $v_1 \neq 0$, so können wir die obige Gleichung umformen zu $y_1 = -\frac{y_2}{v_1} v_2$ und wir erhalten damit

$$-\frac{y_2}{v_1} v^\perp = -\frac{y_2}{v_1} \begin{pmatrix} v_2 \\ -v_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y.$$

Fall 2: Ist $v_1 = 0$, so kann v_2 nicht null sein, weil $v \neq 0$. Damit können wir y_2 ausdrücken als

$$y_2 = -\frac{y_1}{v_2} v_1$$

und damit ist $y = \frac{y_1}{v_2} v^\perp$.

In beiden Fällen gibt es also ein $\lambda \in \mathbb{R}$ mit $y = \lambda v^\perp$ und wir erhalten, dass

$$x = \frac{\mu}{||v||^2} v + y = \frac{\mu}{||v||^2} v + \lambda v^\perp$$

und dies ist ein Element von $G_{\frac{\mu}{||v||^2} v, v^\perp}$. $\square$

KAPITEL II

Algebraische Grundbegriffe

II.1. Gruppen

Definition II.1.1. Eine *Gruppe* ist ein Paar $(G, \cdot)$ bestehend aus einer Menge G und einer Abbildung

$$\cdot\colon G \times G \to G, \quad (g, h) \mapsto g \cdot h,$$

genannt *Verknüpfung*, so dass gilt:

(G1) Für alle $g, h, k \in G$ ist

$$g \cdot (h \cdot k) = (g \cdot h) \cdot k.$$

(G2) Es gibt ein $e \in G$ mit der Eigenschaft, dass $e \cdot g = g$ ist für alle $g \in G$.

(G3) Für alle $g \in G$ gibt es ein $g' \in G$, so dass $g' \cdot g = e$.

Bemerkung II.1.2. Gruppen haben wegen (G2) immer mindestens ein Element; insbesondere sind sie nie die leere Menge. Wir nennen e *neutrales Element* der Gruppe.

Wir fordern *nicht*, dass $g \cdot h = h \cdot g$ gilt!

Wegen (G1) ist es legitim, Klammern wegzulassen.

Beispiele II.1.3.

- Die euklidische Ebene $(\mathbb{R}^2, +)$ ist eine Gruppe mit $e = 0$.
- Ist $\mathbf{n} = \{1, \ldots, n\}$, so ist die Menge aller bijektiven Abbildungen $\{f\colon \mathbf{n} \to \mathbf{n}, f \text{ bijektiv}\}$ eine Gruppe mit der Komposition von Abbildungen. Sie heißt die symmetrische Gruppe auf n Elementen und wir benutzen die Notation Σ_n.

Satz II.1.4. *Es sei $(G, \cdot)$ eine Gruppe mit neutralem Element e.*

(a) *Für alle $g \in G$ gilt: $g \cdot e = g$.*

(b) *Das neutrale Element ist eindeutig.*

(c) *Das Element g' in (G3) mit $g' \cdot g = e$ ist eindeutig. Wir schreiben g^{-1} dafür.*

(d) *Für alle $g \in G$ gilt: $g \cdot g' = e$.*

Beweis. Zu (d): Es sei $g \in G$ beliebig. Aus (G3) folgt: $\exists g : g' \cdot g = e$. Zu diesem g' gibt es wiederum nach (G3) ein $g'' \in G$ mit $g'' \cdot g' = e$. Damit erhalten wir aber

$$\begin{aligned} g \cdot g' &= e \cdot (g \cdot g') \\ &= (g'' \cdot g') \cdot (g \cdot g') \\ &= g'' \cdot ((g' \cdot g) \cdot g') \\ &= g'' \cdot (e \cdot g') \\ &= g'' \cdot g' = e. \end{aligned}$$

Das zeigt (d).

Zu (a): Wir erhalten mit (d)

$$g \cdot e = g \cdot (g' \cdot g) = (g \cdot g') \cdot g = e \cdot g = g$$

und das beweist (a).

Zu (b): Wir nehmen an, $\tilde{e}$ sei ein weiteres neutrales Element von G, also gilt $\tilde{e} \cdot g = g = g \cdot \tilde{e}$ für alle $g \in G$ nach (a). Mit $g = e$ folgt

$$\tilde{e} \cdot e = e = e \cdot \tilde{e}$$

aber da e neutral ist, ist $e \cdot \tilde{e} = \tilde{e}$, also $\tilde{e} = e$.

Zu (c): Es seien g' und g'' Elemente in G mit $g' \cdot g = e = g'' \cdot g$. Dann folgt aus (a) und (d) und (G1):

$$g'' = g'' \cdot e = g'' \cdot (g \cdot g') = (g'' \cdot g) \cdot g' = e \cdot g' = g'$$

also $g'' = g'$. □

Satz II.1.5. *Es sei $(G, \cdot)$ eine Gruppe und $g, h \in G$. Dann gilt:*

(a) *Es gibt ein eindeutiges $x \in G$ mit $x \cdot g = h$.*
(b) *Es gibt ein eindeutiges $y \in G$ mit $g \cdot y = h$.*
(c) $(g^{-1})^{-1} = g$.
(d) $(g \cdot h)^{-1} = h^{-1} \cdot g^{-1}$.

Die Gleichheit (d) kann man sich merken, wenn man an Pullover und Jacken (oder Socken und Schuhe) denkt: Wenn Sie sich anziehen, dann ziehen Sie erst den Pullover an und dann die Jacke. Beim Ausziehen ist es umgekehrt: Sie ziehen erst die Jacke aus und danach den Pullover.

Beweis. Bei (a) zeigen wir erst die Eindeutigkeit: Wenn es ein solches x gibt, dann erfüllt es die Gleichung

$$x = x \cdot e = x \cdot (g \cdot g^{-1}) = (x \cdot g) \cdot g^{-1} = h \cdot g^{-1}$$

und damit haben wir keine Wahl für x. Dies zeigt ebenso die Existenz, weil $x = h \cdot g^{-1}$ die Bedinung erfüllt:

$$x \cdot g = (h \cdot g^{-1}) \cdot g = h \cdot (g^{-1} \cdot g) = h \cdot e = h.$$

Behauptung (b) zeigen Sie analog zu (a).

Zu (c): Das Element $(g^{-1})^{-1}$ ist das eindeutige Inverse zu g^{-1}, aber es gilt auch $e = g \cdot g^{-1}$ und damit müssen beide Elemente gleich sein.

Teil (d) rechnen wir nach:

$$\begin{aligned}(h^{-1} \cdot g^{-1}) \cdot (g \cdot h) =& h^{-1} \cdot (g^{-1} \cdot (g \cdot h)) \\ =& h^{-1} \cdot ((g^{-1} \cdot g) \cdot h) \\ =& h^{-1} \cdot (e \cdot h) \\ =& h^{-1} \cdot h = e.\end{aligned}$$

□

Weitere wichtige Beispiele für Gruppen sind:

Beispiele II.1.6.

(a) $(\mathbb{Z}, +)$ ist eine Gruppe mit $e = 0$ und das Inverse zu $n \in \mathbb{Z}$ ist $-n$.
(b) Analog sind $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ Gruppen.
(c) $(\mathbb{N}_0, +)$ ist *keine* Gruppe, weil es zu $n > 0$ kein Inverses in $\mathbb{N}_0$ gibt.
(d) $(\mathbb{Q} \setminus \{0\}, \cdot)$ ist eine Gruppe mit $e = 1$. Dagegen ist $(\mathbb{Q}, \cdot)$ *keine* Gruppe, weil $0 \in \mathbb{Q}$ kein Inverses hat.
(e) Es sei $\varnothing \neq X$ eine beliebige Menge. Dann ist $\mathsf{Sym}(X) := \{f \colon X \to X, f \text{ bijektiv}\}$ mit der Verkettung von Abbildungen eine Gruppe mit $e = \mathrm{id}_X$. Für $X = \mathbf{n}$ hatten wir diese Gruppe Σ_n genannt.
(f) $G = \{e\}$ mit $e \cdot e = e$ ist eine Gruppe, die *triviale Gruppe*.

Definition II.1.7. Eine Gruppe $(G, \cdot)$ heißt *abelsch*, falls für alle $g, h \in G$ gilt:

$$g \cdot h = h \cdot g.$$

Niels Henrik Abel, 1802–1829.

Bemerkung II.1.8. In der obigen Beispielliste sind (a), (b), (d) und (f) abelsche Gruppen. Dagegen ist $\mathsf{Sym}(X)$ *nicht* abelsch, falls $|X| > 2$. Überlegen Sie sich bitte explizit ein Gegenbeispiel.

Definition II.1.9. Es sei $(G, \cdot)$ eine Gruppe. Eine Teilmenge $H \subset G$ heißt *Untergruppe*, falls gilt:

(UG1) $H \neq \varnothing$,
(UG2) Für alle $h, h' \in H$ gilt: $h \cdot h' \in H$.

(UG3) Für alle $h \in H$ ist $h^{-1} \in H$.

Das heißt, dass H unter der Verknüpfung und Inversenbildung abgeschlossen ist.

Satz II.1.10. *Ist $(G, \cdot)$ eine Gruppe mit neutralem Element e und $H \subset G$ sei Untergruppe von G. Dann gilt:*

- $e \in H$,
- *Mit der Einschränkung von $\cdot$*
$$\cdot|_{H\times H} \colon H \times H \to H$$
ist H eine Gruppe.

Beweis. Da $H \neq \varnothing$ gibt es ein $h \in H$. Mit (UG3) ist dann h^{-1} auch in H und mit (UG2) auch $h \cdot h^{-1} = e$.

Mit (UG2) erhalten wir, dass die Einschränkung von $\cdot$ auf $H \times H$ wiederum Werte in H annimmt. Die Axiome (G1) und (G2) gelten für alle $g \in G$ also auch insbesondere für alle $h \in H$. Axiom (UG3) impliziert (G3) für H. □

Bemerkung II.1.11. Ist $H \subset G$ und $K \subset H$ jeweils eine Untergruppe, so auch $K \subset G$.

Untergruppen abelscher Gruppen sind abelsch.

Beispiele II.1.12.

- Für jede Gruppe $(G, \cdot)$ mit neutralem Element e sind $(\{e\}, \cdot)$ und $(G, \cdot)$ Untergruppen.
- Für $(G, \cdot) = (\mathbb{R}, +)$ ist sowohl $(\mathbb{Q}, +)$ als auch $(\mathbb{Z}, +)$ eine Untergruppe und $(\mathbb{Z}, +) \subset (\mathbb{Q}, +)$ ist Untergruppe.
- $(\mathbb{N}_0, +) \subset (\mathbb{Z}, +)$ ist *keine* Untergruppe.

Wir betrachten Abbildungen zwischen Gruppen, welche die Gruppenstruktur respektieren:

Definition II.1.13. Es seien $(G, \cdot)$ und $(G', *)$ Gruppen. Eine Abbildung $f \colon G \to G'$ heißt *Gruppenhomomorphismus* (oder kurz *Homomorphismus*), falls für alle $g_1, g_2 \in G$ gilt:
$$f(g_1 \cdot g_2) = f(g_1) * f(g_2).$$

Beispiele II.1.14.

- Für $(G, \cdot) = (G', *) = (\mathbb{Z}, +)$ und für ein festes $m \in \mathbb{Z}$ ist die Abbildung
$$f \colon \mathbb{Z} \to \mathbb{Z}, \quad x \mapsto mx$$
ein Homomorphismus, weil
$$f(x + y) = m(x + y) = mx + my = f(x) + f(y).$$
- Es sei $\mathbb{R}_{>0} := \{x \in \mathbb{R}, x > 0\}$. Für $(G, \cdot) = (\mathbb{R}, +)$ und $(G', *) = (\mathbb{R}_{>0}, \cdot)$ definiert die Exponentialfunktion $f(x) = e^x$ einen Homomorphismus, weil
$$f(x + y) = e^{x+y} = e^x \cdot e^y = f(x) \cdot f(y).$$

Satz II.1.15. *Sind $(G, \cdot)$ und $(G', *)$ Gruppen mit neutralem Element e_G beziehungsweise $e_{G'}$ und ist $f \colon G \to G'$ ein Homomorphismus, so gilt:*

(a) $f(e_G) = e_{G'}$.
(b) *Für alle $g \in G$: $f(g)^{-1} = f(g^{-1})$.*
(c) *Ist f zusätzlich bijektiv, so ist $f^{-1} \colon G' \to G$ ebenfalls ein Homomorphismus.*

Beweis. (a): Wir multiplizieren die Gleichung $f(e_G) = f(e_G \cdot e_G) = f(e_G) * f(e_G)$ von links mit dem Inversen von $f(e_G)$ in G' und erhalten
$$\begin{aligned} e_{G'} &= f(e_G)^{-1} * f(e_G) \\ &= f(e_G)^{-1} * (f(e_G) * f(e_G)) \\ &= (f(e_G)^{-1} * f(e_G)) * f(e_G) \\ &= f(e_G). \end{aligned}$$

Für (b) rechnen wir nach, dass

$$f(g^{-1}) * f(g) = f(g^{-1} \cdot g) = f(e_G)$$

und mit (a) folgt, dass dies gleich $e_{G'}$ ist. Da das Inverse eines Elementes eindeutig ist, folgt $f(g^{-1}) = f(g)^{-1}$.

Bei (c) ist zu zeigen, dass $f^{-1}(g_1') \cdot f^{-1}(g_2') = f^{-1}(g_1' * g_2')$ ist für alle $g_1', g_2' \in G'$. Da $f^{-1} \circ f = \mathrm{id}_G$ ist, erhalten wir:

$$\begin{aligned} f^{-1}(g_1') \cdot f^{-1}(g_2') =& (f^{-1} \circ f)(f^{-1}(g_1') \cdot f^{-1}(g_2')) \\ =& f^{-1}(f(f^{-1}(g_1') \cdot f^{-1}(g_2'))). \end{aligned}$$

und da f ein Homomorphismus ist und da $f \circ f^{-1} = \mathrm{id}_{G'}$ ist, stimmt dies überein mit

$$f^{-1}(f(f^{-1}(g_1')) * f(f^{-1}(g_2'))) = f^{-1}(g_1' * g_2').$$

□

Beispiel II.1.16. Die Exponentialabbildung $e\colon (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$ ist ein Homomorphismus und bijektiv. Die Umkehrabbildung, der natürliche Logarithmus, $\ln\colon (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +)$ ist also ebenfalls ein bijektiver Homomorphismus, also gilt für alle $\mathbb{R} \ni x, y \neq 0$:

$$\ln(x \cdot y) = \ln(x) + \ln(y).$$

Definition II.1.17. Zwei Gruppen $(G, \cdot)$ und $(G', *)$ heißen *isomorph*, falls es einen bijektiven Homomorphismus $f\colon G \to G'$ gibt. Die Notation dafür ist $G \cong G'$.

Weisen Sie nach, dass die Isomorphie von Gruppen eine Äquivalenzrelation ist.

Beispiele II.1.18.

- Wir wissen, dass $(\mathbb{R}_{>0}, \cdot) \cong (\mathbb{R}, +)$. Es sei $\alpha \in \mathbb{R} \setminus \{0\}$. Ist dann $f_\alpha\colon (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$, $x \mapsto e^{\alpha x}$ auch ein Isomorphismus?
- Es sei $m \in \mathbb{Z}$ gewählt. Die Abbildung $f\colon (\mathbb{Z}, +) \to (\mathbb{Z}, +)$, $x \mapsto mx$ ist nur für $m \in \{-1, +1\}$ ein Isomorphismus.

Definition II.1.19. Es seien $(G, \cdot)$ und $(G', *)$ zwei Gruppen und $f\colon G \to G'$ sei ein Homomorphismus.

- $\ker(f) := f^{-1}(e_{G'}) = \{g \in G, f(g) = e_{G'}\}$ heißt der *Kern von* f und
- $\mathrm{Bild}(f) = f(G)$ heißt das *Bild von* f.

Satz II.1.20. *Es seien $(G, \cdot)$ und $(G', *)$ zwei Gruppen und $f\colon G \to G'$ sei ein Homomorphismus. Dann gilt:*

(a) *Der Kern von f ist eine Untergruppe von G.*
(b) *Die Abbildung f ist genau dann injektiv, wenn $\ker(f) = \{e_G\}$.*
(c) *Das Bild von f ist eine Untergruppe von G'.*
(d) *Die Abbildung f ist genau dann surjektiv, wenn Bild$(f) = G'$.*

Beweis. Die Aussage (d) ist genau die Definition der Surjektivität.

Bei (a) rechnen wir die Untergruppenaxiome nach.

(UG1): Da immer gilt $f(e_G) = e_{G'}$ ist der Kern von f nicht die leere Menge, weil $e_G \in \ker(f)$.

(UG2): Sind $g_1, g_2 \in \ker(f)$, also $f(g_1) = e_{G'} = f(g_2)$, so ist auch

$$f(g_1 \cdot g_2) = f(g_1) * f(g_2) = e_{G'} * e_{G'} = e_{G'}.$$

(UG3): Ist $g \in \ker(f)$, so ist $f(g^{-1}) = f(g)^{-1} = e_{G'}^{-1} = e_{G'}$.

Zu (b): Wir nehmen an, dass f injektiv ist und dass $g \in \ker(f)$. Wir wissen aber auch, dass $e_G \in \ker(f)$, also $f(g) = e_{G'} = f(e_G)$. Aus der Injektivität folgt $g = e_G$.

Umgekehrt sei $\ker(f) = \{e_G\}$ und für g_1, g_2 gelte $f(g_1) = f(g_2)$. Dann ist

$$e_{G'} = f(g_2) * f(g_1)^{-1} = f(g_2 \cdot g_1^{-1}),$$

so dass $g_2 \cdot g_1^{-1} \in \ker(f)$. Nach Voraussetzung ist $g_2 \cdot g_1^{-1} = e_G$ also $g_1 = g_2$.

Zu (c):

(UG1): Das Bild von f ist nicht leer, weil $f(e_G) = e_{G'}$, also $e_{G'} \in \mathrm{Bild}(f)$.

(UG2): Sind g_1', g_2' im Bild von f, so gibt es $g_1, g_2 \in G$ mit $f(g_1) = g_1'$ und $f(g_2) = g_2'$. Daraus folgt

$$g_1' * g_2' = f(g_1) * f(g_2) = f(g_1 \cdot g_2)$$

und damit ist $g_1' * g_2'$ ebenfalls im Bild von f.

(UG3): Ist g' im Bild von f, so gibt es ein $g \in G$ mit $f(g) = g'$. Dann gilt auch

$$(g')^{-1} = f(g)^{-1} = f(g^{-1})$$

und $(g')^{-1}$ ist im Bild von f. □

Zum Abschluss dieses Abschnitts betrachten wir eine sehr wichtige Beispielklasse abelscher Gruppen, die wir in der Vorlesung oft benutzen werden: Es sei $m \in \mathbb{N}_0$ und R_m sei die Relation auf $\mathbb{Z}$, die definiert ist durch

$$(x, y) \in R_m \Leftrightarrow m \text{ teilt } y - x.$$

Dies ist eine Äquivalenzrelation. Man schreibt $\mathbb{Z}/m\mathbb{Z}$ für die Quotientenmenge $\mathbb{Z}/R_m$. Für die Äquivalenzklassen benutzen wir die Notation:

$$r + m\mathbb{Z} := \{x \in \mathbb{Z}, m \text{ teilt } x - r\}.$$

Oft kürzen wir diese Menge mit $\overline{r}$ ab. Vorsicht, $\overline{r}$ hängt von m ab, auch wenn das in dieser Notation nicht sichtbar ist!

Die Menge $r + m\mathbb{Z}$ heißt *Restklasse von r modulo m*.

Für x und x' aus $\mathbb{Z}$ schreiben wir

$$x \equiv x' \bmod m$$

und sagen, dass x *kongruent* ist zu x' modulo m, falls x und x' in der gleichen Restklasse modulo m liegen.

Beispiele II.1.21.

- Der Fall $m = 0$ ist nicht sonderlich interessant: $(x, y) \in R_0$ gilt genau dann, wenn $x = y$ ist. Damit ist
$$\mathbb{Z}/R_0 = \mathbb{Z}/0\mathbb{Z} = \mathbb{Z}.$$
- Das andere Extrem ist $m = 1$, weil 1 jede Zahl teilt, also ist
$$\mathbb{Z}/R_1 = \mathbb{Z}/1\mathbb{Z}$$
eine Menge mit einem Element.

Interessant sind die Beispiele $\mathbb{Z}/m\mathbb{Z}$ für $m \geqslant 2$. Wir machen $\mathbb{Z}/m\mathbb{Z}$ zu einer Gruppe:

Wir definieren die Verknüpfung durch

$$\overline{r} + \overline{s} := \overline{r + s}.$$

Wir müssen zeigen, dass diese Definition nicht von der Wahl der Repräsentanten abhängt: Ist $\overline{r}_1 = \overline{r}_2$ und $\overline{s_1} = \overline{s}_2$, so ist zu beweisen, dass $\overline{r_1 + s_1} = \overline{r_2 + s_2}$ gilt.

Nach Definition gibt es $a, b \in \mathbb{Z}$, so dass $r_1 - r_2 = ma$ und $s_1 - s_2 = mb$ gilt. Damit können wir $r_1 + s_1$ ausdrücken als

$$r_1 + s_1 = r_2 + ma + s_2 + mb = r_2 + s_2 + m(a + b)$$

und somit ist $\overline{r_1 + s_1} = \overline{r_2 + s_2}$.

Satz II.1.22. *Für alle $m \in \mathbb{N}_0$ ist $\mathbb{Z}/m\mathbb{Z}$ eine abelsche Gruppe und die Abbildung*

$$\pi \colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, \quad r \mapsto \pi(r) = \overline{r} = r + m\mathbb{Z}$$

ist ein surjektiver Homomorphismus mit $\ker(\pi) = m\mathbb{Z}$.

Beweis. Wir rechnen zunächst die Gruppenaxiome nach: (G1): Sind $r, s, t \in \mathbb{Z}$ so gilt

$$\begin{aligned}(\overline{r} + \overline{s}) + \overline{t} =& \overline{r + s} + \overline{t}\\ =& \overline{(r + s) + t}\\ =& \overline{r + (s + t)}\\ =& \overline{r} + \overline{s + t}\\ =& \overline{r} + (\overline{s} + \overline{t}).\end{aligned}$$

(G2): Das Element $\overline{0}$ ist das neutrale Element, weil

$$\overline{0}+\overline{r}=\overline{0+r}=\overline{r}=\overline{r+0}$$

gilt für alle $r \in \mathbb{Z}$.

(G3): Da $\overline{r}+\overline{-r}=\overline{r-r}=\overline{0}$ gilt, ist $\overline{-r}$ das Inverse zu $\overline{r}$ für alle $r \in \mathbb{Z}$.

Da in $\mathbb{Z}/m\mathbb{Z}$ für alle $r, s \in \mathbb{Z}$ gilt

$$\overline{r}+\overline{s}=\overline{r+s}=\overline{s+r}=\overline{s}+\overline{r},$$

ist $\mathbb{Z}/m\mathbb{Z}$ abelsch.

Wir haben in Lemma I.5.11 gesehen, dass π surjektiv ist. Der Kern von π ist

$$\begin{aligned}\ker(\pi) =& \{x \in \mathbb{Z}, \overline{x}=\overline{0}\}\\ =& \{x \in \mathbb{Z}, x+m\mathbb{Z}=0+m\mathbb{Z}\}\\ =& \{x \in \mathbb{Z}, x+m\mathbb{Z}=m\mathbb{Z}\}\\ =& \{x \in \mathbb{Z}, m \text{ teilt } x\}\\ =& m\mathbb{Z}.\end{aligned}$$

□

Definition II.1.23. Die Gruppe $\mathbb{Z}/m\mathbb{Z}$ heißt *zyklische Gruppe der Ordnung m.*

Wir werden später ein geometrischen Modell für $\mathbb{Z}/m\mathbb{Z}$ kennenlernen, welches komplexe Zahlen benutzt.

II.2. Ringe und Körper

Wir betrachten algebraische Strukturen mit zwei Verknüpfungen, die zueinander passen. Zum Beispiel können Sie ganze Zahlen addieren und multiplizieren und die Distributivgesetze klären, wie sich beide Verknüpfungen miteinander vertragen.

Definition II.2.1.

(a) Eine Menge R mit zwei Verknüpfungen

$$\begin{aligned} +&\colon R \times R \to R \text{ und}\\ \cdot&\colon R \times R \to R\end{aligned}$$

heißt ein *Ring*, falls gilt:

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) $\cdot$ ist assoziativ, das heißt, dass für alle $a, b, c \in R$ gilt

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

(R3) Es gelten die Distributivgesetze. Für alle $a, b, c \in R$:

$$\begin{aligned} a \cdot (b+c) =& a \cdot b + a \cdot c,\\ (a+b) \cdot c =& a \cdot c + b \cdot c.\end{aligned}$$

(b) Ein Element $1 \in R$ heißt *Einselement*, falls für alle $a \in R$ gilt:

$$a \cdot 1 = a = 1 \cdot a.$$

(c) Ein Ring heißt *kommutativ*, falls für alle $a, b \in R$ gilt

$$a \cdot b = b \cdot a.$$

Bemerkung II.2.2.

- Es gibt den Nullring $R=\{0\}$, der nur aus dem Element 0 besteht mit $0+0=0$ und $0 \cdot 0 = 0$. Hier ist das Einselement gleich 0. Vorsicht: Viele allgemeine Eigenschaften von Ringen gelten für den Nullring nicht und wir werden ihn oft ausschließen aus unseren Betrachtungen.
- Wir benutzen „Punkt vor Strichrechnung“ in Ringen.

- Ist 0 das neutrale Element in $(R,+)$, so gilt:

$$0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a$$

und damit ist $0 \cdot a = 0$. Analog zeigen Sie, dass auch $a \cdot 0 = 0$ gilt.

Beispiele II.2.3.

(a) $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ sind kommutative Ringe und das Einselement ist die Zahl 1.

(b) Für $m \in \mathbb{Z}$, $m \neq \pm 1$ ist $m\mathbb{Z}$ ein Ring ohne Einselement.

(c) Ist $X \neq \varnothing$ eine Menge und ist R ein Ring, so trägt die Menge aller Abbildungen von X nach R eine Ringstruktur: Wir setzen $\mathrm{Abb}(X,R)$ als die Menge aller Abbildungen von X nach R und definieren für $f,g\colon X \to R$:

$$\begin{aligned}(f+g)(x) :=& f(x)+g(x),\\ (f \cdot g)(x) :=& f(x) \cdot g(x).\end{aligned}$$

Hierbei sind $+$ und $\cdot$ auf der rechten Seite die Verknüpfungen in R. Dies definiert eine Ringstruktur. Hat R ein Einselement 1, so ist die konstante Abbildung mit Wert 1 ein Einselement für $\mathrm{Abb}(X,R)$. Ist R kommutativ, so ist $\mathrm{Abb}(X,R)$ mit der obigen Struktur es auch.

Wir können Abbildungen mit Zusatzstruktur betrachten. In der Analysis werden Sie oft Beispiele der Art betrachten, bei denen $X = [a,b]$ ein abgeschlossenes Intervall mit $[a,b] \subset \mathbb{R}$ ist und $R = \mathbb{R}$. Dann haben wir neben $\mathrm{Abb}([a,b],\mathbb{R})$ die wichtigen Ringe

$$\{f\colon [a,b] \to \mathbb{R}, f \text{ stetig}\}$$

und

$$\{f\colon [a,b] \to \mathbb{R}, f \text{ stetig und stetig differenzierbar}\}.$$

(d) Wir betrachten wieder ein festes $m \in \mathbb{N}_0$ und $\mathbb{Z}/m\mathbb{Z}$. Zusätzlich zur Addition von Restklassen können wir eine Multiplikation definieren durch

$$\overline{r} \cdot \overline{s} := \overline{rs}.$$

Diese Multiplikation ist wohldefiniert, also unabhängig von der Wahl der Repräsentanten: Ist $\overline{r}_1 = \overline{r}_2$ und $\overline{s}_1 = \overline{s}_2$, so gibt es wiederum ganze Zahlen a,b mit

$$r_1 - r_2 = ma \text{ und } s_1 - s_2 = mb.$$

Damit ist

$$r_1 s_1 = (r_2 + ma)(s_2 + mb) = r_2 s_2 + m(mab + as_2 + r_2 b)$$

und daher ist

$$\overline{r}_1 \cdot \overline{s}_1 = \overline{r_1 s_1} = \overline{r_2 s_2} = \overline{r}_2 \cdot \overline{s}_2.$$

Mit den Verknüpfungen $+$ und $\cdot$ ist $\mathbb{Z}/m\mathbb{Z}$ ein kommutativer Ring mit Einselement für $m \neq 1$.

Für $m = 4$ betrachten wir konkret die Additions- und die Multiplikationstabelle.

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Definition II.2.4. Ein Ring $R \neq \{0\}$ heißt *nullteilerfrei*, falls aus $a \cdot b = 0$ in R folgt, dass $a = 0$ oder $b = 0$.

Wir haben oben gesehen, dass $\mathbb{Z}/4\mathbb{Z}$ *nicht* nullteilerfrei ist, weil $\overline{2} \neq \overline{0}$, aber $\overline{2} \cdot \overline{2} = \overline{0}$. Dies gilt allgemeiner:

Satz II.2.5. *Es sei $1 \neq m \in \mathbb{N}$. Dann ist $\mathbb{Z}/m\mathbb{Z}$ genau dann nullteilerfrei, wenn m eine Primzahl ist.*

Beweis. Ist m keine Primzahl, so gibt es natürliche Zahlen k, ℓ mit $1 < k, \ell < m$ und $m = k\ell$. Damit ist $\overline{k} \neq \overline{0}$ und $\overline{\ell} \neq \overline{0}$, aber

$$\overline{0} = \overline{m} = \overline{k} \cdot \overline{\ell}$$

und damit ist $\mathbb{Z}/m\mathbb{Z}$ nicht nullteilerfrei.

Ist m eine Primzahl und ist $\overline{k} \cdot \overline{\ell} = \overline{0}$, so ist $k\ell = am$ für ein $a \in \mathbb{Z}$. Wegen der Eindeutigkeit der Primfaktorzerlegung muss dann m die Zahl k oder ℓ teilen. Dann ist aber deren Restklasse null. □

Wir definieren wieder geeignete Unterobjekte und Abbildungen:

Definition II.2.6.

(a) Ist R ein Ring und $R' \subset R$ eine Teilmenge, so heißt R' ein *Unterring*, falls $(R', +) \subset (R, +)$ eine Untergruppe ist und wenn für alle $a, b \in R'$ das Element $a \cdot b$ ein Element von R' ist.

(b) Hat R ein Einselement 1_R und ist $R' \subset R$, dann verlangt man zusätzlich, dass 1_R auch ein Einselement für R' ist.

(c) Sind $(R, +, \cdot)$ und $(S, +, *)$ Ringe und ist $f\colon R \to S$ eine Abbildung, so heißt f ein *Ringhomomorphismus*, falls für alle $a, b \in R$ gilt:

$$f(a + b) = f(a) + f(b) \text{ und } f(a \cdot b) = f(a) * f(b).$$

(d) Haben R und S Einselemente 1_R beziehungsweise 1_S, so verlangt man zusätzlich, dass $f(1_R) = 1_S$ gilt.

Definition II.2.7. Ein Körper K ist eine Menge K mit zwei Verknüpfungen $+$ und $\cdot$

$$+, \cdot\colon K \times K \to K,$$

so dass gilt

(K1) $(K, +)$ ist eine abelsche Gruppe und

(K2) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

(K3) Die Distributivgesetze

$$\begin{aligned}(a + b) \cdot c =& a \cdot c + b \cdot c \text{ und}\\ a \cdot (b + c) =& a \cdot b + a \cdot c\end{aligned}$$

gelten für alle $a, b, c \in K$.

Bemerkung II.2.8.

- Körper sind insbesondere nullteilerfreie, kommutative Ringe.
- Ist $(K \setminus \{0\}, \cdot)$ nur eine nicht-abelsche Gruppe, aber es gelten alle anderen Körperaxiome, so heißt K *Schiefkörper*.
- Oft schreibt man statt a^{-1} kurz $\frac{1}{a}$ für $a \in K \setminus \{0\}$.
- Wegen (K2) ist $K \setminus \{0\} \neq \varnothing$. Insbesondere ist ein Körper niemals der Nullring und $1 \neq 0$ in K.

Beispiel II.2.9. Die Ringe $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper, aber $(\mathbb{Z}, +, \cdot)$ ist *kein* Körper.

Satz II.2.10. *Ist p eine Primzahl, so ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper.*

Beweis. Wir wissen schon von Satz II.2.5, dass $\mathbb{Z}/p\mathbb{Z}$ ein nullteilerfreier, kommutativer Ring mit Einselement ist. Es bleibt zu zeigen, dass alle $\overline{r} \in \mathbb{Z}/p\mathbb{Z} \setminus \{\overline{0}\}$ ein multiplikatives Inverses haben.

Wir betrachten für ein festes r mit $\overline{r} \neq \overline{0}$ die Abbildung

$$f_r\colon \mathbb{Z}/p\mathbb{Z} \setminus \{0\} \to \mathbb{Z}/p\mathbb{Z} \setminus \{0\}, \quad \overline{s} \mapsto \overline{rs}.$$

Die Abbildung f_r ist injektiv: Ist $\overline{rs} = \overline{rs'}$, dann ist das genau dann der Fall, wenn $\overline{r}(\overline{s} - \overline{s'}) = 0$. Da $\mathbb{Z}/p\mathbb{Z}$ ein nullteilerfreier Ring ist, folgt $\overline{s} - \overline{s'} = 0$. Also ist $\overline{s} = \overline{s'}$.

Damit ist f_r automatisch bijektiv, weil $\mathbb{Z}/p\mathbb{Z} \setminus \{0\}$ eine endliche Menge ist. Insbesondere ist $\overline{1}$ im Bild von f_r. □

Definition II.2.11. Für jede Primzahl p schreiben wir $\mathbb{F}_p$ für $\mathbb{Z}/p\mathbb{Z}$.

Das $\mathbb{F}$ steht hierbei für *field*, also Körper auf Englisch.

II.2.1. Der Körper der komplexen Zahlen. Wir versehen die reelle Ebene mit einer Körperstruktur. Wir wissen schon, dass $(\mathbb{R}^2, +)$ eine abelsche Gruppe ist.

Wir definieren ein Multiplikation auf $\mathbb{R}^2$: Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ definieren wir

$$x \cdot y := \begin{pmatrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix}. \tag{II.2.1}$$

Damit ist $(\mathbb{R}^2 \setminus \{0\}, \cdot)$ eine abelsche Gruppe:

(a) (G1) ist eine mühselige Rechnung, die Sie selbst durchführen.

(b) Das Element $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist das neutrale Element:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \cdot y_1 - 0 \cdot y_2 \\ 1 \cdot y_2 + 0 \cdot y_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

(c) Ist $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, so ist

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{-1} = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = \frac{1}{||x||^2} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

das multiplikative Inverse:

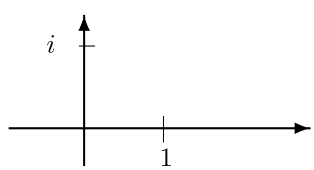
$$\frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1^2 - ((-x_2)x_2) \\ x_1 x_2 - x_1 x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(d) Dass die Distributivgesetze (K3) gelten, folgt mit einer langwierigen Rechnung, die wir hier unterschlagen.

Wir schreiben $\mathbb{C}$ für $\mathbb{R}^2$ mit dieser Körperstruktur, $\mathbb{C}$ heißt auch *komplexe Zahlenebene*.

Definition II.2.12.

(a) Das Element $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{C}$ heißt *imaginäre Einheit* und wird mit i bezeichnet.



(b) Ist $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$, so heißt x_1 der *Realteil von x* und x_2 der *Imaginärteil von x*. Wir kürzen dies ab mit $\mathsf{Re}(x) = x_1$ und $\mathsf{Im}(x) = x_2$.

Bemerkung II.2.13.

- Es gilt $i^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, also erfüllt i die Gleichung

$$i^2 = -1.$$

Wir können also i als eine Quadratwurzel aus -1 auffassen, aber Vorsicht: Was ist falsch an der Gleichungskette

$$-1 = i^2 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1?$$

- Wir können jedes $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$ schreiben als

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1 \cdot 1 + x_2 \cdot i$$

oder kurz $x_1 + x_2 i$.

- Mit dieser Umformulierung sieht die Multiplikation in $\mathbb{C}$ schon freundlicher aus: Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ist

$$x \cdot y = (x_1 + x_2 i)(y_1 + y_2 i) = x_1 y_1 + x_1 y_2 i + x_2 y_1 i + x_2 y_2 i^2 = (x_1 y_1 - x_2 y_2) + (x_1 y_2 + x_2 y_1) i.$$

  Wir brauchen also nur die Rechenregeln in $\mathbb{R}$, $i^2 = -1$ und die Distributivgesetze.

Definition II.2.14.

(a) Die Abbildung

$$\overline{} \colon \mathbb{C} \to \mathbb{C}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \overline{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

heißt *komplexe Konjugation*. Wir schreiben auch $\overline{x_1 + x_2 i} = x_1 - x_2 i$.

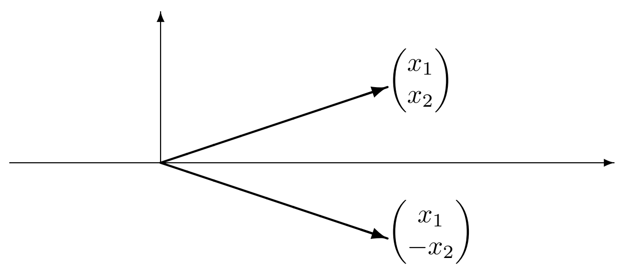
(b) Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$ heißt

$$|x| = \sqrt{x_1^2 + x_2^2}$$

der Absolutbetrag von x. Dies ist nichts anderes als die euklidische Norm des Vektors $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

Bemerkung II.2.15.

(a) Geometrisch ist die komplexe Konjugation die Spiegelung an der x_1-Achse



(b) Traditionell werden Elemente in $\mathbb{C}$ oft mit $z = x + iy$ notiert. Wir werden diese Schreibweise ab jetzt benutzen.

Satz II.2.16. *Für alle $z, w \in \mathbb{C}$ gilt:*

(a) $\overline{\overline{z}} = z$,
(b) $\overline{z + w} = \overline{z} + \overline{w}$,
(c) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$,
(d) $\overline{0} = 0, \overline{1} = 1, \overline{i} = -i$,
(e) $|z + w| \leqslant |z| + |w|$,
(f) $|z \cdot w| = |z||w|$,
(g) $|z|^2 = z \cdot \overline{z}$. *Insbesondere gilt für alle $z \neq 0$, dass*

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

Beweis. Die Beweise der obigen Behauptungen bestehen aus elementaren Rechnungen. Wir führen exemplarisch (b) und (g) aus; (e) hatten wir schon in Satz I.8.9 gesehen.

Zu (b): Wir schreiben $z = x + iy$ und $w = u + iv$ und erhalten

$$\begin{aligned}\overline{z+w} =& \overline{x+iy+u+iv}\\ =& \overline{(x+u)+i(y+v)}\\ =& (x+u)-i(y+v)\\ =& x-iy+u-iv\\ =& \overline{z}+\overline{w}.\end{aligned}$$

Zu (g): Auch hier rechnen wir explizit nach:

$$z \cdot \overline{z} = (x+iy)\cdot(x-iy) = x^2 + iyx - ixy - i^2y^2 = x^2 + y^2 = |z|^2.$$

□

Wir beschreiben nun komplexe Zahlen durch ihre Länge und den Winkel:

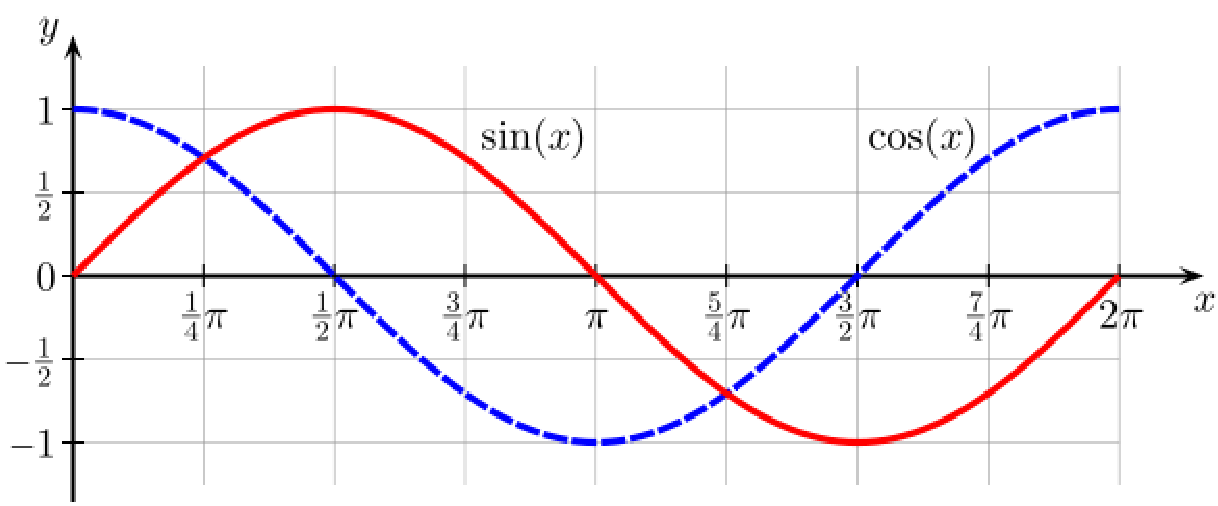
Definition II.2.17. Es sei $0 \neq z \in \mathbb{C}$. Dann heißt die Zahl aus $[0, 2\pi)$, die durch

$$\mathsf{arg}(z) := \begin{cases} \measuredangle(z,1), & \text{falls } \mathsf{Im}(z) \geqslant 0,\\ 2\pi - \measuredangle(z,1), & \text{falls } \mathsf{Im}(z) < 0,\end{cases}$$

festgelegt ist, das *Argument von z*.

Wir messen also den Winkel von 1 aus im mathematisch positiven Sinn, also gegen den Uhrzeigersinn. Zur Wiederholung: Wir hatten $\measuredangle$ in Definition I.8.11 definiert, also erhalten wir hier

$$\measuredangle(z,1) = \arccos\left(\frac{\langle z,1\rangle}{|z|}\right).$$



Quelle: Wiki Commons.

Satz II.2.18. *Ist $0 \neq z \in \mathbb{C}$, so gilt:*

(a)

$$\mathsf{Re}(z) = |z|\cos(\mathsf{arg}(z)), \quad \mathsf{Im}(z) = |z|\sin(\mathsf{arg}(z)).$$

(b) *Das Argument von z ist die eindeutig bestimmt reelle Zahl $\theta \in [0, 2\pi)$, für die gilt:*

$$z = |z|(\cos(\mathsf{arg}(z)) + i\sin(\mathsf{arg}(z))).$$

Beweis. Für (a) betrachten wir zunächst den Fall, dass der Imaginärteil von z nicht negativ ist, und damit haben wir $\mathsf{arg}(z) = \measuredangle(z, 1)$. Wir berechnen

$$\begin{aligned}
|z| \cos(\mathsf{arg}(z)) =& |z| \cos(\measuredangle(z, 1)) \\
=& |z| \cos(\arccos\left(\frac{\langle z, 1\rangle}{|z|}\right)) \\
=& |z| \frac{\langle z, 1\rangle}{|z|} \\
=& \langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle \\
=& x = \mathsf{Re}(z).
\end{aligned}$$

Ebenso ist

$$|z| \sin(\mathsf{arg}(z)) = |z| \sin(\arccos\left(\frac{\langle z, 1\rangle}{|z|}\right)).$$

Wir kürzen im Folgenden $(\cos(t))^2$ mit $\cos^2(t)$ und $(\sin(t))^2$ mit $\sin^2(t)$ ab. Da für alle $t \in \mathbb{R}$ gilt

$$\cos^2 t + \sin^2 t = 1$$

erhalten wir

$$\sin(t) = \pm\sqrt{1 - \cos^2 t}.$$

Wir nehmen an, dass $\mathsf{Im}(z) \geqslant 0$ und damit ist das Argument von z ein Element von $[0, \pi]$ mit der Konsequenz, dass $\sin(\mathsf{arg}(z)) \geqslant 0$. Dies liefert

$$\begin{aligned}
|z| \sin(\arccos\left(\frac{\langle z, 1\rangle}{|z|}\right)) =& |z| \sqrt{1 - \cos^2(\arccos\left(\frac{\langle z, 1\rangle}{|z|}\right))} \\
=& \sqrt{|z|^2 - \langle z, 1\rangle^2} \\
=& \sqrt{\mathsf{Re}(z)^2 + \mathsf{Im}(z)^2 - \mathsf{Re}(z)^2} \\
=& \sqrt{\mathsf{Im}(z)^2} \\
=& |\mathsf{Im}(z)|.
\end{aligned}$$

Aber da $\mathsf{Im}(z) \geqslant 0$, ist dies gleich $\mathsf{Im}(z)$.

Im zweiten Fall mit $\mathsf{Im}(z) < 0$ besteht der Beweis aus einer analogen Rechnung, bei der Sie beachten müssen, dass sowohl der Sinus als auch der Kosinus 2π-periodisch sind.

Zu (b): Ist $\theta = \mathsf{arg}(z)$, so gibt $|z| \cos(\theta)$ gerade den Realteil von z und $|z| \sin(\theta))$ gerade den Imaginärteil von z an. Zu zeigen ist die Eindeutigkeit von θ: Wir nehmen also an, dass wir z schreiben können als

$$z = |z|(\cos(\theta) + i \sin(\theta)) = |z|(\cos(\varrho) + i \sin(\varrho)),$$

wobei sowohl θ als auch ϱ in $[0, 2\pi)$ liegen. Da $|z| \neq 0$ gilt, ergibt ein Koeffizientenvergleich, dass

$$\cos(\theta) = \cos(\varrho) \text{ und } \sin(\theta) = \sin(\varrho) \tag{II.2.2}$$

gilt. Wir benutzen das Additionstheorem für den Sinus und schreiben

$$\sin(\theta - \varrho) = \sin(\theta) \cos(\varrho) - \cos(\theta) \sin(\varrho).$$

Wegen der Identitäten aus (II.2.2) ist dieser Ausdruck 0 und $\theta - \varrho$ is ein Vielfaches von π. Da beide in $[0, 2\pi)$ liegen, ergibt dies

$$\theta - \varrho = k\pi, \quad k \in \{-1, 0, 1\}.$$

Für den Kosinus ergibt das

$$\cos(\theta - \varrho) = \cos(k\pi) = (-1)^k.$$

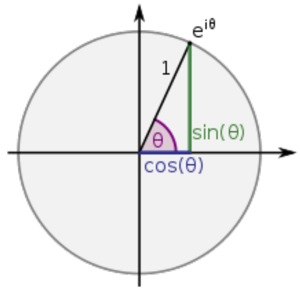
Das Additionstheorem für den Kosinus ergibt aber auch

$$\cos(\theta - \varrho) = \cos(\theta) \cos(\varrho) + \sin(\theta) \sin(\varrho) = \cos^2(\theta) + \sin^2(\theta) = 1.$$

Daher muss $k = 0$ sein und damit ist $\theta = \varrho$. □

Die Darstellung komplexer Zahlen in der Form $z = |z|(\cos(\theta) + i\sin(\theta))$ heißt *Polarkoordinatendarstellung.*

Notation II.2.19. Wir folgen der üblichen Notation und setzen für $\theta \in \mathbb{R}$
$$e^{i\theta} = \cos(\theta) + i\sin(\theta).$$



Quelle: Wiki Commons, Stephan Kulla

Dass dies mehr ist als eine bloße Konvention, lernen Sie in der Analysis oder Sie sind neugierig und gucken schon mal `https://en.wikipedia.org/wiki/Euler%27s_formula#Using_power_series`.

Damit können wir jede komplexe Zahl $z \neq 0$ schreiben als $z = re^{i\theta}$, wobei $|z| = r$ ist.

Beispiel II.2.20. Für $\theta = \pi$ erhalten Sie die berühmte *Eulersche Formel*:
$$e^{i\pi} = \cos(\pi) + i\sin(\pi) = -1 + i \cdot 0 = -1.$$

Satz II.2.21. *Ist $z = re^{i\theta}$ und $w = se^{i\varrho}$, so ist*
$$z \cdot w = rse^{i(\theta+\varrho)}.$$

Bemerkung II.2.22. Die Multiplikation komplexer Zahlen hat also die folgende geometrische Bedeutung:

- Die Längen multiplizieren sich.
- Die Winkel addieren sich.

Beweis. Dies ist eine direkte Rechnung, die wiederum die Additionstheoreme für Sinus und Kosinus benutzt:
$$\begin{aligned} z \cdot w =& |z|(\cos(\theta) + i\sin(\theta))|w|(\cos(\varrho) + i\sin(\varrho)) \\ =& |z||w|(\cos(\theta)\cos(\varrho) - \sin(\theta)\sin(\varrho) + i(\sin(\theta)\cos(\varrho) + \cos(\theta)\sin(\varrho))) \\ =& |z||w|(\cos(\theta+\varrho) + i\sin(\theta+\varrho)). \end{aligned}$$

□

II.3. Vektorräume

Der Begriff des Vektorraums ist grundlegend für die gesamte Lineare Algebra. Er ist eine Abstraktion der Rechenregeln, die wir schon im $\mathbb{R}^2$ beobachtet haben. Vergleichen Sie bitte die folgenden Axiome mit Bemerkung I.8.2.

Definition II.3.1. Es sei K ein Körper. Ein K-*Vektorraum* (oder auch *Vektorraum über* K) besteht aus einer Menge V zusammen mit zwei Abbildungen
$$\begin{aligned} +&: V \times V \to V \\ \cdot&: K \times V \to V, \end{aligned}$$
so dass die folgenden Axiome erfüllt sind:

(V1) $(V, +)$ ist eine abelsche Gruppe.

(V2) Für alle $\lambda, \mu \in K$ und alle $v, w \in V$ gilt:

   (a) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$.

   (b) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

   (c) $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$.

(d) $1 \cdot v = v$.

Bemerkung II.3.2.
- Die Elemente von V werden auch *Vektoren* genannt.
- Das neutrale Element 0_V der abelschen Gruppe $(V, +)$ heißt der *Nullvektor*.
- Die Elemente aus K heißen in diesem Zusammenhang oft *Skalare*.
- Wir werden im Folgenden $\cdot$ häufig weglassen und zum Beispiel λv für $\lambda \cdot v$ schreiben.

Sie kennen schon Beispiele von Vektorräumen:

Beispiele II.3.3.
- Es sei K ein beliebiger Körper und $n \in \mathbb{N}$. Wir definieren auf
$$V = K^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, x_i \in K \right\}$$
die Addition komponentenweise:
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$
und setzen für $\lambda \in K$
$$\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$
Damit ist K^n ein K-Vektorraum.
- Speziell für $n = 1$ erhalten wir, dass jeder Körper K ein K-Vektorraum ist.
- $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum.
- $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum.
- Ist $X \neq \varnothing$ eine Menge und K ein Körper, so ist die Menge aller Abbildungen von X nach K, $V = \operatorname{Abb}(X, K)$, ein K-Vektorraum, wenn wir $\operatorname{Abb}(X, K)$ mit der folgenden Struktur versehen: Es seien $f, g \in \operatorname{Abb}(X, K)$ und $\lambda \in K$. Wir setzen
$$\begin{aligned} (f + g)(x) :=& f(x) + g(x), \\ (\lambda \cdot f)(x) :=& \lambda \cdot f(x). \end{aligned}$$
Der Nullvektor ist die Nullabbildung $X \to K$, die alle $x \in X$ auf 0_K abbildet und $-f$ ist das additive Inverse zu f.
- Ist V ein K-Vektorraum, so ist V insbesondere nicht die leere Menge, weil $(V, +)$ eine abelsche Gruppe ist. Der kleinstmögliche Vektorraum ist der *Nullvektorraum*: Er besitzt ein einziges Element, 0, und hat die Struktur $0 + 0 = 0$ und $\lambda \cdot 0 = 0$ für alle $\lambda \in K$.

Satz II.3.4. *Es sei K ein Körper und V ein K-Vektorraum. Dann gilt für alle $\lambda \in K$ und alle $v, w \in V$:*

(a) $0_K \cdot v = 0_V$,
(b) $\lambda \cdot 0_V = 0_V$.
(c) *Aus $\lambda \cdot v = 0_V$ folgt, dass $\lambda = 0_K$ oder $v = 0_V$.*
(d) $(-1) \cdot v = -v$.

Beweis. Das Argument für (a) kennen Sie bereits: Da $0_K \cdot v = (0_K + 0_K) \cdot v = 0_K \cdot v + 0_K \cdot v$ gilt, können wir $0_K \cdot v$ abziehen und erhalten $0_K \cdot v = 0_V$.

(b) ist ähnlich: $\lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V$ ergibt wiederum $\lambda \cdot 0_V = 0_V$.

(c) Ist $\lambda \cdot v = 0_V$ und $\lambda \neq 0$, so ist
$$v = 1 \cdot v = \lambda^{-1} \lambda v = \lambda^{-1}(\lambda v) = 0_V.$$

(d) Da $(-1)v + v = ((-1) + 1)v = 0_K \cdot v = 0_V$ gilt, ist $(-1)v$ das additive Inverse von v, also $-v$. $\square$

Beachten Sie, dass wir in (c) benutzt haben, dass alle $0_K \neq \lambda \in K$ invertierbar sind.

Oft schreiben wir im Folgenden nur 0 sowohl für 0_K als auch für 0_V.

Wir definieren wieder geeignete Unterobjekte:

Definition II.3.5. Es sei V ein K-Vektorraum. Eine Teilmenge $U \subset V$ heißt ein *Untervektorraum von V*, falls gilt:

(UV1) $U \neq \varnothing$

(UV2) Für alle $u_1, u_2 \in U$ ist $u_1 + u_2 \in U$.

(UV3) Für alle $\lambda \in K$ und $u \in U$ ist $\lambda u \in U$.

Die Axiome (UV2) und (UV3) besagen, dass U unter der Addition und der skalaren Multiplikation abgeschlossen ist.

Die Vektorraumstruktur von V vererbt sich auf Untervektorräume:

Satz II.3.6. *Es sei V ein K-Vektorraum und $U \subset V$ sei ein Untervektorraum. Dann ist U mit der Einschränkung von $+$ und $\cdot$ auf U*

$$+|_{U\times U} \colon U \times U \to U,$$
$$\cdot|_{K\times U} \colon K \times U \to U$$

selbst ein K-Vektorraum mit $0_U = 0_V$.

Beweis. Wir zeigen zunächst, dass $(U, +)$ eine abelsche Gruppe ist: Das Axiom (UV1) impliziert (UG1) und (UG2) folgt aus (UV2). Mit (UV3) folgt, dass für alle $u \in U$ auch $(-1)u = -u$ in U ist, also gilt (UG3), damit ist $(U, +)$ eine Untergruppe von $(V, +)$ und insbesondere abelsch.

Das Axiom (V2) überträgt sich von V auf U. □

Beispiele II.3.7.

(a) In jedem K-Vektorraum V sind $\{0_V\} \subset V$ und $V \subset V$ Untervektorräume.

(b) Wir betrachten im K^n für ein festes k mit $0 \leqslant k \leqslant n$ die Teilmenge

$$U = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}, x_1, \ldots, x_k \in K \right\}.$$

Dies ist immer ein Untervektorraum des K^n.

(c) Ist $K = \mathbb{R}$ und $V = \mathbb{C}$, so ist

$$U_{\mathbb{R}} := \{x + i \cdot 0, x \in \mathbb{R}\}$$

ein Untervektorraum von $\mathbb{C}$. Ebenso ist

$$U_{i\mathbb{R}} := \{0 + i \cdot y, y \in \mathbb{R}\}$$

ein Untervektorraum.

(d) Vorsicht: $\mathbb{C}$ ist natürlich auch ein $\mathbb{C}$-Vektorraum, aber in dieser Situation sind $U_{\mathbb{R}}$ und $U_{i\mathbb{R}}$ keine Untervektorräume, weil die skalare Multiplikation in beiden Fällen hinausführt: $1 + i \cdot 0 \in U_{\mathbb{R}}$ aber $i(1 + i \cdot 0) = i \notin U_{\mathbb{R}}$. Ebenso ist $i(0 + i \cdot 1) = -1 \notin U_{i\mathbb{R}}$.

(e) Ist $X \neq \varnothing$ eine Menge und ist $Y \subset X$ eine Teilmenge, so ist

$$U := \{f \colon X \to K, f(y) = 0_K \text{ falls } y \in Y\}$$

ein Untervektorraum von $\mathrm{Abb}(X, K)$. Man nennt U auch den *Annihilator von Y*.

Für (UV1) bemerken wir, dass die konstante Nullabbildung ein Element von U ist.

Sind $f, g \in U$ und ist $y \in Y$, dann ist

$$(f + g)(y) = f(y) + g(y) = 0_K + 0_K = 0_K$$

und damit ist auch $f + g \in U$. Das zeigt (UV2).

Ist $f \in U$ und $\lambda \in K$, so gilt für alle $y \in Y$:

$$(\lambda f)(u) = \lambda \cdot f(u) = \lambda \cdot 0_K = 0_K$$

und das zeigt (UV3).

(f) In $\mathrm{Abb}(\mathbb{N}_0, K)$ ist

$$U := \{f\colon \mathbb{N}_0 \to K, f(n) \neq 0 \text{ nur für endlich viele } n \in \mathbb{N}_0\}$$

ein Untervektorraum.

Bemerkung II.3.8.

- Sind $W \subset U$ und $U \subset V$ Untervektorräume, so auch $W \subset V$.
- Ist V ein K-Vektorraum und sind $U_1, U_2 \subset V$ Untervektorräume, so ist auch $U_1 \cap U_2$ ein Untervektorraum.

  Vorsicht: $U_1 \cup U_2$ ist im Allgemeinen *kein* Untervektorraum von V. Zum Beispiel können wir in $V = \mathbb{R}^2$ die Untervektorräume

$$U_1 := \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix}, x \in \mathbb{R} \right\}, \quad U_2 := \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix}, y \in \mathbb{R} \right\}$$

  betrachten. Dann sind $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in $U_1 \cup U_2$, aber die Summe ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und dies ist kein Element von $U_1 \cup U_2$.

II.4. Linearkombinationen

Definition II.4.1. Es sei V ein K-Vektorraum.

(a) Sind $v_1, \ldots, v_n$ Elemente von V und sind $\lambda_1, \ldots, \lambda_n \in K$, dann heißt

$$w = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

eine *Linearkombination der Vektoren* $v_1, \ldots, v_n$.

(b) Die Menge

$$\mathsf{Span}_K(v_1, \ldots, v_n) := \{\lambda_1 v_1 + \ldots + \lambda_n v_n, \lambda_i \in K\}$$

heißt der *von* $v_1, \ldots, v_n$ *aufgespannte Raum* oder das *Erzeugnis von* $v_1, \ldots, v_n$.

(c) Ist $\varnothing \neq X \subset V$ eine beliebige Teilmenge von V, so heißt

$$\mathsf{Span}_K(X) := \{\lambda_1 v_1 + \ldots + \lambda_m v_m, \lambda_i \in K, v_i \in X, m \in \mathbb{N}_0\}$$

der *von* X *aufgespannte Raum* oder das *Erzeugnis von* X.

Beachten Sie bitte, dass alle auftretenden Summen endlich sind!

Satz II.4.2. *Es sei V ein K-Vektorraum, $v_1, \ldots, v_n \in V$ und $\varnothing \neq X \subset V$. Dann gilt:*

(a) $\mathsf{Span}_K(v_1, \ldots, v_n)$ *ist ein Untervektorraum von V.*

(b) $\mathsf{Span}_K(X)$ *ist ein Untervektorraum von V.*

(c) *Ist $U \subset V$ ein Untervektorraum und ist $\varnothing \neq X \subset U$, so ist auch* $\mathsf{Span}_K(X) \subset U$.

Bemerkung II.4.3. (c) besagt, dass $\mathsf{Span}_K(X)$ der (bezüglich Inklusion) kleinste Untervektorraum von V ist, der X enthält.

Beweis. Zu (a): Da $0_V = 0_K v_1 + \ldots + 0_K v_n \in \mathsf{Span}_K(v_1, \ldots, v_n)$, gilt (UV1).

Sind $v, w \in \mathsf{Span}_K(v_1, \ldots, v_n)$, so gibt es $\lambda_1, \ldots, \lambda_n \in K$ und $\mu_1, \ldots, \mu_n \in K$ mit

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n, w = \mu_1 v_1 + \ldots + \mu_n v_n.$$

(Notfalls ergänzen wir mit Skalaren 0_K, um auf die gleiche Länge der Linearkombination zu kommen.) Damit ist dann auch

$$v + w = (\lambda_1 + \mu_1)v_1 + \ldots + (\lambda_n + \mu_n)v_n \in \mathsf{Span}_K(v_1, \ldots, v_n).$$

Mit v wie oben und $\mu \in K$ ist

$$\mu v = \mu(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \mu\lambda_1 v_1 + \ldots + \mu\lambda_n v_n \in \mathsf{Span}_K(v_1, \ldots, v_n).$$

Damit gelten auch (UV2) und (UV3) für $\mathsf{Span}_K(v_1, \ldots, v_n)$.

Der Beweis von (b) ist völlig analog.

Zu (c): Ist $\varnothing \neq X \subset U \subset V$, so liegen wegen (UV2) und (UV3) auch alle Linearkombinationen von Elementen von X in U und somit $\mathsf{Span}_K(X) \subset U$. $\square$

Beispiele II.4.4.

- Ist K ein Körper, so seien e_1 bis e_n die Vektoren im K^n der Form
$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$
Dann ist für alle r mit $1 \leqslant r \leqslant n$
$$\mathsf{Span}_K(e_1, \ldots, e_r) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}, x_i \in K \right\}$$
und insbesondere gilt $\mathsf{Span}_K(e_1, \ldots, e_n) = K^n$.
- Für $n = 1$ erhalten wir, dass für alle $x \in K \setminus \{0\}$ gilt:
$$\mathsf{Span}_K(x) = K$$
aber
$$\mathsf{Span}_K(0_K) = 0_K.$$
- Was ist $\mathsf{Span}_{\mathbb{Q}}(\pi)$ für $V = \mathbb{R}$ als $\mathbb{Q}$-Vektorraum?

Wir hatten bisher immer vorausgesetzt, dass $\varnothing \neq X$. Die sinnvolle Konvention ist,
$$\mathsf{Span}_K(\varnothing) = \{0\}$$
zu setzen, weil der Nullvektorraum der kleinstmögliche Vektorraum ist.

Definition II.4.5. Es seien X und Y Mengen. Eine Abbildung
$$\Phi \colon X \to Y$$
heißt eine *durch X indizierte Familie von Elementen von Y*.

Die Menge X heißt dann *Indexmenge*. Statt $\Phi(x)$ schreibt man häufig auch y_x und statt der Familie oft kurz $(y_x)_{x \in X}$.

Ist X eine endliche Menge, so heißt die Familie *endlich*.

Notation II.4.6. Ist $X = \mathbf{n} = \{1, \ldots, n\}$, so ist jede Familie $\Phi \colon \mathbf{n} \to Y$ durch die Ordnung auf $\mathbf{n}$ geordnet. Für Φ schreiben wir dann auch $(y_1, \ldots, y_n)$.

Beispiele II.4.7.

- Eine Familie $\Phi \colon \mathbf{n} \to \mathbb{R}$ entspricht einem Vektor $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.
- Eine Familie $\Phi \colon \mathbf{m} \to \mathbb{R}^n$ entspricht m Vektoren im $\mathbb{R}^n$: $(v_1, \ldots, v_m)$, $v_i \in \mathbb{R}^n$. Wir hatten vorher schon die Familie $\Phi \colon \mathbf{n} \to \mathbb{R}^n$ betrachtet, die $i \in \mathbf{n}$ auf den Vektor e_i abbildet.

Definition II.4.8. Es sei V ein K-Vektorraum.

(a) Eine endliche Familie $(v_1, \ldots, v_r)$ von Vektoren $v_i \in V$ heißt *linear unabhängig*, falls gilt:
Sind $\lambda_1, \ldots, \lambda_r \in K$ und ist $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$, so folgt $\lambda_1 = \ldots = \lambda_r = 0$.
Andernfalls heißt die Familie *linear abhängig*.

(b) Eine beliebige Teilmenge $\varnothing \neq Y \subset V$ heißt *linear unabhängig*, falls für jede endliche Teilmenge $\{v_1, \ldots, v_m\} \subset Y$ die Vektoren $(v_1, \ldots, v_m)$ linear unabhängig sind. Andernfalls heißt Y *linear abhängig*.

(c) Die Konvention ist, dass die leere Menge linear unabhängig ist.

Beispiele II.4.9.

- Im K^3 ist die Familie
$$\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right)$$
linear abhängig, weil
$$1 \cdot \begin{pmatrix}1\\0\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\\0\end{pmatrix} + (-1) \cdot \begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$
gilt.
- Im K^n ist jede Teilmenge von $\{e_1, \ldots, e_n\}$ linear unabhängig.
- Enthält eine Familie von Vektoren eines Vektorraums eine linear abhängige Unterfamilie, so ist sie selbst linear abhängig.
- Jede Teilmenge $Z \subset Y$ einer linear unabhängigen Teilmenge $Y \subset V$ eines Vektorraums V ist selbst linear unabhängig.
- Ist $(v_x)_{x \in X}$ eine Familie mit $v_x \in V$ für alle $x \in X$ und gibt es $x, y \in X$ mit $v_x = v_y$, so ist $(v_x)_{x \in X}$ linear abhängig, weil
$$1 \cdot v_x + (-1) \cdot v_y = 0_V$$
gilt.

Satz II.4.10. *Es sei V ein K-Vektorraum.*

(a) *Für eine Familie $(v_1, \ldots, v_n)$ von Vektoren in V sind äquivalent:*

(i) $(v_1, \ldots, v_n)$ *ist linear unabhängig.*

(ii) *Jeder Vektor $v \in \mathsf{Span}_K(v_1, \ldots, v_n)$ läßt sich in eindeutiger Weise als Linearkombination der Vektoren $v_1, \ldots, v_n$ schreiben.*

(b) *Enthält eine Familie von Vektoren mehr als einen Vektor, so ist sie genau dann linear abhängig, wenn ein Vektor der Familie eine Linearkombination anderer Vektoren der Familie ist.*

Im Beweis benutzen wir das Summenzeichen: $\sum_{i=1}^n \lambda_i v_i$ steht für $\lambda_1 v_1 + \ldots + \lambda_n v_n$.

Beweis. Zu (a): Wir zeigen (ii) $\Rightarrow$ (i), indem wir $\neg$ (i) $\Rightarrow$ $\neg$ (ii) beweisen:
Ist die Familie linear abhängig, so kann man den Vektor 0_V auf mindestens zwei Arten als Linearkombination schreiben.
(i) $\Rightarrow$ (ii): Ist
$$v = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i \text{ mit } \lambda_i, \mu_i \in K$$
so ist
$$0_V = v - v = \sum_{i=1}^n (\lambda_i - \mu_i) v_i.$$
Nach Annahme von (i) impliziert dies, dass für alle i gilt, dass $\lambda_i - \mu_i = 0_K$ ist, also dass $\lambda_i = \mu_i$ gilt.

Zu (b): Es sei $(v_1, \dots, v_n)$ eine linear abhängige Familie. Dann gibt es $\lambda_1, \dots, \lambda_n \in K$ mit $\sum_{i=1}^n \lambda_i v_i = 0_V$, aber nicht alle λ_i sind null. Es sei $\lambda_i \neq 0$. Dann ist

$$v_i = -\sum_{\substack{j=1\\ j\neq i}}^{n} \lambda_i^{-1}\lambda_j v_j.$$

Ist umgekehrt $v_i = \mu_1 v_1 + \ldots + \mu_{i-1}v_{i-1} + \mu_{i+1}v_{i+1} + \ldots + \mu_n v_n$, so ist

$$0_V = \mu_1 v_1 + \ldots + \mu_{i-1}v_{i-1} + (-1)v_i + \mu_{i+1}v_{i+1} + \ldots + \mu_n v_n$$

aber $-1 \neq 0_K$. □

II.5. Basis und Dimension

Definition II.5.1. Es sei K ein Körper und V sei ein K-Vektorraum.

(a) Eine Teilmenge $Y \subset V$ heißt *Erzeugendensystem von V*, falls $V = \mathsf{Span}_K(Y)$.

(b) Eine Teilmenge $Y \subset V$ heißt eine *Basis von V*, falls Y ein linear unabhängiges Erzeugendensystem von V ist.

A priori ist nicht klar, dass jeder Vektorraum eine Basis besitzt. Wir kommen auf diese Frage zurück.

Beispiele II.5.2.

- Ist $V = K^n$, so ist $\{e_1, \dots, e_n\}$ eine Basis von V. Sie heißt *Standardbasis des K^n*. Oft brauchen wir geordnete Basen und dann betrachten wir die Familie $(e_1, \dots, e_n)$ als geordnete Basis des K^n.
- Ist $V = \mathbb{C}$ als $\mathbb{R}$-Vektorraum, so ist $\{1, i\}$ eine Basis.

Satz II.5.3. *Es sei K ein Körper und $V \neq \{0\}$ sei ein K-Vektorraum. Dann sind äquivalent:*

(a) *Y ist eine Basis von V.*

(b) *Y ist ein minimales Erzeugendensystem von V. (Das heißt: Y ist ein Erzeugendensystem von V und für alle $y \in Y$ ist $Y \setminus \{y\}$ kein Erzeugendensystem von V.)*

(c) *Y ist eine maximale linear unabhängige Teilmenge von V. (Das heißt: Y ist linear unabhängig und für alle $v \in V \setminus Y$ ist $Y \cup \{v\}$ linear abhängig.)*

Beweis. (a) ⇒ (b): Wir nehmen an, dass es ein $y \in Y$ gibt, so dass auch $Y \setminus \{y\}$ ein Erzeugendensystem ist. Dann können wir y darstellen als Linearkombination

$$y = \sum_{i=1}^{n} \lambda_i y_i$$

für geeignete $\lambda_i \in K$, $y_i \in Y \setminus \{y\}$. Damit ist aber

$$0 = (-1)y + \sum_{i=1}^{n} \lambda_i y_i$$

und Y wäre linear abhängig.
(b) ⇒ (a): Es sei Y ein minimales Erzeugendensystem. Wir nehmen an, dass die Familie $Y = (y_i)_{i\in I}$ linear abhängig ist, also gibt es $y_1, \dots, y_{n+1} \in Y$ und $\lambda_1, \dots, \lambda_{n+1} \in K$ mit

$$\sum_{i=1}^{n+1} \lambda_i y_i = 0$$

so dass nicht alle $\lambda_i = 0$ sind. Ohne Einschränkung nehmen wir an, dass $\lambda_{n+1} \neq 0$. Dann gilt:

$$y_{n+1} = \sum_{i=1}^{n} -\lambda_{n+1}^{-1}\lambda_i y_i$$

und damit wäre auch $Y \setminus \{y_{n+1}\}$ ein Erzeugendensystem im Widerspruch zur Minimalität von Y.
(a) ⇒ (c): Es sei $v \in V \setminus Y$. Da Y Erzeugendensystem ist, gilt

$$v = \sum_{i\in I} \lambda_i y_i,$$

mit $y_i \in Y$ und wobei nur endlich viele $\lambda_i \neq 0$ sind. Damit ist $Y \cup \{v\}$ linear abhängig.
(c) $\Rightarrow$ (a): Es sei Y eine maximale linear unabhängige Teilmenge von V. Wir zeigen, dass Y dann V erzeugt. Es sei $v \in V$ beliebig.

Ist $v \in Y$, so ist $v \in \mathsf{Span}_K(Y)$.

Ist $v \in V \setminus Y$, so ist nach Voraussetzung $Y \cup \{v\}$ linear abhängig und $\exists \mu, \lambda_1, \ldots, \lambda_n$:

$$\mu v + \sum_{i=1}^{n} \lambda_i y_i = 0,$$

so dass nicht alle $\mu, \lambda_1, \ldots, \lambda_n$ gleich 0 sind und mit $y_1, \ldots, y_n \in Y$. Wäre nun $\mu = 0$, so ist also ein $\lambda_i \neq 0$, aber es gilt trotzdem

$$\sum_{i=1}^{n} \lambda_i y_i = 0.$$

Das ist ein Widerspruch zur linearen Unabhängigkeit von Y. Also ist $\mu \neq 0$, aber damit ist

$$v = -\mu^{-1} \sum_{i=1}^{n} \lambda_i y_i.$$

Da v beliebig war, zeigt dies, dass Y ein Erzeugendensystem ist. □

Definition II.5.4. Ein K-Vektorraum V heißt *endlich erzeugt*, falls es eine endliche Teilmenge $Y \subset V$ gibt, so dass $V = \mathsf{Span}_K(Y)$.

