Albert-Ludwigs-Universität Freiburg

Vorlesungsskript

Lineare Algebra I und II

Angelika Rohde

Wintersemester 2023/24 und Sommersemester 2024

Inhaltsverzeichnis

1 Grundlagen

1.1 Mengen

Beispiel 1.1 (Zahlbereiche)**.**

- $\mathbb{N} = \{1, 2, 3, \dots\}$ *Menge der natürlichen Zahlen*
- $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ *Menge der natürlichen Zahlen mit Null*
- $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ *Menge der ganzen Zahlen*
- $\mathbb{Q} = \{\frac{p}{q} | \, p \in \mathbb{Z}, q \in \mathbb{N}\}$ *Menge der rationalen Zahlen*
- $\mathbb{R}$ *Menge der reellen Zahlen (→ Vorlesung Analysis)*
- $\mathbb{C}$ *Menge der komplexen Zahlen (→ Vorlesung Funktionentheorie)*

Wir möchten hier nicht formal den subtilen des Begriff der Menge erörtern, das ist Gegenstand der Mengenlehre. Für unsere Zwecke besteht eine Menge M wie in Beispiel 1.1 aus unterschiedlichen Elementen; wir schreiben $a \in M$ ("a Element M"), falls das Element a in M enthalten ist, andernfalls gilt $a \notin M$ ("a nicht Element M"). $\emptyset$ bezeichnet die leere Menge, die dadurch ausgezeichnet ist, dass sie keine Elemente enthält.

Notation 1.2 (Mengenangabe mit charakterisierender Eigenschaft)**.**

$$M = \{x \in A | \, x \text{ hat Eigenschaft } E\}$$

Beispiele. $\{x \in \mathbb{N} | \, 3 < x < 7\} = \{4, 5, 6\}$ *und* $\{x \in \mathbb{Z} | \, x^2 = -1\} = \emptyset$ *(leere Menge).*

Definition 1.3. *Seien M und N Mengen.*

M heißt Teilmenge von N ($M \subset N$), wenn gilt: $a \in M \Rightarrow a \in N$.

M heißt echte Teilmenge von N ($M \subsetneq N$), falls $M \subset N$ und $M \neq N$.

Beispiel 1.4. $\mathbb{N} \subsetneq \mathbb{N}_0 \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}$.

Definition 1.5. *Seien M, N Mengen. Wir definieren (":=" bedeutet "ist definiert als")*

$M \cup N := \{a | \, a \in M \text{ oder } a \in N\}$ *(Vereinigung)*

$M \cap N := \{a | \, a \in M \text{ und } a \in N\}$ *(Durchschnitt)*

$M \setminus N := \{a | \, a \in M \text{ und } a \notin N\}$ *(Differenz "M ohne N").*

Lemma 1.6. *Seien A, B, C Mengen. Dann gilt*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Für "=" müssen wir "$\subset$" und "$\supset$" beweisen. Wir zeigen eine Behauptung der Form $D \subset E$, indem wir für beliebiges $x \in D$ folgern, dass auch $x \in E$, also formal: $x \in D \Rightarrow x \in E$ ("$\Rightarrow$" bedeutet "daraus folgt").

Beweis.

"$\subset$": Sei $x \in A \cap (B \cup C)$.
$\Rightarrow x \in A$ und $x \in B \cup C$.
1. Fall: $x \in A$ und $x \in B \Rightarrow x \in A \cap B \subset (A \cap B) \cup (A \cap C)$
2. Fall: $x \in A$ und $x \in C \Rightarrow x \in A \cap C \subset (A \cap B) \cup (A \cap C)$.

"$\supset$": Sei $x \in (A \cap B) \cup (A \cap C)$.
($x \in A$ und $x \in B$) oder ($x \in A$ und $x \in C$)
$\Rightarrow x \in A$ und ($x \in B$ oder $x \in C$)
$\Rightarrow x \in A \cap (B \cup C)$. □

Bemerkung 1.7. *Wir haben hier die Pfeile "$\Rightarrow$" verwendet, wenn von einer Aussage auf eine andere geschlossen wurde. Im Folgenden wird "$\Leftrightarrow$" für eine Äquivalenzaussage verwendet, d.h., wenn beide Richtungen, "$\Rightarrow$" und "$\Leftarrow$", gelten.*

Lemma 1.8. *Seien A, B Mengen. Dann sind äquivalent:*

(i) $A \cup B = B$;

(ii) $A \subset B$.

In Kurzschreibweise: Es gilt $A \cup B = B \Leftrightarrow A \subset B$.

Beweis. Wir zeigen für die Äquivalenz "$\Leftrightarrow$" die beiden Richtungen "$\Rightarrow$" und "$\Leftarrow$".

"$\Rightarrow$": (D.h. wir nehmen an, dass $A \cup B = B$ gilt und müssen $A \subset B$ beweisen.)
Sei $x \in A$. $\Rightarrow x \in A \cup B \overset{(i)}{=} B$, d.h. $A \subset B$.

"$\Leftarrow$": (D.h. wir nehmen an, dass $A \subset B$ gilt und müssen $A \cup B = B$ beweisen, also $A \cup B \subset B$ sowie $B \subset A \cup B$.)
Sei $x \in A \cup B \Rightarrow x \in A$ oder $x \in B \overset{(ii)}{\Rightarrow} x \in B$, also haben wir $A \cup B \subset B$ gezeigt.
$B \subset A \cup B$ ist aber klar, womit $B = A \cup B$. □

Zusammenfassung 1.9 (Direkte Beweismethode)**.** *Die Aussage eines Satzes besteht immer aus einer (oder mehreren) Implikationen*

$$\textit{Aussage } A \Rightarrow \textit{Aussage } B$$

oder Äquivalenzen

$$\textit{Aussage } A \Leftrightarrow \textit{Aussage } B.$$

(Die Notation "⇔" bedeutet "genau dann wenn" und entspricht den beiden Implikationen Aussage A ⇒ Aussage B und Aussage B ⇒ Aussage A.) Die direkte Beweismethode besteht darin, eine Implikation durch eine Abfolge von direkten Schlüssen (Aussage A ⇒ Aussage A1 ⇒ Aussage A2 ⇒ ... ⇒ Aussage B) zu beweisen.

Beispiel 1.10. *Wir wollen zeigen:*

x_0 *ist eine Lösung von* $(x^2 - px + q) = 0$

$\Rightarrow$

$$x_0 = \frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} \quad \textit{und} \quad \frac{p^2}{4} - q \geq 0.$$

Anfängerfehler: Forme $x_0 = \frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$ *so lange um, bis die erste Aussage dasteht:*

$$\begin{aligned} & x_0 - \frac{p}{2} = \pm\sqrt{\frac{p^2}{4} - q} \\ \Rightarrow \quad & \left(x_0 - \frac{p}{2}\right)^2 = \frac{p^2}{4} - q \\ \Rightarrow \quad & x_0^2 - px_0 + \frac{p^2}{4} = \frac{p^2}{4} - q \\ \Rightarrow \quad & x^2 - px + q = 0. \end{aligned}$$

Dies wäre als Beweis obiger Aussage falsch, weil die falsche Schlussrichtung gezeigt wird.

Richtig: x_0 *Lösung von* $x^2 - px + q = 0$

$$\begin{aligned} \Rightarrow \quad & x_0^2 - px_0 + \frac{p^2}{4} = \frac{p^2}{4} - q \\ \Rightarrow \quad & \left(x_0 - \frac{p}{2}\right)^2 = \frac{p^2}{4} - q \\ \Rightarrow \quad & \frac{p^2}{4} - q \geq 0 \quad \textit{und} \quad x_0 - \frac{p}{2} = \pm\sqrt{\frac{p^2}{4} - q} \\ \Rightarrow \quad & \frac{p^2}{4} - q \geq 0 \quad \textit{und} \quad x_0 = \frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \end{aligned}$$

□

1.2 Abbildungen

Definition 1.11. *Seien M und N Mengen. Eine Abbildung f von M nach N ist eine Vorschrift, die jedem Element* $m \in M$ *genau ein Element aus N zuordnet. Notation:*

$$\begin{aligned} f : M &\to N \\ m &\mapsto f(m). \end{aligned}$$

Beispiel. $f : \mathbb{R} \to \mathbb{R}^2,\ x \mapsto (x^2, x+1)$.

Bemerkung 1.12. *Seien M, M', N, N' Mengen.*

(i) Zwei Abbildungen $f : M \to N$ und $g : M' \to N'$ sind gleich, wenn $M = M'$, $N = N'$ und $f(m) = g(m)$ für alle $m \in M$.

(ii) $id_M : M \to M$, $m \mapsto m$, heißt Identität (oder identische Abbildung) auf M.

(iii) Für $A \subset M$ heißt $f_{|A} : A \to N$, $a \mapsto f(a)$, die Einschränkung von f auf A.

Bemerkung 1.13. *In obigem Beispiel haben wir die Notation $(x^2, x+1)$ verwendet. (m, n) heißt Tupel (geordnetes Paar). Formal: $M \times N = \{(m, n) |\, m \in M, n \in N\}$ heißt kartesisches Produkt von M und N. Es gilt $(m, n) = (m', n')$ genau dann, wenn $m = m'$ und $n = n'$.*

Beispiel. $\mathbb{R} \times \mathbb{R} = \{(a, b) |\, a \in \mathbb{R}, b \in \mathbb{R}\} = \mathbb{R}^2$ *(reelle Ebene).*

Analog definiert man $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R} := \{(a_1, \ldots, a_n) |\, a_i \in \mathbb{R} \text{ für } i = 1, \ldots, n\}$.

Definition 1.14. *Seien M, N Mengen, $f : M \to N$ eine Abbildung. f heißt*

- *injektiv, falls gilt: Für alle $m_1, m_2 \in M$ folgt aus $f(m_1) = f(m_2)$ stets $m_1 = m_2$ $\Leftrightarrow$ für alle $m_1, m_2 \in M$ folgt aus $m_1 \neq m_2$ stets $f(m_1) \neq f(m_2)$;*
- *surjektiv, falls gilt: Für jedes $n \in N$ existiert ein $m \in M$ mit $f(m) = n$;*
- *bijektiv, falls f injektiv und surjektiv ist.*

Beispiele 1.15.

(i) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist nicht injektiv, denn $f(2) = f(-2)$, aber $2 \neq -2$. f ist auch nicht surjektiv, denn es existiert kein $x \in \mathbb{R}$ mit $f(x) = -1$.

(ii) $f : [0, \infty) \to \mathbb{R}$, $x \mapsto x^2$, ist injektiv, denn sind $x_1, x_2 \in [0, \infty)$ mit $x_1^2 = f(x_1) = f(x_2) = x_2^2$, so folgt $x_1 = x_2$. Es gibt aber weiterhin kein x mit $f(x) = -1$, d.h. f ist nicht surjektiv.

(iii) $f : [0, \infty) \to [0, \infty)$, $x \mapsto x^2$, ist injektiv (wie oben) und auch surjektiv, denn für alle $y \geq 0$ gilt $f(\sqrt{y}) = (\sqrt{y})^2 = y$. Damit ist f auch bijektiv.

Definition 1.16. *Seien L, M, N Mengen und $f : L \to M$, $g : M \to N$ Abbildungen. $g \circ f : L \to N$, $l \mapsto g(f(l))$, heißt die Verknüpfung (oder Hintereinanderschaltung, Verkettung, Komposition) von f und g.*

Beispiel 1.17. *Seien $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ und $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto x + 1$. Dann ist $g \circ f : \mathbb{R} \to \mathbb{R}$, $x \mapsto g(f(x)) = x^2 + 1$.*

Lemma 1.18. *Seien L, M, N, O Mengen, $f : L \to M$, $g : M \to N$, $h : N \to O$ Abbildungen. Dann gilt:*

$$h \circ (g \circ f) = (h \circ g) \circ f,$$

d.h. die Verknüpfung von Abbildungen ist assoziativ.

Beweis. Für $x \in L$ gilt

$$\big(h \circ (g \circ f)\big)(x) = h\big((g \circ f)(x)\big) = h\big(g(f(x))\big) = (h \circ g)(f(x)) = \big((h \circ g) \circ f\big)(x).$$

□

Wir verwenden ab nun folgende *Abkürzung:* " $\forall$ " bedeutet "für alle".

Lemma und Definition 1.19. *Seien M, N Mengen, $f : M \to N$ eine Abbildung. Dann sind folgende Aussagen äquivalent:*

(i) *f ist bijektiv.*

(ii) *Zu jedem $n \in N$ gibt es genau ein $m \in M$ mit $f(m) = n$.*

(iii) *Es gibt genau eine Abbildung $g : N \to M$ mit $g \circ f = id_M$ und $f \circ g = id_N$. In diesem Fall bezeichnen wir die Abbildung g mit f^{-1} und nennen f^{-1} die Umkehrabbildung (oder inverse Abbildung) zu f, d.h. es gelten $f^{-1}(f(m)) = m \ \forall m \in M$ und $f(f^{-1}(n)) = n \ \forall n \in N$.*

Beweis. Statt (i) $\Leftrightarrow$ (ii) und (ii) $\Leftrightarrow$ (iii) zeigen wir (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii): Sei f bijektiv.
Zu zeigen: Zu jedem $n \in N$ existiert genau ein $m \in M$ mit $f(m) = n$.

- "existiert" folgt aus der Surjektivität,
- "genau ein": Seien $m_1, m_2 \in M$ mit $f(m_1) = f(m_2) = n \Rightarrow m_1 = m_2$ wegen der Injektivität.

(ii) $\Rightarrow$ (iii): Angenommen, zu jedem $n \in N$ existiert genau ein $m \in M$ mit $f(m) = n$.
Zu zeigen: Es existiert genau eine Umkehrabbildung g mit $g \circ f = id_M$ und $f \circ g = id_N$.

- "existiert": Wir definieren

$$\begin{aligned} g : N &\to M, \\ n &\mapsto \text{ das eindeutig bestimmte } m \text{ mit } f(m) = n. \end{aligned}$$

  Dann gilt einerseits für $m \in M$: $(g \circ f)(m) = g(f(m)) = m$, d.h. $g \circ f = id_M$, andererseits für $n \in N$: $(f \circ g)(n) = f(g(n)) = n$, d.h. $f \circ g = id_N$.

- "genau eine": Seien $g_1, g_2 : N \to M$ mit $f \circ g_1 = f \circ g_2 = id_N$ und $g_1 \circ f = g_2 \circ f = id_M$. Zu zeigen: $g_1 = g_2$.
  Nach Lemma 1.18 ist $g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2$. Damit gilt

$$\begin{array}{rl} & g_1 \circ (f \circ g_2) = g_1 \circ id_N = g_1 \\ & \quad \| \\ & (g_1 \circ f) \circ g_2 = id_M \circ g_2 = g_2, \qquad \text{d.h. } g_1 = g_2. \end{array}$$

(iii) ⇒ (i): Wir setzen (iii) voraus. Zu zeigen: f ist bijektiv.

- "injektiv": Seien $m_1, m_2 \in M$ mit $f(m_1) = f(m_2)$.

$$\begin{array}{rl} \Rightarrow & f^{-1}(f(m_1)) = f^{-1}(f(m_2)) \\ \Rightarrow & (f^{-1} \circ f)(m_1) = (f^{-1} \circ f)(m_2) \\ & \quad \| \\ & \quad id_M \\ \Rightarrow & m_1 = m_2. \end{array}$$

- "surjektiv": Sei $n \in N \Rightarrow id_N(n) = n \;\Rightarrow\; (f \circ f^{-1})(n) = n \;\Rightarrow f(f^{-1}(n)) = n$. Mit $m = f^{-1}(n)$ ist dann $f(m) = n$.

[Bemerkung: Wir haben hier für die Implikation (iii) ⇒ (i) nicht die Eindeutigkeit von g verwendet und damit eine etwas stärkere Aussage bewiesen.] □

Beispiel 1.20. *$f : [0, \infty) \to [0, \infty)$, $x \mapsto x^2$, ist bijektiv (siehe oben). Die Umkehrabbildung f^{-1} ist gegeben durch $f^{-1}[0, \infty) \to [0, \infty)$, $x \mapsto \sqrt{x}$.*

Definition 1.21. *Seien M, N Mengen, $f : M \to N$ eine Abbildung, $A \subset M$, $B \subset N$.*

- *$f(A) := \{f(a) | \, a \in A\} \subset N$ heißt Bild von A (unter f).*
- *$f^{-1}(B) := \{m \in M | \, f(m) \in B\} \subset M$ heißt Urbild von B (unter f).*

Bemerkung. *Die Notation $f(A)$ bzw. $f^{-1}(B)$ ist zwar üblich, aber nicht unbedingt glücklich, weil einerseits f dabei quasi zu einer Abbildung auf Teilmengen von M "erweitert" wird, andererseits eine Bezeichnungskollision mit der inversen Abbildung f^{-1} besteht. Das Urbild ist immer definiert, auch dann, wenn die Abbildung f nicht bijektiv ist und die inverse Abbildung nicht existiert. Falls f aber invertierbar ist mit inverser Abbildung f^{-1}, dann gilt $f^{-1}(\{n\}) = \{f^{-1}(n)\}$, $n \in N$.*

1.3 Äquivalenzrelationen

Mitunter ist es zweckmäßg, Relationen zwischen zwei Elementen m_1, m_2 einer Menge M zu studieren. Notation: $m_1 \sim m_2$.

Beispiel 1.22.

(i) M *Menge der Freiburger Mathe-Studenten,* $m_1 \sim m_2 :\Leftrightarrow m_1$ *kennt* m_2.

(ii) $M = \mathbb{R}$, $m_1 \sim m_2 :\Leftrightarrow m_1 \leq m_2$.

(iii) $M = \mathbb{R}^2$, $x, y \in M$, *wobei* $x = (x_1, x_2)$, $y = (y_1, y_2)$. $x \sim y :\Leftrightarrow x_1^2 + x_2^2 = y_1^2 + y_2^2$.

Man kann eine Relation beschreiben durch ihren Graphen $R \subset M \times M$, wobei

$$(m_1, m_2) \in R \;\Leftrightarrow\; m_1 \sim m_2. \tag{1.1}$$

Entsprechend kann man eine Relation defnieren durch eine Teilmenge R von $M \times M$ und das Zeichen $\sim$ durch (1.1).

Definition 1.23. *Eine Relation* $\sim$ *auf einer Menge* M *heißt* Äquivalenzrelation, *wenn für beliebige* $m_1, m_2, m_3 \in M$ *gilt:*

(i) $m_1 \sim m_1$ *(Reflexivität)*

(ii) $m_1 \sim m_2 \Rightarrow m_2 \sim m_1$ *(Symmetrie)*

(iii) $m_1 \sim m_2$ *und* $m_2 \sim m_3 \Rightarrow m_1 \sim m_3$ *(Transitivität).*

Die Relation aus Beispiel 1.22 (iii) definiert eine Äquivalenzrelation, (i) und (ii) jedoch nicht.

Definition 1.24. *Ist eine Äquivalenzrelation* $\sim$ *auf einer Menge* M *gegeben, so heißt eine Teilmenge* $A \subset M$ Äquivalenzklasse *(bezüglich* $\sim$*), falls gilt:*

- $A \neq \emptyset$;
- $m, n \in A \Rightarrow m \sim n$;
- $m \in A$, $n \in M$, $m \sim n \Rightarrow n \in A$.

Lemma 1.25. *Ist* $\sim$ *eine Äquivalenzrelation auf einer Menge* M*, so gehört jedes Element* $a \in M$ *zu genau einer Äquivalenzklasse. Insbesondere gilt für zwei beliebige Äquivalenzklassen* A, A' *entweder* $A = A'$ *oder* $A \cap A' = \emptyset$.

Beweis. Für $a \in M$ sei $A \subset M$ definiert als $A := \{m \in M | \, m \sim a\}$. Wir zeigen, dass A eine Äquivalenzklasse ist, die a enthält. Wegen $a \sim a$ ist $a \in A \Rightarrow A \neq \emptyset$. Sind $m, n \in A$, also $m \sim a$ und $n \sim a$, so folgt $m \sim n$ nach Definition 1.23 (ii) und (iii). Ist $m \in A$, $n \in M$ und $m \sim n$, so ist wegen $m \sim a$ auch $n \sim a$ nach Definition 1.23 (ii) und (iii). Aber dann gilt $n \in A$. Damit ist A eine Äquivalenzklasse, die a enthält.

Es bleibt der Nachweis, dass für zwei beliebige Äquivalenzklassen A, A' entweder $A = A'$ oder $A \cap A' = \emptyset$ gilt. Angenommen, $A \cap A' \neq \emptyset$ und $a \in A \cap A'$. Ist $m \in A$, so ist $m \sim a$ und wegen $a \in A'$ auch $m \in A' \Rightarrow A \subset A'$. Genauso zeigt man $A' \subset A$, woraus $A = A'$ folgt. □

Eine Äquivalenzrelation $\sim$ auf einer Menge M liefert also eine Zerlegung von M in disjunkte Äquivalenzklassen. Diese Äquivalenzklassen kann man nun als Elemente einer neuen Menge auffassen.

Definition 1.26. *Sei M eine Menge, $\sim$ eine Äquivalenzrelation darauf. Die Menge der Äquivalenzklassen (bezüglich $\sim$) heißt Quotientenmenge von M nach der Äquivalenzrelation $\sim$ und wird mit $M/\sim$ bezeichnet.*

Indem man jedem Element $a \in M$ diejenige Äquivalenzklasse A_a zuordnet, in der es enthalten ist, erhält man eine kanonische Abbildung $M \to M/\sim$, $a \mapsto A_a$. Das Urbild der einelementigen Menge $\{A\}$ für ein *Element* A aus $M/\sim$ ist wieder A, aber diesmal aufgefasst als *Teilmenge* von M. Jedes $a \in A$ heißt Repräsentant der Äquivalenzklasse A.

Beispiel 1.27. *Wir betrachten noch einmal $\mathbb{R}^2$ mit der Äquivalenzrelation $\sim$ aus Beispiel 1.22 (iii). Die Äquivalenzklassen sind konzentrische Kreislinien um $(0,0)$ (die Äquivalenzklasse $\{(0,0)\}$ wird dabei als Kreislinie mit Radius Null aufgefasst).*

2 Algebraische Grundbegriffe

2.1 Gruppen und Gruppenhomomorphismen

Definition 2.1. *Sei M eine Menge. Eine (innere) Verknüpfung auf M ist eine Abbildung $* : M \times M \to M$, $(a,b) \mapsto a * b$.*

Beispiele 2.2.

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \ (a,b) \mapsto a + b,$$
$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \ (a,b) \mapsto a \cdot b.$$

Definition 2.3. *Eine Gruppe $(G, *)$ besteht aus einer Menge G und einer Verknüpfung $* : G \times G \to G$, $(a, b) \mapsto a * b$, mit folgenden Eigenschaften:*

(i) Die Verknüpfung $*$ ist assoziativ, d.h. $a * (b * c) = (a * b) * c$ für alle $a, b, c \in G$.

(ii) Es existiert ein neutrales Element $e \in G$, d.h. ein Element $e \in G$ mit $e*a = a*e = a$ für alle $a \in G$.

(iii) Für jedes $a \in G$ existiert ein inverses Element a', d.h. ein Element $a' \in G$ mit $a' * a = a * a' = e$.

Die Gruppe heißt abelsch (oder kommutativ), wenn für alle $a, b \in G$ gilt $a * b = b * a$.

Lemma 2.4. *Sei $(G, *)$ eine Gruppe. Dann gilt:*

(i) Es gibt genau ein neutrales Element in G.

(ii) Für jedes $a \in G$ gibt es ein eindeutig bestimmtes inverses Element.

Beweis. (i) Die Existenz folgt aus der Definition der Gruppe. Eindeutigkeit: Sind e und $\tilde{e}$ neutrale Elemente von G, dann folgt $e = e * \tilde{e} = \tilde{e}$ aus Definition 2.3 (ii).

(ii) Die Existenz folgt aus der Definition der Gruppe. Eindeutigkeit: Seien $b \in G$ und $\tilde{b} \in G$ inverse Elemente zu $a \in G$. Dann gilt

$$b \overset{2.3(ii)}{=} e * b \overset{2.3(iii)}{=} (\tilde{b} * a) * b \overset{2.3(i)}{=} \tilde{b} * (a * b) \overset{2.3(iii)}{=} \tilde{b} * e \overset{2.3(ii)}{=} \tilde{b}. \qquad \square$$

Bemerkung 2.5 (Quantoren)**.** *Vor allem in Beweisen, aber oft auch in Formulierungen von Aussagen, verwendet man die Quantoren $\forall$ ("für alle"), $\exists$ ("es existiert"), $\exists!$ ("es gibt genau ein") - häufig in Verbindung mit einem ":".*

Beispielsweise:

- Assoziativität Gruppe: $\forall\, a, b, c \in G$: $a * (b * c) = (a * b) * c$
- inverses Element: $\forall\, a \in G\ \exists!\, a' \in G$ mit $a' * a = a * a' = e$.

Beispiele 2.6.

(i) $(\mathbb{R}, +)$ ist eine abelsche Gruppe:

- *Assoziativgesetz: $a + (b + c) = (a + b) + c\ \forall\, a, b, c \in \mathbb{R}$*
- *neutrales Element: $e = 0$, denn $a + 0 = 0 + a = a\ \forall\, a \in \mathbb{R}$*
- *inverses Element: $a + (-a) = (-a) + a = 0\ \forall a \in \mathbb{R}$*
- *Kommutativgesetz: $a + b = b + a\ \forall\, a, b \in \mathbb{R}$.*

(ii) $(\mathbb{R} \setminus \{0\}, \cdot)$ *ist eine abelsche Gruppe*

- *Assoziativgesetz:* $a \cdot (b \cdot c) = (a \cdot b) \cdot c \; \forall\, a, b, c \in \mathbb{R} \setminus \{0\}$
- *neutrales Element:* $e = 1$*, denn* $a \cdot 1 = 1 \cdot a = a \; \forall\, a \in \mathbb{R} \setminus \{0\}$
- *inverses Element:* $a \cdot a^{-1} = a^{-1} \cdot a = 1 \; \forall a \in \mathbb{R} \setminus \{0\}$
- *Kommutativgesetz:* $a \cdot b = b \cdot a \; \forall\, a, b \in \mathbb{R} \setminus \{0\}$.

Lemma 2.7. $(\mathbb{R}, \cdot)$ *ist keine Gruppe.*

Beweis. Angenommen, $(\mathbb{R}, \cdot)$ sei eine Gruppe. Dann wäre 1 das neutrale Element, da $a \cdot 1 = 1 \cdot a = a \; \forall\, a \in \mathbb{R}$. Sei nun n' das inverse Element zu 0, dann folgt $n' \cdot 0 = 0 \cdot n' = 1$. Das ist aber falsch, d.h. wir erhalten einen Widerspruch. Deshalb ist $(\mathbb{R}, \cdot)$ keine Gruppe. □

Bemerkung 2.8 (Widerspruchsbeweis)**.** *Wir haben hier einen Widerspruchsbeweis geführt. Dabei nehmen wir an, dass eine zu beweisende Aussage falsch ist und zeigen, dass das zu einem Widerspruch führt. Hat man den Widerspruch hergeleitet, verwendet man häufig das Symbol* $\lightning$ *.*

Bemerkung 2.9 (Negation von Aussagen mit Quantoren)**.** *Aussagen mit Quantoren sind manchmal von der Form*

$$\forall\, x \; \exists\, y \; : \; \textit{Aussage } A(x, y) \textit{ gilt.}$$

Die Negation dieser Aussage ist:

$$\exists\, x \; \forall\, y \; : \; \textit{Aussage } A(x, y) \textit{ gilt nicht.}$$

Beispiel (inverses Element bei der Gruppe)**.** *Sei* G *eine Menge und* $* : G \times G \to G$ *eine Verknüpfung mit neutralem Element* $e \in G$*. (iii) aus Definition 2.3 bedeutet*

$$\forall\, a \in G \; \exists\, a' \in G \; : \; a' * a = e$$

Die Negation ist

$$\exists\, a \in G \; \forall\, a' \in G \; : \; a' * a \neq e.$$

Genau das haben wir in Lemma 2.7 für $(G, *) = (\mathbb{R}, \cdot)$ *gezeigt.*

Analog ist die Negation von

$$\exists\, x \; \forall\, y \; : \; \text{Aussage } A(x, y) \text{ gilt.}$$

$$\forall\, x \; \exists\, y \; : \; \text{Aussage } A(x, y) \text{ gilt nicht.}$$

Lemma 2.10 (Kürzungsregel)**.** *Sei $(G, *)$ eine Gruppe, $a, b, c \in G$. Gilt $a * b = a * c$ oder $b * a = c * a$, so folgt $b = c$.*

Beweis. Gelte $a * b = a * c$.

$$\begin{aligned} &\Rightarrow && a' * (a * b) = a' * (a * c) \\ &\Rightarrow && (\underbrace{a' * a}_{=e}) * b = (a' * a) * c \\ &\Rightarrow && b = c. \end{aligned}$$

Die Aussage für $b * a = c * a$ folgt analog. □

Lemma und Definition 2.11 (Symmetrische Gruppe)**.** *Sei M eine Menge und*

$$S(M) := \{f : M \to M | \, f \text{ ist bijektiv}\}.$$

(i) Dann gilt: $(S(M), \circ)$ ist eine Gruppe, die sogenannte symmetrische Gruppe.

(ii) Ist $M = \{1, \ldots, n\}$, so heißt

$$(S(\{1, \ldots, n\}), \circ) := (S_n, \circ)$$

symmetrische Gruppe auf n Ziffern. Die Elemente von S_n heißen Permutationen und werden mit π bezeichnet (d.h. $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$).

Beweis von (i). Die Verknüpfung $\circ$ ist als Abbildung von $S(M) \times S(M)$ nach $S(M)$ *wohldefiniert*, denn die Komposition bijektiver Abbildungen ist wieder bijektiv. Zudem gilt für $f, g, h \in S(M)$ das Assoziativgesetz $f \circ (g \circ h) = (f \circ g) \circ h$ nach Lemma 1.18; wegen $id_M \circ f = f \circ id_M = f \; \forall f \in S(M)$ ist id_M neutrales Element und für jedes $f \in S(M)$ ist die zugehörige Umkehrabbildung f^{-1} inverses Element: $f \circ f^{-1} = f^{-1} \circ f = id_M$. □

Bemerkung 2.12. *(i) Die Gruppe $(S(M); \circ)$ ist in der Regel nicht abelsch, da die Verknüpfung $f \circ g$ nicht kommutativ ist.*

(ii) Der Begriff "wohldefiniert" wird bei der Definition einer Abbildung $f : M \to N$ verwendet und bedeutet, dass die Zuordnung $m \mapsto f(m)$ eindeutig ist und alle Funktionswerte $f(m)$ im Wertebereich N der Abbildung liegen.

(iii) Die Elemente $\pi \in S_n$ kann man in der Form

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}$$

schreiben (Permutationsschreibweise).

Beispiele.

$$S_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

$$S_2 = \left\{ \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}_{=id_{\{1,2\}}}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$$

$$S_3 = \left\{ \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}}_{=id_{\{1,2,3\}}}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \right.$$
$$\left. \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}$$

Es gilt

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

d.h. $(S_3, \circ)$ *ist nicht abelsch.*

Bemerkung 2.13 (Zyklenschreibweise von S_n)**.**
Setze $\pi^j = \underbrace{\pi \circ \cdots \circ \pi}_{j-mal}$ *und* $\pi^{-j} = \underbrace{\pi^{-1} \circ \cdots \circ \pi^{-1}}_{j-mal}$*. Sei* $m \in \{1, \ldots, n\}$ *fest. Es gilt*

$$\left\{\pi(m), \pi^2(m), \pi^3(m), \ldots\right\} \subset \{1, \ldots, n\}.$$

$\Rightarrow \exists\, i, j, i \neq j,$ *mit* $\pi^i(m) = \pi^j(m)$*. Sei OE ("ohne Einschränkung")* $i > j$*.*

$$\Rightarrow \quad \underbrace{\pi^{-j}(\pi^i(m))}_{=\pi^{i-j}(m)} = \pi^{-j}(\pi^j(m)) = m$$

Sei nun $k \in \mathbb{N}$ *die kleineste natürliche Zahl mit* $\pi^k(m) = m$*. Wir stellen die Elemente* $m, \pi(m), \ldots, \pi^{k-1}(m)$ *in einem* Zyklus *dar:*

$$\underset{\longleftarrow}{\overset{\longrightarrow}{\Big(m}} \quad \overset{\longrightarrow}{\pi(m)} \quad \overset{\longrightarrow}{\pi^2(m)} \quad \overset{\longrightarrow}{\ldots} \quad \overset{\longrightarrow}{\pi^{k-1}(m)\Big)}$$

Die Elemente in einem Zyklus sind alle unterschiedlich (sonst wäre obiges k nicht das kleinste $k \in \mathbb{N}$ mit $\pi^k(m) = m$).

Beispiel:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 4 & 1 & 2 & 6 & 3 & 7 \end{pmatrix}$$
$$= (\ 1\ \ 5\ \ 6\ \ 3\ )(\ 2\ \ 4\ )(\ 7\ )$$

Die Elemente, die auf sich selbst abgebildet werden, lässt man dann noch weg und erhält

$$\pi = (\ 1\ \ 5\ \ 6\ \ 3\ )(\ 2\ \ 4\ ).$$

Konvention: $id = (\ )$.

Beispiel.

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \right.$$
$$\left. \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}$$
$$= \{(\ ), (\ 2\ \ 3\ ), (\ 1\ \ 2\ ), (\ 1\ \ 2\ \ 3\ ), (\ 1\ \ 3\ \ 2\ ), (\ 1\ \ 3\ )\}$$

Wir zeigen noch (mithilfe eines Widerspruchsbeweises), dass die Zyklen disjunkt sind:

Angenommen, die Zyklen sind nicht disjunkt.
$\Rightarrow \exists\, \bar{m} \notin \{m, \ldots, \pi^{k-1}(m)\}$ und $\exists\, i, j \in \mathbb{N}$ mit $\pi^i(\bar{m}) = \pi^j(m)$.
$\Rightarrow \underbrace{\pi^{-i}(\pi^i(\bar{m}))}_{=\bar{m}} = \pi^{-i+j}(m)$. $\Rightarrow \bar{m}$ ist Element des Zyklus' von m. $\lightning$
Der letzte Schluss gilt auch, wenn $j - i < 0$ ist, da $\pi^{-1}(m) = \pi^{k-1}(m)$ (mit $\pi^{-1}(m), \pi^{-2}(m)$, ... durchläuft man den Zyklus rückwärts).

Definition 2.14 (Gruppenhomomorphismus)**.** *Seien $(G, *)$, $(H, \circledast)$ Gruppen. Eine Abbildung $\varphi : G \to H$ heißt Gruppenhomomorphismus, wenn für alle $a, b \in G$ gilt*

$$\varphi(a * b) = \varphi(a) \circledast \varphi(b).$$

Ein Gruppenhomomorphismus heißt Gruppenisomorphismus, wenn er bijektiv ist.

Wir verwenden nachfolgend die Bezeichnung $\mathbb{R}_{>0} := \{x \in \mathbb{R} | \, x > 0\} = (0, \infty)$.

Beispiele 2.15. *(i) $\varphi : \mathbb{Z} \to \mathbb{Z}$, $a \mapsto 2a$, ist ein Gruppenhomomorphismus von $(\mathbb{Z}, +)$ nach $(\mathbb{Z}, +)$, denn $\varphi(a + b) = 2(a + b) = 2a + 2b = \varphi(a) + \varphi(b)$ $\forall\, a, b \in \mathbb{Z}$. φ ist aber kein Gruppenisomorphismus, denn φ ist nicht surjektiv ($1 \notin \varphi(\mathbb{Z})$).*

(ii) $\varphi : \mathbb{Z} \to \mathbb{Z}$, $a \mapsto a+1$, *ist kein Gruppenhomomorphismus von* $(\mathbb{Z}, +)$ *nach* $(\mathbb{Z}, +)$, *denn* $\varphi(2) = 3 \neq \varphi(1) + \varphi(1) = 4$.

(iii) $\varphi : \mathbb{R} \to \mathbb{R}_{>0}$, $x \mapsto e^x$, *ist ein Gruppenisomorphismus von* $(\mathbb{R}, +)$ *nach* $(\mathbb{R}_{>0}, \cdot)$. *Denn einerseits gilt* $\phi(x+y) = e^{x+y} = e^x \cdot e^y = \phi(x) \cdot \phi(y)$ $\forall\, x, y \in \mathbb{R}$, *andererseits ist* $\varphi : \mathbb{R} \to \mathbb{R}_{>0}$ *bijektiv.*

Lemma und Definition 2.16. *Seien* $(G, *)$ *und* $(H, \circledast)$ *Gruppen mit neutralen Elementen* $e_G \in G$ *bzw.* $e_H \in H$ *sowie* $\varphi : G \to H$ *ein Gruppenhomomorphismus. Dann gilt:*

(i) $\varphi(e_G) = e_H$.

(ii) Für alle $a \in G$ *ist* $\varphi(a') = \varphi(a)'$.

(iii) Ist φ *ein Gruppenisomorphismus, dann ist auch* $\varphi^{-1} : H \to G$ *ein Gruppenisomorphismus.*

$(G, *)$ *und* $(H, \circledast)$ *heißen* <u>isomorph</u> *(*$(G, *) \cong (H, \circledast)$*), wenn es einen Gruppenisomorphismus* $\varphi : G \to H$ *gibt.*

Beweis. (i) $e_H \circledast \varphi(e_G) = \varphi(e_G) = \varphi(e_G * e_G) = \varphi(e_G) \circledast \varphi(e_G) \overset{\text{Lemma } 2.10}{\Rightarrow} e_H = \varphi(e_G)$.

(ii) Es gilt $\varphi(a) \circledast \varphi(a') = \varphi(a * a') = \varphi(e_G) \overset{(i)}{=} e_H$, analog $\varphi(a') \circledast \varphi(a) = e_H$.
$\Rightarrow \varphi(a') = \varphi(a)'$.

(iii) φ ist bijektiv, d.h. φ^{-1} existiert und ist ebenfalls bijektiv. Zu zeigen: φ^{-1} ist Gruppenhomomorphismus.
Seien $c, d \in H$. Dann gilt

$$\begin{aligned}
\varphi^{-1}(c \circledast d) &= \varphi^{-1}\Big(\varphi\big(\varphi^{-1}(c)\big) \circledast \varphi\big(\varphi^{-1}(c)\big)\Big) \\
&= \varphi^{-1}\Big(\varphi\big(\varphi^{-1}(c) * \varphi^{-1}(d)\big)\Big) \qquad \text{(da } \varphi \text{ Gruppenhomomorphismus ist)} \\
&= \varphi^{-1}(c) * \varphi^{-1}(d).
\end{aligned}$$

□

Bemerkung. *Insbesondere gilt* $(\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \cdot)$ *nach Beispiel 2.15 (iii).*

Definition 2.17 (Untergruppe)**.** *Sei* $(G, *)$ *eine Gruppe und* $G' \subset G$ *eine nichtleere Teilmenge.* $(G', *)$ *heißt* <u>Untergruppe</u> *von* $(G, *)$*, wenn* $(G', *)$ *selbst eine Gruppe ist.*

In jeder Gruppe $(G, *)$ bildet die einelementige Teilmenge, die nur aus dem neutralen Element besteht, mit $*$ eine Untergruppe. Wir nennen sie die triviale Untergruppe. Ebenso ist natürlich die ganze Gruppe stets eine Untergruppe von sich selber.

Bemerkung 2.18. *Ist* $(G, *)$ *eine Gruppe und* $G' \subset G$ *eine nichtleere Teilmenge, so ist* $(G', *)$ *genau dann eine Untergruppe von* $(G, *)$*, wenn für beliebige* $a, b \in G'$ *auch* $a * b \in G'$ *sowie* $a' \in G'$ *gilt (Übungsblatt 4, Aufgabe 1 (i)).*

2.2 Ringe und Körper

Kurze Übersicht. Ringe und Körper sind Mengen mit zwei Verknüpfungen – einer Addition mit Gruppenstruktur und einer Multiplikation. Dabei gibt es im Ring bei der Multiplikation nicht zwangsläufig ein inverses Element der Multiplikation (wie in $\mathbb{Z}$, wo $3^{-1} = 1/3 \notin \mathbb{Z}$), während beim Körper $(K \setminus \{0\}, \cdot)$, d.h. die Menge K ohne das neutrale Element 0 der Addition zusammen mit der Multiplikation, auch eine Gruppe ist.

Definition 2.19 (Ring)**.** *Ein Ring ist ein Tripel $(R, +, \cdot)$ bestehend aus einer Menge R und zwei Verknüpfungen $+$ und $\cdot$,*

$$+ : R \times R \to R,\ (a, b) \mapsto a + b \quad \text{“Addition”}$$
$$\cdot : R \times R \to R,\ (a, b) \mapsto a \cdot b \quad \text{“Multiplikation”},$$

welche den folgenden drei Bedingungen genügen:

(i) *$(R, +)$ ist eine abelsche Gruppe. Das neutrale Element der Addition wird mit $0 = 0_R$ bezeichnet. Das inverse Element der Addition zu a wird mit $-a$ bezeichnet.*

(ii) *Die Multiplikation ist assoziativ, d.h. für alle $a, b, c \in R$ gilt*

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

(iii) *Es gelten die Distributivgesetze, d.h. für alle $a, b, c \in R$ gelten*

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und}$$
$$(b + c) \cdot a = b \cdot a + c \cdot a.$$

Ein Ring heißt kommutativ, wenn $a \cdot b = b \cdot a$ für alle $a, b \in R$. Ein Ring mit Einselement ist ein Ring, in dem ein Element $1 = 1_R$ existiert mit $1_R \cdot a = a \cdot 1_R = a$ für alle $a \in R$.

Ohne Klammerung gilt die Konvention “$\cdot$” vor “$+$”.

Bemerkung. *Ist aus dem Zusammenhang klar, welche Verknüpfungen gemeint sind, werden diese bei der Angabe von Gruppen oder Ringen oft weggelassen. Man schreibt dann bspw. kurz “Sei R ein Ring.” statt “Sei $(R, +, \cdot)$ ein Ring.”.*

Lemma 2.20. *Sei $(R, +, \cdot)$ ein Ring. Dann gilt:*

(i) *$0 \cdot a = a \cdot 0 = 0$ für alle $a \in R$ und*

(ii) *$a \cdot (-b) = -a \cdot b = (-a) \cdot b$ für alle $a, b \in R$.*

Beweis. (i) Da 0 das neutrale Element der Ringaddition ist, gilt

$$0 \cdot a + 0 = 0 \cdot a = (0+0) \cdot a \overset{2.19(iii)}{=} 0 \cdot a + 0 \cdot a \overset{\text{Lemma } 2.10}{\Rightarrow} 0 = 0 \cdot a.$$

(ii) Da $-a$ das Inverse der Ringaddition zu a ist, gilt

$$0 \overset{\text{Lemma } 2.20(i)}{=} 0 \cdot b = (a + (-a)) \cdot b \overset{2.19(iii)}{=} a \cdot b + (-a) \cdot b \;\Rightarrow\; -a \cdot b = (-a) \cdot b.$$

Analog zeigt man $-a \cdot b = a \cdot (-b)$. □

Beispiel 2.21. *$(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit Einselement. Ebenso sind $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ kommutative Ringe mit Einselement. In $\mathbb{Q} \setminus \{0\}$ und $\mathbb{R} \setminus \{0\}$ gibt es darüberhinaus auch jeweils das inverse Element der Multiplikation ($\rightarrow$ Körper (s.u.)).*

Restklassenringe. Wir wollen jetzt noch ein weiteres Beispiel diskutieren: Sei $m \in \mathbb{N}$ beliebig. Zu jedem $a \in \mathbb{Z}$ gibt es dann bekanntlich eindeutig bestimmte ganze Zahlen q und r mit

$$a = qm + r \quad \text{und} \quad 0 \leq r < m \quad \text{(Übungsblatt 4, Aufgabe 3)}.$$

Setze $r_m(a) := r$ (Rest von a bei Division durch m). Sei $F_m := \{0, 1, 2, \ldots, m-1\}$. Wir definieren eine Addition $+_m$ sowie eine Multiplikation $\cdot_m$ auf F_m durch

$$a +_m b := r_m(a + b) \tag{2.1}$$

$$a \cdot_m b := r_m(a \cdot b). \tag{2.2}$$

Lemma 2.22. *(i) Es gelten folgende Rechenregeln. Für $a, b \in F_m$ ist*

$$r_m(a+b) = r_m(r_m(a) + b) = r_m(a + r_m(b)) = r_m(r_m(a) + r_m(b))$$
$$r_m(a \cdot b) = r_m(r_m(a) \cdot b) = r_m(a \cdot r_m(b)) = r_m(r_m(a) \cdot r_m(b)).$$

(ii) $(F_m, +_m, \cdot_m)$ ist ein kommutativer Ring mit Einselement. Er wird als Restklassenring modulo m bezeichnet.

Beweis. Übungsblatt 4, Aufgabe 2. □

Definition 2.23. *Sei $(R, +, \cdot)$ ein Ring. Er heißt nullteilerfrei, wenn für alle $a, b \in R$ gilt: Aus $a \cdot b = 0_R$ folgt $a = 0_R$ oder $b = 0_R$.*

Satz 2.24. *Sei $m \in \mathbb{N}$, $m > 1$. Dann sind äquivalent:*

(i) $(F_m, +_m, \cdot_m)$ ist nullteilerfrei.

(ii) m ist eine Primzahl.

Beweis. (i) $\Rightarrow$ (*ii*): Wir zeigen die Umkehrung $\neg$(ii) $\Rightarrow \neg$(i) (dies ist äquivalent zur Implikation (i) $\Rightarrow$ (ii) – ein sogenannter indirekter Beweis). Sei m keine Primzahl. Dann gibt es $a, b \in \mathbb{N}$ mit $1 < a$, $b < m$ und $m = a \cdot b$. Es folgt $a \cdot_m b = r_m(a \cdot b) = r_m(m) = 0$, d.h. F_m ist nicht nullteilerfrei.

(ii) $\Rightarrow$ (i): Seien m eine Primzahl und $a, b \in F_m$ mit $a \cdot_m b = 0$. Zz: $a = 0$ oder $b = 0$. Wegen $a \cdot_m b = r_m(a \cdot b) = 0$ gibt es ein $q \in \mathbb{Z}$ mit $a \cdot b = q \cdot m$. $\overset{m \text{ Primzahl}}{\Rightarrow}$ m teilt a oder m teilt b. $\overset{a,b<m}{\Rightarrow}$ $a = 0$ oder $b = 0$. □

Definition 2.25 (Körper)**.** *Ein Körper ist ein kommutativer Ring $(K, +, \cdot)$ mit Einselement, in dem zusätzlich gilt: $(K \setminus \{0_K\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1_K. Damit besitzt jedes von 0_K veschiedene Element $a \in K$ ein Inverses bzgl. der Verknüpfung "$\cdot$", welches wir mit a^{-1} (oder auch $1/a$) bezeichnen.*

Bemerkung. *Per definitionem gilt $0_K \neq 1_K$.*

Beispiele 2.26. *(i) $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper.*

(ii) $(\mathbb{Z}, +, \cdot)$ ist kein Körper, da es zu $a \notin \{-1, 1\}$ kein inverses Element der Multiplikation gibt.

Satz 2.27. *Sei K ein Körper. Dann ist K nullteilerfrei.*

Beweis (durch Widerspruch). Angenommen, K ist nicht nullteilerfrei.
Dann gibt es $a, b \in K$ mit $a \neq 0$, $b \neq 0$ und $a \cdot b = 0$.
$\Rightarrow \exists\, a^{-1}$ mit $a \cdot a^{-1} = 1$ und $\exists\, b^{-1}$ mit $b \cdot b^{-1} = 1$.
$\Rightarrow \underbrace{a \cdot b}_{=0} \cdot a^{-1} \cdot b^{-1} \overset{\text{Kommutativität}}{=} a \cdot a^{-1} \cdot b \cdot b^{-1} = 1 \cdot 1 = 1$. ↯

□

Satz 2.28. *Sei $(F_m, +_m, \cdot_m)$ wie oben der Restklassenring modulo m. F_m ist genau dann ein Körper, wenn m eine Primzahl ist.*

Beweis.

"$\Rightarrow$": Sei F_m ein Körper. $\overset{\text{Satz } 2.27}{\Rightarrow}$ F_m ist nullteilerfrei. $\overset{\text{Satz } 2.24}{\Rightarrow}$ m ist eine Primzahl.

"$\Leftarrow$": Sei nun m eine Primzahl.

$\overset{\text{Lemma } 2.22 \text{ und Satz } 2.24}{\Rightarrow}$ F_m ist nullteilerfreier, kommutativer Ring mit Einselement.

Zz: F_m ist ein Körper. Dafür fehlt nur die Existenz des Inversen a^{-1} zu $a \neq 0$.

Sei nun $a \in F_m \setminus \{0\}$. Betrachte die Abbildung

$$\begin{aligned} F_m &\to F_m \\ x &\mapsto x \cdot_m a. \end{aligned}$$

Die Abbildung ist injektiv, da aus $x \cdot_m a = y \cdot_m a$ folgt, dass $(x-y) \cdot_m a = 0$ (nach dem Distributivgesetz und Lemma 2.20 (ii)) und somit wegen der Nullteilerfreiheit $x - y = 0$, d.h. $x = y$. Damit sind die Elemente $0 \cdot_m a, \ldots, (m-1) \cdot_m a$ alle verschieden. Da F_m nur m Elemente enthält, ist die Abbildung folglich surjektiv und es existiert insbesondere ein $x \in F_m$ mit $x \cdot_m a = 1$.

□

Bemerkung 2.29. *In unserem Beispiel besteht F_m aus den Zahlen $\{0, 1, \ldots, m-1\}$ und $+_m$ sowie $\cdot_m$ sind Verknüpfungen darauf. Der sogenannte Restklassenring F_m wird aber häufig anders eingeführt. Und zwar definiert man auf $\mathbb{Z}$ die Äquivalenzrelation $\sim_m$ durch*

$$x \sim_m y :\Leftrightarrow \; r_m(x) = r_m(y).$$

Die m-elementige Quotientenmenge $\mathbb{Z}/\sim_m$ wird mit $\mathbb{Z}/m\mathbb{Z}$ bezeichnet; ihre Elemente – die Äquivalenzklassen – sind Teilmengen ganzer Zahlen, die jeweils bei Division mit m denselben Rest besitzen. D.h. zwei Zahlen $x, y \in \mathbb{Z}$ liegen genau dann in derselben Äquivalenzklasse, wenn ihre Differenz $x-y$ durch m teilbar ist. Auf der Quotientenmenge $\mathbb{Z}/\sim_m$ definiert man nun die Verknüpfungen $+$ und $\cdot$ von Äquivalenzklassen mithilfe ihrer Repräsentanten durch $+_m$ aus (2.1) und $\cdot_m$ aus (2.2) (wobei hier die Wohldefiniertheit zu prüfen ist!) → Übungsblatt 4, Aufgabe 2. Da aber die m Elemente von F_m gerade Repräsentanten der m verschiedenen Äquivalenzklassen von $\mathbb{Z}/\sim_m$ sind, gelten alle obigen Ergebnisse für F_m entsprechend dann auch für $\mathbb{Z}/m\mathbb{Z}$. Für eine Primzahl p ist $\mathbb{Z}/p\mathbb{Z}$ nach Satz 2.28 ein Körper und heißt Primkörper der Charakteristik p.

2.2.1 Der Körper $\mathbb{C}$ der komplexen Zahlen

In $\mathbb{R}$ kann man Wurzeln aus allen positiven Zahlen ziehen, nicht jedoch aus negativen. Das bringt die Idee auf, die reellen Zahlen zu erweitern.

Definition 2.30. *Die komplexen Zahlen sind definiert als $\mathbb{C} = \mathbb{R}^2$ mit den Verknüpfungen*

$$(a,b) + (c,d) := (a+c, b+d) \quad \text{und} \tag{2.3}$$

$$(a,b) \cdot (c,d) := (ac - bd, ad + bc) \tag{2.4}$$

für alle $a, b, c, d \in \mathbb{R}$.

Satz 2.31. *$\mathbb{C}$ bildet mit der Addition (2.3) und der Multiplikation (2.4) einen Körper.*

Beweis. Dass $(\mathbb{C}, +)$ eine abelsche Gruppe mit neutralem Element $(0,0)$ ist, ist klar. Dass $\mathbb{C}\setminus\{(0,0)\}$ eine abelsche Gruppe bezüglich der Multiplikation ist, wurde auf Übungsblatt 3, Aufgabe 1c) gezeigt. Die Distributivgesetze rechnet man unmittelbar nach. □

Das neutrale Element der Addition ist $0_{\mathbb{C}} = (0,0)$, das neutrale Element der Multiplikation $1_{\mathbb{C}} = (1,0)$. Für (a,b) ist das additive Inverse $-(a,b) = (-a,-b)$, für $(a,b) \neq 0_{\mathbb{C}}$ das multiplikative Inverse

$$(a,b)^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right).$$

Die Abbildung $\mathbb{R} \to \mathbb{C}$, $a \mapsto (a,0)$, definiert einen Körperhomomorphismus (Übungsblatt 4, Aufgabe 4). Wir können folglich $\mathbb{R}$ mit komplexen Zahlen identifizieren, deren zweite Komponente gleich Null ist. Definieren wir schließlich $i = (0,1)$ (die sogenannte imaginäre Einheit), erhalten wir die Darstellung $(a,b) = (a,0) + b \cdot (0,1) = a + bi$ für alle $a, b \in \mathbb{R}$. Insbesondere können wir damit 0 und 1 für $0_{\mathbb{C}}$ und $1_{\mathbb{C}}$ schreiben. Für die imaginäre Einheit i gilt $i^2 = (0,1) \cdot (0,1) = (-1,0) = -1_{\mathbb{C}} = -1$, d.h. -1 besitzt nun eine Wurzel in $\mathbb{C}$.

2.2.2 Der Polynomring $R[t]$

Sei R ein Ring und t eine sogenannte Unbestimmte. Unter einem Polynom über R in der Unbestimmten t versteht man einen Ausdruck der Form

$$P(t) = a_0 + a_1 \cdot t + \cdots + a_n \cdot t^n$$

mit $a_0, \ldots, a_n \in R$ und $n \in \mathbb{N}_0$. Dabei kann als Unbestimmte t all das eingesetzt werden, für was die rechte Seite sinnvoll ist (insbesondere müssen Vielfache und Potenzen definiert sein). $a_0, a_1, \ldots, a_n$ heißen Koeffizienten des Polynoms. Die Menge dieser Polynome wird mit $R[t]$ bezeichnet. Oft lässt man bei einem Polynom die Angabe der Unbestimmten weg und schreibt kurz P. Sind alle Koeffizienten des Polynoms gleich Null, nennt man es Nullpolynom ($P = 0$). Der Grad des Polynoms P wird definiert als

$$deg(P) := \begin{cases} \infty & \text{falls } P = 0, \\ \max\{\nu \in \mathbb{N}_0 |\, a_\nu \neq 0\} & \text{sonst,} \end{cases}$$

wobei die Festlegung für den Grad des Nullpolynoms in der Literatur nicht einheitlich ist. Natürlich kann man für die Unbestimmte t ein Element von R selbst einsetzen. Ist $\lambda \in R$, so ist auch $P(\lambda) = a_0 + a_1 \cdot \lambda + \cdots + a_n \cdot \lambda^n \in R$. Damit erhält man eine Abbildung

$$\begin{aligned} \tilde{P}: \; & R \to R \\ & \lambda \mapsto P(\lambda). \end{aligned}$$

Warum man zwischen P und $\tilde{P}$ so penibel unterscheidet, demonstriert folgendes Beispiel.

Beispiel 2.32. *Sei $F_2 = \{0,1\}$ der Restklassenring modulo 2 und $P(t) = 1 \cdot_m t +_m 1 \cdot_m t^2$.*

Dann gelten

$$P(0) = 0 +_m 0 \cdot_m 0 \overset{\text{Lemma } 2.20}{=} 0 +_m 0 = 0,$$
$$P(1) = 1 +_m 1 \cdot_m 1 = 1 +_m 1 = 0.$$

Damit ist zwar $\tilde{P}$ die Nullabbildung, d.h. $\tilde{P}(x) = 0 \ \forall x \in F_2$, aber P ist nicht das Nullpolynom, d.h. $P \neq 0$.

Auf $R[t]$ kann man nun zwei Verknüpfungen, eine Addition und eine Multiplikation, einführen. Um die Notation zu vereinfachen, verwenden wir auch für die Polynomring-Addition und -Multiplikation die Zeichen $+$ und $\cdot$. Seien

$$P(t) = a_0 + a_1 \cdot t + \cdots + a_n \cdot t^n \quad \text{und} \quad Q(t) = b_0 + b_1 \cdot t + \cdots + b_m \cdot t^m$$

zwei Elemente aus $R[t]$. Ohne Einschränkung gelte für die nachfolgenden Definitionen $m = n$ (andernfalls ergänzen wir $b_{m+1} = \cdots = b_n = 0$ falls $m < n$ und $a_{n+1} = \ldots a_m = 0$ falls $m > n$). Wir definieren

$$P(t) + Q(t) := (a_0 + b_0) + (a_1 + b_1) \cdot t + \cdots + (a_n + b_n) \cdot t^n$$
$$P(t) \cdot Q(t) := c_0 + c_1 \cdot t + \cdots + c_{m+n} \cdot t^{n+m} \quad \text{mit} \quad c_k = \sum_{\substack{i,j \in \{0,\ldots,n\}: \\ i+j=k}} a_i \cdot b_j.$$

Satz und Definition 2.33. *Sei R ein Ring. Dann gelten folgende Aussagen:*

(i) *$(R[t], +, \cdot)$ ist ein Ring; er heißt Polynomring über R.*

(ii) *Ist R kommutativ, so auch $R[t]$.*

(iii) *Ist R nullteilerfrei, so gilt $deg(P \cdot Q) = deg(P) + deg(Q)$.*
(Hier gilt die Konvention $n + \infty = \infty + m = \infty + \infty = \infty \ \forall m, n \in \mathbb{N}_0$.)