Lemma II.5.5. *Es sei V ein K-Vektorraum, der nicht endlich erzeugt ist. Dann gibt es für jedes $n \in \mathbb{N}$ Vektoren $v_1, \ldots, v_n \in V$, die linear unabhängig sind.*

Beweis. Wir machen vollständige Induktion über $n \in \mathbb{N}$.

(IA) Ist V der Nullvektorraum, so ist V endlich erzeugt. Sei also ohne Beschränkung der Allgemeinheit $V \neq \{0\}$. Dann gibt es ein $v_1 \in V$ mit $v_1 \neq 0_V$. Damit ist $\{v_1\}$ linear unabhängig.

(IS) Es sei $n \in \mathbb{N}$ beliebig und es seien $v_1, \ldots, v_n$ linear unabhängig. Es ist $\mathsf{Span}_K(v_1, \ldots, v_n)$ eine echte Teilmenge von V, weil ansonsten V endlich erzeugt wäre. Es gibt also ein $v_{n+1} \in V \setminus \mathsf{Span}_K(v_1, \ldots, v_n)$. Wir behaupten, dass $(v_1, \ldots, v_n, v_{n+1})$ linear unabhängig ist.

Es sei $0 = \sum_{i=1}^{n+1} \lambda_i v_i$ mit $\lambda_i \in K$. Ist $\lambda_{n+1} \neq 0$, so können wir v_{n+1} ausdrücken als

$$v_{n+1} = -\lambda_{n+1}^{-1} \sum_{i=1}^{n} \lambda_i v_i$$

aber damit ist v_{n+1} in $\mathsf{Span}_K(v_1, \ldots, v_n)$. Widerspruch.

Also muss $\lambda_{n+1} = 0$ gelten und wir haben $0 = \sum_{i=1}^{n} \lambda_i v_i$. Da aber $(v_1, \ldots, v_n)$ linear unabhängig ist, müssen alle $\lambda_i = 0$ sein. Also gilt insgesamt

$$0 = \lambda_1 = \ldots = \lambda_n = \lambda_{n+1}.$$

□

Satz II.5.6 (Basisauswahlsatz)**.** *Ist V ein K-Vektorraum und ist $Y \subset V$ ein endliches Erzeugendensystem, so gibt es eine Teilmenge $B \subset Y$, die eine Basis von V ist. Insbesondere hat jeder endlich erzeugte Vektorraum eine Basis.*

Beweis. Ist Y selbst schon linear unabhängig, so ist Y selbst eine Basis. Wir können also ohne Einschränkung annehmen, dass $|Y| > 1$ gilt.

Ist Y linear abhängig, so gibt es nach Satz II.4.10 ein $v \in Y$ mit der Eigenschaft, dass $v \in \mathsf{Span}_K(Y \setminus \{v\})$. Damit ist $Y \setminus \{v\}$ immer noch ein Erzeugendensystem.

Iterieren Sie diesen Prozess. Nach endlich vielen Schritten erhalten Sie ein minimales Erzeugendensystem und dieses ist nach Satz II.5.3 eine Basis. □

Bemerkung II.5.7. Natürlich ist *nicht* jeder Vektorraum endlich erzeugt. Beispiele sind unter anderem $\mathrm{Abb}(\mathbb{N}_0, K)$ oder

$$K[X] := \{f\colon \mathbb{N}_0 \to K, f(n) = 0 \text{ für fast alle } n \in \mathbb{N}_0\}.$$

Wir sehen später mit Zorns Lemma, dass jeder Vektorraum eine Basis besitzt (Max August Zorn (1906–1993)).

Lemma II.5.8 (Austauschlemma)**.** *Ist V ein K-Vektorraum und ist $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V, so ist für $w = \sum_{i=1}^n \lambda_i v_i$ mit $\lambda_i \in K$ die Menge $\mathcal{B}'_k = \{v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n\}$ ebenfalls eine Basis von V, falls $\lambda_k \neq 0$.*

Beispiel II.5.9. Wir betrachten $e_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$ und $e_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ im K^3 und

$$w = \begin{pmatrix}1\\1\\0\end{pmatrix} = e_1 + e_2.$$

Dann sind sowohl $\{e_1, w, e_3\}$ als auch $\{w, e_2, e_3\}$ Basen des K^3.

Beweis. Wir können die $v_1, \ldots, v_n$ gegebenenfalls umnummerieren und können daher annehmen, dass $k = 1$ ist in

$$w = \sum_{i=1}^n \lambda_i v_i. \tag{II.5.1}$$

Wir zeigen, dass $\mathcal{B}' := \mathcal{B}'_1 = \{w, v_2, \ldots, v_n\}$ eine Basis von V ist.

Die Menge $\mathcal{B}'$ ist ein Erzeugendensystem von V: Ist $v \in V$ beliebig, so gibt es $\mu_1, \ldots, \mu_n \in K$ mit

$$v = \sum_{i=1}^n \mu_i v_i,$$

weil $\mathcal{B}$ ein Erzeugendensystem ist. Da $\lambda_1 \neq 0$ gilt in (II.5.1), folgt, dass

$$v_1 = \lambda_1^{-1} w - \left(\sum_{i=2}^n \lambda_1^{-1} \lambda_i v_i\right).$$

Einsetzen ergibt damit für v:

$$\begin{aligned} v =& \mu_1 \left(\lambda_1^{-1} w - \sum_{i=2}^n \lambda_1^{-1} \lambda_i v_i\right) + \sum_{i=2}^n \mu_i v_i \\ =& \mu_1 \lambda_1^{-1} w + \sum_{i=2}^n (\mu_i - \mu_1 \lambda_1^{-1} \lambda_i) v_i \end{aligned}$$

und damit ist $v \in \mathsf{Span}_K(w, v_2, \ldots, v_n)$.

Die Menge $\mathcal{B}'$ ist linear unabhängig: Wir nehmen an, dass

$$\beta_1 w + \sum_{i=1}^2 \beta_i v_i = 0.$$

Einsetzen von (II.5.1) ergibt

$$\sum_{i=1}^n \beta_1 \lambda_i v_i + \sum_{i=1}^2 \beta_i v_i = 0$$

also

$$\beta_1 \lambda_1 v_1 + \sum_{i=2}^n (\beta_1 \lambda_i + \beta_i) v_i = 0.$$

Da $\{v_1, \dots, v_n\}$ linear unabhängig ist, bekommen wir für die Koeffizienten die Gleichungen

$$\beta_1\lambda_1 = 0 \text{ und } \beta_1\lambda_i + \beta_i = 0 \text{ für } 2 \leqslant i \leqslant n.$$

Nach Annahme ist $\lambda_1 \neq 0$, also muss $\beta_1 = 0$ gelten. Damit folgt aber, dass für $2 \leqslant i \leqslant n$ alle $\beta_i = 0$ sind. $\square$

Satz II.5.10 (Austauschsatz)**.** *Ist V ein K-Vektorraum, ist $\mathcal{B} = \{v_1, \dots, v_n\}$ eine Basis von V und ist $\{w_1, \dots, w_r\}$ eine linear unabhängige Teilmenge von V. Dann gilt:*

(a) $r \leqslant n$

(b) $\exists i_1, \dots, i_r \in \{1, \dots, n\}$*, so dass der Austausch von v_{i_1} gegen w_1, …, v_{i_r} gegen w_r eine neue Basis von V ergibt.*

Nach einer Umnummerierung der v_i können wir also $(w_1, \dots, w_r, v_{r+1}, \dots, v_n)$ als neue Basis betrachten.

Beweis. Wir zeigen die Behauptungen mit Induktion über $r \in \mathbb{N}_0$. Für $r = 0$ ist nichts zu zeigen. Es sei die Aussage für alle $k \leqslant r-1$ gültig. Zu zeigen ist, dass sie dann auch für $k = r$ gilt.

Es sei also $\{w_1, \dots, w_r\}$ linear unabhängig. Damit ist auch die Teilmenge $\{w_1, \dots, w_{r-1}\}$ linear unabhängig. Wir erhalten also, dass

- $r - 1 \leqslant n$, und dass
- $\overline{\mathcal{B}} := \{w_1, \dots, w_{r-1}, v_r, \dots, v_n\}$ eine Basis von V ist.

Wir behaupten, dass dann auch $r \leqslant n$ gilt. Wir nehmen an, dass $r > n$ ist. Damit muss $r - 1 = n$ sein und $\overline{\mathcal{B}} = \{w_1, \dots, w_n\}$ ist eine maximale linear unabhängige Teilmenge. Aber nach Annahme ist

$$\{w_1, \dots, w_{n+1}\} = \{w_1, \dots, w_r\}$$

ebenfalls linear unabhängig und dies ist ein Widerspruch zur Maximalität. Also ist $r \leqslant n$.

Wir zeigen nun, dass es ein $i_r \in \{r, \dots, n\}$ gibt, so dass wir v_{i_r} gegen w_r austauschen können. Da $\overline{\mathcal{B}}$ eine Basis ist, finden wir $\lambda_1, \dots, \lambda_{r-1}$ und $\mu_r, \dots, \mu_n$ mit

$$w_r = \sum_{i=1}^{r-1} \lambda_i w_i + \sum_{j=r}^{n} \mu_j v_j.$$

Wären alle $\mu_j = 0$, so wäre w_r eine Linearkombination der $w_1, \dots, w_{r-1}$. Das ist ein Widerspruch zur linearen Unabhängigkeit der $\{w_1, \dots, w_r\}$. Also muss es ein $i_r \in \{r, \dots, n\}$ geben mit $\mu_{i_r} \neq 0$. Mit dem Austauschlemma ist dann $\{w_1, \dots, w_r, v_r, \dots, v_n\} \setminus \{v_{i_r}\}$ eine Basis. $\square$

Definition II.5.11. Ist V ein K-Vektorraum und ist $\mathcal{B} = \{v_1, \dots, v_n\}$ eine Basis von V, so heißt n die *Länge der Basis.*

Korollar II.5.12.

(a) *Hat ein K-Vektorraum V eine endliche Basis, so ist jede Basis von V endlich.*

(b) *Je zwei Basen eines endlich erzeugten K-Vektorraums haben die gleiche Länge.*

Beweis. Zu (a): Ist $\mathcal{B} = \{v_1, \dots, v_n\}$ eine Basis von V und ist $\mathcal{B}'$ eine weitere Basis von V, die nicht endliche Länge hat, so gibt es eine linear unabhängige Teilmenge $\{w_1, \dots, w_{n+1}\} \subset \mathcal{B}'$. Dies ist ein Widerspruch zum Austauschsatz.

Zu (b): Sind $\mathcal{B} = \{v_1, \dots, v_n\}$ und $\mathcal{B}' = \{w_1, \dots, w_m\}$ zwei Basen von V, so gilt mit dem Austauschsatz, weil $\mathcal{B}'$ linear unabhängig ist und $\mathcal{B}$ eine Basis ist, dass $m \leqslant n$ ist. Vertauschen wir die Rollen von $\mathcal{B}'$ und $\mathcal{B}$ und setzen $\mathcal{B}$ als die linear unabhängige Teilmenge an und $\mathcal{B}'$ als Basis, erhalten wir $n \leqslant m$, so dass insgesamt Gleichheit gilt. $\square$

Damit kommen wir zu einer der zentralen Begriffe der linearen Algebra:

Definition II.5.13. Es sei V ein K-Vektorraum. Wir definieren die *Dimension von V über K* als

$$\dim_K(V) := \begin{cases} 0, & \text{falls } V = 0, \\ n, & \text{falls } V \text{ eine Basis der Länge } n \text{ hat,} \\ \infty, & \text{falls keine Basis endlicher Länge für } V \text{ existiert.} \end{cases}$$

Bemerkung II.5.14. Wir betrachten die leere Menge als Basis des Nullvektorraums. Damit ist die obige Definition konsistent.

Beispiele II.5.15.
- Für alle $n \in \mathbb{N}$ ist $\dim_K K^n = n$, weil $\{e_1, \ldots, e_n\}$ eine Basis der Länge n ist.
- $\dim_{\mathbb{R}} \mathbb{C} = 2$, weil $\{1, i\}$ eine Basis der Länge zwei ist.
- $\dim_K K[X] = \infty$.

Satz II.5.16. *Ist V ein endlich erzeugter K-Vektorraum und ist $U \subset V$ ein Untervektorraum, so ist U ebenfalls endlich erzeugt und $\dim_K U \leqslant \dim_K V$. Gilt $\dim_K U = \dim_K V$, so ist $U = V$.*

Beweis. Ist $\dim_K V = n$ und wäre U nicht endlich erzeugt, so gibt es $n+1$ linear unabhängige Vektoren $v_1, \ldots, v_{n+1} \in U \subset V$ und das steht im Widerspruch zum Austauschsatz. Damit hat U eine endliche Basis $\{w_1, \ldots, w_r\}$ und mit dem Austauschsatz folgt $r \leqslant n$.

Nehmen wir an, dass $\dim_K U = \dim_K V < \infty$ und dass $\mathcal{B} = \{w_1, \ldots, w_n\}$ eine Basis von U ist. Gibt es dann ein $v \in V \setminus U = V \setminus \mathsf{Span}_K(w_1, \ldots, w_n)$, so ist $\mathcal{B}' = \{w_1, \ldots, w_n, v\}$ linear unabhängig, aber dann wäre $\dim_K V \geqslant n + 1 > \dim_K U$. $\square$

Korollar II.5.17. *Ist V ein K-Vektorraum mit $\dim_K V = n < \infty$, so bilden je n linear unabhängige Vektoren eine Basis.*

Beweis. Ist $X = \{v_1, \ldots, v_n\}$ linear unabhängig, so ist X Basis von $U := \mathsf{Span}_K(v_1, \ldots, v_n)$. Damit ist aber $\dim_K U = \dim_K V$ und natürlich ist $U \subset V$ ein Untervektorraum. $\square$

Korollar II.5.18. *(Basisergänzungssatz) Es sei V ein K-Vektorraum mit $\dim_K V = n < \infty$. Ist $\{w_1, \ldots, w_m\}$ eine linear unabhängige Menge in V, so gibt es Vektoren $w_{m+1}, \ldots, w_n$, so dass $\{w_1, \ldots, w_n\}$ eine Basis von V ist.*

Beweis. Wir wissen, dass es eine Basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ der Länge n von V gibt. Wir können m Vektoren in $\mathcal{B}$ gegen $w_1, \ldots, w_m$ austauschen und erhalten die gewünschte Basis. $\square$

II.6. Exkurs: Existenz von Basen im allgemeinen Fall

Definition II.6.1. Es sei X eine Menge

(a) Eine Relation R auf X heißt *antisymmetrisch*, falls aus $x \sim y$ und $y \sim x$ für $x, y \in X$ folgt, dass $x = y$ ist.
(b) Eine Relation heißt *partielle Ordnungsrelation*, falls sie reflexiv, transitiv und antisymmetrisch ist.
(c) Eine Menge X mit einer partiellen Ordnungsrelation heißt *partiell geordnete Menge* (im Englischen: *poset* für *partially ordered set*). Oft schreiben wir dann kurz $(X, \leqslant)$

Beispiele II.6.2.
- Natürlich sind $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$ mit der üblichen Relation $x \sim y \Leftrightarrow x \leqslant y$ partiell geordnete Mengen.
- Ist X eine beliebige Menge, so ist die Potenzmenge $\mathcal{P}(X)$ eine partiell geordnete Menge, indem wir setzen:
$$Y \leqslant Z \Leftrightarrow Y \subset Z.$$
- $X = \{2, 3, 4, \ldots\}$ ist eine partiell geordnete Menge mit der Relation
$$x \sim y \Leftrightarrow x \text{ teilt } y.$$

Definition II.6.3. Ist $(X, \leqslant)$ eine partiell geordnete Menge, so heißt ein Element $m \in X$ *maximal*, falls für alle $x \in X$ gilt: Ist $m \leqslant x$, so ist $x = m$. Ein $m \in X$ heißt *minimal*, falls für alle $x \in X$ gilt: $x \leqslant m \Rightarrow x = m$.

Beispiele II.6.4.
- $(\mathbb{R}, \leqslant)$ hat weder ein maximales noch ein minimales Element. Was ist mit $\mathbb{N}_0$?
- $\varnothing$ ist minimal in $\mathcal{P}(X)$ und X ist maximal in $\mathcal{P}(X)$.
- In $X = \{2, 3, 4, \ldots\}$ mit der partiellen Ordnung gegeben durch die Teilbarkeit sind alle Primzahlen minimal. Es gibt also kein eindeutiges minimales Element.

Definition II.6.5. Es sei $(X, \leqslant)$ eine partiell geordnete Menge.

(a) $(X, \leqslant)$ heißt *total oder linear geordnet*, falls für alle $x, y \in X$ gilt, dass $x \leqslant y$ oder dass $y \leqslant x$ gilt.
(b) Ist $(X, \leqslant)$ partiell geordnet und ist $S \subset X$, so heißt S eine *Kette*, falls S mit der Einschränkung von $\leqslant$ total geordnet ist.
(c) Ist $S \subset X$, so heißt ein $x \in X$ eine *obere Schranke für S*, falls für alle $s \in S$ gilt: $s \leqslant x$.

Beispiele II.6.6.

- $(\mathbb{R}, \leqslant)$ ist linear geordnet.
- $(\mathcal{P}(X), \subset)$ ist genau dann linear geordnet, wenn $|X| \leqslant 1$ ist.
- In $X = \{2, 3, 4, \dots\}$ mit der Teilbarkeitsrelation ist nicht linear geordnet. Zum Beispiel teilt 2 weder 3 noch teilt 3 die 2. Gibt es eine unendliche Kette $S \subset X$, so dass $2 \in S$?

Das folgende Lemma wurde unabhängig von Max Zorn im Jahr 1935 und von Kazimierz Kuratowski im Jahr 1922 bewiesen (1896–1980). Es heißt trotzdem meistens nur *Zorns Lemma*

Lemma II.6.7 (Zorns Lemma)**.** *Ist $(X, \leqslant)$ eine partiell geordnete Menge, $X \neq \varnothing$, so dass jede Kette in X eine obere Schranke hat, dann hat X ein maximales Element.*

Dieses Lemma ist äquivalent zum *Auswahlaxiom*, das wir schon einmal benutzt haben, als wir gezeigt haben, dass surjektive Abbildungen ein Rechtsinverses haben. Es besagt, dass es für jede Familie $\mathcal{F}$ nichtleerer Mengen eine Menge S gibt, so dass für alle $F \in \mathcal{F}$ gilt, $|F \cap S| = 1$.

Wir benutzen Zorns Lemma nun für den Existenzsatz für Basen:

Satz II.6.8. *Es sei K ein Körper. Jeder K-Vektorraum besitzt eine Basis.*

Beweis. Wir zeigen sogar mehr: Ist V ein beliebiger K-Vektorraum und ist $X \subset V$ eine beliebige linear unabhängige Teilmenge, so gibt es eine Basis $\mathcal{B}$ von V mit $X \subset \mathcal{B}$.

Wir definieren
$$\mathscr{F} := \{Y \subset V, X \subset Y \text{ und } Y \text{ ist linear unabhängig}\}.$$
Wir prüfen nach, dass $\mathscr{F}$ die Voraussetzung von Zorns Lemma erfüllt:

- $\mathscr{F} \neq \varnothing$, denn $X \in \mathscr{F}$.
- $\mathscr{F}$ ist partiell geordnete Menge mit $\subset$.
- Es sei $S \subset \mathscr{F}$ eine Kette bezüglich $\subset$. Wir betrachten die Vereinigung
$$Z := \bigcup_{Y \in S} Y.$$
Wir behaupten, dass Z eine linear unabhängige Teilmenge von V ist. Wir nehmen dazu beliebige Elemente $v_1, \dots, v_n \in Z$. Für jedes $i \in \{1, \dots, n\}$ sei Y_i so gewählt, dass $v_i \in Y_i$. Alle $Y_1, \dots, Y_n$ sind aus S und S ist total geordnet. Damit folgt, dass es ein i_0 gibt, so dass Y_{i_0} das größte Element in $\{Y_1, \dots, Y_n\}$ bezüglich $\subset$ ist. Damit gilt für alle i, dass $v_i \in Y_i \subset Y_{i_0}$, also sind alle $v_1, \dots, v_n \in Y_{i_0}$. Aber Y_{i_0} selbst ist eine linear unabhängige Menge. Damit ist $\{v_1, \dots, v_n\}$ linear unabhängig.
- Damit haben wir gezeigt, dass $Z = \bigcup_{Y \in S} Y \subset V$, $X \subset Z$ und Z ist linear unabhängig. Nach Konstruktion ist $Z \in \mathscr{F}$.
- Ebenfalls nach Konstruktion ist Z eine obere Schranke von S.

Zorns Lemma impliziert, dass $\mathscr{F}$ ein maximales Element hat.

Wir behaupten, dass dieses maximale Element, welches wir schon einmal $\mathcal{B}$ nennen, eine Basis von V ist. Da $\mathcal{B} \in \mathscr{F}$ ist, ist $\mathcal{B}$ linear unabhängig.

Wir nehmen an, dass es ein $v \in V \setminus \mathsf{Span}_K(\mathcal{B})$ gibt. Dann ist die Menge $\mathcal{B} \cup \{v\}$ eine linear unabhängige Teilmenge von V, die nach Konstruktion X enthält. Aber damit ist $\mathcal{B} \cup \{v\}$ in $\mathscr{F}$ und $\mathcal{B} \subsetneq \mathcal{B} \cup \{v\}$. Dies ist ein Widerspruch zur Maximalität von $\mathcal{B}$. Daher kann es kein solches v geben und $\mathcal{B}$ ist ein Erzeugendensystem von V. $\square$

Zorns Lemma gibt Ihnen keinen Algorithmus, um eine Basis von V zu konstruieren. Wenn Sie also eine Basis von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum suchen, dann wissen Sie jetzt zwar, dass diese existiert (wenn Sie an das Auswahlaxiom glauben), aber hinschreiben können Sie diese noch lange nicht.

II.7. Summen von Untervektorräumen

Wir hatten gesehen, dass für Untervektorräume $U_1, \dots, U_n$ eines Vektorraums V die Vereinigung $U_1 \cup \dots \cup U_n$ im Allgemeinen *kein* Untervektorraum von V ist. Die Summe ist ein geeigneter Ersatz:

Definition II.7.1. Ist V ein K-Vektorraum und sind $U_1, \dots, U_n$ Untervektorräume von V, so heißt

$$U_1 + \dots + U_n := \{v \in V, \exists u_1, \dots, u_n, u_i \in U_i \text{ mit } v = u_1 + \dots + u_n\}.$$

die *Summe der Untervektorräume* $U_1, \dots, U_n$.

Satz II.7.2. *Ist V ein K-Vektorraum und sind $U_1, \dots, U_n$ Untervektorräume von V, so gilt:*

(a) $U_1 + \dots + U_n = \mathsf{Span}_K(U_1 \cup \dots \cup U_n)$*, insbesondere ist* $U_1 + \dots + U_n$ *ein Untervektorraum von* V*.*

(b)
$$\dim_K(U_1 + \dots + U_n) \leqslant \sum_{i=1}^{n} \dim_K U_i.$$

Für (b) benutzen wir die Konventionen

$$\begin{aligned} \infty + \infty &= \infty, \\ \infty + n &= \infty = n + \infty \text{ für alle } n \in \mathbb{N}_0, \\ n &< \infty \text{ für alle } n \in \mathbb{N}_0. \end{aligned}$$

Beweis. Zu (a): Nach Konstruktion gilt $U_1 + \dots + U_n \subset \mathsf{Span}_K(U_1 \cup \dots \cup U_n)$, weil die Elemente von $U_1 + \dots + U_n$ gerade als Summen von Elementen aus den U_i definiert sind. Umgekehrt sei $v \in \mathsf{Span}_K(U_1 \cup \dots \cup U_n)$, also

$$v = \sum_{i=1}^{m} \lambda_i v_i,$$

wobei die $\lambda_i \in K$ und $v_i \in U_1 \cup \dots \cup U_n$ sind. Wir sortieren die v_i danach, in welchem U_j sie liegen. Da wir nicht vorausgesetzt haben, dass $U_i \cap U_j = \{0_V\}$ ist, ist diese Darstellung nicht eindeutig, aber es existieren $v_1^{(j)}, \dots, v_{m_j}^{(j)} \in U_j$, so dass wir die obige Summe umsortieren können zu

$$v = \sum_{j=1}^{m} \left(\sum_{k=1}^{m_j} \lambda_k^{(j)} v_k^{(j)} \right)$$

und dies ist eine Summendarstellung wie in Definition II.7.1.

Für (b) überlegen wir zunächst, dass die Behauptung wahr ist, falls eines der U_i unendliche Dimension hat. Wir können also ohne Beschränkung der Allgemeinheit annehmen, dass alle U_i endlich-dimensional sind. Wir wählen eine Basis $\mathcal{B}_i$ von U_i. Damit ist

$$\mathsf{Span}_K(\mathcal{B}_1 \cup \dots \cup \mathcal{B}_n) = \mathsf{Span}_K(U_1 \cup \dots \cup U_n) = U_1 + \dots + U_n.$$

Damit ist $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_n$ ein Erzeugendensystem, aus dem man eine Basis als Teilmenge auswählen kann, also gilt die Abschätzung:

$$\dim_K(U_1 + \dots + U_n) \leqslant |\mathcal{B}_1 \cup \dots \cup \mathcal{B}_n| \leqslant \sum_{i=1}^{n} |\mathcal{B}_i| = \sum_{i=1}^{n} \dim_K U_i.$$

$\square$

Beispiel II.7.3. Für $V = \mathbb{R}^3$ betrachten wir die Untervektorräume

$$U_1 = \mathsf{Span}_{\mathbb{R}}(e_1, e_2) \text{ und } U_2 = \mathsf{Span}_{\mathbb{R}}(e_2, e_3).$$

Hier ist jeweils $\dim_{\mathbb{R}} U_i = 2$. Die beiden Unterräume schneiden sich in $\mathsf{Span}_{\mathbb{R}}(e_2)$, also

$$\dim_{\mathbb{R}} U_1 \cap U_2 = 1.$$

Es gilt außerdem

$$\dim_{\mathbb{R}}(U_1 + U_2) = \dim_{\mathbb{R}} \mathbb{R}^3 = 3,$$

weil $\mathsf{Span}_{\mathbb{R}}(e_1, e_2, e_3) = \mathbb{R}^3$.

Das Verhalten im obigen Beispiel ist ein allgemeines Phänomen:

Satz II.7.4. *Ist V ein K-Vektorraum und sind U_1, U_2 Untervektorräume endlicher Dimension, so gilt:*

$$\dim_K(U_1 + U_2) = \dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2).$$

Im obigen Beispiel ist dies $3 = 2 + 2 - 1$.

Beweis. Es sei $\mathcal{B} = (v_1, \ldots, v_m)$ eine Basis von $U_1 \cap U_2$. Wir ergänzen diese zu einer Basis
$\mathcal{B}_1 := \{v_1, \ldots, v_m, w_1, \ldots, w_k\}$ von U_1 und
$\mathcal{B}_2 := \{v_1, \ldots, v_m, u_1 \ldots, u_\ell\}$ von U_2.
Wir behaupten, dass dann

$$\tilde{\mathcal{B}} := \{v_1, \ldots, v_m, w_1, \ldots, w_k, u_1 \ldots, u_\ell\}$$

eine Basis von $U_1 + U_2$ ist. Das zeigt die Behauptung, weil

$$\dim_K(U_1 + U_2) = m + k + \ell = m + k + m + \ell - m = \dim_K U_1 + \dim_K U_2 - \dim_K(U_1 \cap U_2).$$

ist.

Die Menge $\tilde{\mathcal{B}}$ ist ein Erzeugendensystem von $U_1 + U_2$: Wir wissen, dass für $i = 1, 2$ gilt, dass $U_i = \mathsf{Span}_K(\mathcal{B}_i)$. Damit folgt aber sofort, dass $U_1 \cup U_2 \subset \mathsf{Span}_K(\mathcal{B}_1 \cup \mathcal{B}_2)$ ist und ebenso

$$U_1 + U_2 = \mathsf{Span}_K(U_1 \cup U_2) \subset \mathsf{Span}_K(\mathcal{B}_1 \cup \mathcal{B}_2) \subset U_1 + U_2.$$

Mit $\tilde{\mathcal{B}} = \mathcal{B}_1 \cup \mathcal{B}_2$ folgt die Behauptung.

Die Menge $\tilde{\mathcal{B}}$ ist linear unabhängig: Wir nehmen an, dass es eine Linearkombination der Null gibt:

$$\sum_{r=1}^{m} \lambda_r v_r + \sum_{s=1}^{k} \mu_s w_s + \sum_{t=1}^{\ell} \nu_t u_t = 0. \tag{II.7.1}$$

Wir setzen

$$v := \sum_{r=1}^{m} \lambda_r v_r + \sum_{s=1}^{k} \mu_s w_s.$$

Damit ist $v \in U_1$ und wegen der Gleichung ist $v = -\sum_{t=1}^{\ell} \nu_t u_t$ ein Element aus U_2. Also ist $v \in U_1 \cap U_2$. Wegen der Eindeutigkeit der Darstellung von Elementen als Linearkombinationen von Basiselementen muss gelten, dass $\mu_1, \ldots, \mu_k = 0$. Ebenso liefert der Vergleich von

$$\sum_{r=1}^{m} \lambda_r v_r = -\sum_{t=1}^{\ell} \nu_t u_t,$$

dass die Koeffizienten $\lambda_1, \ldots, \lambda_m$ und $\nu_1, \ldots, \nu_\ell$ alle null sein müssen. Damit sind alle Koeffizienten in (II.7.1) null. $\square$

Lemma II.7.5. *Ist V ein endlich-dimensionaler K-Vektorraum und sind U_1 und U_2 Untervektorräume mit $U_1 + U_2 = V$, so sind äquivalent:*

(a) $U_1 \cap U_2 = \{0_V\}$.
(b) *Jedes $v \in V$ läßt sich eindeutig als Summe $v = u_1 + u_2$ schreiben mit $u_1 \in U_1$ und $u_2 \in U_2$.*
(c)
$$\dim_K(U_1 + U_2) = \dim_K U_1 + \dim_K U_2.$$

Beweis. (c) $\Rightarrow$ (a): Mit Satz II.7.4 folgt, dass in der Situation in (c) gilt, dass $\dim_K(U_1 \cap U_2) = 0$ ist, also $U_1 \cap U_2 = \{0_V\}$.

(a) $\Rightarrow$ (b): Da $V = U_1 + U_2$ können wir jedes $v \in V$ schreiben als $v = u_1 + u_2$ mit $u_i \in U_i$. Wir nehmen an, dass wir zwei solcher Darstellungen finden können, also zusätzlich noch $v = u_1' + u_2'$ mit $u_i' \in U_i$. Wir stellen

$$v = u_1 + u_2 = u_1' + u_2'$$

um und erhalten

$$u_1 - u_1' = u_2' - u_2.$$

Also liegen sowohl $u_1 - u_1'$ als auch $u_2' - u_2$ in $U_1 \cap U_2$ aber dieser Schnitt ist $\{0_V\}$. Somit gilt

$$u_1 = u_1' \text{ und } u_2 = u_2'.$$

(b) $\Rightarrow$ (c): Ist $\dim_K(U_1 \cap U_2) > 0$, so existiert ein $0_V \neq u \in U_1 \cap U_2$. Dann hat aber der Vektor 0_V zwei verschiedene Darstellungen:

$$0_V = 0_V + 0_V = u - u = u + (-u)$$

im Widerspruch zur Annahme. $\square$

Definition II.7.6. Ist $V = U_1 + U_2$ mit $U_1 \cap U_2 = \{0_V\}$, so heißt V die *(innere) direkte Summe von U_1 und U_2*. In diesem Fall schreiben wir $U_1 \oplus U_2$ für $U_1 + U_2$.

Beispiel II.7.7. Ist $V = \mathbb{R}^3$ mit $U_1 = \mathsf{Span}_{\mathbb{R}}(e_1, e_3)$ und $U_2 = \mathsf{Span}_{\mathbb{R}}(e_2)$, so ist $U_1 + U_2 = U_1 \oplus U_2$, weil der Schnitt von U_1 und U_2 nur aus dem Nullvektor besteht.

Das nächste Resultat folgt aus Lemma II.7.5.

Korollar II.7.8. *Ist V ein endlich-dimensionaler K-Vektorraum und sind U_1, U_2 Untervektorräume, so sind äquivalent:*

(a) $V = U_1 \oplus U_2$,
(b) *Für jede Basis $\mathcal{B}_1$ von U_1 und $\mathcal{B}_2$ von U_2 ist $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ eine disjunkte Vereinigung und ergibt eine Basis von V.*
(c) $V = U_1 + U_2$ *und* $\dim_K V = \dim_K U_1 + \dim_K U_2$.

Beweis. (a) $\Rightarrow$ (c) folgt aus der Dimensionsformel für die Summe aus Satz II.7.4.

(c) $\Rightarrow$ (b): $\mathcal{B}$ erzeugt $\mathsf{Span}_K(U_1 \cup U_2) = U_1 + U_2 = V$. Damit folgt, dass

$$\dim_K V \leqslant |\mathcal{B}| \leqslant |\mathcal{B}_1| + |\mathcal{B}_2| = \dim_K U_1 + \dim_K U_2 = \dim_K V.$$

Also ist $\dim_K V = |\mathcal{B}|$ und $\mathcal{B}$ ist eine Basis von V. Es muss auch die Gleichheit $|\mathcal{B}_1| + |\mathcal{B}_2| = |\mathcal{B}_1 \cup \mathcal{B}_2|$ gelten. Das geht nur, wenn $\mathcal{B}_1 \cap \mathcal{B}_2 = \varnothing$ ist.

(b) $\Rightarrow$ (a): Es sei $\mathcal{B}_i$ jeweils eine Basis für U_i für $i \in \{1, 2\}$. Mit (b) ist dann $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ eine Basis von V und $\mathcal{B}_1 \cap \mathcal{B}_2 = \varnothing$. Es ist

$$V = \mathsf{Span}_K(\mathcal{B}) = \mathsf{Span}_K(\mathcal{B}_1 \cup \mathcal{B}_2) = \mathsf{Span}_K(U_1 \cup U_2) = U_1 + U_2$$

und damit folgt

$$\dim_K V = |\mathcal{B}| = |\mathcal{B}_1| + |\mathcal{B}_2| = \dim_K U_1 + \dim_K U_2,$$

weil $\mathcal{B}_1 \cap \mathcal{B}_2 = \varnothing$. Damit ist $\dim_K(U_1 \cap U_2) = 0$ und $U_1 \cap U_2 = 0_V$. $\square$

KAPITEL III

Lineare Abbildungen

III.1. Definition und Dimensionsformel

Lineare Abbildungen erhalten die Struktur von Vektorräumen. Sie respektieren die Addition und die Multiplikation mit Skalaren

Definition III.1.1. Es sei K ein Körper und V und W seien K-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *K-linear* (oder *Vektorraumhomomorphismus*), falls gilt:

(L1) f ist ein Gruppenhomomorphismus $f\colon (V,+) \to (W,+)$.
(L2) Für alle $v \in V$ und $\lambda \in K$ ist $f(\lambda v) = \lambda f(v)$.

Bemerkung III.1.2.

- $f\colon V \to W$ ist genau dann K-linear, falls gilt:

$$\forall \lambda, \mu \in K, \forall v, v' \in V : f(\lambda v + \mu v') = \lambda f(v) + \mu f(v'). \tag{L}$$

  Dass (L) (L1) impliziert, folgt mit $\lambda = \mu = 1_K$ und (L2) folgt mit $\mu = 0_K$. Umgekehrt, falls (L1) und (L2) gelten, dann erhalten wir wegen (L1), dass $f(\lambda v + \mu v') = f(\lambda v) + f(\mu v')$ ist und mit (L2) können wir dies weiter umformen zu $\lambda f(v) + \mu f(v')$.
- Ist f K-linear, dann gilt auch für endliche Summen

$$f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i f(v_i).$$

Satz III.1.3. *Es sei $f\colon V \to W$ eine K-lineare Abbildung. Dann gilt:*

(a) $f(0_V) = 0_W$ *und* $f(v - v') = f(v) - f(v')$ *für alle* $v, v' \in V$.
(b) *Ist $U \subset V$ ein Untervektorraum, so ist das Bild $f(U)$ ein Untervektorraum von W. Insbesondere ist $f(V)$ ein Untervektorraum von W.*
(c) *Ist $W' \subset W$ ein Untervektorraum, so ist das Urbild $f^{-1}(W')$ ein Untervektorraum von V. Insbesondere ist* $\ker(f) = f^{-1}(0_W)$ *ein Untervektorraum von V.*
(d) *Ist f zusätzlich bijektiv, so ist f^{-1} eine K-lineare Abbildung.*

Beweis. (a) hatten wir schon in Satz II.1.15 gesehen.

Zu (b): Da $0_v \in U$ gilt, ist $f(0_V) \in f(U)$, also ist $f(U) \neq \varnothing$. Sind $w_1, w_2 \in f(U)$, so gibt es $u_1, u_2 \in U$ mit $f(u_1) = w_1$ und $f(u_2) = w_2$. Damit gilt für $\lambda, \mu \in K$:

$$f(\lambda u_1 + \mu u_2) = \lambda f(u_1) + \mu f(u_2) = \lambda w_1 + \mu w_2 \in f(U).$$

Damit sind die Untervektorraumaxiome für $f(U)$ erfüllt.

Zu (c): Ist $W' \subset W$ ein Untervektorraum, so ist $0_W \in W'$. Damit ist $0_V \in f^{-1}(0_W) \subset f^{-1}(W')$, insbesondere ist $f^{-1}(W)$ nicht leer. Sind $v_1, v_2 \in f^{-1}(W')$ und sind $\lambda_1, \lambda_2 \in K$, so ist

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) \in W'$$

und damit ist $\lambda_1 v_1 + \lambda_2 v_2 \in f^{-1}(W')$.

Zu (d): Nach Satz II.1.15 ist f^{-1} ein Gruppenhomomorphismus bzüglich der Addition. Für $w \in W$ und $\lambda \in K$ ist

$$\lambda f^{-1}(w) = f^{-1}(f(\lambda f^{-1}(w))) = f^{-1}(\lambda f(f^{-1}(w))) = f^{-1}(\lambda w).$$

□

Beispiele III.1.4.

- Ist $V = W$, so ist die identische Abbildung K-linear.
- Ist $V = W = \mathbb{R}$, dann ist id_V die Abbildung, die $x \in \mathbb{R}$ auf x abbildet. Dagegen sind die Abbildungen $x \mapsto x^n$ für $n \in \mathbb{N}, n > 1$ *nicht* linear! Genausowenig sind $x \mapsto \sin(x)$, $x \mapsto \cos(x)$ oder $x \mapsto e^x$ linear.
- Sind V und W beliebige K-Vektorräume, so ist die Nullabbildung K-linear: Da $f(v) = 0_W$ für alle v gilt, ist auch $f(\lambda v + \mu v') = 0_W$ und das ist gleich $\lambda 0_W + \mu 0_W = \lambda f(v) + \mu f(v')$.
- Es sei $\lambda \in K$ und $V = W$. Dann ist die Streckung um λ eine K-lineare Abbildung: Wir haben $f\colon V \to V$, $f(v) = \lambda v$. Dann folgt mit den Vektorraumaxiomen

  $$f(\mu v + \nu v') = \lambda(\mu v + \nu v') = \lambda\mu v + \lambda\nu v' = \mu\lambda v + \nu\lambda v' = \mu f(v) + \nu f(v')$$

  für alle $\mu, \nu \in K$, $v, v' \in V$.

  Der Kern von f ist

  $$\ker(f) = \{v \in V, \lambda v = 0_V\} = \begin{cases} 0_V, & \lambda \neq 0_K \\ V, & \lambda = 0_K \end{cases}$$

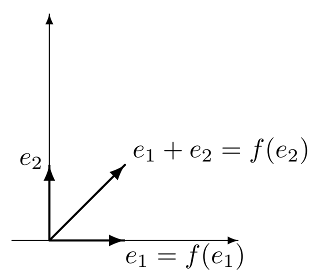
  und das Bild von f ist

  $$f(V) = \{v \in V, \exists v' \in V : \lambda v' = v\} = \begin{cases} 0_V, & \lambda = 0_K \\ V, & \lambda \neq 0_K. \end{cases}$$

  Im zweiten Fall benutzen Sie $\lambda\lambda^{-1}v = v$.
- Eine einfache aber wichtige $\mathbb{R}$-lineare Abbildung des $\mathbb{R}^2$ ist eine *Scherungsabbildung*: Für $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ ist

  $$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ y \end{pmatrix}.$$

  Rechnen Sie nach, dass die Scherungsabbildung $\mathbb{R}$-linear ist. Sie bildet die Standardbasisvektoren wie folgt ab: $f(e_1) = e_1$ und $f(e_2) = e_1 + e_2$.



- Eine weitere wichtige Beispielklasse linearer Abbildungen sind *Drehungen*. Im $\mathbb{R}^2$ kann man Drehungen explizit beschreiben. Für höhere Dimensionen lernen Sie eine einfache Darstellung im zweiten Semester kennen.

  Wir halten einen Winkel $\theta \in \mathbb{R}$ fest. Wir nennen die Abbildung

  $$R_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos(\theta)\cdot x - \sin(\theta)\cdot y \\ \sin(\theta)\cdot x + \cos(\theta)\cdot y \end{pmatrix}$$

  die *Drehung um* θ.

  Setzen wir wiederum die Standardbasisvektoren ein, so erhalten wir

  $$R_\theta(e_1) = R_\theta\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

und
$$R_\theta(e_2) = R_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}.$$
Wir rechnen die $\mathbb{R}$-Linearität von R_θ nach.
Für die Additivität benutzen wir die Definition der Vektoraddition im $\mathbb{R}^2$ und erhalten:
$$\begin{aligned} R_\theta(\begin{pmatrix} x+x' \\ y+y' \end{pmatrix}) &= \begin{pmatrix} \cos(\theta)\cdot(x+x') - \sin(\theta)\cdot(y+y') \\ \sin(\theta)\cdot(x+x') + \cos(\theta)\cdot(y+y') \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta)\cdot x - \sin(\theta)\cdot y \\ \sin(\theta)\cdot x + \cos(\theta)\cdot y \end{pmatrix} + \begin{pmatrix} \cos(\theta)\cdot x' - \sin(\theta)\cdot y' \\ \sin(\theta)\cdot x' + \cos(\theta)\cdot y' \end{pmatrix} \\ &= R_\theta(\begin{pmatrix} x \\ y \end{pmatrix}) + R_\theta(\begin{pmatrix} x' \\ y' \end{pmatrix}). \end{aligned}$$
Für ein $\lambda \in \mathbb{R}$ und $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ erhalten wir:
$$R_\theta(\lambda \begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} \cos(\theta)\cdot(\lambda x) - \sin(\theta)\cdot(\lambda y) \\ \sin(\theta)\cdot(\lambda x) + \cos(\theta)\cdot(\lambda y) \end{pmatrix}$$
Hier können wir λ jeweils ausklammern und mit der Definition der Streckung von Vektoren im $\mathbb{R}^2$ können wir λ herausziehen und erhalten $\lambda R_\theta \begin{pmatrix} x \\ y \end{pmatrix}$.

- Ist K beliebig und ist $V = W = K[X]$, so definieren wir die *formale Ableitung von Polynomen, D,* als K-lineare Abbildung
$$D\colon K[X] \to K[X], \quad D\left(\sum_{i=0}^{n} a_i X^i\right) := \sum_{i=0}^{n} i a_i X^{i-1}.$$
Für $K = \mathbb{R}$ sollte Ihnen das bekannt vorkommen. Vorsicht! Für andere Körper K verhält sich D anders, als Sie das vielleicht gewöhnt sind. So ist zum Beispiel für $K = \mathbb{F}_p$ die formale Ableitung des Polynoms $X^p + X^{p^2}$
$$D(X^p + X^{p^2}) = \overline{p} X^{p-1} + \overline{p^2} X^{p^2-1}.$$
Aber da $\overline{p} = \overline{0} = \overline{p^2}$, verschwindet die formale Ableitung, also
$$D(X^p + X^{p^2}) = 0_{\mathbb{F}_p[X]}.$$
Sie können sich leicht weitere Beispiele überlegen.

Die Komposition linearer Abbildungen ist wiederum linear:

Lemma III.1.5. *Es seien V, W, Z K-Vektorräume und es seien $f\colon V \to W$ und $g\colon W \to Z$ K-lineare Abbildungen. Dann ist auch $g \circ f$ K-linear.*

Beweis. Für $\lambda, \mu \in K$ und $v, v' \in V$ rechnen wir nach:
$$\begin{aligned} (g \circ f)(\lambda v + \mu v') &= g(f(\lambda v + \mu v') && \text{nach Definition von } g \circ f \\ &= g(\lambda f(v) + \mu f(v')) && \text{weil } f\, K\text{-linear ist} \\ &= \lambda g(f(v)) + \mu g(f(v')) && \text{weil } g\, K-\text{linear ist} \\ &= \lambda (g \circ f)(v) + \mu (g \circ f)(v') && \text{nach Definition von } g \circ f. \end{aligned}$$
□

Definition III.1.6.

(a) Eine K-lineare Abbildung $f\colon V \to W$ heißt
  - *Monomorphismus*, falls f injektiv ist,
  - *Epimorphismus*, falls f surjektiv ist,
  - *Isomorphismus*, falls f bijektiv ist,
  - *Endomorphismus*, falls $V = W$ ist, und

- *Automorphismus*, falls f ein bijektiver Endomorphismus ist.

(b) Zwei K-Vektorräume V und W heißen *isomorph*, $V \cong W$, falls es einen Isomorphismus $f\colon V \to W$ gibt.

(c) Die Zahl $\dim_K(\mathrm{Bild}(f)) = \dim_K f(V)$ heißt der *Rang der Abbildung* f. Wir schreiben dafür kurz $\mathsf{rg}(f)$.

Bemerkung III.1.7.

- Isomorphie ist eine Äquivalenzrelation.
- Ist V ein K-Vektorraum, so bilden die Automorphismen von V eine Gruppe, $Aut(V)$. Machen Sie sich bitte die Gruppenstruktur im Detail klar.
- Ist $f\colon V \to W$ eine K-lineare Abbildung und X ein Erzeugendensystem von V, so ist $f(X)$ ein Erzeugendensystem von $f(V)$: Ist $w \in f(V)$, so gibt es ein v in V mit $f(v) = w$. Da X ein Erzeugendensystem von V ist, ist $\mathsf{Span}_K(X) = V$, also gibt es $x_1, \ldots, x_n \in X$ und $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \sum_{i=1}^n \lambda_i x_i$. Damit ist aber auch
$$w = f(v) = f\left(\sum_{i=1}^n \lambda_i x_i\right) = \sum_{i=1}^n \lambda_i f(x_i).$$

Der folgende Satz beschreibt eine wichtige Beziehung zwischen der Dimension eines Definitionsbereiches einer linearen Abbildung und der Dimension von Kern und Bild dieser Abbildung. Wir werden ihn sehr oft benutzen.

Satz III.1.8 (Dimensionsformel)**.** *Ist $f\colon V \to W$ eine K-lineare Abbildung und ist V endlich-dimensional, so gilt:*
$$\dim_K V = \dim_K \ker(f) + \mathsf{rg}(f) = \dim_K \ker(f) + \dim_K \mathit{Bild}(f).$$

Beweis. Da $\ker(f) \subset V$ ein Untervektorraum ist, gilt auch $\dim_K \ker(f) < \infty$. Es sei $\mathcal{B}' = \{v_1, \ldots, v_k\}$ eine Basis von $\ker(f)$. Wir können diese mit dem Basisergänzungssatz zu einer Basis
$$\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$$
von V ergänzen. Wir wissen, dass $f(\mathcal{B})$ ein Erzeugendensystem von $f(V)$ ist, aber wir wissen auch, dass $f(v_1) = \ldots = f(v_k) = 0$.

Also ist $\{f(v_{k+1}), \ldots, f(v_n)\}$ ein Erzeugendensystem von $f(V)$. Nehmen wir an, dass gilt
$$\sum_{i=k+1}^n \lambda_i f(v_i) = 0_W$$
für $\lambda_i \in K$. Da f linear ist, ist dies gleichbedeutend mit
$$f\left(\sum_{i=k+1}^n \lambda_i v_i\right) = 0_W$$
aber damit ist $\sum_{i=k+1}^n \lambda_i v_i$ ein Element im Kern von f. Es gibt also $\lambda_1, \ldots, \lambda_k \in K$, so dass gilt
$$\sum_{i=k+1}^n \lambda_i v_i = -\sum_{i=1}^k \lambda_i v_i$$
also
$$\sum_{i=1}^n \lambda_i v_i = 0_V.$$
Aber $\{v_1, \ldots, v_n\}$ ist eine Basis von V, also müssen alle $\lambda_i = 0_K$ sein. Insbesondere ist dann auch
$$\lambda_{k+1} = \ldots = \lambda_n = 0_K$$
und damit ist $(f(v_{k+1}), \ldots, f(v_n)\}$ linear unabhängig, also eine Basis des Bildes von f. Abzählen ergibt nun
$$n = \dim_K V = k + (n-k) = \dim_K \ker(f) + \dim_K \mathrm{Bild}(f).$$
$\square$

Korollar III.1.9. *Ist $f\colon V \to V$ ein Endomorphismus und gilt $\dim_K V < \infty$, dann sind äquivalent.*

(a) *f ist injektiv,*
(b) *f ist surjektiv,*
(c) *f ist bijektiv.*

Beweis. Die Abbildung f ist genau dann injektiv, wenn $\ker(f) = \{0_V\}$, und dies ist äquivalent dazu, dass $\dim_K \ker(f) = 0$ gilt. Mit der Dimensionsformel aus Satz III.1.8 ergibt dies

$$\dim_K V = \dim_K \mathrm{Bild}(f)$$

also ist $\mathrm{Bild}(f) \subset V$ ein Untervektorraum gleicher Dimension, und dies ist äquivalent zu $\mathrm{Bild}(f) = V$, also zur Surjektivität von f. □

Beispiel III.1.10. Vorsicht, Korollar III.1.9 gilt *nicht* für unendlich-dimensionale Vektorräume. Zum Beispiel ist die formale Ableitung $D\colon \mathbb{R}[X] \to \mathbb{R}[X]$ surjektiv, weil

$$D(\frac{1}{n}X^n) = X^{n-1}, \text{ für } n \geqslant 1$$

und damit sind alle Basiselemente im Bild. Aber konstante Polynome haben triviale Ableitung, also ist D nicht injektiv.

Satz III.1.11. *Es seien V und W K-Vektorräume.*

(a) *Die Menge*

$$\mathsf{Hom}_K(V,W) := \{f\colon V \to W, f \textit{ ist } K\textit{-linear}\}$$

ist ein K-Vektorraum. Er ist ein Untervektorraum von $Abb(V,W)$.
(b) *Ist $V = W$, so ist $\mathsf{End}_K(V) := \mathsf{Hom}_K(V,V)$ ein Ring mit 1.*

Der Ring $\mathsf{End}_K(V)$ heißt der *Endomorphismenring von V*.

Beweis. Zu (a): Sind $f, g \in \mathrm{Abb}(V,W)$, so setzen wir für alle $v \in V$ und für alle $\lambda \in K$

$$(f+g)(v) := f(v) + g(v),$$
$$(\lambda f)(v) := \lambda f(v).$$

Damit wird $\mathrm{Abb}(V,W)$ zu einem K-Vektorraum. Sind $f, g \in \mathsf{Hom}_K(V,W)$, so zeigt eine Rechnung, dass auch $\lambda f + \mu g \in \mathsf{Hom}_K(V,W)$ sind für alle $\lambda, \mu \in K$.

Zu (b): Als multiplikative Verknüpfung auf $\mathsf{End}_K(V)$ setzen wir für $f, g\colon V \to V$ und $v \in V$

$$(gf)(v) := (g \circ f)(v).$$

Damit wird $\mathsf{End}_K(V)$ zu einem Ring. Das Einselement ist die identische Abbildung, id_V. □

III.2. Matrizen

Unser Ziel ist eine möglichst explizite Beschreibung linearer Abbildungen durch die Wahl von Basen und die Angabe ihrer Bilder.

Wodurch sind lineare Abbildungen festgelegt?

Satz III.2.1. *Es seien V und W K-Vektorräume endlicher Dimension, $v_1, \ldots, v_n \in V$ und $w_1, \ldots, w_n \in W$ seien Vektoren. Dann gilt:*

(a) *Ist die Familie $(v_1, \ldots, v_n)$ linear unabhängig, so gibt es mindestens eine lineare Abbildung $f\colon V \to W$ mit $f(v_i) = w_i$ für $1 \leqslant i \leqslant n$.*
(b) *Ist $(v_1, \ldots, v_n)$ eine Basis von V, so gibt es genau eine lineare Abbildung $f\colon V \to W$ mit $f(v_i) = w_i$ für $1 \leqslant i \leqslant n$. Dieses f erfüllt dann, dass $Bild(f) = \mathsf{Span}_K(w_1, \ldots, w_n)$ und dass f genau dann injektiv ist, wenn die Vektoren $(w_1, \ldots, w_n)$ linear unabhängig sind.*

Beweis. Wir beweisen zunächst (b): Jedes $v \in V$ hat eine eindeutige Darstellung als

$$v = \sum_{i=1}^{n} \lambda_i v_i.$$

Haben wir die Werte $f(v_i) = w_i$ gewählt, ist also

$$f(v) = \sum_{i=1}^{n} \lambda_i w_i$$

festgelegt. Sie rechnen leicht nach, dass das so definierte f K-linear ist.

Dies zeigt, dass $\mathrm{Bild}(f) \subset \mathsf{Span}_K(w_1, \ldots, w_n)$. Ist umgekehrt ein $w \in \mathsf{Span}_K(w_1, \ldots, w_n)$, also

$$w = \sum_{i=1}^{n} \mu_i w_i$$

für passende $\mu_i \in K$, so ist

$$w = \sum_{i=1}^{n} \mu_i f(v_i) = f\left(\sum_{i=1}^{n} \mu_i v_i\right)$$

Also $\mathsf{Span}_K(w_1, \ldots, w_n) \subset \mathrm{Bild}(f)$. Damit gilt

$$\mathrm{Bild}(f) = \mathsf{Span}_K(w_1, \ldots, w_n).$$

Ist $(w_1, \ldots, w_n)$ linear abhängig, so gibt es $\mu_1, \ldots, \mu_n \in K$, die nicht alle null sind, mit

$$0 = \sum_{i=1}^{n} \mu_i w_i.$$

Damit ist $v := \sum_{i=1}^{n} \mu_i v_i \neq 0_V$ aber $f(v) = 0_W$. Also ist f dann nicht injektiv.

Ist umgekehrt $f(v) = 0$, so schreiben wir v eindeutig als

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

mit $\lambda_i \in K$. Also gilt

$$0_W = f(v) = \sum_{i=1}^{n} \lambda_i f(v_i) = \sum_{i=1}^{n} \lambda_i w_i.$$

Ist $(w_1, \ldots, w_n)$ linear unabhängig, so folgt $\lambda_1 = \ldots = \lambda_n = 0$ und damit war $v = 0_V$ und f ist injektiv.

Zu (a): Ist $(v_1, \ldots, v_n)$ lediglich linear unabhängig, so ergänzen wir $(v_1, \ldots, v_n)$ zu einer Basis $\mathcal{B} = (v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$ von V. Wir wählen beliebige weitere Elemente $w_{n+1}, \ldots, w_m \in W$ und definieren $f(v_i) := w_i$ für alle $1 \leqslant i \leqslant m$. $\square$

Lineare Abbildungen sind also durch die Werte auf einer Basis festgelegt. Wir benutzen dies zu einem Vergleich von V mit $\dim_K V = n < \infty$ mit K^n.

Korollar III.2.2. *Ist V ein K-Vektorraum und ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V, so ist V isomorph zu K^n. Insbesondere ist jeder endlich-dimensionale K-Vektorraum isomorph zu einem K^n mit einem eindeutig bestimmten $n \geqslant 0$.*

Beweis. Wir wissen, dass $(e_1, \ldots, e_n)$ eine Basis des K^n ist. Wir definieren $\Phi_{\mathcal{B}} \colon K^n \to V$ durch $\Phi_{\mathcal{B}}(e_i) = v_i$ für $1 \leqslant i \leqslant n$. Da $(v_1, \ldots, v_n)$ eine Basis ist und damit insbesondere linear unabhängig, ist $\Phi_{\mathcal{B}}$ injektiv. Nach Konstruktion ist $\Phi_{\mathcal{B}}$ surjektiv, also bijektiv.

Isomorphe Vektorräume haben die gleiche Dimension und für $n \neq m$ gilt $\dim_K K^n = n \neq m = \dim_K K^m$. Damit folgt die Eindeutigkeit. $\square$

Definition III.2.3.

(a) Es sei K ein Körper. Ein rechteckiges Schema

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in K$ heißt eine *$m \times n$-Matrix mit Einträgen in K*.

(b) Die Menge aller $m \times n$-Matrizen mit Einträgen in K bezeichnen wir mit $M(m \times n, K)$.

(c) Ist $f\colon K^n \to K^m$ eine K-lineare Abbildung und ist

$$f(e_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, f(e_n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

so heißt

$$M(f) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

die *darstellende Matrix von f*.

Bemerkung III.2.4. Satz III.2.1 besagt, dass f durch $M(f)$ eindeutig bestimmt ist und umgekehrt.

Ist $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in K^n$ ein beliebiger Vektor, so ist

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^n v_i e_i\right) \\ &= \sum_{i=1}^n v_i f(e_i) \\ &= v_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + v_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^n v_j a_{1j} \\ \vdots \\ \sum_{j=1}^n v_j a_{mj} \end{pmatrix} \end{aligned}$$

Damit können wir die Anwendung einer linearen Abbildung durch Multiplikation einer Matrix mit einem Vektor ausdrücken:

Definition III.2.5. Es sei $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ und $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$.

Die *Multiplikation von A mit v, Av*, ist

$$Av = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^n a_{1j} v_j \\ \vdots \\ \sum_{j=1}^n a_{mj} v_j \end{pmatrix}.$$

Sind $f\colon K^n \to K^m$ und $g\colon K^m \to K^\ell$ K-lineare Abbildungen, dann ist auch $g \circ f$ K-linear. Im Folgenden wollen wir die darstellende Matrix von $g \circ f$, $M(g \circ f)$, bestimmen. Es sei $(e_1, \dots, e_n)$ die Standardbasis des

K^n und $(e'_1, \ldots, e'_m)$ sei die Standardbasis des K^m. Die darstellenden Matrizen, $M(f)$ und $M(g)$, seien

$$M(f) = \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \ldots & b_{mn} \end{pmatrix}, \quad M(g) = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{\ell 1} & \ldots & a_{\ell m} \end{pmatrix}.$$

Für $M(g \circ f)$ gilt dann:

$$\begin{aligned} \begin{pmatrix} c_{1j} \\ \vdots \\ c_{\ell j} \end{pmatrix} &= (g \circ f)(e_j) \\ &= g(f(e_j)) \\ &= g\left(\sum_{s=1}^{m} b_{sj} e'_s\right) \\ &= \sum_{s=1}^{m} b_{sj} g(e'_S) \\ &= \sum_{s=1}^{m} b_{sj} \begin{pmatrix} a_{1s} \\ \vdots \\ a_{\ell s} \end{pmatrix}. \end{aligned}$$

Also ist

$$c_{ij} = \sum_{s=1}^{m} a_{is} b_{sj}.$$

Da wir die Verknüpfung linearer Abbildungen durch Matrizenmultiplikation ausdrücken möchten, definieren wir diese dementsprechend:

Definition III.2.6. Ist $A \in M(\ell \times m, K)$ und $B \in M(m \times n, K)$, so definieren wir das *Produkt von A mit B, AB,* dadurch, dass wir den Eintrag an der Stelle (i, j) von AB angeben:

$$(AB)_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Mit dieser Definition ist dann

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

Beispiel III.2.7. Die Verknüpfung von Abbildungen ist nicht kommutativ und das spiegelt sich in der Nichtkommutativität der Matrizenmultiplikation wider. Ist K beliebig und $1 = 1_K$, $0 = 0_K$, und ist $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, so gilt

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

wohingegen

$$BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Was sind die zugehörigen K-linearen Abbildungen, die durch A und B dargestellt werden?

Beispiel III.2.8. Wir hatten die Drehung um den Winkel θ auf dem $\mathbb{R}^2$ betrachtet mit

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) \cdot x - \sin(\theta) \cdot y \\ \sin(\theta) \cdot x + \cos(\theta) \cdot y \end{pmatrix}$$

und wir hatten die Bilder der Standardbasisvektoren schon ausgerechnet. Damit ist

$$M(R_\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Für zwei Winkel θ_1, θ_2 wissen wir, dass

$$M(R_{\theta_2} \circ R_{\theta_1}) = M(R_{\theta_2})M(R_{\theta_1})$$

gilt, also können wir direkt berechnen, was $M(R_{\theta_2} \circ R_{\theta_1})$ ist:

$$\begin{aligned}
M(R_{\theta_2} \circ R_{\theta_1}) =& \begin{pmatrix} \cos(\theta_2) & -\sin(\theta_2) \\ \sin(\theta_2) & \cos(\theta_2) \end{pmatrix} \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{pmatrix} \\
=& \begin{pmatrix} \cos(\theta_2)\cos(\theta_1) - \sin(\theta_2)\sin(\theta_1), & -\cos(\theta_2)\sin(\theta_1) - \sin(\theta_2)\cos(\theta_1) \\ \sin(\theta_2)\cos(\theta_1) + \cos(\theta_2)\sin(\theta_1), & -\sin(\theta_2)\sin(\theta_1) + \cos(\theta_2)\cos(\theta_1) \end{pmatrix} \\
=& \begin{pmatrix} \cos(\theta_2 + \theta_1) & -\sin(\theta_2 + \theta_1) \\ \sin(\theta_2 + \theta_1) & \cos(\theta_2 + \theta_1) \end{pmatrix} \\
=& M(R_{\theta_2+\theta_1}).
\end{aligned}$$

Hier haben wir wieder die Additionstheoreme für den Kosinus und Sinus benutzt. Die Kommata in der mittleren Matrix dienen nur der besseren Lesbarkeit. Anschaulich war das Ergebnis zu erwarten: Wenn wir erst um den Winkel θ_1 drehen und danach um den Winkel θ_2, so drehen wir insgesamt um den Winkel $\theta_1 + \theta_2$.

Bemerkung III.2.9. Da die Verkettung von Abbildungen assoziativ ist, ist auch die Multiplikation von Matrizen assoziativ.

Ist $f\colon K^n \to K^m$ eine K-lineare Abbildung, so ist

$$\mathsf{rg}(f) = \dim_K \operatorname{Bild}(f) = \dim_K \mathsf{Span}_K(f(e_1), \ldots, f(e_n)).$$

Dies ist gleich der Anzahl linear unabhängiger Spaltenvektoren in $M(f)$.

Beispiel III.2.10. Es seien $f, g\colon \mathbb{R}^3 \to \mathbb{R}^2$ die beiden folgenden Abbildungen:

$$M(f) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \end{pmatrix}, \quad M(g) = \begin{pmatrix} -1 & 2 & 1 \\ 5 & 0 & 2 \end{pmatrix}.$$

Dann ist $\mathsf{rg}(f) = 1$ und $\mathsf{rg}(g) = 2$. Es gilt zum Beispiel

$$\begin{pmatrix} -1 \\ 5 \end{pmatrix} = \frac{-7}{4} \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

In Analogie zu Satz III.1.11 definieren wir:

Definition III.2.11. Es seien $A, B \in M(m \times n, K)$ mit Einträgen a_{ij}, b_{ij} und $\lambda \in K$:

(a)

$$(A + B)_{ij} := (a_{ij} + b_{ij}) \quad \text{und } (\lambda A)_{ij} := \lambda a_{ij}.$$

(b) Die *Transponierte einer Matrix A, A^t*, ist die Matrix

$$(A^t)_{ij} = a_{ji}.$$

(c) Die Matrix

$$E_n = M(\mathrm{id}_{K^n}) = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

heißt die *$n \times n$-Einheitsmatrix*.

Bemerkung III.2.12. Man schreibt die Koordinaten für die Einheitsmatrix auch gerne in der geschlossenen Form

$$(E_n)_{ij} = \delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Die Funktion δ_{ij} heißt die *Kronecker-Delta-Funktion*.

Beispiel III.2.13. Ist

$$A = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ \frac{2}{5} & 1 & 4 & 2 \end{pmatrix} \in M(3 \times 4, \mathbb{Q})$$

so ist die Transponierte von A, A^t

$$A^t = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ \frac{2}{5} & 1 & 4 & 2 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 & \frac{2}{5} \\ 0 & 1 & 1 \\ 0 & 0 & 4 \\ \frac{1}{3} & 0 & 2 \end{pmatrix} \in M(4 \times 3, \mathbb{Q}).$$

Lemma III.2.14 (Rechenregeln für Matrizen)**.** *Sind $A, A' \in M(m \times n, K)$, $B, B' \in M(n \times r, K)$, $C \in M(r \times s, K)$ und $\lambda \in K$, so gilt.*

(a) $A(B + B') = AB + AB'$ *und* $(A + A')B = AB + A'B$.
(b) $(\lambda A)B = \lambda(AB) = A(\lambda B)$.
(c) $(AB)C = A(BC)$.
(d) $(AB)^t = B^t A^t$.
(e) $A = E_m A = AE_n$.

Beweis. Die Behauptungen (a), (b) und (e) zeigt man durch stures Nachrechnen, (c) hatten wir schon begründet.

Zu (d): A habe Einträge a_{ij} und B habe die Einträge b_{jk}. Aus der Formel für die Multiplikation von Matrizen erhalten wir

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} =: c_{ij}.$$

Wir wissen nach Definition der Transponierten, dass

$$((AB)^t)_{ij} = (AB)_{ji} = c_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Da $(B^t)_{jk} = b_{kj}$ und $(A^t)_{ij} = a_{ji}$, erhalten wir für den Eintrag ij in $B^t A^t$:

$$(B^t A^t)_{ij} = \sum_{k=1}^{n} b_{ki} a_{jk}$$

und da die Multiplikation im Körper K kommutativ ist, ist dies gleich

$$\sum_{k=1}^{n} a_{jk} b_{ki} = ((AB)^t)_{ij}.$$

□

Damit erhalten wir insbesondere:

Satz III.2.15. *Die Menge aller $n \times n$-Matrizen über einem Körper K, $M(n \times n, K)$, bildet mit der Addition und Multiplikation von Matrizen einen Ring mit $1 = E_n$.*

Bemerkung III.2.16. Für $n = 1$ ist $M(1 \times 1, K) = K$, aber für alle $n \geqslant 2$ ist $M(n \times n, K)$ nicht kommutativ.

Für $n \geqslant 2$ ist $M(n \times n, K)$ nicht nullteilerfrei. Zum Beispiel ist

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0_{M(2 \times 2, K)}.$$

Definition III.2.17. Eine Matrix $A \in M(n \times n, K)$ heißt *invertierbar*, falls es ein $A' \in M(n \times n, K)$ gibt mit $AA' = E_n = A'A$.

Satz III.2.18. *Die Menge*

$$GL_n(K) := \{A \in M(n \times n, K), A \text{ invertierbar}\}$$

bildet mit der Matrizenmultiplikation eine Gruppe mit neutralem Element E_n.

Beweis. Wir müssen zeigen, dass mit $A, B \in GL_n(K)$ auch $AB \in GL_n(K)$ ist. Es sei A' eine Matrix mit $AA' = E_n = A'A$ und ebenso sei B' eine Matrix mit $BB' = E_n = B'B$. Dann ist auch

$$(B'A')(AB) = B'(A'A)B = B'E_nB = B'B = E_n$$

und

$$(AB)(B'A') = A(BB')A = AE_nA' = AA' = E_n.$$

Wir wissen schon, dass Matrizenmultiplikation assoziativ ist und dass E_n ein neutrales Element ist. □

Definition III.2.19. Die Gruppe $GL_n(K)$ heißt die *allgemeine lineare Gruppe* der $n \times n$-Matrizen.

Hierbei kommt GL von *general linear*. Wir schreiben ab jetzt wieder A^{-1} für A'.

Satz III.2.20. *Es seien V und W K-Vektorräume, so dass $\mathcal{B}_1 = (v_1, \dots, v_n)$ eine geordnete Basis von V ist und $\mathcal{B}_2 = (w_1, \dots, w_m)$ eine geordnete Basis von W. Dann gibt es zu jeder linearen Abbildung $f\colon V \to W$ genau eine Matrix $A = (a_{ij}) \in M(m \times n, K)$, so dass*

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

gilt. Diese Matrix nennen wir $M_{\mathcal{B}_2}^{\mathcal{B}_1}(f)$.

Die Abbildung

$$M_{\mathcal{B}_2}^{\mathcal{B}_1}\colon \mathsf{Hom}_K(V, W) \to M(m \times n, K), \quad f \mapsto M_{\mathcal{B}_2}^{\mathcal{B}_1}(f)$$

ist ein Isomorphismus von K-Vektorräumen.

Definition III.2.21. Die Matrix $M_{\mathcal{B}_2}^{\mathcal{B}_1}(f)$ heißt die *Matrixdarstellung von f bezüglich der Basen $\mathcal{B}_1$ und $\mathcal{B}_2$*.

Nach Wahlen von geordneten Basen von V und W kann man also jede lineare Abbildung eindeutig durch eine Matrix beschreiben.

Beweis. Da $f(v_j) \in W$, können wir $f(v_j)$ als Linearkombination in den $w_1, \dots, w_m$ schreiben, also

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Da $(w_1, \dots, w_m)$ eine geordnete Basis von W ist, ist $f(v_j)$ und damit die j-te Spalte von A eindeutig bestimmt.

Ist $g\colon V \to W$ K-linear, also $f, g \in \mathsf{Hom}_K(V, W)$, und ist $M_{\mathcal{B}_2}^{\mathcal{B}_1}(g) = (b_{ij})$, so gilt:

$$\begin{aligned}(f+g)(v_j) &= f(v_j) + g(v_j) \\ &= \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i \\ &= \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i\end{aligned}$$

und für $\lambda \in K$ gilt

$$(\lambda f)(v_j) = \lambda \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \lambda a_{ij} w_i.$$

Damit gilt:

$$\begin{aligned}M_{\mathcal{B}_2}^{\mathcal{B}_1}(f+g) &= M_{\mathcal{B}_2}^{\mathcal{B}_1}(f) + M_{\mathcal{B}_2}^{\mathcal{B}_1}(g) \text{ und} \\ M_{\mathcal{B}_2}^{\mathcal{B}_1}(\lambda f) &= \lambda M_{\mathcal{B}_2}^{\mathcal{B}_1}(f)\end{aligned}$$

und somit ist $M_{\mathcal{B}_2}^{\mathcal{B}_1}$ eine K-lineare Abbildung. Satz III.2.1 impliziert, dass $M_{\mathcal{B}_2}^{\mathcal{B}_1}$ bijektiv ist. □

Beispiel III.2.22. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ die lineare Abbildung, die $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ auf $\begin{pmatrix} x_1 \\ x_1 + x_2 \end{pmatrix}$ abbildet. Es sei $S = (e_1, e_2)$ die geordnete Standardbasis und $\mathcal{B}_1 = S$, $\mathcal{B}_2 = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = (w_1, w_2)$. Wir berechnen $f(e_1) = e_1 + e_2 = w_1 = 1 \cdot w_1 + 0 \cdot w_2$ und $f(e_2) = e_2 = w_2 = 0 \cdot w_1 + 1 \cdot w_2$ und damit gilt

$$M_S^S(f) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad M_{\mathcal{B}_2}^{\mathcal{B}_1}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Korollar III.2.23. *Es seien V und W wiederum K-Vektorräume mit geordneten Basen $\mathcal{B}_1 = (v_1, \ldots, v_n)$ und $\mathcal{B}_2 = (w_1, \ldots, w_m)$. Dann bilden die linearen Abbildungen*

$$\varphi_j^i\colon V \to W, \quad \varphi_j^i(v_k) = \begin{cases} w_j, & \textit{falls } k = i, \\ 0, & \textit{sonst,} \end{cases}$$

für $1 \leqslant i \leqslant n$ und $1 \leqslant j \leqslant m$ eine Basis von $\mathsf{Hom}_K(V, W)$. Insbesondere ist

$$\dim_K \mathsf{Hom}_K(V, W) = \dim_K V \cdot \dim_K W.$$

Beweis. Wir setzen $E_j^i := M_{\mathcal{B}_2}^{\mathcal{B}_1}(\varphi_j^i)$. Dann ist E_j^i eine Matrix mit

$$(E_j^i)_{k\ell} = \begin{cases} 1, & k = j, \ell = i, \\ 0, & \text{sonst.} \end{cases}$$

Die E_j^i bilden eine Basis von $M(m \times n, K)$, weil wir jedes $A \in M(m \times n, K)$ eindeutig schreiben können als

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} E_i^j.$$

Mit Satz III.2.20 folgt die Behauptung. □

Beispiel III.2.24.

$$\begin{aligned} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} &= 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 4 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= 1 \cdot E_1^1 + 2 \cdot E_1^2 + 3 \cdot E_2^1 + 4 \cdot E_2^2. \end{aligned}$$

Bemerkung III.2.25.
(a) Ist $f \in \mathsf{End}_K(V)$ und ist $\mathcal{B}_1 = \mathcal{B}_2 = \mathcal{B}$, wobei $\mathcal{B} = (v_1, \ldots, v_n)$ eine festgewählte geordnete Basis von V ist, so schreiben wir kurz $M_{\mathcal{B}}(f)$ für $M_{\mathcal{B}}^{\mathcal{B}}(f)$.
(b) Die lineare Abbildung $M_{\mathcal{B}}\colon \mathsf{End}_K(V) \to M(n \times n, K)$ ist dann durch

$$f \mapsto (a_{ij}), \text{ mit } f(v_j) = \sum_{i=1}^{n} a_{ij} v_i$$

gegeben.
(c) Speziell für die identische Abbildung id_V erhalten wir $M_{\mathcal{B}}(\mathrm{id}_V) = E_n$.

III.3. Affine Unterräume und Abbildungen

Wir hatten die Beispiele affiner Geraden im $\mathbb{R}^2$ betrachtet, wobei $G_{p,v}$ für $p \in \mathbb{R}^2$ und $0 \neq v \in \mathbb{R}^2$ gegeben war durch

$$G_{p,v} = \{p + \lambda v, \lambda \in \mathbb{R}\}.$$

Wir werden allgemeine affine Unterräume dazu benutzen, um die Lösungsmengen linearer Gleichungssysteme zu beschreiben.

Definition III.3.1. Es sei V ein K-Vektorraum.
(a) Ist $M \subset V$ und ist $v \in V$, so ist $v + M$ die Menge

$$v + M := \{v + m, m \in M\}.$$

(b) Ist $U \subset V$ ein Untervektorraum, so heißt $v + U$ ein *affiner Unterraum* von V.

Lemma III.3.2. *Ist $A = v + U$ ein affiner Unterraum von V, so gilt:*

(a) $U = \{a - a', a, a' \in A\}$.
(b) $v + U = v' + U \Leftrightarrow v - v' \in U$.

Beweis. Zu (a): Für $u \in U$ setzen wir $a = v + u$ und $a' = v = v + 0_U$ und damit ist $u = a - a'$, also $U \subset \{a - a', a, a' \in A\}$. Umgekehrt seien a und a' aus A, also gibt es $u_1, u_2 \in U$ mit $a = v + u_1$ und $a' = v + u_2$. Damit ist dann $a - a' = u_1 - u_2 \in U$.

Zu (b): Ist $v + U = v' + U$, so ist $v = v + 0 \in v + U = v' + U$, also gibt es ein $u \in U$ mit $v = v' + u$. Damit ist dann $v - v' = u \in U$.

Es $v - v' \in U$ und $v + u \in v + U$, so ist

$$v + u = v' + (v - v') + u \in v + U.$$

Ist $v' + u \in v' + U$, so ist

$$v' + u = v + (v' - v) + U \in v' + U.$$

□

Definition III.3.3. Ist $A = v + U$ ein affiner Unterraum von V, so heißt $\dim_K U$ die *Dimension von A*:

$$\dim_K A := \dim_K U.$$

Bemerkung III.3.4. Lemma III.3.2 besagt, dass U durch A eindeutig bestimmt ist, das heißt die obige Definition der Dimension macht Sinn.

Ist $A = v + U$ und ist $v' \in A$ beliebig, so ist $v - v' \in U$ und damit ist $v + U = v' + U$. Damit eignet sich jedes beliebige $v' \in A$ als Fußpunkt.

Beispiel III.3.5. Es sei $V = \mathbb{R}^2$. Was sind mögliche affine Unterräume des $\mathbb{R}^2$?

- $\dim_{\mathbb{R}} A = 0$: In dem Fall ist $A = v + U$ mit $\dim_{\mathbb{R}} U = 0$. Damit muss aber U der Nullvektorraum $\{0\} \subset \mathbb{R}^2$ sein und $A = v + \{0\} = \{v\}$ ist ein Punkt.
- $\dim_{\mathbb{R}} A = 1$: Hier ist $A = v + U$ mit $\dim_{\mathbb{R}} U = 1$. Damit hat U eine Basis der Form $\{u\}$ mit $\mathbb{R}^2 \ni u \neq 0$ und $U = \{\lambda u, \lambda \in \mathbb{R}\}$ ist eine Gerade durch den Nullpunkt. In diesem Fall ist
$$A = v + U = \{v + \lambda u, \lambda \in \mathbb{R}\}$$
und dies ist die Gerade in Parameterform $G_{v,u}$. Solche Geraden nennen wir auch *affine Geraden*.
- $\dim_{\mathbb{R}} U = 2$: In diesem Fall ist also $U \subset \mathbb{R}^2$ ein Untervektorraum mit $\dim_{\mathbb{R}} U = \dim_{\mathbb{R}} \mathbb{R}^2 = 2$ und damit ist schon $U = \mathbb{R}^2$ und ebenso ist $v + U = U = \mathbb{R}^2$.

Bemerkung III.3.6.

- 2-dimensionale affine Unterräume nennt man *affine Ebenen*.
- Ist $\dim_K V = n$, so heißt ein $(n-1)$-dimensionaler affiner Unterraum von V eine *affine Hyperebene*.

Lemma III.3.7. *Es sei V ein K-Vektorraum.*

(a) *Ein affiner Unterraum $A = v + U$ von V ist genau dann ein Untervektorraum, wenn $0 \in A$.*
(b) *Ist $(A_i \subset V)_{i \in I}$ eine Familie affiner Unterräume von V, so ist ihr Schnitt $\bigcap_{i \in I} A_i$ entweder ein affiner Unterraum von V oder leer.*

Beweis. Zu (a): Ist $v + U = W$ ein Untervektorraum von V, so ist $0 \in v + U$, weil $0 \in W$.

Ist umgekehrt $0 \in v + U$, so ist $v \in U$ und damit ist $v + U = U$.

Zu (b): Ist $\bigcap_{i \in I} A_i \neq \varnothing$, so gibt es ein $v \in \bigcap_{i \in I} A_i$. Also ist $v \in A_i$ für alle $i \in I$. Damit können wir jedes A_i schreiben als $A_i = v + U_i$, wobei U_i ein Untervektorraum von V ist. Aber $\bigcap_{i \in I} U_i$ ist wiederum ein Untervektorraum von V, somit ist

$$\bigcap_{i \in I} A_i = \bigcap_{i \in I} (v + U_i) = v + \bigcap_{i \in I} U_i$$

wieder ein affiner Unterraum von V. □

Definition III.3.8. Es seien V, V' zwei K-Vektorräume. Eine Abbildung $f\colon V \to V'$ heißt *affine Abbildung*, falls es eine lineare Abbildung $\varphi\colon V \to V'$ gibt und ein $w \in V'$, so dass für alle $v \in V$ gilt:

$$f(v) = w + \varphi(v).$$

Bemerkung III.3.9. Ist f gegeben, so sind φ und w eindeutig bestimmt: $f(0) = w$ und $\varphi(v) = f(v) - w$.

Satz III.3.10. *Es seien V_1, V_2, V_3 K-Vektorräume und $f\colon V_1 \to V_2$, $g\colon V_2 \to V_3$ seien affine Abbildungen.*

(a) *Dann ist die Komposition $g \circ f$ ebenfalls affin.*
(b) *Ist $A \subset V_1$ ein affiner Unterraum von V_1 so ist $f(A)$ ein affiner Unterraum von V_2.*
(c) *Ist $B \subset V_2$ ein affiner Unterraum von V_2, so ist das Urbild $f^{-1}(B)$ entweder affin oder leer.*
(d) *Ist $f(v) = w + \varphi(v)$ für ein K-lineares φ, so gilt: f ist genau dann injektiv oder surjektiv, wenn φ es ist.*

Beweis. (a) ist eine einfache Rechnung: Ist $f(v) = w + \varphi(v)$ für alle $v \in V_1$ mit einem $w \in V_2$ und ist $g(v_2) = z + \psi(v_2)$ für alle $v_2 \in V_2$ mit einem $z \in V_3$, so ist

$$(g \circ f)(v) = g(w + \varphi(v)) = z + \psi(w + \varphi(v)) = z + \psi(w) + \psi(\varphi(v)).$$

Wir setzen $u := z + \psi(w) \in V_3$ und $\xi := \psi \circ \varphi$. Dann ist ξ K-linear und $g \circ f$ ist affin.

Zu (b): Ist $A = v + U$ ein affiner Unterraum von V und ist $f(v) = w + \varphi(v)$ wie oben, so ist

$$f(A) = w + \varphi(v + U) = w + \varphi(v) + \varphi(U).$$

Da $\varphi(U)$ ein Untervektorraum von V_2 ist und da $w + \varphi(v) \in V_2$, ist $f(A)$ wiederum affin.

Zu (c): Wir nehmen an, dass das Urbild $f^{-1}(B)$ nicht leer ist und es sei $v_0 \in f^{-1}(B)$. Wir setzen $w_0 := f(v_0) \in B$ und schreiben $B = w_0 + U_2$ mit einem Untervektorraum $U_2 \subset V_2$. Dann ist

$$\begin{aligned}
f^{-1}(B) &= \{v \in V_1, f(v) \in B\} \\
&= \{v \in V_1, w + \varphi(v) \in w_0 + U_2\} \\
&= \{v \in V_1, w + \varphi(v) \in f(v_0) + U_2\} \\
&= \{v \in V_1, w + \varphi(v), w + \varphi(v) \in w + \varphi(v_0) + U_2\} \\
&= \{v \in V_1, \varphi(v - v_0) \in U_2\} \\
&= \{v \in V_1, v - v_0 \in \varphi^{-1}(U_2)\} \\
&= v_0 + \varphi^{-1}(U_2).
\end{aligned}$$

Zu (d): Ist $f(v) = w + \varphi(v)$ für alle $v \in V_1$, so ist $f = T_w \circ \varphi$, wobei $T_w\colon V_2 \to V_2$, $T_w(v_2) = w + v_2$. Die Translation um w ist eine Bijektion mit Inversem T_{-w} und damit folgt die Behauptung. □

Bemerkung III.3.11. Da K-lineare Abbildungen insbesondere affine Abbildungen sind, können wir den obigen Satz auf $B = \{b\}$ anwenden und erhalten mit $f = \varphi$, dass das Urbild $\varphi^{-1}(b)$ entweder leer ist oder affin. Ist $a \in \varphi^{-1}(b)$, so ist

$$\varphi^{-1}(b) = a + \ker(\varphi).$$

Denn ist $v \in \ker(\varphi)$, so ist $\varphi(a + v) = \varphi(a) + \varphi(v) = b + 0 = b$ und ist $a' \in \varphi^{-1}(b)$, so ist $a' = a + (a' - a)$ und $\varphi(a' - a) = b - b = 0$. Damit ist $a' - a \in \ker(\varphi)$.

III.4. Lineare Gleichungssysteme

Definition III.4.1. Es sei K ein Körper.

(a) Ein *lineares Gleichungssystem, (LGS)* über K, ist ein System von Gleichungen der Form

$$\begin{array}{ccc}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\
\vdots & & \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m
\end{array} \qquad \text{(III.4.1)}$$

mit $a_{ij}, b_i \in K$. Gesucht sind $x_1, \ldots, x_n \in K$.
(b) Ist $b_1 = \ldots = b_m = 0$, so heißt das LGS *homogen*; ansonsten heißt es *inhomogen*.
(c) Ersetzt man in einem LGS (III.4.1) alle $b_1, \ldots, b_m$ durch 0, so erhält man das *zugehörige homogene LGS*.

(d) Die Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

heißt die *Koeffizientenmatrix des LGS* (III.4.1). Mit $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$ suchen wir also die *Lösungsmenge des LGS*

$$L(A;b) := \{x \in K^n, Ax = b\}.$$

(e) Die Matrix

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

heißt die *erweiterte Koeffizientenmatrix* des LGS (III.4.1).

Bemerkung III.4.2.

- Zu einem LGS $Ax = b$ wie in (III.4.1) betrachten wir die K-lineare Abbildung $f\colon K^n \to K^m$, die durch Matrixmultiplikation mit A gegeben ist.
- Die Menge $L(A;b)$ ist dann gerade das Urbild von b unter f:

  $$L(A;b) = f^{-1}(b).$$

  Mit Satz III.3.10 folgt, dass $L(A;b)$ entweder leer ist oder ein affiner Unterraum des K^n.
- Ist $v \in L(A;b)$, so ist

  $$L(A;b) = v + L(A;0) = v + \ker(f).$$

  Damit erhält man die allgemeine Lösung des LGS durch Addition der Lösungen des homogenen LGS zu einer speziellen Lösung.
- Wann ist $L(A;b) \neq \varnothing$?

  Es sei $\mathsf{rg}(A)$ die Dimension des von den Spalten von A aufgespannten Unterraums des K^m. Dann gilt:

  $$\mathsf{rg}(A|b) = \begin{cases} \mathsf{rg}(A), & \text{falls } b \text{ Linearkombination der Spaltenvektoren von } A \text{ ist,} \\ \mathsf{rg}(A) + 1, & \text{sonst.} \end{cases}$$

  Im ersten Fall gibt es also $x_1, \dots, x_n \in K$, so dass

  $$b = \sum_{i=1}^{n} x_i f(e_i) = f(\sum_{i=1}^{n} x_i e_i)$$

  und das heißt, dass

  $$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in L(A;b).$$

  Ist umgekehrt $x \in L(A;b)$, so ist $b = f(x) = \sum_{i=1}^{n} x_i f(e_i)$ und somit ist b eine Linearkombination der Spaltenvektoren von A.

Satz III.4.3. *Ist $f = A\cdot\colon K^n \to K^m$ mit $A \in M(m \times n, K)$ und ist $b \in K^m$, so gilt:*

(a) $\mathsf{rg}(A) = \mathsf{rg}(A|b) \Leftrightarrow L(A;b) \neq \varnothing \Leftrightarrow b \in \mathrm{Bild}(f)$.
(b) $L(A;b) = \varnothing \Leftrightarrow \mathsf{rg}(A|b) = \mathsf{rg}(A) + 1 \Leftrightarrow b \notin \mathrm{Bild}(f)$.
(c) *Ist $L(A;b) \neq \varnothing$, so ist $L(A;b)$ ein affiner Unterraum des K^n der Dimension $n - \mathsf{rg}(A)$.*

Beweis. Die Punkte (a) und (b) hatten wir vorher schon begründet. Für (c) schreiben wir $L(A;b) = v + \ker(f)$ und berechnen, dass

$$\dim_K L(A;b) = \dim_K(v + \ker(f)) = \dim_K(\ker(f)) = n - \dim_K \operatorname{Bild}(f) = n - \mathsf{rg}(A),$$

wobei wir die Dimensionsformel für lineare Abbildungen aus Satz III.1.8 benutzt haben. □

Definition III.4.4. Eine Matrix $A \in M(m \times n, K)$ ist in *Zeilenstufenform*, falls für $i = 2, \ldots, m$ gilt: Sind die ersten $(k-1)$ Einträge der $(i-1)$-ten Zeile 0, so sind die ersten k Einträge der i-ten Zeile gleich 0 (für $k \in \{1, \ldots, n\}$).

Eine Matrix in Zeilenstufenform sieht schematisch so aus:

$$\begin{pmatrix} * & * & * & * & * & * & * & * \\ & & * & * & * & * & * & * \\ & & & * & * & * & * & * \\ & & & & & * & * & * \\ & & & & & & * & * \end{pmatrix},$$

wobei $*$ für einen nicht-trivialen Eintrag steht. Da wo nichts steht, ist eine 0. Ein konkretes Beispiel über $\mathbb{R}$ ist

$$\begin{pmatrix} -1 & 0 & \sqrt{2} & \frac{1}{3} & 0 & 1 & \pi \\ 0 & 0 & \frac{2}{5} & \frac{7}{13} & 1 & \sqrt{5} & 2 \\ 0 & 0 & 0 & 4 & 2 & \frac{1}{7} & \frac{\pi}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & e \end{pmatrix}.$$

Lemma III.4.5. *Ist A in Zeilenstufenform, so ist die Lösungsmenge $L(A;b)$ des LGS $Ax = b$ genau dann leer, falls es ein $1 \leqslant i \leqslant m$ gibt mit $a_{ij} = 0$ für alle $1 \leqslant j \leqslant n$, aber $b_i \neq 0$.*

Beweis. Gibt es ein $1 \leqslant i \leqslant m$ gibt mit $a_{ij} = 0$ für alle $1 \leqslant j \leqslant n$, aber $b_i \neq 0$, so lautet die i-te Gleichung des LGS

$$0 = a_{i1}x_1 + \ldots + a_{in}x_n = b_i \neq 0$$

und das ist nicht lösbar.

Gibt es kein solches i, so kann man das LGS suksessive lösen. □

Beispiel III.4.6. Es sei $K = \mathbb{Q}$ und $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ und $b = \begin{pmatrix} 4 \\ 7 \\ 0 \end{pmatrix}$. Die Gleichung $Ax = b$ ergibt sofort $x_3 = 7$ und damit erhalten wir über die erste Gleichung

$$x_1 + 2x_2 + 21 = 4$$

also $x_1 + 2x_2 = -17$. Die Lösungsmenge ist

$$L(A;b) = \{x \in \mathbb{Q}^3, x_3 = 7, x_1 = -2x_2 - 17\}$$

und dies ist ein affiner Unterraum des $\mathbb{Q}^3$ der Dimension 1.

Wenn wir lineare Gleichungssysteme lösen möchten, dann sollten wir sie also in Zeilenstufenform überführen. Wie geht das?

Lemma III.4.7. *Ist $T \in GL_m(K)$, dann ist $L(A;b) = L(TA;Tb)$.*

Beweis. Ein $x \in K^m$ ist genau dann in $L(A;b)$ wenn $Ax = b$ gilt. Da T eine invertierbare $m \times m$-Matrix ist, gilt dies genau dann, wenn $TAx = Tb$ ist, also wenn $x \in L(TA;Tb)$ ist. □

Wir brauchen also invertierbare Matrizen, die uns helfen, ein beliebiges LGS in Zeilenstufenform zu bringen.

Definition III.4.8. Es sei K ein beliebiger Körper.

(a) Für $0 \neq \lambda \in K$ sei $\Delta^i(\lambda)$ die Matrix in $M(m \times m, K)$, die nur auf der Diagonalen von 0 verschiedene Elemente als Eintrag hat, und zwar

$$\Delta^i(\lambda)_{k\ell} = \begin{cases} 1, & k = \ell, k \neq i, \\ \lambda, & k = \ell = i, \\ 0, & \text{sonst.} \end{cases}$$

Schematisch sieht $\Delta^i(\lambda)$ wie folgt aus

$$\Delta^i(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

wobei leere Stellen wiederum für 0 stehen und λ der Eintrag an der Stelle (i,i) ist.

(b) Für $1 \leqslant i \neq j \leqslant m$ sei $\tau(i,j)$ die Matrix aus $M(m \times m, K)$ mit

$$\tau(i,j)_{kk} = \begin{cases} 1, & k \neq i, j, \\ 0, & k = i \text{ oder } k = j, \end{cases}$$

und für $k \neq \ell$:

$$\tau(i,j)_{k\ell} = \begin{cases} 1, & (k,\ell) = (i,j) \text{ oder } (k,\ell) = (j,i), \\ 0, & \text{sonst.} \end{cases}$$

Schematisch:

$$\begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix},$$

wobei die beiden Einträge außerhalb der Diagonalen mit 1 in den Koordinaten (i,j) und (j,i) stehen.

Für $\lambda \in K$ und $i \neq j$ sei

$$\delta(i,j,\lambda) = E_m + \lambda \cdot E_i^j,$$

das heißt, dass die einzigen nichttrivialen Einträge von $\delta(i,j,\lambda)$ Einsen auf der Diagonale und ein λ in Koordinate (i,j) ist.
Matrizen der Form $\Delta^i(\lambda)$, $\tau(i,j)$ und $\delta(i,j,\lambda)$ heißen *Elementarmatrizen.*

Bemerkung III.4.9. Es sei $A \in M(m \times n, K)$.

- Ist $\lambda \neq 0$, so ist $\Delta^i(\lambda) \cdot A$ die Matrix, in der die i-te Zeile von A mit λ multipliziert wurde.
- Die Matrix $\tau(i,j) \cdot A$ unterscheidet sich von A durch die Vertauschung der i-ten und j-ten Zeile.
- $\delta(i,j,\lambda) \cdot \delta(i,j,-\lambda) = E_m$ und $\delta(i,j,\lambda) \cdot A$ entsteht aus A, indem man das λ-fache der j-ten Zeile zur i-ten Zeile von A dazuaddiert.

- Matrizenumformungen wie oben heißen *elementare Zeilenumformungen.* Da Elementarmatrizen invertierbar sind, ändert sich die Lösungsmenge nicht. Geht das LGS $\tilde{A}x = \tilde{b}$ aus $Ax = b$ durch elementare Zeilenumformungen hervor, so ist $L(\tilde{A};\tilde{b}) = L(A;b)$.

Wir erhalten somit einen Algorithmus, der ein LGS in Zeilenstufenform überführt:

(a) Vertauschen Sie die Zeilen so, dass in der ersten Zeile das erste von 0 verschiedene Element nicht weiter rechts steht als bei allen anderen Zeilen.
(b) Multiplizieren Sie alle Zeilen (inklusive der ersten Zeile), bei denen der erste nicht-triviale Eintrag in der gleichen Spalte wie in der ersten Zeile steht mit geeigneten $\lambda \in K \setminus \{0\}$, so dass diese Einträge zu 1 werden.
(c) Subtrahieren Sie die erste Zeile von diesen Zeilen.
(d) Ist noch keine Zeilenstufenform erreicht, so wenden Sie (a)-(c) auf die Matrix an, die durch Streichen der ersten Zeile entsteht.

Bemerkung III.4.10. Die Normierung der ersten Einträge auf 1 ist nicht notwendig zur Lösung, aber sie macht den Vorgang im Allgemeinen übersichtlicher.

Beispiel III.4.11. Wir betrachten die Matrix $\begin{pmatrix} 0 & 1 & 1 \\ 5 & -10 & 15 \\ 1 & 4 & 2 \end{pmatrix}$.

Vertauschen der Zeilen gemäß (a) gibt $\begin{pmatrix} 5 & -10 & 15 \\ 1 & 4 & 2 \\ 0 & 1 & 1 \end{pmatrix}$.

Normierung wie in (b) liefert $\begin{pmatrix} 1 & -2 & 3 \\ 1 & 4 & 2 \\ 0 & 1 & 1 \end{pmatrix}$.

Subtraktion wie in (c) ergibt $\begin{pmatrix} 1 & -2 & 3 \\ 0 & 6 & -1 \\ 0 & 1 & 1 \end{pmatrix}$.

Wir machen jetzt weiter mit den Zeilen 2 und 3, weil noch keine Zeilenstufenform erreicht wurde, normieren die zweite Zeile wie in (b) und erhalten $\begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & -\frac{1}{6} \\ 0 & 1 & 1 \end{pmatrix}$.

Mit (c) bekommen wir die Zeilenstufenform $\begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & -\frac{1}{6} \\ 0 & 0 & \frac{7}{6} \end{pmatrix}$.

Normieren wir noch einmal nach (b) bekommen wir $\begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & -\frac{1}{6} \\ 0 & 0 & 1 \end{pmatrix}$.

Wäre ein Vektor b mit zu berücksichtigen gewesen, hätten wir alle Schritte für die erweiterte Koeffizientenmatrix $(A|b)$ durchführen müssen, um die neue erweiterte Koeffizientenmatrix $(\tilde{A}|\tilde{b})$ zu erhalten.

Wir erhalten also einen Algorithmus zur Lösung inhomogener LGS $Ax = b$. Dies ist der *Gauß-Algorithmus*:

(a) Überführen Sie die Matrix $(A|b)$ wie oben in eine Matrix in Zeilenstufenform $(\tilde{A}|\tilde{b})$.
(b) Überprüfen Sie mit Lemma III.4.5, ob die Lösungsmenge $L(A;b) = L(\tilde{A};\tilde{b})$ leer ist.
(c) Ist $L(\tilde{A};\tilde{b}) \neq \varnothing$, so lösen Sie das LGS schrittweise auf.

III.5. Koordinatentransformationen

Wir hatten im Korollar III.2.2 gesehen, dass es für einen K-Vektorraum V mit Basis $\mathcal{B} = (v_1, \ldots, v_n)$ einen Isomorphismus $\Phi_{\mathcal{B}}\colon K^n \to V$ gibt mit $\Phi_{\mathcal{B}}(e_i) = v_i$.

Zur Erinnerung: Ist $g\colon K^n \to K^m$ eine K-lineare Abbildung, so bezeichnet $M(g)$ die darstellende Matrix zu Standardbasis.

Lemma III.5.1. *Es seien V und W zwei K-Vektorräume und $f\colon V \to W$ sei K-linear. Für geordnete Basen $\mathcal{A} = (v_1, \ldots, v_n)$ von V und $\mathcal{B} = (w_1, \ldots, w_m)$ gilt für die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$:*

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = M(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}).$$

$$\begin{tikzcd}
K^n \arrow[rr, "\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}"] \arrow[d, "\Phi_{\mathcal{A}}"'] & & K^m \\
V \arrow[rr, "f"] & & W \arrow[u, "\Phi_{\mathcal{B}}"']
\end{tikzcd}$$

Beweis. Nach Definition von $M_{\mathcal{B}}^{\mathcal{A}}(f)$ ist

$$f(v_i) = \sum_{j=1}^{m} M_{\mathcal{B}}^{\mathcal{A}}(f)_{ji} w_j.$$

Aber $\Phi_{\mathcal{A}}(e_i) = v_i$ und $\Phi_{\mathcal{B}}(e_j) = w_j$, so dass wir erhalten

$$\begin{aligned}
f(v_i) &= f(\Phi_{\mathcal{A}}(e_i)) \\
&= \Phi_{\mathcal{B}}(\Phi_{\mathcal{B}}^{-1}(f(\Phi_{\mathcal{A}}(e_i)))) \\
&= \Phi_{\mathcal{B}}(\sum_{j=1}^{m} M(\Phi_{\mathcal{B}}^{-1}(f(\Phi_{\mathcal{A}}(e_i))))_{ji} e_j) \\
&= \sum_{j=1}^{m} M(\Phi_{\mathcal{B}}^{-1}(f(\Phi_{\mathcal{A}}(e_i))))_{ji} w_j,
\end{aligned}$$

wobei wir bei der letzten Umformung ausnutzen, dass $\Phi_{\mathcal{B}}$ K-linear ist. □

Damit erhalten wir eine explizite Formel für die darstellende Matrix einer Komposition linearer Abbildungen:

Satz III.5.2. *Es seien V, W und Z endlich-dimensionale K-Vektorräume mit geordneten Basen $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$. Dann gilt für alle K-linearen Abbildungen $f\colon V \to W$ und $g\colon W \to Z$:*

$$M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) = M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f), \tag{III.5.1}$$

wobei $\cdot$ das Matrizenprodukt ist.

Sie können sich das als eine Kürzungsregel merken, indem Sie bei der Komposition von Abbildung die Basis streichen, die doppelt vorkommt:

$$M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) = M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f)$$

Beweis. Wir fügen wieder einmal eine identische Abbildung ein und erhalten:

$$\begin{aligned}
M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) &= M(\Phi_{\mathcal{C}}^{-1} \circ g \circ f \circ \Phi_{\mathcal{A}}) \\
&= M(\Phi_{\mathcal{C}}^{-1} \circ g \circ \Phi_{\mathcal{B}} \circ \Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}) \\
&= M(\Phi_{\mathcal{C}}^{-1} \circ g \circ \Phi_{\mathcal{B}}) \cdot M(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}) \\
&= M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f).
\end{aligned}$$

□

Oft benutzen wir nicht die Standardbasis und wählen stattdessen eine Basis, die an die lineare Abbildung angepasst ist. Betrachten die zum Beispiel eine Spiegelung im $\mathbb{R}^2$ an einer Ursprungsgeraden, dann kann es sinnvoll sein, einen erzeugenden Vektor der Geraden als Basiselement zu betrachten. Wie verhält sich $M_{\mathcal{B}}^{\mathcal{A}}(f)$, wenn wir die Basis wechseln?

Definition III.5.3. Es seien $\mathcal{A}$ und $\mathcal{A}'$ zwei geordnete endliche Basen eines K-Vektorraums V. Dann heißt

$$T_{\mathcal{A}'}^{\mathcal{A}} := M(\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}})$$

die *Transformationsmatrix des Basiswechsels von $\mathcal{A}$ nach $\mathcal{A}'$*.

Bemerkung III.5.4.

- Sie kennen $T^{\mathcal{A}}_{\mathcal{A}'}$ schon, weil
$$T^{\mathcal{A}}_{\mathcal{A}'} = M(\Phi^{-1}_{\mathcal{A}'} \circ \Phi_{\mathcal{A}}) = M(\Phi^{-1}_{\mathcal{A}'} \circ \mathrm{id}_V \circ \Phi_{\mathcal{A}}) = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V),$$
also ist $T^{\mathcal{A}}_{\mathcal{A}'}$ nichts anderes als die darstellende Matrix der identischen Abbildung bezüglich der Basen $\mathcal{A}$ und $\mathcal{A}'$.
- Es ist
$$T^{\mathcal{A}}_{\mathcal{A}'} \cdot T^{\mathcal{A}'}_{\mathcal{A}} = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V) \cdot M^{\mathcal{A}'}_{\mathcal{A}}(\mathrm{id}_V) = M^{\mathcal{A}'}_{\mathcal{A}'}(\mathrm{id}_V \circ \mathrm{id}_V) = M^{\mathcal{A}'}_{\mathcal{A}'}(\mathrm{id}_V)$$
und $M^{\mathcal{A}'}_{\mathcal{A}'}(\mathrm{id}_V)$ ist die Einheitsmatrix E_n, wenn $\dim_K V = n$.
Genauso erhalten wir $T^{\mathcal{A}'}_{\mathcal{A}} \cdot T^{\mathcal{A}}_{\mathcal{A}'} = E_n$ und damit sind beide Matrizen $T^{\mathcal{A}'}_{\mathcal{A}}$ und $T^{\mathcal{A}}_{\mathcal{A}'}$ invertierbar, also Elemente in $GL_n(K)$.

Satz III.5.5 (Transformationssatz)**.** *Es seien V und W endlich-dimensionale K-Vektorräume, $\mathcal{A}, \mathcal{A}'$ seien geordnete Basen von V und $\mathcal{B}, \mathcal{B}'$ seien geordnete Basen von W. Es sei $f\colon V \to W$ eine K-lineare Abbildung.*

Dann gilt:
$$M^{\mathcal{A}'}_{\mathcal{B}'}(f) = T^{\mathcal{B}}_{\mathcal{B}'} \cdot M^{\mathcal{A}}_{\mathcal{B}}(f) \cdot (T^{\mathcal{A}}_{\mathcal{A}'})^{-1} = T^{\mathcal{B}}_{\mathcal{B}'} \cdot M^{\mathcal{A}}_{\mathcal{B}}(f) \cdot T^{\mathcal{A}'}_{\mathcal{A}}.$$
Insbesonder gilt: Ist $f\colon V \to V$ ein Endomorphismus, so ist
$$M_{\mathcal{A}'}(f) = T^{\mathcal{A}}_{\mathcal{A}'} \cdot M_{\mathcal{A}}(f) \cdot (T^{\mathcal{A}}_{\mathcal{A}'})^{-1} = T^{\mathcal{A}}_{\mathcal{A}'} \cdot M_{\mathcal{A}}(f) \cdot T^{\mathcal{A}'}_{\mathcal{A}}.$$

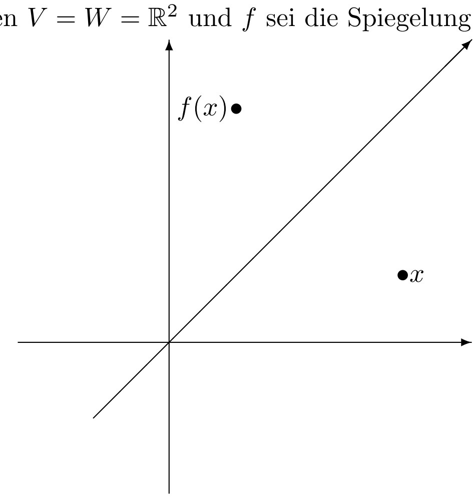
Wenn Sie Diagramme mögen, dann können Sie sich den Satz so veranschaulichen:

$$\begin{tikzcd}
K^n \arrow[rrr, "M^{\mathcal{A}}_{\mathcal{B}}(f)\cdot"] \arrow[dr, "\Phi_{\mathcal{A}}"] \arrow[dd, "T^{\mathcal{A}}_{\mathcal{A}'}\cdot"'] & & & K^m \arrow[dl, "\Phi_{\mathcal{B}}"'] \arrow[dd, "T^{\mathcal{B}}_{\mathcal{B}'}\cdot"] \\
& V \arrow[r, "f"] & W & \\
K^n \arrow[ur, "\Phi_{\mathcal{A}'}"] \arrow[rrr, "M^{\mathcal{A}'}_{\mathcal{B}'}(f)\cdot"] & & & K^m \arrow[ul, "\Phi_{\mathcal{B}'}"']
\end{tikzcd}$$

Beweis. Wir schmuggeln wiederum identische Abbildungen hinein:
$$\begin{aligned}
M^{\mathcal{A}'}_{\mathcal{B}'}(f) &= M^{\mathcal{A}'}_{\mathcal{B}'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) \\
&= M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W) \cdot M^{\mathcal{A}}_{\mathcal{B}}(f) \cdot M^{\mathcal{A}'}_{\mathcal{A}}(\mathrm{id}_V) \\
&= T^{\mathcal{B}}_{\mathcal{B}'} \cdot M^{\mathcal{A}}_{\mathcal{B}}(f) \cdot T^{\mathcal{A}'}_{\mathcal{A}} \\
&= T^{\mathcal{B}}_{\mathcal{B}'} \cdot M^{\mathcal{A}}_{\mathcal{B}}(f) \cdot (T^{\mathcal{A}}_{\mathcal{A}'})^{-1}.
\end{aligned}$$

□

Beispiel III.5.6. Wir betrachten $V = W = \mathbb{R}^2$ und f sei die Spiegelung an der Winkelhalbierenden.



Als geordnete Basis wählen wir

$$\mathcal{A} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right).$$

Dann liegt $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ auf der Spiegelachse und ist invariant unter f, wohingegen

$$f \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = - \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Damit ist die darstellende Matrix

$$M_{\mathcal{A}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Benutzen wir stattdessen die Standardbasis $S = (e_1, e_2)$, so ist die Transformationsmatrix

$$T_S^{\mathcal{A}} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

weil $\begin{pmatrix} 1 \\ 1 \end{pmatrix} = e_1 + e_2$ und $\begin{pmatrix} -1 \\ 1 \end{pmatrix} = e_2 - e_1$. Das Inverse ist

$$(T_S^{\mathcal{A}})^{-1} = T_{\mathcal{A}}^S = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Entweder berechnen Sie das Inverse oder Sie überlegen sich, dass

$$e_1 = \frac{1}{2}(\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ 1 \end{pmatrix}) \text{ und } e_2 = \frac{1}{2}(\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix}).$$

Damit erhalten wir

$$\begin{aligned} M_S(f) &= T_S^{\mathcal{A}} \cdot M_{\mathcal{A}}(f) \cdot T_{\mathcal{A}}^S \\ &= \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

Da $f(e_1) = e_2$ und $f(e_2) = e_1$, stimmt obige Rechnung mit der direkten Bestimmung von $M_S(f)$ überein.

Definition III.5.7.

(a) Zwei Matrizen $X, Y \in M(m \times n, K)$ heißen *äquivalent*, falls es ein $S \in GL_m(K)$ und ein $T \in GL_n(K)$ gibt mit

$$Y = SXT^{-1}.$$

(b) Zwei Matrizen $X, Y \in M(n \times n, K)$ heißen *ähnlich*, falls es ein $T \in GL_n(K)$ gibt mit

$$Y = TXT^{-1}.$$

Bemerkung III.5.8.

- Ähnliche Matrizen sind auch äquivalent mit $S = T$, aber die Umkehrung gilt im Allgemeinen nicht, selbst wenn die beteiligten Matrizen quadratisch sind:

  Als Beispiel betrachten wir eine Matrix $Y \neq E_m \in GL_m(K)$ und wir nehmen $X = E_m$, $S = Y$, $T = E_m$. Dann gilt

  $$SXT^{-1} = YE_mE_m = Y$$

  also ist *jedes* $Y \in GL_m(K)$ äquivalent zu E_m. Aber für alle $T \in GL_n(K)$ gilt

  $$TE_mT^{-1} = TT^{-1} = E_m$$

  und damit ist nur E_m ähnlich zu E_m.

- Sowohl *ähnlich* als auch *äquivalent* definieren eine Äquivalenzrelation auf der Menge der $n \times n$- beziehungsweise $m \times n$-Matrizen. Wir begründen das für *äquivalent*:

  Die Relation ist reflexiv, weil für $S = E_m$ und $T = E_n$ gilt, dass $SXT^{-1} = X$ ist. Die Symmetrie gilt, weil aus $Y = SXT^{-1}$ folgt, dass $S^{-1}YT = X$ gilt und $T = (T^{-1})^{-1}$.

  Ist $Y = S_1XT_1^{-1}$ und $Z = S_2YT_2^{-1}$, so folgt
$$Z = S_2YT_2^{-1} = S_2S_1XT_1^{-1}T_2^{-1} = (S_2S_1)X(T_2T_1)^{-1}.$$

Satz III.5.9.

(a) *Zwei Matrizen X und Y sind genau dann äquivalent, wenn sie die gleiche lineare Abbildung bezüglich verschiedener Basen darstellen. Das heißt: Es gibt einen n-dimensionalen K-Vektorraum V mit zwei geordneten Basen $\mathcal{A}$ und $\mathcal{A}'$ und einen m-dimensionalen K-Vektorraum W mit geordneten Basen $\mathcal{B}$ und $\mathcal{B}'$ und es gibt eine K-lineare Abbildung $f\colon V \to W$, so dass*
$$X = M_{\mathcal{B}}^{\mathcal{A}}(f), \quad Y = M_{\mathcal{B}'}^{\mathcal{A}'}(f).$$

(b) *Zwei quadratische Matrizen X und Y sind genau dann ähnlich, wenn sie den gleichen Endomorphismus bezüglich verschiedener Basen darstellen.*

Beweis. Wir zeigen (a) und überlassen die Anpassung für (b) Ihnen.

Gibt es solche V, W mit entsprechenden Basen und f, so folgt die Äquivalenz von $M_{\mathcal{B}}^{\mathcal{A}}(f)$ und $M_{\mathcal{B}'}^{\mathcal{A}'}(f)$ mit dem Transformationssatz III.5.5. Es sei umgekehrt $Y = (y_{ij})$ äquivalent zu $X = (x_{ij})$ mit $Y = SXT^{-1}$ und $X, Y \in M(m \times n, K)$. Es sei $\mathcal{A} = (v_1, \ldots, v_n)$ eine beliebige geordnete Basis des K^n und $\mathcal{B} = (w_1, \ldots, w_m)$ sei eine beliebige Basis des K^m. Wir definieren $f\colon K^n \to K^m$ durch
$$f(v_i) := \sum_{j=1}^{m} x_{ji} w_j.$$

Damit ist $M_{\mathcal{B}}^{\mathcal{A}}(f) = X$.

Da die Matrizen $S = (s_{ij})$ und $T = (t_{ij})$ invertierbar sind, gibt es eine Basis $\mathcal{A}' = (v'_1, \ldots, v'_n)$ mit
$$v_j = \sum_{i=1}^{n} t_{ij} v'_i$$
und es gibt eine geordnete Basis $\mathcal{B}' = (w'_1, \ldots, w'_m)$ mit
$$w_j = \sum_{i=1}^{m} s_{ij} w'_i.$$

Wir rechnen nach, dass $M_{\mathcal{B}'}^{\mathcal{A}'}(f) = Y$ ist: Es sei T^{-1} das Inverse von T mit Einträgen $(t^{-1})_{ij}$.
$$\begin{aligned}
f(v'_i) &= f(\sum_j ((t^{-1})_{ji} v_j) = \sum_j (t^{-1})_{ji} f(v_j) \\
&= \sum_{j,k} (t^{-1})_{ji} x_{kj} w_k \\
&= \sum_{j,k} x_{kj} (t^{-1})_{ji} w_k \\
&= \sum_{j,k,\ell} x_{kj} (t^{-1})_{ji} s_{\ell k} w'_\ell \\
&= \sum_\ell \left(\sum_{j,k} (s_{\ell k} x_{kj} (t^{-1})_{ji} \right) w'_\ell \\
&= \sum_\ell y_{\ell i} w'_\ell.
\end{aligned}$$

□

Bemerkung III.5.10. Sie zeigen in einer Übungsaufgabe, dass eine Matrix $X \in M(m \times n, K)$ vom Rang r äquivalent ist zu $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Hierbei stehen die Nullen für Nullmatrizen der passenden Größe. Damit sind die Äquivalenzklassen von Matrizen durch den Rang festgelegt. Ähnlichkeitsklassen sind dagegen komplizierter!

Wir hatten den Rang einer linearen Abbildung als die Dimension ihres Bildes definiert und uns in Bemerkung III.2.9 überlegt, dass der Rang einer linearen Abbildung $f\colon K^n \to K^m$ mit der maximalen Anzahl linear unabhängiger Spalten der darstellenden Matrix übereinstimmt. Daher ist die folgende Definition des Spaltenrangs konsistent mit unserer bisherigen Definition des Rangs.

Definition III.5.11. Es sei $X \in M(m \times n, K)$. Die maximale Anzahl linear unabhängiger Spalten von X heißt der *Spaltenrang von* X, $\mathsf{rg}(X)$, und die maximale Anzahl linear unabhängiger Zeilen von X heißt der *Zeilenrang von* X, $\widetilde{\mathsf{rg}}(X)$.

Satz III.5.12. *Sind V, W endlich-dimensionale K-Vektorräume und ist $f\colon V \to W$ eine K-lineare Abbildung, so ist für jede geordnete Basis $\mathcal{A}$ von V und $\mathcal{B}$ von W:*

$$\mathsf{rg} M^{\mathcal{A}}_{\mathcal{B}}(f) = \mathsf{rg}(f).$$

Beweis. Es seien $\Phi_{\mathcal{A}}\colon K^n \to V$ und $\Phi_{\mathcal{B}}\colon K^m \to W$ die Isomorphismen, die durch die geordenten Basen $\mathcal{A}$ und $\mathcal{B}$ definiert sind. Dann gilt:

$$\begin{aligned} \mathsf{rg} M^{\mathcal{A}}_{\mathcal{B}}(f) &= \mathsf{rg}(M(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}})) \\ &= \mathsf{rg}(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}) \\ &= \dim_K \mathrm{Bild}(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}) \end{aligned}$$

und da $\Phi_{\mathcal{A}}$ und $\Phi_{\mathcal{B}}$ Isomorphismen sind, ist dies gleich $\dim_K \mathrm{Bild}(f) = \mathsf{rg}(f)$. □

Lemma III.5.13. *Ist $X \in M(m \times n, K)$ und $S \in GL_m(K)$, so ist*

(a) $(S^{-1})^t = (S^t)^{-1}$
(b) $\mathsf{rg}(X) = \widetilde{\mathsf{rg}}(X^t)$ *und* $\widetilde{\mathsf{rg}}(X) = \mathsf{rg}(X^t)$.

Beweis. Zu (a) erinnern wir uns an Lemma III.2.14 und rechnen nach, dass

$$(S^{-1})^t \cdot S^t = (S \cdot S^{-1})^t = E_m^t = E_m.$$

Teil (b) ist klar, weil die Spalten von X genau die Zeilen von X^t sind und die Zeilen von X^t genau die Spalten von X sind. □

Lemma III.5.14. *Sind X und Y äquivalent, so gilt:*

$$\mathsf{rg}(X) = \mathsf{rg}(Y) \text{ und } \widetilde{\mathsf{rg}}(X) = \widetilde{\mathsf{rg}}(Y).$$

Beweis. Wir wissen mit Satz III.5.9 dass es ein $f \in \mathsf{Hom}_K(K^n, K^m)$ und geordnete Basen $\mathcal{A}, \mathcal{A}'$ von K^n und $\mathcal{B}, \mathcal{B}'$ von K^m gibt, so dass

$$X = M^{\mathcal{A}}_{\mathcal{B}}(f) \text{ und } Y = M^{\mathcal{A}'}_{\mathcal{B}'}(f).$$

Mit Satz III.5.12 folgt damit

$$\mathsf{rg}(X) = \mathsf{rg}(f) = \mathsf{rg}(Y).$$

Das zeigt (a). Für (b) schreiben wir Y als SXT^{-1} mit invertierbaren Matrizen S und T. Damit ist das Transponierte von Y:

$$Y^t = (SXT^{-1})^t = (T^{-1})^t X^t S^t = (T^t)^{-1} X^t S^t$$

und damit sind Y^t und X^t ebenfalls äquivalent und wir erhalten

$$\widetilde{\mathsf{rg}}(X) = \mathsf{rg}(X^t) = \mathsf{rg}(Y^t) = \widetilde{\mathsf{rg}}(Y).$$

□

Satz III.5.15. *Es sei $X \in M(m \times n, K)$. Dann stimmt der Spaltenrang von X mit dem Zeilenrang von X über.*

Beweis. Ist $\mathsf{rg}(X) = r$, so ist X äquivalent zu $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$, also gibt es invertierbare Matrizen S und T, so dass $X = S \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} T^{-1}$. Damit ist dann die Transponierte von X:

$$X^t = \left(S \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} T^{-1} \right) = (T^t)^{-1} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} S^t$$

und somit ist auch X^t äquivalent zu $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ und daher gilt

$$\widetilde{\mathsf{rg}}(X) = \mathsf{rg}(X^t) = r = \mathsf{rg}(X).$$

□

Beispiel III.5.16. Hat eine Matrix weniger Zeilen als Spalten, dann kann man den Zeilenrang oft einfacher bestimmen als den Spaltenrang. So ist zum Beispiel der Zeilenrang von

$$\begin{pmatrix} 10 & 5 & 15 \\ 2 & 1 & 3 \end{pmatrix}$$

sichtbar gleich 1.

III.6. Quotientenvektorräume

Es sei V ein K-Vektorraum und $U \subset V$ sei ein Untervektorraum. Wir setzen für $v, w \in V$:

$$v \sim w :\Leftrightarrow v - w \in U. \tag{III.6.1}$$

Lemma III.6.1. *Dies definiert eine Äquivalenzrelation.*

Beweis. Da $0_V \in U$ ist, ist $v \sim v$, also ist die Relation reflexiv. Ist $v \sim w$, so ist $v - w \in U$. Aber damit ist auch $(-1)(v - w) = w - v \in U$, also ist die Relation symmetrisch. Sind $v - w \in U$ und $w - z \in U$, so ist auch die Summe in U, aber die ist

$$v - w + w - z = v - z$$

also ist die Relation transitiv. □

Wir notieren die Äquivalenzklasse eines Elementes $v \in V$ als $[v]$.

Satz III.6.2. *Ist V ein K-Vektorraum und U ein Untervektorraum von V, so bildet die Quotientenmenge V/U einen K-Vektorraum mit den Verknüpfungen*

$$[v] + [w] := [v + w]$$
$$\lambda[v] := [\lambda v]$$

für $v, w \in V$ und $\lambda \in K$.

Beweis. Wir müssen zeigen, dass die Verknüpfungen wohldefiniert sind. Sind also $v_1 \sim v_2$, $w_1 \sim w_2$, so sind $v_1 - v_2$ und $w_1 - w_2$ aus U. Damit ist dann aber auch

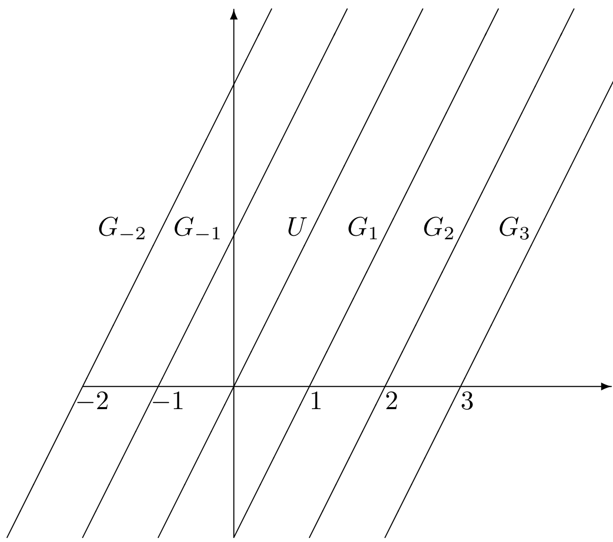
$$(v_1 + w_1) - (v_2 + w_2) = (v_1 - v_2) + (w_1 - w_2) \in U$$

und $v_1 + w_1 \sim v_2 + w_2$. Das zeigt die Wohldefiniertheit das Addition. Ist $v \sim w$ und ist $\lambda \in K$, so ist $\lambda(v - w) \in U$ und somit gilt $\lambda v \sim \lambda w$.

Die Vektorraumaxiome übertragen sich von V auf V/U. Für den Nullvektor $0_{V/U}$ gilt: $0_{V/U} = [u]$, wobei u ein beliebiger Vektor aus U ist. □

Definition III.6.3. Der K-Vektorraum V/U heißt der *Quotientenvektorraum von V nach U*. Die Projektion $\pi \colon V \to V/U$ heißt die *kanonische Projektion.*

Beispiel III.6.4. Ist $V = \mathbb{R}^2$ und ist $U = \mathsf{Span}_K \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$, dann entsprechen die Äquivalenzklassen affine Geraden, die parallel sind zu U: Die Gerade U entspricht der Äquivalenzklasse von des Nullpunkts. Die affine Gerade G_1 mit Richtungsvektor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und Schnittpunkt $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ mit der x-Achse enthält alle Vektoren, die in der Äquivalenzklasse des Vektors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ liegen. Die Äquivalenzklasse des Vektors $\begin{pmatrix} 3,5 \\ 1 \end{pmatrix}$ ist die gleiche wie die Äquivalenzklasse des Vektors $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$ und entspricht der affinen Gerade G_3 mit Richtungsvektor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und Schnittpunkt $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$ mit der x-Achse.



Bemerkung III.6.5. Die kanonische Projektion ist K-linear: Für $\lambda, \mu \in K$ und $v, w \in V$ gilt

$$\pi(\lambda v + \mu w) = [\lambda v + \mu w] = \lambda[v] + \mu[w] = \lambda\pi(v) + \mu\pi(w).$$

Der Kern von π ist genau U: Ist $\pi(v) = 0_{V/U}$, so ist $[v] = 0_{V/U}$ und dies ist genau dann der Fall, wenn $v \in U$ ist.

Wir können den Quotientenvektorraum dazu benutzen, um jede beliebige K-lineare Abbildung durch einen Isomorphismus zu ersetzen:

Satz III.6.6. *Es seien V, W zwei K-Vektorräume und $f\colon V \to W$ sei eine K-lineare Abbildung. Dann existiert ein eindeutiger Isomorphismus $\overline{f}\colon V/\ker(f) \to \mathrm{Bild}(f)$, so dass $f = i \circ \overline{f} \circ \pi$, wobei $i\colon \mathrm{Bild}(f) \hookrightarrow W$ die Inklusion ist.*

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[d, "\pi"'] & & W \\ V/\ker(f) \arrow[rr, "\overline{f}"] & & \mathrm{Bild}(f) \arrow[u, "i"'] \end{tikzcd}$$

Beweis. Falls $\overline{f}$ existiert, ist es eindeutig festgelegt, weil für alle $v \in V$ gelten muss, dass

$$\overline{f}[v] = \overline{f}(\pi(v)) = f(v).$$

Wir haben also gar keine Wahl und müssen es mit der Definition $\overline{f}[v] := f(v)$ versuchen.

Ist $[v_1] = [v_2]$, so ist $v_1 - v_2 \in \ker(f)$ nach Definition und damit gilt

$$\overline{f}[v_1] - \overline{f}[v_2] = f(v_1) - f(v_2) = f(v_1 - v_2) = 0_W$$

also $\overline{f}[v_1] = \overline{f}[v_2]$. Damit ist die Abbildung wohldefiniert und da f K-linear war, ist $\overline{f}$ auch K-linear.

Da das Bild von f nach Konstruktion gleich dem Bild von $\overline{f}$ ist, ist $\overline{f}$ surjektiv als Abbildung $\overline{f}\colon V/\ker(f) \to \operatorname{Bild}(f)$.

Wir haben $\overline{f}$ so definiert, dass es injektiv ist: Ist $\overline{f}[v] = f(v) = 0_W$, so ist $v \in \ker(f)$ und damit ist $[v] = 0_{V/\ker(f)}$. $\square$

Satz III.6.7. *Ist V ein endlich-dimensionaler K-Vektorraum und ist $U \subset V$ ein Untervektorraum, so gilt*

$$\dim_K(V/U) = \dim_K V - \dim_K U.$$

Beweis. Wir betrachten die kanonische Projektion $\pi\colon V \to V/U$ und erhalten mit der Dimensionsformel für K-lineare Abbildungen aus Satz III.1.8

$$\dim_K V = \dim_K \operatorname{Bild}(\pi) + \dim_K \ker(\pi).$$

Der Kern von π ist genau U und da π surjektiv ist, ist $V/U = \operatorname{Bild}(\pi)$. $\square$

Bemerkung III.6.8. Sind V, W endlich-dimensionale K-Vektorräume und ist $f\colon V \to W$ K-linear, so können wir eine geordnete Basis $(v_1, \ldots, v_\ell)$ des Kerns von f zu einer Basis $\mathcal{A} = (v_1, \ldots, v_n)$ von V ergänzen. Es sei $\mathcal{A}' = (v_{\ell+1}, \ldots, v_n)$. Es ist nicht schwierig zu sehen, dass dann $\mathcal{B}' = (f(v_{\ell+1}), \ldots, f(v_n))$ eine Basis des Bildes von f ist, die wir ebenfalls zu einer Basis $\mathcal{B} = (f(v_{\ell+1}), \ldots, f(v_n), w_1, \ldots, w_{m-r})$ von W ergänzen. Die Abbildung $\overline{f}\colon V/\ker(f) \to \operatorname{Bild}(f)$ ist dann festgelegt durch

$$\overline{f}[v_i] = f(v_i), \quad \ell + 1 \leqslant i \leqslant n.$$

Die Matrixdarstellung von f ist $M^{\mathcal{A}}_{\mathcal{B}}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ und die von $\overline{f}$ ist $M^{\mathcal{A}'}_{\mathcal{B}'}(\overline{f}) = E_r$.

Satz III.6.9 (Universelle Eigenschaft des Quotientenvektorraums)**.** *Es sei V ein K-Vektorraum, $U \subset V$ sei ein Untervektorraum und $\pi\colon V \to V/U$ sei die kanonische Projektion.*

(a) *Für alle K-Vektorräume W und alle K-linearen Abbildungen $f\colon V \to W$ mit $f|_U = 0$ gibt es eine eindeutig bestimmte lineare Abbildung $\overline{f}\colon V/U \to W$ mit $\overline{f} \circ \pi = f$:*

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "\pi"'] & & W \\ & V/U \arrow[ur, dotted, "\exists!\overline{f}"'] & \end{tikzcd}$$

(b) *Ist Q ein K-Vektorraum und ist $\hat{\pi}\colon V \to Q$ eine K-lineare Abbildung, so dass gilt*
- $\hat{\pi}|_U = 0$.
- *Zu jeder linearen Abbildung $f\colon V \to W$ mit $f|_U = 0$ gibt es eine eindeutig bestimmte lineare Abbildung $\widetilde{f}\colon Q \to W$ mit $\widetilde{f} \circ \pi = f$.*

Dann gibt es einen eindeutig bestimmten Isomorphismus $\varphi\colon Q \to V/U$, so dass $\varphi \circ \hat{\pi} = \pi$:

$$\begin{tikzcd} & V \arrow[dl, dotted, "\hat{\pi}"'] \arrow[dr, "\pi"] & \\ Q \arrow[rr, "\exists!\varphi"] & & V/U \end{tikzcd}$$

Beweis. Zu (a): Für ein solches $\overline{f}$ gilt wiederum $\overline{f}[v] = \overline{f} \circ \pi(v) = f(v)$ und damit ist $\overline{f}$ eindeutig bestimmt. Da $f(u) = 0_W$ ist für alle $u \in U$, ist $f(v + u) = f(v)$ für alle $v \in V$ und $\overline{f}$ ist wohldefiniert. Die Linearität von f überträgt sich wieder von f auf $\overline{f}$.

Zu (b): Wir wenden (a) auf $\hat{\pi}$ an und erhalten eine eindeutig bestimmte K-lineare Abbildung $\overline{\hat{\pi}}\colon V/U \to Q$ mit $\overline{\hat{\pi}} \circ \pi = \hat{\pi}$. Ebenso gibt es wegen der Eigenschaft von Q zu der Abbildung $\pi\colon V \to V/U$ ein eindeutig bestimmtes $\widetilde{\pi}\colon Q \to V/U$ mit

$$\widetilde{\pi} \circ \hat{\pi} = \pi.$$

Damit erhalten wir insgesamt, dass

$$\overline{\hat{\pi}} \circ \widetilde{\pi} \circ \hat{\pi} = \overline{\hat{\pi}} \circ \pi = \hat{\pi}.$$

Es gilt aber auch $\mathrm{id}_Q \circ \hat{\pi} = \hat{\pi}$. Aus der Eindeutigkeit der Abbildung folgt, dass $\overline{\hat{\pi}} \circ \widetilde{\pi} = \mathrm{id}_Q$ ist. Analog beweist man, dass $\widetilde{\pi} \circ \overline{\hat{\pi}} = \mathrm{id}_{V/U}$. Damit ist $\varphi := \widetilde{\pi}$ der gesuchte eindeutige Isomorphismus. □

Bemerkung III.6.10. Man benutzt den Term *universelle Eigenschaft*, um zu signalisieren, dass etwas durch seine Eigenschaft bis auf einen eindeutigen Isomorphismus bestimmt ist.

Haben wir ein $f\colon V \to W$ mit $f|_U = 0$ und erhalten wir $f = \overline{f} \circ \pi$, so sagt man auch, dass f über V/U faktorisiert.

III.7. Produkte und äußere direkte Summen

Definition III.7.1. Es sei $(V_i)_{i\in I}$ eine Familie von K-Vektorräumen. Dann ist

$$\prod_{i\in I} V_i := \{(v_i)_{i\in I}, v_i \in V_i\}$$

das *Produkt der Vektorräume* V_i und

$$\bigoplus_{i\in I} V_i := \{(v_i)_{i\in I}, v_i \in V_i, v_i = 0 \text{ für fast alle } i \in I\}$$

ist die *äußere direkte Summe der* V_i.

Bemerkung III.7.2.

- Ist $|I| < \infty$, so ist $\bigoplus_{i\in I} V_i = \prod_{i\in I} V_i$. Aber zum Beispiel ist
$$\bigoplus_{n\in\mathbb{N}_0} \mathbb{R} \neq \prod_{n\in\mathbb{N}_0} \mathbb{R} :$$
Die konstante Einsfolge $(1, 1, \ldots)$ ist ein Element von $\prod_{n\in\mathbb{N}_0} \mathbb{R}$, liegt aber nicht in $\bigoplus_{n\in\mathbb{N}_0} \mathbb{R}$, weil unendlich viele Koordinaten nicht-trivial sind.
- Sowohl $\bigoplus_{i\in I} V_i$ als auch $\prod_{i\in I} V_i$ sind K-Vektorräume, wobei die Vektorraumstruktur komponentenweise definiert ist:
$$(v_i)_{i\in I} + (w_i)_{i\in I} := (v_i + w_i)_{i\in I}, \quad \lambda \cdot (v_i)_{i\in I} := (\lambda \cdot v_i)_{i\in I}.$$
- Ist für alle $i \in I$ die Dimension $\dim_K V_i < \infty$, so gilt:
$$\dim_K \left(\bigoplus_{i\in I} V_i \right) = \sum_{i\in I} \dim_K V_i.$$
Die Gleichheit gilt natürlich auch, falls ein V_i unendlich dimensional ist, aber dann ist sie nicht sehr hilfreich.
- Es gilt immer, dass
$$\bigoplus_{i\in I} V_i \subset \prod_{i\in I} V_i$$
ein Untervektorraum ist.

Lineare Abbildungen aus äußeren direkten Summen sind einfach zu beschreiben, ebenso lineare Abbildungen in Produkte. Wir betrachten die Monomorphismen für $j \in I$:

$$\iota_j\colon V_j \to \bigoplus_{i\in I} V_i, \quad (\iota_j(v_j))_i = \begin{cases} v_j, & i = j, \\ 0. & i \neq j. \end{cases}$$

Die Abbildung bildet also ein v_j auf die Folge ab, die nur an der Stelle j v_j ist und ansonsten 0.

Satz III.7.3. *Ist W ein K-Vektorraum und ist $(V_i)_{i\in I}$ eine Familie von K-Vektorräumen, so ist die Abbildung*

$$\mathsf{Hom}_K(\bigoplus_{i\in I} V_i, W) \to \prod_{i\in I} \mathsf{Hom}_K(V_i, W), \quad f \mapsto (f \circ \iota_i)_{i\in I}$$

ein Isomorphismus von K-Vektorräumen.

Das heißt: Ist für jedes $j \in I$ eine lineare Abbildung $g_j\colon V_j \to W$ gegeben, so gibt es ein eindeutiges lineares $g\colon \bigoplus_{i\in I} V_i \to W$ mit $g \circ \iota_i = g_i$ für alle $i \in I$.

Beweis. Ist $v \in \bigoplus_{i \in I} V_i$, so ist $v = \sum_{j \in I} \iota_j(v_j)$, wobei $v_j \in V_j$ und $v_j = 0$ für fast alle $j \in I$.

Ist $f \circ \iota_j = 0$ für alle $j \in I$, so ist

$$f(v) = f(\sum_{j \in I} \iota_j(v_j)) = \sum_{j \in I} f(\iota_j(v_j)) = 0$$

und damit ist f die konstante Nullabbildung. Also ist die Abbildung injektiv.

Ist $g_j : V_j \to W$ gegeben für alle $j \in I$, so setzen wir

$$g(\sum_{j \in I} \iota_j(v_j)) := \sum_{j \in I} g_j(v_j).$$

Die Abbildung g ist dann K-linear und $g \circ \iota_j = g_j$ für alle $j \in I$. Also ist die Abbildung auch surjektiv. □

Satz III.7.4 (Universelle Eigenschaft der direkten Summe)**.** *Ist $(V_i)_{i \in I}$ eine Familie von K-Vektorräumen und ist W ein K-Vektorraum, für den es eine Familie K-linearer Abbildungen $h_j : V_j \to W$ gibt mit der Eigenschaft: Für alle K-Vektorräume Z ist die K-lineare Abbildung*

$$\mathsf{Hom}_K(W, Z) \to \prod_{i \in I} \mathsf{Hom}_K(V_i, Z), \quad f \mapsto (f \circ h_i)_{i \in I}$$

ein Isomorphismus.

Dann gibt es einen eindeutigen Isomorphismus $\varphi : \bigoplus_{i \in I} V_i \to W$ mit $\varphi \circ \iota_j = h_j$ für alle $j \in I$.

Beweis. Wegen der Eigenschaft der äußeren direkten Summe gibt es ein eindeutiges $h : \bigoplus_{i \in I} V_i \to W$ mit $h \circ \iota_j = h_j$ für alle $j \in I$.

Umgekehrt gibt es wegen der geforderten Eigenschaft von W ein eindeutiges $\iota : W \to \bigoplus_{i \in I} V_i$ mit $\iota \circ h_j = \iota_j$ für alle $j \in J$.

Wir betrachten $\iota \circ h$ und $h \circ \iota$.

Die Verknüpfung $\iota \circ h$ erfüllt

$$\iota \circ h \circ \iota_j = \iota \circ h_j = \iota_j$$

aber ebenso gilt natürlich, dass $\mathrm{id}_{\bigoplus_{i \in I} V_i} \circ \iota_j = \iota_j$. Wegen der Eindeutigkeit der Abbildungen folgt

$$\iota \circ h = \mathrm{id}_{\bigoplus_{i \in I} V_i}.$$

Für $h \circ \iota$ erhalten wir in analoger Weise

$$h \circ \iota \circ h_j = h \circ \iota_j = h_j$$

und da ebenfalls $\mathrm{id}_W \circ h_j = h_j$ gilt, impliziert die Eindeutigkeit wiederum $h \circ \iota = \mathrm{id}_W$.

Damit ist $\varphi = h$ der gesuchte Isomorphismus. □

Korollar III.7.5 (Äußere versus innere direkte Summe)**.** *Ist V ein K-Vektorraum und sind $U_1, U_2 \subset V$ Untervektorräume, so dass $V = U_1 \oplus U_2$. Dann ist V isomorph zur äußeren direkten Summe von U_1 und U_2.*

Das Symbol $\oplus$ ist also nicht mehrdeutig. Die direkte Summe von Untervektorräumen stimmt überein mit der äußeren direkten Summe. Da wir hier nur zwei Komponenten haben (also $|I| = 2$), ist $U_1 \oplus U_2 = U_1 \times U_2$.

Beweis. Wir zeigen, dass V die universelle Eigenschaft der äußeren direkten Summe von U_1 und U_2 hat. Es seien $h_1 : U_1 \to V$ und $h_2 : U_2 \to V$ die Inklusionen der Untervektorräume $U_1 \subset V$, $U_2 \subset V$. Jedes $v \in V$ läßt sich nach Lemma II.7.5 eindeutig schreiben als $v = u_1 + u_2$ mit $u_i \in U_i$. Ist W ein beliebiger K-Vektorraum und sind $h_1 : U_1 \to W$ und $h_2 : U_2 \to W$ K-linear, so setzen wir

$$h(v) := h_1(u_1) + h_2(u_2).$$

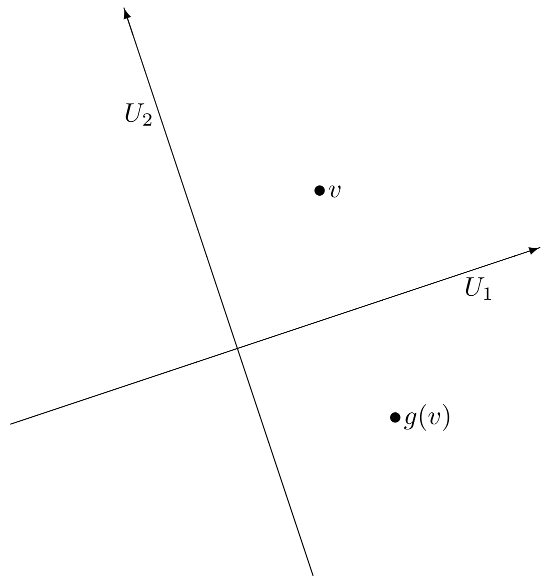
Dieses h ist wohldefiniert, K-linear und erfüllt die Eigenschaften wie in Satz III.7.4. Damit ist V isomorph zur äußeren direkten Summe von U_1 und U_2. □

Beispiel III.7.6. Betrachten wir $V = \mathbb{R}^2$ und sind $U_1 \neq U_2 \subset \mathbb{R}^2$ Untervektorräume jeweils der Dimension 1, dann ist $\mathbb{R}^2 = U_1 \oplus U_2$. Wir können diese Zerlegung benutzen, um lineare Abbildungen auf $\mathbb{R}^2$ zu konstruieren, indem wir sie auf U_1 und U_2 festlegen. Sind $\iota_1 : U_1 \to \mathbb{R}^2$ und $\iota_2 : U_2 \to \mathbb{R}^2$ die Inklusionen, dann können wir zum Beispiel $g_j : U_j \to \mathbb{R}^2$ definieren durch

$$g_1 = \iota_1, \quad g_2 = -\iota_2.$$

Wir erhalten eine eindeutig bestimmte lineare Abbildung $g\colon \mathbb{R}^2 = U_1 \oplus U_2 \to \mathbb{R}^2$ mit $g|_{U_1} = \iota_1$ und $g|_{U_2} = -\iota_2$.

Steht die Gerade U_1 zum Beispiel senkrecht zu U_2, so beschreibt g genau die Spiegelung an U_1 im $\mathbb{R}^2$.



KAPITEL IV

Determinanten

IV.1. Das Vektorprodukt des $\mathbb{R}^3$

Definition IV.1.1. Die Abbildung

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$
$$\left(\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right) = (v, w) \mapsto v \times w = \begin{pmatrix} v_2 w_w - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$

heißt das *Vektorprodukt oder das Kreuzprodukt auf dem* $\mathbb{R}^3$.

Beispiele IV.1.2. Die Vektorprodukte der Standardbasisvektoren sind einfach zu berechnen:

$$e_1 \times e_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

und

$$e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2.$$

Aber Sie erhalten auch:

$$e_2 \times e_1 = -e_3, \quad e_3 \times e_2 = -e_1, \quad e_1 \times e_3 = -e_2 \text{ und } e_1 \times e_1 = e_2 \times e_2 = e_3 \times e_3 = 0.$$

Zusammen mit dem Skalarprodukt im $\mathbb{R}^3$, welches gegeben ist durch

$$\langle -, - \rangle \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}, \quad \left\langle \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right\rangle = v_1 w_1 + v_2 w_2 + v_3 w_3$$

erhalten wir folgende Rechenregeln:

Lemma IV.1.3. *Für alle $u, v, w, v', w' \in \mathbb{R}^3$ und $\lambda, \mu \in \mathbb{R}$ gilt:*

(a) *Das Vektorprodukt ist bilinear:*

$$(\lambda v + \mu v') \times w = \lambda(v \times w) + \mu(v' \times w),$$
$$v \times (\lambda w + \mu w') = \lambda(v \times w) + \mu(v \times w').$$

(b) *Das Vektorprodukt ist antisymmetrisch:*

$$v \times w = -w \times v.$$

(c) *Es gilt die Graßmann-Identität:*

$$u \times (v \times w) = \langle u, w \rangle \cdot v - \langle u, v \rangle \cdot w.$$

(d) *Es gilt die Jacobi-Identität:*

$$u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0.$$

(e) $\langle u \times v, w \rangle = \langle u, v \times w \rangle$.

Beweis. (a), (b) und (e) beweisen wir hier nicht, (c) ist eine Übungsaufgabe. Wir leiten (d) aus (c) her:

$$\begin{aligned}
&u \times (v \times w) + v \times (w \times u) + w \times (u \times v) \\
=&\langle u, w\rangle \cdot v - \langle u, v\rangle \cdot w + \langle v, u\rangle \cdot w - \langle v, w\rangle \cdot u + \langle w, v\rangle \cdot u - \langle w, u\rangle \cdot v \\
=&(\langle u, w\rangle - \langle w, u\rangle) \cdot v + (\langle v, u\rangle - \langle u, v\rangle) \cdot w + (\langle w, v\rangle - \langle v, w\rangle) \cdot u \\
=&0.
\end{aligned}$$

□

Bemerkung IV.1.4. Die Graßmann-Identität ist in der Physik auch als *bac-cab-Regel* bekannt, weil die rechte Seite sich so liest, wenn man die Vektoren $\vec{a}, \vec{b}, \vec{c}$ nennt und die Skalare von rechts an die Vektoren heranmultipliziert:

$$\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}\langle \vec{a}, \vec{c}\rangle - \vec{c}\langle \vec{a}, \vec{b}\rangle.$$

Über das Vektorprodukt kann man die Cauchy-Schwarzsche Ungleichung zu einer Gleichung machen, indem man konkret den Fehlerterm angibt:

Satz IV.1.5. *Für alle $v, w \in \mathbb{R}^3$ gilt:*

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

Beweis. Nach Definition ist $||v \times w||^2 = \langle v \times w, v \times w\rangle$ und mit Lemma IV.1.3 (e) können wir dies umformen:

$$\langle v \times w, v \times w\rangle = \langle v, w \times (v \times w)\rangle.$$

Mit der Graßmann-Identität erhalten wir daraus

$$\begin{aligned}
\langle v, w \times (v \times w)\rangle &= \langle v, \langle w, w\rangle \cdot v - \langle w, v\rangle \cdot w\rangle \\
&= \langle w, w\rangle\langle v, v\rangle - \langle w, v\rangle\langle v, w\rangle \\
&= ||w||^2||v||^2 - \langle w, v\rangle^2.
\end{aligned}$$

□

Bemerkung IV.1.6. Wegen der Antisymmetrie gilt für alle $v \in \mathbb{R}^3$, dass

$$v \times v = -v \times v$$

und damit muss $v \times v = 0$ sein. Damit gilt aber ebenfalls

$$\langle v, v \times w\rangle = \langle v \times v, w\rangle = 0$$

also steht $v \times w$ immer senkrecht auf v. Genauso zeigen Sie, dass $v \times w$ senkrecht auf w steht. Wir benutzen die Notation $x \perp y$ wenn x senkrecht auf y steht:

$$v \perp v \times w \perp w.$$

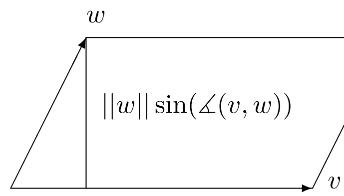
Welche Länge hat $v \times w$? Wir erhalten mit Satz IV.1.5, dass

$$||v \times w||^2 = ||v||^2||w||^2 - \langle v, w\rangle^2$$

ist. Wir formen dies weiter um zu

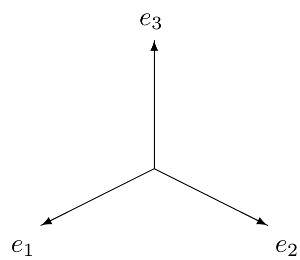
$$\begin{aligned}
||v \times w||^2 &= ||v||^2||w||^2 - \langle v, w\rangle^2 \\
&= ||v||^2||w||^2(1 - \cos^2(\measuredangle(v, w))) \\
&= ||v||^2||w||^2(\sin^2(\measuredangle(v, w))).
\end{aligned}$$

Damit ist $||v \times w||$ der Flächeninhalt des von v und w aufgespannten Parallelograms.