Beweis. Übungsblatt 5, Aufgabe 1. □

In Analogie zur Division mit Rest bei ganzen Zahlen verfährt man beim Polynomring $K[t]$ über einem Körper K.

Satz 2.34. *Seien K ein Körper und $K[t]$ der Polynomring über K. Dann gibt es zu $P, Q \in K[t]$ eindeutig bestimmte Polynome $q, r \in K[t]$ mit folgenden Eigenschaften:*

(i) *$P = Q \cdot q + r$ sowie*

(ii) *$deg(r) < deg(Q)$, falls $r \neq 0$.*

Beweis. Eindeutigkeit: Seien $q, r, q', r' \in K[t]$ mit

$$P = Q \cdot q + r \quad \text{und } deg(r) < deg(Q), \text{ falls } r \neq 0$$
$$P = Q \cdot q' + r' \quad \text{und } deg(r') < deg(Q), \text{ falls } r' \neq 0.$$

Dann gilt nach Lemma 2.20 (ii) und dem Distributivgesetz $Q \cdot (q + (-q')) = (r' + (-r))$. Sind $r \neq 0$ und $r' \neq 0$, folgt

$$deg(r' + (-r)) \leq \max\big(deg(r), deg(-r')\big) < deg(Q) \text{ falls } r \neq r'.$$

Aber $q + (-q') \neq 0$ impliziert

$$deg(r' + (-r)) = deg\big(Q \cdot (q + (-q'))\big) = deg(Q) + deg\big(q + (-q')\big) \geq deg(Q),$$

was nicht sein kann, womit $q + (-q') = 0$ und damit $r' + (-r) = 0$.

Sind $r = 0$ und $r' \neq 0$, ist einerseits $deg(r') < deg(Q)$, andererseits folgt wie oben $deg(r') \geq deg(Q)$, was unmöglich ist, womit auch $r' = 0$. Aber mit $r = r' = 0$ folgt aus $\infty = deg(Q \cdot (q + (-q')) = deg(Q) \cdot deg(q + (-q'))$ auch $q - q' = 0$.

Existenz: Existiert ein $q \in K[t]$ mit $P = Q \cdot q$, folgt die Aussage mit $r = 0$. Andernfalls gilt $P + (-Q \cdot p) \neq 0$ für alle Polynome $p \in K[t]$, insbesondere $deg(P + (-Q \cdot p)) \geq 0$. Wir wählen nun ein $q \in K[t]$ mit der Eigenschaft

$$deg\big(P + (-Q \cdot q)\big) \leq deg\big(P + (-Q \cdot p)\big) \quad \text{für alle } p \in K[t].$$

Mit $r := P + (-Q \cdot q)$ gilt dann (i). Zu zeigen bleibt hierfür (ii), was wir nachfolgend durch Widerspruch beweisen. Angenommen, $deg(r) \geq deg(Q)$. Ist

$$Q = b_0 + b_1 \cdot t + \cdots + b_m \cdot t^m \quad \text{und} \quad r = c_0 + c_1 \cdot t + \cdots + c_k \cdot t^k$$

mit $b_m \neq 0$ und $c_k \neq 0$ für $k \geq m$, definieren wir $p := q + \frac{c_k}{b_m} \cdot t^{k-m}$. Aber dann folgt

$$r + \Big(- Q \cdot \Big[\frac{c_k}{b_m} \cdot t^{k-m}\Big]\Big) = P + \Big(- Q \cdot q + Q \cdot \Big[\frac{c_k}{b_m} \cdot t^{k-m}\Big]\Big) = P + (-Q \cdot p).$$

Da r und $Q \cdot [\frac{c_k}{b_m} \cdot t^{k-m}]$ denselben höchsten (von Null verschiedenen) Koeffizienten haben, ergibt sich weiter

$$deg\Big(r + \Big(- Q \cdot \Big[\frac{c_k}{b_m} \cdot t^{k-m}\Big]\Big)\Big) < deg(r),$$

also $deg(P + (-Q \cdot p)) < deg(r)$. ↯ □

Nach diesen Überlegungen kommen wir nun zu einer wesentlichen Frage – nämlich der nach der Existenz von Nullstellen.

Das nächste Lemma zeigt, dass man bei einem Polynom den Linearfaktor $(t - \lambda)$ abspalten kann, wenn λ eine Nullstelle ist.

Lemma 2.35. *Ist $\lambda \in K$ eine Nullstelle von $P \in K[t]$, $P \neq 0$, so existiert ein eindeutiges Polynom $Q \in K[t]$ mit*

$$P = (t - \lambda) \cdot Q \quad \text{und} \quad deg(Q) = deg(P) - 1.$$

Beweis. Nach Satz 2.34 gibt es eindeutig bestimmte $Q, r \in K[t]$ mit $P = (t - \lambda) \cdot Q + r$ und $deg(r) < deg(t - \lambda) = 1$ falls $r \neq 0$. Damit ist $r = a_0$ mit $a_0 \in K$. Wegen $P(\lambda) = 0$ folgt

$$0 = P(\lambda) = (\lambda - \lambda) \cdot Q(\lambda) + a_0 = a_0,$$

d.h. $r = 0$ und $P = (t - \lambda) \cdot Q$ ist gezeigt. $deg(Q) = deg(P) - 1$ folgt dann aus

$$deg(P) = deg((t - \lambda) \cdot Q) = deg(t - \lambda) + deg(Q) = 1 + deg(Q). \qquad \square$$

Korollar 2.36. *Seien K ein Körper, $P \in K[t]$ ein Polynom und k die Anzahl der Nullstellen von P. Ist $P \neq 0$, so gilt $k \leq deg(P)$.*

Beweis. Wir beweisen die Aussage durch vollständige Induktion nach dem Grad des Polynoms.

Induktionsanfang: Ist $deg(P) = 0$, so ist $P = a_0 \neq 0$ ein konstantes Polynom. Dieses hat aber keine Nullstelle, also ist die Behauptung korrekt.

Induktionsschritt: Sei die Aussage für Polynome $Q \in K[t]$ mit $deg(Q) \leq n - 1$, $n \in \mathbb{N}$, bereits bewiesen. Sei $P \in K[t]$ mit $deg(P) = n$. Besitzt P keine Nullstelle, so ist die Behauptung richtig. Ist andernfalls $\lambda \in K$ Nullstelle von P, so existiert nach Lemma 2.35 $Q \in K[t]$ mit $P = (t-\lambda)\cdot Q$ und $deg(Q) = n-1$. Alle von λ verschiedenen Nullstellen von P müssen folglich auch welche von Q sein. Nach Induktionsannahme besitzt Q aber höchstens $n - 1$ Nullstellen, womit $k \leq (n - 1) + 1 = n$ ist. $\square$

Über dem Körper $K = \mathbb{R}$ gibt es Polynome P mit $deg(P) > 0$, die keine Nullstelle besitzen.

Beispiel 2.37. *Sei $K = \mathbb{R}$ und $P(t) = t^2 + 1$ für $t \in \mathbb{R}$, so ist $P(t) \geq 1$ für alle $t \in \mathbb{R}$ und P besitzt insbesondere in $\mathbb{R}$ keine Nullstelle.*

Über dem Körper $\mathbb{C}$ gibt es ein solches Beispiel jedoch nicht.

Satz 2.38 (Fundamentalsatz der Algebra)**.** *Jedes Polynom $P \in \mathbb{C}[t]$ mit $deg(P) > 0$ hat mindestens eine Nullstelle.*

Beweis. Später. $\square$

3 Vektorräume

Hier wird noch eine Zeichnung eingefügt.

Definition 3.1. *Sei K ein Körper. Eine Menge V mit einer Addition*

$$+ : V \times V \to V,\ (v, w) \mapsto v + w$$

und einer skalaren Multiplikation

$$\cdot : K \times V \to V,\ (\lambda, v) \mapsto \lambda \cdot v$$

heißt <u>K-Vektorraum</u> (K-VR), falls Folgendes gilt:

(i) $(V,)$ *ist eine abelsche Gruppe. Das neutrale Element wird mit* $0 = 0_V$*, das zu* $v \in V$ *inverse Element mit* $-v$ *bezeichnet.*

(ii) Für die skalare Multiplikation gelten folgende Axiome: Für alle $\lambda, \mu \in K$ *und alle* $v, w \in V$ *gilt*

$$\begin{aligned}
&(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v \\
&\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w \\
&\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v \\
&1 \cdot v = v.
\end{aligned}$$

Konventionen: "+" vor "·", $\lambda v := \lambda \cdot v$.

Bemerkung. *Es ist wichtig, zwischen skalarer Multiplikation und Multiplikation in K sowie zwischen Addition in V und Addition in K zu unterscheiden.*

Beispiele 3.2. *(i) K Körper, $n \in \mathbb{N}$, $K^n = \{(x_1, \ldots, x_n) |\ x_1, \ldots, x_n \in K\}$ mit*

Addition $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ *und*

Skalarmultiplikation $\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$

ist ein Vektorraum über K (der sog. <u>Standard-Vektorraum</u>). Es gelten $0 = (0, \ldots, 0)$ *und* $-(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n)$.

[Die Axiome aus der obigen Definition muss man nachrechnen.]

(ii) Seien M eine Menge und K ein Körper. Wir definieren

$$\begin{aligned}
Abb(M, K) :=& \{f |\ f \text{ ist eine Abbildung } f : M \to K\} \\
=& \{f : M \to K\} \text{ (Kurzschreibweise)}
\end{aligned}$$

$V = Abb(M, K)$ wird mit folgenden Verknüpfungen zu einem Vektorraum:

$$(f+g)(x) := f(x) + g(x) \ \forall\, x \in M \quad \textit{(Addition)}$$
$$(\lambda \cdot f)(x) := \lambda \cdot f(x) \ \forall\, x \in M \quad \textit{(Skalarmultiplikation).}$$

Bemerkung 3.3 (Notation). *(i) Vektoren (die Elemente) eines Vektorraums V werden oft besonders gekennzeichnet, bspw. mit einem Pfeil $\vec{v}$ oder fett gedruckt $\mathbf{v}$. Wir verzichten auf weitere Kennzeichnungen und verwenden vor allem die Buchstaben u, v, w.*

(ii) Vektoren (vor allem) aus dem K^n werden oft als Spalten geschrieben, d.h.

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \textit{anstelle der Zeilenform} \quad (x_1, \ldots, x_n).$$

Lemma 3.4. *Sei K ein Körper, V ein K-VR. Dann gilt:*

(i) $0_K \cdot v = 0_V \ \forall\, v \in V$

(ii) $\lambda \cdot 0_V = 0_V \ \forall\, \lambda \in K$

(iii) $\lambda \cdot v = 0 \Rightarrow \lambda = 0_K$ *oder* $v = 0_V$

(iv) $(-1) \cdot v = -v \ \forall\, v \in V$.

Beweis. (i) und (ii) folgen mit der Kürzungsregel (Lemma 2.10), (iv) ist einfach. (iii): Seien $\lambda \in K, v \in V$ mit $\lambda \cdot v = 0$. Ist $\lambda \neq 0_K$, so folgt $v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0_V \overset{(ii)}{=} 0_V$. □

3.1 Untervektorräume und lineare Hülle

Definition 3.5. *Seien K ein Körper und V ein K-Vektorraum. $W \subset V$ mit $W \neq \emptyset$ heißt Untervektorraum (kurz UVR) von V, falls W mit den eingeschränkten Verknüpfungen selbst ein Vektorraum ist.*

Lemma 3.6. *Eine Teilmenge $W \neq \emptyset$ eines K-VRs V ist genau dann ein Untervektorraum von V, falls gilt:*

(i) $u, v \in W \Rightarrow u + v \in W$
(d.h. W ist abgeschlossen unter der Addition)

(ii) $\lambda \in K,\ w \in W \Rightarrow \lambda \cdot w \in W$
(d.h. W ist abgeschlossen unter der skalaren Multiplikation).

Beweis. "⇒": Ist W selbst ein K-VR, so gelten natürlich die beiden Bedingungen (i) und (ii).

"⇐": Wir nehmen an, dass (i) und (ii) gelten. Für die Aussage "W ist K-VR" müssen wir die Existenz des inversen Elements und des neutralen Elements der Addition zeigen:

- Sei $w \in W$. $\Rightarrow -w \overset{\text{Lemma } 3.4(iv)}{=} (-1) \cdot w \in W$ nach (ii).
- Sei w ein beliebiges Element aus $W \Rightarrow -w \in W \Rightarrow 0_V = w + (-w) \in W$ nach (i). 0_V ist aber auch das neutrale Element in W, da ja $w + 0_V = w$. $0_V + w = w$ $\forall\, w \in W$.

□

Beispiele 3.7. *Was sind Untervektorräume von $V = \mathbb{R}^2$?*

(i) $W = \{(0,0)\}$,

(ii) $W = V = \mathbb{R}^2$,

(iii) alle Geraden durch den Ursprung.

Beweis von (iii). Bis auf die senkrechte Gerade kann man alle Geraden durch den Ursprung als Menge

$$W = \{(x, mx) |\, x \in \mathbb{R}\}$$

schreiben mit $m \in \mathbb{R}$. Hier gilt für alle $v = (x, mx)$ und $w = (y, my) \in W$:

$(x, mx) + (y, my) = (x + y, m(x + y)) \in W$ sowie

$\lambda \cdot (x, mx) = (\lambda x, m \cdot (\lambda x)) \in W$, womit nach Lemma 3.6 W UVR von $\mathbb{R}^2$ ist.

Für die senkrechte Gerade durch den Ursprung $\{(0, y) | y \in \mathbb{R}\}$ geht der Beweis analog.

□

Der nächste Satz sagt, dass $\mathbb{R}^2$ neben diesen Beispielen keine weiteren Untervektorräume besitzt.

Satz 3.8. *Sei $V = \mathbb{R}^2$ und $W \subset V$ mit $W \neq \emptyset$. Dann ist W ein UVR von V genau dann, wenn $W = \{(0,0\}$, $W = \mathbb{R}^2$ oder W eine Gerade durch den Ursprung ist, d.h. entweder $W = \{(x, \lambda_0 x) |\, x \in \mathbb{R}\}$ für ein $\lambda_0 \in \mathbb{R}$ oder $W = \{(0, y) |\, y \in \mathbb{R}\}$.*

Beweis. "⇐": Haben wir gezeigt.

"⇒": Sei W ein anderer Vektorraum als $W = \{(0,0)\}$, $W = \mathbb{R}^2$ oder $W = \{(0, y) |\, y \in \mathbb{R}\}$ und $(w_2^*, w_2^*) \in W$ ein Punkt mit $w_1^* \neq 0$. Ein solcher muss existieren, denn außer $\{(0,0)\}$ bildet keine echte Teilmenge von $\{(0, y) |\, y \in \mathbb{R}\}$ einen UVR. Setze $\lambda_0 := w_2^*/w_1^*$.

Wir zeigen: $W = \{(x, \lambda_0 x) |\, x \in \mathbb{R}\}$.

"$\supset$": Sei $x \in \mathbb{R}$.

$$\Rightarrow \qquad (x, \lambda_0 x) = \left(x, \frac{w_2^*}{w_1^*} x\right) = \frac{x}{w_1^*}(w_1^*, w_2^*) \in W,$$

da W ein UVR ist.

"$\subset$": Sei nun (w_1, w_2) ein beliebiges anderes Element aus W. Wir müssen zeigen, dass (w_1, w_2) auf der Geraden liegt, d.h., dass $w_2 = \lambda_0 w_1$ gilt.

Widerspruchsbeweis:

Annahme: Gelte $w_2 \neq \lambda_0 w_1$. Wir zeigen, dass dann bereits $W = \mathbb{R}^2$ gelten muss, d.h., dass dann jedes $(z_1, z_2) \in \mathbb{R}^2$ in W liegt.

Sei dafür $(z_1, z_2) \in \mathbb{R}^2$ beliebig. Wir setzen

$$a = \frac{\lambda_0 z_1 - z_2}{w_1 \lambda_0 - w_2}, \quad \text{und} \quad b = \frac{w_1 z_2 - w_2 z_1}{w_1 \lambda_0 - w_2}$$

(der Nenner ist $\neq 0$ wegen $w_2 \neq \lambda_0 w_1$). Damit ist aber

$$(z_1, z_2) = \underbrace{a \cdot (w_1, w_2)}_{\in W} + \underbrace{b \cdot (1, \lambda_0)}_{\in W},$$

also $(z_1, z_2) \in W$. $\lightning$ Somit gilt $w_2 = \lambda_0 w_1$. □

Bemerkung 3.9. *Die Frage "Wie kommt man auf obiges a und b?" ist für die Gültigkeit des Beweises nicht relevant – wohl aber für das "Finden" des Beweises: Man muss dafür das Gleichungssystem*

$$z_1 = a w_1 + b$$
$$z_2 = a w_2 + b \lambda_0$$

mit a, b unbekannt bei bekannten w_1, w_2, z_1, z_2 lösen (z. Bsp. indem man $b = z_1 - a w_1$ in die zweite Gleichung einsetzt).

Bemerkung 3.10. *Für Abbildungen $f : \mathbb{R} \to \mathbb{R}$ haben wir damit gezeigt:*

$$\{(x, f(x)) \,|\, x \in \mathbb{R}\} \quad \textit{ist UVR von } \mathbb{R}^2 \;\Leftrightarrow\; f(x) = \lambda_0 x \quad \textit{für ein } \lambda_0 \in \mathbb{R}.$$

Satz 3.11. *Sei K ein Körper, V ein K-VR, I eine Indexmenge und $(W_i)_{i \in I}$ eine Familie von Untervektorräumen von V (d.h. für jedes $i \in I$ ist W_i ein UVR von V). Dann gilt*

$$W = \bigcap_{i \in I} W_i = \{w \in V \,|\, w \in W_i \textit{ für alle } i \in I\}$$

ist ein UVR von V.
("Beliebige Durchschnitte von Untervektorräumen sind wieder Untervektorräume").

Beweis. $0_V \in W$, d.h. $W \neq \emptyset$. Wir weisen (i) und (ii) aus Lemma 3.6 nach.

(i) Seien $v, w \in W$.
$\Rightarrow v, w \in W_i \; \forall\, i \in I$
$\Rightarrow u + v \in W_i \; \forall\, i \in I$
$\Rightarrow v + w \in \cap_{i \in I} W_i = W$.

(ii) Seien $\lambda \in K$, $v \in W$.
$\Rightarrow v \in W_i \; \forall\, i \in I$
$\Rightarrow \lambda v \in W_i \; \forall\, i \in I$
$\Rightarrow \lambda v \in \cap_{i \in I} W_i = W$. □

Um die Notation weiter zu vereinfachen, lassen wir nachfolgend (wie gerade bereits im Beweis) den Punkt $\cdot$ bei der Skalarmultiplikation meistens weg.

Beispiel 3.12. *Die Vereinigung von UVRs ist im Allgemeinen kein UVR. Man betrachte z.Bsp.* $K = \mathbb{R}$, $V = \mathbb{R}^2$

$$W_1 = \{(x_1, x_2) \in \mathbb{R}^2 | \, x_1 = x_2\}$$
$$W_2 = \{(x_1, x_2) \in \mathbb{R}^2 | \, x_1 = 0\}.$$

W_1 und W_2 sind UVRs. Aber $W_1 \cup W_2$ ist kein UVR, da bspw. $(1,1) \in W_1 \subset W_1 \cup W_2$ *und* $(0,1) \in W_2 \subset W_1 \cup W_2$, *aber* $(1,1) + (0,1) = (1,2) \notin W_1 \cup W_2$.

Dass $W_1 \cup W_2$ kein UVR sein kann, folgt natürlich auch bereits aus Satz 3.8.

Definition 3.13. *Seien K ein Körper und V ein K-VR.*

(i) Seien $r \in \mathbb{N}$, $v_1, \ldots, v_r \in V$ und $a_1, \ldots, a_r \in K$. Der Ausdruck

$$v = a_1 v_1 + \cdots + a_r v_r$$

heißt Linearkombination von $v_1, \ldots, v_r$. Es gilt $v \in V$.

(ii) Ist $M = \{v_1, \ldots, v_r\} \subset V$ für ein $r \in \mathbb{N}$, so heißt

$$\mathrm{Lin}(M) := \{a_1 v_1 + \cdots + a_r v_r | \, a_1, \ldots, a_r \in K\}$$

lineare Hülle von M.

(iii) Ist $M \subset V$ beliebig, $M \neq \emptyset$, so heißt

$$\mathrm{Lin}(M) := \bigcup_{\substack{L \subset M: \\ L \text{ endlich}}} \mathrm{Lin}(L)$$

die lineare Hülle von M. Wir setzen $\mathrm{Lin}(\emptyset) = \{0\}$.

Bemerkung 3.14. *Für unendliche Mengen M (d.h. Mengen M mit mehr als endlich vielen Elementen) gilt damit:*

$$v \in \mathrm{Lin}(M)$$

$$\Leftrightarrow$$

$\exists$ *endliche Teilmenge* $\{v_1, \ldots, v_s\} \subset M$ *und* $\exists\, a_1, \ldots, a_s \in K$ *mit* $v = a_1 v_1 + \cdots + A_s v_s$.

D.h. die lineare Hülle von unendlich vielen Vektoren besteht aus allen Linearkombinationen von je endlich vielen.

Beispiele 3.15. *(i) In Satz 3.8 haben wir die Gerade* $W = \{(x, \lambda_0 x) | x \in \mathbb{R}\}$ *betrachtet und dann im Beweis den Vektor* $(1, \lambda_0) \in W$. *Wegen* $(x, \lambda_0 x) = x(1, \lambda_0)$ *gilt* $W = \mathrm{Lin}(\{(1, \lambda_0)\})$.

(ii) Im Beweis von Satz 3.8 wurde dann gezeigt, dass jeder beliebige Vektor $(z_1, z_2) \in \mathbb{R}^2$ *als Linearkombination der beiden Vektoren* (w_1, w_2) *und* $(1, \lambda_0)$ *dargestellt werden kann, d.h. dass gilt*

$$\mathrm{Lin}(\{(1, \lambda_0), (w_1, w_2)\}) = \mathbb{R}^2.$$

Voraussetzung war, dass (w_1, w_2) *nicht auf der Geraden* $\{(x, \lambda_0 x) | x \in \mathbb{R}\}$ *liegt* ($w_2 \neq \lambda_0 w_1$), *d.h. dass* $(w_1, w_2) \notin \mathrm{Lin}(\{(1, \lambda_0)\})$.

(iii) Wegen $(x_1, x_2) = x_1(1, 0) + x_2(0, 1)\ \forall\, x_1, x_2 \in \mathbb{R}$ *gilt* $\mathrm{Lin}(\{(1, 0), (0, 1)\}) = \mathbb{R}^2$.

(iv) Im $\mathbb{R}^n$ *gilt*

$$(x_1, \ldots, x_n) = x_1 \underbrace{(1, 0, \ldots, 0)}_{=:e_1} + x_2 \underbrace{(0, 1, 0, \ldots, 0)}_{=:e_2} + \cdots + x_n \underbrace{(0, \ldots, 0, 1)}_{=:e_n}.$$

Hier ist $e_i \in \mathbb{R}^n$ *für* $i \in \{1, \ldots, n\}$ *derjenige Vektor, dessen i-ter Eintrag eine* 1 *ist und der sonst überall den Eintrag Null hat.* $e_1, \ldots, e_n$ *heißen die Einheitsvektoren im* $\mathbb{R}^n$. *Damit ist* $\mathbb{R}^n = \mathrm{Lin}(\{e_1, \ldots, e_n\})$.

Satz 3.16. *Sei K ein Körper, V ein K-VR und $M \subset V$. Dann gilt:*

(i) $\mathrm{Lin}(M)$ *ist ein UVR von V.*

(ii) Ist W ein UVR von V mit $M \subset W$, dann gilt $\mathrm{Lin}(M) \subset W$.

(iii) Es gilt

$$\mathrm{Lin}(M) = \bigcap_{\substack{W \text{ UVR von } V \\ \text{mit } M \subset W}} W,$$

d.h. $\mathrm{Lin}(M)$ *ist der kleinste UVR von V, der alle Vektoren v aus M enthält.*

Beweis. (i) Heuristisch (und auch inhaltlich) ist der Beweis klar: Wenn man zwei Linearkombinationen addiert bzw. mit einem Skalar multipliziert, erhält man wieder eine Linearkombination. Das Problem ist, den Beweis formal korrekt aufzuschreiben:

Wir wollen Lemma 3.6 verwenden, müssen also entsprechend die Voraussetzungen überprüfen. Es gilt

- $\mathrm{Lin}(M) \neq \emptyset$.
- Seien $u_1, u_2 \in \mathrm{Lin}(M)$.

  $\Rightarrow$ Es existieren endliche Mengen $L_1, L_2 \subset M$ mit $u_i \in \mathrm{Lin}(L_i)$, $i = 1, 2$.

  Mit $L^* = L_1 \cup L_2$ gilt $u_1.u_2 \in \mathrm{Lin}(L^*)$. Sei nun $L^* = \{v_1, \ldots, v_r\}$ für ein $r \in \mathbb{N}$.

  $\Rightarrow u_1 = \alpha_1 v_1 + \cdots + \alpha_r v_r$ mit $\alpha_i \in K$
  $u_2 = \beta_1 v_1 + \cdots + \beta_r v_r$ mit $\beta_i \in K$

  $\Rightarrow u_1 + u_2 = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_n + \beta_n)v_n \in \mathrm{Lin}(L^*) \subset \mathrm{Lin}(M)$.
- Seien $\lambda \in K$, $u \in \mathrm{Lin}(M)$, d.h. es existiert eine endliche Menge $L = \{v_1, \ldots, v_r\}$, $L \subset M$, mit $u = \alpha_1 v_1 + \cdots + \alpha_r v_r$ für $\alpha_1, \ldots, \alpha_r \in K$.

  $\Rightarrow \lambda u = (\lambda\alpha_1)v_1 + \cdots + (\lambda\alpha_r)v_r \in \mathrm{Lin}(L) \subset \mathrm{Lin}(M)$.

(ii) Sei $W \subset V$ ein UVR mit $M \subset W$. Es gilt

$$\mathrm{Lin}(M) = \bigcup_{\substack{L \subset M: \\ L \text{ endlich}}} \mathrm{Lin}(L),$$

d.h. wir müssen zeigen: $\mathrm{Lin}(L) \subset W$ für alle endlichen Teilmengen L von M. Aber für endliches $L = \{v_1, \ldots, v_r\}$ ist

$$\mathrm{Lin}(L) = \{\alpha_1 v_1 + \cdots + \alpha_r v_r | \alpha_1, \ldots, \alpha_r \in K\} \in W,$$

da W Vektorraum ist und alle $v_i \in L \subset M \subset W$, $i = 1, \ldots, r$.

(iii) Nach (i) ist $\mathrm{Lin}(M)$ ein UVR von V und damit einer der UVRs, über die der Durchschnitt gebildet wird, also

$$\bigcap_{\substack{W \text{ UVR von } V \\ \text{mit } M \subset W}} W \subset \mathrm{Lin}(M).$$

Andererseits folgt für alle $M \subset W$ nach (ii) auch $\mathrm{Lin}(M) \subset W$, womit

$$\mathrm{Lin}(M) \subset \bigcap_{\substack{W \text{ UVR von } V \\ \text{mit } M \subset W}} W.$$

□

Wir wollen abschließend noch einige Eigenschaften der linearen Hülle festhalten.

Lemma 3.17. *Seien K ein Körper, V ein K-Vektorraum, $M \subset V$. Dann gelten folgende Aussagen:*

(i) $M \subset \mathrm{Lin}(M)$.

(ii) $M \subset M' \subset V \Rightarrow Lin(M) \subset \mathrm{Lin}(M')$.

(iii) $M = \mathrm{Lin}(M) \Leftrightarrow M$ *ist UVR von* V.

(iv) $\mathrm{Lin}(\mathrm{Lin}(M)) = \mathrm{Lin}(M)$.

Beweis. (i) folgt unmittelbar aus der Definition der linearen Hülle.

(ii) $M \subset M' \Rightarrow M \subset \mathrm{Lin}(M')$. $\mathrm{Lin}(M')$ ist damit einer der UVRs W in der Darstellung

$$\mathrm{Lin}(M) = \bigcap_{\substack{W \text{ UVR von } V \\ \text{mit } M \subset W}} W$$

aus Satz 3.16 (iii). $\Rightarrow \mathrm{Lin}(M) \subset \mathrm{Lin}(M')$.

(iii) "$\Rightarrow$": Klar. "$\Leftarrow$": Da M ein UVR ist, ist er einer der Ws im Durchschnitt aus Satz 3.16 (iii). $\Rightarrow M = \mathrm{Lin}(M)$.