Man kann sich merken, in welche Richtung $v \times w$ zeigt, indem man die rechte-Hand-Regel benutzt: Entspricht v dem Daumen und w dem Zeigefinger der rechten Hand, dann entspricht $v \times w$ dem Mittelfinger.

Wem das zu kompliziert ist, der kann sich auch lediglich den Fall $e_1 \times e_2 = e_3$ merken:



Definition IV.1.7. Es sei (x, y, z) eine geordnete Basis des $\mathbb{R}^3$. Für ein $p \in \mathbb{R}^3$ heißt

$$P := \{p + \lambda x + \mu y + \nu z, 0 \leqslant \lambda, \mu, \nu \leqslant 1\}$$

das *von* (x, y, z) *aufgespannte Parallelotop (oder Spat) an* p.

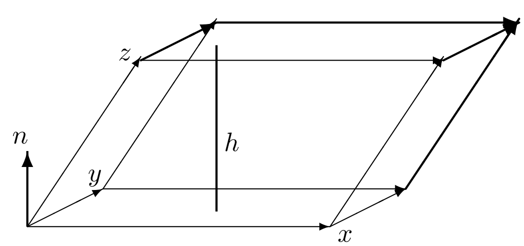
Satz IV.1.8. *Das Volumen des von* (x, y, z) *aufgespannten Parallelotops ist*

$$\mathsf{Vol}(P) = |\langle x \times y, z \rangle|.$$

Das Vektorprodukt wird deshalb auch manchmal *Spatprodukt* genannt.

Beweis. Das Volumen von P ist natürlich unabhängig von p. Wir können also ohne Einschränkung annehmen, dass $p = 0$ ist. Das Volumen errechnet sich aus dem Produkt aus Grundfläche und Höhe. Die Grundfläche ist genau das Parallelogram, welches von x und y aufgespannt wird. Für die Fläche kennen wir den Flächeninhalt und der ist $||x \times y||$.

Wir betrachten $n := \frac{x \times y}{||x \times y||}$. Dieser Vektor hat Länge 1 und steht senkrecht auf x und y.



Damit ist die Höhe des Parallelotops $h = |\langle n, z \rangle|$ und wir erhalten insgesamt für das Volumen von P:

$$\mathsf{Vol}(P) = ||x \times y|| \cdot h = ||x \times y|| \cdot |\langle n, z \rangle| = ||x \times y|| \cdot \frac{|\langle x \times y, z \rangle|}{||x \times y||} = |\langle x \times y, z \rangle|.$$

□

IV.2. Die Determinantenabbildung

Definition IV.2.1. Es sei K ein beliebiger Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\det\colon M(n \times n, K) \to K, \quad A \mapsto \det(A),$$

heißt eine *Determinantenabbildung*, falls gilt:

(D1) Die Abbildung det ist linear in jeder Zeile, das heißt

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & \dots & \lambda a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \lambda \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

und

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + a'_{i1} & \dots & a_{in} + a'_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} + \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a'_{i1} & \dots & a'_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

(D2) Die Abbildung det ist *alternierend*, das heißt, dass $\det(A) = 0$, falls zwei Zeilen von A übereinstimmen.

(D3) Die Abbildung det ist normiert durch $\det(E_n) = 1$.

Satz IV.2.2. *Ist $n \in \mathbb{N}$, K ein Körper und ist $\det\colon M(n \times n, K) \to K$ eine Determinantenabbildung, so gilt für alle $A, B \in M(n \times n, K)$ und alle $\lambda \in K$:*

(a) $\det(\lambda A) = \lambda^n \det(A)$.

(b) *Ist eine Zeile von A gleich 0, so ist $\det(A) = 0$.*

(c) *Entsteht B aus A durch das Vertauschen zweier Zeilen, so ist $\det(B) = -\det(A)$.*

(d) *Entsteht B aus A durch die Addition eines Vielfachen einer Zeile von A, so ist $\det(B) = \det(A)$.*

(e) *Ist A eine obere Dreiecksmatrix, also ist $A = (a_{ij})$ und $a_{ij} = 0$ für alle $i > j$, so ist $\det(A) = \prod_{i=1}^n a_{ii}$.*

Eine obere Dreiecksmatrix ist also von der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

Unterhalb der Diagonalen stehen also nur Nullen.

Beweis. Zu (a): Jede der n Zeilen von A ist mit λ multipliziert und daher erhalten wir mit (D1), dass $\det(\lambda A) = \lambda^n \det(A)$ ist.

Zu (b): Ist die i-te Zeile von A null, so wählen Sie $a_{i1}, \dots, a_{in}$ beliebig aus K (zum Beispiel $a_{ij} = 1$) für $1 \leqslant j \leqslant n$. Dann ist

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ 0 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ 0 \cdot a_{i1} & \dots & 0 \cdot a_{in} \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = 0 \cdot \det(A) = 0.$$

Zu (c): Es sei $a_i = (a_{i1}, \dots, a_{in})$ die i-te Zeile von A. Dann können wir A schreiben als

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Es sei B die Matrix, die aus A entsteht, indem man Zeile i und Zeile j vertauscht für $i < j$. Nach (D2) gilt

$$0 = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i + a_j \\ a_{i+1} \\ \vdots \\ a_{j-1} \\ a_i + a_j \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix}$$

und nach (D1) und (D2) können wir diese Determinante umformen zu

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_j \\ a_{i+1} \\ \vdots \\ a_{j-1} \\ a_i \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_{j-1} \\ a_j \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix},$$

weil

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_{j-1} \\ a_i \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix} = 0 = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_j \\ a_{i+1} \\ \vdots \\ a_{j-1} \\ a_j \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix}$$

und wir erhalten $\det(B) + \det(A)$.

Zu (d): Wir schreiben die Matrix B als $\begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i + \lambda a_j \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$ und erhalten wegen (D1):

$$\det B = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i + \lambda a_j \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ \lambda a_j \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \det(A) + \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_j \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$$

aber im letzten Term kommt die Zeile a_j doppelt vor und daher bleibt nur $\det(A)$ übrig.

Zu (e): Ist $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$ und sind alle $a_{ii} \neq 0$, so kann man A durch wiederholte Addition von Vielfachen von Zeilen in die Matrix transformieren, die als nicht-triviale Einträge nur die Diagonaleinträge $a_{11}, \dots, a_{nn}$ hat:

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

Die Determinante dieser Matrix ist aber nach (D1)

$$\det \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} = a_{11} \cdot \ldots \cdot a_{nn} \det(E_n)$$

und wegen der Normierung in (D3) ist dies gleich $\prod_{i=1}^n a_{ii}$.

Sind nicht alle $a_{ii} \neq 0$, so gibt es ein größtes i_0, so dass $a_{i_0 i_0} = 0$. Durch wiederholte Addition von Vielfachen der Zeilen $i_0 + 1, \dots, n$ kann man A transformieren zu einer Matrix, deren i_0-te Zeile nur aus Nullen besteht. Damit ist aber die Determinante von A wegen (a) trivial. □

Lemma IV.2.3. *Ist $A \in M(n \times n, K)$, dann ist die Determinante von A genau dann trivial, wenn $\mathsf{rg}(A) < n$ gilt.*

Beweis. Durch allgemeine Zeilenumformungen können wir A zu A' transformieren, wobei A' eine obere Dreiecksmatrix ist:

$$A' = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Hierbei steht $*$ für einen beliebigen Teil der Matrix A', der uns nicht weiter interessiert und 0 steht für ein unteres Dreieck voller Nullen. Damit ist die Determinante von A

$$\det(A) = \pm \det(A') = \pm \prod_{i=1}^{n} \lambda_i$$

und die Determinante von A verschwindet genau dann, wenn es ein $i \in \{1, \ldots, n\}$ gibt, so dass $\lambda_i = 0$. Dies ist genau dann der Fall, wenn A nicht invertierbar ist, und dies wiederum ist äquivalent dazu, dass A nicht vollen Rang hat. $\square$

Korollar IV.2.4 (Eindeutigkeit der Determinantenabbildung)**.** *Für jeden Körper K und für alle $n \geqslant 1$ gibt es höchstens eine Determinantenabbildung*

$$\det\colon M(n \times n, K) \to K.$$

Beweis. Jedes $A \in M(n \times n, K)$ kann durch elementare Zeilenumformungen auf obere Dreiecksgestalt

$$A' = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

gebracht werden und es gilt $\det(A) = \pm \det(A')$, wobei das Vorzeichen bestimmt ist durch die Anzahl der Zeilenvertauschungen. Damit ist

$$\det(A) = \pm \prod_{i=1}^{n} \lambda_i$$

eindeutig bestimmt. $\square$

Satz IV.2.5. *Für jeden Körper K und für alle $n \geqslant 1$ gibt es eine Determinantenabbildung*

$$\det = \det_n\colon M(n \times n, K) \to K.$$

Beweis. Wir beweisen die Behauptung mit vollständiger Induktion über n. Für $n = 1$ müssen wir $\det_1(a) = a$ setzen und diese Abbildung erfüllt die Axiome (D1), (D2) und (D3).

Für den Induktionsschritt definieren wir A'_{ij} als die Matrix, die aus A entsteht, wenn wir die i-te Zeile und die j-te Spalte streichen.

$$A'_{ij} = \begin{pmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{nn} \end{pmatrix}$$

Damit ist $A'_{ij} \in M((n-1) \times (n-1), K)$ und $\det_{n-1} A'_{ij}$ existiert.

Wir wählen ein beliebiges aber festes $j \in \{1, \ldots, n\}$ und definieren

$$\det_n(A) := \textstyle\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \cdot \det_{n-1}(A'_{ij}).$$

Wir müssen (D1), (D2) und (D3) für dieses $\det_n$ nachrechnen.

(D1): Ist $\tilde{A} = (\tilde{a}_{ij})$ mit

$$\tilde{a}_{ij} = \begin{cases} \lambda a_{kj}, & \text{für } k = i, \\ a_{ij}, & \text{für } k \neq i. \end{cases}$$

Damit ist $A'_{kj} = \tilde{A}'_{kj}$ und für $i \neq k$ entsteht $\tilde{A}'_{ij}$ durch Multiplikation einer Zeile von A'_{ij} mit $\lambda \in K$. Damit ist $\det_{n-1}(\tilde{A}'_{ij}) = \lambda \det_{n-1}(A'_{ij})$ und für $\det_n$ erhalten wir

$$\begin{aligned}
\det_n \tilde{A} &= \sum_{\substack{i=1 \\ i\neq k}}^{n} (-1)^{i+j} \tilde{a}_{ij} \det_{n-1}(\tilde{A}'_{ij}) + (-1)^{k+j} \tilde{a}_{kj} \det_{n-1}(\tilde{A}'_{kj}) \\
&= \sum_{\substack{i=1 \\ i\neq k}}^{n} (-1)^{i+j} \tilde{a}_{ij} \lambda \det_{n-1}(A'_{ij}) + (-1)^{k+j} \lambda a_{kj} \det_{n-1}(A'_{kj}) \\
&= \lambda \det_n(A).
\end{aligned}$$

Die Additivität von $\det_n$ zeigen Sie analog.

Zu (D2): In A sei die k-te Zeile gleich der ℓ-ten für $k \neq \ell$.

Für $i \neq k$ oder $i \neq \ell$ hat A'_{ij} auch zwei übereinstimmende Zeilen und somit ist nach Induktionsannahme $\det_{n-1}(A'_{ij}) = 0$ für alle $i \neq k$ und alle $i \neq \ell$. Damit bleiben nur zwei Terme für $\det_n(A)$:

$$\det_n(A) = (-1)^{k+j} a_{k\ell} \det_{n-1}(A'_{kj}) + (-1)^{\ell+j} a_{\ell j} \det_{n-1}(A'_{\ell j}).$$

Es gilt $a_{kj} = a_{\ell j}$. Ohne Beschränkung der Allgemeinheit ist $k < \ell$. Die Matrix $A'_{\ell j}$ geht aus A'_{kj} durch $\ell - k - 1$ Zeilenvertauschungen hervor. Damit ist

$$\det_{n-1}(A'_{\ell j}) = (-1)^{\ell-k-1} \det_{n-1}(A'_{kj})$$

und wir erhalten

$$\det_n(A) = (-1)^{k+j} a_{kj} \det_{n-1}(A'_{kj}) + (-1)^{\ell+j}(-1)^{\ell-k-1} \det_{n-1}(A'_{kj}) = 0,$$

weil

$$(-1)^{\ell-k-1+\ell-j} = (-1)^{2\ell-k-j-1} = (-1)^{k+j+1}.$$

Zu (D3): Wir rechnen einfach nach:

$$\det_n(E_n) = \textstyle\sum_{i=1}^{n} (-1)^{i+j} (E_n)_{ij} \det_{n-1}((E_n)'_{ij})$$

Da $(E_n)_{ij} = \delta_{ij}$, bleibt nur der Term für $i = j$ übrig und die obige Summe ist

$$(-1)^{2j} \det_{n-1}(E_{n-1}) = 1.$$

□

Wir schreiben ab jetzt wieder det statt $\det_n$. Aus dem Beweis von Satz IV.2.5 erhält man sofort die Folgerung:

Korollar IV.2.6 (Spaltenentwicklungssatz von Laplace)**.** *Für $A \in M(n \times n, K)$ ist*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \tag{IV.2.1}$$

für jedes beliebige aber fest gewählte $j \in \{1, \ldots, n\}$.

Beispiel IV.2.7. Wir wollen die Determinante von $\begin{pmatrix} 1 & 2 & 3 \\ 1 & i & 1 \\ 0 & 1 & i \end{pmatrix} \in M(3 \times 3, \mathbb{C})$ berechnen. Die Vorzeichen kann man sich mit dem Schachbrettmuster $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$ merken. Für $j = 1$ ergibt sich für $\det_3 \begin{pmatrix} 1 & 2 & 3 \\ 1 & i & 1 \\ 0 & 1 & i \end{pmatrix}$

$$(-1)^2 \cdot 1 \cdot \det_2 \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix} + (-1)^3 \cdot 1 \cdot \det_2 \begin{pmatrix} 2 & 3 \\ 1 & i \end{pmatrix} + (-1)^4 \cdot 0 \cdot \det_2 \begin{pmatrix} 2 & 3 \\ i & 1 \end{pmatrix}.$$

Determinanten von 2×2-Matrizen sind mit der Formel aus dem Beweis elementar berechenbar:

$$\det_2 \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Damit erhalten wir für die obige Determinante:

$$i^2 - 1 - (2i - 3) = -2 - 2i + 3 = 1 - 2i.$$

Beispiel IV.2.8. Ist $n = 3$ und $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ so ergibt die Formel für die Determinante

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} \\ &\quad - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} \\ &\quad + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22} \end{aligned}$$

Das gibt die sogenannte *Sarrus-Merkregel* für die Determinante von 3×3-Matrizen. Schreiben Sie die Matrix hin und zusätzlich wiederholen Sie die ersten beiden Spalten:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Sie summieren die Terme mit der durchgezogenen Linie mit dem Faktor $(+1)$ und die mit der gestrichelten Linie mit (-1) auf. Aber Vorsicht! Diese Formel gilt nur für $n = 3$. Für allgemeines n hat die Determinante $n!$ Summanden.

Satz IV.2.9. *Die Determinantenfunktion ist linear in jeder Spalte.*

Beweis. Sind $a_1, \dots, a_n, a'_j, a''_j \in K^n$ und ist $\lambda \in K$, so folgt mit der Entwicklung nach der j-ten Spalte:

$$\begin{aligned} \det(a_1, \dots, a_{j-1}, \lambda a_j, a_{j+1}, \dots, a_n) &= \sum_{i=1}^{n} (-1)^{i+j} (\lambda a_{ij}) \det(A'_{ij}) \\ &= \lambda \det(a_1, \dots, a_n) \end{aligned}$$

und

$$\begin{aligned} \det(a_1, \dots, a_{j-1}, a'_j + a''_j, a_{j+1}, \dots, a_n) &= \sum_{i=1}^{n} (-1)^{i+j} (a'_{ij} + a''_{ij}) \det(A'_{ij}) \\ &= \sum_{i=1}^{n} (-1)^{i+j} a'_{ij} \det(A'_{ij}) + \sum_{i=1}^{n} (-1)^{i+j} a''_{ij} \det(A'_{ij}) \\ &= \det(a_1, \dots, a_{j-1}, a'_j, a_{j+1}, \dots, a_n) \\ &\quad + \det(a_1, \dots, a_{j-1}, a''_j, a_{j+1}, \dots, a_n). \end{aligned}$$

$\square$

Satz IV.2.10. *Für alle $A \in M(n \times n, K)$ gilt*

$$\det(A) = \det(A^t).$$

Beweis. Wir zeigen, dass die Abbildung $A \mapsto \det(A^t)$ eine Determinantenabbildung ist. Wegen der Eindeutigkeit folgt dann die Behauptung. Das Axiom (D1) gilt wegen IV.2.9. Hat A zwei übereinstimmende Zeilen, so hat A^t zwei übereinstimmende Spalten. Somit ist der Rang von A^t nicht maximal und somit verschwindet die Determinante $\det(A^t)$. Dies zeigt (D2). Schließlich gilt (D3):

$$\det(E_n^t) = \det(E_n) = 1,$$

weil $E_n^t = E_n$ ist. $\square$

Korollar IV.2.11 (Zeilenentwicklungssatz)**.** *Es seien $A, B \in M(n \times n, K)$.*

(a) *Für jedes $1 \leqslant i \leqslant n$ gilt:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}).$$

(b) *Entsteht B aus A durch das Vertauschen zweier Spalten, so ist* $\det(B) = -\det(A)$.

Beweis. Zu (a): Dies ist genau die Berechnung der Determinante $\det(A^t) = \det(A)$ durch die Entwicklung nach der i-ten Spalte.

Zu (b): In diesem Fall entsteht B^t aus A^t durch das Vertauschen zweier Zeilen. Damit gilt

$$\det(B) = \det(B^t) = -\det(A^t) = -\det(A).$$

□

Satz IV.2.12 (Multiplikationssatz)**.** *Für alle $A, B \in M(n \times n, K)$ gilt*

$$\det(AB) = \det(A)\det(B).$$

Beweis. Gilt $\det(B) = 0$, so ist der Rang von B echt kleiner als n. Damit gibt es ein $x \in K^n$ mit $x \neq 0$, aber $Bx = 0$. Damit ist dann aber auch $(AB)x = A(Bx) = 0$, so dass der Rang von AB auch nicht maximal ist, also ist dann auch $\det(AB) = 0$ und die behauptete Gleichheit gilt in diesem Fall.

Wir können also annehmen, dass $\det(B) \neq 0$. Wir zeigen, dass die Abbildung

$$A \mapsto \frac{\det(AB)}{\det(B)}$$

eine Determinantenabbildung ist. Das beweist die Behauptung.

Die Normierung (D3) ist einfach:

$$\frac{\det(E_n B)}{\det(B)} = \frac{\det(B)}{\det(B)} = 1.$$

Ebenso einfach ist (D2): Hat A zwei übereinstimmende Zeilen, so ist der Rang von A echt kleiner n. Damit ist aber auch der Rang von AB nicht maximal und $\det(AB) = 0$.

Die Arbeit steckt im Nachweis von (D1): Wir betrachten die Elementarmatrix $\Delta^i(\lambda)$ aus Definition III.4.8 und bilden $\tilde{A} = \Delta^i(\lambda)A$. Dann entsteht $\tilde{A}$ aus A durch Multiplikation der i-ten Zeile mit $\lambda \in K$. Aber $\det(\Delta^i(\lambda)) = \lambda$. Damit ist

$$\frac{\det(\tilde{A}B)}{\det(B)} = \frac{\det(\Delta^i(\lambda)AB)}{\det(B)} = \lambda \frac{\det(AB)}{\det(B)}.$$

Für die Additivität schreiben wir $A = \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$ mit Zeilenvektoren a^i und $B = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix}$ mit Spaltenvektoren b_i. Das Produkt AB ist

$$\begin{pmatrix} a^1 \cdot b_1 & \dots & a^1 \cdot b_n \\ \vdots & & \vdots \\ a^n \cdot b_1 & \dots & a^n \cdot b_n \end{pmatrix},$$

wobei $a^i \cdot b_j$ das Matrixprodukt einer $1 \times n$ mit einer $n \times 1$-Matrix ist. Ist $a^i = (a^i_{(1)} + a^i_{(2)})$, so erhalten wir in der i-ten Zeile

$$((a^i_{(1)} + a^i_{(2)}) \cdot b_1, \dots, (a^i_{(1)} + a^i_{(2)}) \cdot b_n) = ((a^i_{(1)}) \cdot b_1, \dots, (a^i_{(1)}) \cdot b_n) + ((a^i_{(2)}) \cdot b_1, \dots, (a^i_{(2)}) \cdot b_n).$$

Damit ist dann $\det(AB) = \det(A_{(1)}B) + \det(A_{(2)}B)$, wobei

$$A_{(j)} = \begin{pmatrix} a^1 \\ \vdots \\ a^{i-1} \\ a^i_{(j)} \\ a^{i+1} \\ \vdots \\ a^n \end{pmatrix}.$$

Insgesamt erhalten wir

$$\frac{\det(AB)}{\det(B)} = \frac{\det(A_{(1)}B) + \det(A_{(2)}B)}{\det(B)} = \frac{\det(A_{(1)}B)}{\det(B)} + \frac{\det(A_{(2)}B)}{\det(B)}.$$

□

Korollar IV.2.13. *Ist* $A \in GL_n(K)$*, so ist* $\det(A^{-1}) = \det(A)^{-1}$*.*

Beweis. Das folgt sofort wegen

$$1 = \det(E_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}).$$

□

Bemerkung IV.2.14. Ähnliche Matrizen haben übereinstimmende Determinanten: Ist $B = TAT^{-1}$ mit $A, B \in M(n \times n, K)$ und $T \in GL_n(K)$, so ist

$$\det(B) = \det(TAT^{-1}) = \det(T)\det(A)\det(T^{-1}) = \det(T)\det(A)\det(T)^{-1} = \det(A).$$

Definition IV.2.15. Für einen Endomorphismus $f\colon V \to V$ eines endlich-dimensionalen K-Vektorraums V und für eine beliebige geordnete Basis $\mathcal{B}$ von V heißt

$$\det(f) := \det M^{\mathcal{B}}_{\mathcal{B}}(f)$$

die *Determinante von* f.

Korollar IV.2.16. *Für Endomorphismen* $f, g\colon V \to V$ *mit* V *wie oben gilt:*

(a) *Die Abbildung* f *ist genau dann ein Automorphismus, wenn* $\det(f) \neq 0$*.*
(b) *Für alle* $f \in \mathrm{Aut}(V)$ *gilt:* $\det(f^{-1}) = \det(f)^{-1}$*.*
(c) $\det(g \circ f) = \det(g)\det(f)$*.*

Beispiele IV.2.17.

(a) Ist $R_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um den Winkel θ, so ist

$$\det(R_\theta) = \det \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \cos^2\theta + \sin^2\theta = 1.$$

(b) Ist $S_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Ursprungsgerade mit Winkel θ, so hat S_θ die Matrixdarstellung $\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$, weil $S_\theta(e_1) = \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \end{pmatrix}$ und $S_\theta(e_2) = \begin{pmatrix} \sin 2\theta \\ -\cos 2\theta \end{pmatrix}$. Damit ist die Determinante

$$\det(S_\theta) = \det \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} = -\cos^2 2\theta - \sin^2 2\theta = -1.$$

Beachten Sie, dass in beiden Beispielen der Wert der Determinante unabhängig vom Winkel ist!

Satz IV.2.18. *Für ein beliebiges* $A \in M(n \times n, K)$ *setzen wir* $C = (c_{ij}) \in M(n \times n, K)$ *mit* $c_{ij} = (-1)^{i+j}\det(A'_{ji})$*. Dann gilt*

$$AC = CA = \det(A)E_n.$$

Insbesondere gilt für $A \in GL_n(K)$*:*

(IV.2.2) $$A^{-1} = \frac{1}{\det(A)}C.$$

Beweis. Wir überprüfen zunächst die Diagonaleinträge: Für beliebiges $i \in \{1, \ldots, n\}$ ist

$$(AC)_{ii} = \sum_{j=1}^{n} a_{ij} c_{ji} = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} \det(A'_{ij}) = \det(A).$$

Für $i \neq k$ gilt

$$(AC)_{ik} = \sum_{j=1}^{n} a_{ij} c_{jk} = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{j+k} \det(A'_{kj}).$$

Ersetzen wir in A die k-te Zeile durch die i-Zeile von A und nennen wir die resultierende Matrix $\tilde{A}$, so ist $A'_{kj} = \tilde{A}'_{kj}$ und somit

$$\sum_{j=1}^{n} a_{ij} \cdot (-1)^{j+k} \det(A'_{kj}) = \sum_{j=1}^{n} \tilde{a}_{kj} \cdot (-1)^{j+k} \det(\tilde{A}'_{kj}) = \det(\tilde{A})$$

aber da $\tilde{A}$ zwei übereinstimmende Zeilen hat, ist $\det(\tilde{A}) = 0$. Damit ist $(AC)_{ik} = 0$ für alle $k \neq i$ und $AC = \det(A)E_n$. Analog folgert man, dass $CA = \det(A)E_n$ gilt. □

Beispiel IV.2.19. Ist K beliebig und ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(K)$, so ist $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ mit $c_{11} = (-1)^{1+1} \det(d) = d$, $c_{12} = (-1)^{1+2} \det(b) = -b$, $c_{21} = (-1)^{2+1} \det(c) = -c$ und $c_{22} = (-1)^{2+2} \det(a) = a$, also

$$C = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Die Determinante von A ist $ad - bc$ und somit ist

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

Satz IV.2.20 (Cramersche Regel)**.** *Es seien $a_1, \ldots, a_n, b \in K^n$ und die Matrix $A = (a_1, \ldots, a_n) \in M(n \times n, K)$ sei invertierbar. Dann ist das LGS $Ax = b$ eindeutig lösbar durch $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ mit*

$$x_i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det(A)}.$$

Bemerkung IV.2.21. Die Cramersche Regel ist für numerische Berechnungen hochgradig ungünstig. Wir werden aber die Existenz und Form der Lösung benutzen.

Beweis. Da A invertierbar ist, ist klar, dass $Ax = b$ eindeutig lösbar ist durch $x = A^{-1}b$. Wir wissen schon, dass

$$(A^{-1})_{ij} = \frac{1}{\det(A)} (-1)^{i+j} \det(A'_{ji}),$$

und somit erhalten wir für x_i

$$x_i = \sum_{j=1}^{n} (A^{-1})_{ij} \cdot b_j = \sum_{j=1}^{n} \frac{1}{\det(A)} (-1)^{i+j} \det(A'_{ji}) \cdot b_j.$$

Wir betrachten die Matrix $(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n)$. Die Entwicklung nach der i-ten Spalte liefert für die Determinante

$$\det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) = (-1)^{i+j} \det(A'_{ji}).$$

Da $b = \sum_{j=1}^{n} b_j e_j$, folgt

$$\begin{aligned} x_i &= \sum_{j=1}^{n} \frac{1}{\det(A)} (-1)^{i+j} \det(A'_{ji}) \cdot b_j \\ &= \frac{1}{\det(A)} \sum_{j=1}^{n} \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) \cdot b_j \\ &= \frac{1}{\det(A)} \det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n). \end{aligned}$$

□

IV.3. Orientierung und Volumen

Orientierung und Volumen sind Konzepte, die Anwendungen von Determinanten bei geometrischen Fragestellungen liefern, zum Beispiel in höherer Analysis, Topologie und Differentialgeometrie, wo man die Orientierbarkeit und das Volumen von sogenannten Mannigfaltigkeiten studiert. Der Grundkörper K ist für diese Betrachtungen $K = \mathbb{R}$.

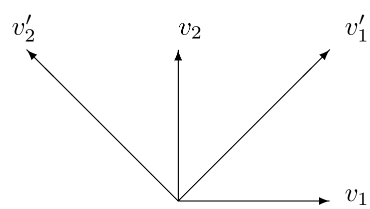
Definition IV.3.1. Es sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum. Zwei geordnete Basen $\mathcal{B}_1, \mathcal{B}_2$ von V heißen *gleich orientiert*, falls für die Transformationsmatrix $T_{\mathcal{B}_2}^{\mathcal{B}_1}$ gilt, dass $\det(T_{\mathcal{B}_2}^{\mathcal{B}_1}) > 0$ ist. Andernfalls heißen $\mathcal{B}_1$ und $\mathcal{B}_2$ *entgegengesetzt orientiert.*

Beispiele IV.3.2.

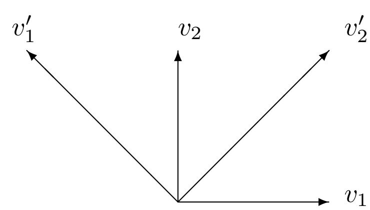
(a) Ist $V = \mathbb{R}$ und ist $\mathcal{B}_1 = (b_1)$ und ist $\mathcal{B}_2 = (b_2)$ mit $b_1, b_2 \in \mathbb{R} \setminus \{0\}$, so sind $\mathcal{B}_1$ und $\mathcal{B}_2$ genau dann gleich orientiert, falls es ein positives $\lambda \in \mathbb{R}$ gibt mit $b_2 = \lambda b_1$.

Ist $\lambda < 0$, so sind $\mathcal{B}_1$ und $\mathcal{B}_2$ entgegengesetzt orientiert. Dies folgt, weil λ^{-1} in diesem Fall die Determinante der Basiswechselmatrix ist.

(b) Ist $V = \mathbb{R}^2$, so sind in



$\mathcal{B}_1 = (v_1, v_2)$ und $\mathcal{B}_2 = (v'_1, v'_2)$ gleich orientiert, während in der Situation



$\mathcal{B}_1 = (v_1, v_2)$ und $\mathcal{B}_2 = (v'_1, v'_2)$ entgegengesetzt orientiert sind.

Lemma IV.3.3. *Die Relation* gleich orientiert *ist eine Äquivalenzrelation auf der Menge aller geordneten Basen eines $\mathbb{R}$-Vektorraums V endlicher Dimension.*

Beweis. Die Dimension von V sei n. Dann ist $\det(T^{\mathcal{B}}_{\mathcal{B}}) = \det(E_n) = 1 > 0$ für jede beliebige geordnete Basis $\mathcal{B}$ von V.

Da
$$\det(T^{\mathcal{B}_1}_{\mathcal{B}_2}) = \det(T^{\mathcal{B}_2}_{\mathcal{B}_1})^{-1},$$
gilt, dass $\det(T^{\mathcal{B}_1}_{\mathcal{B}_2})$ genau dann positiv ist, wenn $\det(T^{\mathcal{B}_2}_{\mathcal{B}_1})$ positiv ist.

Mit
$$\det(T^{\mathcal{B}_1}_{\mathcal{B}_3}) = \det(T^{\mathcal{B}_1}_{\mathcal{B}_2})\det(T^{\mathcal{B}_2}_{\mathcal{B}_3})$$
folgt die Transitivität. □

Definition IV.3.4. Ist V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum, so heißt eine Äquivalenzklasse geordneter Basen eine *Orientierung von V*.

Satz IV.3.5. *Jeder endlich-dimensionaler $\mathbb{R}$-Vektorraum $V \neq \{0\}$ besitzt genau zwei Orientierungen.*

Beweis. Es gibt mindestens zwei Orientierungen: Ist $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ eine geordnete Basis und sei $\mathcal{B}' = (-v_1, v_2, \ldots, v_n)$, dann ist
$$\det(T^{\mathcal{B}}_{\mathcal{B}'}) = \det\begin{pmatrix} -1 & 0 \\ 0 & E_{n-1} \end{pmatrix} = -1$$
und damit sind $\mathcal{B}$ und $\mathcal{B}'$ entgegengesetzt orientiert.

Es gibt genau zwei Orientierungen: Sind $\mathcal{B}_1$ und $\mathcal{B}_2$ entgegengesetzt orientiert und sind $\mathcal{B}_2$ und $\mathcal{B}_3$ ebenfalls entgegengesetzt orientiert, so ist $\det(T^{\mathcal{B}_1}_{\mathcal{B}_2}) < 0$ und $\det(T^{\mathcal{B}_2}_{\mathcal{B}_3}) < 0$. Damit ist aber $\det(T^{\mathcal{B}_1}_{\mathcal{B}_3}) > 0$, so dass $\mathcal{B}_1$ und $\mathcal{B}_3$ gleich orientiert sind. □

Definition IV.3.6. Es sei $V \neq \{0\}$ ein $\mathbb{R}$-Vektorraum und $\dim_{\mathbb{R}} V < \infty$. Wir nennen einen Automorphismus $f \in \mathrm{Aut}(V)$ *orientierungserhaltend*, falls $\det(f) > 0$ ist. Wir schreiben $\mathrm{Aut}(V)^+$ für
$$\mathrm{Aut}(V)^+ = \{f \in \mathrm{Aut}(V), \det(f) > 0\}.$$

Bemerkung IV.3.7. Rechnen Sie nach, dass $\mathrm{Aut}(V)^+ \subset \mathrm{Aut}(V)$ eine Untergruppe ist!

Beispiel IV.3.8. Die Drehungen R_θ des $\mathbb{R}^2$ sind orientierungserhaltend, weil $\det(R_\theta) = 1$. Aber Spiegelungen S_θ haben $\det(S_\theta) = -1$, sind also nicht orientierungserhaltend.

IV.4. Determinanten und Permutationen

Wir hatten für ein beliebiges $n \in \mathbb{N}$ die symmetrische Gruppe Σ_n kennengelernt als die Gruppe der bijektiven Selbstabbildungen der Menge $\mathbf{n} = \{1, \ldots, n\}$. Sie hatten sich überlegt, dass $|\Sigma_n| = n!$ gilt und dass Σ_n nicht abelsch ist, falls $n \geqslant 3$ ist.

Im Folgenden benutzen wir die folgende Notation: Ist $\sigma \in \Sigma_n$, so notieren wir σ durch die Wertetabelle $\begin{pmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{pmatrix}$. Sind $\tau, \sigma \in \Sigma_n$, so ist damit
$$\begin{pmatrix} 1 & \ldots & n \\ \tau(1) & \ldots & \tau(n) \end{pmatrix}\begin{pmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{pmatrix} = \begin{pmatrix} 1 & \ldots & n \\ \tau(\sigma(1)) & \ldots & \tau(\sigma(n)) \end{pmatrix}.$$
Zum Beispiel ist
$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$
Die Elemente in Σ_n nennt man auch *Permutationen*.

Definition IV.4.1. Ein $\tau \in \Sigma_n$ heißt *Transposition*, falls τ genau zwei Elemente aus $\mathbf{n}$ vertauscht und alle anderen Elemente fest läßt.

Beispiel IV.4.2. Die Elemente
$$\tau_{12} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau_{13} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \text{ und } \tau_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$
sind die Transpositionen in Σ_3.

Wir schreiben im Folgenden id für $\mathrm{id}_{\mathbf{n}}$.

Lemma IV.4.3.
(a) *Für jede Transposition $\tau \in \Sigma_n$ gilt $\tau \circ \tau =: \tau^2 = \mathrm{id}$.*
(b) *Jede Permutation $\sigma \in \Sigma_n$ für $n \geqslant 2$ läßt sich als endliches Produkt von Transpositionen schreiben:*

$$\sigma = \tau_1 \circ \ldots \circ \tau_k, \quad \tau_i \text{ ist Transposition.}$$

(c) *Jede Transposition $\tau \in \Sigma_n$ ist zur Transposition*

$$\tau_{12} = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 2 & 1 & 3 & \ldots & n \end{pmatrix}$$

konjugiert, das heißt es gibt ein $\sigma \in \Sigma_n$, so dass

$$\tau = \sigma \circ \tau_{12} \circ \sigma^{-1}$$

gilt.

Beweis. Behauptung (a) gilt nach Definition einer Transposition.
Zu (b): Ist $\sigma = \mathrm{id}$, so setzen wir $\sigma = \tau \circ \tau$ mit einer beliebigen Transposition τ.
Ist $\sigma \neq \mathrm{id}$, so gibt es ein $i_1 \in \{1, \ldots, n\}$, so dass

$$\sigma(i) = \begin{cases} i, & \text{falls } i \in \{1, \ldots, i_1 - 1\}, \\ \neq i_1, & \text{für } i = i_1. \end{cases}$$

Damit ist dann $\sigma(i_1) > i_1$. Wir betrachten die Transposition τ_1, die genau i_1 und $\sigma(i_1)$ vertauscht und wir setzen $\sigma_1 = \tau_1 \circ \sigma$. Dann ist $\sigma(i) = i$ für alle $i \in \{1, \ldots, i_1\}$. Entweder ist $\sigma_1 = \mathrm{id}$ oder wir iterieren den obigen Prozess und erhalten τ_2, σ_2 bis τ_k, σ_k, so dass $\sigma_k = \mathrm{id}$ mit $k \leqslant n$ und

$$\sigma_k = \tau_k \circ \ldots \circ \tau_1 \circ \sigma = \mathrm{id}.$$

Dann ist aber

$$\sigma = (\tau_k \circ \ldots \circ \tau_1)^{-1} = \tau_1^{-1} \circ \ldots \circ \tau_k^{-1} = \tau_1 \circ \ldots \circ \tau_k.$$

Zu (c): Die Transposition τ vertausche k und ℓ in $\mathbf{n}$. Wir wählen ein $\sigma \in \Sigma_n$ mit $\sigma(1) = k$ und $\sigma(2) = \ell$. Die anderen Werte von σ sind beliebig.
Dann gilt

$$\sigma \circ \tau_{12} \circ \sigma^{-1}(k) = \sigma(\tau_{12}(1)) = \sigma(2) = \ell \text{ und } \sigma \circ \tau_{12} \circ \sigma^{-1}(\ell) = \sigma(\tau_{12}(2)) = \sigma(1) = k.$$

Ist $k \neq i \neq \ell$, so ist $\sigma \circ \tau_{12} \circ \sigma^{-1}(i) = \sigma \circ \sigma^{-1}(i) = i$. □

Bemerkung IV.4.4. Zu einer Permutation $\sigma \in \Sigma_n$ betrachten wir die Matrix E_σ mit

$$E_\sigma = \begin{pmatrix} (e_{\sigma^{-1}(1)})^t \\ \vdots \\ (e_{\sigma^{-1}(n)})^t \end{pmatrix}.$$

Da E_σ aus E_n durch das Vertauschen von Zeilen hervorgeht, muss gelten

$$\det(E_\sigma) \in \{\pm 1\}.$$

Ist $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, so ist

$$E_\sigma = \begin{pmatrix} (e_{\sigma^{-1}(1)})^t \\ (e_{\sigma^{-1}(2)})^t \\ (e_{\sigma^{-1}(3)})^t \end{pmatrix} = \begin{pmatrix} (e_3)^t \\ (e_1)^t \\ (e_2)^t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Lemma IV.4.5. *Für $\tau, \sigma \in \Sigma_n$ gilt:*
(a) $(E_\sigma)_{ij} = \delta_{i,\sigma(j)}$
(b) $E_{\sigma \circ \tau} = E_\sigma E_\tau$.

Beweis. Zu (a): Die j-te Spalte von E_σ enthält nur eine einzige 1 und sonst nur Nullen. In der i-ten Zeile steht $(e_{\sigma^{-1}(i)})^t$. Das hat nur dann eine 1 an der Stelle j, wenn $\sigma^{-1}(i) = j$ ist, also $i = \sigma(j)$.

Für (b) vergleichen wir die Matrixeinträge:

$$\begin{aligned}(E_\sigma E_\tau)_{ij} &= \sum_{k=1}^{n} (E_\sigma)_{ik}(E_\tau)_{kj} \\ &= \sum_{k=1}^{n} \delta_{i,\sigma(k)}\delta_{k,\tau(j)} \\ &= \delta_{i,\sigma(\tau(j))} \\ &= (E_{\sigma\circ\tau})_{ij}.\end{aligned}$$

□

Definition IV.4.6. Die Abbildung

$$\mathrm{sign}\colon \Sigma_n \to \{\pm 1\}, \quad \sigma \mapsto \det(E_\sigma) =: \mathrm{sign}(\sigma)$$

heißt die *Signumsabbildung*. Der Kern der Signumsabbildung heißt die *alternierende Gruppe* A_n, also

$$A_n = \ker(\mathrm{sign}) = \{\sigma \in \Sigma_n, \mathrm{sign}(\sigma) = +1\}.$$

Bemerkung IV.4.7.

- Da gilt:
  $$\mathrm{sign}(\sigma\circ\tau) = \det(E_{\sigma\circ\tau}) = \det(E_\sigma E_\tau) = \det(E_\sigma)\det(E_\tau) = \mathrm{sign}(\sigma)\mathrm{sign}(\tau),$$
  ist sign ein Gruppenhomomorphismus $\mathrm{sign}\colon (\Sigma_n, \circ) \to (\{\pm 1\}, \cdot)$.
- Ist τ eine Transposition, so ist $\mathrm{sign}(\tau) = -1$.
- Die Gruppe A_n hat $\frac{n!}{2}$ Elemente.

Satz IV.4.8 (Leibnizregel)**.** *Ist K ein beliebiger Körper und ist $A = (a_{ij}) \in M(n\times n, K)$, so gilt*

$$\det(A) = \sum_{\sigma\in\Sigma_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}.$$

Beweis. Die i-te Zeile von A ist $(a_{i1}, \ldots, a_{in})$ und das können wir schreiben als

$$a_{i1}e_1^t + \ldots + a_{in}e_n^t.$$

Entwickeln wir die Determinante von A nach der ersten Zeile, erhalten wir damit mit (D1)

$$\det(A) = \sum_{i_1=1}^{n} a_{1i_1} \det\begin{pmatrix}(e_{i_1})^t \\ (a_2)^t \\ \vdots \\ (a_n)^t\end{pmatrix},$$

wobei wir $(a_j)^t$ abkürzend für die j-te Spalte von A schreiben. Wir iterieren diesen Prozess, indem wir die zweite Zeile ebenso als Summe schreiben, und danach die weiteren Zeilen bis zur letzten. Damit ergibt sich für die Determinante von A:

$$\begin{aligned}\det(A) &= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} a_{1i_1}a_{2i_2} \det\begin{pmatrix}(e_{i_1})^t \\ (e_{i_2})^t \\ (a_3)^t \\ \vdots \\ (a_n)^t\end{pmatrix} \\ &= \ldots = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\ldots\sum_{i_n=1}^{n} a_{1i_1}\cdot\ldots\cdot a_{ni_n} \det\begin{pmatrix}(e_{i_1})^t \\ \vdots \\ (e_{i_n})^t\end{pmatrix}.\end{aligned}$$

Ist $i_j = i_k$ für $1 \leqslant j \neq k \leqslant n$, so ist $\det \begin{pmatrix} (e_{i_1})^t \\ \vdots \\ (e_{i_n})^t \end{pmatrix} = 0$, also müssen wir nur diejenigen $(i_1, \ldots, i_n)$ berücksichtigen, die paarweise verschieden sind. Für solche $(i_1, \ldots, i_n)$ gibt es aber genau ein $\sigma \in \Sigma_n$ mit $\sigma(j) = i_j$. Damit folgt schließlich

$$\det(A) = \sum_{\sigma \in \Sigma_n} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \det \begin{pmatrix} (e_{\sigma(1)})^t \\ \vdots \\ (e_{\sigma(n)})^t \end{pmatrix} = \sum_{\sigma \in \Sigma_n} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \operatorname{sign}(\sigma^{-1}).$$

Aber $\operatorname{sign}(\sigma^{-1})$ stimmt mit $\operatorname{sign}(\sigma)$ überein. □

Korollar IV.4.9. *Ist $A \in M(n \times n, K)$ mit $n = n_1 + n_2$, $n_1, n_2 \geqslant 1$ und so dass es $B \in M(n_1 \times n_1, K)$ und $C \in M(n_2 \times n_2, K)$ gibt mit*

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$$

mit einer Matrix D. Dann gilt $\det(A) = \det(B)\det(C)$.

Beweis. Wir betrachten die Abbildung $i\colon \Sigma_{n_1} \times \Sigma_{n_2} \to \Sigma_{n_1+n_2}$,

$$i(\sigma_1, \sigma_2)(k) = \begin{cases} \sigma_1(k), & k \leqslant n_1 \\ \sigma_2(k - n_1) + n_1, & k > n_1. \end{cases}$$

Es gilt:

$$i(\sigma_1, \sigma_2) = i(\sigma_1, \mathrm{id}_{\mathbf{n_2}}) \circ i(\mathrm{id}_{\mathbf{n_1}}, \sigma_2).$$

Damit ist

$$\operatorname{sign}(i(\sigma_1, \sigma_2)) = \operatorname{sign}(i(\sigma_1, \mathrm{id}_{\mathbf{n_2}}))\operatorname{sign}(i(\mathrm{id}_{\mathbf{n_1}}, \sigma_2)) = \operatorname{sign}(\sigma_1)\operatorname{sign}(\sigma_2).$$

In der Determinante von A

$$\det(A) = \sum_{\sigma \in \Sigma_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$$

sind nur die Summanden nicht null, für die gilt

$$\sigma(i) \in \{1, \ldots, n_1\} \Rightarrow i \in \{1, \ldots, n_1\}.$$

Diese Permutationen entsprechen aber genau dem Bild von i und wir erhalten

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in i(\Sigma_{n_1} \times \Sigma_{n_2})} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \\
&= \sum_{(\sigma_1, \sigma_2) \in \Sigma_{n_1} \times \Sigma_{n_2}} \operatorname{sign}(i(\sigma_1, \sigma_2)) a_{1 i(\sigma_1,\sigma_2)(1)} \cdot \ldots \cdot a_{n i(\sigma_1,\sigma_2)(n)} \\
&= \sum_{\sigma_1 \in \Sigma_{n_1}} \operatorname{sign}(\sigma_1) a_{1\sigma_1(1)} \cdot \ldots \cdot a_{n_1 \sigma_1(n_1)} \cdot \sum_{\sigma_2 \in \Sigma_{n_2}} \operatorname{sign}(\sigma_2) a_{n_1+1, \sigma_2(1)+n_1} \cdot \ldots \cdot a_{n_1+n_2, \sigma_2(n_2)+n_1} \\
&= \sum_{\sigma_1 \in \Sigma_{n_1}} \operatorname{sign}(\sigma_1) b_{1\sigma_1(1)} \cdot \ldots \cdot b_{n_1 \sigma_1(n_1)} \cdot \sum_{\sigma_2 \in \Sigma_{n_2}} \operatorname{sign}(\sigma_2) c_{1, \sigma_2(1)} \cdot \ldots \cdot a_{n_2, \sigma_2(n_2)} \\
&= \det(B) \cdot \det(C).
\end{aligned}$$

□

Beispiel IV.4.10. Es sei $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & 4 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 3 & 0 \end{pmatrix} \in M(3 \times 3, \mathbb{Q})$. Dann ist

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & 4 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 3 & 0 \end{pmatrix} &= \det \begin{pmatrix} 1 & 1 & 2 & 0 \\ 3 & 0 & 4 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \\ &= (-3) \cdot (3-2) = -3. \end{aligned}$$

Beim ersten Schritt haben wir Spaltenvertauschungen durchgeführt (Spalte vier mit Spalte drei und dann Spalte drei mit Spalte zwei).

IV.5. Minoren

Minoren liefern ein Determinantenkriterium für den Rang einer Matrix.

Definition IV.5.1. Ist $A \in M(m \times n, K)$ und ist $k \leqslant m, n$, so heißt $A' \in M(k \times k, K)$ eine *$k \times k$-Untermatrix von A*, falls A' aus A durch das Streichen von $m-k$ Zeilen und $n-k$ Spalten hervorgeht.

Die Determinanten $\det(A')$ heißt dann ein *Minor k-ter Ordnung von A, oder kurz ein k-Minor.*

Satz IV.5.2 (Minorenkriterium für den Rang)**.** *ist $A \in M(m \times n, K)$, so sind äquivalent:*

(a) $\mathsf{rg}(A) \geqslant k$

(b) *Es gibt eine $k \times k$-Untermatrix A' von A mit* $\det(A') \neq 0$.

Beweis. Gilt $\det(A') \neq 0$, so ist der Rang von A' genau k, aber damit gilt für den Rang von A

$$\mathsf{rg}(A) \geqslant \mathsf{rg}(A') = k.$$

Ist umgekehrt $\mathsf{rg}(A) \geqslant k$, so gibt es k linear unabhängige Zeilen in A. Wir wählen diese Zeilen aus und schreiben sie in eine Matrix $B \in M(k \times n, K)$. Da der Zeilenrang von B gleich dem Spaltenrang von B ist, muss es in B auch k linear unabhängige Spalten geben. Wir schreiben diese in eine Matrix $C \in M(k \times k, K)$ und dieses C erfüllt $\det(C) \neq 0$, weil C vollen Rang hat. □

Korollar IV.5.3. *Für $A \in M(m \times n, K)$ und $r \in \mathbb{N}$ sind äquivalent:*

(a) $\mathsf{rg}(A) = r$

(b) *Es gibt einen Minor der Ordnung r, der nicht verschwindet und alle Minoren höherer Ordnung sind trivial.*

Beispiel IV.5.4. Der Rang von $\begin{pmatrix} 3 & 6 & 0 \\ 2 & 4 & 1 \end{pmatrix}$ ist zwei, weil zum Beispiel gilt, dass $\det \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix} = 3 \neq 0$ ist.

KAPITEL V

Eigenwerttheorie

V.1. Eigenwerte und Eigenvektoren

Beispiel V.1.1. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $M(f) = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}$, also $f(e_1) = 4e_1$. In Richtung e_1 verursacht f also lediglich eine Streckung um den Faktor 4.

Definition V.1.2. Es sei V ein K-Vektorraum und $f\colon V \to V$ sei ein Endomorphismus.

(a) Ein $\lambda \in K$ heißt ein *Eigenwert von* f, falls es ein $v \in V \setminus \{0\}$ gibt, so dass

$$f(v) = \lambda v.$$

(b) Ein Vektor $v \in V \setminus \{0\}$ mit $f(v) = \lambda v$ heißt ein *Eigenvektor von* f *zum Eigenwert* λ.

(c) Für $A \in M(n \times n, K)$ heißt ein $\lambda \in K$ ein *Eigenwert von* A, falls es ein $v \in V \setminus \{0\}$ gibt, so dass $Av = \lambda v$ und dieses v heißt dann *Eigenvektor von* A *zum Eigenwert* λ.

(d) Der Endomorphismus f heißt *diagonalisierbar*, falls es eine Basis von V gibt, die aus Eigenvektoren besteht.

Bemerkung V.1.3. Vorsicht: 0_V ist *niemals* ein Eigenvektor, aber $\lambda = 0_K$ kann als Eigenwert vorkommen.

Ist f diagonalisierbar und ist V endlich-dimensional, so ist $M_{\mathcal{B}}^{\mathcal{B}}(f)$ mit $\mathcal{B}$ wie oben eine Diagonalmatrix.

Beispiele V.1.4.

- Ist $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $M(f) = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalisierbar?

  Es gilt $f\begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ und $\mathcal{B} = \left(e_1, \begin{pmatrix} 1 \\ -3 \end{pmatrix}\right)$ ist eine Basis aus Eigenvektoren, also ist f diagonalisierbar mit Eigenwerten 4 und 1.
- Es sei R_θ die Drehung im $\mathbb{R}^2$ um den Winkel θ, also $M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Speziell für $\theta = 0, \pi$ erhalten wir

  $$M(R_0) = E_2 \text{ und } M(R_\pi) = -E_2,$$

  und damit sind R_0 und R_π diagonalisierbar mit $\mathcal{B} = (e_1, e_2)$. Was passiert für andere Winkel?

Satz V.1.5. *Ist $f \in \mathsf{End}_K(V)$ und sind $v_1, \ldots, v_n$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$, so ist die Familie $(v_1, \ldots, v_n)$ linear unabhängig.*

Beweis. Wir beweisen die Behauptung mit Induktion über n. Für $n = 1$ ist also v_1 ein Eigenvektor zum Eigenwert λ_1 und damit ist $v_1 \neq 0$, so dass (v_1) linear unabhängig ist.

Für den Induktionsschluss von $n-1$ auf n betrachten wir

$$\sum_{i=1}^{n} \mu_i v_i = 0$$

mit $\mu_i \in K$. Damit ist dann auch

$$0 = f(0) = \sum_{i=1}^{n} \mu_i f(v_i) = \sum_{i=1}^{n} \mu_i \lambda_i v_i.$$

Andererseits ist auch

$$\lambda_1 \sum_{i=1}^{n} \mu_i v_i = \sum_{i=1}^{n} \lambda_1 \mu_i v_i = 0,$$

so dass

$$\sum_{i=2}^{n} \mu_i(\lambda_i - \lambda_1)v_i = 0$$

gilt. Nach Induktionsannahme sind allerdings $(v_2, \dots, v_n)$ linear unabhängig, so dass wir für alle $2 \leqslant i \leqslant n$ die Gleichung $\mu_i(\lambda_1 - \lambda_i) = 0$ erhalten. Die λ_i waren aber paarweise verschieden, so dass $\mu_i = 0$ sein muss für $2 \leqslant i \leqslant n$. In der Ausgangsgleichung ist dann aber auch $\mu_1 v_1 = 0$. Da $v_1 \neq 0$ als Eigenvektor, muss damit $\mu_1 = 0$ sein. Somit sind alle μ_i trivial. □

Korollar V.1.6. *Es sei* $\dim_K V = n < \infty$. *Dann gilt:*

(a) *Jedes* $f \in \mathsf{End}_K(V)$ *hat höchstens* n *verschiedene Eigenwerte.*
(b) *Hat* f *genau* n *verschiedene Eigenwerte, so ist* f *diagonalisierbar.*

Beweis. Da jede Familie von Vektoren mit mehr als n Elementen linear abhängig ist, folgt (a). Ist v_i jeweils ein Eigenvektor zum Eigenwert λ_i, so ist nach Satz V.1.5 die Familie $\mathcal{B} = (v_1, \dots, v_n)$ linear unabhängig und damit ist $\mathcal{B}$ eine Basis. Die Matrixdarstellung von f bezüglich $\mathcal{B}$ ist

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

□

Bemerkung V.1.7. Die Bedingung 1.6 (b) ist nur eine Implikation, keine Äquivalenz. Die Einheitsmatrix E_n ist natürlich diagonalisierbar, aber alle n Eigenwerte stimmen überein und sind gleich 1.

Definition V.1.8. Für ein $f \in \mathsf{End}_K(V)$ und ein $\lambda \in K$ heißt

$$\mathrm{Eig}(f; \lambda) := \{v \in V, f(v) = \lambda v\}$$

der *Eigenraum von* f *bezüglich* λ.

Bemerkung V.1.9.

- Der Nullvektor 0_V ist zwar niemals ein Eigenvektor, aber es gilt $0_V \in \mathrm{Eig}(f; \lambda)$ für alle $f \in \mathsf{End}_K(V)$ und alle $\lambda \in K$.
- Ein $\lambda \in K$ ist genau dann ein Eigenwert von f, wenn der Eigenraum nicht nur aus dem Nullvektor besteht.
- Ist $\lambda_1 \neq \lambda_2$, so ist
$$\mathrm{Eig}(f; \lambda_1) \cap \mathrm{Eig}(f; \lambda_2) = \{0_V\} :$$
Ist v ein Vektor im Schnitt, so gilt $\lambda_1 v = f(v) = \lambda_2 v$ und damit $(\lambda_1 - \lambda_2)v = 0$. Ist $v \neq 0_V$, so muss $\lambda_1 - \lambda_2 = 0$ gelten, also $\lambda_1 = \lambda_2$.
- Für ein $v \in V$ gilt:
$$v \in \mathrm{Eig}(f; \lambda) \Leftrightarrow v \in \ker(f - \lambda \mathrm{id}_V),$$
wobei $f - \lambda \mathrm{id}_V \in \mathsf{End}_K(V)$. Damit ist $\mathrm{Eig}(f; \lambda)$ ein Untervektorraum von V.
- Mit dieser Überlegung erhalten wir, dass $\lambda \in K$ genau dann ein Eigenwert von f ist, falls $f - \lambda \mathrm{id}_V$ nicht injektiv ist. Ist V endlich-dimensional, so ist das äquivalent dazu, dass
$$\det(f - \lambda \mathrm{id}_V) = 0.$$
Wir haben also ein Determinantenkriterium dafür, dass λ ein Eigenwert ist.

Definition V.1.10. Es sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathsf{End}_K(V)$. Die Abbildung

$$P_f \colon K \to K, \quad \lambda \mapsto \det(f - \lambda \mathrm{id}_V)$$

heißt das *charakteristische Polynom von* f.

Bemerkung V.1.11.

- Eigentlich müsste P_f die *charakteristische Polynomfunktion* heißen, aber die Bezeichnung ist so üblich.
- Die Eigenwerte von f sind genau die Nullstellen von P_f.
- Ist $\mathcal{B}$ eine geordnete Basis von V und ist $A = M_{\mathcal{B}}^{\mathcal{B}}(f) \in M(n \times n, K)$, so ist
$$M_{\mathcal{B}}^{\mathcal{B}}(f - \lambda \mathrm{id}_V) = A - \lambda E_n$$
und für das charakteristische Polynom gilt:
$$P_f(\lambda) = \det(f - \lambda \mathrm{id}_V) = \det(M_{\mathcal{B}}^{\mathcal{B}}(f - \lambda \mathrm{id}_V)) = \det(A - \lambda E_n).$$
Wir setzen daher $P_A(\lambda) := P_f(\lambda)$.

Lemma V.1.12. *Sind A und A' ähnliche Matrizen, so ist $P_A(\lambda) = P_{A'}(\lambda)$.*

Beweis. Ähnliche Matrizen stellen den gleichen Endomorphismus $f \in \mathsf{End}_K(V)$ bezüglich verschiedener Basen dar. Für dieses f ist dann
$$P_A(\lambda) = P_f(\lambda) = P_{A'}(\lambda).$$
□

Definition V.1.13. Ist $f \in \mathsf{End}_K(V)$ und $\lambda \in K$. Gilt $\dim_K \mathrm{Eig}(f; \lambda) \neq 0$, so heißt
$$\mu_{\mathrm{geo}}(f; \lambda) := \dim_K \mathrm{Eig}(f; \lambda)$$
die *geometrische Vielfachheit von λ bezüglich f.*

Beispiel V.1.14. Es sei $R_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$ wiederum die Drehung um den Winkel θ mit
$$A_\theta := M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$
Dann ist
$$\begin{aligned} P_{A_\theta}(\lambda) &= \det \begin{pmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{pmatrix} \\ &= (\cos\theta - \lambda)^2 + \sin^2\theta = \cos^2\theta - 2\lambda\cos\theta + \lambda^2 + \sin^2\theta \\ &= \lambda^2 - 2\lambda\cos\theta + 1. \end{aligned}$$
Diese Polynomfunktion hat die komplexen Nullstellen
$$\lambda_1 = \cos\theta + i\sin\theta, \quad \lambda_2 = \cos\theta - i\sin\theta.$$
Diese Nullstellen sind nur dann reell, wenn $\sin\theta = 0$, und für $\theta \in [0, 2\pi)$ ist das nur für $\theta = 0, \pi$ der Fall. Die Drehnung um den Winkel 0, R_0, ist die identische Abbildung und $A_0 = E_2$ ist eine Diagonalmatrix. Wir haben
$$\dim_{\mathbb{R}} \mathrm{Eig}(R_0; 1) = 2 = \mu_{\mathrm{geo}}(R_0; 1).$$
Die Drehung um 180 Grad, R_π, hat $A_\pi = -E_2$ und hier sind
$$\dim_{\mathbb{R}} \mathrm{Eig}(R_\pi; -1) = 2 = \mu_{\mathrm{geo}}(R_\pi; -1).$$

Drehungen sind also nur für die Ausnahmewinkel $0, \pi$ diagonalisierbar als $\mathbb{R}$-lineare Abbildungen.

Bemerkung V.1.15. Wir erhalten folgenden Algorithmus zur Bestimmung von Eigenwerten und Eigenvektoren von $A \in M(n \times n, K)$:

(a) Berechnen Sie die Nullstellen der Funktion P_A.

(b) Ist $P_A(\lambda) = 0$, so bestimmen Sie mit dem Gauß-Algorithmus die Eigenvektoren:
$$v \in \mathrm{Eig}(A; \lambda) \Leftrightarrow (A - \lambda E_n)v = 0.$$

(c) Hat der K^n eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$, die aus Eigenvektoren von A besteht, so setzen wir
$$S^{-1} := (v_1, \ldots, v_n) \in GL_n(K).$$
Damit gilt dann
$$SAS^{-1}e_i = SAv_i = S\lambda_i v_i = \lambda_i S v_i = \lambda_i e_i$$
und SAS^{-1} ist damit eine Diagonalmatrix.

Beispiel V.1.16. Es sei $S_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Ursprungsgeraden mit Winkel θ. Dann ist

$$M(S_\theta) = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

und

$$\begin{aligned} P_{S_\theta}(\lambda) &= \det \begin{pmatrix} \cos 2\theta - \lambda & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta - \lambda \end{pmatrix} \\ &= (\cos 2\theta - \lambda)(-\cos 2\theta - \lambda) - \sin^2 2\theta \\ &= -\cos^2 2\theta + \lambda^2 - \sin^2 2\theta \\ &= \lambda^2 - 1. \end{aligned}$$

Wir erhalten also die Nullstellen $\lambda_1 = 1$ und $\lambda_2 = -1$. Der Eigenraum zu 1 ist:

$$\operatorname{Eig}(S_\theta; 1) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x\cos 2\theta + y \sin 2\theta \\ x \sin 2\theta - y \cos 2\theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \right\}$$

Mit den Additionstheoremen erhalten wir, dass die Lösungsmenge genau aus den Vektoren

$$\left\{ \begin{pmatrix} t\cos\theta \\ t \sin\theta \end{pmatrix}, t \in \mathbb{R} \right\}$$

besteht und dies ist genau die Gerade, an der S_θ spiegelt. Analog bekommen wir

$$\operatorname{Eig}(S_\theta; -1) = \left\{ \begin{pmatrix} -t\sin\theta \\ t \cos\theta \end{pmatrix}, t \in \mathbb{R} \right\},$$

und dies ist die Urspungsgerade, die senkrecht auf der Spiegelungsgeraden steht.

Damit ist zum Beispiel für $t = 1$ $S^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, also genau $M(R_\theta)$, und

$$SM(S_\theta)S^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Beispiel V.1.17. Es sei $A = \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix} \in M(2 \times 2, \mathbb{R})$. Dann ist

$$P_A(\lambda) = \det \begin{pmatrix} 2-\lambda & 2 \\ 0 & 2-\lambda \end{pmatrix} = (2-\lambda)^2.$$

Wir haben also nur $\lambda = 2$ als Nullstelle. Was ist $\operatorname{Eig}(A; 2)$? Das passende LGS ist

$$A \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

also $2y = 0$. Damit bekommen wir $y = 0$ und $\operatorname{Eig}(A; 2) = \mathsf{Span}_{\mathbb{R}}(e_1)$ also $\dim_{\mathbb{R}} \operatorname{Eig}(A; 2) = 1$. Die Matrix A ist also nicht diagonalisierbar.

V.2. Polynome

Wir hatten schon in den Übungen zur Linearen Algebra und Analytischen Geometrie I Polynome kennengelernt. Für das Weitere brauchen wir etwas mehr Hintergrund. In diesem Abschnitt sei K ein kommutativer Ring mit 1.

Definition V.2.1. Eine Menge A mit Verknüpfungen

$$\begin{aligned} +&\colon A \times A \to A \\ *&\colon A \times A \to A \\ \cdot&\colon K \times A \to A \end{aligned}$$

heißt eine K-*Algebra*, falls gilt:

(A1) $(A, +, *)$ ist ein Ring.

(A2) Für alle $a, b \in A$ und $\lambda, \mu \in K$ gilt:

$$\begin{aligned}(\lambda + \mu) \cdot a &= \lambda \cdot a + \mu \cdot a,\\ \lambda \cdot (a+b) &= \lambda \cdot a + \lambda \cdot b,\\ (\lambda\mu) \cdot a &= \lambda \cdot (\mu \cdot a),\\ 1 \cdot a &= a.\end{aligned}$$

(A3) Für alle $a, b \in A$ und $\lambda, \mu \in K$ gilt:

$$(\lambda \cdot a) * (\mu \cdot b) = (\lambda\mu) \cdot (a * b).$$

Bemerkung V.2.2.

- Ist K ein Körper, so ist A ein K-Vektorraum.
- Wir schreiben meistens ab statt $a * b$.
- Ist $(A, +, *)$ ein kommutativer Ring (oder ein Ring mit 1), so heißt A eine *kommutative K-Algebra* (oder eine *unitäre K-Algebra*).
- Sie kennen schon Beispiele von K-Algebren: Für jedes $n \in \mathbb{N}$ ist $M(n \times n, K)$ eine unitäre K-Algebra, die allerdings für $n \geqslant 2$ nicht kommutativ ist.
- Für jeden K-Vektorraum V ist $\mathsf{End}_K(V)$ eine unitäre K-Algebra.

Definition V.2.3. Sind A und B zwei K-Algebren, so heißt ein Ringhomomorphismus $f\colon A \to B$ ein *K-Algebrahomomorphismus*, wenn für alle $\lambda \in K$ und alle $a \in A$ gilt:

$$f(\lambda \cdot a) = \lambda \cdot f(a).$$

Sind A und B unitär mit Einselementen $1_A \in A$ und $1_B \in B$, so verlangen wir zusätzlich, dass gilt:

$$f(1_A) = 1_B.$$

Wir wiederholen die Konstruktion von Polynomalgebren.

Definition V.2.4. Es sei K ein kommutativer Ring mit 1.

(a) Als Menge ist $K[X]$ die Menge aller Abbildungen $f\colon \mathbb{N}_0 \to K$ mit der Eigenschaft, dass $f(n) = 0$ ist für fast alle $n \in \mathbb{N}_0$.
(b) Das Nullpolynom $0 \in K[X]$ ist die Abbildung $0\colon \mathbb{N}_0 \to K$ mit $0(n) = 0_K$ für alle $n \in \mathbb{N}_0$.
(c) Für $f, g \in K[X]$ sei $(f+g)(n) := f(n) + g(n)$ und für $\lambda \in K$ sei $(\lambda f)(n) := \lambda f(n)$.
(d) Für $f, g \in K[X]$ sei

$$(fg)(n) = \sum_{i=0}^{n} f(i) g(n-i)$$

und $1_{K[X]}$ sei die Abbildung

$$1_{K[X]}(n) = \begin{cases} 1_K, & n = 0 \\ 0_K, & n \neq 0. \end{cases}$$

Satz V.2.5. *Für $K \neq 0$ ist $K[X]$ eine unitäre und kommutative K-Algebra.*

Beweis. Sie haben schon gezeigt, dass $K[X]$ ein kommutativer Ring mit Eins ist, somit gilt (A1). Sie haben ebenfalls gezeigt, dass $K[X]$ ein K-Vektorraum ist; das impliziert (A2). Wir rechnen (A3) nach: Es seien also $f, g \in K[X]$ und $\lambda, \mu \in K$. Dann gilt $(\lambda f)(i) = \lambda f(i)$ und ebenso für g. Da K kommutativ ist, erhalten wir:

$$\begin{aligned}(\lambda f)(\mu g)(n) &= \sum_{i=0}^{n} (\lambda f)(i)(\mu g)(n-i)\\ &= \lambda\mu \sum_{i=0}^{n} f(i) g(n-i)\\ &= \lambda\mu (fg)(n).\end{aligned}$$

□

Wir benutzen wieder die Schreibweise, bei der wir $f\colon \mathbb{N}_0 \to K$ mit der Folge $(f(0), f(1), \ldots)$ identifizieren und X für die Folge $(0, 1, 0, \ldots)$ schreiben. Dann ist $X^i \cdot X^j = X^{i+j}$ für alle $i, j \in \mathbb{N}_0$. Ist $f \in K[X]$ ein beliebiges Element mit $f(n) = a_n$, so entspricht f der Folge $(a_0, a_1, \ldots)$, die wir ausdrücken können als

$$a_0 \cdot X^0 + a_1 \cdot X + \ldots + a_N \cdot X^N,$$

wenn N maximal ist, mit $a_N \neq 0$.

Definition V.2.6. Es sei K ein kommutativer Ring mit 1.

(a) Ist $f \in K[X]$, $f = (a_0, a_1, a_2, \ldots)$, so heißt
- a_i der *Koeffizient von* X^i,
- N der *Grad von* f, $\mathrm{Grad}(f)$, wenn N maximal ist mit $a_N \neq 0$; a_N heißt dann der *Höchstkoeffizient von* f.
- Das Element f heißt *normiert*, falls sein Höchstkoeffizient 1 ist.

(b) Das Nullpolynom hat Grad $-\infty$.

(c) Die K-Algebra $K[X]$ heißt die *Polynomalgebra über* K und X heißt *Unbestimmte*.

Satz V.2.7. *Sind $f, g \in K[X]$, so gilt*

(a) $\mathrm{Grad}(f + g) \leqslant \max(\mathrm{Grad}(f), \mathrm{Grad}(g))$,

(b) $\mathrm{Grad}(fg) \leqslant \mathrm{Grad}(f) + \mathrm{Grad}(g)$. *Ist K nullteilerfrei, so ist* $\mathrm{Grad}(fg) = \mathrm{Grad}(f) + \mathrm{Grad}(g)$.

Beispiel V.2.8. Ein drastisches Beispiel dafür, dass in (a) nur eine Abschätzung gilt, ist $\mathrm{Grad}(f + (-f)) = \mathrm{Grad}(0) = -\infty$. Ist $K = \mathbb{Z}/6\mathbb{Z}$, und $f = \overline{3}$, $g = \overline{2}X^3 + \overline{2}$, so erhalten wir

$$\mathrm{Grad}(fg) = \mathrm{Grad}(\overline{3}(\overline{2}X^3 + \overline{2})) = \mathrm{Grad}(0) = -\infty.$$

Daher gilt für kommutative Ringe mit Nullteilern bei (b) auch im Allgemeinen nur eine Abschätzung.

Beweis. Ist $f = 0$ (oder $g = 0$), so ist $f + g = g$ (oder $f + g = f$) und $fg = 0$. In diesem Fall erhalten wir dann (a) und (b) in der Form

$$\mathrm{Grad}(f + g) = \max(\mathrm{Grad}(f), \mathrm{Grad}(g)), \text{ und } \mathrm{Grad}(fg) = -\infty.$$

Wir nehmen also an, dass

$$f = \sum_{i=0}^{N} a_i X^i \neq 0 \text{ und } g = \sum_{j=0}^{M} b_j X^j \neq 0$$

und $N = \mathrm{Grad}(f)$, $M = \mathrm{Grad}(g)$. Die Abschätzung für (a) können Sie direkt ablesen. Für das Produkt ergibt sich

$$fg = \sum_{k=0}^{N+M} c_k X^k \text{ mit } c_{N+M} = a_N b_M.$$

Hat K keine Nullteiler, so ist $c_{N+M} \neq 0$, weil $a_N \neq 0 \neq b_M$ und deshalb gilt Gleichheit. □

Weil nullteilerfreie kommutative Ringe mit 1 so wichtig sind, haben sie einen eigenen Namen.

Definition V.2.9. Ein *Integritätsbereich* ist ein nullteilerfreier kommutativer Ring mit 1, der nicht der Nullring ist.

Satz V.2.10. *Es sei B eine unitäre K-Algebra. Dann gibt es für jede Wahl von $b \in B$ genau einen K-Algebrahomomorphismus $\varphi_b\colon K[X] \to B$ mit $\varphi_b(X) = b$.*

Beweis. Zur Eindeutigkeit: Da $\varphi_b(X) = b$ ist, gilt für ein beliebiges $f = \sum_{i=0}^{N} a_i X^i$, dass

$$\varphi_b(f) = \sum_{i=0}^{N} a_i b^i$$

und damit ist φ_b eindeutig festgelegt durch die Forderung, dass es ein K-Algbrahomomorphismus ist. Das impliziert unter anderem $\varphi_b(1_{K[X]}) = 1_B$. □

Bemerkung V.2.11. Die Abbildung φ_b heißt auch der *zu $b \in B$ gehörende Einsetzungshomomorphismus.*

Ist $B = K$, so erhalten wir für jedes $\lambda \in K$

$$\varphi_\lambda(f) = f(\lambda).$$

Jedes $f \in K[X]$ definiert damit eine Polynomfunktion

$$\varphi_{(-)}(f)\colon \lambda \mapsto f(\lambda).$$

Die Zuordnung $f \mapsto \varphi_{(-)}(f)$ ist im Allgemeinen *nicht* injektiv.

Beispiel V.2.12. Es sei $K = \mathbb{F}_p$ der Körper mit p Elementen für eine Primzahl p. Das Polynom $f = X^p - X$ ist nicht das Nullpolynom, aber für alle $\overline{a} \in \mathbb{F}_p$ gilt:

$$\varphi_{\overline{a}}(f) = \overline{a}^p - \overline{a} = \overline{a} - \overline{a} = \overline{0},$$

weil für alle $\overline{a} \in \mathbb{F}_p$ gilt, dass $\overline{a}^p = \overline{a}$. Also ist die zugehörige Polynomfunktion die Nullabbildung.

Satz V.2.13 (Polynomdivision mit Rest)**.** *Es sei K ein Integritätsbereich. Ist $g \in K[X]$, $g \neq 0_{K[X]}$, so dass der Höchstkoeffizient von g invertierbar ist in K, so gibt es zu jedem Polynom $f \in K[X]$ eindeutig bestimmte Polynome $q, r \in K[X]$, so dass*

$$f = qg + r \textit{ mit } \mathrm{Grad}(r) < \mathrm{Grad}(g).$$

Beweis. Wir betrachten zunächst den Fall, dass $f = 0$ ist oder dass $f \neq 0$ aber $\mathrm{Grad}(f) < \mathrm{Grad}(g)$ ist. In diesem Fall setzen wir $q = 0$ und $r = f$.

Es sei also $f \neq 0$ und $M := \mathrm{Grad}(f) \geqslant \mathrm{Grad}(g) =: N$. Wir beweisen die Behauptung durch absteigende Induktion nach M. Für $M = N$ korrigiert q den Höchstkoeffizienten und r sammelt die abweichenden Anteile niedrigeren Grades auf.

Es sei also $M > N$ und wir schreiben $f = \sum_{i=0}^M a_i X^i$ und $g = \sum_{j=0}^N b_j X^j$ mit $a_M \neq 0 \neq b_N$. Nach Voraussetzung ist b_N invertierbar in K.

Wir definieren

$$h := f - a_M b_N^{-1} X^{M-N} g.$$

Nach Konstruktion ist der Koeffizient bei X^M für h gerade $a_M - a_M b_N^{-1} b_N = 0$ und damit ist der Grad von h echt kleiner als M. Nach Induktionsvoraussetzung gibt es also $q_0, r \in K[X]$ mit $\mathrm{Grad}(r) < N$ und $h = q_0 g + r$. Damit können wir auch f zerlegen:

$$f = h + a_M b_N^{-1} X^{M-N} g = (q_0 + a_M b_N^{-1} X^{M-N}) g + r$$

und wir setzen $q = q_0 + a_M b_N^{-1} X^{M-N}$. Das zeigt die Existenz der Zerlegung.

Zur Eindeutigkeit: Ist $f = qg + r = pg + s$, so dass $\mathrm{Grad}(r), \mathrm{Grad}(s) < \mathrm{Grad}(g)$ gilt, so ist

$$(q - p)g = s - r.$$

Es gilt aber $\mathrm{Grad}((q - p)g) = \mathrm{Grad}(q - p) + \mathrm{Grad}(g)$, während $\mathrm{Grad}(s - r) < \mathrm{Grad}(g)$ ist. Das geht nur, wenn $q - p = 0$ ist und damit ist auch $s - r = 0$, also sind $q = p$ und $r = s$. □

Korollar V.2.14. *Ist K ein Integritätsbereich, $f \in K[X]$ und ist $\lambda \in K$ eine Nullstelle von f (also $\varphi_\lambda(f) = f(\lambda) = 0$), so gibt es genau ein Polynom $g \in K[X]$ mit*

$$f = (X - \lambda)g \in K[X]$$

und $\mathrm{Grad}(g) = \mathrm{Grad}(f) - 1$.

Beweis. Polynomdivision mit Rest liefert

$$f = q(X - \lambda) + r$$

mit $\mathrm{Grad}(r) < 1$, aber $\varphi_\lambda(f) = 0 = \varphi_\lambda(q(X - \lambda) + r) = r(\lambda)$. Also muss $r(\lambda) = 0$ gelten, aber r ist konstant, so dass $r = 0$ ist. □

Korollar V.2.15. *Ist K ein Integritätsbereich, so besitzt ein $f \in K[X] \setminus \{0\}$, höchstens* $\mathrm{Grad}(f)$ *viele Nullstellen.*

Beweis. Wir beweisen die Behauptung durch Induktion über den Grad von f. Ist $\mathrm{Grad}(f) = 0$, so hat f gar keine Nullstellen.

Ist $\lambda \in K$ eine Nullstelle, so zerlegen wir f wie eben als $f = (X - \lambda)g$ mit $\mathrm{Grad}(g) = \mathrm{Grad}(f) - 1$. Nach Induktionsvoraussetzung hat g dann höchstens $\mathrm{Grad}(f)-1$ viele Nullstellen und daher hat f höchstens $\mathrm{Grad}(f)$ viele Nullstellen. $\square$

Korollar V.2.16. *Hat ein Integritätsbereich K unendlich viele Elemente, so ist die Auswertungsabbildung*

$$K[X] \to \mathrm{Abb}(K, K), \quad f \mapsto \varphi_{(-)}(f) = f(-)$$

injektiv.

Beweis. Wäre $\varphi_{(-)}(f)$ die Nullabbildung für $f \neq 0$, so hätte f unendlich viele Nullstellen. $\square$

Das Korollar gilt insbesondere für $K = \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ oder $\mathbb{C}$.

Definition V.2.17. Ist K ein Integritätsbereich und $0 \neq f \in K[X]$, so heißt

$$\mu(f; \lambda) := \max\{r \in \mathbb{N}_0, \exists g \in K[X], f = (X - \lambda)^r g\}$$

die *Vielfachheit der Nullstelle λ von f*.

Bemerkung V.2.18. Ist $\mu(f; \lambda) = 0$, so ist $f = g$, $r = 0$ und damit $f(\lambda) \neq 0$, also ist λ in diesem Fall gar keine Nullstelle von f. Es gilt immer $\mu(f; \lambda) \leqslant \mathrm{Grad}(f)$.

Durch Induktion nach dem Grad von f beweisen Sie direkt mit Korollar V.2.14 die folgende Zerlegung.

Lemma V.2.19. *Ist K ein Integritätsbereich und $0 \neq f \in K[X]$, so dass f mindestens eine Nullstelle in K besitzt. so gibt es $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_i \neq \lambda_j$ für $i \neq j$, $n_1, \ldots, n_m \in \mathbb{N}$ und ein $g \in K[X]$ ohne Nullstelle in K, so dass*

$$f = (X - \lambda_1)^{n_1} \cdot \ldots \cdot (X - \lambda_m)^{n_m} g.$$

Die λ_i, n_i, g sind bis auf Reihenfolge eindeutig.

Definition V.2.20. Für einen Integritätsbereich K sagt man, dass ein $f \in K[X]$ *in Linearfaktoren zerfällt*, falls

$$f = a(X - \lambda_1)^{n_1} \cdot \ldots \cdot (X - \lambda_m)^{n_m}$$

gilt mit $a, \lambda_1, \ldots, \lambda_m \in K$ und $n_i \in \mathbb{N}$.

Beispiel V.2.21. Wir betrachten die Drehung um θ im $\mathbb{R}^2$ und hatten

$$P_{R_\theta}(X) = \det \begin{pmatrix} \cos\theta - X & -\sin\theta \\ \cos\theta & \sin\theta - X \end{pmatrix} = X^2 - 2\cos\theta X + 1.$$

Dann zerfällt $P_{R_\theta}(X)$ für alle θ über $\mathbb{C}$ in Linearfaktoren

$$P_{R_\theta}(X) = (X - e^{i\theta})(X - e^{-i\theta})$$

aber nur für $\theta = 0, \pi$ zerfällt $P_{R_\theta}(X)$ auch über $\mathbb{R}$ in Linearfaktoren als

$$P_{R_0}(X) = (X - 1)^2, \text{ und } P_{R_\pi}(X) = (X + 1)^2.$$

Bemerkung V.2.22. In der Funktionentheorie und der Topologie lernen Sie einen Beweis des **Fundamentalsatzes der Algebra** kennen:

Ist $f \in \mathbb{C}[X]$ mit $\mathrm{Grad}(f) \geqslant 1$, so hat f mindestens eine Nullstelle in $\mathbb{C}$. Damit zerfällt jedes solche $f \in \mathbb{C}[X]$ in Linearfaktoren.

V.3. Diagonalisierbarkeit

Ziel dieses Abschnitts ist es, ein hinreichendes und notwendiges Kriterium dafür zu entwickeln, dass ein Endomorphismus eines endlich-dimensionalen Vektorraums diagonalisierbar ist. Wir betrachten ab jetzt wieder K als Körper.

Bemerkung V.3.1. Ist ein $A \in M(n \times n, K)$ diagonalisierbar, so ist

$$P_A(X) = \det \begin{pmatrix} \lambda_1 - X & 0 & \dots & 0 \\ 0 & \lambda_2 - X & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n - X \end{pmatrix} = \prod_{i=1}^{n} (-1)^n (X - \lambda_i)$$

und P_A zerfällt also in Linearfaktoren. Die Umkehrung gilt allerdings nicht. Wir hatten schon für $B = \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix}$ gesehen, dass zwar $P_B(X) = (X-2)^2$, aber B war nicht diagonalisierbar.

Definition V.3.2. Ist K ein Körper und ist $\lambda \in K$ ein Eigenwert für $A \in M(n \times n, K)$, so heißt die Vielfachheit der Nullstelle λ des charakteristischen Polynoms die *algebraische Vielfachheit des Eigenwerts λ von A*:

$$\mu_{\text{alg}}(A; \lambda) := \mu(P_A(X); \lambda).$$

Ist $f \in \mathsf{End}_K(V)$ und ist V endlich-dimensional, so ist $\mu_{\text{alg}}(f; \lambda) := \mu(P_f(X); \lambda)$ und heißt die *algebraische Vielfachheit des Eigenwerts λ von f.*

Beispiel V.3.3. Für $\lambda \neq 0$ sei $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. Dann ist $\mu_{\text{alg}}(A; \lambda) = 2$, weil $P_A(X) = (\lambda - X)^2$, aber $\mu_{\text{geo}}(A; \lambda) = 1$ mit μ_{geo} wie in Definition V.1.13.

Lemma V.3.4. *Ist $A \in M(n \times n, K)$ und ist $\lambda \in K$ ein Eigenwert von A, so gilt*

$$1 \leqslant \mu_{geo}(A; \lambda) \leqslant \mu_{alg}(A; \lambda) \leqslant n.$$

Beweis. Ist $m = \mu_{\text{geo}}(A; \lambda)$, so gibt es eine Basis $(v_1, \dots, v_m)$ von $\text{Eig}(A; \lambda)$. Wir ergänzen diese zu einer Basis $(v_1, \dots, v_n)$ von K^n. Mit $S^{-1} = (v_1, \dots, v_n)$ ist

$$SAS^{-1} = \begin{pmatrix} \lambda E_m & B \\ 0 & C \end{pmatrix}$$

mit geeigneten Matrizen B, C. Mit Korollar IV.4.9 folgt, dass

$$P_A(X) = (\lambda - X)^m \det(C - \lambda E_{n-m})$$

und somit ist $\mu_{\text{alg}}(A; \lambda) \geqslant m = \mu_{\text{geo}}(A; \lambda)$. □

Satz V.3.5 (Kriterium für Diagonalisierbarkeit)**.** *Es sei V ein K-Vektorraum mit $\dim_K(V) = n$ und $f \in \mathsf{End}_K(V)$. Es seien $\lambda_1, \dots, \lambda_m$ die paarweise verschiedenen Eigenwerte von f. Dann sind äquivalent:*

(a) *f ist diagonalisierbar,*

(b) *$P_f(X)$ zerfällt in Linearfaktoren und $\mu_{geo}(f; \lambda_i) = \mu_{alg}(f; \lambda_i)$ für alle $1 \leqslant i \leqslant m$,*

(c)
$$\sum_{i=1}^{m} \mu_{geo}(f; \lambda_i) = n,$$

(d)
$$V = \bigoplus_{i=1}^{m} \text{Eig}(f; \lambda_i),$$

das heißt, dass sich jedes $v \in V$ eindeutig schreiben läßt als $v = u_1 + \dots + u_m$ mit $u_i \in \text{Eig}(f; \lambda_i)$ für $1 \leqslant i \leqslant m$.

Definition V.3.6. Die Zerlegung von V wie in (d) heißt die *Eigenraumzerlegung von V.*

Beweis. (a) $\Rightarrow$ (b): Wir wählen eine Basis $\mathcal{B}$ aus Eigenvektoren $v_1, \ldots, v_{k_1}$ von λ_1, $v_{k_1+1}, \ldots, v_{k_1+k_2}$ von λ_2, … und $v_{k_1+\ldots,k_{m-1}+1}, \ldots, v_{k_1+\ldots+k_m}$ von λ_m. Dann ist $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ von Diagonalgestalt mit

$$A = \begin{pmatrix} \lambda_1 E_{k_1} & 0 & \ldots & 0 \\ 0 & \lambda_2 E_{k_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_m E_{k_m} \end{pmatrix}$$

und $P_A(X) = \prod_{i=1}^m (\lambda_i - X)^{k_i}$, so dass für alle $1 \leqslant i \leqslant m$ gilt, dass

$$\mu_{\text{geo}}(A; \lambda_i) = \mu_{\text{alg}}(A; \lambda_i).$$

(b) $\Rightarrow$ (c): Zerfällt P_f in Linearfaktoren $P_f(X) = \prod_{i=1}^m (\lambda_i - X)^{k_i}$, so gilt $n = \sum_{i=1}^m k_i$. Aber da $k_i = \mu_{\text{alg}}(A; \lambda_i) = \mu_{\text{geo}}(A; \lambda_i)$ ist für alle i, ist auch

$$n = \sum_{i=1}^m k_i = \sum_{i=1}^m \mu_{\text{geo}}(A; \lambda).$$

(c) $\Rightarrow$ (d): Setze $U = \operatorname{Eig}(f; \lambda_1) + \ldots + \operatorname{Eig}(f; \lambda_m)$. Aber Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. Damit ist U die direkte Summe der Eigenräume und für die Dimension von U erhalten wir

$$\dim_K U = \sum_{i=1}^m \dim_K \operatorname{Eig}(f; \lambda_i) = \sum_{i=1}^m \mu_{\text{geo}}(f; \lambda_i) = n = \dim_K V.$$

Also ist $U = V$.

(d) $\Rightarrow$ (a): Wähle Basen für alle Eigenräume $\operatorname{Eig}(f; \lambda_i)$. Die Vereinigung dieser Basen ist eine Basis von V aus Eigenvektoren. $\square$

Bemerkung V.3.7. Eine Matrix $A \in M(n \times n, K)$ ist damit genau dann diagonalisierbar, falls $P_A(X)$ in Linearfaktoren zerfällt und für alle Eigenwerte λ von A gilt, dass $\mu_{\text{geo}}(A; \lambda) = \mu_{\text{alg}}(A; \lambda)$.

V.4. Triagonalisierbarkeit

Wir haben eben gesehen, dass nicht alle quadratischen Matrizen diagonalisierbar sind. Wir untersuchen jetzt, wann eine Matrix zumindest ähnlich ist zu einer oberen Dreiecksmatrix.

In diesem Abschnitt ist K ein festgewählter Körper und V ist ein K-Vektorraum der Dimension $n < \infty$.

Definition V.4.1. Es sei $f \in \mathsf{End}_K(V)$. Dann heißt ein Untervektorraum $U \subset V$ *f-invariant*, falls $f(U) \subset U$.

Beispiele V.4.2.

- $\{0_V\} \subset V$ ist f-invariant für alle $f \in \mathsf{End}_K(V)$.
- V selbst ist natürlich auch f-invariant für alle $f \in \mathsf{End}_K(V)$.
- Sind U_1, U_2 f-invariant, so auch $U_1 \cap U_2$ und $U_1 + U_2$.
- Jeder Eigenraum von f, $U = \operatorname{Eig}(f; \lambda)$, ist f-invariant.

Ist $U \subset V$ ein f-invarianter Untervektorraum, so können wir f auf U einschränken und erhalten $f|_U \in \mathsf{End}_K(U)$.

Lemma V.4.3. *Ist $f \in \mathsf{End}_K(V)$ und ist U f-invariant, so teilt $P_{f|_U}(X)$ das Polynom $P_f(X)$.*

Beweis. Es sei $\mathcal{A}$ eine geordnete Basis von U und wir ergänzen sie zu einer geordneten Basis $\mathcal{B}$ von V. Dann hat $M_{\mathcal{B}}^{\mathcal{B}}(f)$ die Form

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} M_{\mathcal{A}}^{\mathcal{A}}(f|_U) & D \\ 0 & C \end{pmatrix}$$

und

$$P_f(X) = \det(M_{\mathcal{A}}^{\mathcal{A}}(f|_U) - X E_{\dim_K U}) \det(C - X E_{n-\dim_K U}) = P_{f|_U} \det(C - X E_{n-\dim_K U}).$$

$\square$

Definition V.4.4.

(a) Eine *Fahne* von V ist eine Kette von Untervektorräumen

$$\{0_V\} = V_0 \subset V_1 \subset \ldots \subset V_n = V, \tag{V.4.1}$$

so dass für alle $0 \leqslant i \leqslant n$ gilt: $\dim_K V_i = i$.

(b) Eine Fahne wie in (V.4.1) heißt *f-invariant* für $f \in \mathsf{End}_K(V)$, falls für alle $0 \leqslant i \leqslant n$ gilt: $f(V_i) \subset V_i$.

Bemerkung V.4.5. Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V, so ist $V_i = \mathsf{Span}_K(v_1, \ldots, v_i)$ nicht immer f-invariant. Betrachten Sie zum Beispiel $V = \mathbb{R}^2$ mit der Standardbasis und $f(e_1) = e_2, f(e_2) = e_1$.

Satz V.4.6. *Ist $f \in \mathsf{End}_K(V)$, so sind äquivalent:*

(a) *Es gibt eine f-invariante Fahne von V.*
(b) *Es gibt eine geordnete Basis $\mathcal{B}$ von V, so dass $M_{\mathcal{B}}^{\mathcal{B}}(f)$ eine obere Dreiecksmatrix ist.*

Beweis. (a) $\Rightarrow$ (b): Es sei

$$\{0_V\} = V_0 \subset V_1 \subset \ldots \subset V_n = V$$

eine f-invariante Fahne und es sei (v_1) eine Basis von V_1. Wir können sie ergänzen zu einer Basis (v_1, v_2) von V_2 und iterativ erhalten wir für jedes V_i eine Basis $(v_1, \ldots, v_i)$ bis hin zu $i = n$ und $\mathcal{B} = (v_1, \ldots, v_n)$. Damit ist $M_{\mathcal{B}}^{\mathcal{B}}(f)$ eine obere Dreiecksmatrix.

(b) $\Rightarrow$ (a): $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ sei eine obere Dreiecksmatrix für $\mathcal{B} = (v_1, \ldots, v_n)$. Wir schreiben wieder $A = (a_{ij})$ und setzen $V_i := \mathsf{Span}_K(v_1, \ldots, v_i)$. Es gilt

$$f(v_i) = \sum_{j=1}^{n} a_{ji} v_j = \sum_{j=1}^{i} a_{ji} v_j,$$

weil A obere Dreiecksmatrix ist und somit ist jedes V_i f-invariant mit Dimension i. □

Definition V.4.7. Ein $f \in \mathsf{End}_K(V)$ heißt *triagonalisierbar*, falls es eine geordnete Basis $\mathcal{B}$ von V gibt, so dass $M_{\mathcal{B}}^{\mathcal{B}}(f)$ eine obere Dreiecksmatrix ist. Eine Matrix $A \in M(n \times n, K)$ heißt *triagonalisierbar*, falls sie ähnlich ist zu einer oberen Dreiecksmatrix.

Satz V.4.8 (Triagonalisierbarkeitskriterium)**.** *Es sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathsf{End}_K(V)$. Dann sind äquivalent:*

(a) *f ist triagonalisierbar,*
(b) *$P_f(X)$ zerfällt in Linearfaktoren.*

Beweis. (a) $\Rightarrow$ (b): Es sei $M_{\mathcal{B}}^{\mathcal{B}}(f)$ eine obere Dreiecksmatrix mit Diagonale $(\lambda_1, \ldots, \lambda_n)$, dann ist $P_f(X) = \prod_{i=1}^{n}(\lambda_i - X)$.

(b) $\Rightarrow$ (a): Wir machen Induktion über $n = \dim_K(V)$. Für $n = 1$ ist nichts zu zeigen.

Wir nehmen an, dass $P_f(X)$ zerfällt als $P_f(X) = \prod_{i=1}^{n}(\lambda_i - X)$ für ein $n > 1$. Es sei $0 \neq v_1 \in \mathrm{Eig}(f; \lambda_1)$. Wir ergänzen (v_1) zu einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V. Dann gilt:

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{A} \end{pmatrix}$$

und $P_f(X) = (\lambda_1 - X) P_{\tilde{A}}(X)$. Also zerfällt $P_{\tilde{A}}(X)$ in Linearfaktoren und nach Induktionsvoraussetzung gibt es ein $\tilde{S} \in GL_{n-1}(K)$, so dass $\tilde{S}\tilde{A}\tilde{S}^{-1} =: \tilde{D}$ eine obere Dreiecksmatrix ist. Wir setzen

$$S := \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S} \end{pmatrix}.$$

Dann ist $S^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}^{-1} \end{pmatrix}$ und

$$\begin{aligned} SM_{\mathcal{B}}^{\mathcal{B}}(f)S^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{S}\tilde{A}\tilde{S}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{D} \end{pmatrix} \end{aligned}$$

und dies ist eine obere Dreiecksmatrix. □

Benutzen wir den Fundamentalsatz der Algebra, so erhalten wir:

Korollar V.4.9. *Jede Matrix $A \in M(n \times n, \mathbb{C})$ ist triagonalisierbar.*

Beispiel V.4.10. Es sei $A = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix} \in M(2 \times 2, \mathbb{R})$. Dann ist

$$P_A(X) = (1 - X)(3 - X) + 1 = X^2 - 4X + 4 = (X - 2)^2.$$

Also muss A triagonalisierbar sein. Das LGS

$$\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

ergibt $x = -y$, so dass $\mathrm{Eig}(A; 2) = \mathsf{Span}_{\mathbb{R}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Der Vektor e_2 erfüllt $Ae_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, also können wir $S^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ setzen und bekommen

$$SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}.$$

V.5. Das Minimalpolynom

Wir hatten für jede unitäre K-Algebra B die Abbildung

$$K[X] \to \mathrm{Abb}(B, B), \quad f \mapsto (b \mapsto \varphi_b(f) = f(b))$$

kennengelernt. Wir ordnen sie in den folgenden Kontext ein:

Definition V.5.1.

(a) Ist R ein Ring, so heißt eine Teilmenge $I \subset R$ ein *Linksideal* (beziehungsweise ein *Rechtsideal*), falls gilt:
  (I1) $(I, +) \subset (R, +)$ ist eine Untergruppe.
  (I2) Für alle $r \in R$ und alle $x \in I$ ist $rx \in I$ (beziehungsweise $xr \in I$).

(b) $I \subset R$ heißt *(beidseitiges) Ideal*, falls es sowohl Rechts- als auch Linksideal von R ist.

(c) Für eine K-Algebra B über einem Körper K ist ein Ideal $I \subset B$ ein Untervektorraum, der zusätzlich (I2) erfüllt.

(d) Ist K ein Körper und B eine unitäre K-Algebra, so ist

$$\mathrm{Ann}(b) := \{f \in K[X], \varphi_b(f) = f(b) = 0\}$$

der *Annihilator von b*.

Lemma V.5.2. *Der Annihilator $\mathrm{Ann}(b) \subset K[X]$ ist ein Ideal für alle $b \in B$.*

Beweis. Sind $f, g \in \mathrm{Ann}(b)$, also $f(b) = g(b) = 0$ und sind $\lambda, \mu \in K$, so ist auch

$$(\lambda f + \mu g)(b) = \lambda f(b) + \mu g(b) = 0$$

und somit ist $\mathrm{Ann}(b) \subset K[X]$ ein Untervektorraum.

Ist $f \in \operatorname{Ann}(b)$ und ist $h \in K[X]$ beliebig, so gilt

$$(fh)(b) = f(b)h(b) = 0 \cdot h(b) = 0 = h(b) \cdot 0 = h(b)f(b).$$

Also ist sowohl fh also auch hf ein Element des Annihilators. □

Beispiele V.5.3. Wir bestimmen im Folgenden Annihilatoren zu $B = M(n \times n, K)$. Hierbei ist die Auswertung eines Polynoms auf einer Matrix so zu verstehen, dass wir für ein $A \in M(n \times n, K)$ und ein $f \in K[X]$ mit $f = a_m X^m + \ldots + a_1 X + a_0$ die Matrix $f(A)$ als $a_m A^m + \ldots + a_1 A + a_0 E_n$ definieren.

Der Annihilator der Einheitsmatrix ist

$$\operatorname{Ann}(E_n) = \{f \in K[X], f(E_n) = 0\}.$$

Dies sind also alle Polynome $f = a_m X^m + \ldots + a_1 X + a_0$ mit

$$a_m E_n^m + \ldots + a_1 E_n + a_0 E_n = 0.$$

Dies können wir zusammenfassen zu der Bedingung

$$(a_m + \ldots + a_1 + a_0)E_n = 0.$$

Dies kann nur wahr sein, wenn $a_m + \ldots + a_1 + a_0 = 0$ ist, aber dies ist genau $f(1)$. Damit gilt:

$$f \in \operatorname{Ann}(E_n) \Leftrightarrow f(1) = 0 \Leftrightarrow f = (X-1)g \text{ für ein } g \in K[X].$$

Betrachten wir dagegen die $n \times n$-Nullmatrix 0_n, so ist

$$\operatorname{Ann}(0_n) = \{f \in K[X],\ f(0_n) = 0_n\}.$$

Schreiben wir f wieder wie oben, so ist $f(0_n) = a_0 E_n$, also ist

$$\operatorname{Ann}(0_n) := \{f \in K[X],\ f \text{ hat keinen konstanten Term}\}.$$

Diese Polynome können wir schreiben als $X \cdot g$ mit $g \in K[X]$, weil sie 0 als Nullstelle haben.

Für $n = 2$ betrachten wir die Matrix $\psi = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Es gilt $\psi^2 = 0$. Setzen wir ψ in $f = a_m X^m + \ldots + a_1 X + a_0$ ein, so bleibt nur $f(\psi) = a_1\psi + a_0 E_2$ übrig, also die Matrix $\begin{pmatrix} a_0 & a_1 \\ 0 & a_0 \end{pmatrix}$. Also sind alle Polynome im Annihilator, die weder einen nicht-trivialen konstanten Term noch eine nicht-trivialen linearen Term haben.

Satz V.5.4. *Ist K ein Körper und ist $0 \neq I \subset K[X]$ ein Ideal, so gibt es ein eindeutig bestimmtes normiertes $h \in K[X]$ mit*

$$I = \{f \in K[X], \exists g \in K[X] : f = gh\}.$$

Die übliche Notation ist $I = h \cdot K[X]$ oder kurz $I = (h)$ und solche Ideale heißen *Hauptideale*. Man sagt auch, dass $K[X]$ ein Hauptidealring ist.

Beweis. Wir zeigen erst die Eindeutigkeit: Es sei also $(h_2) = I = (h_1)$. Dann gilt für alle $0 \neq g \in I$, dass $g = h_1 f_1 = h_2 f_2$ mit $f_1, f_2 \in K[X]$. Also gilt für den Grad von g:

$$\operatorname{Grad}(g) = \operatorname{Grad}(h_i) + \operatorname{Grad}(f_i) \geqslant \operatorname{Grad}(h_i),$$

Das Element $h_1 - h_2$ ist ein Element von I, weil $(I, +)$ eine abelsche Gruppe ist. Mit $g = h_1 - h_2$ erhalten wir also

$$\operatorname{Grad}(h_1 - h_2) \geqslant \operatorname{Grad}(h_i),$$

falls $h_1 - h_2 \neq 0$ ist. Da aber h_1 und h_2 beide normiert sind, ist $\operatorname{Grad}(h_1 - h_2) < \operatorname{Grad}(h_i)$, also muss $h_1 - h_2 = 0$ sein und somit $h_1 = h_2$.

Wir zeigen die Existenz von h über die Polynomdivision mit Rest. Es sei $0 \neq h' \in I$ so gewählt, dass für alle $g \in I$ gilt: $\operatorname{Grad}(g) \geqslant \operatorname{Grad}(h')$. Wir schreiben h' wieder als

$$h' = a_n X^n + \ldots + a_1 X + a_0$$

mit $a_n \neq 0$. Da I ein Ideal ist, liegt auch $a_n^{-1} h' =: h$ in I.

Es ist klar, dass $hK[X] \subset I$, weil $h \in I$. Ist umgekehrt $f \in I \setminus \{0\}$, so können wir f schreiben als

$$f = qh + r$$

mit Grad$(r) <$ Grad(h). Dann ist aber $r = f - qh$ und $f - qh \in I$, weil h und f in I sind. Der Grad von h war aber minimal und daher muss $r = 0$ gelten. Wir können also jedes solche f als Vielfaches von h schreiben und damit ist auch $I \subset (h)$. □

Definition V.5.5. Es sei K ein Körper, B sei eine unitäre K-Algebra und $b \in B$. Das eindeutig bestimmte normierte Polynom $m_b \in K[X]$ mit $\mathrm{Ann}(b) = (m_b) \subset K[X]$ heißt das *Minimalpolynom von b*.

Speziell für $B = M(n \times n, K)$ erhalten wir für jede Matrix $A \in M(n \times n, K)$ das *Minimalpolynom von A, m_A*.

Arthur Cayley (1821–1895), William Rowan Hamilton (1805–1865)

Satz V.5.6 (Satz von Cayley-Hamilton)**.** *Ist $A \in M(n \times n, K)$, so gilt*

$$\varphi_A(P_A(X)) = P_A(A) = 0_n.$$

Also ist das charakteristische Polynom von A, $P_A(X)$, ein Element des Annihilators von A und wird daher vom Minimalpolynom von A, $m_A(X)$, geteilt.

Beweis. Wir betrachten für ein $0 \neq v \in K^n$ die Vektoren

$$v_0 = A^0 v = E_n v = v, v_1 = Av, v_2 = A^2 v, \ldots$$

also $v_i = A^i v$ für $i \in \mathbb{N}_0$. Da $\dim_K K^n = n$, gibt es ein $m \leqslant n$, so dass $(v_0, \ldots, v_{m-1})$ linear unabhängig ist, aber $(v_0, \ldots, v_{m-1}, v_m)$ ist linear abhängig. Es sei $U = \mathsf{Span}_K(v_0, \ldots, v_{m-1})$ und wir schreiben

$$v_m = -\alpha_0 v_0 - \ldots - \alpha_{m-1} v_{m-1}. \tag{V.5.1}$$

Damit ist die Darstellung von $A|_U$ bezüglich der Basis $(v_0, \ldots, v_{m-1})$ gleich

$$\begin{pmatrix} 0 & 0 & \ldots & \ldots & \ldots & 0 & -\alpha_0 \\ 1 & 0 & & & & \vdots & \vdots \\ 0 & 1 & 0 & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & 1 & -\alpha_{m-1} \end{pmatrix}.$$

Das charakteristische Polynom von $A|_U$ ist daher

$$P_{A|_U}(X) = \det(A|_U - XE_m) = \det \begin{pmatrix} -X & 0 & \ldots & \ldots & \ldots & 0 & -\alpha_0 \\ 1 & -X & & & & \vdots & \vdots \\ 0 & 1 & -X & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & -X & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & 1 & -\alpha_{m-1} - X \end{pmatrix}.$$

Wir entwickeln die Determinante nach der letzten Spalte und erhalten

$$(-1)^{m+2}\alpha_0 \det \begin{pmatrix} 1 & -X & 0 & \dots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -X \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} + (-1)^{m+3}\alpha_1 \det \begin{pmatrix} -X & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & -X & & & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & -X \\ 0 & 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

$$+ (-1)^{2m}(-\alpha_{m-1} - X) \det \begin{pmatrix} -X & 0 & \dots & \dots & \dots & 0 \\ 1 & -X & & & & \vdots \\ 0 & 1 & -X & & & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 & -X \end{pmatrix}$$

$$= (-1)^m(\alpha_0 + \alpha_1 X + \ldots + \alpha_{m-1}X^{m-1} + X^m).$$

Da U A-invariant ist, gibt es ein $q \in K[X]$ mit $P_A(X) = q(X)P_{A|_U}(X)$. Wir wenden diese Gleichung auf v an und erhalten

$$P_A(A)v = q(A)P_{A|_U}(A)v = q(A)((-1)^m(\alpha_0 v_0 + \alpha_1 v_1 + \ldots + \alpha_{m-1}v_{m-1} + v_m)).$$

Da aber $\alpha_0 v_0 + \alpha_1 v_1 + \ldots + \alpha_{m-1}v_{m-1} + v_m = 0$ gilt, ist dieser Ausdruck trivial. Da wir ein beliebiges $v \neq 0$ gewählt hatten, folgt damit dass $P_A(A) = 0_n \in M(n \times n, K)$ ist. □

Korollar V.5.7. *Für jedes $A \in M(n \times n, K)$ haben $m_A(X)$ und $P_A(X)$ übereinstimmende Nullstellen und für alle $\lambda \in K$ gilt*

$$\mu(m_A; \lambda) \leqslant \mu(P_A; \lambda).$$

Beweis. Nach Cayley-Hamilton gilt

$$P_A(X) = q m_A(X) \tag{V.5.2}$$

für ein $q \in K[X]$. Somit ist jede Nullstelle von m_A auch eine Nullstelle von P_A. Ist umgekehrt λ eine Nullstelle von P_A, ist also λ ein Eigenwert von A, so gibt es ein $0 \neq v \in K^n$ mit $Av = \lambda v$. Wir schreiben

$$m_A(X) = X^m + \alpha_{m-1}X^{m-1} + \ldots + \alpha_1 X + \alpha_0.$$

Da $m_A(A) = 0_n$ ist, ist auch

$$0 = 0_n v = m_A(A)v = ((A^m + \alpha_{m-1}A^{m-1} + \ldots + \alpha_1 A + \alpha_0)E_n)v = \lambda^m v + \alpha_{m-1}\lambda^{m-1}v + \ldots + \alpha_0 v = m_A(\lambda)v.$$

Da aber $v \neq 0$ vorausgesetzt war, ist $m_A(\lambda) = 0$. Die Abschätzung über die Vielfachheit der Nullstellen folgt direkt aus (V.5.2). □

Satz V.5.8. *Es sei $A \in M(n \times n, K)$. Ist A diagonalisierbar, so zerfällt m_A in paarweise verschiedene Linearfaktoren.*

Bemerkung V.5.9. Es gilt auch die Umkehrung, aber den Beweis dafür sehen wir erst später.

Beweis. Es sei A diagonalisierbar und es seien $\lambda_1, \ldots, \lambda_m$ die paarweise verschiedenen Eigenwerte von A, also gilt $m \leqslant n$. Wir setzen

$$g_A(X) = \prod_{i=1}^{m}(X - \lambda_i) \in K[X].$$

Ist $g_A(\mu) = 0$, so ist auch $P_A(\mu) = 0$, weil μ dann einer der Eigenwerte sein muss. Damit muss g_A das Minimalpolynom m_A teilen.

Da A diagonalisierbar ist, gilt

$$K^n = \bigoplus_{i=1}^{m} \mathrm{Eig}(A; \lambda_i).$$

Wir schreiben ein beliebiges $0 \neq v$ eindeutig als

$$v = v_1 + \ldots + v_m$$

mit $v_i \in \mathrm{Eig}(A; \lambda_i)$. Damit ist dann $g_A(A)v_i = 0$ für alle i, weil

$$g_A(A)v_i = \left(\prod_{\substack{j=1 \\ j \neq i}}^{m} (A - \lambda_j E_n) \right) \cdot \underbrace{(A - \lambda_i E_n)v_i}_{=0}.$$

Damit ist $g_A(A)v$ ebenfalls trivial. Da v beliebig gewählt war, folgt, dass $g_A(A) = 0_n$ ist, aber damit ist g_A ein Vielfaches von m_A. Zusammen erhalten wir $g_A = m_A$. $\square$

KAPITEL VI

Normalformen

Das Ziel dieses Kapitels ist es, Ähnlichkeitsklassen quadratischer Matrizen zu beschreiben.

VI.1. Charakteristische Matrizen

Im Folgenden betrachten wir Matrizen über $K[X]$. Allgemeiner sei R ein kommutativer Ring mit Eins, zum Beispiel $\mathbb{Z}$, $\mathbb{Z}/6\mathbb{Z}$ oder $K[X]$.

Definition VI.1.1. Zwei Matrizen $A, B \in M(n \times n, R)$ heißen *äquivalent* über R, falls es invertierbare Matrizen $P, Q \in M(n \times n; R)$ gibt, so dass $B = P^{-1}AQ$.

Sie heißen *ähnlich* über R, falls es eine invertierbare Matrix $S \in GL_n(R)$ gibt mit $B = S^{-1}AS$.

Definition VI.1.2. Für einen Körper K und $A \in M(n \times n, K)$ sei $M_A(X) := A - XE_n \in M(n \times n, K[X])$ die *charakteristische Matrix von A*.

Das charakteristische Polynom $P_A(X)$ ist also die Determinante der charakteristischen Matrix:

$$P_A(X) = \det(A - XE_n) = \det M_A(X) \in K[X].$$

Ferdinand Georg Frobenius (1849–1917)

Satz VI.1.3 (Satz von Frobenius)**.** *Es sei K ein Körper. Zwei Matrizen $A, B \in M(n \times n, K)$ sind genau dann ähnlich über K, wenn die charakteristischen Matrizen $M_A(X)$ und $M_B(X)$ äquivalent sind über $K[X]$.*

Wir können also anstelle des Ähnlichkeitsproblems für Matrizen über K das Äquivalenzproblem über $K[X]$ untersuchen.

Beweis. Die eine Richtung der Äquivalenz ist einfach: Ist $B = S^{-1}AS$ mit $S \in GL_n(K)$, so ist

$$M_B(X) = B - XE_n = S^{-1}AS - S^{-1}E_nSX = S^{-1}(A - E_nX)S = S^{-1}M_A(X)S$$

und da $GL_n(K) \subset GL_n(K[X])$ eine Untergruppe ist, gibt das sogar die Ähnlichkeit, nicht nur die Äquivalenz, der charakteristischen Matrizen.

Nehmen wir nun an, dass $M_A(X)$ äquivalent ist zu $M_B(X)$. Es gibt also invertierbare Matrizen $P(X), Q(X) \in GL_n(K[X])$ mit

$$P(X)(B - XE_n) = (A - XE_n)Q(X).$$

Wie sehen $P(X)$ und $Q(X)$ aus? Dies sind Matrizen mit Einträgen in $K[X]$. Wir können die Summen in den Einträgen von $P(X)$ und $Q(X)$ auseinanderziehen und erhalten

$$P(X) = \sum_{i=0}^{m} P_iX^i, \text{ und } Q(X) = \sum_{i=0}^{m} Q_iX^i \text{ mit } P_i, Q_i \in M(n \times n, K).$$

Hier bringen wir die Summen auf die gleiche Länge, indem wir gegebenenfalls Nullen einfügen. Beachten Sie, dass wir X^0 mit E_n identifizieren. Damit ist

$$P(X)(B - XE_n) = \sum_{i=0}^{m} P_iBX^i - \sum_{i=0}^{m} P_iX^{i+1}$$

und

$$(A - XE_n)Q(X) = \sum_{i=0}^{m} AQ_iX^i - \sum_{i=0}^{m} Q_iX^{i+1}$$

Wir vergleichen die Koeffizienten und erhalten die Gleichungen

$$P_iB - P_{i-1} = AQ_i - Q_{i-1}, \quad 1 \leqslant i \leqslant m \tag{VI.1.1}$$

und

$$P_0B = AQ_0 \text{ und } P_m = Q_m. \tag{VI.1.2}$$

Wir multiplizieren die Gleichung (VI.1.1) mit A^i durch und summieren auf. Dies gibt

$$\begin{aligned}\sum_{i=1}^{m} A^iP_iB - \sum_{i=1}^{m} A^iP_{i-1} &= \sum_{i=1}^{m} A^{i+1}Q_i - \sum_{i=1}^{m} A^iQ_{i-1}\\ &= A^{m+1}Q_m - AQ_0, \quad \text{weil sich die restlichen Terme wegheben}\\ &= A^{m+1}P_m - P_0B \quad \text{wegen (VI.1.2)}\end{aligned}$$

Ein Vergleich der Anfangs- und Endterme ergibt

$$\sum_{i=0}^{m} A^iP_iB = \sum_{i=0}^{m} A^{i+1}P_i$$

und wenn wir B beziehungsweise A ausklammern, erhalten wir:

$$\left(\sum_{i=0}^{m} A^iP_i\right)B = A\left(\sum_{i=0}^{m} A^iP_i\right). \tag{VI.1.3}$$

Wir definieren $S = \sum_{i=0}^{m} A^iP_i \in M(n \times n, K)$. Dann besagt (VI.1.3) gerade, dass $SB = AS$. Es bleibt also zu zeigen, dass S invertierbar ist. Nach Voraussetzung war $P(X)$ invertierbar, das heißt, es gibt ein $R(X) \in M(n \times n, K[X])$ mit

$$P(X)R(X) = E_n = E_nX^0. \tag{VI.1.4}$$

Wir schreiben $R(X)$ wieder explizit als $R(X) = \sum_{i=0}^{m} R_iX^i$ mit $R_i \in M(n\times n, K)$. Es sei $T = \sum_{j=0}^{m} B^jR_j \in M(n \times n, K)$. Wir rechnen nach, dass

$$ST = S\sum_{j=0}^{m} B^jR_j = \sum_{j=0}^{m} A^jSR_j,$$

weil mit (VI.1.3) gilt, dass $SB = AS$ ist. Wir setzen die Definition von S ein und bekommen

$$ST = \sum_{j=0}^{m}\sum_{i=0}^{m} A^{i+j}P_iR_j.$$

Da $R(X)$ das Inverse zu $P(X)$ ist, bleibt von der ganzen Doppelsumme nur E_n übrig. $\square$

Wir werden also im Folgenden Matrizen über $K[X]$ betrachten. Vorsicht: $K[X]$ ist natürlich *kein* Körper, aber wir werden die folgende Eigenschaft von $K[X]$ benutzen:

Definition VI.1.4. Ein kommutativer Integritätsbereich R heißt ein *euklidischer Ring*, falls es eine Abbildung

$$\nu\colon R \to \mathbb{N}_0$$

gibt, mit $\nu(0) = 0$, so dass es für alle $a, b \in R$ mit $a \neq 0$ Elemente $q, r \in R$ gibt mit

$$b = qa + r, \text{ so dass } \nu(r) < \nu(a).$$

Die Abbildung ν heißt *euklidische Norm*.

Sie kennen zwei Beispiele euklidischer Ringe:

Beispiele VI.1.5.

(a) Wir hatten in Satz V.2.13 gezeigt, dass $K[X]$ die Polynomdivision mit Rest zuläßt. Hier ist

$$\nu(f) = \begin{cases} \mathrm{Grad}(f) + 1, & f \neq 0 \\ 0, & f = 0 \end{cases}$$

die euklidische Norm.

(b) In den ganzen Zahlen, $\mathbb{Z}$, können Sie ebenfalls mit Rest teilen. Hier ist ν der Absolutbetrag.

Bemerkung VI.1.6. Sie können den Beweis von Satz V.5.4 auf beliebige euklidische Ringe übertragen. Damit ist *jedes* Ideal $I \subset R$ in einem euklidischen Ring ein Hauptideal, also von der Form $I = (a) = aR$ für ein $a \in R$. Solche Ringe nennt man *Hauptidealringe*.

Definition VI.1.7. Es sei R ein Integritätsbereich.

(a) Sind $a, b \in R$ und ist $q \in R$ gegeben mit $b = qa$, so *teilt a b*. Wir sagen auch, dass b ein *Vielfaches von a* ist. Die Notation ist $a|b$.

(b) Für eine Familie von Elementen $(a_i \in R)_{i \in I}$ heißt $d \in R$ ein *gemeinsamer Teiler der a_i*, falls $d|a_i$ gilt für alle $i \in I$. Ein gemeinsamer Teiler heißt *größter gemeinsamer Teiler, (ggT)*, falls er von allen gemeinsamen Teilern geteilt wird.

(c) Für zwei Ideale $I_1, I_2 \subset R$ ist

$$I_1 + I_2 := \{x_1 + x_2, x_1 \in I_1, x_2 \in I_2\}$$

die *Summe von I_1 und I_2*.

(d) Ein $p \in R$ heißt *irreduzibel*, falls gilt: p ist keine Einheit und ist $p = ab$ mit $a, b \in R$, so ist a oder b invertierbar.

Bemerkung VI.1.8.

- Das Ideal $I_1 + I_2$ ist das kleinste Ideal mit $I_1 \subset I_1 + I_2$ und $I_2 \subset I_1 + I_2$.
- Ist R ein Hauptidealring so gilt:

  $$a|b \Leftrightarrow (b) \subset (a) :$$

  Wenn es ein $q \in R$ gibt mit $aq = b$, ist klar, dass $b \in (a)$ gilt. Damit aber auch $(b) \subset (a)$.
  Ist umgekehrt $(b) \subset (a)$, so ist $b \in (a)$ und damit gibt es ein solches q.
- Ist R ein Hauptidealring, so ist für $a_1, \ldots, a_n \in R$ das Ideal $(a_1) + \ldots + (a_n)$ wiederum ein Ideal, also gibt es ein $d \in R$ mit

  $$(a_1) + \ldots + (a_n) = (d).$$

  Für alle i ist damit $(a_i) \subset (d)$, also teilt d alle a_i. Ist e ein weiterer Teiler aller a_i, so ist $(a_i) \subset (e)$ für alle i und somit auch $(d) = (a_1) + \ldots + (a_n) \subset (e)$. Somit gilt $d|e$. Also ist d der ggT von $a_1, \ldots, a_n$.
- Ist $R = \mathbb{Z}$, so sind die irreduziblen Elemente die Zahlen $\pm p$, wobei p eine Primzahl ist.
- Ist $R = \mathbb{R}[X]$, so sind die Polynome $aX + b$ mit $a \neq 0$ und die Polynome $aX^2 + bx + c$ mit $b^2 - 4ac < 0$ irreduzibel. Der Fundamentalsatz der Algebra sagt, dass für $R = \mathbb{C}[X]$ nur die Polynome $aX + b$ mit $a \neq 0$ irreduzibel sind.
- Ist R ein euklidischer Ring, so gilt $\nu(x) = 0$ genau dann, wenn $x = 0$ ist. Wäre es nicht null, so könnten wir die 1 schreiben als $1 = qx + r$ mit $\nu(r) < \nu(x) = 0$. Es gibt aber kein Element mit kleinerer Bewertung als null.

VI.2. Invariantenteiler

Ziel dieses Abschnitts ist eine Normalform für quadratische Matrizen über einem euklidischen Ring R.

Satz VI.2.1. *Es sei R ein euklidischer Ring und $A \in M(n \times n, R)$. Wir erlauben als Zeilen- und Spaltentransformationen das Vertauschen zweier Zeilen oder Spalten und die Addition eines Vielfachen einer Spalte* *oder Zeile zu einer anderen Spalte oder Zeile. Dann läßt sich A immer umformen zu einer Diagonalmatrix*

$$\begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_n \end{pmatrix}$$

so dass $c_1|c_2|\dots|c_n$.

Beachten Sie, dass jedes c_i die 0 teilt.

Beweis. Ohne Einschränkung sei $A \neq 0$. Wir können weiterhin annehmen, dass $a_{11} \neq 0$ ist und dass für alle Einträge a_{ij} gilt, dass $\nu(a_{11}) \leqslant \nu(a_{ij})$ gilt. Ansonsten vertauschen wir die Zeilen und Spalten so lange, bis oben links ein solches Element steht.

Unser Ziel ist es A in eine Matrix der Form $\begin{pmatrix} \tilde{a}_{11} & 0 \\ 0 & \tilde{A} \end{pmatrix}$ zu transformieren, so dass $\tilde{a}_{11}$ alle Einträge von $\tilde{A}$ teilt. Dann können wir iterativ weitermachen.

Nehmen wir an, es gibt in der ersten Spalte ein Element $a_{i1} \neq 0$ und i sei minimal mit dieser Eigenschaft und $i > 1$. Da R euklidisch ist und da $a_{11} \neq 0$ ist, finden wir ein $q \in R$ mit $\nu(a_{i1} - qa_{11}) < \nu(a_{11})$. Wir addieren das $-q$-fache der ersten Zeile zur i-ten Zeile und erhalten eine Matrix in der der neue Eintrag x_{i1} an der Stelle $(i, 1)$ eine echt kleinere euklidische Norm hat als vorher. Ist dieser Eintrag nicht null, so vertauschen wir wieder Zeilen und Spalten und erhalten eine Matrix A', in der gilt, dass

$$a'_{11} \neq 0 \text{ und } \nu(a'_{11}) \leqslant \nu(a'_{ij}) \text{ für alle } i, j.$$

Da x_{i1} unter den Einträgen von A' ist, haben wir in diesem Schritt erreicht, dass $\nu(a'_{11}) < \nu(a_{11})$. Solange also in der ersten Zeile oder Spalte ein Element außerhalb des Eintrags in $(1, 1)$ nicht verschwindet, können wir die euklidische Norm des ersten Diagonalelements echt verkleinern. Dieses Verfahren muss dadurch nach endlich vielen Schritten abbrechen, weil $\nu(a_{11})$ eine feste natürliche Zahl war. Wir finden also nach endlich vielen Schritten eine Matrix der Form

$$B = \begin{pmatrix} b_{11} & 0 \\ 0 & \tilde{B} \end{pmatrix}$$

mit $\nu(b_{11}) \leqslant \nu(b_{ij})$. Da wir $A \neq 0$ angenommen haben, ist $b_{11} \neq 0$. Wir müssen jetzt noch dafür sorgen, dass wir diese Matrix ersetzen können durch eine, in der der Eintrag in $(1, 1)$ alle anderen Einträge teilt.

Nehmen wir an, dass es ein Paar (i, j) gibt, so dass b_{ij} nicht von b_{11} geteilt wird. Wir teilen wieder mit Rest und erhalten ein $w \in R$ mit $\nu(b_{ij} - wd_{11}) < \nu(b_{11})$, aber nach Annahme ist $b_{ij} - wd_{11} \neq 0$.

Wir addieren die erste Zeile von B zur i-ten Zeile.

Wir Subtrahieren das w-fache der ersten Spalte der so entstandenen Matrix von der j-ten Spalte der Matrix. Zusätzlich löschen wir den nicht-trivialen Eintrag in der 1-ten Spalte durch Addition eines Vielfachen der ersten Zeile. Dann haben wir eine neue Matrix B', die von der Form so ausieht wie B, in der $b'_{ij} \neq 0$ ist, aber $\nu(b'_{ij}) < \nu(b_{ij})$. Diesen Prozess können wir wiederum nur endlich oft durchführen wegen der Beschränktheit der Norm von b_{ij}. Somit können wir B in endlich vielen Schritten in eine Matrix $\tilde{A}$ der Form $\begin{pmatrix} \tilde{a}_{11} & 0 \\ 0 & \tilde{A} \end{pmatrix}$ transformieren, so dass $\tilde{a}_{11}|\tilde{a_{ij}}$ für alle i, j gilt. □

Korollar VI.2.2 (Invariantenteilersatz). *Über einem euklidischen Ring ist jede Matrix $A \in M(n \times n, R)$ äquivalent zu einer Diagonalmatrix*

$$\begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_n \end{pmatrix}$$

so dass $c_1|c_2|\dots|c_n$.

Beweis. Die Spalten- und Zeilenumformungen im Beweis entsprechen der Multiplikation mit invertierbaren Matrizen, nämlich Matrizen der Form $\tau(i,j)$ und $\delta(i,j,\lambda)$ wie in Definition III.4.8. $\square$

Bemerkung VI.2.3. Die Elemente $c_1, \dots, c_n$ sind nicht eindeutig, sie sind nur bis auf Multiplikation mit invertierbaren Elementen festgelegt. Für $R = K[X]$ kann man die Eindeutigkeit erzwingen, indem man verlangt, dass die Polynome c_i normiert sind. Das werden wir im Folgenden tun.

Definition VI.2.4. Die $c_i \in R$ wie in Korollar VI.2.2 heißen *Invariantenteiler von A*.

Wir wollen die c_i mit anderen Größen in Beziehung setzen. Minoren hatten wir in Definition IV.5.1.

Definition VI.2.5. Ist $A = (a_{ij}(X)) \in M(n \times n, K[X])$, so sei $d_j(A)$ der größte gemeinsame Teiler aller Minoren j-ter Ordnung von A. Dann heißt $d_j(A)$ der *j-te Determinantenteiler von A*.

Beispiel VI.2.6. Einige Determinantenteiler können Sie sofort ohne Nachdenken hinschreiben: Für $j = n$ erhalten wir $d_n(A) = \det(A)$ und $d_1(A)$ ist der ggT aller Matrixeinträge $a_{ij}(X)$.

Determinantenteiler geben ein notwendiges Kriterium für Äquivalenz:

Lemma VI.2.7. *Sind $A_1, A_2 \in M(n \times n, K[X])$ äquivalent, so haben A_1 und A_2 übereinstimmende Determinantenteiler.*

Beweis. Wir zeigen, dass $d_j(A_1)$ alle Minoren j-ter Ordnung von A_2 und damit $d_j(A_2)$ teilt. Wegen der Symmetrie der Behauptung, folgt dann die Gleichheit.

Ist $S \in M(n \times n, K[X])$ beliebig, so sind die Zeilen von SA_1 Linearkombinationen von Zeilen von A_1: Wir schreiben S als Matrix mit Zeilen s_i, also

$$S = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$$

und wir schreiben A_1 als Matrix mit Spalten a_i, also $A_1 = (a_1, \dots, a_n)$.

Dann ist

$$SA_1 = \begin{pmatrix} \langle s_1, a_1 \rangle & \dots & \langle s_1, a_n \rangle \\ \vdots & & \vdots \\ \langle s_n, a_1 \rangle & \dots & \langle s_n, a_n \rangle \end{pmatrix}.$$

Damit sind Minoren j-ter Ordnung von SA_1 Linearkombinationen der Minoren j-ter Ordnung von A_1. Ebenso überlegt man sich, dass für jedes T die Spalten von SA_1T Linearkombinationen der Spalten von SA_1 sind und damit sind die Minoren j-ter Ordnung von $A_2 = SA_1T$ Linearkombinationen der Minoren j-ter Ordnung von A_1; insbesondere teilt $d_j(A_1)$ die Minoren j-ter Ordnung von A_2. Damit teilt auch $d_j(A_1)$ diese Minoren. $\square$

Definition VI.2.8. Ist $A \in M(n \times n, K)$ und ist $M_A(X)$ die charakteristische Matrix von A, dann heißen die normierten Polynome

$$c_A^{(j)} := c_j(M_A(X)) \text{ und } d_A^{(j)} := d_j(M_A(X))$$

die *Invarianten- und Determinantenteiler von A*.

Bemerkung VI.2.9. Für Diagonalmatrizen in Invariantenteilerform sind die Determinantenteiler natürlich einfach hinzuschreiben: Ist

$$A = \begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_n \end{pmatrix}$$

mit $c_1 | \dots | c_n$, so ist $d_j(A) = c_1 \cdot \ldots \cdot c_j$.

Insbesondere ist also $d_A^{(j)} = c_A^{(1)} \cdot \ldots \cdot c_A^{(j)}$ für alle $1 \leqslant j \leqslant n$.

Satz VI.2.10 (Klassifikation über Invarianten- und Determinantenteiler)**.** *Zwei Matrizen $A, B \in M(n \times n, K)$ über einem Körper K sind genau dann ähnlich, wenn sie übereinstimmende Invarianten- und Determinantenteiler haben.*

Beweis. Wir hatten gesehen, dass A und B genau dann ähnlich sind, wenn ihre charakteristischen Matrizen $M_A(X)$ und $M_B(X)$ äquivalent sind. Diese Matrizen wiederum sind äquivalent zu Diagonalmatrizen

$$\begin{pmatrix} c_A^{(1)} & 0 & \dots & 0 \\ 0 & c_A^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_A^{(n)} \end{pmatrix} \text{ und } \begin{pmatrix} c_B^{(1)} & 0 & \dots & 0 \\ 0 & c_B^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_B^{(n)} \end{pmatrix}$$

Diese sind genau dann äquivalent, wenn $c_A^{(i)} = c_B^{(i)}$ gilt für alle $1 \leqslant i \leqslant n$. Da wir die Determiantenteiler ausdrücken können als

$$d_A^{(1)} = c_A^{(1)} \text{ und } d_A^{(j)} = c_A^{(j)} d_A^{(j-1)} \text{ für } j > 1$$

und ebenso für B, sind die Invariantenteiler von A und B genau dann gleich, wenn die Determinantenteiler übereinstimmen. □

Beispiel VI.2.11. Wir betrachten $A = \begin{pmatrix} 4 & 0 & -3 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \end{pmatrix} \in M(3 \times 3, \mathbb{R})$ mit charakteristischer Matrix

$$M_A(X) = \begin{pmatrix} 4-X & 0 & -3 \\ 1 & 2-X & 1 \\ 1 & 2 & 2-X \end{pmatrix}.$$

Wir vertauschen die erste mit der zweite Zeile und erhalten

$$\begin{pmatrix} 1 & 2-X & 1 \\ 4-X & 0 & -3 \\ 1 & 2 & 2-X \end{pmatrix}.$$

Wir addieren das $-(2-X)$-fache der ersten Spalte zur zweiten Spalte und ziehen die erste Spalte zusätzlich von der dritten ab:

$$\begin{pmatrix} 1 & 0 & 0 \\ 4-X & -(4-X)(2-X) & -3-4+X \\ 1 & 2-(2-X) & 2-X-1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 4-X & -(4-X)(2-X) & X-7 \\ 1 & X & 1-X \end{pmatrix}.$$

Wir ziehen das $(4-X)$-fache der ersten Zeile von der zweiten ab und ziehen die erste Zeile von der dritten ab:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -(4-X)(2-X) & X-7 \\ 0 & X & 1-X \end{pmatrix}.$$

Wir addieren nun die dritte Spalte zur zweiten:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -(4-X)(2-X)+X-7 & X-7 \\ 0 & 1 & 1-X \end{pmatrix}.$$

Vertauschen von Zeile 2 und 3 gibt

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1-X \\ 0 & -(4-X)(2-X)+X-7 & X-7 \end{pmatrix}.$$

Diese Matrix können wir schließlich durch Ausräumen der Einträge in $(2,3)$ und $(3,2)$ durch die benachbarte Eins umformen zu

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -X^3+8X^2-21X+8 \end{pmatrix}.$$

Sie wissen, dass der Eintrag unten rechts das charakteristische Polynom sein muss.

VI.3. Frobenius-, Weierstraß- und Jordannormalform

Das Ziel dieses Abschnitts ist es, handhabbare und übersichtliche Repräsentanten der Ähnlichkeitsklassen quadratischer Matrizen zu erhalten. Im gesamten Abschnitt sei K ein Körper.

Definition VI.3.1. Es sei $g = X^n + \alpha_{n-1}X^{n-1} + \ldots + \alpha_1 X + \alpha_0 \in K[X]$ ein normiertes Polynom. Ist der Grad von g gleich 1, so sei $B_g := (-\alpha_0)$. Für $n > 1$ sei

$$B_g := \begin{pmatrix} 0 & 0 & \ldots & \ldots & \ldots & 0 & -\alpha_0 \\ 1 & 0 & & & & \vdots & \vdots \\ 0 & 1 & 0 & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & 1 & -\alpha_{n-1} \end{pmatrix} \in M(n \times n, K)$$

die *Begleitmatrix von g*.

Bemerkung VI.3.2. Im Beweis des Satzes von Cayley-Hamilton haben wir gesehen, dass $P_{B_g}(X) = (-1)^n g(X)$ gilt. Die charakteristische Matrix von B_g ist

$$M_{B_g}(X) := \begin{pmatrix} -X & 0 & \ldots & \ldots & \ldots & 0 & -\alpha_0 \\ 1 & -X & & & & \vdots & \vdots \\ 0 & 1 & -X & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & -X & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & 1 & -\alpha_{n-1} - X \end{pmatrix} \in M(n \times n, K[X])$$

Streichen wir die erste Zeile und die n-te Spalte, so bleibt

$$\begin{pmatrix} 1 & -X & 0 & \ldots & 0 \\ 0 & 1 & -X & \ddots & \vdots \\ \vdots & 0 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & -X \\ 0 & \ldots & \ldots & 0 & 1 \end{pmatrix} \in M(n \times n, K[X])$$

mit Determinante 1, so dass die Determinantenteiler von B_g gerade $1, \ldots, 1, g$ sind und

$$c_{B_g}^{(i)} = d_{B_g}^{(i)} = \begin{cases} 1, & i < n, \\ g, & i = n. \end{cases}$$

Damit ist $B_g - XE_n = M_{B_g}(X)$ äquivalent zur Matrix

$$\begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & g \end{pmatrix}$$

Satz VI.3.3 (Frobenius-Normalform)**.** *Ist $A \in M(n \times n, K)$, dann ist A zu genau einer Matrix der Form*

$$\begin{pmatrix} B_{g_1} & 0 & \dots & \dots & 0 \\ 0 & B_{g_2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & B_{g_{r-1}} & 0 \\ 0 & \dots & \dots & 0 & B_{g_r} \end{pmatrix} \in M(n \times n, K)$$

ähnlich, wobei die $g_1, \dots, g_r$ normierte Polynome über K vom Grad $\geqslant 1$ sind, für die $g_1 | \dots | g_r$ gilt.

Beweis. Die charakteristische Matrix von

$$B_{g_1,\dots,g_r} := \begin{pmatrix} B_{g_1} & 0 & \dots & \dots & 0 \\ 0 & B_{g_2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & B_{g_{r-1}} & 0 \\ 0 & \dots & \dots & 0 & B_{g_r} \end{pmatrix}$$

ist äquivalent zu einer Diagonalmatrix mit Diagonaleinträgen $(1, \dots, g_1, \dots, 1, \dots, 1, g_r)$ und diese ist wiederum äquivalent zu einer Diagonalmatrix mit Diagonaleinträgen $(1, \dots, 1, g_1, \dots, g_r)$. Sind $1, \dots, 1, g_1, \dots, g_r$ die Invariantenteiler von A, so sind $1, \dots, 1, g_1, \dots, g_r$ ebenfalls die Invariantenteiler von $B_{g_1,\dots,g_r}$ und mit VI.2.10 folgt, dass A ähnlich ist zu $B_{g_1,\dots,g_r}$. Da die Invariantenteiler eindeutig sind, sind auch die Matrizen $B_{g_1}, \dots, B_{g_r}$ eindeutig. □

Die Matrizen B_{g_1} bis B_{g_r} kann man gegebenenfalls weiter aufspalten:

Lemma VI.3.4. *Ist $g = h_1 \cdot \ldots \cdot h_k$, so dass die h_i paarweise teilerfremd sind und normiert vom Grad $\geqslant 1$, so ist B_g ähnlich zu*

$$\begin{pmatrix} B_{h_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_{h_k} \end{pmatrix}$$

Beweis. Die charakteristische Matrix von

$$\begin{pmatrix} B_{h_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_{h_k} \end{pmatrix}$$

ist über $K[X]$ äquivalent zu

$$H(X) := \begin{pmatrix} E_m & 0 & \dots & 0 \\ 0 & h_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & h_k \end{pmatrix}$$

für ein passendes m. Zu zeigen ist, dass $H(X)$ und $G(X) = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & g \end{pmatrix}$ die gleichen Invarianten- und Determinantenteiler haben. Wir berechnen die Determinantenteiler: Für d_n gilt

$$d_n(G(X)) = g = h_1 \cdot \dots \cdot h_k = d_n(H(X))$$

und $d_j(G(X)) = 1$ für alle $1 \leqslant j \leqslant n-1$. Für $H(X)$ ist

$$d_{n-1}(H(X)) = \operatorname{ggT}\left(\prod_{\substack{i=1\\ i\neq 1}}^{k} h_i, \dots, \prod_{\substack{i=1\\ i\neq k}}^{k} h_i\right).$$

Aber da die h_i paarweise teilerfremd sind, ist $d_{n-1}(H(X)) = d_j(H(X)) = 1$ für $1 \leqslant j \leqslant n-1$.

$\square$

Beispiel VI.3.5. Betrachten wir das Polynom $X^2 + 1 = (X - i)(X + i) \in \mathbb{C}[X]$, so ist $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ die Begleitmatrix zu $X^2 + 0 \cdot X + 1$ und diese Matrix ist ähnlich zu $\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$ über $\mathbb{C}$, weil

$$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} B_{h_1} & 0 \\ 0 & B_{h_2} \end{pmatrix}$$

mit $h_1 = X + i$ und $h_2 = X - i$.

Karl Theodor Wilhelm Weierstraß (1815–1897).

Korollar VI.3.6 (Weierstraß-Normalform)**.** *Ist $A \in M(n \times n, K)$ dann gibt es eine bis auf Reihenfolge eindeutig bestimmte Folge $h_1, \dots, h_k$ von Potenzen irreduzibler, normierter Polynome in $K[X]$, so dass A ähnlich ist zu* $\begin{pmatrix} B_{h_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_{h_k} \end{pmatrix}$.

Die h_i heißen *Weierstraßsche Elementarteiler*. Der folgende Beweis bleibt unvollständig, weil wir nicht gezeigt haben, dass sich jedes Polynom in $K[X]$ vom Grad $\geqslant 1$ in irreduzible Faktoren zerlegen läßt.

Beweis. Wir wissen, dass A ähnlich ist zu

$$\begin{pmatrix} B_{g_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_{g_r} \end{pmatrix},$$

wenn $1, \dots, 1, g_1, \dots, g_r$ die Invariantenteiler von A sind. Wir benutzen die hier unbewiesene Zerlegung der g_i als Produkt von Potenzen irreduzibler Polynome in $K[X]$. $\square$

Wir untersuchen den Spezialfall, bei dem alle h_i von der Form $(X-\lambda)^m$ sind:

Lemma VI.3.7. *Ist $\lambda \in K$ und ist $m \in \mathbb{N}$, so ist die Begleitmatrix von $(X-\lambda)^m$ ähnlich zur Matrix*

$$J(\lambda; m) := \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & \lambda \end{pmatrix}.$$

Beweis. Wir betrachten die charakteristische Matrix von $J(\lambda; m)$:

$$M_{J(\lambda;m)}(X) = \begin{pmatrix} \lambda - X & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & \lambda - X \end{pmatrix}.$$

und erhalten

$$\det M_{J(\lambda;m)}(X) = d^{(m)}_{J(\lambda;m)} = (\lambda - X)^m = (-1)^m (X-\lambda)^m.$$

Das Streichen der ersten Zeile und der m-ten Spalte gibt eine Matrix mit Determinante 1 und somit stimmen die Determinantenteiler von $J(\lambda; m)$ mit denen der Begleitmatrix von $(X-\lambda)^m$ überein:

$$d^{(i)}_{J(\lambda;m)} = d^{(i)}_{B_{(X-\lambda)^m}} \text{ für } 1 \leqslant i \leqslant m.$$

□

Definition VI.3.8. Die Matrix $J(\lambda; m)$ heißt die *$m \times m$-Jordan Matrix zu λ*.

Marie Ennemond Camille Jordan, genannt Camille Jordan (1838–1922).

Satz VI.3.9 (Jordansche Normalform)**.** *Es sei $A \in M(n \times n, K)$ und das charakteristische Polynome von A, $P_A(X)$, zerfalle über K in Linearfaktoren. Dann gibt es eine bis auf Reihenfolge eindeutige Folge von Jordan-Matrizen $J_1, \dots, J_r$ über K, so dass A ähnlich ist zu*

$$\begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_r \end{pmatrix}.$$

Beweis. Da $c^{(j)}_A$ immer $d^{(j)}_A$ teilt und da diese wiederum $P_A(X) = d^{(n)}_A$ teilen, zerfallen alle Invarianten- und Determinantenteiler von A ebenfalls in Linearfaktoren. Damit ist A ähnlich zu

$$\begin{pmatrix} B_{(X-\lambda_1)^{m_1}} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_{(X-\lambda_r)^{m_r}} \end{pmatrix}$$

und damit auch ähnlich zu

$$\begin{pmatrix} J(\lambda_1; m_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J(\lambda_r; m_r) \end{pmatrix}.$$

□

Beispiel VI.3.10. Es sei $A = \begin{pmatrix} 9 & -7 & 0 & 2 \\ 7 & -5 & 0 & 2 \\ 4 & -4 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in M(4 \times 4; \mathbb{Q})$. Wir berechnen $P_A(X)$ durch Entwicklung nach der letzten Zeile und erhalten

$$\begin{aligned} P_A(X) &= \det \begin{pmatrix} 9-X & -7 & 0 & 2 \\ 7 & -5-X & 0 & 2 \\ 4 & -4 & 2-X & 1 \\ 0 & 0 & 0 & 2-X \end{pmatrix} \\ &= (2-X)^2((9-X)(-5-X)+49) \\ &= (2-X)^2(X^2-4X+4) \\ &= (2-X)^4 \end{aligned}$$

und somit ist $\lambda = 2$ eine vierfache Nullstelle. Wir müssen die Eigenvektoren von A bestimmen:

$$(A - 2E_4) \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} = 0 \Leftrightarrow \begin{pmatrix} 7 & -7 & 0 & 2 \\ 7 & -7 & 0 & 2 \\ 4 & -4 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} = 0.$$

Die Koeffizientenmatrix hat Rang 2: Sie kann nicht mehr als Rang 2 haben wegen der 0-Spalte und der beiden sichtbar linear abhängigen ersten und zweiten Spalte. Da $\det \begin{pmatrix} -7 & 2 \\ -4 & 1 \end{pmatrix} = -7 + 8 = 1 \neq 0$ hat sie genau Rang 2.

Damit ist die Dimension des Eigenraums $\dim_{\mathbb{Q}} \operatorname{Eig}(A; 2) = 2$ und als Jordansche Normalformen kommen a priori

$$A_1 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} J(2;1) & 0 \\ 0 & J(2;3) \end{pmatrix}$$

und

$$A_2 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} J(2;2) & 0 \\ 0 & J(2;2) \end{pmatrix}$$

in Betracht.

Wir bestimmen das Minimalpolynom von A. Es ist $A - 2E_4 \neq 0$. Das hatten wir oben schon gesehen. Aber $(A - 2E_4)^2$ ist

$$\begin{pmatrix} 7 & -7 & 0 & 2 \\ 7 & -7 & 0 & 2 \\ 4 & -4 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 7 & -7 & 0 & 2 \\ 7 & -7 & 0 & 2 \\ 4 & -4 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 0.$$

Also ist das Minimalpolynom $m_A(X) = (X-2)^2$. Ähnliche Matrizen haben übereinstimmende Minimalpolynome. Wir bestimmen also die Minimalpolynome von A_1 und A_2:

Die Matrix $(A_1 - 2)$ ist $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ und damit ist $(A_1 - 2)^2 \neq 0$. Daher muss die Jordansche Normalform von A gleich A_2 sein. Wir machen den Test:

$$(A_2 - 2) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = 0.$$

Bemerkung VI.3.11.

(a) Oft finden Sie Jordan-Matrizen in der Form

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

Dies liefert eine äquivalente Theorie, weil

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} = J(\lambda; m)^t$$

und für $A \in M(n \times n; K)$ gilt für alle i: $c_A^{(i)} = c_{A^t}^{(i)}$ und $d_A^{(i)} = d_{A^t}^{(i)}$.

(b) Ist A ähnlich zu $\begin{pmatrix} J(\lambda_1; m_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J(\lambda_r; m_r) \end{pmatrix}$ mit paarweise verschiedenen λ_i, so ist

$$P_A(X) = (\lambda_1 - X)^{m_1} \cdot \ldots \cdot (\lambda_r - X)^{m_r}$$

und damit gilt $\mu_{\text{alg}}(A; \lambda_i) = m_i$.