(iv) Es gilt $\mathrm{Lin}(M) \subset \mathrm{Lin}(\mathrm{Lin}(M))$. Außerdem ist $\mathrm{Lin}(M)$ einer der Vektorräume W bei der $\cap$-Bildung von $\mathrm{Lin}(\mathrm{Lin}(M))$ aus Satz 3.16 (iii). $\Rightarrow$ Behauptung.
Alternativ: Nach Satz 3.16 (i) ist $\mathrm{Lin}(M)$ UVR und nach Satz 3.16 (iii) ist $\mathrm{Lin}(\mathrm{Lin}(M))$ der kleinste UVR, der $\mathrm{Lin}(M)$ enthält. □

3.2 Lineare Unabhängigkeit, Basis und Dimension

Notation 3.18 (Summenzeichen)**.** *Wir verwenden ab jetzt das Symbol "$\sum$" als Abkürzung für eine Summe von endlich vielen Elementen, z. Bsp.*

$$\sum_{i=1}^{n} x_i = x_1 + \cdots + x_n$$

oder für $I = \{i_1, \ldots, i_r\} \subset \mathbb{N}$ (Indexmenge)

$$\sum_{i \in I} y_i = y_{i_1} + \cdots + y_{i_r}.$$

Nimmt man aus einer endlichen Indexmenge I ein Element i_0 heraus und summiert über $I \setminus \{i_0\}$, kennzeichnet man dies häufig mit $\sum_{i \neq i_0}$ anstelle von $\sum_{i \in I \setminus \{i_0\}}$.

Bemerkung 3.19. *Bei Vektorräumen ist es üblich, von einer "Familie" oder einem "System" $(v_1, \dots, v_r)$ von Vektoren zu sprechen. Für eine beliebige Indexmenge I ist eine Familie $(v_i)_{i\in I}$ formal gegeben durch eine Abbildung $I \to V$, $i \mapsto v_i$. Eine Familie muss nicht endlich sein – grundsätzlich studieren wir auch Familien mit unendlich (abzählbar) vielen Vektoren $(v_1, v_2, \dots)$ oder allgemeiner $(v_i)_{i\in I}$ mit einer beliebigen Indexmenge I. Im Spezialfall $I = \mathbb{N}$ nennt man eine Familie "Folge" – der Begriff ist aus der Analysis I bereits bekannt. Im Gegensatz zur Menge der Mitglieder der Familie $\{v_i | i \in I\}$ ist bei einer Familie die Zuordnung $i \mapsto v_i$ fest, also insbesondere $(v_1, \dots, v_r) \neq (v_{\pi(1)}, \dots, v_{\pi(r)})$ für jede Permutation $\pi \in S_r \setminus \{id\}$ (sofern die v_is alle verschieden sind).*

Definition 3.20. *Sei V ein K-Vektorraum.*

(i) Eine endliche Familie von Vektoren $(v_1, \dots, v_r)$ heißt linear unabhängig, falls aus

$$\sum_{i=1}^{r} \lambda_i v_i = 0 \quad \text{mit } \lambda_1, \dots, \lambda_r \in K$$

folgt $\lambda_1 = \cdots = \lambda_r = 0$.

(ii) Eine unendliche Familie von Vektoren $(v_i)_{i\in I}$ heißt linear unabhängig, wenn für jede endliche Teilmenge $J \subset I$ die Familie $(v_i)_{i\in J}$ linear unabhängig ist.

(iii) Eine Familie von Vektoren heißt linear abhängig, falls sie nicht linear unabhängig ist.

Zum besseren Verständnis halten wir fest:

Lemma 3.21. *Ist $r \geq 2$, so gilt: $(v_1, \dots, v_r)$ ist linear abhängig*

$$\Leftrightarrow$$

$\exists\, i_0 \in \{1, \dots, r\}$, so dass v_{i_0} eine Linearkombination aus $\{v_j | \, j \in \{1, \dots, r\} \setminus \{i_0\}\}$ ist.

Beweis. "$\Rightarrow$": $(v_1, \dots, v_r)$ linear abhängig $\Rightarrow \exists\ \lambda_1, \dots, \lambda_r \in K$ (nicht alle $= 0$) mit

$$\sum_{i=1}^{r} \lambda_i v_i = 0.$$

Ist nun aber i_0 ein Index mit $\lambda_{i_0} \neq 0$, folgt

$$v_{i_0} = -\lambda_0^{-1} \sum_{i\in\{1,\dots,r\}\setminus\{i_0\}} \lambda_i v_i = \sum_{i\in\{1,\dots,r\}\setminus\{i_0\}} (\lambda_{i_0}^{-1}\lambda_i) v_i.$$

"$\Leftarrow$": Gelte

$$v_{i_0} = \sum_{i\in\{1,\dots,r\}\setminus\{i_0\}} \lambda_i v_i.$$

für ein $i_0 \in \{1, \ldots, r\}$. Setze $\lambda_{i_0} = -1$. $\Rightarrow \sum_{i=1}^{r} \lambda_i v_i = 0 \Rightarrow (v_1, \ldots, v_r)$ sind linear abhängig. $\square$

Beispiele 3.22. *(i) Sei V der $\mathbb{R}$-VR $\mathbb{R}^n$. Die Familie der Einheitsvektoren $(e_1, \ldots, e_n)$ ist linear unabhängig, denn sind $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\sum_{i=1}^{n} \lambda_i e_i = 0$, d.h.*

$$0 = \lambda_1(1, 0, \ldots, 0) + \cdots + \lambda_n(0, \ldots, 0, 1) = (\lambda_1, \ldots, \lambda_n),$$

so folgt $\lambda_1 = \cdots = \lambda_n = 0$.

(ii) Sei $V = \mathbb{R}^2$. Die Familie $((1,1), (1,0), (0,1))$ ist linear abhängig, denn

$$1 \cdot (1,1) + (-1) \cdot (0,1) + (-1) \cdot (1,0) = 0.$$

(iii) Enthält die Familie $(v_1, \ldots, v_r)$ die 0 (d.h. das neutrale Element $0 = 0_V$ der Addition in V), so ist sie linear abhängig (weil $1 \cdot 0_V = 0_V$ mit $1 \neq 0 = 0_K$ gilt).

(iv) Enthält die Familie $(v_1, \ldots, v_r)$ zweimal denselben Vektor, so ist sie linear abhängig (weil $1 \cdot v + (-1) \cdot v = 0$ ist).

Zur Vereinfachung der Notation setzen wir $\mathrm{Lin}\big((v_i)_{i \in I}\big) := \mathrm{Lin}(\{v_i | i \in I\})$ und für endliche Familien $(v_1, \ldots, v_r)$ auch $\mathrm{Lin}(v_1, \ldots, v_r) := \mathrm{Lin}(\{v_1, \ldots, v_r\})$.

Satz 3.23. *Sei V ein K-VR, $(v_i)_{i \in I}$ eine Familie von Vektoren. Dann sind äquivalent:*

(i) $(v_i)_{i \in I}$ ist linear unabhängig.

(ii) Jeder Vektor $v \in \mathrm{Lin}\big((v_i)_{i \in I}\big)$ lässt sich in eindeutiger Weise als Linearkombination aus Vektoren der Familie $(v_i)_{i \in I}$ linear darstellen.

Beweis. (i) $\Rightarrow$ (ii): Sei $(v_i)_{i \in I}$ linear unabhängig und $v \in \mathrm{Lin}\big((v_i)_{i \in I}\big)$.
$\Rightarrow$ Es existieren $J \subset I$, J endlich, und $\lambda_i \in K$ für $i \in J$ mit

$$v = \sum_{i \in J} \lambda_i v_i \quad \text{(Existenz einer Darstellung)}.$$

Eindeutigkeit: Angenommen, es gibt eine weitere Darstellung, d.h. es existieren $H \subset I$, H endlich, und $\mu_i \in K$ für $i \in H$ mit

$$v = \sum_{i \in H} \mu_i v_i.$$

Bilde $G := H \cup J$ und setze $\lambda_i = 0$ für $i \in G \setminus J$ sowie $\mu_i = 0$ für $i \in G \setminus H$.

$$\Rightarrow \qquad 0 = v - v = \sum_{i \in G} (\lambda_i - \mu_i) v_i.$$

Da $(v_i)_{i\in I}$ linear unbahängige Familie ist, folgt $\lambda_i = \mu_i \ \forall\, i \in G$ und die Darstellung ist damit eindeutig.

(ii) $\Rightarrow$ (i): Gelte (ii). Zu zeigen: $(v_i)_{i\in I}$ ist linear unabhängig. (Beweisidee: Man nutzt die Eindeutigkeit der Darstellung für $v \in \mathrm{Lin}\big((v_i)_{i\in I}\big)$ für $v = 0$.)

Angenommen, $(v_i)_{i\in I}$ ist linear abhängig.
$\Rightarrow \exists\ J \subset I$, J endlich mit $(v_i)_{i\in J}$ ist linear abhängig
$\Rightarrow \forall\, i \in J\ \exists\, \lambda_i \in K$ (nicht alle $= 0$) mit

$$\sum_{i\in J} \lambda_i v_i = 0.$$

$\Rightarrow$ Die Darstellung der 0 als Element von $\mathrm{Lin}\big((v_i)_{i\in I}\big)$ ist nicht eindeutig (die andere Darstellung ist $0 = \sum_{i\in J} 0 \cdot v_i$). $\lightning$

Damit ist $(v_i)_{i\in I}$ linear unabhängig. □

Definition 3.24. *Sei V ein K-VR und $(v_i)_{i\in I}$ eine Familie von Vektoren.*

(i) $(v_i)_{i\in I}$ *heißt* <u>Erzeugendensystem</u> *(kurz: ES) von V, wenn $V = \mathrm{Lin}\big((v_i)_{i\in I}\big)$.*

(ii) V *heißt* <u>endlich erzeugt</u>*, wenn V ein endliches Erzeugendensystem besitzt.*

(iii) $(v_i)_{i\in I}$ *heißt* <u>Basis</u> *von V, wenn $(v_i)_{i\in I}$ ein linear unabhängiges Erzeugendensystem von V ist.*

(iv) Ist $B = (v_1, \ldots, v_r)$ eine endliche Basis von V, dann heißt r die <u>Länge</u> *von B.*

Bemerkung. *Jeder Vektorraum besitzt ein Erzeugendensystem, da $V = \mathrm{Lin}(V)$.*

Frage: Besitzt auch jeder Vektorraum eine Basis?
Wir wollen das zunächst für Vektorräume mit endlichem ES untersuchen.

Satz 3.25. *Sei $V \neq \{0\}$, $M = (v_1, \ldots, v_n)$ sei eine endliche Familie von Vektoren aus V. Dann sind äquivalent:*

(i) M ist eine Basis von V.

(ii) M ist ein unverkürzbares ES von V, d.h. M ist ein ES und für jedes $i \in \{1, \ldots, n\}$ ist $(v_j)_{j\in\{1,\ldots,n\}\setminus i}$ kein ES von V.

(iii) Zu jedem $v \in V$ gibt es eindeutig bestimmte $\lambda_1, \ldots, \lambda_n \in K$ mit

$$v = \sum_{i=1}^{n} \lambda_i v_i.$$

(iv) *M ist unverlängerbar linear unabhängig, d.h. M ist linear unabhängig und für jedes $v \in V$ ist die Familie $(v_1, \ldots, v_n, v)$ linear abhängig.*

Beweis. (i) $\Rightarrow$ (ii): Sei M Basis von V. Angenommen, M ist verkürzbar.
$\Rightarrow \exists\, v_i \in M$ mit $v_i = \sum_{j \neq i} \lambda_j v_j$
$\Rightarrow (v_1, \ldots, v_n)$ ist linear abhängig. $\lightning$

(ii) $\Rightarrow$ (iii): Sei $v \in V$. M Erzeugendensystem $\Rightarrow \exists\, \lambda_1, \ldots, \lambda_n :\ v = \sum_{i=1}^n \lambda_i v_i$.
Angenommen, die λ_i sind nicht eindeutig.
$\Rightarrow \exists\, \mu_1, \ldots, \mu_n :\ v = \sum_{i=1}^n \mu_i v_i$ und für mindestens ein $i_0 \in \{1, \ldots, n\}$ ist $\mu_{i_0} \neq \lambda_{i_0}$.
$\Rightarrow 0 = \sum_{i=1}^n (\mu_i - \lambda_i) v_i = \sum_{i \neq i_0} (\mu_i - \lambda_i) v_i + (\mu_{i_0} - \lambda_{i_0}) v_{i_0}$.
$\Rightarrow v_{i_0} = -(\mu_{i_0} - \lambda_{i_0})^{-1} \sum_{i \neq i_0} (\mu_i - \lambda_i) v_i$.
$\Rightarrow v_{i_0}$ kann weggelassen werden, d.h. M ist verkürzbar. $\lightning$

(iii) $\Rightarrow$ (iv): Nach Satz 3.23 ist die Familie $(v_1, \ldots, v_n)$ linear unabhängig. Sei $v \in V$.
$\Rightarrow \exists$ (eindeutige) $\lambda_1, \ldots, \lambda_n :\ v = \sum_{i=1}^n \lambda_i v_i$.
$\Rightarrow (v_1, \ldots, v_n, v)$ ist linear abhängig.
$\Rightarrow M$ ist unverlängerbar linear unabhängig.

(iv) $\Rightarrow$ (i): Sei M unverlängerbar linear unabhängig und sei $v \in V$.
$\Rightarrow (v_1, \ldots, v_n, v)$ ist linear abhängig.
$\Rightarrow$ Es existieren $\lambda_1, \ldots, \lambda_n, \lambda \in K$, nicht alle gleich 0, mit

$$\lambda v + \sum_{i=1}^n \lambda_i v_i = 0.$$

Es folgt $\lambda \neq 0$, andernfalls wäre $\sum_{i=1}^n \lambda_i v_i = 0$ und damit auch $\lambda_1 = \cdots = \lambda_n = 0$, denn M ist linear unabhängig. $\Rightarrow v = -\lambda^{-1} \sum_{i=1}^n \lambda_i v_i$. Aber damit ist M eine Basis. □

Korollar 3.26 (Korollar bedeutet "Folgerung")**.** *Seien K ein Körper und $V \neq \{0\}$ ein K-VR. Sei $M = (v_1, \ldots, v_n)$ eine endliche, linear unabhängige Familie von Vektoren aus V. Ist M keine Basis, so ist M verlängerbar linear unabhängig, d.h. es gibt ein $v \in V$, so dass $(v_1, \ldots, v_n, v)$ linear unabhängig ist.*

Beweis. Nach Satz 3.25 gilt (iv) $\Rightarrow$ (i). Die Umkehrung ergibt $\neg$(i) $\Rightarrow \neg$(iv). □

Eine weitere Folgerung ist der

Satz 3.27 (Basisauswahlsatz)**.** *Besitzt V ein endliches Erzeugendensystem, dann kann man daraus eine Basis auswählen, d.h. zu einem ES $(v_1, \ldots, v_n)$ gibt es eine Teilmenge $\{i_1, \ldots, i_r\} \subset \{1, \ldots, n\}$, so dass $(v_{i_1}, \ldots, v_{i_r})$ eine Basis ist.*

Beweis. Man entferne von dem ES nacheinander solange Elemente, bis die resultierende Familie ein unverkürzbares ES und somit nach Satz 3.25 auch eine Basis ist. □

Bemerkung. *Beim "Verkürzen" des ES' kann man im Allgemeinen nicht beliebige Elemente entfernen, sondern muss sukzessive überprüfen, ob das ES ohne jeden einzelnen der Vektoren noch ein ES ist. Insofern sagt der Satz nichts über die beste Strategie aus, aus einem gegebenen ES eine Basis zu finden.*

Korollar 3.28. *Jeder endlich erzeugte K-VR besitzt eine Basis von endlicher Länge.*

Fragen: Ist V ein endlich erzeugter K-VR.

- Ist dann jede Basis von endlicher Länge?
- Sind verschiedene Basen von V gleich lang?

Lemma 3.29 (Basisaustauschlemma)**.** *Sei V ein endlich erzeugter K-VR, $B = (v_1, \dots, v_r)$ eine Basis von V, $w = \sum_{i=1}^{r} \lambda_i v_i$. Dann gilt: Ist $\lambda_k \neq 0$ für ein $k \in \{1, \dots, r\}$, so ist*

$$B' = (v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r)$$

ebenfalls eine Basis von V (d.h. man kann v_k gegen w austauschen).

Beweis. Damit die Formulierung des Beweises nicht zu technisch wird, nehmen wir OE ("ohne Einschränkung" – man schreibt manchmal auch OBdA = "ohne Beschränkung der Allgemeinheit") an, dass $k = 1$ ist. Wegen $\lambda_1 \neq 0$ gilt dann

$$v_1 = -\frac{1}{\lambda_1} \sum_{i=2}^{r} \lambda_i v_1 + \frac{1}{\lambda_1} w. \tag{3.1}$$

Wir zeigen jetzt, dass auch $B' = (w, v_2, \dots, v_r)$ eine Basis ist.

- B' ist ein ES von V:

  Sei $v \in V$. Da $(v_1, \dots, v_r)$ Basis von V ist, existieren $\mu_1, \dots, \mu_r \in K$ mit

  $$v = \sum_{i=1}^{r} \mu_i v_i \overset{(3.1)}{=} \frac{\mu_1}{\lambda_1} w + \sum_{i=2}^{r} \Big(\mu_i - \mu_1 \frac{\lambda_i}{\lambda_1}\Big) v_i,$$

  also ist $v \in \mathrm{Lin}(w, v_2, \dots, v_r)$.

- B' ist linear unabhängig:

  Seien μ und $\mu_2, \dots, \mu_r \in K$ mit $\mu w + \sum_{i=2}^{r} \mu_i v_i = 0$. Wegen $w = \sum_{i=1}^{r} \lambda_i v_i$ mit $\lambda_1 \neq 0$ folgt $\mu\lambda_1 v_1 + \sum_{i=2}^{r} (\mu_i + \mu\lambda_i) v_i = 0$.
  $\overset{B \text{ Basis}}{\Rightarrow} \mu\lambda_1 = 0$ und $\mu_i + \mu\lambda_i = 0 \ \forall\, i \in \{2, \dots, r\}$.
  $\overset{\lambda_1 \neq 0}{\Rightarrow} \mu = 0$ und $\mu_i = 0 \ \forall\, i \in \{2, \dots, r\}$.
  $\Rightarrow B'$ ist Basis. □

Satz 3.30 (Basisaustauschsatz)**.** *Seien V ein endlich erzeugter K-VR, $(w_1, \dots, w_m)$ eine linear unabhängige Familie in V. Dann gilt:*

(i) Ist $B = (v_1, \dots, v_r)$ eine Basis, dann ist $r \geq m$.

(ii) Es gibt Indizes $i_{m+1}, \dots, i_r \in \{1, \dots, r\}$, so dass $B^* = (w_1, \dots, w_m, v_{i_{m+1}}, \dots, v_{i_r})$ wieder eine Basis von V ist, d.h. man kann m Elemente der Basis B gegen die $w_1, \dots, w_m$ austauschen.

Beweisidee: Wende das Austauschlemma 3.29 sukzessive auf $w_1, \dots, w_m$ an. Formal machen wir das mit einem Induktionsbeweis nach $n = 1, \dots, m$.

Beweis. Wir zeigen (ii) per Induktion nach n. Aus dem Beweis ergibt sich dann auch (i).

Wir halten zunächst fest, dass aus $(w_1, \dots, w_m)$ linear unabhängig auch $(w_1, \dots, w_n)$ linear unabhängig für alle $n \leq m$ folgt.

Induktionsanfang: Da B Basis ist, gibt es $\lambda_1, \dots, \lambda_r \in K$ mit $w_1 = \sum_{i=1}^{r} \lambda_i v_i$. Da $w_1 \neq 0$ gibt es ein $k \in \{1, \dots, r\}$ mit $\lambda_k \neq 0$. Nach dem Austauschlemma kann man v_k durch w_1 ersetzen und erhält eine neue Basis.

Induktionsschritt $n - 1 \to n$ für $n \leq m$: Sei die Behauptung für $n - 1$ bereits bewiesen, d.h. per Induktionsvoraussetzung gibt es Indizes $i_n, \dots, i_r \in \{1, \dots, r\}$, so dass $B^* = (w_1, \dots, w_{n-1}, v_{i_n}, \dots, v_{i_r})$ eine Basis von V ist. Im Falle $n - 1 = r$ wäre $\tilde{B} = (w_1, \dots, w_{n-1})$ eine Basis von V, d.h. nach Satz 3.25 (iv) unverlängerbar linear unabhängig. $\lightning$ (zu $(w_1, \dots, w_n)$ linear unabhängig), d.h. $n - 1 < r$. Da B^* eine Basis ist, gibt es $\lambda_1, \dots, \lambda_r \in K$ mit

$$w_n = \sum_{k=1}^{n-1} \lambda_k w_k + \sum_{k=n}^{r} \lambda_k v_{i_k}.$$

Falls $\lambda_n = \dots = \lambda_r = 0$, dann wäre $(w_1, \dots, w_n)$ linear abhängige Familie $\lightning$. Also gibt es ein $\lambda_k \neq 0$ mit $k \in \{n, \dots, r\}$. Nach dem Austauschlemma kann man das v_{i_k} gegen das w_n austauschen und erhält, dass $\tilde{B}^* := (w_1, \dots, w_n, v_{j_{n+1}}, \dots, v_{j_r})$ mit $j_{n+1}, \dots, j_r \in \{i_n, \dots, i_r\} \backslash \{i_k\} \subset \{1, \dots, r\}$ eine Basis von V ist. Aus dem letzten Schritt der Induktion für $n = m$ folgt auch $r \geq m$, d.h. (i). □

Satz 3.31. *Seien K ein Körper und V ein K-VR. Dann gilt:*

(i) Ist V endlich erzeugt, so ist jede Basis von V von endlicher Länge, und alle Basen von V haben dieselbe Länge.

(ii) Ist V nicht endlich erzeugt, dann existert von V keine Basis endlicher Länge.

Beweis. (i) Sei V endlich erzeugt. $\overset{\text{Korollar 3.28}}{\Rightarrow}$ $\exists$ endliche Basis $(v_1, \dots, v_r)$ von V.
Sei $(w_i)_{i \in I}$ eine beliebige Basis von V.
Ist I unendlich oder endlich mit Länge $s \geq r+1$, so existieren $i_1, \dots, i_{r+1} \in I$, so dass $(v_{i_1}, \dots, v_{i_{r+1}})$ linear unabhängige Familie ist.
$\overset{\text{Austauschsatz}}{\Rightarrow}$ Länge der Basis $(= r) \geq$ Anzahl der unabhängigen Vektoren $(= r+1)$ $\lightning$
$\Rightarrow I$ endlich mit Länge $s \leq r$.
Vertauscht man die Basen im obigen Argument, erhält man $r \leq s$ und damit $s = r$.

(ii) Angenommen, es existiert eine Basis endlicher Länge. Dann ist diese insbesondere ein endliches Erzeugendensystem. $\lightning$ □

Definition 3.32. *Die Dimension eines K-Vektorraums V ist definiert als*

$$\dim_K V := \begin{cases} \text{Länge einer Basis} & \text{falls } V \text{ endlich erzeugt ist} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Wir setzen $\dim_K \{0\} := 0$.

Wegen Satz 3.31 (i) ist $\dim_K V$ wohldefiniert. Falls keine Verwechslungsgefahr besteht, lässt man den Körper K auch weg und schreibt kurz $\dim V$.

Beispiele 3.33. *(i) Die Standardbasis $(e_1, \dots, e_n)$ von K^n hat die Länge n, also ist* $\dim_K K^n = n$.

(ii) Die von einem Vektor $(1, \lambda_0)$ erzeugte Gerade V im $\mathbb{R}^2$ ist ein $\mathbb{R}$-VR mit Dimension $\dim V = 1$.

(iii) Seien v_1 und v_2 zwei linear unabhängige Vektoren in $\mathbb{R}^3$. Dann ist der UVR $U =$ $\mathrm{Lin}(v_1, v_2)$ eine Ebene durch den Nullpunkt mit $\dim U = 2$.

Korollar 3.34. *Seien V ein endlichdimensionaler K-VR und $U \subset V$ ein UVR. Dann gilt:*

(i) U ist endlichdimensional.

(ii) $\dim_K U \leq \dim_K V$.

(iii) Es gilt $U = V \Leftrightarrow \dim_K U = \dim_K V$.

Beweis. Sei $\dim V = r \in \mathbb{N}$. Es existiert also eine Basis der Länge r.

(i) Angenommen, U ist nicht endlichdimensional (insbesondere $U \neq \{0\}$). Wir zeigen per Induktion: $\forall\, m \in \mathbb{N}$ gibt es dann linear unabhängige $(u_1, \dots, u_m)$ in U.

Induktionsanfang $m = 1$: Wegen $U \neq \{0\}$ $\exists\, u_1 \in U \setminus \{0\}$ und (u_1) ist linear unabhängige Familie.

Induktionsschritt $m \to m+1$: Seien $(u_1, \ldots, u_m)$ linear unabhängig. $(u_1, \ldots, u_m)$ kann keine Basis sein, andernfalls wäre U endlichdimensional. $\overset{\text{Korollar 3.26}}{\Rightarrow}$ $(u_1, \ldots, u_m)$ ist verlängerbar linear unabhängig, d.h. $\exists\, u_{m+1} \in U$: $(u_1, \ldots, u_{m+1})$ ist linear unabhängig.

$\Rightarrow (u_1, \ldots, u_{r+1})$ linear unabhängig in V. $\lightning$ zum Austauschsatz.

(ii) Seien $s = \dim_K U$ und $(u_1, \ldots, u_s)$ eine Basis von U. $\Rightarrow (u_1, \ldots, u_s)$ linear unabhängig $\overset{\text{Austauschsatz}}{\Rightarrow}$ $s \leq r = \dim_K V$.

(iii) "$\Rightarrow$": klar. "$\Leftarrow$": Angenommen, $U \subset V$, aber $U \neq V$. Sei $(u_1, \ldots, u_r)$ eine Basis von U. Wegen $U \neq V$ ist $(u_1, \ldots, u_r)$ zwar linear unabhängig, aber keine Basis von V. Nach Korollar 3.26 ist $(u_1, \ldots, u_r)$ dann in V verlängerbar linear unabhängig, d.h. $\exists\, v \in V$, so dass $(u_1, \ldots, u_r, v)$ linear unabhängig ist. $\overset{\text{Austauschsatz}}{\Rightarrow}$ $r+1 \leq \dim V = \dim U = r$. $\lightning$ □

Satz 3.35 (Basisergänzungssatz)**.** *Seien K ein Körper, V ein endlichdimensionaler K-VR mit $r = \dim_K V$ und $(u_1, \ldots, u_n)$ eine linear unabhängige Familie in V. Dann existieren $u_{n+1}, \ldots, u_r \in V$, so dass $(u_1, \ldots, u_r)$ eine Basis von V ist, d.h. $(u_1, \ldots, u_n)$ kann zu einer Basis ergänzt werden.*

Beweis. Nach Korollar 3.28 und Satz 3.31 besitzt V eine Basis $(v_1, \ldots, v_r)$. Die Behauptung folgt nun aus dem Basisaustauschsatz (d.h. die $u_{n+1}, \ldots, u_r$ sind $n - r$ Vektoren von $(v_1, \ldots, v_r)$). □

3.3 Auswahlaxiom, Zornsches Lemma und Basisexistenzsatz allgemein

Zu einer Menge M bezeichnen wir mit $\mathcal{P}(M)^\times = \{S \subset M | S \neq 0\}$ die Menge der nicht-leeren Teilmengen von M, d.h. die Potenzmenge von M ohne die leere Menge.

Definition 3.36. *Eine Auswahlfunktion auf einer Menge M ist eine Abbildung $f : \mathcal{P}(M)^\times \to M$ mit der Eigenschaft, dass $f(A) \in A$ für alle $A \in \mathcal{P}(M)^\times$.*

Die Funktion f "wählt" also für jede nicht-leere Teilmenge A von M ein Element aus A "aus".

Auswahlaxiom: Zu jeder Menge gibt es eine Auswahlfunktion.

Wenngleich die Existenz einer Auswahlfunktion plausibel erscheint und in Einzelfällen auch ohne axiomatische Forderung eine Auswahlfunktion schlicht angegeben werden kann, muss ihre Existenz im Allgemeinen tatsächlich axiomatisch gefordert werden.

Es gibt mehrere zum Auswahlaxiom logisch äquivalente Postulierungen – eine ist das Zornsche Lemma. Um dieses formulieren zu können, benötigen wir noch die Klärung einiger Begriffe.

Definition 3.37. *Eine (partielle) Ordnung auf einer Menge M ist eine Relation $\preceq$ auf der Menge M, so dass für alle $x, y, z \in M$ gilt:*

(i) $x \preceq y$ *und* $y \preceq z \Rightarrow x \preceq z$ *(Transitivität)*

(ii) $x \preceq y$ *und* $y \preceq x \Rightarrow x = y$ *(Antisymmetrie)*

(iii) $x \preceq x$ *(Reflexivität).*

Eine totale Ordnung auf einer Menge ist eine partielle Ordnung auf M*, so dass für alle* $x, y \in M$ *zusätzlich*

(iv) $x \preceq y$ *oder* $y \preceq x$ *gilt,*

also zwei Elemente aus M *immer zueinander in Relation stehen.*

Eine Menge mit partieller bzw. totaler Ordnung bezeichnet man als partiell bzw. total geordnete Menge.