(c) Für die geometrische Vielfachheit gilt dagegen:

$$\mu_{\text{geo}} \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & \lambda \end{pmatrix} = 1$$

Allgemeiner gilt: Die geometrische Vielfachheit eines Eigenwerts ist die Anzahl der Jordan-Blöcke zu diesem Eigenwert.

(d) Das Minimalpolynom können Sie ebenfalls direkt aus der Jordanschen Normalform ablesen: Sind $\lambda_1, \dots, \lambda_m$ die verschiedenen Eigenwerte von A und zerfällt $P_A(X)$ in Linearfaktoren, so sei ℓ_i jeweils die Größe eines maximalen Jordanblocks zum Eigenwert λ_i, dann gilt:

$$((J(\lambda_i, \ell_i) - \lambda_i E_{\ell_i})^k = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix}^k = \begin{cases} 0, & k \geqslant \ell_i, \\ \neq 0, & k < \ell_i. \end{cases}$$

Damit gilt für das Minimalpolynom von A:

$$m_A(X) = (X - \lambda_1)^{\ell_1} \cdot \ldots \cdot (X - \lambda_m)^{\ell_m}.$$

Wir erhalten als Spezialfall: Zerfällt das Minimalpolynom von A in paarweise verschiedene Linearfaktoren, so ist $\ell_i = 1$ für alle $1 \leqslant i \leqslant m$ und damit gibt es nur Jordanblöcke der Größe 1. In diesem Fall ist A diagonalisierbar. Das ist die fehlende Rückrichtung von Satz V.5.8:

Satz VI.3.12. *Es sei $A \in M(n\times n, K)$. Dann ist A genau dann diagonalisierbar, wenn das Minimalpolynom $m_A(X)$ in paarweise verschiedene Linearfaktoren zerfällt.*

Bemerkung VI.3.13. Ist V ein K-Vektorraum endlicher Dimension und es sei A die darstellende Matrix von $f \in \mathsf{End}_K(V)$, so ist $P_f(X) = P_A(X)$. Zerfällt $P_A(X)$ in Linearfaktoren, so ist A ähnlich zu $J := \begin{pmatrix} J(\lambda_1; m_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J(\lambda_r; m_r) \end{pmatrix}$. Wir schreiben diese Matrix als Summe $J = D + N$, wobei D eine Diagonalmatrix ist der Form

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \lambda_1 & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \lambda_r & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & \lambda_r \end{pmatrix}$$

und N hat nur Einsen unterhalb der Diagonalen:

$$N = \begin{pmatrix} N_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & N_m \end{pmatrix}$$

mit $N_i = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$, wobei wir wieder die Konvention benutzen, dass dort wo nichts steht, Nullen sind.

Man kann also jedes solche f zerlegen als $f = f_D + f_N$, wobei f_D diagonalisierbar und f_N nilpotent ist. Es gilt

$$f_N \circ f_D = f_D \circ f_N.$$

Dies liefert eine *multiplikative Version der Jordanschen Normalform*: Ist $f \in \mathrm{Aut}_K(V)$, so setzen wir $f_U = \mathrm{id}_V + f_D^{-1} \circ f_N$. Damit erhalten wir

$$f_D \circ f_U = f_D \circ (\mathrm{id}_V + f_D^{-1} \circ f_N) = f_D + f_N = f.$$

Die Abbildung f_U heißt *unipotent*. Wir können also jeden solchen Automorphismus f multiplikativ zerlegen in einen Diagonalanteil f_D und einen unipotenten Anteil f_U.

VI.4. Exkurs: Anwendung der Jordanschen Normalform auf Differentialgleichungen

In diesem Abschnitt sei $K = \mathbb{R}$ oder $\mathbb{C}$. Wir suchen ein n-mal stetig differenzierbares $f\colon \mathbb{R} \to K$, welches die Gleichung

$$f^{(n)} + a_{n-1} f^{(n-1)} + \ldots + a_1 f' + a_0 f = 0 \tag{VI.4.1}$$

erfüllt. Hierbei seien die a_i aus K und $f^{(i)}$ bezeichne die i-te Ableitung von f. Die obige Gleichung nennt man eine *homogene Differenzialgleichung mit konstanten Koeffizienten.*

Wir setzen $x_1 = f, \ldots x_n = f^{(n-1)}$. Dann ist (VI.4.1) äquivalent zu

$$x_1' = x_2, x_2' = x_3, \ldots, x_{n-1}' = x_n \text{ und } x_n' = f^{(n)},$$

also $x_n' = -\sum_{i=1}^n a_{i-1}x_i$, so dass die x_i einmal stetig differenzierbar sind. Wir benutzen die Abkürzung $C^i(\mathbb{R}, K)$ für den K-Vektorraum der i-mal stetig differenzierbaren Funktionen auf $\mathbb{R}$. Die x_i sind also aus $C^1(\mathbb{R}, K)$ und

$$X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in C^1(\mathbb{R}, K^n).$$

Insgesamt ist (VI.4.1) äquivalent zu

$$X' = AX \text{ mit } A = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ \vdots & 0 & 1 & 0 & \ldots & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \\ 0 & \ldots & \ldots & \ldots & 0 & 1 \\ -a_0 & -a_1 & \ldots & \ldots & \ldots & -a_{n-1} \end{pmatrix}.$$

Wir können also eine homogene Differenzialgleichung n-ter Ordnung mit konstanten Koeffizienten auf eine Matrixgleichung der Form $X' = AX$ bringen mit $X \in C^1(\mathbb{R}, K^n)$.

Wir betrachten leicht allgemeiner $X' = AX$ mit $X \in C^1(\mathbb{R}, K^n)$ und beliebigem $A \in M(n \times n, K)$. Schreiben wir aus, was das heißt, so erhalten wir

$$\begin{aligned} x_1' &= a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \quad & \quad \vdots \\ x_n' &= a_{n1}x_1 + \ldots + a_{nn}x_n \end{aligned}$$

Wir betrachten zunächst einfache Spezialfälle: Ist A zum Beispiel eine Diagonalmatrix, so reduzieren sich die obigen Gleichungen zu

$$\begin{aligned} x_1' &= a_{11}x_1 \\ \vdots \quad & \quad \vdots \\ x_n' &= a_{nn}x_n \end{aligned}$$

und in der Analysisvorlesung lernen Sie, dass Sie in diesem Fall

(VI.4.2) $$x_i(t) = c_i e^{a_{ii}t}$$

als Lösung erhalten. Da $e^0 = 1$ gilt, ist $c_i = x_i(0)$.

Ein ähnlich einfacher Fall ist der, bei dem A diagonalisierbar ist. Es gibt also ein $S \in GL_n(K)$, so dass $SAS^{-1} = D$ ist, wobei $D = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & d_n \end{pmatrix}$ eine Diagonalmatrix ist. Damit ist $SA = DS$ und wir erhalten:

$$(Sx)' = Sx' = SAx = D(Sx).$$

Aber mit $y := Sx$ wissen wir nach dem ersten Fall, wie eine Lösung aussieht für $y' = Dy$. Wir erhalten also y und damit auch $x = S^{-1}y$.

Definition VI.4.1. Es sei $K = \mathbb{R}$ oder $\mathbb{C}$. Für $A \in M(n \times n, K)$ ist

$$e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!}. \tag{VI.4.3}$$

Bemerkung VI.4.2. Wir fassen einige Eigenschaften der Exponentialabbildung für Matrizen zusammen:

- Die Reihe in (VI.4.3) konvergiert für alle $A \in M(n \times n, K)$.
- Für $0 \in M(n \times n, K)$ erhalten wir $e^0 = E_n$.
- Sind $A, B \in M(n \times n, K)$, so ist
$$e^{A+B} = e^A e^B.$$
- Damit ist klar, dass e^A invertierbar ist, mit $(e^A)^{-1} = e^{-A}$, weil
$$e^A e^{-A} = e^{A-A} = e^0 = E_n = e^{-A+A} = e^{-A} e^A.$$
- Für Diagonalmatrizen $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ ist die Exponentialfunktion beschreibbar als

$$\begin{aligned} e^{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}} &= \sum_{i=0}^{\infty} \frac{1}{i!} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}^i \\ &= \sum_{i=0}^{\infty} \begin{pmatrix} \frac{1}{i!}\lambda_1^i & & \\ & \ddots & \\ & & \frac{1}{i!}\lambda_n^i \end{pmatrix} \\ &= \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} \end{aligned}$$

Damit wird das System (VI.4.2) mit $\lambda_i := a_{ii}$ zu

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} = e^{\begin{pmatrix} \lambda_1 t & & \\ & \ddots & \\ & & \lambda_n t \end{pmatrix}} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Da $x_i(0) = c_i$, können wir das weiter umformen zu

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = e^{\begin{pmatrix} \lambda_1 t & & \\ & \ddots & \\ & & \lambda_n t \end{pmatrix}} \cdot \begin{pmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{pmatrix}.$$

Dies ist die eindeutige Lösung des Anfangswertproblems

$$X' = AX, \quad X(0) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Ist $K = \mathbb{R}$, so wird nicht jedes charakteristische Polynom in Linearfaktoren zerfallen, aber für $K = \mathbb{C}$ gilt dies nach dem Fundamentalsatz der Algebra. Wir beschränken uns jetzt auf den Fall, wo A nur aus einem Jordanblock besteht.

Es sei also $X' = J(\lambda; n)X$:

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & \lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ x_1 + \lambda x_2 \\ \vdots \\ x_{n-1} + \lambda x_n \end{pmatrix}.$$

Auch in diesem Fall gibt es eine explizite Lösung. Wir leiten die ersten Koordinaten her:
Es ist klar, dass $x_1 = c_1 e^{\lambda t}$ gilt mit Anfangswert $c_1 = x_1(0)$. Für x_2 erhalten wir

$$x_2' = x_1 + \lambda x_2 = c_1 e^{\lambda t} + \lambda x_2.$$

Für $x_2 := c_2 e^{\lambda t} + c_1 t e^{\lambda t}$ ist

$$x_2' = \lambda c_2 e^{\lambda t} + c_1 \lambda t e^{\lambda t} + c_1 e^{\lambda t} = \lambda x_2 + c_1 e^{\lambda t} = \lambda x_2 + x_1.$$

Setzen wir für $x_3' = x_2 + \lambda x_2$ ein, was wir über x_2 wissen, so ist

$$x_3' = c_1 t e^{\lambda t} + c_2 e^{\lambda t} + \lambda x_3,$$

so dass man hier $x_3 = c_3 e^{\lambda t} + c_2 t e^{\lambda t} + \frac{c_1}{2} t^2 e^{\lambda t}$ raten kann.
Für allgemeines n erhält man:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = e^{J(\lambda;n)t} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Für $n = 3$ ist

$$\begin{aligned} e^{J(\lambda;3)t} &= e^{\begin{pmatrix} \lambda t & 0 & 0 \\ t & \lambda t & 0 \\ 0 & t & \lambda t \end{pmatrix}} \\ &= e^{\begin{pmatrix} \lambda t & 0 & 0 \\ 0 & \lambda t & 0 \\ 0 & 0 & \lambda t \end{pmatrix}} \cdot e^{\begin{pmatrix} 0 & 0 & 0 \\ t & 0 & 0 \\ 0 & t & 0 \end{pmatrix}} \\ &= \begin{pmatrix} e^{\lambda t} & 0 & 0 \\ 0 & e^{\lambda t} & 0 \\ 0 & 0 & e^{\lambda t} \end{pmatrix} \left(E_3 + \begin{pmatrix} 0 & 0 & 0 \\ t & 0 & 0 \\ 0 & t & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{t^2}{2} & 0 & 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} e^{\lambda t} & 0 & 0 \\ 0 & e^{\lambda t} & 0 \\ 0 & 0 & e^{\lambda t} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ \frac{t^2}{2} & t & 1 \end{pmatrix} \end{aligned}$$

und

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = e^{J(\lambda;3)t} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & 0 & 0 \\ 0 & e^{\lambda t} & 0 \\ 0 & 0 & e^{\lambda t} \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_1 t + c_2 \\ c_1 \frac{t^2}{2} + c_2 t + c_3 \end{pmatrix}.$$

VI.5. Verallgemeinerte Eigenräume

Es sei K wieder ein Körper, V sei ein K-Vektorraum mit $1 \leqslant \dim_K V = n < \infty$ und $f \in \mathsf{End}_K(V)$. Wir nehmen an, dass das charakteristische Polynom von f über K in Linearfaktoren zerfällt, also

$$P_f(X) = (-1)^n (X - \lambda_1)^{m_1} \cdot \ldots \cdot (X - \lambda_r)^{m_r}$$

mit paarweise verschiedenen $\lambda_i \in K$ und $m_i \geqslant 1$.

Definition VI.5.1. Der Vektorraum $\ker((f - \lambda_i \mathrm{id}_V)^{m_i}) =: v\mathrm{Eig}(f; \lambda_i)$ heißt der *verallgemeinerte Eigenraum von f zum Eigenwert λ_i.*

Es ist auch die Bezeichnung *Hauptraum von f zu λ_i* üblich.

Bemerkung VI.5.2. Der Eigenraum $\mathrm{Eig}(f;\lambda_i)$ ist immer im verallgemeinerten Eigenraum $v\mathrm{Eig}(f;\lambda_i)$ enthalten. Allgemeiner gilt: Ist $g \in \mathsf{End}_K(V)$, so ist

$$\ker(g) \subset \ker(g \circ g) \subset \ldots \subset \ker(\underbrace{g \circ \ldots \circ g}_{m}) = \ker(g^m) \subset \ldots$$

für alle $m \geqslant 1$.

Verallgemeinerte Eigenräume kann man zur Bestimmung der Basiswechselmatrizen für Jordannormalformen benutzen. Ist $A \in M(n \times n, K)$, $P_A(X) = (-1)^n(X-\lambda_1)^{\ell_1} \cdot \ldots \cdot (X-\lambda_r)^{\ell_r}$ und $m_A(X) = (X-\lambda_1)^{m_1} \cdot \ldots \cdot (X-\lambda_r)^{m_r}$, so haben wir

$$\mathrm{Eig}(A;\lambda_i) = \ker(A-\lambda_i E_n) \subset \ker((A-\lambda_i E_n)^2) \subset \ldots \subset \ker((A-\lambda_i E_n)^{m_i}) = v\mathrm{Eig}(A;\lambda_i).$$

Wir starten mit einem Vektor $0 \neq v_0^1 \in \ker((A-\lambda_1 E_n)^{m_1})$ mit $(A-\lambda_1 E_n)^i(v_0^1) \neq 0$ für alle $i < m_1$. Wir setzen

$$\begin{aligned} v_1^1 &:= (A-\lambda_1 E_n)(v_0^1), \\ \vdots \quad & \qquad \vdots \\ v_{m_1-1}^1 &:= (A-\lambda_1 E_n)(v_{m_1-2}^1). \end{aligned}$$

Aus $(X-\lambda_1)^{j+1} = X(X-\lambda_1)^j - \lambda_1(X-\lambda_1)^j$ folgt

$$X(X-\lambda_1)^j = (X-\lambda_1)^{j+1} + \lambda_1(X-\lambda_1)^j. \tag{VI.5.1}$$

Lemma VI.5.3. *Es gilt:*

$$Av_i^1 = \begin{cases} v_{i+1}^1 + \lambda_1 v_i^1, & 0 \leqslant i < m_1 - 1 \\ \lambda_1 v_{m_1-1}^1, & i = m_1 - 1. \end{cases}$$

Beweis. Es gilt $(A-\lambda_1 E_n)^{m_1}(v_0^1) = 0$ nach Voraussetzung und mit (VI.5.1) erhalten wir

$$Av_i^1 = A \cdot (A-\lambda_1 E_n)^i(v_0^1) = (A-\lambda_1 E_n)^{i+1}(v_0^1) + \lambda_1(A-\lambda_1 E_n)^i(v_0^1).$$

Für $i < m_1 - 1$ ergibt sich damit

$$A \cdot v_i^1 = v_{i+1}^1 + \lambda_1 v_i^1,$$

während wir für $i = m_1 - 1$

$$(A-\lambda_1 E_n)^{m_1}(v_0^1) = 0$$

erhalten und damit

$$A \cdot v_{m_1-1}^1 = \lambda_1 v_{m_1-1}^1.$$

□

Wir können damit die Basiswechselmatrix für die Jordannormalform bestimmen. Es sei λ_j ein Eigenwert und λ_j habe Multiplizität m_j in $m_A(X)$.

(a) Wir suchen zu λ_j ein v_0^j, also ein $v_0^j \in K^n$ mit

$$(A-\lambda_j E_n)^{m_j}(v_0^j) = 0 \text{ aber } (A-\lambda_j E_n)^i(v_0^j) \neq 0 \text{ für } 0 < i < m_j.$$

(b) Wir berechnen die Vektoren

$$v_0^j, v_1^j = (A-\lambda_j E_n)v_0^j, \ldots, v_{m_j-1}^j = (A-\lambda_j E_n)^{m_j-1}v_0^j.$$

(c) Wir wiederholen die Schritte (a) und (b) gegebenenfalls, falls es noch weitere Jordanblöcke zu λ_j gibt. Diese treten mit Multiplizität $\leqslant m_j$ in $m_A(X)$ auf. Wir achten darauf, dass die neuen Vektoren linear unabhängig von den bisher gefundenen sind.

(d) Haben wir λ_j abgearbeitet, so machen wir mit den anderen Eigenwerten von A weiter.

Beispiel VI.5.4. Wir betrachten wieder das Beispiel

$$A = \begin{pmatrix} 9 & -7 & 0 & 2 \\ 7 & -5 & 0 & 2 \\ 4 & -4 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Sie wissen, dass 2 ein Eigenwert ist und dass

$$P_A(X) = (X-2)^4 \text{ und } m_A(X) = (X-2)^2.$$

Wir suchen also ein $v_0^1 \in \mathbb{Q}^4$ mit $(A - 2E_4)v_0^1 \neq 0$. Dazu schreiben wir hin

$$A - 2E_4 = \begin{pmatrix} 7 & -7 & 0 & 2 \\ 7 & -7 & 0 & 2 \\ 4 & -4 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Da die erste Spalte von $A - 2E_4$ nicht verschwindet, eignet sich zum Beispiel e_1 als v_0^1 und wir erhalten als v_1^1

$$v_1^1 = (A - 2E_4)e_1 = \begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix}.$$

Damit ist $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix} \right)$ der Beginn einer passenden Basis.

Wir brauchen nun einen von e_1 und $\begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix}$ linear unabhängigen Vektor w_0^1 wiederum mit $(A-2E_4)w_0^1 \neq 0$.

Hierzu können wir zum Beispiel $w_0^1 = e_4$ nehmen: e_4 liegt nicht im Erzeugnis von e_1 und $\begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix}$, weil es eine nicht-triviale 4. Koordinate hat.

Damit erhalten wir als Basis

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 7 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix} \right).$$

Wir nehmen diese Vektoren als Spaltenvektoren der Matrix S:

$$S = \begin{pmatrix} 1 & 7 & 0 & 2 \\ 0 & 7 & 0 & 2 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Dann gilt

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

VI.6. Zyklische Unterräume

Es sei K wieder ein Körper, V sei ein K-Vektorraum und $f \in \mathsf{End}_K(V)$.

Definition VI.6.1. Der Vektorraum V heißt *zyklisch bezüglich* f oder f*-zyklisch*, falls es ein $v \in V$ gibt mit

$$V = \mathsf{Span}_K(v, f(v), f^2(v), \ldots).$$

Jeder Vektor $v \in V$ mit dieser Eigenschaft heißt ein f*-zyklischer Vektor von* V.

Die folgenden Matrizen sollten Ihnen bekannt vorkommen:

Satz VI.6.2. *Es sei* $1 < \dim_K V = n < \infty$ *und* $f \in \mathsf{End}_K(V)$. *Dann gilt, dass* V *genau dann zyklisch ist bezüglich* f*, wenn es eine geordnete Basis* $\mathcal{B}$ *von* V *gibt, so dass*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & \dots & \dots & \dots & 0 & -a_0 \\ 1 & 0 & & & & \vdots & \vdots \\ 0 & 1 & 0 & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Beweis. Ist V f-zyklisch und ist v ein f-zyklischer Vektor von V, so setzen wir

$$\mathcal{B} = (v, f(v), \ldots, f^{n-1}(v)).$$

Wir haben früher schon gesehen, dass $\mathcal{B}$ linear unabhängig ist, aber dass $(v, f(v), \ldots, f^{n-1}(v), f^n(v))$ natürlich linear abhängig sein muss, weil wir hier mehr Vektoren haben als die Dimension von V. Also gibt es $a_0, \ldots, a_{n-1} \in K$ mit

$$f^n(v) + a_{n-1}f^{n-1}(v) + \ldots + a_1 f(v) + a_0 = 0.$$

Damit ist

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & \dots & \dots & \dots & 0 & -a_0 \\ 1 & 0 & & & & \vdots & \vdots \\ 0 & 1 & 0 & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Hat f umgekehrt eine Matrixdarstellung $M_{\mathcal{B}}(f)$ wie oben mit $\mathcal{B} = (v_1, \ldots, v_n)$, so ist v_1 f-zyklisch. $\square$

Bemerkung VI.6.3. Wir können aus einer Matrixdarstellung wie oben wieder direkt das charakterische Polynom ablesen: Ist

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & \dots & \dots & \dots & 0 & -a_0 \\ 1 & 0 & & & & \vdots & \vdots \\ 0 & 1 & 0 & & & \vdots & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & 0 & \dots & \dots & 0 & 1 & -a_{n-1} \end{pmatrix},$$

so ist $P_f(X) = (-1)^n(X^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0)$.

Korollar VI.6.4. *Ist $\dim_K V = n < \infty$ und ist V f-zyklisch für ein $f \in \mathsf{End}_K(V)$, so ist*

$$(-1)^n m_f(X) = P_f(X).$$

Beweis. Da $(v, f(v), \ldots, f^{n-1}(v))$ linear unabhängig ist, kann kein Polynom $p \in K[X]$, welches einen Grad strikt kleiner als n hat, f annulieren. □

Korollar VI.6.5. *Ist $\dim_K V = n < \infty$ und $f \in \mathsf{End}_K(V)$, so stimmt das Minimalpolynom von f mit dem n-ten Invariantenteiler überein:*

$$m_f(X) = c^{(n)}(f).$$

Beweis. Es seien $1, \ldots, 1, g_1, \ldots, g_r$ die Invariantenteiler von f mit Begleitmatrizen $B_{g_1}, \ldots, B_{g_r}$ und $g_1 \mid \ldots \mid g_r$. Wir wissen, dass jede darstellende Matrix $M_{\mathcal{B}}(f)$ von f ähnlich ist zu

$$B = \begin{pmatrix} B_{g_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & B_{g_r} \end{pmatrix}.$$

Wir müssen also zeigen, dass $g_r = m_f(X)$ gilt.

Ist $g \in K[X]$ beliebig, so ist

$$g(B) = \begin{pmatrix} g(B_{g_1}) & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & g(B_{g_r}) \end{pmatrix}.$$

Wir wissen, dass g_i das Minimalpolynom von B_{g_i} ist für alle $1 \leqslant i \leqslant r$. Damit gilt

$$g(B) = 0 \Leftrightarrow g_i \mid g \text{ für alle } 1 \leqslant i \leqslant r.$$

Da aber $g_1 \mid \ldots \mid g_r$ gilt, erhalten wir

$$g(B) = 0 \Leftrightarrow g_r \mid g.$$

Aber dies ist genau die Charakterisierung des Minimalpolynoms als normiertes Polynom, welches den Annihilator von B, und damit von f, erzeugt. Also ist $g_r = m_f(X)$. □

Bemerkung VI.6.6. Es sei wieder $f \in \mathsf{End}_K(V)$ für ein endlich-dimensionales V.

(a) Da gilt

$$(-1)^n P_f(X) = d_f^{(n)} = c_f^{(1)} \cdot \ldots \cdot c_f^{(n)}$$

erhalten wir wiederum mit $m_f(X) = c_f^{(n)}$, dass $m_f(X) \mid P_f(X)$.

(b) Gilt $m_f(X) = (-1)^n P_f(X)$, so sind die Invariantenteiler von f genau $1, \ldots, 1, m_f(X)$ und dies sind die Invariantenteiler von $B_{m_f(X)}$. Begleitmatrizen sind gerade die darstellenden Matrizen von Endomorphismen zyklischer Vektorräume. Damit erhalten wir die folgende Verschärfung von Korollar VI.6.4:

Satz VI.6.7. *Ist $f \in \mathsf{End}_K(V)$ und $\dim_K V = n < \infty$, dann ist V genau dann f-zyklisch, wenn $m_f(X) = (-1)^n P_f(X)$ gilt.*

Definition VI.6.8. Ist $U \subset V$ ein Untervektorraum eines K-Vektorraums V und ist U f-invariant für ein $f \in \mathsf{End}_K(V)$, so heißt U *f-zyklisch*, falls U $f|_U$-zyklisch ist.

$$\begin{array}{ccc} U & \xrightarrow{f|_U} & U \\ \cap & & \cap \\ V & \xrightarrow[f]{} & V \end{array}$$

Satz VI.6.9. *Ist $f \in \mathsf{End}_K(V)$ mit $\dim_K V < \infty$, so kann man V zerlegen als*

$$V = U_1 \oplus \ldots \oplus U_r,$$

so dass $U_i \neq 0$ jeweils ein f-zyklischer Unterraum von V ist. Zusätzlich erfüllen die Minimalpolynome g_i von $f|_{U_i}$, dass $g_i \mid g_{i+1}$ für alle $1 \leqslant i \leqslant r-1$. Die g_i sind dann die Invariantenteiler von f und $g_r = m_f(X)$.

Beweis. Es gibt eine geordnete Basis $\mathcal{B}$ von V, so dass

$$M_{\mathcal{B}}(f) = \begin{pmatrix} B_{g_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & B_{g_r} \end{pmatrix}.$$

Jedem B_i entspricht ein f-invarianter Untervektorraum U_i, so dass $f|_{U_i}$ durch B_{g_i} dargestellt wird. Damit ist U_i automatisch f-zyklisch und $V = U_1 \oplus \ldots \oplus U_r$. Aus der Eindeutigkeit der Invariantenteiler folgt, dass die g_i genau die nicht-konstanten Invariantenteiler von f sind. $\square$

KAPITEL VII

Dualräume und Bilinearformen

VII.1. Der Dualraum eines Vektorraums

Definition VII.1.1. Es sei K ein Körper und V sei ein K-Vektorraum. Dann heißt

$$V^* := \mathsf{Hom}_K(V, K) = \{\varphi\colon V \to K,\ \varphi \text{ ist } K\text{-linear}\}$$

der *Dualraum von* V. Seine Elemente heißen *Linearformen* oder *(lineare) Funktionale*.

Beispiele VII.1.2.

(a) Ist $V = K^n$ und ist $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n$ fest gewählt, so ist

$$\varphi^a\colon K^n \to K, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^n a_i x_i$$

eine Linearform. Für $n = 2$ sind $\varphi^{e_1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1$ und $\varphi^{e_2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2$ konkrete Beispiele.

(b) Ist $V = C^0([0,1], \mathbb{R})$ der Vektorraum der stetigen reellwertigen Funktionen auf dem Einheitsintervall, so ist $\varepsilon\colon V \to \mathbb{R}$, $\varepsilon(f) = f(0)$ eine Linearform.

(c) Ebenfalls auf $V = C^0([0,1], \mathbb{R})$ definiert das Integral eine Linearform:

$$\psi\colon C^0([0,1], \mathbb{R}) \to \mathbb{R}, \quad \psi(f) = \int_0^1 f(x)dx.$$

Bemerkung VII.1.3. Wir hatten allgemein in Korollar III.2.23 die Dimensionsformel $\dim_K \mathsf{Hom}_K(V, W) = \dim_K V \cdot \dim_K W$, falls $\dim_K V, \dim_K W < \infty$. Sie wissen also, dass $\dim_K V^* = \dim_K V$ gilt für endlich-dimensionales V.

Definition VII.1.4. Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine geordnete Basis von V, so sei die *zu $\mathcal{B}$ duale Basis $\mathcal{B}^* = (v_1^*, \dots, v_n^*)$* definiert durch

$$v_j^*(v_i) = \delta_{ij}.$$

Bemerkung VII.1.5. Die duale Basis verdient ihren Namen: Ist $\sum_{i=1}^n \lambda_i v_i^* = 0$ mit $\lambda_i \in K$, so gilt insbesondere, dass

$$0 = \sum_{i=1}^n \lambda_i v_i^*(v_j) = \lambda_j$$

ist für alle $1 \leqslant j \leqslant n$.

Sie können mit Funktionalen testen, ob ein Vektor trivial ist:

Lemma VII.1.6. *Es sei V ein beliebiger K-Vektorraum und es sei $v \in V \setminus \{0\}$. Dann gibt es ein $\varphi \in V^*$ mit $\varphi(v) \neq 0$.*

Beweis. Da $v \neq 0$, ist (v) linear unabhängig. Wir ergänzen (v) zu einer Basis von V: $\mathcal{B} = (v, v_i)_{i \in I}$. (Wenn V endlich-dimensional ist, geht das mit dem Basisergänzungssatz II.5.18; sonst brauchen Sie Satz

II.6.8.) Wie jede K-lineare Abbildung sind Linearformen durch die Werte auf einer Basis bestimmt. Wir definieren

$$\varphi\colon V \to K, \quad \varphi(v) = 1,\, \varphi(v_i) = 0 \text{ für alle } i \in I.$$

□

Satz VII.1.7. *Ist V ein endlich-dimensionaler K-Vektorraum, so liefert jede Wahl einer geordneten Basis $\mathcal{B}$ von V einen Isomorphismus*

$$g_{\mathcal{B}}\colon V \to V^*.$$

Beweis. Ist $\mathcal{B} = (v_1, \ldots, v_n)$, so ist $\mathcal{B}^* = (v_1^*, \ldots, v_n^*)$ und wir setzen natürlich $g_{\mathcal{B}}(v_i) = v_i^*$. □

Bemerkung VII.1.8. Da dieser Isomorphismus von der Basiswahl abhängt, betont man das gerne und sagt, dass dieser Isomorphismus *nicht kanonisch* ist.

Beispiel VII.1.9. Reskaliert man zum Beispiel die Basis $\mathcal{B}$ und betrachtet für $\lambda \neq 0$ die Basis $\mathcal{B}' = (\lambda v_1, \ldots, \lambda v_n)$, so gilt nach Definition der dualen Basis

$$v_i^*(v_j) = \delta_{ij} = (\lambda v_i)^*(\lambda v_j).$$

Die Linearität von $(\lambda v_i)^*$ ergibt dann $(\lambda v_i)^*(\lambda v_j) = \lambda(\lambda v_i)^*(v_j) = \delta_{ij}$. Also ist

$$(\lambda v_i)^* = \lambda^{-1} \cdot v_i^*.$$

Für die Isomorphismen $g_{\mathcal{B}}$ und $g_{\mathcal{B}'}$ erhalten wir damit die folgende Beziehung:

$$\begin{aligned} g_{\mathcal{B}'}(v_i) &= g_{\mathcal{B}'}(\lambda^{-1}\lambda v_i) \\ &= \lambda^{-1} g_{\mathcal{B}'}(\lambda v_i) \\ &= \lambda^{-1}(\lambda v_i)^* \\ &= \lambda^{-2}(v_i^*) \\ &= \lambda^{-2} g_{\mathcal{B}}(v_i). \end{aligned}$$

Die Reskalierung schlägt sich also quadratisch auf die zugehörigen Isomorphismen durch. Damit ist $g_{\mathcal{B}'} \neq g_{\mathcal{B}}$ für alle $\lambda \neq \pm 1$.

Definition VII.1.10. Sind V und W zwei K-Vektorräume und ist $f\colon V \to W$ eine K-lineare Abbildung, so heißt $f^*\colon W^* \to V^*$, $f^*(\psi) = \psi \circ f$, die *zu f duale Abbildung*.

Bemerkung VII.1.11. Beachten Sie, dass sich beim Übergang zur dualen Abbildung die Laufrichtung der Abbildung ändert!

$$\mathsf{Hom}_K(V,W) \times \mathsf{Hom}_K(W,K) \to \mathsf{Hom}_K(V,K), \quad (f,\psi) \mapsto \psi \circ f,$$

$$\begin{tikzcd} V \arrow[rr, "f^*(\psi)"] \arrow[dr, "f"'] & & K. \\ & W \arrow[ur, "\psi"'] & \end{tikzcd}$$

Lemma VII.1.12.

(a) *Die Abbildung f^* ist K-linear.*

(b) *Sind U, V, W K-Vektorräume und sind $U \xrightarrow{f} V \xrightarrow{g} W$ K-lineare Abbildungen, so ist $(g \circ f)^* = f^* \circ g^*$.*

Beweis. Wir rechnen (a) nach:

$$\begin{aligned} f^*(\lambda\psi_1 + \mu\psi_2) &= (\lambda\psi_1 + \mu\psi_2) \circ f \quad \text{(nach Definition)} \\ &= \lambda\psi_1 \circ f + \mu\psi_2 \circ f \\ &= \lambda(\psi_1 \circ f) + \mu(\psi_2 \circ f) \\ &= \lambda f^*(\psi_1) + \mu f^*(\psi_2). \end{aligned}$$

Zu (b): Für $\psi \in W^*$ ist

$$\begin{aligned}(g \circ f)^*(\psi) &= \psi \circ (g \circ f) \\ &= (\psi \circ g) \circ f \\ &= f^*(\psi \circ g) \\ &= f^*(g^*(\psi)) \\ &= f^* \circ g^*(\psi).\end{aligned}$$

$\square$

Beispiel VII.1.13. Ist $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ $\mathbb{R}$-linear mit darstellender Matrix $M(f) = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ und ist $\psi\colon \mathbb{R}^2 \to \mathbb{R}$ die Linearform $\psi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1 - x_2$, so ist $M(\psi) = (2, -1)$. Die Abbildung $f^*(\psi)$ hat die darstellende Matrix

$$M(f^*(\psi)) = M(\psi \circ f) = M(\psi) \cdot M(f) = (2, -1) \cdot \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} = (5, 4, 1).$$

Sie bildet e_1 auf 5, e_2 auf 4 und e_3 auf 1 ab.

Satz VII.1.14. *(Matrixdarstellung dualer Abbildungen) Es seien V und W endlich-dimensionale K-Vektorräume mit geordneten Basen $\mathcal{B}_1$ und $\mathcal{B}_2$. Ist $f\colon V \to W$ eine K-lineare Abbildung, so gilt*

$$M_{\mathcal{B}_1^*}^{\mathcal{B}_2^*}(f^*) = M_{\mathcal{B}_2}^{\mathcal{B}_1}(f)^t. \tag{VII.1.1}$$

Beweis. Es seien $\mathcal{B}_1 = (v_1, \ldots, v_n)$ und $\mathcal{B}_2 = (w_1, \ldots, w_m)$. Die darstellenden Matrizen für f und f^* seien $M_{\mathcal{B}_2}^{\mathcal{B}_1}(f) = (a_{ij})$ und $M_{\mathcal{B}_1^*}^{\mathcal{B}_2^*}(f^*) = (\alpha_{ij})$. Zu zeigen ist $a_{ij} = \alpha_{ji}$. Wir wissen, dass

$$f(v_j) = \sum_{k=1}^{m} a_{kj} w_k$$

und dass

$$f^*(w_j^*) = \sum_{\ell=1}^{n} \alpha_{\ell j} v_\ell^*.$$

Wir wenden ein festes w_i^* an und erhalten

$$w_i^*(f(v_j)) = w_i^* \sum_{k=1}^{m} a_{kj} w_k = \sum_{k=1}^{m} a_{kj} w_i^*(w_k) = a_{ij}.$$

Aber es gilt auch $w_i^*(f(v_j)) = f^*(w_i^*)(v_j)$ und somit

$$w_i^*(f(v_j)) = f^*(w_i^*)(v_j) = \sum_{\ell=1}^{n} \alpha_{\ell i} v_\ell^*(v_j) = \alpha_{ji}.$$

$\square$

Beispiel VII.1.15. Wir betrachten wieder das Beispiel von vorher. Es sei also f die lineare Abbildung mit $M(f) = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ und $\psi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1 - x_2$. Wir betrachten die duale Basis zur Standardbasis (e_1^*, e_2^*). Dann ist $\psi = 2e_1^* - e_2^*$ und es hat als Vektor des $(\mathbb{R}^2)^*$ die Darstellung $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Multiplizieren wir die Transponierte von $M(f)$ mit diesem Vektor, so erhalten wir

$$M(f)^t \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix}$$

und das ist der Vektor $5e_1^* + 4e_2^* + 1e_3^*$ im $(\mathbb{R}^3)^*$.

Definition VII.1.16. Ist V ein K-Vektorraum und ist M eine Teilmenge von V, so heißt
$$M^0 := \{\varphi \in V^*, \varphi(m) = 0 \text{ für alle } m \in M\}$$
der *Annulator von M*.

Lemma VII.1.17.
(a) *M^0 ist für jedes $M \subset V$ ein Untervektorraum von V^*.*
(b) *Ist $\dim_K V < \infty$, so gilt $\mathsf{Span}_K(M) = \{v \in V, \varphi(v) = 0$ für alle $\varphi \in M^0\}$.*
(c) *Ist $M = U$ ein Untervektorraum von V und ist $\dim_K V < \infty$, so gilt*
$$U = \{v \in V, \varphi(v) = 0 \textit{ für alle } \varphi \in U^0\}.$$
(d) *Ist $\dim_K V < \infty$ und ist $U \subset V$ ein Untervektorraum in V, so gilt $\dim_K U^0 = \dim_K V - \dim_K U$.*

Beweis. Die Behauptung (a) folgt durch direktes Nachrechnen. Für (b) nehmen wir ein $v \in \mathsf{Span}_K(M)$, also $v = \sum_{i=1}^n \lambda_i m_i$ mit $m_i \in M$. Ist $\varphi \in M^0$, so gilt
$$\varphi(v) = \sum_{i=1}^n \lambda_i \varphi(m_i) = 0,$$
also gilt $\mathsf{Span}_K(M) \subset \{v \in V, \varphi(v) = 0$ für alle $\varphi \in M^0\}$. Wir nehmen an, dass es ein $v \in V$ gibt mit $\varphi(v) = 0$ für alle $\varphi \in M^0$, aber mit $v \notin \mathsf{Span}_K(M)$. Es sei $(u_1, \ldots, u_\ell)$ eine geordnete Basis $\mathsf{Span}_K(M)$. Dann ist nach Annahme $(u_1, \ldots, u_\ell, v)$ linear unabhängig und kann daher zu einer Basis $(u_1, \ldots, u_\ell, v, w_1, \ldots, w_r)$ von V ergänzt werden. Wir setzen
$$\begin{aligned} \varphi(u_i) &= 0, 1 \leqslant i \leqslant \ell, \\ \varphi(v) &= 1, \\ \varphi(w_i) &= 1, 1 \leqslant i \leqslant r. \end{aligned}$$
Dann ist $\varphi \in M^0$ nach Konstruktion, aber $\varphi(v) = 1$. Das ist ein Widerspruch, also muss v im Erzeugnis von M liegen. Behauptung (c) folgt aus (b), weil $\mathsf{Span}_K U = U$ gilt in (c). Für (d) wählen wir eine geordnete Basis $\mathcal{B}_1 = (u_1, \ldots, u_k)$ von U. Wir ergänzen diese zu einer geordneten Basis $\mathcal{B} = (u_1, \ldots, u_k, w_{k+1}, \ldots, w_n)$ von V. Die zu $\mathcal{B}$ duale Basis von V^* ist also $\mathcal{B}^* = (u_1^*, \ldots, u_k^*, w_{k+1}^*, \ldots, w_n^*)$.

Es sei $\varphi \in V^*$. Dann können wir φ schreiben als
$$\varphi = \sum_{i=1}^k \lambda_i u_i^* + \sum_{j=k+1}^n \mu_j w_j^*. \tag{VII.1.2}$$
Mit dieser Darstellung ist φ genau dann in U^0, wenn $\varphi(u_j) = 0$ ist für alle $1 \leqslant j \leqslant k$ und das wiederum ist äquivalent dazu, dass in (VII.1.2) $\lambda_1 = \ldots = \lambda_k = 0$ ist. Dies ist genau dann der Fall, wenn $\varphi \in \mathsf{Span}_K(w_{k+1}^*, \ldots, w_n^*)$ gilt.

Damit ist $(w_{k+1}^*, \ldots, w_n^*)$ eine Basis von U^0 und
$$\dim_K U^0 = n - k = \dim_K V - \dim_K U.$$
□

Eine konkrete Anwendung ist die Beschreibung von Hyperebenen durch Funktionale: Ist V ein n-dimensionaler Vektorraum und ist $U \subset V$ ein $(n-1)$-dimensionaler Untervektorraum, so wissen wir mit Lemma VII.1.17, dass $\dim_K U^0 = n - (n-1) = 1$ ist. Damit ist $U^0 = \mathsf{Span}_K(\varphi)$ für jedes $\varphi \in U^0 \setminus \{0\}$. Die Hyperebene U ist dann
$$U = \{v \in V, \varphi(v) = 0\}.$$

Beispiel VII.1.18. Es sei $U \subset \mathbb{R}^3$,
$$U = \left\{ \begin{pmatrix} x_1 \\ x_1 \\ x_3 \end{pmatrix}, x_1, x_3 \in \mathbb{R} \right\}.$$

Dann können wir U auch beschreiben durch das Funktional

$$\varphi \in \mathbb{R}^3, \quad \varphi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 - x_2,$$

weil

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, x_i \in \mathbb{R}, \varphi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \right\}.$$

Eine ähnliche Beschreibung erhalten wir für affine Hyperebenen, also für affine Räume der Form $v_0 + U$, wobei $U \subset V$ eine Hyperebene ist. Ein $v \in V$ ist genau dann in $v_0 + U$, wenn $v - v_0 \in U$ ist. Ist $\varphi \in U^0 \setminus \{0\}$, so gilt:

$$v - v_0 \in U \Leftrightarrow \varphi(v - v_0) = 0 \Leftrightarrow \varphi(v) = \varphi(v_0).$$

Damit ist also φ auf $v_0 + U$ konstant mit Wert $\varphi(v_0)$. Dies liefert eine koordinatenfreie Variante der Hesse-Normalform affiner Hyperebenen: Ist $c := \varphi(v_0)$, so gilt

$$v_0 + U = \{v \in V, \varphi(v) = c\}.$$

Beispiel VII.1.19. Ist $v_0 = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} \in \mathbb{R}^3$ und nehmen wir wieder das Funktional aus dem obigen Beispiel, so ist

$$\varphi \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} = 4 - 2 = 2.$$

Damit ist

$$v_0 + U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3, \varphi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \right\} = \left\{ \begin{pmatrix} x_1 \\ x_1 - 2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \right\}.$$

Wir benutzen Annulatoren auch, um das Bild und den Kern dualer Abbildungen zu beschreiben:

Satz VII.1.20. *(Bild-Kern-Korrespondenz) Es seien V und W endlich-dimensionale K-Vektorräume und $f\colon V \to W$ sei K-linear. Dann gilt.*

(a) $\mathrm{Bild}(f^*) = (\ker(f))^0$.
(b) $\ker(f^*) = (\mathrm{Bild}(f))^0$.

Beweis. Für die erste Enthaltenseinsrelation in (a) brauchen wir nicht, dass V endlich-dimensional ist: Ist $\varphi \in \mathrm{Bild}(f^*)$, so gibt es ein $\psi \in W^*$ mit $\varphi = f^*(\psi) = \psi \circ f$. Ist $v \in \ker(f)$, so folgt

$$\varphi(v) = (\psi \circ f)(v) = \psi(f(v)) = \psi(0) = 0$$

und damit ist $\varphi \in (\ker(f))^0$.

Ist umgekehrt $\varphi \in (\ker(f))^0$ und ist $r = \mathsf{rg}(f)$, so wählen wir geordnete Basen $\mathcal{B}_1 = (v_1, \ldots, v_n)$ von V und $\mathcal{B}_2 = (w_1, \ldots, w_m)$ von W, so dass

$$M^{\mathcal{B}_1}_{\mathcal{B}_2}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

ist. Wir definieren $\psi \in W^*$ als

$$\psi(w_j) := \begin{cases} \varphi(v_j), & 1 \leqslant j \leqslant r, \\ 0, & r < j. \end{cases}$$

Dann ist

$$f^*(\psi)(v_j) = \psi(f(v_j)) = \psi(w_j) = \begin{cases} \varphi(v_j), & 1 \leqslant j \leqslant r, \\ 0, & r < j. \end{cases}$$

Da $\varphi \in (\ker(f))^0 = (\mathsf{Span}_K(v_{r+1}, \ldots, v_n))^0$, ist φ ebenfalls trivial auf $v_{r+1}, \ldots, v_n$ und damit ist

$$\varphi = \psi \circ f = f^*(\psi)$$

und $(\ker(f))^0 \subset \mathrm{Bild}(f^*)$.

Für (b) rechnen wir lediglich nach:

$$\begin{aligned}
(\mathrm{Bild}(f))^0 &= \{\varphi \in W^*, \varphi(w) = 0 \text{ für alle } w \in \mathrm{Bild}(f)\} \\
&= \{\varphi \in W^*, \varphi(f(v)) = 0 \text{ für alle } v \in V\} \\
&= \{\varphi \in W^*, f^*(\varphi)(v) = 0 \text{ für alle } v \in V\} \\
&= \ker(f^*).
\end{aligned}$$

□

Korollar VII.1.21. *Es seien V und W endlich-dimensionale K-Vektorräume und $f\colon V \to W$ sei K-linear. Dann ist $\mathsf{rg}(f) = \mathsf{rg}(f^*)$.*

Dies liefert wiederum, dass der Zeilen- und Spaltenrang einer Matrix übereinstimmen. Insbesondere wissen wir schon, dass das obige Resultat wahr ist. Wir beweisen es trotzdem noch einmal:

Beweis.

$$\begin{aligned}
\mathsf{rg}(f^*) &= \dim_K \mathrm{Bild}(f^*) \quad \text{(nach Definition)} \\
&= \dim_K (\ker(f))^0 \quad \text{(nach Satz VII.1.20)} \\
&= \dim_K V - \dim_K \ker(f) \quad \text{(nach Lemma VII.1.17)} \\
&= \dim_K \mathrm{Bild}(f) \quad \text{(mit der Dimensionsformel)} \\
&= \mathsf{rg}(f).
\end{aligned}$$

□

Definition VII.1.22. Für einen K-Vektorraum V heißt

$$V^{**} := (V^*)^* = \mathsf{Hom}_K(V^*, K)$$

der *Bidualraum von V*.

Bemerkung VII.1.23. Dual- und Bidualräume sind wichtige Objekte zum Beispiel in der Funktionalanalysis und der Physik.

Zu einem festen $v \in V$ betrachten wir die Abbildung

$$i_V^v\colon V^* \to K, \quad i_V^v(\varphi) = \varphi(v),$$

das heißt, i_V^v wertet einfach ein Funktional auf v aus.

- Die Abbildung i_V^v ist linear: Für $\lambda, \mu \in K$ und $\varphi, \psi \in V^*$ ist
$$i_V^v(\lambda\varphi + \mu\psi) = (\lambda\varphi + \mu\psi)(v) = \lambda\varphi(v) + \mu\psi(v) = \lambda i_V^v(\varphi) + \mu i_V^v(\psi).$$
- Wir können aber auch das $v \in V$ variieren und dann gilt für $\lambda, \mu \in K$ und $v_1, v_2 \in V$:
$$i_V^{\lambda v_1 + \mu v_2}(\varphi) = \varphi(\lambda v_1 + \mu v_2) = \lambda\varphi(v_1) + \mu\varphi(v_2) = \lambda i_V^{v_1}(\varphi) + \mu i_V^{v_2}(\varphi).$$
- Wir erhalten also insgesamt eine lineare Abbildung
$$\begin{aligned} i_V\colon V &\to V^{**}, \\ v &\mapsto i_V^v. \end{aligned}$$

Satz VII.1.24. *(Vergleich von Vektorräumen mit ihren Bidualräumen)*

(a) *i_V ist immer ein Monomorphismus.*
(b) *Ist $\dim_K V < \infty$, so ist i_V ein Isomorphismus.*

(c) *Ist $f \in \mathsf{Hom}_K(V,W)$ für einen K-Vektorraum W, so ist*

$$f^{**} \circ i_V = i_W \circ f$$

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "i_V"'] & W \arrow[d, "i_W"] \\ V^{**} \arrow[r, "f^{**}"] & W^{**} \end{tikzcd}$$

Beweis. Zu (a): Ist $i_V^v = 0$, so gilt für alle $\varphi \in V^*$, dass

$$0 = i_V^v(\varphi) = \varphi(v)$$

gilt. Mit Lemma VII.1.6 folgt dann aber schon, dass $v = 0$ ist, also ist i_V immer ein Monomorphismus.

Zu (b): Ist $\dim_K V < \infty$, so gilt

$$\dim_K V^{**} = \dim_K (V^*)^* = \dim_K V^* = \dim_K V.$$

Da i injektiv ist, ist es somit auch ein Isomorphismus.

Zu (c): Wir schreiben aus, was die Definitionen ergeben: Für $f \in \mathsf{Hom}_K(V,W)$, $v \in V$ und $\psi \in W^*$ ist

$$\begin{aligned} ((f^{**}) \circ i_V^v)(\psi) &= i_V^v \circ f^*(\psi) \\ &= i_V^v(\psi \circ f) \\ &= (\psi \circ f)(v) \end{aligned}$$

und

$$(i_W \circ f)(v)(\psi) = i_W^{f(v)}(\psi) = \psi(f(v)).$$

□

Bemerkung VII.1.25. Ist V unendlich-dimensional, so ist $i_V(V)$ in der Regel ein echter Untervektorraum von V^{**}. Für endlich-dimensionales V heißt i_V auch *kanonischer Isomorphismus*, weil i_V von keinerlei Wahlen, insbesondere von keiner Basiswahl, abhängig ist.

VII.2. Bilinearformen

Wir hatten schon das Standardskalarprodukt auf dem K^n kennengelernt und verallgemeinern diesen Begriff wie folgt:

Definition VII.2.1.

(a) Sind V und W zwei K-Vektorräume, so heißt eine Abbildung

$$\beta \colon V \times W \to K$$

eine *Bilinearform*, falls für alle $\lambda \in K$, $v, v' \in V$ und $w, w' \in W$ gilt:

$$\begin{aligned} \beta(v + v', w) &= \beta(v,w) + \beta(v',w), \\ \beta(v, w + w') &= \beta(v,w) + \beta(v,w') \\ \beta(\lambda v, w) &= \lambda\beta(v,w) = \beta(v, \lambda w). \end{aligned}$$

(b) Eine Bilinearform β heißt *nicht-ausgeartet im ersten (beziehungsweise zweiten) Argument*, falls aus $\beta(v,w) = 0$ für alle $w \in W$ folgt, dass $v = 0$ gilt (beziehungsweise, falls aus $\beta(v,w) = 0$ für alle $v \in V$ folgt, dass $w = 0$ gilt). Gilt beides, so heißt β *nicht-ausgeartet*.

(c) Wir verwenden die Notation $\mathrm{Bil}_K(V,W)$ für

$$\mathrm{Bil}_K(V,W) := \{\beta \colon V \times W \to K, \beta \text{ ist Bilinearform}\}.$$

Beispiele VII.2.2.

- Ist $A \in M(m \times n, K)$, so ist die Abbildung

  $$\beta_A \colon K^m \times K^n \to K, \quad (x,y) \mapsto x^t \cdot A \cdot y$$

  eine Bilinearform.

- Ein wichtiges Beispiel ist die *Minkowski-Form*: Ist $K = \mathbb{R}$ und $m = n = 4$, so gibt $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ die Bilinearform
$$\beta_A\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}\right) = (x_1, x_2, x_3, x_4)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4.$$
Der $\mathbb{R}^4$ wird hierbei in drei Raumkoordinaten (die ersten drei) und eine Zeitkoordinate (die letzte) aufgeteilt. Vorsicht: Die Konventionen, welches die Zeit- und welches die Raumkoordinaten sind, ist nicht einheitlich.
Hermann Minkowski, 1864–1909.
- Für $n = m$ und $A = E_n$ bekommen wir als β_A genau das Standardskalarprodukt auf dem K^n:
$$\beta_{E_n}\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{i=1}^n x_iy_i.$$
- Ist V ein K-Vektorraum, so ist die Auswertungsabbildung
$$\varepsilon\colon V^* \times V \to V, \quad (\varphi, v) \mapsto \varphi(v) \tag{VII.2.1}$$
eine Bilinearform. Lemma VII.1.6 impliziert, dass ε im zweiten Argument nicht-ausgeartet ist: Ist $\varphi(v) = 0$ für alle $\varphi \in V^*$, so ist $v = 0$. Umgekehrt ist natürlich $\varphi = 0$, falls $\varphi(v) = 0$ ist für alle $v \in V$, also ist ε auch nicht-ausgeartet im ersten Argument, also insgesamt nicht-ausgeartet.

Ist die Minkowski-Form nicht-ausgeartet? Was ist mit dem Standardskalarprodukt?

Bemerkung VII.2.3. Die Menge $\mathrm{Bil}_K(V, W)$ ist selbst ein K-Vektorraum.
Ist $\beta \in \mathrm{Bil}_K(V, W)$, so erhalten wir für festes $v \in V$ eine Abbildung
$$\beta(v, -)\colon W \to K,$$
die K-linear ist. Damit ist $\beta(v, -) \in W^*$ für alle $v \in V$ und wir können v ebenfalls variieren und erhalten:
$$\beta_V\colon V \to W^*, \quad \beta_V(v) := \beta(v, -).$$
Sie können natürlich auch das zweite Argument benutzen und bekommen
$$\beta_W\colon W \to V^*, \quad \beta_W(w) := \beta(-, w).$$

Lemma VII.2.4. *Es seien V und W zwei K-Vektorräume. Dann gilt:*
(a) *Die Abbildung*
$$\delta\colon \mathrm{Bil}_K(V, W) \to \mathsf{Hom}_K(V, W^*)$$
$$\beta \mapsto \beta_V$$
ist ein Isomorphismus von K-Vektorräumen.
(b) *Gilt zusätzlich, dass $\dim_K V, \dim_K W < \infty$ sind, so ist*
$$\dim_K \mathrm{Bil}_K(V, W) = \dim_K V \cdot \dim_K W.$$
(c) *Die Bilinearform β ist genau dann nicht-ausgeartet im ersten Argument, falls β_V injektiv ist.*

Beweis. Für (a) definieren wir die Abbildung
$$\gamma\colon \mathsf{Hom}_K(V, W^*) \to \mathrm{Bil}_K(V, W)$$
als $\gamma(f) = \varepsilon(f(-), -)$, so dass $\gamma(f)(v, w) = \varepsilon(f(v), w) = f(v)(w)$ ist. Hierbei ist ε die Auswertungabbildung aus (VII.2.1). Wir zeigen, dass γ invers zu δ ist. Die Komposition $\gamma \circ \delta$ ergibt auf $\beta \in \mathrm{Bil}_K(V, W)$ die Bilinearform mit Werten
$$(\gamma \circ \delta)(\beta)(v, w) = \gamma(\delta(\beta))(v, w) = \delta(\beta)(v)(w) = \beta(v, w)$$

und daher ist $\gamma \circ \delta(\beta) = \beta$.

Umgekehrt, ist $\delta \circ \gamma$ ausgewertet auf einem $f \in \mathsf{Hom}_K(V, W^*)$ das Element in $\mathsf{Hom}_K(V, W^*)$ mit Wert

$$(\delta \circ \gamma)(f)(v) = \delta(\gamma(f))(v) = \gamma(f)(v, -) = f(v).$$

Für (b) berechnen wir die Dimension mit (a) als

$$\dim_K \operatorname{Bil}_K(V, W) = \dim_K \mathsf{Hom}_K(V, W^*) = \dim_K V \cdot \dim_K W^* = \dim_K V \cdot \dim_K W.$$

Zu (c): Die Abbildung β_V ist genau dann nicht injektiv, falls es ein $0 \neq v \in V$ gibt mit $\beta_V(v) = 0$ und das ist äquivalent zu

$$\beta_V(v)(w) = 0 \text{ für alle } w \in W.$$

Das ist genau dann der Fall, wenn β im ersten Argument ausgeartet ist. □

Bemerkung VII.2.5. Ist $\beta \in \operatorname{Bil}_K(V, W)$, ist $\mathcal{B}_1 = (v_1, \ldots, v_n)$ eine geordnete Basis von V und ist $\mathcal{B}_2 = (w_1, \ldots, w_m)$ eine geordnete Basis von W, so können wir die Matrix $B = (\beta_{ij})$ mit $\beta_{ij} = \beta(v_i, w_j)$ bilden.

- Für die Matrix $B = (\beta_{ij})$ gilt, dass $B \in M(n \times m, K)$ ist.
- Die Bilinearform β ist durch die Matrix B eindeutig bestimmt, weil für $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{j=1}^m \mu_j w_j$ wegen der Bilinearität von β gilt:

$$\begin{aligned}
\beta(v, w) &= \beta\Big(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^m \mu_j w_j\Big) \\
&\sum_{i=1}^n \lambda_i \sum_{j=1}^m \mu_j \beta(v_i, w_j) \\
&= \sum_{i=1}^n \lambda_i \sum_{j=1}^m \mu_j \beta_{ij} \\
&= (\lambda_1, \ldots, \lambda_n) \cdot B \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}.
\end{aligned}$$

Definition VII.2.6. Für $\beta \in \operatorname{Bil}_K(V, W)$ heißt $B = (\beta_{ij})$ die *darstellende Matrix der Bilinearform β bezüglich der Basen $\mathcal{B}_1$ und $\mathcal{B}_2$.*

Notation VII.2.7. Um diese darstellenden Matrizen von darstellenden Matrizen linearer Abbildungen abzugrenzen, schreibt man oft

$$B = (\beta_{ij}) = M_{\mathcal{B}_1, \mathcal{B}_2}(\beta).$$

Sie ahnen schon was kommt. Wir machen uns klar, was bei einem Basiswechsel passiert:

Satz VII.2.8. *(Transformationsformel) Es seien V und W endlich-dimensionale K-Vektorräume, $\mathcal{B}_1, \mathcal{B}_1'$ seien geordnete Basen von V und $\mathcal{B}_2, \mathcal{B}_2'$ seien geordnete Basen von W. Ist $\beta \in \operatorname{Bil}_K(V, W)$, so gilt:*

$$M_{\mathcal{B}_1, \mathcal{B}_2}(\beta) = (T^{\mathcal{B}_1}_{\mathcal{B}_1'})^t \cdot M_{\mathcal{B}_1', \mathcal{B}_2'}(\beta) \cdot T^{\mathcal{B}_2}_{\mathcal{B}_2'}.$$

Insbesondere gilt für $V = W$ und $\mathcal{B}_1 = \mathcal{B}_2$, $\mathcal{B}_1' = \mathcal{B}_2'$:

$$M_{\mathcal{B}_1, \mathcal{B}_1}(\beta) = (T^{\mathcal{B}_1}_{\mathcal{B}_1'})^t \cdot M_{\mathcal{B}_1', \mathcal{B}_1'}(\beta) \cdot T^{\mathcal{B}_1}_{\mathcal{B}_1'}.$$

Beweis. Wir legen die Notation fest: Es seien

$$\begin{aligned}
\mathcal{B}_1 &= (v_1, \ldots, v_n), \quad & \mathcal{B}_1' &= (v_1', \ldots, v_n'), \\
\mathcal{B}_2 &= (w_1, \ldots, w_m), \quad & \mathcal{B}_2' &= (w_1', \ldots, w_m'),
\end{aligned}$$

also

$$v_i = \sum_{\ell=1}^n (T^{\mathcal{B}_1}_{\mathcal{B}_1'})_{\ell i} v_\ell' \text{ und } w_j = \sum_{k=1}^m (T^{\mathcal{B}_2}_{\mathcal{B}_2'})_{kj} w_k'.$$

Stures Nachrechnen ergibt dann die behauptete Formel:

$$\begin{aligned}
M_{\mathcal{B}_1,\mathcal{B}_2}(\beta)_{ij} &= \beta(v_i, w_j)\\
&= \beta\left(\sum_{\ell=1}^{n}(T_{\mathcal{B}_1'}^{\mathcal{B}_1})_{\ell i}v_\ell', \sum_{k=1}^{m}(T_{\mathcal{B}_2'}^{\mathcal{B}_2})_{ki}w_k'\right)\\
&= \sum_{\ell=1}^{n}\sum_{k=1}^{m}(T_{\mathcal{B}_1'}^{\mathcal{B}_1})_{\ell i}\beta(v_\ell', w_k')(T_{\mathcal{B}_2'}^{\mathcal{B}_2})_{ki}\\
&= \sum_{\ell=1}^{n}\sum_{k=1}^{m}(T_{\mathcal{B}_1'}^{\mathcal{B}_1})_{\ell i}M_{\mathcal{B}_1',\mathcal{B}_2'}(\beta)_{\ell k}(T_{\mathcal{B}_2'}^{\mathcal{B}_2})_{ki}\\
&= \left((T_{\mathcal{B}_1'}^{\mathcal{B}_1})^t \cdot M_{\mathcal{B}_1',\mathcal{B}_2'}(\beta)\cdot T_{\mathcal{B}_2'}^{\mathcal{B}_2}\right)_{ij}.
\end{aligned}$$

□

Für Matrixdarstellungen von Endomorphismen hatten wir die Äquivalenzrelation der Ähnlichkeit betrachtet. Für Bilinearformen erhalten wir einen analogen, aber doch verschiedenen, Begriff.

Definition VII.2.9. Für $B, C \in M(n \times n, K)$ heißt B *kongruent zu* C, falls es ein $S \in GL_n(K)$ gibt mit

$$C = S^t \cdot B \cdot S.$$

Bemerkung VII.2.10.

- Kongruenz ist eine Äquivalenzrelation auf der Menge $M(n \times n, K)$. Rechnen Sie das bitte nach.
- Ist B kongruent zu C mit $C = S^t \cdot B \cdot S$, so gilt

  $$\det(C) = \det(S^t BS) = \det(S^t)\det(B)\det(S) = \det(S)^2\det(B).$$

  Da $S \in GL_n K$ ist, ist $\det(S) \neq 0$.
- B ist genau dann kongruent zu C, wenn B und C die gleiche Bilinearform β bezüglich verschiedener Basen darstellen.

VII.3. Bilinearformen mit speziellen Eigenschaften

Im gesamten Abschnitt sei K ein Körper und V ein K-Vektorraum.

Definition VII.3.1. Eine Bilinearform $\beta\colon V \times V \to K$ heißt

- *symmetrisch*, falls $\beta(v, w) = \beta(w, v)$ für alle $v, w \in V$,
- *alternierend*, falls $\beta(v, w) = -\beta(w, v)$ für alle $v, w \in V$.

Nur falls die Charakteristik von K 2 ist, stimmen beide Begriffe überein.

Beispiele VII.3.2.

(a) Das Standardskalarprodukt auf dem K^n

$$\langle -, - \rangle\colon K^n \times K^n \to K, \quad \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i$$

ist symmetrisch.

(b) Ist $V = C^0([0,1], \mathbb{R})$, so ist

$$\beta(f, g) := \int_0^1 f(x)g(x)dx$$

symmetrisch.

(c) Ist $V = K^{2n}$ und ist B die Matrix

$$B = \begin{pmatrix} 0 & -E_n \\ E_n & 0 \end{pmatrix} \in M(2n \times 2n, K),$$

so ist $B^t = -B$. Die assoziierte Bilinearform

$$\beta_B(v,w) = v^t \cdot B \cdot w$$

ist alternierend. Im Spezialfall $n=1$ erhalten wir

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M(2\times 2, K)$$

und

$$\beta_B\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_2 y_1 - x_1 y_2.$$

Bemerkung VII.3.3.

- Ist $B \in M(n\times n, K)$ und ist $B^t = -B$, so gilt
$$\det(B) = \det(B^t) = \det(-B) = (-1)^n \det(B).$$
Falls $2 \neq 0$ ist in K und falls n ungerade ist, erzwingt dies, dass $\det(B) = 0$ ist.
- Ist $B \in M(n\times n, K)$ und ist β_B die assoziierte Bilinearform mit $\beta_B(v,w) = v^t B w$, so ist β_B genau dann symmetrisch, falls $B^t = B$ gilt: Ist $B = (\beta_{ij})$, $\mathcal{B} = (v_1, \ldots, v_n)$, $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{j=1}^n \mu_j v_j$, dann ist
$$\beta(v,w) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \beta_{ij}.$$
Dies ist genau dann gleich $\beta(w,v)$ für alle v, w, wenn $\beta_{ij} = \beta_{ji}$ ist.
- Vorsicht: Symmetrische Matrizen, also $B \in M(n\times n, K)$ mit $B^t = B$ bilden keine Untergruppe mit der Matrizenmultiplikation. Zum einen muss B nicht invertierbar sein und zum anderen ist $(AB)^t = B^t A^t$, also selbst, wenn $A^t = A$ und $B^t = B$ gilt, ist $(AB)^t = BA$ und nicht gleich AB.

Definition VII.3.4. Ist $\beta\colon V\times V \to K$ eine symmetrische Bilinearform, so heißt

$$q\colon V \to K, \quad q(v) := \beta(v,v)$$

die zu β assoziierte quadratische Form.

Bemerkung VII.3.5. Das *quadratisch* kommt daher, dass die Bilinearität von β dafür sorgt, dass für ein $\lambda \in K$ gilt

$$q(\lambda v) = \beta(\lambda v, \lambda v) = \lambda^2 \beta(v,v) = \lambda^2 q(v).$$

Beispiel VII.3.6. Es sei $V = \mathbb{R}^2$ und $B = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ mit $a, b \in \mathbb{R}$. Dann ist

$$\beta_B\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1, x_2)\begin{pmatrix} a & b \\ b & a \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = a(x_1y_1 + x_2y_2) + b(y_1x_2 + x_1y_2).$$

Die assoziierte quadratische Form ist

$$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a(x_1^2 + x_2^2) + 2bx_1x_2.$$

(1) Betrachten wir den Spezialfall, bei dem $b = 0$ ist, aber $a \neq 0$, so ist

$$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a(x_1^2 + x_2^2).$$

Diese quadratische Form können Sie sich anhand ihrer *Niveaumengen*

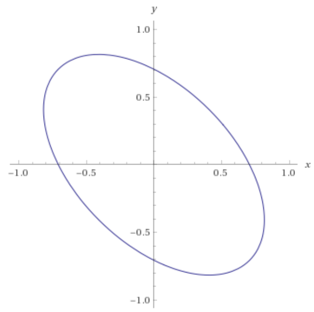
$$N_r := \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = r\}$$

veranschaulichen. Hier gilt

$$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = r \Leftrightarrow x_1^2 + x_2^2 = \frac{r}{a}.$$

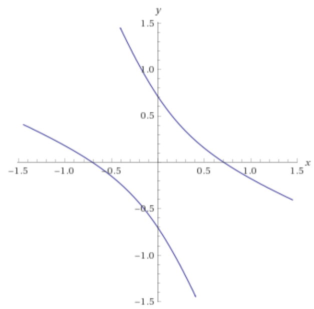
Die ergibt als Niveaumenge eine Kreislinie mit Radius $\frac{r}{a}$.

(2) Der Fall $a = 2$, $b = 1$ ergibt für $r = 1$ eine Ellipse.



Wolfram Alpha

(3) Für $a = 2$, $b = 3$ und $r = 1$ erhalten Sie als Niveaumenge eine Hyperbel.



Wolfram Alpha

Satz VII.3.7. *(Polarisierungsformel) Ist $2 \neq 0$ in K, so gilt für jede symmetrische Bilinearform β auf V und ihre assoziierte quadratische Form q:*

$$\beta(v, w) = \frac{1}{2}(q(v + w) - q(v) - q(w)).$$

Beweis. In der folgenden Rechnung benutzen wir erst die Definition von q, dann die Bilinearität und dann die Symmetrie von β:

$$\begin{aligned} q(v + w) - q(v) - q(w) &= \beta(v + w, v + w) - \beta(v, v) - \beta(w, w) \\ &= \beta(v, v) + \beta(v, w) + \beta(w, v) + \beta(w, w) - \beta(v, v) - \beta(w, w) \\ &= 2\beta(v, w). \end{aligned}$$

□

Bemerkung VII.3.8. Ist $V = K^n$ und ist $B = (\beta_{ij})$ die darstellende Matrix von β bezüglich der Standardbasis, so ist

$$\begin{aligned} q\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= (x_1, \ldots, x_n) \cdot B \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \sum_{i,j=1}^{n} x_i \beta_{ij} x_j \\ &= \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{1 \leqslant i < j \leqslant n} 2\beta_{ij} x_i x_j. \end{aligned}$$

Bei der letzten Umformung haben wir benutzt, dass $B = B^t$ gilt.

Wir wollen einen angemessenen Skalarproduktsbegriff für den $\mathbb{C}^n$ haben. Wir haben die Längenmessung im $\mathbb{R}^n$ mit

$$||x|| = \sqrt{\langle x, x \rangle}$$

für $x \in \mathbb{R}^n$. Für $\mathbb{C}$ hatten wir die Norm einer komplexen Zahl z definiert als

$$|z| = \sqrt{z \cdot \bar{z}}.$$

Wir setzen diese Definition auf den $\mathbb{C}^n$ für $n \in \mathbb{N}$ fort:

Definition VII.3.9. Das *Standardskalarprodukt des* $\mathbb{C}^n$ ist

$$\langle z, w \rangle := z_1 \bar{w}_1 + \ldots + z_n \bar{w}_n \text{ für } z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \text{ und } w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}. \tag{VII.3.1}$$

Bemerkung VII.3.10.

(a) Dieses Skalar ist additiv in beiden Komponenten. Es gilt:

$$\langle z + z', w \rangle = \langle z, w \rangle + \langle z', w \rangle$$

und

$$\langle z, w + w' \rangle = \langle z, w \rangle + \langle z, w' \rangle$$

für alle $z, z', w, w' \in \mathbb{C}^n$.

(b) Ist $\lambda \in \mathbb{C}$, so gilt auch

$$\langle \lambda z, w \rangle = \lambda \langle z, w \rangle$$

für alle $z, w \in \mathbb{C}^n$, weil sich in (VII.3.1) eine Streckung mit λ im ersten Argument aus der Summe herauszieht. Dahingegen gilt

$$\langle z, \lambda w \rangle = \overline{\lambda} \langle z, w \rangle,$$

weil die Einträge der zweiten Komponente konjugiert werden in (VII.3.1).

(c) Dieses Skalarprodukt liefert einen Längenbegriff, indem wir setzen:

$$||z|| := \sqrt{\langle z, z \rangle} = \sqrt{\sum_{i=1}^{n} z_i \bar{z}_i}.$$

(d) Die Inklusion $\mathbb{R}^n \subset \mathbb{C}^n$

$$\mathbb{R}^n \ni x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 + 0 \cdot i \\ \vdots \\ x_n + 0 \cdot i \end{pmatrix} \in \mathbb{C}^n$$

ist verträglich mit der Definition der Skalarprodukte im $\mathbb{R}^n$ und $\mathbb{C}^n$, weil für $x, y \in \mathbb{R}^n$ gilt, dass das Skalarprodukt im $\mathbb{C}^n$ angewandt auf x und y den Wert

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i = \sum_{i=1}^{n} x_i y_i$$

ergibt, weil für reelle Zahlen r gilt, dass $\bar{r} = r$ ist. Also stimmt der obige Ausdruck mit dem Wert des Standardskalarprodukts auf dem $\mathbb{R}^n$ von x und y überein.

(e) Aber Vorsicht: Wir haben $\mathbb{C}$ mit $\mathbb{R}^2$ identifiziert und genauso können Sie den $\mathbb{C}^n$ mit dem $\mathbb{R}^{2n}$ identifizieren, indem Sie ein

$$a = \begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \\ y_n \end{pmatrix} \in \mathbb{R}^{2n}$$

auf

$$\begin{pmatrix} x_1 + iy_1 \\ \vdots \\ x_n + iy_n \end{pmatrix} = z \in \mathbb{C}^n$$

abbilden.

Ist $b = \begin{pmatrix} x_1' \\ y_1' \\ \vdots \\ x_n' \\ y_n' \end{pmatrix} \in \mathbb{R}^{2n}$, so erhalten wir für das Skalarprodukt im $\mathbb{R}^{2n}$

$$\langle a, b \rangle = \sum_{i=1}^{n} (x_i x_i' + y_i y_i'),$$

wohingegen das Skalarprodukt im $\mathbb{C}^n$ von z und $w = \begin{pmatrix} x_1' + iy_1' \\ \vdots \\ x_n' + iy_n' \end{pmatrix}$ den Wert

$$\langle z, w \rangle = \sum_{i=1}^{n} (x_i + iy_i)(x_i' - iy_i') = \sum_{i=1}^{n} (x_i x_i' + y_i y_i') + i \sum_{i=1}^{n} (x_i y_i' - x_i' y_i) = \langle z, w \rangle + i \sum_{i=1}^{n} (x_i y_i' - x_i' y_i)$$

hat. Der Isomorphismus $\mathbb{C}^n \cong \mathbb{R}^{2n}$ von $\mathbb{R}$-Vektorräumen ist also *nicht* mit den jeweiligen Skalarprodukten kompatibel.