Beispiele 3.38. *(i) Die Potenzmenge einer Menge* M *ist durch die Inklusionsrelation* $\subset$ *partiell geordnet.*

(ii) Die ganzen Zahlen $\mathbb{Z}$ *sind mit der* $\leq$*-Relation total geordnet.*

Definition 3.39. *(i) Sei* M *eine bezüglich* $\preceq$ *partiell geordnete Menge. Ein Element* $x \in M$ *heißt obere Schranke für die Teilmenge* $S \subset M$*, wenn* $y \preceq y$ *für alle* $y \in S$ *gilt.*

(ii) Die Menge M *heißt bezüglich der Ordnung* $\preceq$ *induktiv geordnet, wenn jede total geordnete Teilmenge* $S \subset M$ *eine obere Schranke in* M *besitzt.*

(iii) Ein Element $x \in M$ *einer bezüglich* $\preceq$ *partiell geordneten Menge* M *heißt maximales Element, wenn für alle* $y \in M$ *mit* $x \preceq y$ *bereits* $x = y$ *gilt.*

Für eine Menge M ist beispielsweise die Potenzmenge bezüglich der Inklusion induktiv geordnet. Mithilfe eines Fixpunktsatzes von Bourbaki kann man beweisen, dass folgende Aussage aus dem Auswahlaxiom folgt – tatsächlich ist letzteres dazu sogar äquivalent.

Lemma von Zorn: Eine nicht-leere induktiv geordnete Menge besitzt ein maximales Element.

Satz 3.40 (Basisergänzungssatz für unendlichdimensionale VR)**.** *Seien* K *ein Körper,* V *ein* K*-VR und* $(u_j)_{j \in J}$ *eine linear unabhängige Familie in* V*. Dann kann* $(u_j)_{j \in J}$ *zu einer Basis von* V *ergänzt werden, d.h. es existiert* I *mit* $J \subset I$ *und eine Familie* $(v_i)_{i \in I}$ *mit* $v_j = u_j$ *für alle* $j \in J$*, so dass* $(v_i)_{i \in I}$ *eine Basis von* V *ist. Insbesondere besitzt jeder* K*-VR eine Basis.*

Zum Beweis muss man zeigen, dass es eine maximal linear unabhängige Familie gibt, die $(u_j)_{j \in J}$ enthält. Dazu verwenden wir das Zornsche Lemma.

Beweis. Sei $\mathcal{M}$ das System aller Teilmengen $A \subset V$, die $\{u_j | j \in J\}$ enthalten und selbst eine linear unabhängige Familie von Vektoren bilden:

$$\mathcal{M} = \{A \subset V | A \text{ ist linear unabhängig und } \{u_j | j \in J\} \subset A\}.$$

$\mathcal{M}$ ist partiell geordnet durch die Inklusion $\subset$. Diese Ordnung ist aber induktiv, denn bei einer total geordneten Teilmenge $\mathcal{M}' \subset \mathcal{M}$ ist auch $\cup_{A \in \mathcal{M}'} A$ linear unabhängig. Sind nämlich sind $v_1, \ldots, v_r \in \cup_{A \in \mathcal{M}'} A$ paarweise verschieden, so gibt es ein $A \in \mathcal{M}'$ mit $v_1, \ldots, v_r \in A$. Nach dem Lemma von Zorn besitzt $\mathcal{M}$ ein maximales Element. Aber jedes solche ist eine Basis: Wäre ein maximales Element M keine Basis, gäbe es ein $v \in V \setminus \mathrm{Lin}(M)$. Aber dann wäre $M \cup \{v\}$ linear unabhängig, im Widerspruch zur Maximalität von M. □

3.4 Matrizen

Das Arbeiten mit Matrizen spielt eine zentrale Rolle in der linearen Algebra. Nach einer Einführung werden wir als erste Anwendung eine Basis von einem UVR $U = \mathrm{Lin}(v_1, \ldots, v_n)$ bestimmen.

Definition 3.41. *Sei K ein Körper, $m, n \in \mathbb{N}$. Eine $m \times n$-Matrix A (mit Elementen aus K) ist eine Familie*

$$A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \quad \textit{(oder kurz } (a_{ij})\textit{).}$$

Die Menge aller $m \times n$-Matrizen mit Einträgen aus K wird mit $M(m \times n, K)$ bezeichnet.

Bemerkung 3.42. *Wir schreiben ab jetzt die Elemente aus K^n immer als Spaltenvektoren, also*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n.$$

Definition 3.43 (Verknüpfungen auf $M(m \times n, K)$)**.**

(i) Für $(a_{ij}), (b_{ij}) \in M(m \times n, K)$ definieren wir die Addition durch

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}) \tag{3.2}$$

und die skalare Multiplikation durch

$$\lambda \cdot (a_{ij}) := (\lambda \cdot a_{ij}) \quad \textit{für } \lambda \in K.$$

(ii) Wir definieren die Multiplikation von Matrizen durch

$$\begin{aligned} \cdot : M(m \times n, K) \times M(n \times r, K) &\to M(m \times r, K) \\ (a_{ij})_{\substack{1 \le i \le m \\ 1 \le k \le n}} \cdot (b_{jk})_{\substack{1 \le j \le n \\ 1 \le k \le t}} &\mapsto (c_{ik})_{\substack{1 \le i \le m \\ 1 \le k \le r}} \end{aligned} \tag{3.3}$$

mit $c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}$.

$$i\text{-te Zeile} \begin{pmatrix} & * & \\ a_{i1} & \dots & a_{in} \\ & * & \end{pmatrix} \cdot \begin{pmatrix} & b_{1k} & \\ * & \vdots & * \\ & b_{nk} & \end{pmatrix} = \begin{pmatrix} & c_{ik} & \end{pmatrix}$$

k-te Spalte

Bemerkung. *Es ist leicht nachzurechnen, dass $M(m \times n, K)$ mit den Verknüpfungen aus (i) ein K-Vektorraum mit Dimension $\dim_K M(m \times n, K) = m \cdot n$ ist. Für den Nachweis der Dimension verwendet man, dass die $m \times n$-Matrizen E_{ij}, $i, = 1, \dots, m$, $j = 1, \dots, n$, mit*

$$(E_{ij})_{i'j'} = \ (i', j')\text{-ter Eintrag von } E_{ij} \ = \begin{cases} 1 & \text{falls } (i', j') = (i, j) \\ 0 & \text{andernfalls} \end{cases}$$

eine Basis bilden.

Lemma 3.44. *Seien K ein Körper, $m, n, r, s \in \mathbb{N}$, $A, A_1, A_2 \in M(m \times n, K)$ und $B, B_1, B_2 \in M(n \times r, K)$, $C \in M(r \times s, K)$.*

(i) Es gelten folgende Rechenregeln:

$$\begin{aligned} A \cdot (B_1 + B_2) &= A \cdot B_1 + A \cdot B_2 \\ (A_1 + A_2) \cdot B &= A_1 \cdot B + A_2 \cdot B \\ A \cdot (\lambda \cdot B) &= \lambda \cdot (A \cdot B) = (\lambda \cdot A) \cdot B \\ A \cdot (B \cdot C) &= (A \cdot B) \cdot C \end{aligned}$$

aber in der Regel <u>nicht</u>: $A \cdot B = B \cdot A$.

(ii) Mit der <u>Einheitsmatrix</u>

$$E_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in M(n \times n, K)$$

gilt $A \cdot E_n = A = E_m \cdot A$.

Beweis. Nachrechnen! □

Der Fall $m = n$ von "quadratischen" Matrizen ist besonders wichtig.

Lemma 3.45. *$M(n \times n, K)$ ist mit den inneren Veerknüpfungen $+$ aus (3.2) und $\cdot$ aus (3.3) ein Ring mit Einselement E_n. Für $n > 1$ ist dieser Ring nicht kommutativ, d.h. im Allgemeinen ist $A \cdot B \neq B \cdot A$.*

Beweis. Einfaches Nachrechnen der Eigenschaften (Übungsaufgabe)! Für den Nachweis der fehlenden Kommutativität muss man ein Beispiel angeben:

$$\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & \\ \hline 0 & & 0 \end{array}\right) \cdot \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & \\ \hline 0 & & 0 \end{array}\right) = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & \\ \hline 0 & & 0 \end{array}\right),$$

aber

$$\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & \\ \hline 0 & & 0 \end{array}\right) \cdot \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & \\ \hline 0 & & 0 \end{array}\right) = \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & \\ \hline 0 & & 0 \end{array}\right).$$

□

Definition 3.46. *$A \in M(n \times n, K)$ heißt invertierbar, wenn es $B \in M(n \times n, K)$ gibt mit $A \cdot B = B \cdot A = E_n$.*

Lemma 3.47. *Es gilt*

$$GL(n, K) := \left\{ A \in M(n \times n, K) \,\middle|\, A \text{ ist invertierbar} \right\}$$

ist bezüglich der Matrizenmultiplikation eine Gruppe, die allgemeine lineare Gruppe. Das neutrale Element ist E_n. Das zu $A \in GL(n, K)$ inverse Element bezeichnen wir mit A^{-1}. Für $A, B \in GL(n, K)$ gilt $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis. Wir müssen zunächst zeigen, dass für $A_1, A_2 \in GL(n, K)$ das Produkt $A_1 \cdot A_2$ wieder in $GL(n, K)$ liegt, d.h. auch invertierbar ist. Betrachte $A_2^{-1} \cdot A_1^{-1}$. Es gilt nach dem Assoziativgesetz der Matrizenmultiplikation (Lemma 3.44 (i))

$$(A_1 \cdot A_2)(A_2^{-1} \cdot A_1^{-1}) = A_1 \cdot A_1^{-1} = E_n$$

und

$$(A_2^{-1}A_1^{-1}) \cdot (A_1 \cdot A_2) = A_2^{-1} \cdot A_2 = E_n,$$

d.h. $A_1 \cdot A_2$ ist invertierbar mit Inversem $A_2^{-1} \cdot A_1^{-1}$. Damit gilt $A_1 \cdot A_2 \in GL(n, K)$. Das neutrale Element ist E_n. Mit $A \in GL(n, K)$ liegt auch A^{-1} in $GL(n, K)$, da A^{-1} das Inverse A besitzt: $A^{-1} \cdot A = A \cdot A^{-1} = E_n$. □

Definition 3.48. *Sei $A = (a_{ij}) \in M(m \times n, K)$. Dann heißt*

$$A^t := \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

die zu A transponierte Matrix (Transponierte von A). Wir schreiben oft A' anstelle von A^t.

Für $n = m$ ensteht A^t aus A durch "Spiegelung" an der Diagonalen $(a_{11}, a_{22}, \dots, a_{nn})$.

Beispiel 3.49.

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

Lemma 3.50. *Seien $A, A_1, A_2 \in M(m \times n, K)$, $B \in M(n \times r, K)$, $\lambda \in K$. Dann gilt*

(i) $(A_1 + A_2)^t = A_1^t + A_2^t$

(ii) $(\lambda \cdot A)^t = \lambda \cdot A^t$

(iii) $(A^t)^t = A$

(iv) $(A \cdot B)^t = B^t \cdot A^t$.

Beweis. Wir beweisen nur (iv), (i) – (iii) sind klar. Mit $(a_{ij}^t) := A^t$ und $(b_{ij}^t) := B^t$ ist

$$(B^t \cdot A^t)_{ki} = \sum_{j=1}^{n} b_{kj}^t a_{ji}^t = \sum_{j=1}^{n} b_{jk} a_{ij} = \sum_{j=1}^{n} a_{ij} b_{jk} = (A \cdot B)_{ik} = ((A \cdot B)^t)_{ki}.$$

Also ist $(A \cdot B)^t = B^t \cdot A^t$. □

3.4.1 Zeilenstufenform und Gaußalgorithmus

Beispiel 3.51. *Wir wollen folgendes Gleichungssystem lösen:*

$$\begin{array}{rcrcrcl} x & - & y & + & 2z & = & 4 \\ 3x & - & 3y & + & z & = & 2 \\ 2x & + & y & - & z & = & 5. \end{array}$$

Man kann das Gleichungssystem auch in Matrizenform schreiben

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & -3 & 1 \\ 2 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix}. \tag{3.4}$$

$$
\begin{array}{rcrcrcrll}
x & - & y & + & 2z & = & 4 & & \\
3x & - & 3y & + & z & = & 2 & | & \textit{Zeile 2 - 3 · Zeile 1} \\
2x & + & y & - & z & = & 5 & | & \textit{Zeile 3 - 2 · Zeile 1} \\
\hline
x & - & y & + & 2z & = & 4 & & \\
 & & & - & 5z & = & -10 & | & \textit{Vertauschen der 2.} \\
 & & 3y & - & 5z & = & -3 & | & \textit{und 3. Zeile} \\
\hline
x & - & y & + & 2z & = & 4 & & \\
 & & 3y & - & 5z & = & -3 & | & \cdot 1/3 \\
 & & & - & 5z & = & -10 & | & \cdot(-1/5) \\
\hline
x & - & y & + & 2z & = & 4 & & \\
 & & y & - & \frac{5}{3}z & = & -1 & & \\
 & & & & z & = & 2 & &
\end{array}
$$

In der Matrizenformschreibweise (3.4) können wir die gleichen Operationen an den Zeilen von

$$
A := \begin{pmatrix} 1 & -1 & 2 \\ 3 & -3 & 1 \\ 2 & 1 & -1 \end{pmatrix}
$$

durchführen. Die Zeilenoperationen wollen wir in diesem Abschnitt untersuchen. Wir halten noch die Matrixschreibweise als vorletzten Schritt fest:

$$
\underbrace{\begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & -5 \\ 0 & 0 & -5 \end{pmatrix}}_{\textit{Zeilenstufenform von } A} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ -10 \end{pmatrix}.
$$

Definition 3.52. *Sei $A \in M(m \times n, K)$ mit Zeilen (!) $a_1, \ldots, a_m \in K^m$. Wir definieren folgende elementaren Zeilenumformungen von A:*

(I) Multiplikation $ZM(i, \lambda)$ der i-ten Zeile mit $\lambda \in K \setminus \{0\}$

$$
\begin{pmatrix} \cdots \\ a_i \\ \cdots \end{pmatrix} \to \begin{pmatrix} \cdots \\ \lambda \cdot a_i \\ \cdots \end{pmatrix}
$$

(II) Addition $ZA(i, j, \lambda)$ des λ-fachen der j-ten Zeile zur i-ten Zeile($i \neq j$)

$$
\begin{pmatrix} \cdots \\ a_i \\ \cdots \\ a_j \\ \cdots \end{pmatrix} \to \begin{pmatrix} \cdots \\ a_i + \lambda \cdot a_j \\ \cdots \\ a_j \\ \cdots \end{pmatrix}
$$

(III) Vertauschen $ZV(i,j)$ der i-ten mit der j-ten Zeile ($i \neq j$)

$$\begin{pmatrix} \dots \\ a_i \\ \dots \\ a_j \\ \dots \end{pmatrix} \to \begin{pmatrix} \dots \\ a_j \\ \dots \\ a_i \\ \dots \end{pmatrix}$$

.

Bemerkung 3.53. *(III) kann man auch durch mehrfache Hintereinanderausführung von (I) und (II) erhalten.*

$$\begin{pmatrix} \dots \\ a_i \\ \dots \\ a_j \\ \dots \end{pmatrix} \xrightarrow{(II)} \begin{pmatrix} \dots \\ a_i + a_j \\ \dots \\ a_j \\ \dots \end{pmatrix} \xrightarrow{(I)} \begin{pmatrix} \dots \\ a_i + a_j \\ \dots \\ -a_j \\ \dots \end{pmatrix} \xrightarrow{(II)} \begin{pmatrix} \dots \\ a_i + a_j \\ \dots \\ a_i \\ \dots \end{pmatrix}$$

$$\downarrow (I)$$

$$\begin{pmatrix} \dots \\ a_j \\ \dots \\ a_i \\ \dots \end{pmatrix} \xleftarrow{(I)} \begin{pmatrix} \dots \\ a_j \\ \dots \\ -a_i \\ \dots \end{pmatrix} \xleftarrow{(II)} \begin{pmatrix} \dots \\ a_i + a_j \\ \dots \\ -a_i \\ \dots \end{pmatrix}$$

Lemma 3.54. *Elementare Zeilenumformungen von $A \in M(m \times n, K)$ erhält man durch Multiplikation von A von links mit den zugehörigen $M(m \times m, K)$-Elementarmatrizen, die wir mit $ZM(i,\lambda)$, $ZA(i,j,\lambda)$ und $ZV(i,j)$ bezeichnen. Es gilt:*

$$ZV(i,j) = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \downarrow i\text{-te Spalte} \quad \downarrow j\text{-te Spalte} \\ \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & 0 & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & 1 & & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & 0 & & & & & & & & & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow\ i\text{-te Zeile} \\ \\ \\ \\ \\ \leftarrow\ j\text{-te Zeile} \\ \\ \\ \\ \end{array}$$

$$ZM(i,\lambda) = \begin{array}{c} \\ \downarrow i\text{-te Spalte} \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & 0 & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \leftarrow i\text{-te Zeile}$$

$$ZA(i,j,\lambda) = \begin{array}{c} \\ \downarrow j\text{-te Spalte} \\ \\ \end{array} \begin{pmatrix} 1 & & & 0 \\ & \ddots & \lambda & \\ & 0 & \ddots & \\ & & & 1 \end{pmatrix} \leftarrow i\text{-te Zeile}$$

.

Beweis. Nachrechnen, dass $ZM(i,\lambda)\cdot A$, $ZA(i,j,\lambda)\cdot A$ und $ZV(i,j)\cdot A$ zu der entsprechenden Zeilenumformumg führt. □

Bemerkung 3.55. *In Beispiel 3.51 gilt*

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -\frac{5}{3} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{5} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 2 \\ 3 & -3 & 1 \\ 2 & 1 & -1 \end{pmatrix}}_{= \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 5 \\ 0 & 0 & -5 \end{pmatrix}}$$

.

Definition 3.56. *Sei $A \in M(m\times n, K)$ mit Zeilen $a_1,\ldots,a_m$. Man sagt, A sei in Zeilenstufenform (ZSF), falls folgende Bedingungen erfüllt sind:*

(i) Es gibt ein $r\in\mathbb{N}_0$, $r\le m$, so dass für die Zeilen $a_1,\ldots,a_m$ gilt:

$$a_i \neq 0 \text{ für } i=1,\ldots,r \text{ und } a_{r+1}=\cdots=a_m=0.$$

(ii) Setzen wir $j_i := \min\{j \in \{1,\dots,n\} | \, a_{ij} \neq 0\}$ *für* $i = 1,\dots,r$*, so gilt*

$$j_1 < j_2 < \cdots < j_r \quad \textit{(Stufenbedingung).}$$

Die Elemente $a_{1j_1},\dots,a_{rj_r}$ *heißen Pivots.*

Wir wollen jetzt den von den Zahlen $a_1,\dots,a_m$ aufgespannten Vektorraum betrachten und beweisen, dass die r Zeilen der Zeilenstufenform eine Basis bilden. Können wir dann noch ein Konstruktionsverfahren für die Zeilenstufenform angeben, haben wir eine Methode gefunden, um aus einem Erzeugendensystem eine Basis zu konstruieren.

Definition 3.57. *Sei* $A \in M(m \times n, K)$ *mit Zeilen* $a_1,\dots,a_m$*.*

$$ZR(A) := \mathrm{Lin}(a_1,\dots,a_m) \subset K^n$$

heißt der Zeilenraum von A. $SR(A) := ZR(A^t) \subset K^m$ *wird als Spaltenraum von A bezeichnet.*

Lemma 3.58. *Seien* $A, B \in M(m \times n, K)$*. Dann gilt: Ist B aus A durch eine endliche Folge von elementaren Zeilenumformungen entstanden, dann ist* $ZR(B) = ZR(A)$*.*

Beweis. Wir müssen die Aussage für die Zeilenumformungen (I) und (II) beweisen. Für (I) ist für $\lambda \neq 0$ und $i \in \{1,\dots,m\}$ zu zeigen:

$$\mathrm{Lin}(a_1,\dots,a_m) = \mathrm{Lin}(a_1,\dots,a_{i-1},\lambda a_i, a_{i+1},\dots,a_m).$$

Wegen der Implikation $M \subset \mathrm{Lin}(M') \Rightarrow \mathrm{Lin}(M) \subset \mathrm{Lin}(M')$ nach Lemma 3.17 muss gezeigt werden, dass alle Vektoren aus dem linken Erzeugendensystem in der linearen Hülle auf der rechten Seite enthalten sind und umgekehrt. Das ist aber unmittelbar klar. Für (II) folgt mit demselben Argument

$$\mathrm{Lin}(a_1,\dots,a_m) = \mathrm{Lin}(a_1,\dots,a_{i-1},a_i + \lambda a_j, a_{i+1},\dots,a_m).$$

D.h. die Aussage $ZR(A) = ZR(B)$ gilt nach jeweils einer Umformung vom Typ (I) oder (II) und damit auch für jede endliche Abfolge von elementaren Zeilenumformungen. □

Wir zeigen nun, dass man wie in Beispiel 3.51 jede $m \times n$-Matrix A durch endlich viele elementare Zeilenumformungen in eine Matrix B von ZSF bringen kann.

Gaußalgorithmus (oder gaußsches Eliminationsverfahren) Sei $A \in M(m \times n, K)$ mit $A \neq 0$. Wir wenden in jedem der r Schritte ($r \in \mathbb{N}$, $r \leq \min(n,m)$) die folgenden drei Teilschritte bestehend aus elementaren Zeilenumformungen an, um eine Matrix B in ZFS zu erhalten.

Start: Setze $A_1 = A$.

1. Teilschritt: Sei j_1 die erste Spalte von A_1, die nicht nur Nullen enthält, d.h.

$$j_1 := \min\big\{j \in \{1,\dots,n\} \,\big|\, \exists\, i \in \{1,\dots,m\} \text{ mit } a_{ij} \neq 0\big\}.$$

Wähle ein i_1 aus mit $a_{i_1 j_1} \neq 0$.

2. Teilschritt: Vertausche die i_1-te Zeile mit der ersten Zeile (Umformung vom Typ (III)). Man erhält das erste Pivot

$$\tilde{a}_{1j_1} = a_{i_1 j_1} \neq 0$$

und die transformierte Matrix

$$\tilde{A}_1 = \begin{pmatrix} 0 & \dots & 0 & \tilde{a}_{1j_1} & * & \dots & * \\ & & & * & & & \\ & 0 & & \vdots & & * & \\ & & & * & & & \end{pmatrix}$$

3. Teilschritt: Durch Umformungen vom Typ II können alle Einträge der j-ten Spalte unterhalb der obersten Zeile zu 0 gemacht werden. Man erhält

$$\bar{A}_1 = \left(\begin{array}{cccc|ccc} 0 & \dots & 0 & \tilde{a}_{1j_1} & * & \dots & * \\ \hline & & & 0 & & & \\ & 0 & & \vdots & & A_2 & \\ & & & 0 & & & \end{array}\right).$$

Man wende dann diese drei Teilschritte auf A_2 an. Dabei kann man die Zeilenumformungen von A_2 auf die ersten Spalten von $\bar{A}_1$ ausdehnen, da diese nur Nullen enthalten.

Iteriere dieses Verfahren. Das Verfahren bricht ab, weil die Folge $j_1 < j_2 \dots$ streng monoton wachsend ist und entweder nach r Schritten $j_r = n$ oder $A_{r+1} = 0$ gilt.

Wir bezeichnen mit B die resultierende Matrix in ZSF. Wegen des 1. Teilschrittes ("Wähle ein i_1 aus ...") ist die Matrix B nicht eindeutig.

Lemma 3.59. *Die ersten r Zeilen $b_1, \dots, b_r$ von B bilden eine Basis von $ZR(A)$. Es gilt $ZR(A) = ZR(B) = \mathrm{Lin}(b_1, \dots, b_r)$ und $\dim ZR(A) = \dim ZR(B) = r$.*

Beweis. $ZR(A) = ZR(B)$ und $\dim ZR(A) = \dim ZR(B) = r$ sind klar. Wir müssen nur die lineare Unabhängigkeit von $b_1, \dots, b_r$ zeigen. Seien $\lambda_1, \dots, \lambda_r \in K$ mit $\lambda_1 b_1 + \dots \lambda_r b_r = 0$. In der j_1-ten Komponente von $\lambda_1 b_1 + \dots + \lambda_r b_r$ steht $\lambda_1 b_{1j_1} = 0$. Wegen $b_{1j_1} \neq 0$ folgt $\lambda_1 = 0$. In der zweiten Komponente von $\lambda_1 b_1 + \dots \lambda_r b_r = \lambda_2 b_2 + \dots \lambda_r b_r$

steht $\lambda_2 b_{2j_2} = 0$ und wegen $b_{2j_2} \neq 0$ folgt $\lambda_2 = 0$. Man erhält so iterativ $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. □

Fazit: Wir haben damit eine Methode gefunden, um zu einem endlichen Erzeugendensystem eine Basis und die Dimension des erzeugten Vektorraums zu bestimmen.

Beispiel 3.60. *Sei* $W = \operatorname{Lin}\big((0,0,3,-1),(0,1,2,0),(0,3,0,2)\big) \subset \mathbb{R}^4$.

$$\begin{pmatrix} 0 & 0 & 3 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 3 & 0 & 2 \end{pmatrix} \overset{ZV(1,2)}{\longrightarrow} \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 3 & 0 & 2 \end{pmatrix} \overset{ZA(3,1,-3)}{\longrightarrow} \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -6 & 2 \end{pmatrix}$$

$$\downarrow ZA(3,2,2)$$

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$\Longrightarrow \big((0,1,2,0),(0,0,3,-1)\big)$ *ist eine Basis von* W *mit* $\dim W = 2$.

Definition 3.61. *Sei* $A \in M(m \times n, K)$.
Zeilenrang(A)$:= \dim ZR(A)$ *heißt Zeilenrang von* A.
Spaltenrang(A)$:= \dim SR(A)$ *heißt Spaltenrang von* A.

Beispiel 3.62. *Wir setzen Beispiel 3.60 mit*

$$A = \begin{pmatrix} 0 & 0 & 3 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 3 & 0 & 2 \end{pmatrix}$$

fort. Mithilfe elementarer Zeilenumformungen hatten wir die ZSF

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

erreicht, d.h. Zeilenrang(A)$= 2$. *Analog der Spaltenrang:*

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 3 \\ 3 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 0 & 2 \\ 0 & 1 & 3 \\ 3 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 2 & 6 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

d.h. Spaltenrang(A)$= 2 =$ *Zeilenrang(A).*

Wir werden später zeigen, dass generell Zeilenrang = Spaltenrang gilt.

3.5 Die Summe von Untervektorräumen

Definition 3.63. *Seien $U, W \subset V$ Untervektorräume (UVRs).*

$$U + W := \{u + w \mid u \in U, w \in W\}$$

heißt die Summe von U und W.

Achtung: $U + W$ ist nicht (!) die Vereinigung von U und W als Mengen!

Lemma 3.64. *Seien $U, W \subset V$ UVRs. Dann gilt*

(i) $U + W = \mathrm{Lin}(U \cup W)$, d.h. $U + W$ ist der kleinste UVR, der U und W enthält.

(ii) Sind U und W endlichdimensional, dann ist auch $U + W$ endlichdimensional mit $\dim(U + W) \leq \dim U + \dim W$.

Beweis. (i) "$\subset$": $U + W \subset \mathrm{Lin}(U \cup W)$ ist klar.
"$\supset$": Sei $v \in \mathrm{Lin}(U \cup W) \Rightarrow \exists\, s \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_s \in K$ und $w_1, \ldots, w_s \in U \cup W$ mit

$$v = \sum_{j=1}^{s} \lambda_j w_j.$$

Seien OE für ein $r \leq s$ die Vektoren $w_1, \ldots, w_r \in U$ und $w_{r+1}, \ldots, w_s \in W$. Dann ist

$$v \;=\; \underbrace{\sum_{j=1}^{s} \lambda_j w_j}_{\in U} \;+\; \underbrace{\sum_{j=r+1}^{s} \lambda_j w_j}_{\in W} \;\in\; U + W.$$

(ii) Sind $(u_1, \ldots, u_r)$ eine Basis von U und $(w_1, \ldots, w_s)$ eine Basis von W, so ist die Familie $(u_1, \ldots, u_r, w_1, \ldots, w_s)$ ein Erzeugendensystem von $U + W$.
$\Rightarrow \dim(U + W) \leq r + s = \dim U + \dim W$. □

Beispiele 3.65. *(i)* $K = \mathbb{R}$, $V = \mathbb{R}^2$, $U = \mathrm{Lin}((-1, 1))$, $W = \mathrm{Lin}((1, 1))$.
$\Rightarrow U + W = \mathrm{Lin}((1, -1)) + \mathrm{Lin}((1, 1)) = \mathrm{Lin}\big((-1, 1), (1, 1)\big) = \mathbb{R}^2$.

(ii) $K = \mathbb{R}$, $U = \mathrm{Lin}(e_1, e_2)$, $V = \mathrm{Lin}(e_2, e_3)$.
$\Rightarrow U + W$ *enthält* e_1, e_2, e_3, *d.h.* $U + W = \mathbb{R}^3$.

Satz 3.66. *Seien $U, W \subset V$ endlichdimensionale UVRs. Dann gilt*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Beweis. (i) Sei $(v_1, \dots, v_m)$ eine Basis von $U \cap W$. Nach dem Basisergänzungssatz gibt es $u_1, \dots, u_k \in U$ und $w_1, \dots, w_l \in W$, so dass $B_1 = (v_1, \dots, v_m, u_1, \dots, u_k)$ eine Basis von U ist und $B_2 = (v_1, \dots, v_m, w_1, \dots, w_l)$ eine Basis von W ist.

(ii) Behauptung: $B = (v_1, \dots, v_m, u_1, \dots, u_k, w_1, \dots, w_l)$ ist eine Basis von $U + W$. Klar: B ist eine Erzeugendensystem. Wir zeigen, dass B linear unabhängig ist. Sei

$$\underbrace{\lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 u_1 + \dots + \mu_k u_k}_{=:u \in U} + \nu_1 w_1 + \dots + \nu_l w_l = 0$$

$\Rightarrow u = -\nu_1 w_1 - \dots - \nu_l w_l \in W$
$\Rightarrow u \in U \cap W$
$\Rightarrow \mu_1 = \dots = \mu_k = 0$ (wegen Eindeutigkeit der Darstellung in B_1)
$\Rightarrow \lambda_1 v_1 + \dots + \lambda_m v_m + \nu_1 w_1 + \dots + \nu_l w_l = 0$
$\Rightarrow \lambda_1 = \dots = \lambda_m = \nu_1 = \dots = \nu_l = 0$ (da B_2 Basis).

(iii) Aus (i) und (ii) folgt

$$\begin{array}{ccccc} \dim(U+W) & = & \dim(U) & + \dim(W) & - \dim(U \cap W). \\ \| & & \| & \| & \| \\ m+k+l & & m+k & m+l & m \end{array}$$

□

Definition 3.67. *Seien $U, W \subset V$ UVRs. V heißt direkte Summe von U und W, wenn gilt: $V = U + W$ und für jedes $v \in V$ gibt es eine eindeutige Darstellung $v = u + w$ mit $u \in U$ und $w \in W$. Man schreibt $V = U \oplus W$.*

Bemerkung 3.68. *(i) In Beispiel 3.65 (i) gilt $V = U \oplus W$ und in Beispiel 3.65 (ii) ist $U + V = \mathbb{R}^3$ keine direkte Summe.*

(ii) Die Definition gilt auch für unendlichdimensionale UVRs.

Lemma 3.69. *Seien $U, W \subset V$ UVRs. Dann sind äquivalent:*

(i) $V = U \oplus W$

(ii) $V = U + W$ und $U \cap W = \{0\}$.

Beweis. (i) $\Rightarrow$ (ii): Aus (i) folgt sofort $V = U + W$. Angenommen, $U \cap V \neq \{0\}$. Seien $z \in U \cap W$, $z \neq 0$, und

$$v = u + w = \underbrace{(u - z)}_{\in U} + \underbrace{(w + z)}_{\in W},$$

d.h. die Darstellung ist nicht eindeutig. ↯
$\Rightarrow U \cap W = \{0\}$ und $V = U \oplus W$.

(ii) $\Rightarrow$ (i): Sei $v \in V$. $\Rightarrow \exists\, u \in U$, $w \in W$ mit $v = u + w$. Wir zeigen die Eindeutigkeit der Darstellung. Sei $v = u + w = \tilde{u} + \tilde{w}$ eine weitere Darstelllung mit $\tilde{u} \in U$, $\tilde{w} \in W$.

$$\Rightarrow \quad \underbrace{u - \tilde{u}}_{\in U} = \underbrace{\tilde{w} - w}_{\in W} = 0,$$

da $U \cap W = \{0\}$. Es folgen $u = \tilde{u}$, $w = \tilde{w}$ und die Darstellung ist eindeutig. □

Satz 3.70. *Seien V endlichdimensionaler K-VR, $U, W \subset V$ UVRs. Dann sind äquivalent:*

(i) $V = U \oplus W$

(ii) Für alle Basen $(u_1, \ldots, u_k)$ von U und $(w_1, \ldots, w_l)$ von W ist $(u_1, \ldots, u_k, w_1, \ldots, w_l)$ eine Basis von V.

(iii) Es gibt Basen $(u_1, \ldots, u_k)$ von U und $(w_1, \ldots, w_l)$ von W, so dass die Familie $(u_1, \ldots, u_k, w_1, \ldots, w_l)$ eine Basis von V ist.

(iv) $V = U + W$ *und* $\dim V = \dim U + \dim W$.

Beweis. (i)$\Rightarrow$(ii): Erzeugendensystem ist klar. Die behauptete lineare Unabhängigkeit folgt aus Satz 3.23 oder analog zum zweiten Beweisteil von Satz 3.66.

(ii)$\Rightarrow$(iii) ist klar.

(iii)$\Rightarrow$(iv) ist klar.

(iv)$\Rightarrow$(i) folgt aus Satz 3.66 und Lemma 3.69. □

Satz und Definition 3.71. *Seien V ein VR, $U \subset V$ ein UVR. Dann existiert ein UVR $W \subset V$ mit $V = U \oplus W$. W heißt Komplement zu U in V.*

Man beachte, das V nicht als endlichdimensional vorausgesetzt wurde.

Beweis. Wir wenden den Baisergänzungssatz 3.40 an.; Sei $(u_j)_{j \in J}$ eine Basis von U.
$\Rightarrow \exists\, I$ mit $J \subset I$ und eine Basis $(v_i)_{i \in I}$ von V mit $v_j = u_j \;\forall\, j \in J$. Insbesondere gilt

$$U = \operatorname{Lin}\big((v_i)_{i \in J}\big).$$

Setze $W := \operatorname{Lin}\big((v_i)_{i \in I \setminus J}\big)$. Wir wollen zeigen: $U \cap W = \{0\}$.
Sei $v \in U \cap W$. $\Rightarrow$ Es existieren Darstellungen $v = \lambda_{j_1} v_{j_1} + \cdots + \lambda_{j_k} v_{j_k} = \mu_{i_1} v_{i_1} + \cdots + \mu_{i_l} v_{i_l}$ mit $j_1, \ldots, j_k \in J$ und $i_1, \ldots, i_l \in I \setminus J$.

$$\Rightarrow \quad 0 = \sum_{m=1}^{k} \lambda_{j_m} v_{j_m} - \sum_{m=1}^{l} \mu_{i_m} v_{i_m}.$$

Da $(v_i)_{i\in I}$ Basis von V ist und somit insebseondere der Nullvektor eine eindeutige Darstellung als Linearkombination besitzt, folgt $\lambda_{j_1} = \cdots = \lambda_{j_k} = \mu_{i_1} = \cdots = \mu_{i_l} = 0$.
$\Rightarrow v = 0$. Lemma 3.69 impliziert dann die Behauptung. □

Bemerkung 3.72. *Das Komplement ist nicht eindeutig bestimmt. Beispielsweise gilt für $K = \mathbb{R}$, $V = \mathbb{R}^2$ und $U = \mathrm{Lin}(e_1)$*

$$V = U \oplus \mathrm{Lin}(e_2) = U \oplus \mathrm{Lin}((1,1)).$$

Definition 3.73 (Summe von r Vektorräumen)**.** *Sind $2 \leq r \in \mathbb{N}$ und $U_1, \ldots, U_r \subset V$ UVRs, so definieren wir*

$$U_1 + \ldots U_r := \left\{ u_1 + \cdots + u_r \;\middle|\; u_i \in U_i \text{ für } i = 1, \ldots, r \right\}.$$

Die Summe heißt direkte Summe, falls die Darstellung

$$v = u_1 + \cdots + u_r \quad \text{mit } u_i \in U_i,\ i = 1, \ldots r,$$

für alle $v \in U_1 + \cdots + U_r$ eindeutig ist. Man schreibt dann

$$U_1 \oplus \cdots \oplus U_r = \bigoplus_{i=1}^{r} U_i.$$

Bemerkung 3.74. *(i) Die Definition ist im Falle $r = 2$ identisch mit Definition 3.67.*

(ii) Außerdem ist offensichtlich, dass für $r = 3$ gilt

$$U_1 + U_2 + U_3 = (U_1 + U_2) + U_3$$
$$U_1 \oplus U_2 \oplus U_3 = (U_1 \oplus U_2) \oplus U_3$$

bzw. für allgemeine $r \in \mathbb{N}$, $r \geq 2$, entsprechend

$$U_1 + \cdots + U_r = (\ldots((U_1 + U_2) + U_3) + \ldots) + U_r$$
$$U_1 \oplus \cdots \oplus U_r = (\ldots((U_1 \oplus U_2) \oplus U_3) \oplus \ldots) \oplus U_r.$$

Damit übertragen sich die Ergebnisse des Falls $r = 2$ iterativ auf den allgemeinen Fall, z. Bsp. gilt für einen K-VR V:

Satz 3.75. *Seien $U_1, \ldots, U_r \subset V$ endlich erzeugte UVRs. Dann sind äquivalent:*

(i) $V = U_1 \oplus \cdots \oplus U_r$.

(ii) Für alle Basen $\left(u_1^{(i)}, \ldots, u_{k_i}^{(i)}\right)$ *von* U_i *(*$i = 1, \ldots, r$*) ist*

$$\left(u_1^{(1)}, \ldots, u_{k_1}^{(1)}, u_1^{(2)}, \ldots, u_{k_2}^{(2)}, \ldots, u_1^{(r)}, \ldots, u_{k_r}^{(r)}\right)$$

eine Basis von V*.*

(iii) $v = u_1 + \cdots + u_r$ *mit* $u_i \in U_i$*,* $i = 1, \ldots, r$*, und* $\dim V = \dim U_1 + \cdots + \dim U_r$*.*

Beweis. Der Beweis erfolgt durch iterative Anwendung von Satz 3.70. □

Das nächste Lemma beinhaltet für allgemeines $r \geq 2$ das Analogon der Bedingung (ii) aus Lemma 3.69.

Lemma 3.76. *Seien* $r \geq 2$ *und* $U_1, \ldots, U_r$ *UVRs mit* $V = U_1 + \cdots + U_r$*. Dann gilt* $V = U_1 \oplus \cdots \oplus U_r$ *genau dann, wenn*

$$U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{r} U_j = \{0\} \quad \forall\, i = 1, \ldots, r.$$

Beweis. "⇒": Angenommen, es existiert $i \in \{1, \ldots, r\}$ mit

$$U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{r} U_j \neq \{0\}.$$

$\Rightarrow \exists\, u_j \in U_j$, $j = 1, \ldots, r$, $u_i \neq 0$ mit $u_i = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_r$, womit $0 = u_1 + \cdots + u_{i-1} - u_i + u_{i+1} + \cdots + u_r$. ↯ (zur Eindeutigkeit der Darstellung der 0)

"⇐": Sei $v \in V$, besitzt also eine Darstellung $v = u_1 + \cdots + u_r$ mit $u_i \in U_i$, $i = 1, \ldots, r$. Zu zeigen: Die Darstellung ist eindeutig. Seien $\tilde{u}_j \in U_j$, $j = 1, \ldots, r$, mit $v = u_1 + \cdots + u_r = \tilde{u}_1 + \cdots + \tilde{u}_r$, so dass nicht $u_j = \tilde{u}_j$ für alle j gilt. Es existiert folglich $i_0 \in \{1, \ldots, r\}$ mit $u_{i_0} \neq \tilde{u}_{i_0}$.

$$\Rightarrow \quad 0 \neq u_{i_0} - \tilde{u}_{i_0} = -\sum_{\substack{j=1 \\ j \neq i_0}}^{r} (u_j - \tilde{u}_j) \quad \Rightarrow \quad U_{i_0} \cap \sum_{\substack{j=1 \\ j \neq i_0}}^{r} U_j \neq \{0\}. \text{ ↯}$$

□

Bemerkung 3.77. *(i) Für den Nachweis einer direkten Summe* $U_1 \oplus \cdots \oplus U_r$ $(r \geq 3)$ *reicht es nicht, paarweise* $U_i \cap U_j = \{0\}$ *für alle* $i \neq j$*,* $1 \leq i, j \leq r$ *zu fordern. Ein Gegenbeispiel für* $r = 3$*:* $K = \mathbb{R}$*,* $V = \mathbb{R}^2$*,*

$$U_1 = \mathrm{Lin}(e_1),\ U_2 = \mathrm{Lin}(e_2),\ U_3 = \mathrm{Lin}((1,1)).$$

Dann gilt $V = U_1 + U_2 + U_3 = \mathbb{R}^2$ *mit* $U_1 \cap U_2 = \{0\}$, $U_2 \cap U_3 = \{0\}$ *und* $U_1 \cap U_3 = \{0\}$, *aber*

$$U_1 \cap \underbrace{(U_2 + U_3)}_{=\mathbb{R}^2} = U_1 \neq \{0\},$$

d.h. die Summe ist nicht direkt.

(ii) Zur Warnung vor Rechenfehlern sei bemerkt, dass für die Summe von UVRs kein Distributivgesetz gilt. Mit den Bezeichnungen aus (i) ist bspw.

$$\underbrace{(U_1 \cap U_2)}_{=\{0\}} + \underbrace{(U_1 \cap U_3)}_{=\{0\}} = \{0\} \neq U_1 = U_1 \cap (U_2 + U_3).$$

Der Vollständigkeit halber sei nach angemerkt, dass ja gilt

$$U_1 \oplus \cdots \oplus U_r = (\dots((U_1 \oplus U_2) \oplus U_3) \oplus \dots) \oplus U_r. \tag{3.5}$$

Die Bedingung, dass alle Summen direkt sind, wäre

$$U_i \cap \sum_{j=1}^{i-1} U_j = \{0\} \quad \forall\, i \in \{1, \dots, r\}.$$

Wegen der Gleichheit in (3.5) folgt aus Lemma 3.76, dass dies äquivalent ist zu

$$U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{r} U_j = \{0\} \quad \forall\, i \in \{1, \dots, r\}.$$

Das ist nicht ganz offensichtlich. Man kann diese Äquivalenz aber trotzdem direkt nachrechnen (Übungsblatt 11, Aufgabe 1).

4 Lineare Abbildungen

In diesem Kapitel sind U, V und W stets K-Vektorräume.

Definition 4.1. *Sei* $f : V \to W$ *eine Abbildung.* f *heißt (*K*-)lineare Abbildung oder (Vektorraum-)Homomorphismus, wenn folgende Bedingungen erfüllt sind:*

(L1) $f(u + v) = f(u) + f(v)$ *für alle* $u, v \in V$

(L2) $f(\lambda v) = \lambda f(v)$ *für alle* $v \in V$ *und für alle* $\lambda \in K$.

Beispiele 4.2. *(i) Sei $A = (a_{ij}) \in M(m \times n, K)$. Wir betrachten die Abbildung*

$$\tilde{A} : K^n \to K^m,\ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A \cdot x.$$

Es gelten für $u, v \in K^n$, $\lambda \in K$:

$$\tilde{A}(u+v) = A \cdot (u+v) = A \cdot u + A \cdot v = \tilde{A}(u) + \tilde{A}(v) \text{ sowie}$$
$$\tilde{A}(\lambda v) = A \cdot (\lambda v) = \lambda \cdot (A \cdot v) = \lambda \cdot \tilde{A}(v).$$

Wegen

$$\tilde{A}(e_i) = A \cdot e_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{im} \end{pmatrix}$$

stehen in den Spalten von A die Bilder der Basisvektoren $e_1, \ldots, e_n$ von K^n unter $\tilde{A}$. Sind $A \in M(m \times n, K)$ und $B \in M(n \times r, K)$, $x \in K^r$, dann gilt

$$\widetilde{AB}(x) = (AB)x = A(Bx) = A \cdot \tilde{B}(x) = \tilde{A}\big(\tilde{B}(x)\big) = \big(\tilde{A} \circ \tilde{B}\big)(x),$$

d.h. die Verknüpfung $\tilde{A} \circ \tilde{B} = \widetilde{AB}$ entspricht der Matrix-Multiplikation.

(ii) Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$, $\binom{x_1}{x_2} \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\binom{x_1}{x_2} = \binom{x_1}{-x_2}$. Die Abbildung ist linear nach (i) und beschreibt die Spiegelung an der x_1-Achse.

(iii) Sei $V = \{f : (0,1) \to \mathbb{R} |\ f \text{ ist differenzierbar}\}$. V ist ein $\mathbb{R}$-VR. Dann ist

$$' : V \to \{g : (0,1) \to \mathbb{R}\},$$
$$f \mapsto f'$$

lineare Abbildung, denn für $f, g \in V$ gilt $(f+g)' = f' + g'$ und $(\lambda \cdot f)' = \lambda \cdot f'$ $(\lambda \in \mathbb{R})$.

Lemma 4.3. *Sei $f : V \to W$ eine lineare Abbildung. Dann gilt:*

(i) $f(0) = 0$

(ii) $f\big(\sum_{i=1}^n \lambda_i v_i\big) = \sum_{i=1}^n \lambda_i f(v_i)$ *für alle* $v_i \in V$, $\lambda_i \in K$ $(i = 1, \ldots, n)$

(iii) $V' \subset V$ *UVR* $\Rightarrow f(V')$ *ist UVR*

(iv) $W' \subset W$ *UVR* $\Rightarrow f^{-1}(W) \subset V$ *ist UVR* (f^{-1} *Urbild*)

(v) $(v_i)_{i\in I}$ *linear abhängige Familie in* $V \Rightarrow (f(v_i))_{i\in I}$ *linear abhängige Famile in* W

(vi) $V' = \operatorname{Lin}\big((v_i)_{i\in I}\big) \Rightarrow f(V') = \operatorname{Lin}\big((f(v_i))_{i\in I}\big)$

(vii) W *endlichdim.* $\Rightarrow f(V)$ *endlichdim. UVR von* W *mit* $\dim f(V) \leq \dim W$.

Beweis. (i) Es gilt $f(0) = f(0+0) \overset{(L1)}{=} f(0) + f(0) \overset{\text{Kürzungsregel}}{\Rightarrow} f(0) = 0$.

(ii) folgt aus iterativer Anwendung von (L1) und (L2).

(iii) Sei $V' \subset V$ UVR. Zu zeigen: $f(V')$ it UVR von W. Dazu weisen wir die Bedingungen aus Lemma 3.6 nach.

- Wegen $0 \in V'$ ist $f(0) \overset{(i)}{=} 0 \in f(V')$. Insbesondere ist $f(V') \neq \emptyset$.
- Seien $w_1, w_2 \in f(V')$, d.g. es existieren $v_1, v_2 \in V'$ mit $w_1 = f(v_1)$, $w_2 = F(v_2)$.
  $\Rightarrow w_1 + w_2 = f(v_1) + f(v_2) \overset{(L1)}{=} f(\underbrace{v_1 + v_2}_{\in V'}) \in f(V')$.
- Sei $\lambda \in K$, $w \in f(V')$, d.h. $w = f(v)$ für ein $v \in V'$.
  $\Rightarrow \lambda w = \lambda f(v) \overset{(L2)}{=} f(\underbrace{\lambda v}_{\in V'}) \in f(V')$.

(iv) Sei $W' \subset W$ UVR. Zu zeigen: $f^{-1}(W') \subset V$ ist UVR. Wieder weisen wir dazu die Bedingungen aus Lemma 3.6 nach.

- Wegen $f(0) \overset{(i)}{=} 0 \in W'$ ist $0 \in f^{-1}(\{0\})$, womit $f^{-1}(W') \neq \emptyset$.
- Seien $v_1, v_2 \in f^{-1}(W') \Rightarrow f(v_1), f(v_2) \in W'$
  $\Rightarrow f(v_1 + v_2) \overset{(L1)}{=} f(v_1) + f(v_2) \in W' \Rightarrow v_1 + v_2 \in f^{-1}(W')$.
- Seien $\lambda \in K$, $v \in f^{-1}(W')$. $\Rightarrow f(v) \in W'$
  $\Rightarrow f(\lambda v) \overset{(L2)}{=} \lambda \cdot f(v) \in W' \Rightarrow \lambda v \in f^{-1}(W')$.

(v) Sei $(v_i)_{i\in I}$ linear abhängige Familie. $\Rightarrow \exists\, J \subset I$, J endlich und $\lambda_i \in K$ für $i \in J$ nicht alle gleich Null mit $\sum_{i\in J} \lambda_i v_i = 0$.

$$\Rightarrow \quad 0 \overset{(i)}{=} f(0) = f\left(\sum_{i\in J} \lambda_i v_i\right) \overset{(ii)}{=} \sum_{i\in J} \lambda_i f(v_i).$$

$\Rightarrow \big((f(v_i))_{i\in I}\big)$ ist linear abhängig.

(vi) ‘

“$\subset$”: Sei $w \in f(V')$, d.h. $w = f(v)$ für ein $v \in V'$.
$\Rightarrow \exists\, J \subset I$ endlich und $\lambda_i \in K$ für $i \in J$ mit $v = \sum_{i\in J} \lambda_i v_i$.
$\Rightarrow w = f(v) \overset{(ii)}{=} \sum_{i\in J} \lambda_i f(v_i)$, also $w \in \operatorname{Lin}\big((f(v_i))_{i\in I}\big)$.

"$\supset$" Sei $w \in \mathrm{Lin}\big((f(v_i))_{i\in I}\big)$.
$\Rightarrow \exists\, J \subset I$ endlich und $\lambda_i \in K$ für $i \in J$ mit

$$w = \sum_{i\in J} \lambda_i f(v_i) \overset{(ii)}{=} f\Big(\sum_{i\in J} \lambda_i v_i\Big) \in f(V').$$

(vii) Nach (iii) ist $f(V)$ UVR von W. Die Behauptung ist dann gerade Aussage (ii) aus Korollar 3.34. □

Lemma 4.4. *Seien $f : V \to W$, $g : U \to V$ lineare Abbildungen. Dann ist auch $f \circ g : U \to W$ eine lineare Abbildung.*

Beweis. Es sind (L1) und (L2) zu verifizieren. Seien $u_1, u_2 \in U$. Dann folgt

$$(f \circ g)(u_1 + u_2) = f\big(g(u_1 + u_2)\big) = f\big(g(u_1) + g(u_2)\big) = (f \circ g)(u_1) + (f \circ g)(u_2),$$

also (L1). Analog zeigt man (L2), d.h. $(f \circ g)(\lambda u) = \lambda(f \circ g)(u)$ für $\lambda \in K$. □

Definition 4.5. *Man bezeichnet*

- *eine lineare Abbildung $f : V \to W$ als <u>Homomorphismus</u> und setzt*

$$\mathrm{Hom}_K(V, W) := \big\{f : V \to W \,\big|\, f \text{ ist } K\text{-linear}\big\},$$

- *eine lineare Abbildung $f : V \to V$ als <u>Endomorphismus</u> und setzt*

$$\mathrm{End}_K(V) := \big\{f : V \to V \,\big|\, f \text{ ist } K\text{-linear}\big\}.$$

Lemma 4.6. *(i) $\mathrm{Hom}_k(V, W)$ ist bzgl.*
der Addition $(f, g) \mapsto f + g$ mit $(f + g)(v) = f(v) + g(v)\ \forall\, v \in V$ und
der skalaren Multiplikation $(\lambda, f) \mapsto \lambda \cdot f$ mit $(\lambda \cdot f)(v) = \lambda\big(f(v)\big)\ \forall\, v \in V$
ein Vektorraum.

(i) $\mathrm{End}_K(V)$ ist bzgl. $+ : (f, g) \mapsto f+g$ und $\circ : (f, g) \mapsto f\circ g$ ein Ring mit Einselement id_V.

Beweis. Nachrechnen! □

Definition 4.7. *Man bezeichnet*

- *eine bijektive lineare Abbildung $f : V \to W$ als <u>Isomorphismus</u> und setzt*

$$\mathrm{Iso}_K(V, W) := \big\{f : V \to W \,\big|\, f \text{ ist Isomorphismus}\big\},$$

- *eine bijektive lineare Abbildung $f : V \to V$ als Automorphismus und setzt*

$$\mathrm{End}_K(V) := \{f : V \to V \mid f \text{ ist Automorphismus}\}.$$

Existiert zwischen Vektorräumen V und W ein Isomorphismus, so heißen V und W isomorph (Notation $V \cong W$).

Lemma 4.8. *Sei $f : V \to W$ eine lineare Abbildung. Dann gilt*

$$f \text{ Isomorphismus} \quad \Rightarrow \quad f^{-1} \text{ Isomorphismus.}$$

Beweis. Analog zum Beweis bei Gruppen (Lemma 2.16 (iii)). □

4.1 Bild, Kern und Dimensionsformel

Definition 4.9. *Sei $f : V \to W$ lineare Abbildung. Dann heißen*

- $\mathrm{Bild}\, f := f(V) = \{w \in W \mid \exists\, v \in V \text{ mit } f(v) = w\}$ *das Bild von f,*
- $\mathrm{Kern}\, f := f^{-1}(\{0\}) = \{v \in V \mid f(v) = 0\}$ *der Kern von f.*

Satz 4.10. *Seien $f : V \to W$ lineare Abbildungen. Dann gelten folgende Aussagen.*

(i) $\mathrm{Bild}\, f \subset W$ *und* $\mathrm{Kern}\, f \subset V$ *sind UVRs.*

(ii) f *surjektiv* $\Leftrightarrow \mathrm{Bild}\, f = W$.

(iii) f *injektiv* $\Leftrightarrow \mathrm{Kern}\, f = \{0\}$.

(iv) f *injektiv und* $(v_i)_{i\in I}$ *linear unabhängige Familie in* $V \Rightarrow \big(f(v_i)\big)_{i\in I}$ *ist linear unabhängig.*

Beweis. (i) folgt aus Lemma 4.3 (iii) und (iv).

(ii) ist gerade die Definition von Surjektivität.

(iii)

"$\Rightarrow$": Sei f injektiv. Zu zeigen: $\mathrm{Kern}\, f = \{0\}$.

"$\supset$": $f(0) = 0 \Rightarrow 0 \in \mathrm{Kern}\, f$.

"$\subset$": Sei $u \in \mathrm{Kern}\, f \Rightarrow f(u) = 0 = f(0) \overset{f \text{ injektiv}}{\Rightarrow} u = 0$.

"$\Leftarrow$": Gelte nun $\mathrm{Kern}\, f = 0$. Seien $u, v \in V$ mit $f(u) = f(v)$. $\overset{f \text{ linear}}{\Rightarrow} f(u - v) = 0 \overset{\mathrm{Kern}\, f = \{0\}}{\Rightarrow} u - v = 0$, d.h. f ist injektiv.

(iv) Seien $J \subset I$, J endlich und $\lambda_j \in K$ $(j \in J)$ mit

$$\sum_{j \in J} \lambda_j f(v_j) = 0. \tag{4.1}$$

Zu zeigen: $\lambda_j = 0 \; \forall \, j \in J$. Wegen

$$0 = \sum_{j \in J} \lambda_j f(v_j) \overset{\text{Lemma 4.3}(ii)}{=} f\Big(\sum_{j \in J} \lambda_j v_j \Big)$$

folgt aus (iii) aber $\sum_{j \in J} \lambda_j v_j = 0 \overset{(v_i)_{i \in I} \text{ lin. unabh}}{\Rightarrow} \lambda_j = 0 \; \forall \, j \in J.$ □

Definition 4.11. *Sei $f : V \to W$ eine lineare Abbildung. Dann heißt*

$$\operatorname{Rang} f := \dim \big(\operatorname{Bild}(f) \big)$$

der Rang von f.

Beispiel und Definition 4.12 (Matrizen)**.** *Wir betrachten wie in Beispiel 4.2 (i) die zu $A \in M(m \times n, K)$ gehörende Abbildung $\tilde{A} : K^n \to K^m$, $x \mapsto Ax$. Wegen $K^n = \operatorname{Lin}(e_1, \ldots, e_n)$ folgt aus Lemma 4.3 (vi)*

$$\begin{array}{ccccc} \operatorname{Bild} \tilde{A} & = & \operatorname{Lin} \big(\tilde{A}(e_1), \ldots, \tilde{A}(e_n) \big). \\ & & \| & & \| \\ & & Ae_1 & & Ae_n \end{array}$$

$Ae_1, \ldots, Ae_n$ sind aber genau die Spalten von A, d.h.

$$\operatorname{Rang} \tilde{A} = \dim \big(\operatorname{Bild} \tilde{A} \big) = \dim \big(SR(A) \big) = \; \textit{Spaltenrang}\,(A).$$

Wir definieren den Rang der Matrix A durch

$$\operatorname{Rang}(A) := \operatorname{Rang} \tilde{A} = \; \textit{Spaltenrang}\,(A).$$

Satz 4.13 (Dimensionsformel für lineare Abbildungen)**.** *Seien V endlichdimensionaler K-VR und $f : V \to W$ eine lineare Abbildung. Sei ferner $(v_1, \ldots, v_k)$ eine Basis von* Kern f*, $(w_1, \ldots, w_l)$ eine Basis von* Bild f*. Für $i = 1, \ldots, l$ seien $u_i \in V$ mit $f(u_i) = w_i$. Dann gilt: $B = (v_1, \ldots, v_k, u_1, \ldots, u_l)$ ist eine Basis von V und*

$$\dim V = \dim \big(\operatorname{Kern} f \big) + \dim \big(\operatorname{Bild} f \big).$$

Beweis. Wir zeigen (i) B ist Erzeugendensystem von V und (ii) B ist linear unabhängig.