Definition VII.3.11. Es sei V ein $\mathbb{C}$-Vektorraum.

(a) Eine Abbildung $f\colon V \to V$ heißt *semi-linear*, falls $f(v_1 + v_2) = f(v_1) + f(v_2)$ und $f(\lambda v) = \bar{\lambda} f(v)$ gilt für alle $v_1, v_2, v \in V$ und $\lambda \in \mathbb{C}$.

(b) Eine Abbildung $\beta\colon V \times V \to \mathbb{C}$ heißt eine *Sesquilinearform*, falls β im ersten Argument linear ist und im zweiten Argument semi-linear ist.

(c) Eine Sesquilinearform β heißt *hermitesch*, falls für alle $v, w \in V$ gilt:

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

Beispiel VII.3.12. Das Standardskalarprodukt auf dem $\mathbb{C}^n$ ist eine hermitesche Sesquilinearform.

Bemerkung VII.3.13. Charles Hermite (1822-1901).

Das semi von halb kommt ist klar. Sesqui kann man sich als anderthalb merken: (einfach) linear in der ersten Komponente und halblinear in der zweiten, also anderthalb.

Wir fassen einige Eigenschaften von Sesquilinearformen zusammen:

Satz VII.3.14. *Es sei V ein $\mathbb{C}$-Vektorraum und $\beta\colon V \times V \to \mathbb{C}$ eine Sesquilinearform.*

(a) *Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V, $B = \beta(v_i, v_j) =: M_{\mathcal{B}}(\beta)$, $v = \sum_{i=1}^{n} \lambda_i v_i$ und $w = \sum_{j=1}^{n} \mu_j v_j$, so ist*

$$\beta(v, w) = (\lambda_1, \ldots, \lambda_n) \cdot B \cdot \begin{pmatrix} \overline{\mu}_1 \\ \vdots \\ \overline{\mu}_n \end{pmatrix}.$$

(b) *Ist $\mathcal{B}'$ eine weitere geordnete Basis von V, so gilt für $B' = M_{\mathcal{B}'}(\beta)$:*

$$B' = (T_{\mathcal{B}}^{\mathcal{B}'})^t \cdot B \cdot \overline{T_{\mathcal{B}}^{\mathcal{B}'}}.$$

(c) *Die Sesquilinearform β ist genau dann hermitesch, wenn $\overline{B}^t = B$ gilt für B wie in (a).*

(d) *Ist $q(v) = \beta(v, v)$, so ist*

$$\beta(v, w) = \frac{1}{4}(q(v + w) - q(v - w) + iq(v + iw) - iq(v - iw)).$$

Hierbei müssen Sie nicht voraussetzen, dass β hermitesch ist.

Beweis. Sie beweisen (a) wie in Bemerkung VII.2.5, (b) wie in Satz VII.2.8, (c) wie in Bemerkung VII.3.3 und (d) wie in Satz VII.3.7. $\square$

VII.4. Skalarprodukte und Orthonormalisierung

Unser Ziel ist es, aus einer gegebenen Familie $(v_1, \ldots, v_n)$ von Vektoren eine neue Familie $(w_1, \ldots, w_n)$ zu konstruieren, so dass alle w_i Norm 1 haben und so dass w_i senkrecht steht auf w_j für $i \neq j$. Im Folgenden sei der Grundkörper $K = \mathbb{R}$ oder $\mathbb{C}$.

Definition VII.4.1. Es sei V ein K-Vektorraum.

(a) Eine symmetrische (beziehungsweise hermitesche) Bilinearform β heißt *positiv definit*, falls für alle $0 \neq v \in V$ gilt, dass $\beta(v, v) > 0$ ist.
(b) Eine positiv-definite symmetrische (beziehungsweise hermitesche) Bilinearform auf V heißt ein *Skalarprodukt auf* V.
(c) Ein reeller Vektorraum mit einem Skalarprodukt heißt ein *euklidischer Vektorraum*, und ein komplexer Vektorraum mit einem Skalarprodukt heißt ein *unitärer Vektorraum*.

Bemerkung VII.4.2. Die Eigenschaft positiv-definit zu sein impliziert insbesondere auch für hermitesche β, dass $\beta(v, v) \in \mathbb{R}$ ist, weil $\mathbb{C}$ kein angeordneter Körper ist. Diese Eigenschaft impliziert auch, dass β nicht-ausgeartet ist.

Beispiele VII.4.3.

- Das Standardskalarprodukt auf dem $\mathbb{R}^n$ oder $\mathbb{C}^n$ macht $\mathbb{R}^n$ zu einem euklidischen und $\mathbb{C}^n$ zu einem unitären Vektorraum.
- Es sei $V = C^0([0, 1], \mathbb{R})$ und wir betrachten die Bilinearform
$$\langle f, g\rangle := \int_0^1 f(x)g(x)dx.$$
Diese ist symmetrisch und $\langle f, f\rangle > 0$ für alle $f \neq 0$, also ist V mit dieser Bilinearform euklidisch.

Definition VII.4.4. Es sei V ein euklidischer oder ein unitärer Vektorraum mit Skalarprodukt $\langle -, -\rangle \colon V \times V \to K$. Dann heißt
$$|| - || \colon V \to \mathbb{R}, \quad ||v|| := \sqrt{\langle v, v\rangle}$$
die *Norm auf* V.

Lemma VII.4.5. *Für jeden euklidischen oder unitären Vektorraum gilt die Cauchy-Schwarzsche Ungleichung:*
$$|\langle v, w\rangle| \leqslant ||v||||w|| \text{ für alle } v, w \in V.$$
Es gilt genau dann $|\langle v, w\rangle| = ||v||||w||$, *wenn* v *und* w *linear abhängig sind.*

Beweis. Wir führen den Beweis für $K = \mathbb{C}$; der reelle Fall ist einfacher.

Da das Skalarprodukt positiv-definit ist, erhalten wir

$$0 \leqslant \langle \lambda v + \mu w, \lambda v + \mu w\rangle = \lambda\bar{\lambda}\langle v, v\rangle + \mu\bar{\mu}\langle w, w\rangle + \lambda\bar{\mu}\langle v, w\rangle + \mu\bar{\lambda}\langle w, v\rangle. \tag{VII.4.1}$$

Wir setzen $\lambda := \langle w, w\rangle$. Ist $w \neq 0$, so ist also $0 < \lambda \in \mathbb{R}$. Multiplizieren wir (VII.4.1) mit λ^{-1} durch, so ergibt das
$$0 \leqslant \langle v, v\rangle\langle w, w\rangle + \mu\bar{\mu} + \bar{\mu}\langle v, w\rangle + \mu\langle w, v\rangle.$$
Für $\mu = -\langle v, w\rangle$ erhalten wir
$$\begin{aligned} 0 &\leqslant \langle v, v\rangle\langle w, w\rangle + \langle v, w\rangle\overline{\langle v, w\rangle} - \overline{\langle v, w\rangle}\langle v, w\rangle - \langle v, w\rangle\langle w, v\rangle \\ &= ||v||^2||w||^2 - |\langle v, w\rangle|^2. \end{aligned}$$
Wurzelziehen gibt das Ergebnis.

Ist $w \neq 0$, so folgt aus $v = \lambda w$, dass $|\langle v, w\rangle| = |\langle \lambda w, w\rangle| = |\lambda|||w||^2$ und $||v||||w|| = |\lambda|||w||^2$.

Gilt umgekehrt $|\langle v, w\rangle| = ||v||||w||$, so ist mit $\lambda := \langle w, w\rangle$ und $\mu := -\langle v, w\rangle$
$$0 = \langle \lambda v + \mu w, \lambda v + \mu w\rangle.$$
Aber damit ist $\lambda v + \mu w = 0$. □

Satz VII.4.6. *(Eigenschaften der Norm) In einem euklidischen oder unitären Vektorraum* V *gilt:*

(a) $||v|| = 0 \Leftrightarrow v = 0$.
(b) $||\lambda v|| = |\lambda|||v||$.
(c) $||v + w|| \leqslant ||v|| + ||w||$ *(Dreiecksungleichung).*

Beweis. Eigenschaften (a) und (b) folgen direkt aus der Definition der Norm. Zu (c) rechnen wir nach:

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

Beachten Sie, dass

$$\langle v,w\rangle + \langle w,v\rangle = \langle v,w\rangle + \overline{\langle v,w\rangle} \in \mathbb{R}$$

ist für alle v, w, weil für alle $z \in \mathbb{C}$ gilt, dass $z + \bar{z} \in \mathbb{R}$. Für den vorletzten Schritt haben wir die Cauchy-Schwarzsche Ungleichung benutzt. Wurzelziehen gibt die Behauptung. □

Definition VII.4.7.

(a) Zwei Vektoren v und w eines euklidischen oder eines unitären Vektorraums heißen zueinander *orthogonal* ($v \perp w$), falls $\langle v, w\rangle = 0$.
(b) Zwei Untervektorräume U_1, U_2 von V heißen zueinander *orthogonal*, ($U_1 \perp U_2$), falls $\langle u_1, u_2\rangle = 0$ für alle $u_1 \in U_1$ und alle $u_2 \in U_2$.
(c) Für einen Untervektorraum $U \subset V$ heißt

$$U^\perp := \{v \in V, \langle u, v\rangle = 0 \text{ für alle } u \in U\}$$

das *orthogonale Komplement von U in V*.
(d) Eine Familie von Vektoren $(v_1, \ldots, v_n)$ mit $v_i \in V$ heißt *orthogonal*, falls $\langle v_i, v_j\rangle = 0$ für alle $i \neq j$.
(e) Eine Familie von Vektoren $(v_1, \ldots, v_n)$ mit $v_i \in V$ heißt *orthonormal*, falls sie orthogonal ist und zusätzlich gilt, dass $||v_i|| = 1$ für $1 \leqslant i \leqslant n$.
(f) Ist eine geordnete Basis $(v_1, \ldots, v_n)$ orthonormal, so heißt sie eine *Orthonormalbasis von V*. Wir kürzen das mit *ON-Basis* ab.

Beispiele VII.4.8.

- Die Standardbasis $(e_1, \ldots, e_n)$ ist eine ON-Basis des $\mathbb{R}^n$. Es gilt

$$\mathsf{Span}_{\mathbb{R}}(e_i)^\perp = \mathsf{Span}_{\mathbb{R}}(e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n).$$

- Welche $f \in C^0([-\pi, \pi], \mathbb{R})$ sind orthogonal zu $g(x) = \sin(x)$? Es gilt zum Beispiel

$$\int_{-\pi}^{\pi} \sin(x)\cos(x)dx = 0.$$

  Finden Sie andere Beispiele?

Lemma VII.4.9. *Es sei V wieder ein euklidischer oder unitärer Vektorraum.*

- *Ist $(v_1, \ldots, v_n)$ eine orthogonale Familie mit $v_i \neq 0$, so ist $(v_1, \ldots, v_n)$ linear unabhängig und $(\frac{v_1}{||v_1||}, \ldots, \frac{v_n}{||v_n||})$ ist orthonormal.*
- *Ist $(v_1, \ldots, v_n)$ eine ON-Basis von V, so gilt für jedes $v \in V$*

(VII.4.2) $$v = \langle v, v_1\rangle v_1 + \ldots + \langle v, v_n\rangle v_n.$$

Beweis. Für (a) nehmen wir an, dass $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$ ist. Wir wenden $\langle v_i, -\rangle$ auf diese Gleichung an und erhalten, dass $\lambda_i\langle v_i, v_i\rangle = 0$ ist. Aber $\langle v_i, v_i\rangle \neq 0$, weil nach Voraussetzung $v_i \neq 0$. Also folgt, dass alle $\lambda_i = 0$ sind.

Für die normierte Familie $(\frac{v_1}{||v_1||}, \ldots, \frac{v_n}{||v_n||})$ gilt

$$\langle \frac{v_i}{||v_i||}, \frac{v_j}{||v_j||}\rangle = \frac{1}{||v_i||||v_j||}\langle v_i, v_j\rangle = \delta_{i,j}.$$

Für (b) setzen wir $\lambda_i := \langle v, v_i \rangle$ und rechnen nach, dass

$$\langle \lambda_1 v_1 + \ldots + \lambda_n v_n, v_i \rangle = \lambda_i$$

gilt für alle $1 \leqslant i \leqslant n$. Damit ist

$$\langle v - (\lambda_1 v_1 + \ldots + \lambda_n v_n), v_i \rangle = 0 \text{ für alle } i.$$

Aber da die v_i nach Vorausssetzung eine Basis von V bilden, folgt damit, dass

$$v - (\lambda_1 v_1 + \ldots + \lambda_n v_n) = v - (\langle v, v_1 \rangle v_1 + \ldots + \langle v, v_n \rangle v_n) = 0.$$

□

Jørgen Pedersen Gram, 1850–1916, Erhard Schmidt, 1876–1959.

Satz VII.4.10. *(Satz von Gram-Schmidt) Es sei V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und $U \subset V$ sei ein Untervektorraum mit einer ON-Basis $\mathcal{B} = (u_1, \ldots, u_m)$. Dann gibt es eine Ergänzung $(u_1, \ldots, u_m, v_{m+1}, \ldots, v_n)$, die eine ON-Basis von V ist. Insbesondere besitzt jeder solche Vektorraum V eine ON-Basis.*

Beweis. Ist $U = V$, so ist nichts zu zeigen. Wir nehmen also an, dass $U \subsetneq V$ und es sei $v \in V \setminus U$. Wir setzen

$$\tilde{v} := \langle v, u_1 \rangle u_1 + \ldots + \langle v, u_m \rangle u_m$$

und $w := v - \tilde{v}$. Es gilt

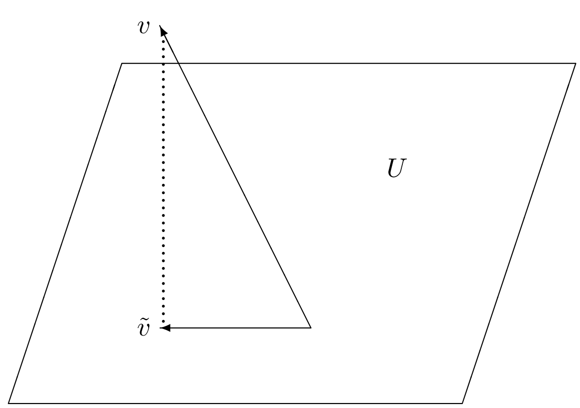
$$\begin{aligned} \langle w, u_i \rangle &= \langle v, u_i \rangle - \langle \tilde{v}, u_i \rangle \\ &= \langle v, u_i \rangle - \langle \langle v, u_i \rangle u_i, u_i \rangle \\ &= 0. \end{aligned}$$

Damit ist w orthogonal zu allen Vektoren aus U. Wir setzen

$$v_{m+1} := \frac{w}{||w||}$$

und betrachten $U' := \mathsf{Span}_K(u_1, \ldots, u_m, v_{m+1})$. Dann hat U' eine ON-Basis und hat eine Dimension mehr als U. Durch Iteration folgt die Behauptung, weil V endlich-dimensional ist. □

Bemerkung VII.4.11. Der Vektor $\tilde{v}$ ist die orthogonale Projektion von v auf den Untervektorraum U.



Bemerkung VII.4.12. Aus dem vorigen Beweis können wir das Gram-Schmidtsche Orthogonalisierungsverfahren herleiten.

Gegeben sei eine Familie $(w_1, \ldots, w_k)$ linear unabhängiger Vektoren eines euklidischen oder unitären Vektorraums.

- Wir setzen $v_1 := w_1$.

- Wir bilden

$$\begin{aligned}
v_2 &:= w_2 - \frac{\langle v_1, w_2\rangle}{\langle v_1, v_1\rangle} v_1 \\
v_3 &:= w_3 - \frac{\langle v_1, w_3\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle v_2, w_3\rangle}{\langle v_2, v_2\rangle} v_2 \\
&\vdots \\
v_k &= w_k - \sum_{i=1}^{k-1} \frac{\langle v_i, w_k\rangle}{\langle v_i, v_i\rangle} v_i.
\end{aligned}$$

- Dann sind die Vektoren $(v_1, \ldots, v_k)$ zueinander orthogonal. Durch Normalisierung erhalten Sie ein ON-System $(\frac{v_1}{||v_1||}, \ldots, \frac{v_k}{||v_k||})$
- Sie können auch zwischendurch Normalisieren und setzen

$$v_j' := w_j - \sum_{i=1}^{j-1} \langle v_i, w_j\rangle v_i, \quad v_j := \frac{v_j'}{||v_j'||}.$$

Definition VII.4.13. Ist V euklidisch oder unitär und ist $V = U_1 \oplus \ldots \oplus U_r$, so heißt die direkte Summenzerlegung *orthogonal*, falls $U_i \perp U_j$ für alle $i \neq j$.

Satz VII.4.14. *Ist U ein Untervektorraum eines euklidischen oder unitären K-Vektorraums V endlicher Dimension, so ist $V = U \oplus U^{\perp}$ und $\dim_K V = \dim_K U + \dim_K U^{\perp}$.*

Beweis. Es sei $(u_1, \ldots, u_m)$ eine ON-Basis von U, die wir zu einer ON-Basis $(u_1, \ldots, u_m, v_{m+1}, \ldots, v_n)$ von V ergänzen.

Wir behaupten, dass $U^{\perp} = \mathsf{Span}_K(v_{m+1}, \ldots, v_n)$ gilt:

Da die $v_{m+1}, \ldots, v_n$ orthogonal sind zu $u_1, \ldots, u_m$, folgt sofort, dass $\mathsf{Span}_K(v_{m+1}, \ldots, v_n) \subset U^{\perp}$ gilt. Ist umgekehrt $w = \sum_{i=1}^{m} \lambda_i u_i + \sum_{i=m+1}^{n} \mu_i v_i$ ein Element von $U^{\perp}$, so folgt, dass $\lambda_i = \langle w, u_i\rangle = 0$ gilt für alle $1 \leqslant i \leqslant m$ und somit ist $w \in \mathsf{Span}_K(v_{m+1}, \ldots, v_n)$. □

Ab jetzt schränken wir uns in diesem Abschnitt auf $K = \mathbb{R}$ an.

Definition VII.4.15. Es sei V ein euklidischer Vektorraum und es seien $v_1, \ldots, v_n \in V$. Dann heißt

$$G(v_1, \ldots, v_n) := \det \begin{pmatrix} \langle v_1, v_1\rangle & \ldots & \langle v_1, v_n\rangle \\ \vdots & & \vdots \\ \langle v_n, v_1\rangle & \ldots & \langle v_n, v_n\rangle \end{pmatrix}$$

die *Gramsche Determinante von* $(v_1, \ldots, v_n)$.

Satz VII.4.16. *Es ist genau dann $G(v_1, \ldots, v_n) > 0$, wenn $(v_1, \ldots, v_n)$ linear unabhängig ist.*

Bemerkung VII.4.17. Man kann zeigen, dass $G(v_1, \ldots, v_n)$ in jedem Fall nicht negativ ist (s. Fischer, Lineare Algebra, 15. Auflage, Vieweg Verlag, Seiten 208–210).

Beweis. Wir wissen, dass $U := \mathsf{Span}_{\mathbb{R}}(v_1, \ldots, v_n)$ ein endlich-dimensionaler euklidischer Vektorraum ist. Satz VII.4.10 impliziert, dass U eine ON-Basis $(w_1, \ldots, w_m)$ besitzt. Wir schreiben $v_i = a_{i1} w_1 + \ldots + a_{im} w_m$ und setzen $A := (a_{ij}) \in M(n \times m, \mathbb{R})$. Damit folgt

$$G(v_1, \ldots, v_n) = \det(A \cdot A^t).$$

Die Familie $(v_1, \ldots, v_n)$ ist genau dann linear unabhängig, wenn $n = m$ ist und A invertierbar ist und dies ist äquivalent zu $\det(A) \neq 0$, also $\det(A)^2 > 0$. □

Definition VII.4.18. Es sei V ein euklidischer endlich-dimensionaler $\mathbb{R}$-Vektorraum und $v_1, \ldots, v_n \in V$. Dann heißt

$$\sqrt{G(v_1, \ldots, v_n)} =: \mathsf{Vol}(v_1, \ldots, v_n)$$

das *Volumen des von* $(v_1, \ldots, v_n)$ *aufgespannten Spats*.

Beispiel VII.4.19. Betrachten wir $v_1, v_2, v_3 \in \mathbb{R}^3$ und setzen A an als die Matrix, welche die Vektoren v_1, v_2, v_3 als Spaltenvektoren hat, so ist $\mathsf{Vol}(v_1, v_2, v_3) = \sqrt{\det(A)^2} = |\det(A)|$ und das stimmt mit unserem vorherigen Volumenbegriff überein.

Bemerkung VII.4.20. Der Volumenbegriff ist von entscheidender Bedeutung in der Analysis und der Differentialgeometrie.

Geometrisch wichtig ist die folgende Tatsache:

Satz VII.4.21. *(Hadamardsche Ungleichung) Es sei $n < \infty$ und V sei ein n-dimensionaler euklidischer Vektorraum. Es seien $v_1, \dots, v_m \in V$ mit $m \leqslant n$. Dann gilt:*

(a) $\mathsf{Vol}(v_1, \dots, v_m) \leqslant ||v_1|| \cdot \ldots \cdot ||v_m||$.
(b) $\mathsf{Vol}(v_1, \dots, v_m) = ||v_1|| \cdot \ldots \cdot ||v_m|| \Leftrightarrow$ *die v_i sind paarweise orthogonal.*

Beweis. Wir zeigen

$$G(v_1, \dots, v_m) \leqslant ||v_1||^2 \cdot \ldots \cdot ||v_m||^2.$$

Genauer: Es sei $r \leqslant m$ beliebig und $U = \mathsf{Span}_{\mathbb{R}}(v_1, \dots, v_{r-1})$. Da V die orthogonale Summe $V = U \oplus U^\perp$ ist, gibt es für jedes v_r genau ein $u \in U$ und ein $\tilde{v}_r \in U^\perp$ mit $v_r = u + \tilde{v}_r$.

Wir zeigen als Zwischenbehauptung, dass gilt:

$$G(v_1, \dots, v_r) = G(v_1, \dots, v_{r-1}) \cdot \langle \tilde{v}_r, \tilde{v}_r \rangle.$$

Wir schreiben dafür einfach mal hin, wie $G(v_1, \dots, v_r)$ aussieht:

$$\begin{aligned}
G(v_1, \dots, v_r) &= G(v_1, \dots, v_{r-1}, u + \tilde{v}_r) \\
&= \det \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_{r-1} \rangle & \langle v_1, u + \tilde{v}_r \rangle \\ \vdots & & \vdots & \vdots \\ \langle v_{r-1}, v_1 \rangle & \dots & \langle v_{r-1}, v_{r-1} \rangle & \langle v_{r-1}, u + \tilde{v}_r \rangle \\ \langle u + \tilde{v}_r, v_1 \rangle & \dots & \langle u + \tilde{v}_r, v_{r-1} \rangle & \langle u + \tilde{v}_r, u + \tilde{v}_r \rangle \end{pmatrix} \\
&= \det \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_{r-1} \rangle & \langle v_1, u \rangle \\ \vdots & & \vdots & \vdots \\ \langle v_{r-1}, v_1 \rangle & \dots & \langle v_{r-1}, v_{r-1} \rangle & \langle v_{r-1}, u \rangle \\ \langle u, v_1 \rangle & \dots & \langle u, v_{r-1} \rangle & \langle u, u \rangle + \langle \tilde{v}_r, \tilde{v}_r \rangle \end{pmatrix}
\end{aligned}$$

- Wir schreiben u als $u = \lambda_1 v_1 + \ldots + \lambda_{r-1} v_{r-1}$.
- Ziehen Sie das λ_1-fache der ersten Spalte von der Spalte r ab.
- Ziehen Sie das λ_2-fache der zweiten Spalte von der Spalte r ab.
- Fahren Sie fort bis zur $(r-1)$sten Spalte und ziehen Sie das λ_{r-1}-fache der Spalte $r-1$ von der Spalte r ab.
- Damit ist die obige Determinante gleich

$$\det \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_{r-1} \rangle & 0 \\ \vdots & & \vdots & \vdots \\ \langle v_{r-1}, v_1 \rangle & \dots & \langle v_{r-1}, v_{r-1} \rangle & 0 \\ * & \dots & * & \langle \tilde{v}_r, \tilde{v}_r \rangle \end{pmatrix} = G(v_1, \dots, v_{r-1}) \cdot \langle \tilde{v}_r, \tilde{v}_r \rangle.$$

In der obigen Formel stehen die $*$-Symbole für irgendwelche Einträge, die uns nicht weiter interessieren.

Damit haben wir die Zwischenbehauptung bewiesen.

Da

$$\langle v_r, v_r \rangle = \langle u + \tilde{v}_r, u + \tilde{v}_r \rangle = \langle u, u \rangle + \langle \tilde{v}_r, \tilde{v}_r \rangle,$$

ist

$$G(v_1, \dots, v_{r-1}) \cdot \langle \tilde{v}_r, \tilde{v}_r \rangle \leqslant G(v_1, \dots, v_{r-1}) \cdot \langle v_r, v_r \rangle.$$

Gleichheit gilt genau dann, wenn $v_r \in U^\perp$ ist, weil genau dann $v_r = \tilde{v}_r$ gilt.

□

VII.5. Orthogonale und unitäre Endomorphismen

In diesem Abschnitt sei V wieder ein euklidischer oder unitärer Vektorraum. Wir arbeiten also wieder über $K = \mathbb{R}$ oder $\mathbb{C}$. Wir möchten Endomorphismen verstehen, die die Norm von Elementen nicht verändern. Spontan fallen einem Drehungen und Spiegelungen ein. Gibt es andere Beispiele?

Definition VII.5.1.

- Ein $f \in \mathsf{End}_K(V)$ heißt *orthogonal*, falls $K = \mathbb{R}$ und

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \quad \text{für alle } v, w \in V.$$

- Ein $f \in \mathsf{End}_K(V)$ heißt *unitär*, falls $K = \mathbb{C}$ und

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \quad \text{für alle } v, w \in V.$$

Lemma VII.5.2. *Es sei $f \in \mathsf{End}_K(V)$ orthogonal oder unitär. Dann gilt:*

(a) *$v \perp w \Leftrightarrow f(v) \perp f(w)$ für alle $v, w \in V$.*
(b) *$||f(v)|| = ||v||$ für alle $v \in V$.*
(c) *Ist $\dim_K V < \infty$, so ist f ein Isomorphismus und f^{-1} ist wiederum orthogonal oder unitär.*
(d) *Ist λ ein Eigenwert von f, so ist $|\lambda| = 1$.*

Als Umkehrung von (b) gilt: Ist $f \in \mathsf{End}_K(V)$ gegeben und gilt $||f(v)|| = ||v||$ für alle $v \in V$, so ist f orthogonal beziehungsweise unitär.

Beweis. Behauptung (a) folgt direkt aus der Definition, weil $\langle f(v), f(w)\rangle = \langle v, w\rangle$ und damit gilt insbesondere

$$\langle f(v), f(w)\rangle = 0 \Leftrightarrow \langle v, w\rangle = 0.$$

Ebenso folgt auch (b) direkt aus der Definition, weil

$$||f(v)||^2 = \langle f(v), f(v)\rangle = \langle v, v\rangle = ||v||^2$$

und mit dem Ziehen der Quadratwurzel erhalten wir (b).

Zu (c): Wir haben gerade gezeigt, dass $||f(v)|| = ||v||$ gilt für alle $v \in V$. Damit folgt aber aus $f(v) = 0$, dass $v = 0$ ist. Somit ist f injektiv. Ist $\dim_K V < \infty$, so folgt daraus schon die Bijektivität von f.

Wir rechnen mit $v = f(v')$ und $w = f(w')$ nach:

$$\begin{aligned}\langle f^{-1}(v), f^{-1}(w)\rangle &= \langle f^{-1}(f(v')), f^{-1}(f(w'))\rangle\\ &= \langle v', w'\rangle\\ &= \langle f(v'), f(w')\rangle\\ &= \langle v, w\rangle.\end{aligned}$$

Zu (d): Ist v ein Eigenvektor zum Eigenwert λ, so ist

$$\begin{aligned}||v|| &= ||f(v)||\\ &= ||\lambda v||\\ &= |\lambda|||v||.\end{aligned}$$

Da $v \neq 0$, muss damit $|\lambda| = 1$ sein.

Nehmen wir nun an, dass $f \in \mathsf{End}_K(V)$ die Norm erhält. Wir setzen $q(v) := \langle v, v\rangle$, so dass nach Voraussetzung $q(f(v)) = q(v)$. Wegen der Polarisierungsformel aus Satz VII.3.7 im euklidischen Fall und Satz VII.3.14 im sesquilinearen Fall läßt sich $\langle v, w\rangle$ durch q ausdrücken, so dass f auch das Skalarprodukt erhält. □

Beispiel VII.5.3. Es sei $R_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um den Winkel θ mit $M(R_\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$. Dann ist R_θ eine orthogonale Abbildung, weil Drehungen die Länge nicht verändern.

Beispiel VII.5.4. Es seien $\zeta_1, \ldots, \zeta_n \in \mathbb{C}$ mit $|\zeta_i| = 1$ für $1 \leqslant i \leqslant n$. Dann ist die Abbildung

$$f\colon \mathbb{C}^n \to \mathbb{C}^n, \quad \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \mapsto \begin{pmatrix} \zeta_1 z_1 \\ \vdots \\ \zeta_n z_n \end{pmatrix}$$

ein unitäre Abbildung, weil

$$\begin{aligned} \left\| \begin{pmatrix} \zeta_1 z_1 \\ \vdots \\ \zeta_n z_n \end{pmatrix} \right\| &= \sqrt{\sum_{i=1}^n (\zeta_i z_i)(\overline{\zeta_i z_i})} \\ &= \sqrt{\sum_{i=1}^n \zeta_i z_i \overline{\zeta_i}\,\overline{z_i}} \\ &= \sqrt{\sum_{i=1}^n \zeta_i \overline{\zeta_i} z_i \overline{z_i}} \end{aligned}$$

Da $|\zeta_i| = 1$, ist aber $\overline{\zeta_i} = \zeta_i^{-1}$, so dass wir oben

$$\sqrt{\sum_{i=1}^n z_i \overline{z_i}} = \left\| \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \right\|$$

erhalten.

Definition VII.5.5.

(a) Eine Matrix $A \in GL_n(\mathbb{R})$ heißt *orthogonal*, falls $A^t = A^{-1}$ gilt.
(b) Eine Matrix $A \in GL_n(\mathbb{C})$ heißt *unitär*, falls

$$\overline{A}^t = A^{-1} \text{ gilt.}$$

Bemerkung VII.5.6.

- Ist $A \in GL_n(\mathbb{R})$ dann ist die Multiplikation mit A als lineare Abbildung

  $$A\cdot\colon \mathbb{R}^n \to \mathbb{R}^n$$

  orthogonal, falls

  $$x^t \cdot y = \langle x, y \rangle = \langle Ax, Ay \rangle = (Ax)^t \cdot (Ay) = x^t \cdot A^t \cdot A \cdot y$$

  gilt für alle $x, y \in \mathbb{R}^n$. Das geht nur dann, wenn $A^t A = E_n$ ist, also wenn A orthogonal ist. Damit sind beide Definitionen von *orthogonal* konsistent.
- Ist $A \in GL_n(\mathbb{C})$, dann drücken wir das Standardskalarprodukt auf dem $\mathbb{C}^n$ aus als $\langle x, y \rangle = x^t \cdot \bar{y}$ und erhalten mit einer analogen Rechnung die Bedingung, dass

  $$x^t \cdot \bar{y} = \langle x, y \rangle = \langle Ax, Ay \rangle = (Ax)^t \cdot \overline{(Ay)} = x^t \cdot A^t \cdot \bar{A} \cdot \bar{y}$$

  gilt für alle $x, y \in \mathbb{C}^n$, also dass $A^t \bar{A} = E_n$. Das ist äquivalent zu $A^{-1} = \bar{A}^t$.
- Ist A orthogonal, so gilt für die Determinante:

  $$\det(A)^2 = \det(A)\det(A^t) = \det(A)\det(A^{-1}) = 1$$

  und damit folgt, dass $\det(A) = \pm 1$ ist. Es kommen also nur zwei mögliche Werte für die Determinante in Frage!
- Ist A unitär, so gilt

  $$1 = \det(A)\det(\bar{A}^t) = \det(A)\det(\bar{A}) = \det(A)\overline{\det(A)} = |\det(A)|^2$$

  und damit ist $|\det(A)| = 1$.

mögliche Determinanten für euklidische bzw. unitäre Matrizen

$= \{z \in \mathbb{C}, |z| = 1\}$

Definition VII.5.7.

(a)
$$O(n) := \{A \in GL_n(\mathbb{R}), A^{-1} = A^t\}$$
heißt die *orthogonale Gruppe* (der $n \times n$-Matrizen).

(b)
$$SO(n) := \{A \in GL_n(\mathbb{R}), A^{-1} = A^t, \det(A) = 1\}$$
heißt die *spezielle orthogonale Gruppe* (der $n \times n$-Matrizen).

(c)
$$U(n) := \{A \in GL_n(\mathbb{C}), A^{-1} = \bar{A}^t\}$$
heißt die *unitäre Gruppe* (der $n \times n$-Matrizen).

(d)
$$SU(n) := \{A \in GL_n(\mathbb{C}), A^{-1} = \bar{A}^t, \det(A) = 1\}$$
heißt die *spezielle unitäre Gruppe* (der $n \times n$-Matrizen).

Bemerkung VII.5.8. Ist $A \in SO(n)$, so ist A insbesondere orientierungserhaltend.

Dies sind alles wirklich Gruppen, genauer Untergruppen der jeweiligen GL_n:
$$SO(n) \subset O(n) \subset GL_n(\mathbb{R}),$$
$$SU(n) \subset U(n) \subset GL_n(\mathbb{C}).$$
Wir begründen dies für $O(n)$. Die anderen Fälle gehen analog.

Sind $A, B \in O(n)$, gilt also $A^{-1} = A^t$ und $B^{-1} = B^t$, so ist auch
$$(AB)(AB)^t = A \cdot B \cdot B^t \cdot A^t = E_n.$$
Die Einheitsmatrix E_n ist in $O(n)$ und mit $A \in O(n)$ ist auch $A^{-1} = A^t \in O(n)$, weil $(A^t)^t = A = (A^{-1})^{-1}$.

Es sei wieder $K = \mathbb{R}$ oder $\mathbb{C}$.

Lemma VII.5.9. *Für ein $A \in M(n \times n, \mathbb{K})$ sind äquivalent:*

(a) *$A \in O(n)$ (für $K = \mathbb{R}$) beziehungsweise $A \in U(n)$ (für $K = \mathbb{C}$).*
(b) *Die Spalten von A sind eine ON-Basis des K^n.*
(c) *Die Zeilen von A sind eine ON-Basis des K^n.*

Beweis. Wir beweisen das Lemma für $K = \mathbb{C}$.

Eine Matrix A ist genau dann ein Element in $U(n)$, wenn $\bar{A}^t \cdot A = E_n$ gilt, und dies ist äquivalent zu (c).

Es gilt genau dann $\bar{A}^t \cdot A = E_n$, wenn $A^t \cdot \bar{A} = E_n$ und dies ist äquivalent zu (b). □

Satz VII.5.10. *Es sei V ein endlich-dimensionaler euklidischer oder unitärer K-Vektorraum und $\mathcal{B}$ sei eine ON-Basis von V. Dann gilt:*

Ein $f \in \mathsf{End}_K(V)$ ist genau dann orthogonal (bzw. unitär), wenn $M_{\mathcal{B}}^{\mathcal{B}}(f)$ orthogonal (bzw. unitär) ist.

Beweis. Für $v, w \in V$ sei $x \in K^n$ die Basisdarstellung von v und y die Basisdarstellung von w. (Ist $v = \sum_{i=1}^n \lambda_i v_i$ für $\mathcal{B} = (v_1, \ldots, v_n)$, so ist $x = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$.) Da $\mathcal{B}$ eine ON-Basis ist, gilt $\langle v, w \rangle = x^t \cdot \bar{y}$. Es sei $A := M_{\mathcal{B}}^{\mathcal{B}}(f) \in M(n \times n, K)$. Dann gilt (für $K = \mathbb{C}$):
$$\langle f(v), f(w) \rangle = \langle v, w \rangle \Leftrightarrow (Ax)^t \cdot \overline{(Ay)} = x^t \cdot \bar{y}.$$
Da v, w beliebig waren, ist dies äquivalent zu $A^t \cdot \bar{A} = E_n$. □

Für kleine n kann man die obigen Gruppen geometrisch beschreiben. Wir beginnen mit $SO(2)$:

Beispiel VII.5.11. Wir hatten
$$SO(2) = \{A \in M(2 \times 2, \mathbb{R}), \det(A) = 1, A^t = A^{-1}\}.$$
Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann erhalten wir die beiden Bedingungen

- $ad - bc = 1$,
- $$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Dies übersetzt sich in die folgenden Gleichungen:

(a) $ad - bc = 1$,
(b) $a^2 + c^2 = 1$,
(c) $b^2 + d^2 = 1$ und
(d) $ab + cd = 0$.

Wir schreiben die Spalten von A als Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}$, $\begin{pmatrix} b \\ d \end{pmatrix} \in \mathbb{R}^2$. Die Bedingungen (b) und (c) sind dann äquivalent zu
$$\left\| \begin{pmatrix} a \\ c \end{pmatrix} \right\|^2 = 1 = \left\| \begin{pmatrix} b \\ d \end{pmatrix} \right\|^2.$$
Diese Vektoren liegen also auf dem Einheitskreis. Somit gibt es ein $\theta \in [0, 2\pi)$ mit
$$\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

Gleichung (d) besagt, dass die Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ zueinander senkrecht stehen. Da die Länge übereinstimmt, bleiben nur die Möglichkeiten
$$\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} -c \\ a \end{pmatrix} \text{ oder } \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} c \\ -a \end{pmatrix}.$$
Die zweite Möglichkeit steht im Widerspruch zu Gleichung (a), weil dann
$$ad - bc = -a^2 - c^2 = -1$$
wäre. Die erste Möglichkeit ist konsistent:
$$ad - bc = a^2 + c^2 = \cos^2\theta + \sin^2\theta = 1.$$
Damit erhalten wir insgesamt
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
und dies ist eine Drehmatrix. Man kann genauer $SO(2)$ mit der Einheitskreislinie identifizieren:
$$SO(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \theta \in [0, 2\pi) \right\} \cong \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \theta \in [0, 2\pi) \right\},$$
weil der zweite Spaltenvektor durch den ersten eindeutig festgelegt ist.

Beispiel VII.5.12. Was passiert, wenn wir ein $A \in O(2)$ mit $\det(A) = -1$ haben? Dann ändert sich oben nur die erste Bedingung zu

(a') $ad - bc = -1$.

Damit muss dann
$$\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} -c \\ a \end{pmatrix}$$
sein und wir erhalten
$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

und als Abbildung schickt A ein $\begin{pmatrix} x \\ y \end{pmatrix}$ auf

$$\begin{pmatrix} \cos\theta x + \sin\theta y \\ \sin\theta x - \cos\theta y \end{pmatrix}.$$

Das ist also eine Spiegelung an der Ursprungsgeraden mit Richtungsvektor

$$\begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}.$$

Zusammen mit elementaren Rechnungen zu $O(1)$ und $U(1)$ erhalten wir insgesamt:

Satz VII.5.13.

$$O(1) = \{\pm 1\}$$

$$O(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \theta \in [0, 2\pi) \right\} \cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \theta \in [0, 2\pi) \right\}$$

$$SO(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \theta \in [0, 2\pi) \right\}$$

$$U(1) = \{z \in \mathbb{C}, z\bar{z} = 1\} = \{z \in \mathbb{C}, |z| = 1\}.$$

Bemerkung VII.5.14. Die Gruppen $SO(2)$ und $U(1)$ sind beide nicht nur als Mengen gleich zu $S^1 := \{z \in \mathbb{C}, |z| = 1\}$, sondern auch als *Liegruppen*. Was das ist, lernen Sie später.

Was passiert für höhere n? Die Normalformen für unitäre Matrizen sind einfach:

Satz VII.5.15. *Jeder unitäre Endomorphismus eines unitären endlich-dimensionalen Vektorraums besitzt eine ON-Basis aus Eigenvektoren.*

Korollar VII.5.16. *Ist $A \in U(n)$, so gibt es ein $S \in U(n)$ mit*

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_i \in \mathbb{C}$ und $|\lambda_i| = 1$ gilt für $1 \leqslant i \leqslant n$.

Beweis von Korollar VII.5.16. Mit Satz VII.5.15 gibt es eine ON-Basis aus Eigenvektoren $(v_1, \dots, v_n)$ für die unitäre lineare Abbildung, welche die Multiplikation mit A ist. Wir setzen S an als die Matrix mit den Spalten $(v_1, \dots, v_n)$. Dass die Eigenwerte von A Betrag 1 haben, wissen wir schon. □

Beweis von Satz VII.5.15. Wir machen Induktion über die komplexe Dimension von V.

Gilt $\dim_{\mathbb{C}} V = 1$, so ist $f(v) = \lambda v$ mit einem $\lambda \in \mathbb{C}$. Da $\lambda^{-1} = \bar{\lambda}^t = \bar{\lambda}$ gelten muss, gilt $|\lambda| = 1$.

Es sei nun $n > 1$ und $P_f(X) = (-1)^n(X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n)$ sei die Zerlegung des charakteristischen Polynoms in Linearfaktoren. (Das geht, weil wir über $\mathbb{C}$ arbeiten.)

Es sei $v_1 \in V$ ein Eigenvektor zu λ_1 mit $||v_1|| = 1$. Wir setzen

$$U := \mathsf{Span}_{\mathbb{C}}(v_1)^{\perp} = \{v \in V, \langle v, v_1 \rangle = 0\}.$$

Da $|\lambda_1| = 1$ gilt, ist λ_1 insbesondere nicht trivial. Wir rechnen

$$\begin{aligned}\lambda_1\langle v_1, f(u)\rangle &= \langle \lambda_1 v_1, f(u)\rangle \\ &= \langle f(v_1), f(u)\rangle \\ &= \langle v_1, u\rangle \\ &= 0,\end{aligned}$$

falls $u \in U$. Damit ist $f(U) \subset U$, also ist U ein f-invarianter Untervektorraum. Da f ein Automorphismus ist, folgt sogar $f(U) = U$. Die Einschränkung von f auf U, $f|_U$, ist ebenfalls unitär und $\dim_{\mathbb{C}} U = \dim_{\mathbb{C}} V - 1$. Nach Induktionsannahme hat U eine ON-Basis $(v_2, \ldots, v_n)$ aus Eigenvektoren und damit ist $(v_1, \ldots, v_n)$ eine ON-Basis von V, die aus Eigenvektoren von f besteht. □

Das analoge Resultat für orthogonale Endomorphismen und Matrizen ist etwas komplizierter:

Satz VII.5.17. *Es sei $f \in \mathsf{End}_{\mathbb{R}}(V)$ orthogonal, wobei V ein endlich-dimensionaler euklidischer Vektorraum ist. Dann gibt es eine ON-Basis $\mathcal{B}$ von V, so dass*

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} E_r & 0 & \ldots & \ldots & 0 \\ 0 & -E_s & \ddots & & \vdots \\ \vdots & \ddots & D_1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & D_k \end{pmatrix}.$$

Hierbei ist $D_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix} \in SO(2)$ mit $\theta_i \in (0, 2\pi) \setminus \{\pi\}$ und $s, r, k \geqslant 0$.

Bemerkung VII.5.18.

Sie wissen, dass wir nichts Besseres erwarten können, weil Drehmatrizen für $\theta \neq 0, \pi$ nicht diagonalisierbar sind.

Die Matrizen E_r sind in $O(r)$ und $-E_s \in O(s)$.

Ist S_θ eine Spiegelung im $\mathbb{R}^2$, so haben wir die Eigenwerte $+1$ und -1 und wir wissen, dass wir eine Basis aus Eigenvektoren finden können, so dass die Matrixdarstellung von der Form $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist. Der Satz besagt also, dass *jede* orthogonale Abbildung eines endlich-dimensionalen euklidischen Vektorraums aus diesen Bausteinen zusammengesetzt ist.

Für den Beweis von Satz VII.5.17 müssen wir etwas ausholen. Wir werden Satz VII.5.15 für den Beweis benutzen. Dazu müssen wir beschreiben, wie wir reelle und komplexe Vektorräume in Beziehung setzen können:

Definition VII.5.19. Es sei V ein $\mathbb{R}$-Vektorraum. Die *Komplexifizierung von V* hat als unterliegende Menge $V_{\mathbb{C}} := V \times V$. Für $z = x + iy$ mit $x, y \in \mathbb{R}$ definieren wir

$$z(v_1, v_2) := (xv_1 - yv_2, yv_1 + xv_2). \tag{VII.5.1}$$

Beispiel VII.5.20. Ist $V = \mathbb{R}$, so ist $V_{\mathbb{C}} = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ mit

$$(x + iy)(x_1, x_2) = (xx_1 - yx_2, yx_1 + xx_2)$$

und damit ist $\mathbb{R}_{\mathbb{C}}$ isomorph zu $\mathbb{C}$.

Bemerkung VII.5.21.

- Mit der skalaren Multiplikation aus (VII.5.1) ist $V_{\mathbb{C}}$ ein $\mathbb{C}$ Vektorraum. Rechnen Sie das bitte nach!
- Für die Dimension gilt:

$$\dim_{\mathbb{R}} V_{\mathbb{C}} = 2\dim_{\mathbb{R}} V, \quad \dim_{\mathbb{C}} V_{\mathbb{C}} = \dim_{\mathbb{R}} V.$$

- Ist V ein euklidischer Vektorraum mit Skalarprodukt $\langle -,- \rangle \colon V \times V \to \mathbb{R}$, so definieren wir
$$\langle (v_1, v_2), (w_1, w_2) \rangle_{\mathbb{C}} := (\langle v_1, w_1 \rangle + \langle v_2, w_2 \rangle) + i(\langle v_2, w_1 \rangle - \langle v_1, w_2 \rangle).$$
Damit ist $\langle -,- \rangle_{\mathbb{C}}$ ein Skalarprodukt auf $V_{\mathbb{C}}$ und $V_{\mathbb{C}}$ ist ein unitärer Vektorraum.
- Ist $f \colon V \to W$ eine $\mathbb{R}$-lineare Abbildung, so sei $f_{\mathbb{C}} \colon V_{\mathbb{C}} \to W_{\mathbb{C}}$ definiert als
$$f_{\mathbb{C}}(v_1, v_2) := (f(v_1), f(v_2)).$$
Dann ist $f_{\mathbb{C}}$ auf $V_{\mathbb{C}}$ eine $\mathbb{C}$-lineare Abbildung. Die Additivität ist klar. Es sei $z = x + iy \in \mathbb{C}$:
$$\begin{aligned} f_{\mathbb{C}}(z(v_1, v_2)) &= f_{\mathbb{C}}(xv_1 - yv_2, yv_1 + xv_2) \\ &= (xf(v_1) - yf(v_2), yf(v_2) + xf(v_2)) \text{ ,weil } f \ \mathbb{R}\text{-linear ist} \\ &= z(f(v_1), f(v_2)). \end{aligned}$$
- Es gilt: Ist f orthogonal auf V, so ist $f_{\mathbb{C}}$ unitär auf $V_{\mathbb{C}}$.

Lemma VII.5.22. *Ist f ein orthogonaler Endomorphismus eines endlich-dimensionalen euklidischen Vektorraums V mit $\dim_{\mathbb{R}} V \geqslant 1$, so gibt es einen f-invarianten Untervektorraum $U \subset V$ mit $1 \leqslant \dim_{\mathbb{R}} U \leqslant 2$.*

Beweis. Wir betrachten $f_{\mathbb{C}} \colon V_{\mathbb{C}} \to V_{\mathbb{C}}$. Da $f_{\mathbb{C}}$ unitär ist, so gibt es einen Eigenvektor $v = (v_1, v_2) \in V_{\mathbb{C}}$ mit $f_{\mathbb{C}}(v) = \lambda v$ und $\lambda \in \mathbb{C}$ mit $|\lambda| = 1$.

Insbesondere ist $\mathsf{Span}_{\mathbb{C}}(v)$ ein $f_{\mathbb{C}}$-invarianter Untervektorraum von $V_{\mathbb{C}}$ mit $\dim_{\mathbb{C}} \mathsf{Span}_{\mathbb{C}}(v) = 1$.

Wir betrachten $U := \mathsf{Span}_{\mathbb{R}}(v_1, v_2) \subset V$; U ist automatisch ein $\mathbb{R}$-Untervektorraum von V und $1 \leqslant \dim_{\mathbb{R}} U \leqslant 2$. □

Damit können wir endlich die Normalformen für orthogonale Endomorphismen herleiten:

Beweis von Satz VII.5.17. Wir machen wieder Induktion, diesmal über die Dimension $\dim_{\mathbb{R}} V = n < \infty$.

Ist $\dim_{\mathbb{R}} V = 1$ und ist f eine orthogonale Abbildung $f \colon V \to V$, so ist f zwangsläufig gleich $\pm \mathrm{id}_V$ und die Normalform ist bewiesen.

Es sei also $n > 1$. Nach Lemma VII.5.22 gibt es einen f-invarianten Untervektorraum $U \subset V$ mit $1 \leqslant \dim_{\mathbb{R}} V \leqslant 2$. Da f ein Automorphismus ist, folgt wiederum $f(U) = U$ und $f^{-1}(U) = U$.

Wir zerlegen V als $V = U \oplus U^{\perp}$. Es sei $u \in U$ und $v \in U^{\perp}$. Da f^{-1} orthogonal ist, gilt
$$\begin{aligned} \langle f(v), u \rangle &= \langle f^{-1}(f(v)), f^{-1}(u) \rangle \\ &= \langle v, f^{-1}(u) \rangle \\ &= 0, \end{aligned}$$
weil $v \in U^{\perp}$ und $f^{-1}(u) \in U$.

Wir betrachten $f|_U \colon U \to U$ und $f|_{U^{\perp}} \colon U^{\perp} \to U^{\perp}$. Beides sind orthogonale Abbildungen und beide haben eine Dimension, die echt kleiner ist als die Dimension von V. Nach Induktionsannahme gibt es für $U^{\perp}$ eine ON-Basis $\mathcal{B}'$, so dass $M^{\mathcal{B}'}_{\mathcal{B}'}(f|_{U^{\perp}})$ aussieht wie gewünscht.

Für $f|_U$ erhalten wir mit Satz VII.5.13, dass $f|_U$ eine Matrixdarstellung der Form
$$M^{\mathcal{B}''}_{\mathcal{B}''}(f|_U) = (\pm 1) \text{ oder } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ oder } \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
hat mit $\theta \in [0, 2\pi)$.

Es gibt daher eine Anordnung von $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$, so dass $M^{\mathcal{B}}_{\mathcal{B}}(f)$ die gewünschte Form hat. □

Bemerkung VII.5.23. Die Matrix zu einem orthogonalen f der Form

$$\begin{pmatrix} E_r & 0 & \dots & \dots & 0 \\ 0 & -E_s & \ddots & & \vdots \\ \vdots & \ddots & D_1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & D_k \end{pmatrix}$$

heißt die *Normalform von* f. Die Zahlen r, s und k sind eindeutig, ebenso die Drehwinkel θ_i, aber die Reihenfolge der Drehmatrizen ist nicht eindeutig. Man nennt die D_i auch oft *Drehkästchen*.

VII.6. Selbstadjungierte Endomorphismen

Neben orthogonalen und unitären Endomorphismen treten symmetrische und hermitesche Matrizen häufig auf. In diesem Abschnitt sei wiederum $K = \mathbb{R}$ oder $\mathbb{C}$ und V sei ein euklidischer oder unitärer Vektorraum endlicher Dimension.

Definition VII.6.1. Ein $f \in \mathsf{End}_K(V)$ heißt *selbstadjungiert*, falls für alle $v, w \in V$ gilt:

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

Satz VII.6.2. *Ist $f \in \mathsf{End}_K(V)$ und ist $\mathcal{B}$ eine ON-Basis von V, so ist f genau dann selbstadjungiert, wenn $M_{\mathcal{B}}^{\mathcal{B}}(f)$ symmetrisch (für $K = \mathbb{R}$) oder hermitesch (für $K = \mathbb{C}$) ist.*

Das heißt also, dass $M_{\mathcal{B}}^{\mathcal{B}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(f)^t$ ist für $K = \mathbb{R}$ oder $\overline{M_{\mathcal{B}}^{\mathcal{B}}(f)}^t = M_{\mathcal{B}}^{\mathcal{B}}(f)$ für $K = \mathbb{C}$.

Beweis. Es sei $\Phi_{\mathcal{B}} \colon K^n \to V$ die Vergleichsabbildung zwischen V und K^n und es sei $\Phi_{\mathcal{B}}(x) = v$, $\Phi_{\mathcal{B}}(y) = w$. Wir setzen $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$. Da $\mathcal{B}$ eine ON-Basis ist, ist mit $x = \sum_{i=1}^n x_i e_i$ und $y = \sum_{j=1}^n y_i e_i$

$$\langle v, w \rangle = \left\langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j \right\rangle = \begin{cases} \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle v_i, v_j \rangle, & K = \mathbb{R}, \\ \sum_{i=1}^n \sum_{j=1}^n x_i \bar{y}_j \langle v_i, v_j \rangle, & K = \mathbb{C} \end{cases} = \begin{cases} \sum_{i=1}^n x_i y_i, & K = \mathbb{R}, \\ \sum_{i=1}^n x_i \bar{y}_i, & K = \mathbb{C} \end{cases}$$

und daher

$$\begin{aligned} \langle v, w \rangle &= \langle \Phi_{\mathcal{B}}(x), \Phi_{\mathcal{B}}(y) \rangle \\ &= \begin{cases} x^t \cdot y, & K = \mathbb{R} \\ x^t \cdot \bar{y}, & K = \mathbb{C}. \end{cases} \end{aligned}$$

Ebenso ist

$$\begin{aligned} \langle f(v), w \rangle &= \begin{cases} (Ax)^t \cdot y, & K = \mathbb{R} \\ (Ax)^t \cdot \bar{y}, & K = \mathbb{C} \end{cases} \\ &= \begin{cases} x^t \cdot A^t \cdot y, & K = \mathbb{R} \\ x^t \cdot A^t \cdot \bar{y}, & K = \mathbb{C} \end{cases} \end{aligned}$$

und

$$\begin{aligned} \langle v, f(w) \rangle &= \begin{cases} x^t \cdot (Ay), & K = \mathbb{R} \\ x^t \cdot \overline{Ay}, & K = \mathbb{C} \end{cases} \\ &= \begin{cases} x^t \cdot A \cdot y, & K = \mathbb{R} \\ x^t \cdot \bar{A} \cdot \bar{y}, & K = \mathbb{C}. \end{cases} \end{aligned}$$

Beide Terme sind für alle x, y nur dann gleich, wenn $A^t = A$ für $K = \mathbb{R}$ oder $A^t = \bar{A}$ ist für $K = \mathbb{C}$ und letzteres ist äquivalent zu $A = \bar{A}^t$. □

Lemma VII.6.3. *Ist f selbstadjungiert, so gilt sowohl für $K = \mathbb{R}$ als auch für $K = \mathbb{C}$, dass alle Eigenwerte reell sind. Insbesondere haben hermitesche Matrizen ausschließlich reelle Eigenwerte.*

Beweis. Ist $0 \neq v \in V$ mit $f(v) = \lambda v$, so ist

$$\begin{aligned}\lambda\langle v, v\rangle &= \langle \lambda v, v\rangle \\ &= \langle f(v), v\rangle \\ &= \langle v, f(v)\rangle \\ &= \langle v, \lambda v\rangle \\ &= \bar{\lambda}\langle v, v\rangle.\end{aligned}$$

Da $v \neq 0$ ist $||v||^2 \neq 0$ und daher muss $\lambda = \bar{\lambda}$ sein. □

Die Normalformen selbstadjungierter Endomorphismen sind sehr einfach und elegant:

Satz VII.6.4. *Zu jedem selbstadjungierten Endomorphismus $f \in \mathsf{End}_K(V)$ eines endlich-dimensionalen K-Vektorraums gibt es eine ON-Basis von V, die aus Eigenvektoren von f besteht. Das heißt für jedes $A \in M(n \times n, K)$ mit*

$$\begin{cases} A = A^t, & K = \mathbb{R}, \\ A = \bar{A}^t, & K = \mathbb{C}, \end{cases}$$

gibt es ein

$$\begin{cases} S \in O(n), & K = \mathbb{R}, \\ S \in U(n), & K = \mathbb{C}, \end{cases}$$

so dass

$$\bar{S}^t \cdot A \cdot S = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

mit reellen λ_i.

Beweis. Wir beginnen mit dem Fall $K = \mathbb{C}$. Wir wissen, dass über $\mathbb{C}$ das charakteristische Polynom $P_f(X)$ in Linearfaktoren zerfällt, also

$$P_f(X) = (-1)^n \prod_{i=1}^{n} (X - \lambda_i)$$

mit $\lambda_i \in \mathbb{C}$, aber wir wissen schon mit Lemma VII.6.3, dass alle λ_i reell sind. Es sei $v_1 \neq 0$ ein Eigenvektor zu λ_1 mit $||v_1|| = 1$. Wir setzen

$$U := \{u \in V, \langle v_1, u\rangle = 0\}.$$

Der Untervektorraum U ist f-invariant, weil wegen der Selbstadjungiertheit von f für alle $u \in U$ gilt

$$\langle v_1, f(u)\rangle = \langle f(v_1), u\rangle = \langle \lambda_1 v_1, u\rangle = \lambda_1 \langle v_1, u\rangle = 0.$$

Wir machen Induktion über $\dim_{\mathbb{C}} V = n$ und nach Induktionsannahme gibt es eine ON-Basis aus Eigenvektoren von $f|_U$, $(v_2, \dots, v_n)$ und somit ist $\mathcal{B} = (v_1, \dots, v_n)$ eine ON-Basis von V mit den gewünschten Eigenschaften.

Wir betrachten nun den Fall $K = \mathbb{R}$ und es sei $\mathcal{B}$ eine beliebige ON-Basis von V. Nach Satz VII.6.2 ist dann $M_{\mathcal{B}}^{\mathcal{B}}(f)$ symmetrisch, also insbesondere auch hermitesch, weil

$$\overline{M_{\mathcal{B}}^{\mathcal{B}}(f)}^t = M_{\mathcal{B}}^{\mathcal{B}}(f)^t = M_{\mathcal{B}}^{\mathcal{B}}(f).$$

Über $\mathbb{C}$ zerfällt $P_f(X)$ in Linearfaktoren mit reellen Faktoren $(X - \lambda_i)$, aber damit zerfällt $P_f(X)$ auch über $\mathbb{R}$ in Linearfaktoren. Damit können wir den Beweis im reellen Fall genauso weiterführen wie im komplexen Fall. □

Beispiel VII.6.5. Wir betrachten $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \in M(2 \times 2, \mathbb{R})$. Diese Matrix hat

$$P_A(X) = (X-1)^2 - 9 = X^2 - 2X - 8 = (X+2)(X-4)$$

und damit hat A die Eigenwerte 4 und -2. Die Matrix A ist somit ähnlich zu $\begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$.

Da

$$\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} - \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix},$$

ist der Eigenraum zum Eigenwert 4 gleich

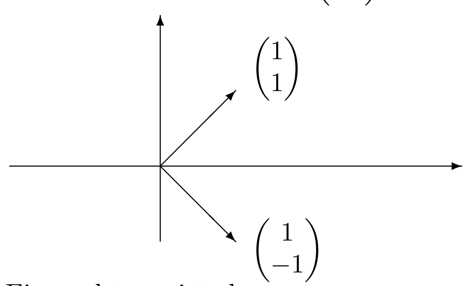
$$\mathrm{Eig}(A; 4) = \mathsf{Span}_{\mathbb{R}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Für den Eigenwert -2 erhalten wir mit

$$\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} - \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix},$$

dass

$$\mathrm{Eig}(A; -2) = \mathsf{Span}_{\mathbb{R}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$



Die zugehörige ON-Basis aus Eigenvektoren ist also

$$\left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right).$$

Korollar VII.6.6. *Ist f ein selbstadjungierter Endomorphismus eines endlich-dimensionalen Vektorraums und sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f, so ist*

$$V = \mathrm{Eig}(f; \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A; \lambda_r)$$

und diese direkte Summe ist orthogonal.

Beweis. Wir haben nach Satz VII.6.4 eine ON-Basis $\mathcal{B}$ aus Eigenvektoren. Durch Umnummerieren können wir $\mathcal{B}$ schreiben als $\mathcal{B} = \mathcal{B}_1 \sqcup \ldots \sqcup \mathcal{B}_r$, so dass $\mathcal{B}_i$ eine ON-Basis für $\mathrm{Eig}(f; \lambda_i)$ ist. $\square$

Zusammengefasst haben wir:

Bemerkung VII.6.7.

- Was wir mehrfach benutzt haben, ist, dass wir $O(n)$ als Untergruppe von $U(n)$ auffassen können: Ist eine Matrix $A \in O(n)$, so gilt $A^t = A^{-1}$. Aber da $A \in GL_n(\mathbb{R})$ ist, gilt automatisch $\bar{A} = A$, also auch
$$\bar{A}^t = A^t = A^{-1}.$$
- Ebenso können wir eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ auffassen als eine Matrix in $M(n \times n, \mathbb{C})$, weil $\mathbb{R} \subset \mathbb{C}$. Dann gilt $\bar{A}^t = A^t$, weil A nur reelle Einträge hat, und $A^t = A$, weil A symmetrisch ist. Damit ist A auch hermitesch.
- Sowohl Matrizen $A \in M(n \times n, \mathbb{C})$ mit $A = \bar{A}^t$ (hermitesch), als auch mit $A^{-1} = \bar{A}^t$ (unitär) sind diagonalisierbar mit ON-Basen aus Eigenvektoren.
- Matrizen $A \in M(n \times n, \mathbb{R})$ mit $A = A^t$ sind ebenfalls diagonalisierbar mit ON-Basen aus Eigenvektoren.

- Für $A \in M(n \times n, \mathbb{R})$ mit $A^{-1} = A^t$ (orthogonal) sind wir den Umweg über die Komplexifizierung gegangen.

Beispiel VII.6.8. Es sei $A = \begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$. Diese Matrix ist sichtbar symmetrisch. Das charakteristische Polynom ist

$$\begin{aligned} P_A(X) &= \det \begin{pmatrix} -1-X & 2 & -3 \\ 2 & 1-X & 0 \\ -3 & 0 & 1-X \end{pmatrix} \\ &= -X^3 + X^2 + 14X - 14 \\ &= (X-1)(X^2-14). \end{aligned}$$

Damit ist A ähnlich zu

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{14} & 0 \\ 0 & 0 & -\sqrt{14} \end{pmatrix}.$$

Eine ON-Basis aus Eigenvektoren zu finden, ist in diesem Beispiel nicht so kompliziert. Wir suchen beliebige Eigenvektoren zu 1, $\sqrt{14}$ und $-\sqrt{14}$ und diese sind nach Korollar VII.6.6 automatisch orthogonal, so dass ihre Normierung eine ON-Basis des $\mathbb{R}^3$ bildet.

Im allgemeinen Fall erhalten wir folgenden Algorithmus:

- Sie zerlegen $P_f(X)$ in Linearfaktoren über $\mathbb{C}$:
$$P_f(X) = (-1)^n (X - \lambda_1)^{m_1} \cdot \ldots \cdot (X - \lambda_r)^{m_r}.$$
- Sie bestimmen die Basen $\mathrm{Eig}(f; \lambda_j)$ für $1 \leqslant j \leqslant r$.
- Sie orthonormalisieren diese mit dem Gram-Schmidt Verfahren. Im komplexen Fall (für f unitär oder selbstadjungiert) war es das.
- Falls f über $\mathbb{R}$ definiert war und f war selbstadjungiert, so geht alles durch.
- Falls f über $\mathbb{R}$ definiert war und f war orthogonal, so müssen Sie aus den $\lambda_j \in \mathbb{C}$ mit $|\lambda_j| = 1$ noch die reellen Drehkästchen $D_j = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$ formen, falls $\lambda_j = \cos\theta_j + i \sin\theta_j$. (Da f über $\mathbb{R}$ definiert war, kommt sowohl λ_j als auch $\bar{\lambda}_j$ als Nullstelle vor. Sie nehmen für das Paar $(\lambda_j, \bar{\lambda}_j)$ ein solches Drehkästchen.)

VII.7. Hauptachsentransformation

Es sei $\beta\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform und $A = (\beta(e_i, e_j))_{ij} \in M(n \times n, \mathbb{R})$. Da $A^t = A$ gilt, ist A diagonalisierbar.

Satz VII.7.1. *Es seien β und A wie oben.*

(a) *Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine ON-Basis des $\mathbb{R}^n$ bezüglich des Standardskalarproduktes des $\mathbb{R}^n$ und sind die v_i Eigenvektoren von A, so ist*
$$M_{\mathcal{B},\mathcal{B}}(\beta) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix},$$
also $\beta(v_i, v_j) = \lambda_i \delta_{ij}$, falls $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A sind.

(b) *Es gibt eine Basis $\mathcal{B}'$ des $\mathbb{R}^n$, so dass*

$$M_{\mathcal{B}',\mathcal{B}'}(\beta) = \begin{pmatrix} E_k & 0 & 0 \\ 0 & -E_\ell & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

das heißt, dass es ein $S \in GL_n(\mathbb{R})$ gibt, so dass

$$S^t \cdot A \cdot S = \begin{pmatrix} E_k & 0 & 0 \\ 0 & -E_\ell & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Bemerkung VII.7.2. Der letzte Punkt im Satz entspricht unserem Kongruenzbegriff aus Definition VII.2.9.

Beweis. Die Existenz einer ON-Basis aus Eigenvektoren ist Satz VII.6.4, also gilt für eine solche ON-Basis $Av_j = \lambda_j v_j$ mit $\lambda_j \in \mathbb{R}$. Wir sortieren die λ_j nach Größe. Es sei also $\lambda_1, \dots, \lambda_k > 0$ und $\lambda_{k+1}, \dots, \lambda_{k+\ell} < 0$. Wir setzen

$$w_j := \begin{cases} |\lambda_j|^{-1/2} v_j, & \text{für } 1 \leqslant j \leqslant k+\ell, \\ v_j, & \text{für } j > k+\ell. \end{cases}$$

Die w_i sind orthogonal zueinander und

$$\beta(w_i, w_j) = \begin{cases} 1, & 1 \leqslant i = j \leqslant k, \\ -1, & k+1 \leqslant i = j \leqslant k+\ell, \\ 0, & \text{sonst.} \end{cases}$$

Wir setzen $\mathcal{B}$ an als die Basis $(v_1, \dots, v_n)$ und $\mathcal{B}'$ ist die Basis $(w_1, \dots, w_n)$.

Es sei S die Matrix mit $(v_1, \dots, v_n)$ als Spaltenvektoren:

$$S^t A S = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}, \quad M_{\mathcal{B}'\mathcal{B}'}(\beta) = \begin{pmatrix} E_k & 0 & 0 \\ 0 & -E_\ell & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

Was hat das Resultat mit Hauptachsen zu tun? Wir betrachten dazu ein Beispiel:

Beispiel VII.7.3. Es sei $A = \begin{pmatrix} u & v \\ v & u \end{pmatrix} \in M(2 \times 2, \mathbb{R})$. Die zugehörige Bilinearform ist

$$\beta(x, y) = x^t A y = u(x_1 y_1 + x_2 y_2) + v(x_1 y_2 + x_2 y_1).$$

Das charakteristische Polynom von A ist

$$\begin{aligned} P_A(X) &= \det \begin{pmatrix} u - X & v \\ v & u - X \end{pmatrix} \\ &= (u - X)^2 - v^2 \\ &= X^2 - 2uX + u^2 - v^2 \\ &= (X - (u+v))(X - (u-v)). \end{aligned}$$

Normierte Eigenvektoren sind

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ und}$$
$$v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Die quadratische Form q zu β ist

$$q(x) = \beta(x, x) = u x_1^2 + u x_2^2 + 2v x_1 x_2.$$

Wir betrachten die geordnete Basis $\mathcal{B} = (v_1, v_2)$ und z_1, z_2 seien die Koordinaten von $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Damit gilt

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} z_1 - z_2 \\ z_1 + z_2 \end{pmatrix}.$$

Wir werten q auf diesem Vektor aus unter erhalten:

$$\begin{aligned} q\left(\frac{1}{\sqrt{2}} \begin{pmatrix} z_1 - z_2 \\ z_1 + z_2 \end{pmatrix}\right) &= \frac{1}{2}((2z_1^2 + 2z_2^2)u + 2v(z_1^2 - z_2^2)) \\ &= z_1^2(u+v) + z_2^2(u-v) \\ &= \lambda_1 z_1^2 + \lambda_2 z_2^2. \end{aligned}$$

Wir betrachten nun die Niveaulinie $N_q(1) = \{z, q(z) = 1\}$ für den Fall, dass $\lambda_1 > 0$ und $\lambda_2 \neq 0$ ist und setzen

$$a := \frac{1}{\sqrt{\lambda_1}}, \quad b := \frac{1}{\sqrt{|\lambda_2|}}.$$

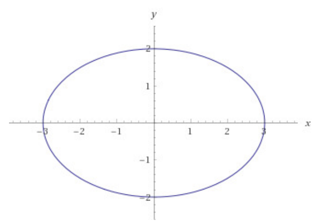
(a) Ist $\lambda_1 > 0$ und $\lambda_2 > 0$, so bekommen wir in $N_q(1)$ alle $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ mit

$$(z_1\sqrt{\lambda_1})^2 + (z_2\sqrt{\lambda_2})^2 = 1$$

also alle z mit

$$\frac{z_1^2}{a^2} + \frac{z_2^2}{b^2} = 1$$

und dies ist eine Ellipse mit Hauptachsen a, b.

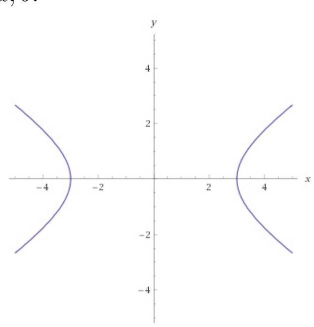


Ellipse mit Hauptachsen $a = 3$ und $b = 2$, Wolfram Alpha

(b) Ist $\lambda_1 > 0$ und $\lambda_2 < 0$, so ist $\lambda_2 = -|\lambda_2|$ und

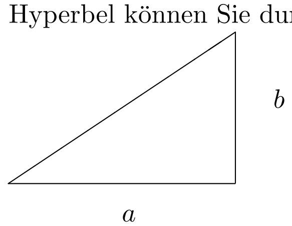
$$N_q(1) = \{z, \frac{z_1^2}{a^2} - \frac{z_2^2}{b^2} = 1\}.$$

Dies ist eine Hyperbel mit Achsen a, b.



Hyperbel mit Achsen $a = 3$ und $b = 2$, Wolfram Alpha

Die Hyperbel können Sie durch die Durchstoßpunkte bei $(\pm a, 0)$ und das Dreieck



beschreiben. Der obere Hyperbelast schmiegt sich an die Gerade mit Steigung $\frac{b}{a}$ an.

Definition VII.7.4. Es sei β eine Bilinearform auf einem beliebigen K-Vektorraum V. Dann heißt

$$V_0 = V_0^\beta := \{v \in V, \beta(v,w) = 0 \text{ für alle } w \in V\}$$

der *Ausartungsraum von* β.

Bemerkung VII.7.5. Da β bilinear ist, ist V_0^β ein Untervektorraum.
Ist die Bilinearform β nicht ausgeartet, so ist $V_0^\beta = \{0\}$.

Korollar VII.7.6. *Ist β eine symmetrische Bilinearform auf dem $\mathbb{R}^n$, so gilt:*

$$\mathbb{R}^n = V_+ \oplus V_0 \oplus V_-,$$

so dass gilt:

- *V_+ und V_- sind Untervektorräume des $\mathbb{R}^n$ und $V_0 = (\mathbb{R}^n)_0^\beta$.*
- *Die obige direkte Summe ist orthogonal, sowohl bezüglich β als auch bezüglich des Standardskalarproduktes auf dem $\mathbb{R}^n$.*
- *Für β gilt: $\beta(v,v) > 0$ für alle $0 \neq v \in V_+$ und $\beta(v,v) < 0$ für alle $0 \neq v \in V_-$.*

Beweis. Es sei A die darstellende Matrix von β und $\mathcal{B} = (v_1, \ldots, v_n)$ sei eine ON-Basis des $\mathbb{R}^n$. Wir setzen

$$\begin{aligned} V_+ &:= \mathsf{Span}_\mathbb{R}(v_1 \ldots, v_k), \\ V_- &= \mathsf{Span}_\mathbb{R}(v_{k+1}, \ldots, v_{k+\ell}), \\ V_0 &= \mathsf{Span}_\mathbb{R}(v_{k+\ell+1}, \ldots, v_n), \end{aligned}$$

wobei k und ℓ sind wie in Satz VII.7.1 (b).