(i) Sei $v \in V$. $\Rightarrow f(v) \in \text{Bild}\, f$. $\Rightarrow \exists\, \mu_1, \ldots, \mu_l \in K$ mit

$$f(v) = \sum_{j=1}^{l} \mu_j w_j.$$

Sei nun $u = \sum_{j=1}^{l} \mu_j u_j$.

$$\Rightarrow \quad f(u) = \sum_{j=1}^{l} \mu_j f(u_j) = \sum_{j=1}^{l} \mu_j w_j = f(v).$$

$\Rightarrow f(u - v) = 0 \Rightarrow v - u \in \text{Kern}\, f$.
$\Rightarrow \exists\, \lambda_1, \ldots, \lambda_k \in K$ mit $v - u = \sum_{j=1}^{k} \lambda_j v_j$.

$$\Rightarrow \quad v = \sum_{j=1}^{k} \lambda_j v_j + \sum_{j=1}^{l} \mu_j u_j.$$

(ii) Seien $\mu_1, \ldots, \mu_l, \lambda_1, \ldots, \lambda_k \in K$ mit

$$\sum_{j=1}^{k} \lambda_j v_j + \sum_{j=1}^{l} \mu_j u_j = 0.$$

Anwendung von f auf beiden Seiten der Identität ergibt nach Lemma 4.3 (i) – (ii)

$$\sum_{j=1}^{k} \lambda_j \underbrace{f(v_j)}_{=0} + \sum_{j=1}^{l} \mu_j \underbrace{f(u_j)}_{=w_j} = 0.$$

$\overset{(w_1,\ldots,w_l) \text{ Basis}}{\Rightarrow} \mu_1 = \cdots = \mu_l = 0.$
$\Rightarrow \sum_{j=1}^{k} \lambda_j v_j = 0.$
$\overset{(v_1,\ldots,v_k) \text{ Basis}}{\Rightarrow} \lambda_1 = \cdots = \lambda_k = 0.$

Damit ist B Basis von $V \Rightarrow \dim V = k + l = \dim(\text{Kern}\, f) + \dim(\text{Bild}\, f)$. □

Korollar 4.14. *Seien V und W endlichdimensionale K-Vektorräume. Dann sind äquivalent:*

(i) $V \cong W$

(ii) $\dim V = \dim W$.

Beweis. (i) $\Rightarrow$ (ii): Sei $V \cong W$, d.h. es existiert ein Isomorphismus $f : V \to W$. Sei $(v_1, \ldots, v_r)$ eine Basis von V. Da f injektiv ist, folgt aus Satz 4.10 (iv), dass $\big(f(v_1), \ldots, f(v_r)\big)$

linear unabhängig ist. Die Surjektivität von f ergibt damit

$$W = \operatorname{Bild} f = \operatorname{Lin}\big(f(v_1), \ldots, f(v_r)\big),$$

d.h. $\big(f(v_1), \ldots, f(v_r)\big)$ ist ein Erzeugendensystem und damit insgesamt eine Basis.
$\Rightarrow \dim W = r = \dim V$.

(ii) $\Rightarrow$ (i): Sei $\dim V = \dim W = r \in \mathbb{N}$. Seien $(v_1, \ldots, v_r)$ Basis von V und $(w_1, \ldots, w_r)$ Basis von W. Wir definieren

$$\begin{aligned} f : V &\to W \\ v = \sum_{j=1}^{r} \lambda_j v_j &\mapsto \sum_{j=1}^{r} \lambda_j w_j. \end{aligned}$$

- f ist wohldefiniert, da $(v_1, \ldots, v_r)$ eine Basis von V ist und damit jedes $v \in V$ eine eindeutige Darstellung als Linearkombination aus $v_1, \ldots, v_r$ besitzt.
- f ist linear (Nachrechnen!).
- Es gilt: $\operatorname{Bild} f = \operatorname{Lin}(w_1, \ldots, w_r) = W$, d.h. f ist surjektiv.
- f ist injektiv, denn aus

$$\underbrace{\dim V}_{=r} = \dim\big(\operatorname{Kern} f\big) + \underbrace{\dim\big(\underbrace{\operatorname{Bild} f}_{W}\big)}_{=r}$$

folgt $\dim\big(\operatorname{Kern} f\big) = 0$, d.h. $\operatorname{Kern} f = \{0\}$.

□

Korollar 4.15. *(i) Seien $n, m \in N_0$. Dann gilt: $K^n \cong K^m \Leftrightarrow n = m$.*

(ii) Ist V ein endlichdimensionaler K-Vektorraum, so gilt $V \cong K^n$ mit $n := \dim V$.

Korollar 4.16. *Seien V und W endlichdimensionale K-Vektorräume mit $\dim V = \dim W$ und $f : V \to W$ linear. Dann sind äquivalent:*

(i) f ist injektiv,

(ii) f ist surjektiv.

(iii) f ist bijektiv.

Beweis. (i) $\Rightarrow$ (ii): f injektiv $\Rightarrow \operatorname{Kern} f = \{0\}$, d.h. $\dim\big(\operatorname{Kern} f\big) = 0$.

$$\Rightarrow \dim\big(\operatorname{Bild} f\big) \overset{\text{Satz } 4.13}{=} \dim V = \dim W \Rightarrow \operatorname{Bild} f = W \Rightarrow f \text{ surjektiv.}$$

(ii) $\Rightarrow$ (iii): Sei f surjektiv. $\Rightarrow \dim\big(\operatorname{Kern} f\big) = \dim(V) - \dim\big(\underbrace{\operatorname{Bild} f}_{=W}\big) = 0.$

$$\Rightarrow \operatorname{Kern} f = \{0\} \overset{\text{Satz } 4.10(iii)}{\Rightarrow} f \text{ injektiv.}$$

Insgesamt ist f also bijektiv.

(iii) $\Rightarrow$ (i): Klar. □

4.2 Affine Unterräume

Um die Urbilder $f^{-1}(\{w\})$ genauer untersuchen zu können, brauchen wir den Begriff des affinen Unterraums.

Definition 4.17. *Sei V ein K-VR. $Z \subset V$ heißt affiner Unterraum von V, wenn es ein $z \in V$ und einen UVR $U \subset V$ gibt mit*

$$Z = z + U := \{z + u \,|\, u \in U\},$$

d.h. affine Unterräume entstehen durch "Parallelverschiebung" von UVRs.

Beachte: Ist $z \notin U$, dann ist $0 \notin z + U$, d.h. in diesem Falle ist Z kein UVR von V.

Lemma 4.18. *Seien $z \in V$ und $U \subset V$ UVR, ferner $Z = z + U$. Dann gilt:*

(i) Für jedes $z' \in Z$ ist $Z = z' + U$.

(ii) Sind $\tilde{z} \in V$ und $\tilde{U} \subset V$ UVR mit $\tilde{z} + \tilde{U} = z + U$, dann ist $U = \tilde{U}$ und $z - \tilde{z} \in U$.

Beweis. (i) Sei $z' \in Z$, d.h. $z' = z + u'$ mit $u' \in U$.

$$\Rightarrow \quad z' + U = z + \underbrace{\big(u' + U\big)}_{=U} = z + U = Z.$$

(ii) Es gilt $U = \{y_1 - y_2 \,|\, y_1, y_2 \in z + U\}$, womit

$$U = \big\{y_1 - y_2 \,\big|\, y_1, y_2 \in z + U\big\} = \big\{y_1 - y_2 \,\big|\, y_1, y_2 \in \tilde{z} + \tilde{U}\big\} = \tilde{U}.$$

$\Rightarrow z + U = \tilde{z} + U \Rightarrow z \in z + U = \tilde{z} + U \Rightarrow z - \tilde{z} \in U.$ □

Konsequenz und Definition. Bei einem affinen Unterraum $Z = z + U$ ist der UVR eindeutig bestimmt, der "Verschiebungspunkt" z kann beliebig aus Z gewählt werden. Da U eindeutig ist, können wir die Dimension definieren durch

$$\dim Z := \dim U.$$

Beispiel 4.19. *Wir vergleichen UVRs und affine Unterräume in* $\mathbb{R}^2$.

UVRs:

$\dim U = 0:\ \{0\}$
$\dim U = 1:\ \mathrm{Lin}(v), 0 \neq v \in \mathbb{R}^2$ *(Geraden durch den Ursprung)*
$\dim U = 2:\ \mathbb{R}^2$.

Affine Unterräume:

$\dim Z = 0:\ \{z\}, z \in \mathbb{R}^2$ *(Punkte)*
$\dim Z = 1:\ z + \mathrm{Lin}(v), 0 \neq v \in \mathbb{R}^2$ *(verschobene Geraden)*
$\dim Z = 2:\ \mathbb{R}^2$.

Lemma 4.20. *Sei* $f : V \to W$ *eine lineare Abbildung,* $w \in \mathrm{Bild}\, f$ *und* $v \in f^{-1}(\{w\})$. *Dann gilt:*

$$f^{-1}(\{w\}) = v + \mathrm{Kern}\, f \quad \textit{und} \quad \dim f^{-1}(\{w\}) = \dim V - \dim\big(\mathrm{Bild}\, f\big).$$

Beweis. "$\subset$": Sei $u \in f^{-1}(\{w\})$. $\Rightarrow f(u) = w = f(v) \Rightarrow f(u) - f(v) = f(u - v) = 0$
$\Rightarrow u - v \in \mathrm{Kern}\, f \Rightarrow u \in v + \mathrm{Kern}\, f$.

"$\supset$": Sei $u \in v + \mathrm{Kern}\, f \Rightarrow \exists\, x \in \mathrm{Kern}\, f$: $u = v + x$

$$\Rightarrow \quad f(u) = f(v + x) = f(v) + \underbrace{f(x)}_{=0} = f(v) = w,$$

da $v \in f^{-1}(\{w\})$. Aber damit ist $u \in f^{-1}(\{w\})$.
Schließlich gilt $\dim f^{-1}(\{w\}) = \dim\big(\mathrm{Kern}\, f\big) \overset{\text{Satz } 4.13}{=} \dim V - \dim\big(\mathrm{Bild}\, f\big)$. □

4.3 Lineare Gleichungssysteme

In diesem Abschnitt seien $A = (a_{ij}) \in M(m \times n, K)$,

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m.$$

Wir untersuchen das lineare Gleichungssystem (LGS)

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & \vdots & & & & \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m, \end{array}$$

in Matrix-Schreibweise $Ax = b$. Wir gehabt induziert die Matrix die lineare Abbildung

$$\tilde{A} : K^n \to K^m, \quad x \mapsto A \cdot x.$$

Wir betrachten das Problem der Lösung zunächst theoretisch und dann algorithmisch.

Definition 4.21. *Das LGS $Ax = b$ heißt*

homogen, wenn $b = 0$,

inhomogen, wenn $b \neq 0$.

Das LGS $Ax = 0$ heißt das zu $Ax = b$ gehörende homogene LGS. A heißt Koeffizientenmatrix.

$$\text{Lös}(A, b) := \{x \in K^n | \, Ax = b\} = \tilde{A}^{-1}(\{b\})$$

heißt Lösungsraum des LGS $Ax = b$.

Es gilt

$$\text{Lös}(A, 0) = \{x \in K^n | \, Ax = 0\} = \text{Kern}\, \tilde{A}.$$

Satz 4.22. *Es gilt:*

(i) $\text{Lös}(A, 0) \subset K^n$ *ist ein UVR der Dimension* $n - \text{Rang}(A)$.

(ii) $\text{Lös}(A, b) \subset K^n$ *ist ein affiner Unterraum von* K^n. *Ist* $\text{Lös}(A, b) \neq \emptyset$, *dann hat dieser die Dimension* $n - \text{Rang}(A)$.

(iii) Sind $\text{Lös}(A, b) \neq \emptyset$ *und* $v \in \text{Lös}(A, b)$, *dann ist*

$$\text{Lös}(A, b) = v + \text{Lös}(A, 0).$$

Bemerkung. *(iii) bedeutet: Falls eine Lösung v von $Ax = b$ existiert, dann erhält man alle Lösungen, indem man zu der speziellen Lösung v alle Lösungen des zugehörigen homogenen Gleichungssystems addiert.*

Beweis. (i) Es ist $\text{Lös}(A, 0) = \text{Kern}\, \tilde{A}$. $\text{Kern}\, \tilde{A}$ ist ein $UVR \subset K^n$ mit

$$\dim\left(\text{Kern}\, \tilde{A}\right) \overset{\text{Dimensionsformel}}{=} \dim K^n - \dim\left(\text{Bild}\, \tilde{A}\right) = n - \text{Rang}\, A.$$

(ii) Es ist $\text{Lös}(A,b) = \tilde{A}^{-1}(\{b\})$ nach Lemma 4.20 ein affiner Unterraum von K^n. Falls $\text{Lös}(A,b) \neq \emptyset$, dann gilt $b \in \text{Bild}\,\tilde{A}$ und

$$\dim\big(\text{Lös}(A,b)\big) = \dim \tilde{A}^{-1}(\{b\}) = \dim\big(\text{Kern}\,\tilde{A}\big) = n - \text{Rang}\,A.$$

(iii) Mit $v \in \text{Lös}(A,b)$ gilt nach Lemma 4.20

$$\text{Lös}(A,b) = \tilde{A}^{-1}(\{b\}) = v + \text{Kern}\,\tilde{A} = v + \text{Lös}(A,0).$$

□

Bemerkung. $\text{Lös}(A,0)$ *enthält immer die triviale Lösung* 0*; nicht-triviale Lösungen von* $Ax = 0$ *existieren wegen (i) genau dann, wenn* $\text{Rang}\,A < n$.

Definition 4.23.

$$A|b := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \in M(m \times (n+1), K)$$

heißt erweiterte Koeffizientenmatrix des LGS $Ax = b$.

Satz 4.24. *Es sind äquivalent:*

(i) $\text{Lös}(A,b) \neq \emptyset$*, d.h. das LGS* $Ax = b$ *besitzt eine Lösung.*

(ii) $\text{Rang}\,A = \text{Rang}(A|b)$.

Beweis. Es gilt $\text{Bild}\,\tilde{A} = SR(A) \subset SR(A|b) = \text{Bild}\,\widetilde{A|b}$. Damit erhalten wir

$$\begin{aligned} \text{Lös}(A,b) \neq \emptyset &\Leftrightarrow b \in \text{Bild}\,\tilde{A} \\ &\Leftrightarrow \dim \text{Bild}\,\tilde{A} = \dim \text{Bild}\,\widetilde{A|b} \\ &\Leftrightarrow \text{Rang}\,\tilde{A} = \text{Rang}\,\widetilde{A|b} \\ &\Leftrightarrow \text{Rang}\,A = \text{Rang}\,A|b. \end{aligned}$$

□

Korollar 4.25. *Es sind äquivalent:*

(i) Das LGS $Ax = b$ *hat genaue eine Lösung.*

(ii) $\text{Rang}\,A = \text{Rang}(A|b) = n$.

Beweis. "(i)⇒(ii)": Nach Satz 5.4 folgt aus der Existenz der Lösung Rang $A = \text{Rang}(A|b)$. Aus der Eindeutigkeit der Lösung folgt $\dim\big(\text{Lös}(A,b)\big) = 0$, also

$$0 = \dim\big(\text{Lös}(A,b)\big) = n - \text{Rang}\, A,$$

d.h. Rang $A = n$.

"(ii)⇒(i)": Sei Rang $A = \text{Rang}(A|b) = n$. Nach Satz 5.4 ist dann $\text{Lös}(A,b) \neq \emptyset$. Nach Satz 4.22 (ii) gilt $\dim\big(\text{Lös}(A,b)\big) = n - \text{Rang}\, A = 0$, d.h. die Lösung ist eindeutig. □

Wir möchten schließlich noch einen Algorithmus zur Bestimmung der Lösungen $\text{Lös}(A,b)$ angeben.

Definition 4.26. *Sei $A \in M(m \times n, K)$. Man sagt, A sei in strenger Zeilenstufenform (SZSF), wenn A in ZSF mit Pivotspalten an $j_1, \ldots, j_r$ ist und wenn gilt:*

(i) $a_{1j_1} = \cdots = a_{rj_r} = 1$ *und*

(ii) $a_{ij_k} = 0$ *für alle* $i \in \{1, \ldots, k-1\}$ *und* $k \in \{1, \ldots, r\}$.

Satz 4.27. *A lässt sich durch elementare Zeilenumformungen auf SZSF bringen.*

Beweis. A lässt sich mit dem Gaußschen Eliminationsverfahren aus Abschnitt 3.4.1 durch elementare Zeilenumformungen auf ZSF

$$B = \left(\begin{array}{ccccccccccc}
0 & \ldots & 0 & b_{1j_1} & * & & \ldots & & & * \\
0 & & \ldots & & 0 & b_{2j_2} & * & & \ldots & * \\
 & & & & \ddots & & & & & \\
0 & & & \ldots & & & & 0 & b_{rj_r} & * \\
\hline
0 & & & 0 & \ldots & 0 & & & & 0
\end{array}\right)$$

bringen, d.h. man erhält aus A damit B. Multipliziere nun die i-te Zeile mit $1/b_{ij_i}$ und annuliere dann die Einträge der j_k-ten Spalte oberhalb des Pivot-Elements durch Subtraktion geeigneter Vielfache der k-ten Zeile. □

Lemma 4.28. *Sei $C \in M(m \times n, K)$, $d \in K^n$. Ist $C|d$ durch eine Folge von elementaren Zeilenumformungen aus $A|b$ entstanden, so ist* $\text{Lös}(C,d) = \text{Lös}(A,b)$.

Beweis. Wegen Bemerkung 3.53 reicht es, Umformungen vom Typ (I) und (II) zu betrachten.

Typ I: Sei $C|d$ durch $A|b$ dadurch entstanden, dass die j-te Zeile mit $\lambda \in K \setminus \{0\}$

multipliziert wurde, d.h.

$$A|b := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \quad \text{und} \quad C|d := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ \lambda a_{j1} & \dots & \lambda a_{jn} & \lambda b_j \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}.$$

so gilt: $x \in \text{Lös}(A|b) \;\Leftrightarrow\; Ax = b \;\Leftrightarrow Cx = d \;\Leftrightarrow\; x \in \text{Lös}(C, d)$.

Typ II: Analog. □

Beispiel 4.29. *Gegeben sei das LGS $Ax = b$ mit*

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 & 3 \\ 3 & 6 & -3 & 3 & 9 \\ 4 & 8 & -4 & 5 & 9 \end{pmatrix} \quad \textit{und einem Vektor } b \in \mathbb{R}^3.$$

Betrachte zunächst das zugehörige homogene LGS. Mit elementaren Zeilenumformungen erhält man

$$A \to \begin{pmatrix} 1 & 2 & -1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 1 & 3 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 0 & 6 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} := S.$$

Setze $B = \left(\begin{smallmatrix} 2 & -1 & 6 \\ 0 & 0 & -3 \end{smallmatrix}\right)$. Es gilt $Sx = 0$ genau dann, wenn $\binom{x_1}{x_4} = -B \begin{pmatrix} x_2 \\ x_3 \\ x_5 \end{pmatrix}$.

x_2, x_3, x_5 können beliebig dann gewählt werden. Der Lösungsraum $\text{Lös}(A, 0)$ des homogenen Systems ist ein UVR (Satz 4.22 (i)), der am besten durch die Angabe einer Basis beschrieben werden kann. Hierzu gibt es natürlich viele Möglickeiten; eine einfache ist, für $(x_2\ x_3\ x_5)'$ die Einheitsvektoren e_1, e_2, e_3 des $\mathbb{R}^3$ zu wählen. Für e_i gilt dann

$$\binom{x_1}{x_4} = -Be_i = -\ i\text{-te Spalte von } B,$$

d.h. wir erhalten durch Einsetzen von e_1, e_2, e_3 folgende Basisvektoren von $\text{Lös}(A, 0)$:

$$w_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad w_3 = \begin{pmatrix} -6 \\ 0 \\ 0 \\ 3 \\ 1 \end{pmatrix}.$$

Algorithmus (Gaußalgorithmus zur Lösung für ein homogenes LGS)

Eingabe: $A \in M(m \times n, K)$

Ausgabe: Basis von $\text{Lös}(A, 0)$.

1. Schritt: Bringe die Matrix durch elementare Zeilenumformungen in SZSF

$$S = \left(\begin{array}{ccccccccccc} 0 & \dots & 0 & 1 & * & 0 & \dots & & 0 & * \\ 0 & & \dots & & 0 & 1 & * & \dots & 0 & * \\ & & & & \ddots & & & & & \\ 0 & & & \dots & & & & 0 & 1 & * \\ \hline 0 & & & 0 & \dots & 0 & & & & 0 \\ & & & j_1 & & j_2 & \cdots & & j_r & \end{array}\right)$$

$r = \text{Zeilenrang}(A)$

2. Schritt: Sei $B \in M(r \times (n-r), K)$ die Matrix, die durch Streichen der Spalten mit den Indizes $j_1, \dots, j_r$ und den Zeilen mit den Indizes $r+1, \dots, n$ entsteht (diese beinhalten nur Nullen als Einträge). Seien $k_1 < k_2 < \cdots < k_{n-r}$ mit $\{1, \dots, n\} = \{j_1, \dots, j_r, k_1, \dots, k_{n-r}\}$ (d.h. die k_i sind die Indizes von Nicht-Pivot-Spalten).

3. Schritt: Eine Basis von $\text{Lös}(A, 0)$ ist gegeben durch $w_1, \dots, w_{n-r} \in K^n$, wobei

$$w_i = \begin{pmatrix} w_{i1} \\ \vdots \\ w_{in} \end{pmatrix}$$

gegeben ist durch

$$\begin{pmatrix} w_{ij_1} \\ \vdots \\ w_{ij_r} \end{pmatrix} = i\text{-te Spalte von } -B, \quad \begin{pmatrix} w_{ik_1} \\ \vdots \\ w_{ik_{n-r}} \end{pmatrix} = e_i \in K^{n-r}.$$

Beweis. Nach Lemma 4.28 gilt $\text{Lös}(A, 0) = \text{Lös}(S, 0)$ und ferner $x \in \text{Lös}(S, 0)$

$$\Leftrightarrow \quad Sx = 0 \quad \Leftrightarrow \quad 0 = \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_r} \end{pmatrix} + B \begin{pmatrix} x_{k_1} \\ \vdots \\ x_{k_{n-r}} \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_r} \end{pmatrix} = -B \begin{pmatrix} x_{k_1} \\ \vdots \\ x_{k_{n-r}} \end{pmatrix}.$$

Nach beliebiger Vorgabe von $x_{k_1}, \dots, x_{k_{n-r}}$ ergeben sich $x_{j_1}, \dots, x_{j_r}$ eindeutig, wenn $x \in K^n$ eine Lösung ist. Um eine Basis des Lösungsraums $\text{Lös}(A, 0) \subset K^n$ zu erhalten,

kann man mit $(x_{k_1}, \ldots, x_{k_{n-r}})'$ eine Basis des K^{n-r} durchlaufen, also bspw.

$$\begin{pmatrix} x_{k_1} \\ \vdots \\ x_{k_{n-r}} \end{pmatrix} = e_i \in K^{n-r} \quad \text{für } i = 1, \ldots, n-r.$$

Dann ist

$$\begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_r} \end{pmatrix} = i\text{-te Spalte von } -B,$$

d.h. wir erhalten obige Vektoren $w_1, \ldots, w_{n-r}$. Nach Konstruktion ist $(w_1, \ldots, w_{n-r})$ ein Erzeugendensystem von $\text{Lös}(S, 0) = \text{Lös}(A, 0)$. Die Familie $(w_1, \ldots, w_{n-r})$ ist aber auch linear unabhängig: Denn sind $\lambda_1, \ldots, \lambda_{n-r} \in K$ mit

$$\lambda_1 w_1 + \cdots + \lambda_{n-r} w_{n-r} = 0,$$

so lautet der Eintrag der k_i-ten Zeile ($i \in \{1, \ldots, n-r\}$)

$$\underbrace{\lambda_1 \underbrace{w_{1k_i}}_{=0} + \ldots + \lambda_i \underbrace{w_{ik_i}}_{=1}}_{=0} + \underbrace{\cdots + \lambda_{n-r} \underbrace{w_{n-r,k_i}}_{=0}}_{=0} = 0$$

$\Rightarrow \lambda_i = 0$ für $i = 1, \ldots, n-r$. □

Satz 4.30 (Korollar aus dem Gaußalgorithmus zur Lösung homogener LGS)**.** *Es sei* $A \in M(m \times n, K)$*. Dann gilt*

$$\textit{Zeilenrang}(A) = \textit{Spaltenrang}(A) = \text{Rang } A.$$

Beweis. Im obigen Algorithmus haben wir gezeigt

$$\dim\big(\text{Lös}(A, 0)\big) = n - \text{Zeilenrang}(A) = n - r.$$

Nach Satz 4.22 (i) ist

$$\dim\big(\text{Lös}(A, 0)\big) = n - \text{Spaltenrang}(A) = \text{Rang } A,$$

also folgt die Behauptung. □

Beispiel 4.31 (Fortsetzung von Beispiel 4.29)**.** *Zur Suche einer speziellen Lösung des inhomogenen Systems* $Ax = b$ *bringen wir die Matrix* $A|b$ *in SZSF und erhalten für*

$b = (3, 9, 13)'$ *bzw.* $b = (3, 10, 13)'$ *mit den Umformungen aus Beispiel 4.29*

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 6 & 0 \\ 0 & 0 & 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) := S|z \quad \textit{bzw.} \quad \left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 6 & 2 \\ 0 & 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) := \tilde{z}.$$

Für z *mit* $z_3 = 1$ *ist klar, dass es keine Lösung von* $Sx = z$ *gibt, d.h.* $\text{Lös}(A, b) = \text{Lös}(S, z) = 0$. *Im Falle* $\tilde{z}$ *ist*

$$\begin{array}{ll} j_1 = 1 \ \rightarrow & \\ & \\ & \\ j_2 = 4 \ \rightarrow & \end{array} \quad \begin{pmatrix} \tilde{z}_1 \\ 0 \\ 0 \\ \tilde{z}_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

eine Lösung. Man beachte, dass mit $\tilde{z}$ *gilt* $r = 2$ *und* $j_2 = 4$ *und mit* z *ist* $r = 3$ *und* $j_r = j_3 = 6 = n + 1$.

Algorithmus (Gaußalgorithmus zur Lösung für ein inhomogenes LGS)

Eingabe: $A \in M(m \times n, K)$, $b \in K^m$.

Ausgabe: Affiner Unterraum $\text{Lös}(A, b)$.

1. Schritt: Bringe die Matrix $A|b$ durch elementare Zeilenumformungen auf SZSF $S|z$. Seien $r = \text{Rang}(A|b) = \text{Rang}(S|z)$ und $j_1, \ldots, j_r$ die Positionen der Pivot-Spalten.

2. Schritt: Falls $j_r = n + 1$, dann gilt $\text{Lös}(A, b) = \emptyset$.

3. Schritt: Ist $j_r < n + 1$, dann gilt $\text{Lös}(A, b) = v + \text{Lös}(A, 0)$ mit einem $v \in \text{Lös}(A, b)$. Eine Basis von $\text{Lös}(A, 0)$ wird ermittelt wie im zugehörigen Algorithmus für ein homogenes LGS. Eine spezielle Lösung ist gegeben durch $v = (v_1, \ldots, v_n)' \in K^n$, wobei

$$\begin{pmatrix} v_{j_1} \\ \vdots \\ v_{j_r} \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_r \end{pmatrix} \quad \text{und} \quad v_i = 0 \text{ für } i \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}.$$

Beweis. Gilt $j_r = n + 1$ ist klar, dass $\text{Lös}(A, b) = \text{Lös}(S, z) = \emptyset$ ist, denn die j_r-te Gleichung des Gleichungssystems lautet $0 = 1$. Im Falle $j_r < n + 1$ ist v aus Schritt 3 eine spezielle Lösung von $Sx = z$, d.h. auch von $Ax = b$. □

4.4 Darstellende Matrizen

In Beispiel 4.2 war $\tilde{A} : K^n \to K^m$, $x \mapsto Ax$ mit $A \in M(m \times n, K)$, ein Beispiel für eine lineare Abbildung. In diesem Abschnitt wird gezeigt, dass man jede lineare Abbildung praktisch so schreiben kann. Nachfolgend seien V und W endlichdimensionale K-VRs.

Lemma 4.32. *Seien $v_1, \ldots, v_r \in V$ und $w_1, \ldots, w_r \in W$. Dann gelten:*

(i) Ist $(v_1, \ldots, v_r)$ eine linear unabhängige Familie, dann gibt es eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, r$.