Wir müssen nur noch zeigen, dass $V_0 = (\mathbb{R}^n)_0^\beta$ ist:

Ist $v \in V_0$, so ist $Av = 0$ und damit ist $\beta(v,w) = 0$ für alle $w \in \mathbb{R}^n$, so dass $V_0 \subset (\mathbb{R}^n)_0^\beta$ gilt.

Umgekehrt nehmen wir ein $v \in (\mathbb{R}^n)_0^\beta$ und stellen es bezüglich der Basis $\mathcal{B}$ dar, also

$$v = \sum_{i=1}^n \mu_i v_i.$$

Wäre $\mu_i \neq 0$ für ein $i \in \{1, \ldots, k+\ell\}$, so wäre

$$\beta(v, v_i) = \mu_i \neq 0.$$

Damit ist $v \in \mathsf{Span}_\mathbb{R}(v_{k+\ell+1}, \ldots, v_n)$. ☐

Satz VII.7.7. *(Trägheitssatz von Sylvester) Es sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und β sei eine symmetrische Bilinearform auf V. Es sei*

$$V = V_+ \oplus V_0 \oplus V_-,$$

wobei

- *$V_0 = V_0^\beta$,*
- *$\beta(v,v) > 0$ für alle $0 \neq v \in V_+$ und*
- *$\beta(v,v) < 0$ für alle $0 \neq v \in V_-$.*

Dann sind die Zahlen $r_+ := \dim_\mathbb{R} V_+$, $r_- := \dim_\mathbb{R} V_-$ und $r_0 := \dim_\mathbb{R} V_0$ Invarianten von β und

$$r_+ + r_- + r_0 = \dim_\mathbb{R} V.$$

Weiterhin gilt:

$$r_+ := \max\{\dim_\mathbb{R} U, U \subset V \text{ Untervektorraum}, \quad \beta(u,u) > 0 \quad \forall u \in U \setminus \{0\}\}$$

und

$$r_- := \max\{\dim_\mathbb{R} U, U \subset V \text{ Untervektorraum}, \quad \beta(u,u) < 0 \quad \forall u \in U \setminus \{0\}\}$$

Beweis. Wir zeigen zunächst, dass für einen Untervektorraum $U \subset V$ mit $\beta(u,u) > 0$ für alle $0 \neq u \in U$ gilt, dass

$$U \cap (V_- \oplus V_0) = \{0\}.$$

Dafür nehmen wir an, dass es ein $0 \neq u \in U$ der Form $u = u_- + u_0$ gibt mit $u_- \in V_-$ und $u_0 \in V_0$. Dann ist $\beta(u,u) > 0$, weil $u \in U$, aber auch

$$\beta(u,u) = \beta(u_- + u_0, u_- + u_0) = \beta(u_-, u_-) \leqslant 0.$$

Damit besteht der Schnitt nur aus der Null.

Da $U \oplus V_- \oplus V_0 \subset V$ gilt, ist

$$\dim_{\mathbb{R}} U + \dim_{\mathbb{R}} V_- + \dim_{\mathbb{R}} V_0 \leqslant \dim_{\mathbb{R}} V.$$

Wir nehmen an, dass $V = V'_+ \oplus V'_- \oplus V_0$ ist mit $\dim_{\mathbb{R}} V'_+ =: k$, $\ell := \dim_{\mathbb{R}} V'_-$ und so dass β positiv ist auf $V'_+ \setminus \{0\}$ und negativ auf $V'_- \setminus \{0\}$.

Wir setzen $U := V'_+$. Dann folgt $k \leqslant r_+$ und aus Symmetrie dann auch $r_+ \leqslant k$, also $r_+ = k$. Damit folgt auch

$$r_- = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} V_0 - \dim V_+.$$

□

Korollar VII.7.8. *Ist V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und ist β eine symmetrische Bilinearform mit Invarianten r_+, r_- und r_0, so gibt es eine ON-Basis $\mathcal{B}$ von V mit*

$$M_{\mathcal{B},\mathcal{B}}(\beta) = \begin{pmatrix} E_{r_+} & 0 & 0 \\ 0 & -E_{r_-} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Insbesondere gibt es eine orthogonale Summenzerlegung

$$V = V_+ \oplus V_- \oplus V_0.$$

Korollar VII.7.9. *Ist $A \in M(n \times n, \mathbb{R})$ eine symmetrische Matrix und ist $S \in GL_n(\mathbb{R})$, so haben A und $S^t \cdot A \cdot S$ die gleichen Anzahlen positiver und negativer Eigenwerte mit Vielfachheiten gezählt.*

Bemerkung VII.7.10. Daher kommt der Name *Trägheitssatz*: Diese Anzahlen verändern sich nicht, wenn man eine andere Basisdarstellung wählt.

Definition VII.7.11. In der Situation von Korollar VII.7.8 sagt man, dass β die *Signatur* (r_+, r_-, r_0) hat. Ist β nicht ausgeartet, das heißt ist $r_0 = 0$, so sagt man auch, dass β *Signatur* (r_+, r_-) hat.

Bemerkung VII.7.12. Allgemeiner kann man Endomorphismen betrachten, die allgemeine symmetrische nicht-ausgeartete Bilinearformen invariant lassen. Dies kommt in Anwendungen oft vor:

Ist $2 \neq 0$ in K und ist V ein endlich-dimensionaler K-Vektorraum mit einer symmetrischen, nicht-ausgearteten Bilinearform β, so betrachten wir wieder die zugehörige quadratische Form q und setzen

$$O(V;q) := \{f \in \mathsf{End}_K(V),\ q(f(v)) = q(v) \text{ für alle } v \in V\}.$$

Dies ist die *orthogonale Gruppe von V bezüglich q*.

Wichtige Spezialfälle sind der $\mathbb{R}^n$ mit $\beta = \sigma_p$, wobei

$$\sigma_p(x,y) := \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{n} x_i y_i.$$

Man schreibt oft kurz $O(p, n-p)$ für $O(\mathbb{R}^n, \sigma_p)$. Für $p = n$ erhalten Sie $O(n,0) = O(n)$.

Ab jetzt sei V ein K-Vektorraum und K sei hierbei ein beliebiger Körper. Wir setzen allerdings voraus, dass $0 \neq 2$ gilt in K.

Satz VII.7.13. *(Orthogonalisierungssatz) Ist V endlich-dimensional und ist β eine symmetrische Bilinearform, so gibt es eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V, so dass $\beta(v_i, v_j) = 0$ ist für alle $i \neq j$. Für $q(v) = \beta(v, v)$ und $\alpha_i := q(v_i)$ folgt dann*

$$q(v) = \alpha_1\lambda_1^2 + \dots + \alpha_n\lambda_n^2 \text{ für } v = \sum_{i=1}^{n} \lambda_i v_i. \tag{VII.7.1}$$

Bemerkung VII.7.14. Die Form von $q(v)$ in (VII.7.1) nennt man auch *rein quadratisch*, weil keine gemischten Terme $\lambda_i\lambda_j$ vorkommen für $i \neq j$.

Beweis. Wir machen Induktion über $\dim_K V = n$.
Für $n = 1$ tut es jedes $v_1 \neq 0$.
Es sei nun $n > 1$. Wir können die Polarisierungsformel aus Satz VII.3.7 benutzen, weil $2 \neq 0$ ist in K.
Ist $q(v) = 0$ für alle $v \in V$, so ist damit auch $\beta(v, w) = 0$ für alle $v, w \in V$ und es ist nichts zu zeigen.
Gibt es ein $v_1 \in V$ mit $q(v_1) \neq 0$, so setzen wir

$$U := \{v \in V, \beta(v_1, v) = 0\}.$$

Wir behaupten, dass $V = \mathsf{Span}_K(v_1) \oplus U$ ist: Ist $v \in \mathsf{Span}_K(v_1) \cap U$, so ist einerseits $v = \lambda_1 v_1$ für ein $\lambda_1 \in K$ aber auch

$$0 = \beta(v_1, v) = \beta(v_1, \lambda_1 v_1) = \lambda_1\beta(v_1, v_1) = \lambda_1 q(v_1).$$

Da nach Annahme $q(v_1) \neq 0$, muss also $\lambda_1 = 0$ sein, also ist $v = 0$. Der Schnitt der beiden Untervektorräume besteht also nur aus der Null.

Wir müssen noch zeigen, dass $V = U + \mathsf{Span}_K(v_1)$ gilt. Für $v \in V$ setzen wir

$$\tilde{v} := \frac{\beta(v_1, v)}{\beta(v_1, v_1)} v_1.$$

Damit ist $v = \tilde{v} + v - \tilde{v}$ und es gilt $\beta(v_1, \tilde{v}) = \beta(v_1, v)$. Dann gilt aber auch

$$\beta(v_1, v - \tilde{v}) = \beta(v_1, v) - \beta(v_1, \tilde{v}) = 0,$$

so dass $v - \tilde{v} \in U$ gilt. Damit ist $V = \mathsf{Span}_K(v_1) \oplus U$.

Wir betrachten $\beta|_{U\times U}$. Es ist $\dim_K U = n-1$ und daher gibt es nach Induktionsvoraussetzung eine Basis $(v_2, \dots, v_n)$, so dass $\beta|_{U\times U}(v_i, v_j) = 0$ ist für $i \neq j$. Somit ist $\mathcal{B} = (v_1, \dots, v_n)$ die gewünschte Basis. $\square$

Für die reellen Zahlen hatten wir das folgende Resultat schon in einer stärkeren Version in VII.6.4. Für allgemeine Körper gilt:

Korollar VII.7.15. *Ist $0 \neq 2$ in K und ist $A \in M(n\times n, K)$ gegeben mit $A^t = A$, so gibt es ein $S \in GL_n(K)$ mit*

$$S^t \cdot A \cdot S = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Definition VII.7.16. Eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ ist *positiv-definit*, falls die zugehörige symmetrische Bilinearform positiv ist, also $x^t \cdot A \cdot x > 0$ für alle $x \neq 0$.

Wann sind symmetrische Matrizen $A \in M(n \times n, \mathbb{R})$ positiv definit? Das zu entscheiden ist zum Beispiel wichtig in der Analysis bei der Untersuchung von Extremalwerten. Mit Satz VII.6.4 folgt, dass dies gilt, falls alle Eigenwerte positiv sind. Falls Sie die Eigenwerte nicht bestimmen möchten, hilft das folgende Kriterium.

Satz VII.7.17. *(Hauptminoren-Kriterium) Ist $A \in M(n \times n, \mathbb{R})$ und ist $A = (a_{ij})$ symmetrisch, so ist A genau dann positiv-definit, wenn $\det(A_k) > 0$ ist für alle $1 \leqslant k \leqslant n$, wobei*

$$A_k := \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}.$$

Beweis. Nehmen wir an, dass A positiv-definit ist. Es gibt zu A mit Satz VII.6.4 ein $S \in GL_n(\mathbb{R})$, so dass S^tAS eine Diagonalmatrix D mit Eigenwerten $\lambda_1, \dots, \lambda_n$ ist. Es gilt $\det(A) = \det(S)^{-2} \prod_{i=1}^n \lambda_i$, also ist $\det(A)$ genau dann positiv, wenn $\prod_{i=1}^n \lambda_i$ positiv ist.

Gibt es ein $\lambda_i \leqslant 0$, so gilt für einen Eigenvektor $0 \neq x$, dass $Dx = \lambda_i x$ ist und somit $x^tDx \leqslant 0$. Dann ist aber auch $(Sx)^tA(Sx) \leqslant 0$. Somit sind alle $\lambda_i > 0$.

Es sei β die zugehörige Bilinearform zu A. Die Matrix A_k ist die darstellende Matrix der Bilinearform $\beta_k := \beta_{\mathbb{R}^k \times \mathbb{R}^k}$, wobei wir hier $\mathbb{R}^k$ mit dem Vektorraum

$$\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}, x_i \in \mathbb{R} \right\}$$

identifizieren. Für $k \geqslant 1$ ist β_k positiv-definit, also ist $\det(A_k) > 0$.

Für die Rückrichtung nehmen wir an, dass $\det(A_k) > 0$ ist für alle $1 \leqslant k \leqslant n$. Wir zeigen die Behauptung über Induktion nach n.

Ist $n = 1$, so ist $A_1 = (a_{11})$ und hier ist die Behauptung klar.

Sei nun $n > 1$. Wir wissen dann, dass A_{n-1} positiv definit ist. Wir finden ein $T \in GL_{n-1}(\mathbb{R})$ mit

$$T^tA_{n-1}T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_{n-1} \end{pmatrix}$$

und wir wissen, dass alle λ_i positiv sind mit dem Argument von eben. Wir bilden die Blockmatrix

$$T' := \begin{pmatrix} T & 0 \\ 0 & 1 \end{pmatrix} \in GL_n\mathbb{R}$$

und bilden $(T')^tAT'$. Diese Matrix ist von der Form

$$(T')^tAT' = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & b_1 \\ 0 & \lambda_2 & \ddots & \vdots & b_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & b_{n-1} \\ b_1 & b_2 & \dots & b_{n-1} & b_n \end{pmatrix} =: B.$$

Nach Voraussetzung ist $\det A = \det(A_n) > 0$ und somit ist auch $\det B = \det(T')^2 \det A > 0$. Mit

$$S := \begin{pmatrix} 1 & 0 & \dots & 0 & -\frac{b_1}{\lambda_1} \\ 0 & 1 & \ddots & \vdots & -\frac{b_2}{\lambda_2} \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -\frac{b_{n-1}}{\lambda_{n-1}} \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

rechnen Sie nach, dass

$$S^tBS = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

gilt mit $\lambda_n = -(\frac{b_1^2}{\lambda_1} + \ldots + \frac{b_{n-1}^2}{\lambda_{n-1}}) + b_n$. Da $0 < \det B = (\det S)^{-2} \prod_{i=1}^n \lambda_i$, ist auch $\lambda_n > 0$. Also sind alle Eigenwerte von A positiv und A ist positiv-definit. $\square$

VII.8. Skalarprodukte und Dualität

Ist allgemein

$$\beta \colon V \times W \to K$$

eine Bilinearform, so können wir jeweils eine Komponente festhalten und erhalten Elemente im Dualraum

$$\beta_v \colon W \to K, \quad \beta_v(w) := \beta(v, w) \text{ für festes } v \in V,$$
$$\beta_w \colon V \to K, \quad \beta_w(v) := \beta(v, w) \text{ für festes } w \in W.$$

Wir haben also $\beta_v \in W^*$ und $\beta_w \in V^*$. Wir lassen nun wieder v und w variieren:

Definition VII.8.1. Es seien V und W beliebige K-Vektorraum und $\beta \colon V \times W \to K$ sei eine Bilinearform. Wir setzen

$$\begin{aligned} \beta_V \colon V &\to W^*, \\ v &\mapsto \beta_v(-), \\ \beta_W \colon W &\to V^*, \\ w &\mapsto \beta_w(-). \end{aligned}$$

Bemerkung VII.8.2. Die Bilinearform β ist genau dann nicht-ausgeartet, wenn β_V und β_W injektiv sind. Beide Abbildungen β_V und β_W sind linear.

Satz VII.8.3. *Sind V und W endlich-dimensionale K-Vektorräume und ist β nicht-ausgeartet, dann sind $\beta_V \colon V \to W^*$ und $\beta_W \colon W \to V^*$ Isomorphismen.*

Beweis. Nach Voraussetzung sind β_V und β_W injektiv, also gilt $\ker(\beta_V) = \{0\}$ und $\ker(\beta_W) = \{0\}$. Mit der Dimensionsformel erhalten wir:

$$\dim_K V = \dim_K \operatorname{Bild}(\beta_V) \leqslant \dim_K W^* \text{ und } \dim_K W = \dim_K \operatorname{Bild}(\beta_W) \leqslant \dim_K V^*.$$

Da aber $\dim_K V^* = \dim_K V$ und $\dim_K W^* = \dim_K W$ sind, folgt $\dim_K V = \dim_K W$ und beide Abbildungen sind auch surjektiv. $\square$

Korollar VII.8.4. *Ist V ein endlich-dimensionaler euklidischer Vektorraum, so ist die Abbildung*

$$\Psi \colon V \to V^*, \quad v \mapsto \langle v, - \rangle$$

ein Isomorphismus.

Beweis. $\beta = \langle -, - \rangle$ ist nicht-ausgeartet und $\langle v, - \rangle = \beta_v$. Somit ist $\Psi = \beta_V$. $\square$

Bemerkung VII.8.5. Dieser Isomorphismus ist basisunabhängig. Somit ist für jeden endlich-dimensionalen euklidischen Vektorraum der Dualraum *kanonisch* isomorph zum ursprünglichen Vektorraum!

Beispiele VII.8.6.

- Ist $V = \mathbb{R}^n$, so bildet $\Psi \colon \mathbb{R}^n \to (\mathbb{R}^n)^*$ den Vektor e_i auf $\langle e_i, - \rangle$ ab. Aber $\langle e_i, - \rangle$ ist genau das Funktional e_i^*.
- Ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ und setzen wir $\mathsf{Span}_{\mathbb{R}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} =: U$, so ist

  $$U^\perp = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, x_1 + x_2 = 0 \right\} = \mathsf{Span}_{\mathbb{R}} \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

  und

  $$\Psi(U^\perp) = \{\varphi \in (\mathbb{R}^2)^*, \varphi = \langle v, - \rangle, v \in U^\perp\}.$$

  Für $0 \neq v \in U^\perp$ gilt aber, dass $\langle v, w \rangle$ genau dann trivial ist, wenn $w \in U$ ist. Also ist $\Psi(U^\perp) = U^0$. (Hierbei ist U^0 der Annulator von U nach Definition VII.1.16.)

Diese Beispiele waren kein Zufall:

Satz VII.8.7. *Ist V ein endlich-dimensionaler euklidischer Vektorraum, so gilt:*
(a) $\Psi(U^\perp) = U^0$ *für alle Untervektorräume* $U \subset V$.
(b) *Ist* $\mathcal{B} = (v_1, \ldots, v_n)$ *eine ON-Basis von* V, *so ist* $\Psi(v_i) = v_i^*$.

Beweis. Für (a) benutzen wir, dass für $v \in U^\perp$ und $\Psi(v) = \langle v, - \rangle$ gilt, dass $\Psi(v)(u) = 0$ ist für alle $u \in U$. Somit ist $\Psi(U^\perp) \subset U^0$. Da

$$\dim_{\mathbb{R}} U^\perp = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} U = \dim_{\mathbb{R}} U^0$$

ist, folgt die Gleichheit.

Für (b) rechnen wir nach:

$$\Psi(v_i)(v_j) = \langle v_i, v_j \rangle = \delta_{ij}$$

und damit ist $\langle v_i, - \rangle = v_i^*$. □

Definition VII.8.8. Sind V und W endlich-dimensionale euklidische Vektorräume und ist $f \colon V \to W$ eine $\mathbb{R}$-lineare Abbildung, so ist die *zu f adjungierte Abbildung*, $f^{ad} \colon W \to V$, definiert als

$$f^{ad} := \Psi_V^{-1} \circ f^* \circ \Psi_W.$$

$$\begin{tikzcd} W \arrow[r, dotted, "f^{ad}"] \arrow[d, "\Psi_W"'] & V \arrow[d, "\Psi_V"] \\ W^* \arrow[r, "f^*"] & V^* \end{tikzcd}$$

Lemma VII.8.9. *Für alle $v \in V$, $w \in W$ gilt:*

$$\langle v, f^{ad}(w) \rangle = \langle f(v), w \rangle.$$

Beweis. Wir setzen die Definitionen ein und erhalten

$$\begin{aligned} f^*(\Psi_W(w)) &= f^* \langle w, - \rangle, \text{ also} \\ f^*(\Psi_W(w))(v) &= \langle w, - \rangle (f(v)) \\ &= \langle w, f(v) \rangle \\ &= \langle f(v), w \rangle. \end{aligned}$$

Ebenso gilt $\Psi_V \circ f^{ad}(w)(v) = \langle v, f^{ad}(w) \rangle$. Da nach Definition $\Psi_V \circ f^{ad} = f^* \circ \Psi_W$ gilt, folgt die Behauptung. □

Korollar VII.8.10. *Es seien V und W endlich-dimensionale euklidische Vektorräume und $f \colon V \to W$ sei eine $\mathbb{R}$-lineare Abbildung. Ist $\mathcal{B}_1$ eine ON-Basis von V und $\mathcal{B}_2$ eine ON-Basis von W, so gilt*

$$M^{\mathcal{B}_2}_{\mathcal{B}_1}(f^{ad}) = (M^{\mathcal{B}_1}_{\mathcal{B}_2}(f))^t.$$

Beweis. Nach Definition ist

$$\begin{aligned} M^{\mathcal{B}_2}_{\mathcal{B}_1}(f^{ad}) &= M^{\mathcal{B}_2}_{\mathcal{B}_1}(\Psi_V^{-1} \circ f^* \circ \Psi_W) \\ &= M^{\mathcal{B}_1^*}_{\mathcal{B}_1}(\Psi_V^{-1}) M^{\mathcal{B}_2^*}_{\mathcal{B}_1^*}(f^*) M^{\mathcal{B}_2}_{\mathcal{B}_2^*}(\Psi_W). \end{aligned}$$

Wir wissen nach Satz VII.1.14, dass $M^{\mathcal{B}_2^*}_{\mathcal{B}_1^*}(f^*) = M^{\mathcal{B}_1}_{\mathcal{B}_2}(f)^t$. Da $\Psi_V^{-1}(v_i^*) = v_i$ und $\Psi_W(w_i) = w_i^*$, gilt $M^{\mathcal{B}_1^*}_{\mathcal{B}_1}(\Psi_V^{-1}) = E_{\dim_{\mathbb{R}} V}$ und $M^{\mathcal{B}_2}_{\mathcal{B}_2^*}(\Psi_W) = E_{\dim_{\mathbb{R}} W}$. □

Bemerkung VII.8.11. Falls f eine selbstadjungierte Abbildung ist, so gilt $f = f^{ad}$, so dass die Notation Sinn macht.

Korollar VII.8.12. *Ist $f \colon V \to W$ eine $\mathbb{R}$-lineare Abbildung und sind V und W endlich-dimensional und euklidisch, so gilt*
(a) $\operatorname{Bild}(f^{ad}) = (\ker(f))^\perp$ *und* $\ker(f^{ad}) = (\operatorname{Bild}(f))^\perp$.

(b) *Ist $V = W$, so gilt*
$$V = \ker(f) \oplus \mathrm{Bild}(f^{ad}) = \ker(f^{ad}) \oplus \mathrm{Bild}(f)$$
und beides sind orthogonale Zerlegungen.

(c) *Im Spezialfall, dass $V = W$ und $f = f^{ad}$ ist, ist V also die orthogonale direkte Summe von* $\ker(f)$ *und* $\mathrm{Bild}(f)$.

Beweis. Die Behauptungen (b) und (c) folgen aus der ersten Behauptung.
Zu (a) rechnen wir stur nach, dass mit Satz VII.1.20 gilt:
$$\begin{aligned}\mathrm{Bild}(f^{ad}) &= \mathrm{Bild}(\Psi_V^{-1} \circ f^* \circ \Psi_W)\\ &= \mathrm{Bild}(\Psi_V^{-1} \circ f^*)\\ &= \Psi_V^{-1}(\mathrm{Bild}(f^*))\\ &= \Psi_V^{-1}(\ker(f)^0)\\ &= (\ker(f))^{\perp}.\end{aligned}$$
Die zweite Behauptung in (a) zeigen Sie auf analoge Weise. □

Bemerkung VII.8.13. Was passiert über $\mathbb{C}$? Es seien V und W also $\mathbb{C}$-Vektorräume mit einer Sesquilinearform $\beta\colon V \times W \to \mathbb{C}$. Wir können versuchen, $\Psi\colon V \to W^*$ zu bilden, indem wir $\Psi(v) = \beta(v, -)$ betrachten, aber $\beta(v, -)$ ist nur halblinear:
$$\Psi(v)(\lambda w) = \beta(v, \lambda w) = \bar{\lambda}\beta(v, w) = \bar{\lambda}\Psi(v)(w),$$
so dass $\Psi(v) \notin W^*$.

Wir können stattdessen die zweite Komponente benutzen und $\Phi\colon W \to V^*$ betrachten:
$$\Phi(w)(v) := \beta(v, w).$$
Diese Abbildung ist linear in v und halblinear in w.

Wir können trotzdem die adjungierte Abbildung zu einem $\mathbb{C}$-linearen $f\colon V \to W$ definieren, wenn V und W endlich-dimensionale unitäre Vektorräume sind:

Wir betrachten wieder $\Phi_V : V \to V^*$ und $\Phi_W : W \to W^*$ mit
$$\Phi_V(v_1)(v_2) = \langle v_2, v_1\rangle \text{ und } \Phi_W(w_1)(w_2) = \langle w_2, w_1\rangle.$$
Dann sind Φ_V und Φ_W halblineare, bijektive Abbildungen.

Definition VII.8.14. Es seien V und W endlich-dimensionale unitäre Vektorräume und $f\colon V \to W$ sei $\mathbb{C}$-linear. Dann ist die *zu f adjungierte Abbildung*, f^{ad}, definiert als
$$f^{ad} := \Phi_V^{-1} \circ f^* \circ \Phi_W : W \to V.$$

Lemma VII.8.15. *Die Abbildung f^{ad} ist $\mathbb{C}$-linear.*

Beweis. Es ist klar, dass f^{ad} die Addition respektiert. Es sei $\lambda \in \mathbb{C}$ und $w \in W$. Dann gilt
$$\begin{aligned}f^{ad}(\lambda w) &= (\Phi_V^{-1} \circ f^* \circ \Phi_W)(\lambda w)\\ &= \Phi_V^{-1}(f^*(\bar{\lambda}\Phi_W(w)))\\ &= \Phi_V^{-1}(\bar{\lambda} f^*(\Phi_W(w)))\\ &= \lambda\Phi_V^{-1}(f^*(\Phi_W(w)))\\ &= \lambda f^{ad}(w).\end{aligned}$$
□

Bemerkung VII.8.16. Lemma VII.8.9 überträgt sich. Korollar VII.8.10 müssen wir anpassen: Sind $\mathcal{B}_1$ und $\mathcal{B}_2$ ON-Basen von V beziehungsweise W, dann gilt für f^{ad}:

(VII.8.1) $$M^{\mathcal{B}_2}_{\mathcal{B}_1}(f^{ad}) = \overline{M^{\mathcal{B}_1}_{\mathcal{B}_2}(f)}^{\,t}\ :$$

Ist $f(v_j) = \sum_{i=1}^m a_{ij} w_i$ und ist $f^{ad}(w_i) = \sum_{j=1}^n b_{ji} v_j$, dann gilt

$$a_{ij} = \langle f(v_j), w_i \rangle = \langle v_j, f^{ad}(w_i) \rangle = \overline{b_{ji}}.$$

Definition VII.8.17. Es sei V ein unitärer, endlich-dimensionaler Vektorraum, so heißt $f \in \mathsf{End}_{\mathbb{C}}(V)$ *normal*, falls $f \circ f^{ad} = f^{ad} \circ f$ gilt.

$$\begin{tikzcd}
 & V \arrow[dl, "f^{ad}"'] \arrow[dr, "f"] & \\
V \arrow[dr, "f"'] & & V \arrow[dl, "f^{ad}"] \\
 & V &
\end{tikzcd}$$

Ein $A \in M(n \times n, \mathbb{C})$ heißt *normal*, falls $A\bar{A}^t = \bar{A}^t A$ gilt.

Bemerkung VII.8.18. Sie kennen schon Spezialfälle normaler Endomorphismen:

- Ist f selbstadjungiert, so ist $f^{ad} = f$ und daher gilt $f^{ad} \circ f = f \circ f = f \circ f^{ad}$.
- Ist f unitär, so ist f^{-1} ebenso unitär, und somit folgt

$$\langle v, f^{ad}(v') \rangle = \langle f(v), v' \rangle = \langle f^{-1}(f(v)), f^{-1}(v') \rangle = \langle v, f^{-1}(v') \rangle.$$

  Dann ist f aber auch normal, weil $f^{ad} = f^{-1}$ und

$$f \circ f^{ad} = f \circ f^{-1} = \mathrm{id}_V = f^{-1} \circ f = f^{ad} \circ f.$$

Damit umfaßt also der Begriff der Normalität sowohl unitäre als auch selbstadjungierte Endomorphismen. Für beide Klassen von Endomorphismen hatten wir Normalformen. Wir wollen diese auf normale Endomorphismen verallgemeinern.

Lemma VII.8.19. *Ist V ein unitärer, endlich-dimensionaler Vektorraum und ist $f \in \mathsf{End}_{\mathbb{C}}(V)$ normal, so gilt:*

(a) $\mathrm{Bild}(f) = \mathrm{Bild}(f^{ad})$,
(b) $\ker(f) = \ker(f^{ad})$

und daher ist $V = \mathrm{Bild}(f) \oplus \ker(f)$, wobei die direkte Summe orthogonal ist.

Beweis. Wir zeigen zunächst (b): Ein $v \in V$ ist genau dann im Kern von f, wenn $f(v) = 0$ und das ist äquivalent zu $||f(v)|| = 0$ und zu $||f(v)||^2 = 0$. Wir können dies umformen zu

$$\begin{aligned}
0 &= \langle f(v), f(v) \rangle \\
&= \langle v, f^{ad}(f(v)) \rangle \\
&= \langle v, f \circ f^{ad}(v) \rangle \\
&= \overline{\langle f \circ f^{ad}(v), v \rangle} \\
&= \overline{\langle f^{ad}(v), f^{ad}(v) \rangle}.
\end{aligned}$$

Dieser Ausdruck ist genau dann trivial, wenn $\langle f^{ad}(v), f^{ad}(v) \rangle = 0$ und dies ist äquivalent zu $f^{ad}(v) = 0$.

Zu (a): Mit Korollar VII.8.12 im komplexen Fall gilt

$$\mathrm{Bild}(f^{ad}) = (\ker(f))^{\perp} \text{ und } \ker(f^{ad}) = (\mathrm{Bild}(f))^{\perp}.$$

Mit (b) folgt dann

$$\begin{aligned}
\mathrm{Bild}(f^{ad}) &= (\ker(f))^{\perp} \\
&= (\ker(f^{ad}))^{\perp} \\
&= ((\mathrm{Bild}(f)^{\perp})^{\perp} \\
&= \mathrm{Bild}(f).
\end{aligned}$$

□

Korollar VII.8.20. *Ist V ein endlich-dimensionaler unitärer Vektorraum und ist $f \in \mathsf{End}_{\mathbb{C}}(V)$ normal, so ist*

$$\mathrm{Eig}(f;\lambda) = \mathrm{Eig}(f^{ad}, \overline{\lambda})$$

für alle $\lambda \in \mathbb{C}$.

Beweis. Wir setzen $g := f - \lambda \mathrm{id}_V$. Das Adjungierte von g ist $g^{ad} = f^{ad} - \bar{\lambda}\mathrm{id}_V$, weil

$$\langle v, g^{ad}(v')\rangle = \langle g(v), v'\rangle = \langle f(v) - \lambda v, v'\rangle = \langle f(v), v'\rangle - \langle \lambda v, v'\rangle = \langle v, (f^{ad} - \bar{\lambda}\mathrm{id}_V)(v')\rangle.$$

Damit gilt auch, dass g normal ist, also $g^{ad} \circ g = g \circ g^{ad}$ und

$$\mathrm{Eig}(f, \lambda) = \ker(g) = \ker(g^{ad}) = \mathrm{Eig}(f^{ad}, \bar{\lambda}).$$

□

Mit diesen Resultaten können wir normale Endomorphismen vollständig charakterisieren.

Satz VII.8.21. *(Hauptsatz über normale Endomorphismen) Ist V ein endlich-dimensionaler unitärer Vektorraum und ist $f \in \mathsf{End}_{\mathbb{C}}(V)$, so sind äquivalent:*

(a) *Es gibt eine ON-Basis von V, die aus Eigenvektoren von f besteht.*
(b) *f ist normal.*

Beweis. Wir nehmen an, dass f normal ist. Da wir über $\mathbb{C}$ arbeiten, zerfällt das charakteristische Polynom von f in Linearfaktoren:

$$P_f(X) = (-1)^n (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n).$$

Ist f die Nullabbildung, so ist nichts zu zeigen.

Es sei also $0 \neq v_1 \in \mathrm{Eig}(v, \lambda_1)$ und $\lambda_1 \neq 0$. Dann ist $U := \mathsf{Span}_{\mathbb{C}}(v_1) \neq \{0\}$ und wir setzen

$$W := \{v \in V, \langle v, v_1\rangle = 0\}.$$

Dann ist V die orthogonale direkte Summe von U und W und $\dim_{\mathbb{C}} W \leqslant n - 1$. Zu zeigen ist, dass W f-invariant ist und dass $f|_W$ wieder normal ist. Dann bekommen wir per Induktion eine Basis aus Eigenvektoren.

Es sei $w \in W$, also $\langle w, v_1\rangle = 0$. Dann ist auch

$$\langle f(w), v_1\rangle = \langle w, f^{ad}(v_1)\rangle = \langle w, \overline{\lambda_1} v_1\rangle = \lambda_1 \langle w, v_1\rangle = 0$$

und somit ist $f(W) \subset W$.

Für jedes w in W gilt natürlich, dass $f^{ad} \circ f(w) = f \circ f^{ad}(w)$, weil f normal ist. Da $f(W) \subset W$, gilt auch $f^{ad}(W) \subset W$, weil:

$$\begin{aligned}
\langle f^{ad}(w), v_1\rangle &= \overline{\langle v_1, f^{ad}(w)\rangle} \\
&= \overline{\langle f(v_1), w\rangle} \\
&= \overline{\lambda_1 \langle v_1, w\rangle} \\
&= \overline{\lambda_1}\langle w, v_1\rangle \\
&= 0.
\end{aligned}$$

Damit ist $f|_W$ normal.

Nehmen wir nun umgekehrt an, dass f eine ON-Basis aus Eigenvektoren besitzt, also ein $\mathcal{B} = (v_1, \ldots, v_n)$ mit $f(v_i) = \lambda_i v_i$. Dann gilt

$$\begin{aligned}
\langle v_j, f^{ad}(v_i)\rangle &= \langle f(v_j), v_i\rangle \\
&= \lambda_j \langle v_j, v_i\rangle \\
&= \lambda_j \delta_{ij}.
\end{aligned}$$

Damit gilt $f^{ad}(v_i) = \bar{\lambda}_i v_i$ und auch

$$f(f^{ad}(v_i)) = f(\overline{\lambda_i} v_i) = \overline{\lambda_i} f(v_i) = \overline{\lambda_i}\lambda_i v_i$$

und
$$f^{ad}(f(v_i)) = f^{ad}(\lambda_i v_i) = \lambda_i f^{ad}(v_i) = \lambda_i \overline{\lambda_i} v_i.$$
Also stimmen $f^{ad} \circ f$ und $f \circ f^{ad}$ auf $\mathcal{B}$ überein; sie sind somit gleich und f ist normal. $\square$

VII.9. Tensorprodukte

Wir haben uns ausführlich mit Bilinearformen beschäftigt. Wir verallgemeinern den Begriff auf beliebige Vektorräume als Bildbereich:

Definition VII.9.1. Es seien U, V, W K-Vektorräume für einen beliebigen Körper K. Eine *K-bilineare Abbildung* $\alpha\colon U \times V \to W$ ist eine Abbildung, die in U und in V K-linear ist. Es gilt also

(VII.9.1) $$\alpha(\lambda u_1 + \mu u_2, v) = \lambda\alpha(u_1, v) + \mu\alpha(u_2, v),$$

(VII.9.2) $$\alpha(u, \lambda v_1 + \mu v_2) = \lambda\alpha(u, v_1) + \mu\alpha(u, v_2),$$

für alle $\lambda, \mu \in K$, $u_1, u_2, u \in U$ und $v, v_1, v_2 \in V$.

Bemerkung VII.9.2. Ist $\alpha\colon U \times V \to W$ K-bilinear und ist $f \in \mathsf{Hom}_K(W, Z)$ für einen K-Vektorraum Z, so ist die Verknüpfung $f \circ \alpha\colon U \times V \to Z$ eine K-bilineare Abbildung.

Definition VII.9.3. Es sei K wieder ein Körper und U und V seien K-Vektorräume. Ein *Tensorprodukt von U und V* ist ein Paar $(U \otimes V, \kappa)$, wobei

- $U \otimes V$ ein K-Vektorraum ist und
- $\kappa\colon U \times V \to U \otimes V$ eine K-bilineare Abbildung ist, so dass $(U \otimes V, \kappa)$ die folgende universelle Eigenschaft hat:

Für jede K-bilineare Abbildung $\alpha\colon U \times V \to W$ in einen K-Vektorraum W gibt es genau eine K-lineare Abbildung $\phi_\alpha\colon U \otimes V \to W$, so dass
$$\alpha = \phi_\alpha \circ \kappa.$$
$$\begin{tikzcd} U \times V \arrow[r, "\alpha"] \arrow[d, "\kappa"'] & W \\ U \otimes V \arrow[ur, dotted, "\exists!\phi_\alpha"'] & \end{tikzcd}$$

Bemerkung VII.9.4.

(a) Für den Wert $\kappa(u, v)$ schreiben wir $u \otimes v$.

(b) Bilineare Abbildungen α lassen sich mithilfe des Tensorprodukts also auf lineare Abbildungen zurückführen, also genau auf die Abbildungen, mit denen wir uns jetzt zwei Semester lang beschäftigt haben.

(c) Möchte man klarstellen, über welchem Grundkörper man arbeitet, so notiert man K mit und schreibt $U \otimes_K V$ statt $U \otimes V$.

Satz VII.9.5. *Ein Tensorprodukt $(U \otimes V, \kappa)$ ist bis auf eindeutige Isomorphie eindeutig.*

Beweis. Wir nehmen an, dass $(U \tilde{\otimes} V, \tilde{\kappa})$ ein weiteres Tensorprodukt von U und V ist. Wir betrachten das Diagramm
$$\begin{tikzcd} U \times V \arrow[r, "\tilde{\kappa}"] \arrow[d, "\kappa"'] & U \tilde{\otimes} V. \\ U \otimes V & \end{tikzcd}$$
Wegen der universellen Eigenschaft von $(U \otimes V, \kappa)$ gibt es genau ein $\phi_{\tilde{\kappa}}$ mit $\phi_{\tilde{\kappa}} \circ \kappa = \tilde{\kappa}$:
$$\begin{tikzcd} U \times V \arrow[r, "\tilde{\kappa}"] \arrow[d, "\kappa"'] & U \tilde{\otimes} V. \\ U \otimes V \arrow[ur, dotted, "\exists!\phi_{\tilde{\kappa}}"'] & \end{tikzcd}$$

Wenn wir das obige Diagramm spiegeln, dann gibt uns die universelle Eigenschaft von $(U\tilde{\otimes}V, \tilde{\kappa})$ ein eindeutiges ϕ_κ mit $\phi_\kappa \circ \tilde{\kappa} = \kappa$:

$$\begin{tikzcd} U \times V \arrow[r, "\kappa"] \arrow[d, "\tilde{\kappa}"'] & U \otimes V. \\ U\tilde{\otimes}V \arrow[ur, dotted, "\exists!\phi_\kappa"'] & \end{tikzcd}$$

Für die Verkettungen $\phi_{\tilde{\kappa}} \circ \phi_\kappa$ und $\phi_\kappa \circ \phi_{\tilde{\kappa}}$ gilt

$$\phi_{\tilde{\kappa}} \circ \phi_\kappa \circ \tilde{\kappa} = \phi_{\tilde{\kappa}} \circ \kappa = \tilde{\kappa} = \mathrm{id}_{U\tilde{\otimes}V} \circ \tilde{\kappa} \text{ und}$$
$$\phi_\kappa \circ \phi_{\tilde{\kappa}} \circ \kappa = \phi_\kappa \circ \tilde{\kappa} = \kappa = \mathrm{id}_{U\otimes V} \circ \kappa.$$

Damit beschreiben $\phi_\kappa \circ \phi_{\tilde{\kappa}} \circ \kappa$ und $\mathrm{id}_{U\otimes V} \circ \kappa$ die gleiche bilineare Abbildung $U \times V \to U \otimes V$. $\phi_{\tilde{\kappa}} \circ \phi_\kappa \circ \tilde{\kappa}$ und $\mathrm{id}_{U\tilde{\otimes}V} \circ \tilde{\kappa}$ beschreiben die gleiche bilineare Abbildung $U \times V \to U\tilde{\otimes}V$. Wegen der Eindeutigkeitsaussage bei der universellen Eigenschaft muss damit gelten, dass

$$\phi_\kappa \circ \phi_{\tilde{\kappa}} = \mathrm{id}_{U\otimes V} \text{ und } \phi_{\tilde{\kappa}} \circ \phi_\kappa = \mathrm{id}_{U\tilde{\otimes}V}.$$

□

Wir zeigen nun die Existenz von Tensorprodukten.

Satz VII.9.6. *Es sei K ein Körper und U, V seien K-Vektorräume. Dann existiert ein Tensorprodukt $(U \otimes V, \kappa)$.*

Beweis. Wir betrachten $U \times V$ und K als Mengen und wir benutzen die Menge aller Abbildungen von $U \times V$ nach K, $\mathrm{Abb}(U \times V, K)$. Wir definieren eine Abbildung

$$\delta\colon U \times V \to \mathrm{Abb}(U \times V, K), \quad \delta(u_1, v_1)(u_2, v_2) := \begin{cases} 1, & u_1 = u_2 \text{ und } v_1 = v_2, \\ 0, & \text{sonst.} \end{cases}$$

Es sei

$$W := \mathsf{Span}_K\{\delta(u, v), u \in U, v \in V\}.$$

Der Vektorraum W enthält also endliche Linearkombinationen von Abbildungen $\delta(u, v)$.

Wir wollen δ dazu zwingen, bilinear zu werden und bilden den Quotientenvektorraum

$$T := W/X,$$

wobei X der Untervektorraum von W ist, der von allen Elementen in W der Form

(VII.9.3) $$\delta(\lambda u_1 + \mu u_2, v) - \lambda\delta(u_1, v) - \mu\delta(u_2, v) \text{ und } \delta(u, \lambda v_1 + \mu v_2) - \lambda\delta(u, v_1) - \mu\delta(u, v_2)$$

erzeugt wird, also $X = \mathsf{Span}_K(M)$, wobei M die Menge

$$\{\delta(\lambda u_1+\mu u_2, v)-\lambda\delta(u_1, v)-\mu\delta(u_2, v), \delta(u, \lambda v_1+\mu v_2)-\lambda\delta(u, v_1)-\mu\delta(u, v_2), u, u_1, u_2 \in U, \lambda, \mu \in K, v, v_1, v_2 \in V\}$$

ist. Es sei $\pi\colon W \to W/X = T$ die kanonische Projektion. Nach Konstruktion ist T ein Vektorraum und $\pi \circ \delta\colon U \times V \to T$ ist bilinear.

Wir behaupten, dass das Paar $(T, \pi \circ \delta)$ ein Tensorprodukt von U und V ist. Wir zeigen dies, indem wir nachweisen, dass $(T, \pi \circ \delta)$ die universelle Eigenschaft des Tensorproduktes hat.

Sei dazu also $\alpha\colon U \times V \to Y$ eine beliebige bilineare Abbildung in einen K-Vektorraum Y. Jedes Element in T läßt sich schreiben als $\pi\left(\sum_{i=1}^n \lambda_i \delta(u_i, v_i)\right)$. Wir setzen $\phi_\alpha\colon T \to Y$ an als

$$\phi_\alpha(\pi(\delta(u, v))) := \alpha(u, v)$$

und allgemein

$$\phi_\alpha\left(\pi\left(\sum_{i=1}^n \lambda_i \delta(u_i, v_i)\right)\right) := \sum_{i=1}^n \lambda_i \alpha(u_i, v_i).$$

Wir müssen zeigen, dass diese Abbildung wohldefiniert ist. Ist x ein Element aus X, so ist x eine Linearkombination von Elementen wie in (VII.9.3). Da aber α bilinear ist, gilt

$$\alpha(\lambda u_1 + \mu u_2, v) - \lambda\alpha(u_1, v) - \mu\alpha(u_2, v) = 0 \text{ und } \alpha(u, \lambda v_1 + \mu v_2) - \lambda\alpha(u, v_1) - \mu\alpha(u, v_2) = 0,$$

so dass $\alpha|_X = 0$ gilt.

Damit ist $\phi_\alpha\colon W/X \to Y$ wohldefiniert und linear. Es gilt nach Konstruktion

$$\phi_\alpha \circ \pi \circ \delta = \alpha,$$

weil wir $\phi_\alpha \circ \pi \circ \delta(u, v)$ genau als $\alpha(u, v)$ definiert hatten.

Da die Elemente $\delta(u, v)$ den Vektorraum W erzeugen, erzeugen die Elemente $\pi(\delta(u, v))$ den Quotientenvektorraum. Wenn wir verlangen, dass $\phi_\alpha \circ \pi \circ \delta = \alpha$ gilt, ist ϕ_α damit eindeutig festgelegt. □

Bemerkung VII.9.7. Die Elemente $u \otimes v$ heißen *Tensoren*. Es gilt insbesondere, dass

$$(\lambda u) \otimes v = \lambda(u \otimes v) = u \otimes \lambda v$$

gilt. Deshalb lassen wir im Folgenden manchmal die Klammern weg.

Wenn Sie in $U \otimes V$ rechnen möchten, müssen Sie immer die Bilinearität von $(-) \otimes (-)$ beachten: Ist $u = \sum_{i=1}^m \lambda_i u_i$ und $v = \sum_{j=1}^n \mu_j v_j$, so ist

$$\begin{aligned} u \otimes v &= \left(\sum_{i=1}^m \lambda_i u_i\right) \otimes \left(\sum_{j=1}^n \mu_j v_j\right) \\ &= \sum_{i=1}^m \lambda_i \left(u_i \otimes \left(\sum_{j=1}^n \mu_j v_j\right)\right) \\ &= \sum_{i=1}^m \lambda_i \sum_{j=1}^n \mu_j u_i \otimes v_j \\ &= \sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j u_i \otimes v_j \end{aligned}$$

Elemente der Form $u \otimes v$ erzeugen $U \otimes V$, aber nicht jedes Element in $U \otimes V$ läßt sich so ausdrücken: In $\mathbb{R}^2 \otimes \mathbb{R}^3$ ist zum Beispiel

$$2e_1 \otimes e_1 + 3e_2 \otimes e_3$$

nicht weiter zu vereinfachen.

Beispiel VII.9.8. In $\mathbb{Q}^3 \otimes \mathbb{Q}^3$ ist

$$-2e_1 \otimes e_1 - \frac{2}{5}e_1 \otimes e_2 - 6e_1 \otimes e_3 + 3e_2 \otimes e_1 + \frac{3}{5}e_2 \otimes e_2 + 9e_2 \otimes e_3 - \frac{1}{4}e_3 \otimes e_1 - \frac{1}{20}e_3 \otimes e_2 - \frac{3}{4}e_3 \otimes e_3$$

gleich $(2e_1 - 3e_2 + \frac{1}{4}e_3) \otimes (-e_1 - \frac{1}{5}e_2 - 3e_3)$.

Lemma VII.9.9. *Es sei K ein Körper und U, V seien K-Vektorräume. Ist $\dim_K U = m < \infty$ und $\dim_K V = n < \infty$, so ist*

$$\dim_K U \otimes V = m \cdot n.$$

Beweis. Es sei $\mathcal{B}_U = (u_1, \ldots, u_m)$ eine Basis von U und $\mathcal{B}_V = (v_1, \ldots, v_n)$ eine Basis von V. Dann ist $(u_i \otimes v_j, i = 1, \ldots, m, j = 1, \ldots, n)$ eine Basis von $U \otimes V$: Nach Konstruktion von $U \otimes V$ sind die $(u_i \otimes v_j, i = 1, \ldots, m, j = 1, \ldots, n)$ linear unabhängig und sie erzeugen $U \otimes V$. □

Bemerkung VII.9.10. Möchten Sie die Basis $(u_i \otimes v_j, i = 1, \ldots, m, j = 1, \ldots n)$ anordnen, so ist die sogenannte *lexikographische Anordnung* üblich:

$$(u_1 \otimes v_1, \ldots, u_1 \otimes v_n, u_2 \otimes v_1, \ldots, u_2 \otimes v_n, \ldots, u_m \otimes v_1, \ldots, u_m \otimes v_n)$$

Wenn Sie sich die Basiselemente in einer Matrix angeordnet vorstellen

$$\begin{pmatrix} u_1 \otimes v_1 & \ldots & u_1 \otimes v_n \\ u_2 \otimes v_1 & \ldots & u_2 \otimes v_n \\ \vdots & \vdots & \vdots \\ u_m \otimes v_1 & \ldots & u_m \otimes v_n \end{pmatrix},$$

dann laufen Sie die Zeilen entlang von oben nach unten.

Lemma VII.9.11. *Sind $f \in \mathsf{Hom}_K(U, U')$ und $g \in \mathsf{Hom}_K(V, V')$, so induzieren f und g eine K-lineare Abbildung*

$$f \otimes g \colon U \otimes V \to U' \otimes V'.$$

Beweis. Wir betrachten das Diagramm

$$\begin{tikzcd} U \times V \arrow[r, "\kappa"] \arrow[d, "{(f,g)}"'] & U \otimes V \\ U' \times V' \arrow[r, "\kappa'"] & U' \otimes V'. \end{tikzcd}$$

Hierbei sind κ und κ' die Strukturabbildungen des jeweiligen Tensorprodukts und $(f, g)(u, v) := (f(u), g(v))$. Die Abbildung $\kappa' \circ (f, g)$ ist K-bilinear. Daher existiert wegen der universellen Eigenschaft des Tensorprodukts $(U \otimes V, \kappa)$ eine eindeutige K-lineare Abbildung von $U \otimes V$ nach $U' \otimes V'$, die wir $f \otimes g$ nennen:

$$\begin{tikzcd} U \times V \arrow[r, "\kappa"] \arrow[d, "{(f,g)}"'] & U \otimes V \arrow[d, dotted, "f\otimes g"] \\ U' \times V' \arrow[r, "\kappa'"] & U' \otimes V'. \end{tikzcd}$$

Nach Konstruktion erfüllt die Abbildung die Gleichung

$$(f \otimes g) \circ \kappa = \kappa' \circ (f, g).$$

□

Bemerkung VII.9.12. Wir können die Abbildung $f \otimes g$ explizit beschreiben. Es sei $(u, v) \in U \times V$. Das kommutative Diagramm von oben ergibt für (u, v):

$$\begin{tikzcd} (u,v) \arrow[r, mapsto, "\kappa"] \arrow[d, mapsto, "{(f,g)}"'] & \kappa(u,v) = u \otimes v \arrow[d, mapsto, "f\otimes g"] \\ (f(u), g(v)) \arrow[r, mapsto, "\kappa'"] & \kappa'(f(u), g(v)) = f(u) \otimes g(v). \end{tikzcd}$$

Es gilt also $(f \otimes g)(u \otimes v) = f(u) \otimes g(v)$.

Beispiel VII.9.13. Für $f \colon \mathbb{R}^2 \to \mathbb{R}^3$ mit $M(f) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 2 \end{pmatrix}$, also $f(e_1) = e_1 + e_3$ und $f(e_2) = 2(e_2 + e_3)$ und für $g \colon \mathbb{R}^2 \to \mathbb{R}$ mit $g(e_1) = g(e_2) = e_1$, also $M(g) = \begin{pmatrix} 1 & 1 \end{pmatrix}$ hat $f \otimes g \colon \mathbb{R}^2 \otimes \mathbb{R}^2 \to \mathbb{R}^3 \otimes \mathbb{R}$ die Werte

$$(f \otimes g)(e_1 \otimes e_1) = (e_1 + e_3) \otimes e_1 = e_1 \otimes e_1 + e_3 \otimes e_1 = (f \otimes g)(e_1 \otimes e_2)$$

und

$$(f \otimes g)(e_2 \otimes e_1) = 2(e_2 + e_3) \otimes e_1 = 2e_2 \otimes e_1 + 2e_3 \otimes e_1 = (f \otimes g)(e_2 \otimes e_2).$$

Die Matrixdarstellung von $f \otimes g$ ergibt also mit der lexikographischen Anordnung der Basisvektoren die Matrix:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix}.$$

Satz VII.9.14. *(Rechenregeln für Tensorprodukte)*

(a) *Sind $f_1, f_2 \colon U \to U'$ und $g_1, g_2 \colon V \to V'$ K-linear, so gilt für alle $\lambda, \mu \in K$:*

$$\begin{aligned} (\lambda f_1 + \mu f_2) \otimes g_1 &= \lambda f_1 \otimes g_1 + \mu f_2 \otimes g_1, \\ f_1 \otimes (\lambda g_1 + \mu g_2) &= \lambda f_1 \otimes g_1 + \mu f_1 \otimes g_2, \\ (\mathrm{id}_{U'} \otimes g_1) \circ (f_1 \otimes \mathrm{id}_V) &= (f_1 \otimes \mathrm{id}_{V'}) \circ (\mathrm{id}_U \otimes g_1), \\ \mathrm{id}_U \otimes \mathrm{id}_V &= \mathrm{id}_{U \otimes V}. \end{aligned}$$

(b) *Ist $U = U_1 \oplus U_2$, so ist*

$$U \otimes V = (U_1 \oplus U_2) \otimes V \cong U_1 \otimes V \oplus U_2 \otimes V.$$

(c) *Es gilt $U \otimes V \cong V \otimes U$.*

(d) *Ist W ein weiterer K-Vektorraum, so ist*

$$(U \otimes V) \otimes W \cong U \otimes (V \otimes W).$$

(e)

$$K \otimes V \cong V \cong V \otimes K.$$

Das Tensorprodukt ist also bis auf Isomorphie eine assoziative und kommutative Verknüpfung, die ein Distributivgesetz bezüglich direkter Summen erfüllt und für die K bis auf Isomorphie ein Einselement ist.

Beweis. Zum Beweis der ersten Gleichung in (a) betrachten wir das Diagramm:

$$\begin{tikzcd}
U \times V \arrow[r, bend left, "(\lambda f_1+\mu f_2, g_1)"] \arrow[r, bend right, "\lambda(f_1,g_1)+\mu(f_2,g_1)"'] \arrow[d, "\kappa"'] & U' \times V' \arrow[r, "\kappa'"] & U' \otimes V' \\
U \otimes V \arrow[urr, bend right=10, "(\lambda f_1+\mu f_2)\otimes g_1"] \arrow[urr, bend right=40, "\lambda(f_1\otimes g_1)+\mu(f_2\otimes g_1)"'] & &
\end{tikzcd}$$

Es gilt aber

$$\begin{aligned}
((\lambda f_1 + \mu f_2) \otimes g_1) \circ \kappa(u, v) &= ((\lambda f_1 + \mu f_2) \otimes g_1)(u \otimes v) \\
&= (\lambda f_1 + \mu f_2)(u) \otimes g_1(v) \\
&= (\lambda f_1(u) + \mu f_2(u)) \otimes g_1(v) \\
&= \lambda f_1(u) \otimes g_1(v) + \mu f_2(u) \otimes g_1(v) \\
&= (\lambda f_1 \otimes g_1 + \mu f_2 \otimes g_1) \circ \kappa(u, v)
\end{aligned}$$

und damit sind die beiden Abbildungen $(\lambda f_1 + \mu f_2) \otimes g_1$ und $\lambda f_1 \otimes g_1 + \mu f_2 \otimes g_1$ gleich.

Die andere Gleichungen in (a) folgen mit analogen Rechnungen.

Für (b) betrachten wir das Diagramm

$$\begin{tikzcd}
U \times V \arrow[r, "\alpha"] \arrow[d, "\kappa"'] & (U_1 \otimes V) \oplus (U_2 \otimes V), \\
U \otimes V \arrow[ur, dotted, "\phi_\alpha"'] &
\end{tikzcd}$$

wobei α die Abbildung ist, die $(u, v) \in U \times V$ auf $u_1 \otimes v + u_2 \otimes v$ schickt, wenn $u = u_1 + u_2$ mit $u_i \in U_i$ für $i = 1, 2$.

Da α bilinear ist, gibt es genau ein K-lineares $\phi_\alpha\colon U \otimes V \to (U_1 \otimes V) \oplus (U_2 \otimes V)$ mit $\phi_\alpha \circ \kappa = \alpha$.

Wir definieren eine Umkehrabbildung ψ_α zu ϕ_α, die auf Erzeugern $u_i \otimes v$ definiert ist als $\psi_\alpha(u_i \otimes v) = u_i \otimes v$. Dann ist

$$\psi_\alpha \circ \phi_\alpha(u \otimes v) = \psi_\alpha(u_1 \otimes v + u_2 \otimes v) = \psi_\alpha(u_1 \otimes v) + \psi_\alpha(u_2 \otimes v) = u_1 \otimes v + u_2 \otimes v = (u_1 + u_2) \otimes v = u \otimes v.$$

Für einen Erzeuger gilt

$$\phi_\alpha \circ \psi_\alpha(u_i \otimes v) = \phi_\alpha(u_i \otimes v) = u_i \otimes v.$$

Damit stimmt $\psi_\alpha \circ \phi_\alpha$ auf Erzeugern mit $\mathrm{id}_{U \otimes V}$ überein und $\phi_\alpha \circ \psi_\alpha$ stimmt auf Erzeugern mit $\mathrm{id}_{(U_1 \otimes V) \oplus (U_2 \otimes V)}$ überein. Damit sind diese Abbildungen gleich.

Zu (c) definieren wir

$$c_{U,V}\colon U \otimes V \to V \otimes U$$

über

$$\begin{tikzcd}
U \times V \arrow[r, "\alpha"] \arrow[d, "\kappa"'] & V \otimes U, \\
U \otimes V \arrow[ur, dotted, "c_{U,V}"'] &
\end{tikzcd}$$

wobei $\alpha(u,v) = v \otimes u$ ist. Es gilt also $c_{U,V}(u \otimes v) = v \otimes u$.

Analog ist $c_{V,U} \colon V \otimes U \to U \otimes V$ definiert über

$$\begin{tikzcd}
V \times U \arrow[r, "\beta"] \arrow[d, "\kappa"'] & U \otimes V \\
V \otimes U \arrow[ur, dotted, "c_{V,U}"'] &
\end{tikzcd}$$

mit $\beta(v,u) = u \otimes v$.

Dann gilt

$$c_{V,U} \circ c_{U,V}(u \otimes v) = c_{V,U}(v \otimes u) = u \otimes v = \mathrm{id}_{U \otimes V}(u \otimes v),$$

und

$$c_{U,V} \circ c_{V,U}(v \otimes u) = c_{U,V}(u \otimes v) = v \otimes u = \mathrm{id}_{V \otimes U}(v \otimes u),$$

so dass die Verkettung jeweils die Identität ergibt.

Für die Assoziativität in (d) definieren wir $\alpha \colon (U \otimes V) \times W \to U \otimes (V \otimes W)$ auf Erzeugern als $\alpha(u \otimes v, w) := u \otimes (v \otimes w)$ und erhalten eine eindeutige K-lineare Abbildung $\gamma \colon (U \otimes V) \otimes W \to U \otimes (V \otimes W)$ mit $\gamma \circ \kappa = \alpha$:

$$\begin{tikzcd}
(U \otimes V) \times W \arrow[r, "\alpha"] \arrow[d, "\kappa"'] & U \otimes (V \otimes W) \\
(U \otimes V) \otimes W \arrow[ur, dotted, "\gamma"'] &
\end{tikzcd}$$

Das Inverse zu γ erhalten wir durch

$$\begin{tikzcd}
U \times (V \otimes W) \arrow[r, "\beta"] \arrow[d, "\kappa"'] & (U \otimes V) \otimes W \\
U \otimes (V \otimes W) \arrow[ur, dotted, "\gamma^{-1}"'] &
\end{tikzcd}$$

mit $\beta(u, v \otimes w) := (u \otimes v) \otimes w$.

Zu (e) betrachten wir die K-lineare Abbildung $f \colon K \otimes V \to V$, die auf Erzeugern $\lambda \otimes v$ gegeben ist durch $f(\lambda \otimes v) = \lambda v$. Umgekehrt setzen wir $g \colon V \to K \otimes V$ an als $g(v) = 1 \otimes v$. Sie rechnen nach, dass diese Abbildungen wohldefiniert sind. Es ist

$$(g \circ f)(\lambda \otimes v) = g(\lambda v) = \lambda g(v) = \lambda(1 \otimes v) = \lambda \otimes v$$

und

$$(f \circ g)(v) = f(1 \otimes v) = 1 \cdot v = v.$$

Analog zeigen Sie, dass $V \otimes K \cong V$ gilt. □

Tensorprodukte interagieren mit Homomorphismenräumen in der folgenden Art und Weise:

Satz VII.9.15. *Sind V und W endlich-dimensionale K-Vektorräume, so gilt:*

(a) $(V \otimes W)^* \cong V^* \otimes W^*$.
(b) $V^* \otimes W = \mathsf{Hom}_K(V, K) \otimes W \cong \mathsf{Hom}_K(V, W)$.

Beweis. Sie können Dimensionen zählen und sehen, dass die beteiligten Vektorräume die gleiche Dimension haben, so dass sie isomorph sein müssen. Wir möchten aber explizite Isomorphismen angeben:

Für (a) definieren wir eine Abbildung $h\colon V^* \otimes W^* \to (V \otimes W)^*$, indem wir für $\varphi \in V^*$ und $\psi \in W^*$ auf dem Erzeuger $\varphi \otimes \psi$ von $V^* \otimes W^*$ setzen

$$h(\varphi \otimes \psi)(v \otimes w) := \varphi(v) \cdot \psi(w).$$

wobei die Multiplikation in K stattfindet. Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $\mathcal{B}' = (w_1, \dots, w_m)$ eine Basis von W, so bildet h das Basiselement $v_i^* \otimes w_j^*$ ab auf

$$h(v_i^* \otimes w_j^*)(v_k \otimes w_\ell) = v_i^*(v_k) \cdot w_j^*(w_\ell) = \delta_{ik} \cdot \delta_{j\ell} = (v_i \otimes w_j)^*(v_k \otimes w_\ell)$$

und somit gilt

$$h(v_i^* \otimes w_j^*) = (v_i \otimes w_j)^*$$

und da $((v_i \otimes w_j)^*, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m)$ eine Basis von $(V \otimes W)^*$ ist, ist h ein Isomorphismus.

Für (b) sei $\tilde{h}\colon \mathsf{Hom}_K(V, K) \otimes W \to \mathsf{Hom}_K(V, W)$ definiert als

$$\tilde{h}(\varphi \otimes w)(v) := \varphi(v) \cdot w$$

für $v \in V$, $w \in W$ und $\varphi \in V^*$. Mit der Notation von eben wissen wir, dass

$$((v_i^* \otimes w_j), 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m)$$

eine Basis von $\mathsf{Hom}_K(V, K) \otimes W$ ist. Für $\mathsf{Hom}_K(V, W)$ hatten wir in Korollar III.2.23 die Basis $(\varphi_j^i, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m)$ betrachtet mit

$$(\varphi_j^i(v_k)) = \begin{cases} w_j, & \text{falls } k = i, \\ 0, & \text{sonst.} \end{cases}$$

Es ist

$$\tilde{h}(v_i^* \otimes w_j)(v_k) = v_i^*(v_k) \cdot w_j = \delta_{ik} \cdot w_j = \varphi_j^i(v_k)$$

und damit bildet $\tilde{h}$ eine Basis von $\mathsf{Hom}_K(V, K) \otimes W$ auf eine Basis von $\mathsf{Hom}_K(V, W)$ ab. □

Bemerkung VII.9.16.

- Wir hatten Matrizen zum einen als darstellende Matrizen linearer Abbildungen betrachtet, aber auch als Matrizen, die eine Bilinearform beschreiben.

  Ist $A \in M(m \times n, K)$, so ist die Multiplikation mit A eine Abbildung

  $$A\cdot\colon K^n \to K^m,$$

  also entspricht A unter dem obigen Isomorphismen einem Element in $\mathsf{Hom}_K(K^n, K^m) \cong (K^n)^* \otimes K^m$.

  Betrachten wir die Bilinearform, die durch A gegeben ist, also

  $$K^m \times K^n \to K, \quad (x, y) \mapsto x^t \cdot A \cdot y,$$

  so entspricht A einem Element in

  $$\mathrm{Bil}(K^m, K^n) \cong \mathsf{Hom}_K(K^m \otimes K^n, K) = (K^m \otimes K^n)^* \cong (K^m)^* \otimes (K^n)^*.$$

- Wir hatten in Definition VII.5.19 die Komplexifizierung $V_\mathbb{C}$ eines reellen Vektorraums V betrachtet. Es gilt, dass

  $$V_\mathbb{C} \cong \mathbb{C} \otimes_\mathbb{R} V$$

  ist. Sie bilden ein $(v_1, v_2) \in V_\mathbb{C}$ auf $1 \otimes v_1 + i \otimes v_2 \in \mathbb{C} \otimes_\mathbb{R} V$ ab. Dies ist ein Isomorphismus von $\mathbb{C}$-Vektorräumen.

VII.10. Alternierende bilineare Abbildungen

Wir hatten schon den Begriff der alternierenden Bilinearform. Wir betrachten zunächst leicht allgemeiner alternierende bilineare Abbildungen:

Definition VII.10.1. Es seien V und W K-Vektorräume. Eine bilineare Abbildung $\beta\colon V \times V \to W$ heißt *alternierend*, falls $\beta(v,v) = 0$ gilt für alle $v \in V$.

Bemerkung VII.10.2. Ist $\beta(v,v) = 0$ für alle $v \in V$, so gilt für alle $v_1, v_2 \in V$:

$$\beta(v_1, v_2) = -\beta(v_2, v_1),$$

weil dann

$$0 = \beta(v_1 + v_2, v_1 + v_2) = \beta(v_1, v_1) + \beta(v_1, v_2) + \beta(v_2, v_1) + \beta(v_2, v_2) = \beta(v_1, v_2) + \beta(v_2, v_1)$$

ist.

Gilt umgekehrt $\beta(v_1, v_2) = -\beta(v_2, v_1)$ für alle $v_1, v_2 \in V$, so ist für $v_1 = v_2 = v$

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

Ist die Charaktertistik von K nicht 2, so impliziert dies, dass $\beta(v,v) = 0$ ist für alle $v \in V$.

Wir können das Tensorprodukt benutzen, um alternierende bilineare Abbildungen zu identifizieren:

Satz VII.10.3. *Es sei $\beta\colon V \times V \to W$ bilinear. Dann ist β genau dann alternierend, wenn gilt*

$$\mathsf{Span}_K(v \otimes v, v \in V) \subset \ker(\phi_\beta\colon V \otimes V \to W).$$

$$\begin{tikzcd} V \times V \arrow[r, "\beta"] \arrow[d, "\kappa"'] & W \\ V \otimes V \arrow[ur, dotted, "\phi_\beta"'] & \end{tikzcd}$$

Beweis. Dies folgt nach Definition, weil

$$\beta(v,v) = \phi_\beta \circ \kappa(v,v) = \phi_\beta(v \otimes v).$$

□

Satz VII.10.4. *Es sei V ein K-Vektorraum mit $\dim_K V \geqslant 2$.*

(a) *Es gibt einen K-Vektorraum $\Lambda^2 V$ zusammen mit einer alternierenden bilinearen Abbildung*

$$\wedge\colon V \times V \to \Lambda^2 V,$$

so dass $(\Lambda^2 V, \wedge)$ die folgende universelle Eigenschaft hat:

Für alle alternierenden bilinearen Abbildungen $f\colon V \times V \to W$ gibt es genau eine K-lineare Abbildung $f_\wedge\colon \Lambda^2 V \to W$ mit $f_\wedge \circ \wedge = f$.

$$\begin{tikzcd} V \times V \arrow[r, "f"] \arrow[d, "\wedge"'] & W \\ \Lambda^2 V \arrow[ur, dotted, "\exists! f_\wedge"'] & \end{tikzcd}$$

(b) *Ist $\dim_K V = n \geqslant 2$ und hat V die Basis $\mathcal{B} = (v_1, \dots, v_n)$, so ist $(\wedge(v_i, v_j), 1 \leqslant i < j \leqslant n)$ eine Basis von $\Lambda^2 V$, also gilt*

$$\dim_K \Lambda^2 V = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Wir schreiben im Folgenden $v \wedge v'$ für $\wedge(v, v')$.

Beweis. Für (a) definieren wir

$$\Lambda^2 V := V \otimes V / \mathsf{Span}_K(v \otimes v, v \in V)$$

und $\pi\colon V \otimes V \to \Lambda^2 V$ sei die kanonische Projektion. Mit $\kappa\colon V \times V \to V \otimes V$ setzen wir

$$\wedge := \pi \circ \kappa,$$

also

$$v \wedge v' = \pi(\kappa(v, v')) = \pi(v \otimes v').$$

Nach Konstruktion ist $\wedge$ bilinear und alternierend. Wir zeigen die universelle Eigenschaft. Dazu sei $f\colon V \times V \to W$ bilinear und alternierend. Wir betrachten das Diagramm

$$\begin{tikzcd}
V \times V \arrow[r, "f"] \arrow[d, "\kappa"'] & W \\
V \otimes V \arrow[d, "\pi"'] & \\
\Lambda^2 V &
\end{tikzcd}$$

Wegen der universellen Eigenschaft des Tensorprodukts $(V \otimes V, \kappa)$ gibt es eine eindeutige K-lineare Abbildung $\phi_f\colon V \otimes V \to W$ mit $\phi_f \circ \kappa = f$.

$$\begin{tikzcd}
V \times V \arrow[r, "f"] \arrow[d, "\kappa"'] & W \\
V \otimes V \arrow[d, "\pi"'] \arrow[ur, dotted, "\phi_f"'] & \\
\Lambda^2 V &
\end{tikzcd}$$

Für alle $v \in V$ gilt

$$\phi_f(v \otimes v) = f(v, v) = 0,$$

also ist $\mathsf{Span}_K(v \otimes v, v \in V) \subset \ker(\phi_f)$. Mit der universellen Eigenschaft des Quotientenvektorraums gibt es eine eindeutige lineare Abbildung $f_\wedge\colon \Lambda^2 V \to W$ mit

$$f_\wedge \circ \pi = \phi_f.$$

$$\begin{tikzcd}
V \times V \arrow[r, "f"] \arrow[d, "\kappa"'] & W \\
V \otimes V \arrow[d, "\pi"'] \arrow[ur, dotted, "\phi_f"'] & \\
\Lambda^2 V \arrow[uur, dotted, "f_\wedge"'] &
\end{tikzcd}$$

Dann gilt aber auch:

$$\begin{aligned}
f_\wedge \circ \wedge &= f_\wedge \circ \pi \circ \kappa \\
&= \phi_f \circ \kappa \\
&= f.
\end{aligned}$$

Für (b) ist klar, dass die Elemente $(v_i \wedge v_j, 1 \leqslant i, j \leqslant n)$ den Vektorraum $\Lambda^2 V$ erzeugen, weil $(v_i \otimes v_j, 1 \leqslant i, j \leqslant n)$ den Vektorraum $V \otimes V$ erzeugen und weil $\Lambda^2 V$ ein Quotient von $V \otimes V$ ist.

Für diese Erzeuger gelten die Relationen

$$v_i \wedge v_i = 0,$$
$$v_i \wedge v_j = -v_j \wedge v_i \text{ für } i \neq j.$$

Damit ist $(v_i \wedge v_j, 1 \leqslant i < j \leqslant n)$ ein Erzeugendensystem. Dieses System ist minimal. Nehmen wir an, wir könnten ein $v_i \wedge v_j$ weglassen für $i < j$. Es ist

$$0 \neq \pi(v_i \otimes v_j) = v_i \wedge v_j$$

und die Äquivalenzklasse von $v_i \otimes v_j$ enthält nur zwei Elemente, $v_i \otimes v_j$ und $-v_j \otimes v_i$, und $v_j \wedge v_i$ ist nicht in unserem Erzeugendensystem, weil $j > i$. □

Bemerkung VII.10.5. Ist V ein n-dimensionaler K-Vektorraum und ist $2 \leqslant k \leqslant n$, dann können Sie analog die k-te äußere Potenz von V, $\Lambda^k V$, definieren und konstruieren, so dass $\Lambda^k V$ eine universelle Eigenschaft hat für Abbildungen

$$f\colon \underbrace{V \times \ldots \times V}_{k} \to W,$$

die alternierend sind, in dem Sinne, dass

$$f(v_1', \ldots, v_k') = 0 \text{ gilt, falls } \exists i \neq j, \text{ so dass } v_i' = v_j'.$$

Dann gilt

$$\dim_K \Lambda^k V = \binom{n}{k}$$

und eine Basis ist gegeben durch $(v_{i_1} \wedge \ldots \wedge v_{i_k}, 1 \leqslant i_1 < \ldots < i_k \leqslant n)$, falls $(v_1, \ldots, v_n)$ eine Basis von V ist.