(ii) Ist $(v_1, \ldots, v_r)$ Basis von V, dann gibt es genau eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, r$. Diese hat folgende Eigenschaften:

(a) $\operatorname{Bild} f = \operatorname{Lin}(w_1, \ldots, w_r)$. Insbesondere gilt
f ist surjektiv $\Leftrightarrow$ $(w_1, \ldots, w_r)$ ist ein Erzeugendensystem von W.

(b) f ist injektiv $\Leftrightarrow$ $(w_1, \ldots, w_r)$ ist linear unabhängig.

Beweis. (i) folgt aus (ii) mit dem Basisergänzungssatz.

(ii) Da $(v_1, \ldots, v_r)$ Basis ist, ist die Darstellung eines jeden Vektors $v \in V$ als Linearkombination aus $v_1, \ldots, v_r$ eindeutig. Somit ist folgende Abbildung $f : V \to W$ wohldefiniert:

$$v = \sum_{i=1}^{r} \lambda_i v_i \mapsto \sum_{i=1}^{r} \lambda_i w_i.$$

Es gilt $f(v_i) = w_i$ für $i = 1, \ldots, r$. Außerdem ist f linear, denn mit $u = \sum_{i=1}^{r} \mu_i v_i$ ist

$$f(u + v) = f\Big(\sum_{i=1}^{r} (\lambda_i + \mu_i) v_i \Big) = \sum_{i=1}^{r} (\lambda_i + \mu_i) w_i = f(u) + f(v),$$

analog gilt $f(\lambda v) = \lambda \cdot f(v)$ für $\lambda \in K$. [Bemerkung: Diese Konstruktion wurde auch schon einmal im Beweis von Korollar 4.14 verwendet.]

Eindeutigkeit von f: Sei $g : V \to W$ eine weitere lineare Abbildung mit $g(v_i) = w_i$. Für $v = \sum_{i=1}^{r} \lambda_i w_i$ gilt dann $g(v) = \sum_{i=1}^{r} \lambda_i g(v_i) = \sum_{i=1}^{r} \lambda_i w_i = f(v)$, d.h. $f = g$.

(a) $\operatorname{Bild} f = \operatorname{Lin}(w_1, \ldots, w_r)$ folgt aus Aussage (vi) aus Lemma 4.3.

(b) Es gilt: f injektiv $\overset{\text{Satz } 4.10(iii)}{\Leftrightarrow}$ $\operatorname{Kern} f = \{0\}$ $\overset{\text{Satz } 4.10(iv)}{\Rightarrow}$ $(\underbrace{f(v_1)}_{=w_1}, \ldots, \underbrace{f(v_r)}_{=w_r})$ linear unabh.

Umgekehrt gilt: $(w_1, \ldots, w_r)$ linear unabhängig

$\Rightarrow \dim\big(\operatorname{Bild} f\big) = r$
$\Rightarrow \dim\big(\operatorname{Kern} f\big) = \dim V - \dim\big(\operatorname{Bild} f\big) = 0$
$\Rightarrow \operatorname{Kern} f = \{0\}$, d.h. f ist injektiv.

□

Bemerkung. *Man braucht in Lemma 4.32 die lineare Unabhängigkeit der v_i, damit die Abbildung f wohldefiniert ist. Ist bspw. (v_1, v_2) mit $v_1 = v_2$, kann man nicht einfach $f(v_1) = w_1$ und $f(v_2) = w_2$ für beliebige w_1, w_2 setzen.*

Korollar und Definition 4.33. *Sei $B = (v_1, \dots, v_n)$ eine Basis von V (entspricht hier dem W aus Lemma 4.32). Dann gibt es genau einen Isomorphismus*

$$\Phi_B : K^n \to V \quad \text{mit} \quad \Phi_B(e_i) = v_i \quad \text{für } i = 1, \dots, n.$$

Die Abbildung Φ_B heißt das durch B bestimmte Koordinatensystem von V. Für $v = \sum_{i=1}^{n} \lambda_i v_i$ gilt

$$\Phi_B \left(\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \right) = \sum_{i=1}^{n} \lambda_i v_i, \quad \text{womit} \quad \Phi_B^{-1}(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

D.h. Φ_B^{-1} ordnet dem Vektor $v \in V$ seine Koordinaten bzgl. B zu.

Bemerkung. *Es mag auf den ersten Blick merkwürdig erscheinen, den Begriff "Koordinatensystem" für eine Abbildung zu verwenden. Allerdings ist klar, dass zu einem Koordinatensystem neben den Basisvektoren, welche die Achsen beschreiben, auch eine Methode gehören muss, wie man einen einzelnen Vektor dann darstellt. Das beschreibt gerade die Abbildung Φ_B^{-1}.*

Beispiel 4.34 (Gedrehtes Koordinatensystem)**.** *Wir möchten den Vektor $\binom{1}{3}$ in zwei verschiedenen Koordinatensystemen darstellen.*

- *Einerseits gilt $\binom{1}{3} = 1 \cdot \binom{1}{0} + 3 \cdot \binom{0}{1}$, d.h. die Koordinaten bzgl. $B = \left(\binom{1}{0}, \binom{0}{1}\right)$ sind $\binom{1}{3}$ und $\Phi_B^{-1}\left(\binom{1}{3}\right) = \binom{1}{3}$.*
  Vektor ↑ *Koordinaten* ↑
- *Wir betrachten nun das gedrehte Koordinatensystem mit den Achsen $\binom{1}{1}$ und $\binom{-1}{1}$. Es gilt $\binom{1}{3} = 2 \cdot \binom{1}{1} + 1 \cdot \binom{-1}{1}$, d.h. die Koordinaten bzgl. $B = \left(\binom{1}{1}, \binom{-1}{1}\right)$ sind $\binom{2}{1}$ und $\Phi_B^{-1}\left(\binom{1}{3}\right) = \binom{2}{1}$.*
  Vektor ↑ *Koordinaten* ↑

Im zweiten Fall sind der dargestellte Vektor und die Koordinaten verschieden. Im ersten Fall sind der Vektor und seine Koordinaten gleich. Das liegt daran, dass als Basis die Einheitsvektoren benutzt worden sind, d.h. $B = (e_1, e_2)$ sowie daran, dass wir ein Beispiel aus dem $\mathbb{R}^2$ (allgemeiner K^n) gewählt haben.

Beispiel 4.35 (Polynomraum)**.** *Wir betrachten den* $\mathbb{R}$*-Vektorraum*

$$\mathbb{R}[t]_n = \{P \in \mathbb{R}[t] | \, deg(P) \leq n\} \cup \{0\}$$

der Polynome mit reellen Koeffizienten vom Grad $\leq n$ *inklusive Nullpolynom von Übungsblatt 8, Aufgabe 4.* $B = (P_0, P_1, \ldots, P_n)$ *mit* $P_0(t) = 1$ *und* $P_k(t) = t^k$, $1 \leq k \leq n$, *bildet eine Basis von* $\mathbb{R}[t]_n$*. Die Abbildung* Φ_B *ist nun*

$$\Phi_B : \mathbb{R}^{n+1} \to \mathbb{R}[t]_n, \quad \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \sum_{i=0}^{n} \lambda_i P_i.$$

$\begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_n \end{pmatrix} = \Phi_B^{-1}\left(\sum_{i=0}^{n} \lambda_i P_i\right)$ *sind die Koordinaten des Polynoms* $\sum_{i=0}^{n} \lambda_i P_i$*.*

Man könnte aber auch die Basis $B' = (P_0, P_0 + P_1, \ldots, \sum_{k=0}^{n} P_k)$ *verwenden mit zugehörigem Koordinatensystem* $\Phi_{B'}$ *und entsprechenden Koordinaten* $\Phi_{B'}^{-1}(P)$ *für ein Polynom* $P \in \mathbb{R}[t]_n$*. Z. Bsp. für* $n = 2$ *gilt dann*

$$\Phi_B^{-1}(P_0 + P_1 + P_2) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \textit{und} \quad \Phi_{B'}^{-1}(P_0 + P_1 + P_2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Setzt man für die Unbekannte t *Elemente des Körpers* $\mathbb{R}$ *ein, ist die einem Polynom* $P = \sum_{i=0}^{n} \lambda_i P_i \in \mathbb{R}[t]_n$ *zugeordnete Abbildung*

$$\tilde{P} : \mathbb{R} \to \mathbb{R}, \quad x \mapsto P(x),$$

differenzierbar und $\tilde{P}'(x) = \sum_{i=1}^{n} i\lambda_i x^{i-1}$*. Sei entsprechend*

$$d: \ \mathbb{R}[t]_n \to \mathbb{R}[t]_n, \quad d(P_i) = \begin{cases} i \cdot P_{i-1} & \textit{falls } i = 1, \ldots, n \\ 0 & \textit{falls } i = 0. \end{cases} \qquad (d \textit{ für "derivative"})$$

Damit ist die lineare Abbildung d *nach Lemma 4.32 eindeutig bestimmt. Ist nun* $\sum_{i=0}^{n} \lambda_i P_i$ *mit Koordinaten*

$$\begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_n \end{pmatrix} = \Phi_B^{-1}\left(\sum_{i=0}^{n} \lambda_i P_i\right)$$

gegeben, so gilt

$$d\Big(\sum_{i=0}^{n}\lambda_i P_i\Big)=\sum_{i=1}^{n} i\lambda_i P_{i-1}=\sum_{i=0}^{n-1}(i+1)\lambda_{i+1}P_i,$$

d.h. die Koordinaten der Abbildung sind

$$\begin{pmatrix} 1\cdot\lambda_1\\ \vdots\\ n\cdot\lambda_n\\ 0\end{pmatrix}.$$

Auf "Koordinatenebene" kann man das schreiben als

$$\begin{pmatrix} 1\cdot\lambda_1\\ \vdots\\ n\cdot\lambda_n\\ 0\end{pmatrix}=\underbrace{\begin{pmatrix} 0 & 1 & & 0\\ & & \ddots & \\ & & \ddots & n\\ 0 & & & 0\end{pmatrix}}_{=:A}\cdot\begin{pmatrix}\lambda_0\\ \lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}$$

(Koeffizientenmatrix)

oder als Abbildung $d=\Phi_B\circ\tilde{A}\circ\Phi_B^{-1}$ von $\mathbb{R}[t]_n$ nach $\mathbb{R}[t]_n$ mit der zu A gehörenden linearen Abbildung $\tilde{A}$.

Satz und Definition 4.36. *Seien $B=(v_1,\dots,v_n)$ eine Basis von V und $C=(w_1,\dots,w_m)$ eine Basis von W. Dann gelten folgende Aussagen:*

(i) Für jede lineare Abbildung $f: V\to W$ gibt es genau eine Matrix $A=(a_{ij})$ aus $M(m\times n,K)$ mit

$$f(v_j)=\sum_{i=1}^{m}a_{ij}w_i \quad \text{für alle } j=1,\dots,n.$$

$M_C^B(f):=A$ heißt die Darstellungsmatrix von f bzgl. der Basen B und C (von V bzw. W). In der j-ten Spalte von $M_C^B(f)$ stehen die Koordinaten von $f(v_j)$ bzgl. der Basis C von W (für $j=1,\dots,n$).

(ii) Es gilt dann $f=\Phi_C\circ\widetilde{M_C^B(f)}\circ\Phi_B^{-1}$.

(iii) Die in (i) enthaltene Abbildung

$$M_C^B:\operatorname{Hom}_K(V,W)\to M(m\times n,K),\quad f\mapsto M_C^B(f)$$

ist ein Isomorphismus von K-Vektorräumen.

Insbesondere gilt

$$\dim \mathrm{Hom}_K(V,W) = m \cdot n.$$

Im Falle $V = W$ und $B = C$ ist die Abbildung

$$M_B : \mathrm{End}_K(V) \to M(n \times n, K), \quad f \mapsto M_B^B(f) =: M_B(f)$$

ein Isomorphismus von K-Vektorräumen.

Beweis. (i) folgt aus Lemma 4.32 (ii), da $B = (v_1, \ldots, v_n)$ eine Basis ist. Da $(w_1, \ldots, w_m)$ eine Basis von W ist, kann man die Bilder $f(v_j)$ als Linearkombination daraus darstellen:

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad j = 1, \ldots, n.$$

(ii) Es gilt mit $e_j \in K^n$

$$\begin{aligned}\Big(\Phi_C \circ \widetilde{M_C^B(f)} \circ \Phi_B^{-1}\Big)(v_j) &= \Big(\Phi_C \circ \widetilde{M_C^B(f)}\Big)(e_j) \\ &= \Phi_C\big(Ae_j\big) = \Phi_C \left(\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \right) = \sum_{i=1}^{m} a_{ij} w_i = f(v_j).\end{aligned}$$

Da f eindeutig bestimmt ist, folgt Gleichheit.

(iii) Nachweis der Linearität: Seien $f, g \in \mathrm{Hom}_K(V,W)$ mit $M_C^B(f) = (a_{ij})$, $M_C^B(g) = (b_{ij})$. $\Rightarrow$ $(f+g)(v_j) = f(v_j) + g(v_j) = \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m b_{ij} w_i = \sum_{i=1}^m \big(a_{ij} + b_{ij}\big) w_i$, womit

$$M_C^B(f+g) = \big(a_{ij} + b_{ij}\big) = M_C^B(f) + M_C^B(g)$$

ist. Analog zeigt man $M_C^B(\lambda f) = \lambda M_C^B(f)$.

Nachweis der Bijektivität: Ist $A = (a_{ij}) \in M(n \times n, K)$, so gibt es nach Lemma 4.32 (ii) genau ein $f \in \mathrm{Hom}_K(V,W)$ mit $f(v_j) = \sum_{i=1}^m a_{ij} w_i$ für $i = 1, \ldots, n$, d.h. M_C^B ist surjektiv und injektiv. □

Bemerkung 4.37. *Für $f \in Hom_K(V,W)$ und $v \in V$ mit $v = \sum_{j=1}^n \lambda_j v_j$ gilt damit*

$$f(v) = \sum_{j=1}^{n} \lambda_j f(v_j) = \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{m} a_{ij} w_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \lambda_j \right) w_i = \sum_{i=1}^{m} \left(A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \right)_i w_i.$$

Dies ist eine andere Schreibweise für Aussage (ii) in Satz 4.36 und verdeutlicht nochmal, dass mit $A = M_C^B(f)$ die Koordinaten bezüglich der Basen B und C transformiert werden.

Was besagt Satz 4.36 für $V = K^n$, $W = K^m$ mit den Einheitsbasen $B = (e_1, \ldots, e_n)$ von K^n und $C = (e_1, \ldots, e_m)$ von K^m? Zu gegebenem $f \in \mathrm{Hom}_K(K^n, K^m)$ ist A definiert durch

$$f(e_j) = \sum_{i=1}^{m} a_{ij} e_i = \underbrace{\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}}_{j\text{-te Spalte von } A} = Ae_j.$$

Das ist dieselbe Beziehung wie zwischen $\tilde{A}$ und A in Beispiel 4.2, d.h. $f = \tilde{A}$. Wir halten dies als Korollar fest.

Korollar 4.38. *Die Abbildung* $M_{(e_1,\ldots,e_m)}^{(e_1,\ldots,e_n)} : \mathrm{Hom}_K(K^n, K^m) \to M(m \times n, K)$ *ist ein Isomorphismus mit Umkehrabbildung*

$$\sim \; : M(m \times n, K) \to \mathrm{Hom}_K(K^n, K^m), \quad A \mapsto \tilde{A}.$$

Man kann dieses Korollar nun verwenden, um eine nicht-triviale Aussage für Matrizen herzuleiten.

Satz 4.39. *Sei $A \in M(n \times n, K)$. Dann sind äquivalent:*

(i) A besitzt eine Rechtsinverse, d.h. es existiert $B \in M(n \times n, K)$ mit $A \cdot B = E_n$.

(ii) $\tilde{A} : K^n \to K^n$ ist ein Isomorphismus.

(iii) A ist invertierbar, d.h. es existiert $B \in M(n \times n, K)$ mit $A \cdot B = B \cdot A = E_n$.

Die Aussage gilt auch, wenn man in (i) die Existenz der Linksinversen verlangt. Die Beweisidee besteht darin, die Frage nach der Inversen anstelle von $M(n \times n, K)$ in $\mathrm{Hom}_K(K^n, K^n)$ zu diskutieren, wo unmittelbar klar ist, dass aus der Existenz der Rechtsinversen auch die Existenz der Inversen folgt.

Beweis. (iii)⇒(i) ist trivial. $\tilde{A} \circ \tilde{B} = \widetilde{AB} = \tilde{E}_n = id_{K^n}$

(i)⇒(ii): Sei $B \in M(n \times n, K)$ mit $AB = E_n$.

$\Rightarrow \; K^n = (\tilde{A} \circ \tilde{B})(K^n) \subset \tilde{A}(K^n) \subset K^n$

$\Rightarrow \; \tilde{A}$ surjektiv

$\overset{\text{Korollar 4.16}}{\Rightarrow}$ $\tilde{A}$ ist bijektiv, d.h. $\tilde{A}$ ist Isomorphismus.

(ii)⇒(iii): Sei $\tilde{A}$ Isomorphismus, d.h. es existiert die inverse Abbildung g mit $\tilde{A} \circ g = g \circ \tilde{A} = id_{K^n}$. Nach Korollar 4.38 gibt es ein $B \in M(n \times n, K)$ mit $g = \tilde{B}$. ⇒ Behauptung. □

Satz 4.40 (Ringisomorphismus). *(i) Die Abbildung*

$$M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)} : \mathrm{End}_K(K^n) \to M(n \times n, K)$$

ist auch ein Ringisomorphismus (also bijektiv und die Ringstruktur erhaltend), d.h.

$$M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}(f+g) = M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}(f) + M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}(g) \quad \textit{und}$$
$$M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}(f \circ g) = M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}(f) \cdot M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}(g) \quad \textit{für alle } f,g \in \mathrm{End}_K(K^n).$$

(ii) Sei "$*$" eine beliebige Multiplikation auf $M(n \times n, K)$, so dass $M(n \times n, K)$ mit den Verknüpfungen $+$ und $*$ einen Ring bildet. Ist $M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}$ dann ein Ringisomorphismus, so folgt "$* = \cdot$", d.h. die Matrixmultiplikation ist gerade so definiert, dass das Matrixprodukt die Matrix der hintereinander ausgeführten Abbildungen ist.

Bemerkung. *Eine analoge Aussage gilt auch für das Matrixprodukt $A \cdot B$ mit Matrizen $A \in M(m \times n, K)$ und $B \in M(n \times r, K)$.*

Beweis. Wir setzen $M := M_{(e_1,\dots,e_n)}^{(e_1,\dots,e_n)}$.

(i) Wir verwenden wie oben die Bezeichnung $\tilde{A}$ für die durch $A \in M(n \times n, K)$ definierte lineare Abbildung von K^n nach K^n. Es gilt für alle $x \in K^n$

$$\widetilde{AB}(x) = (AB)x = A(Bx) = A \cdot \tilde{B}(x) = \tilde{A}\big(\tilde{B}(x)\big) = \big(\tilde{A} \circ \tilde{B}\big)(x),$$

womit $M\big(\tilde{A} \circ \tilde{B}\big) = M\big(\widetilde{AB}\big) = A \cdot B = M\big(\tilde{A}\big) \cdot M\big(\tilde{B}\big)$. Da M nach Satz 4.36 bereits VR-Isomorphismus ist, gilt auch $M\big(\tilde{A} + \tilde{B}\big) = M\big(\tilde{A}\big) + M\big(\tilde{B}\big)$ und M ist bijektiv, also ist M damit Ringisomorphismus.

(ii) Sei $\sim$ die Umkehrabbildung von M aus Korollar 4.38. Nach Satz 4.36 (i) besteht die j-te Spalte von $A = M\big(\tilde{A}\big)$ aus $\tilde{A}(e_j)$, d.h. für $C = A * B$ mit $A, B \in M(n \times n, K)$ gilt

$$c_{ij} = \Big(\widetilde{A * B}(e_j)\Big)_i = \Big(\underbrace{M\big(\tilde{A}\big) * M\big(\tilde{B}\big)}_{\substack{= M(\tilde{A} \circ \tilde{B}) \\ \text{da } M \text{ Ringiso-} \\ \text{morphismus}}}(e_j)\Big)_i = \Big(\big(\tilde{A} \circ \tilde{B}\big)(e_j)\Big)_i$$
$$= \big(\tilde{A}\big(\tilde{B}(e_j)\big)\big)_i = \left(\tilde{A}\left(\begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}\right)\right)_i$$
$$= \left(\sum_{k=1}^{n} b_{kj} \tilde{A}(e_k)\right)_i = \left(\sum_{k=1}^{n} b_{kj} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix}\right)_i = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

□

Im restlichen Teil dieses Abschnittes formulieren wir noch Konsequenzen aus Satz 4.36 und Korollar 4.38.

Satz 4.41. *Sei $U \subset K^n$. Dann sind äquivalent:*

(i) U ist UVR von K^n.

(ii) Es gibt ein $m \in \mathbb{N}$ und ein $A \in M(m \times n, K)$ mit $U = \text{Lös}(A, 0)$.

Beweis. (ii)⇒(i): Aussage (i) aus Satz 4.22.

(i)⇒(ii): Sei $(u_1, \ldots, u_r)$ eine Basis von U. Diese kann nach dem Basisergänzungssatz zu einer Basis $(u_1, \ldots, u_r, u_{r+1}, \ldots, u_n)$ von K^n ergänzt werden. Sei $m = n - r$. Definiere nun $f : K^n \to K^n$ durch

$$f(u_i) = \begin{cases} 0 & \text{falls } i \in \{1, \ldots, r\} \\ e_{i-r} & \text{falls } i \in \{r+1, \ldots, n\}. \end{cases}$$

Wegen $U \subset \text{Kern } f$ und $\dim\big(\text{Kern } f\big) = n - \dim\big(\text{Bild } f\big) = n - (n-r) = r$ (Dimensionsformel) folgt $\text{Kern } f = U$ nach Korollar 3.34 (iii). Nach Satz 4.36 und Korollar 4.38 gibt es genau ein $A \in M(m \times n, K)$ mit $f(x) = Ax$. Es gilt $\text{Lös}(A, 0) = \text{Kern } f = U$. □

Satz 4.42. *Sei $Z \subset K^n$. Dann sind äquivalent:*

(i) Z ist affiner Unterraum von K^n.

(ii) Es gibt $m \in \mathbb{N}$, $A \in M(m \times n, K)$ und $b \in K^m$ mit $Z = \text{Lös}(A, b)$.

Beweis. (ii)⇒(i): Aussage (ii) aus Satz 4.22.

(i)⇒(ii): Sei $Z = z + U$ mit $z \in Z$ und einem UVR $U \subset K^n$. Nach Satz 4.41 existieren $m \in \mathbb{N}$ und $A \in M(m \times n, K)$ mit $U = \text{Lös}(A, 0)$. Mit $b := Az$ ist $z \in \text{Lös}(A, b)$ und nach Satz 4.22 (iii) folgt $\text{Lös}(A, b) = z + U = Z$. □

Bemerkung. *Zusammenfassend ergeben Satz 4.41 und Satz 4.42:*

- *UVRs von K^n sind Lösungsräume homogener linearer Gleichungssysteme und*
- *affine Unterräume von K^n sind Lösungsräume inhomogener linearer Gleichungssysteme.*

Wir geben zu $f \in \text{Hom}_K(V, W)$ schließlich noch eine einfache Darstellungsmatrix an.

Lemma 4.43. *Sei $f : V \to W$ eine lineare Abbildung. Dann gibt es Basen B von V und C von W mit*

$$M_C^B(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}, \quad r = \text{Rang } f.$$

Beweis. Sei $\bar{C} = (w_1, \ldots, w_r)$ Basis von Bild f. Sind $u_1 \in f^{-1}(\{w_1\}), \ldots, u_r \in f^{-1}(\{w_r\})$ und ist $(v_1, \ldots, v_k)$ eine Basis von Kern f, dann ist nach Satz 4.13 (Dimensionsformel) $B = (u_1, \ldots, u_r, v_1, \ldots, v_k)$ eine Basis von V mit $r + k = \dim V$. Nach Konstruktion ist $f(u_1) = w_1, \ldots, f(u_r) = w_r,\ f(v_1) = 0, \ldots, f(v_k) = 0$, d.h. $M_C^B(f) = \left(\begin{smallmatrix} E_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$, wobei C eine von $\bar{C}$ ergänzte Basis von W ist. □

Bemerkung. *Sehr viel schwieriger ist es, für $f \in \mathrm{End}_K(V)$ eine Basis B zu finden, so dass $M_B(f)$ möglichst einfach ist ($\rightarrow$ Jordansche Normalform, Lineare Algebra II).*

4.5 Kommutative Diagramme und Basiswechsel

In diesem Abschnitt soll untersucht werden, wie sich die Darstellungsmatrix $M_C^B(f)$ einer linearen Abbildung f bei einem Basiswechsel verhält. Nach Satz 4.36 (ii) gilt

$$f = \Phi_C \circ \widetilde{M_C^B(f)} \circ \Phi_B^{-1}. \tag{4.2}$$

Bemerkung 4.44 (Kommutative Diagramme)**.** *In der linearen Algebra werden häufig sogenannte kommutative Diagramme verwendet, um die Gleichheit von Abbildungsverknüpfungen zu visualisieren. Schreibt man obige Identität in der Form*

$$f \circ \Phi_B = \Phi_C \circ \widetilde{M_C^B(f)},$$

so erhält man folgendes "kommutatives Diagramm":

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_B} & V \\ {\scriptstyle \widetilde{M_C^B(f)}}\downarrow & & \downarrow{\scriptstyle f} \\ K^m & \xrightarrow{\Phi_C} & W \end{array}$$

Man sagt "das Diagramm ist kommutativ", wenn man jeden Weg entlang der Pfeile "laufen" kann und immer dasselbe Ergebnis erhält ("laufen" bedeutet, die jeweilige Abbildung anzuwenden, d.h. die Abbildungen entlang des Pfads werden verknüpft). Die einzelnen Abbildungen müssen nicht bijektiv und im Allgemeinen kein Homomorphismus sein. Ist eine Abbildung bijektiv, kann man ihre Inverse hinschreiben und den Pfeil umdrehen:

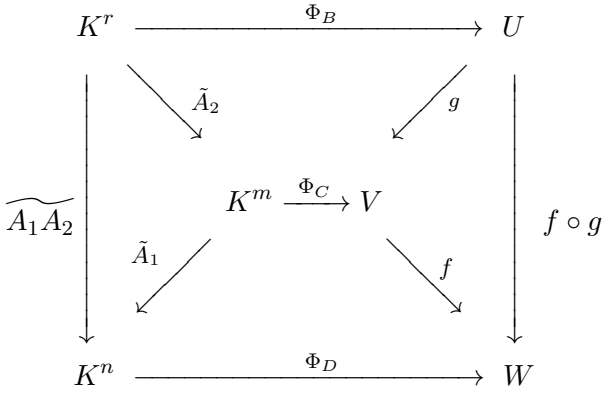
$$\begin{array}{ccc} K^n & \xleftarrow{\Phi_B^{-1}} & V \\ {\scriptstyle \widetilde{M_C^B(f)}}\downarrow & & \downarrow{\scriptstyle f} \\ K^m & \xrightarrow{\Phi_C} & W \end{array}$$

Die Kommutativität des Diagramms liefert dann direkt (4.2).

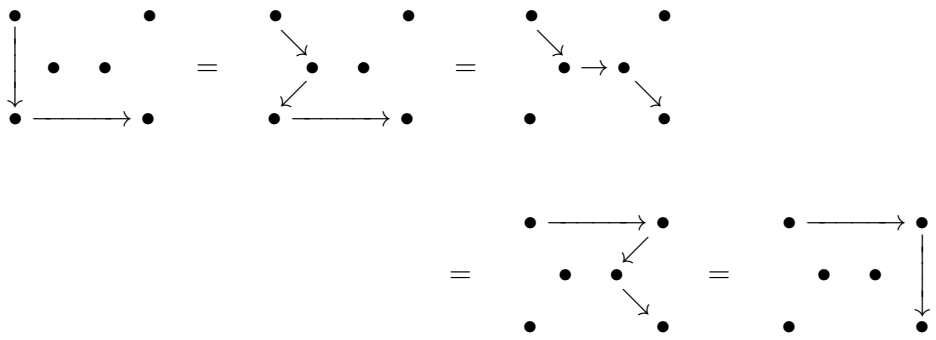
Satz 4.45. *Seien U, V, W endlichdimensionale K-VRs mit Basen B, C, D; $g : U \to V$, $f : V \to W$ lineare Abbildungen. Dann gilt*

$$M_D^B(f \circ g) = M_D^C(f) \cdot M_C^B(g).$$

Beweis. Wir setzen $r := \dim U$, $m := \dim V$, $n := \dim W$ sowie $A_1 := M_D^C(f)$, $A_2 := M_C^B(g)$. Damit ergibt sich folgendes Diagramm:



Die beiden Trapeze sind kommutativ wegen (4.2), das linke Dreieck wegen $\widetilde{A_1 A_2} = \tilde{A}_1 \circ \tilde{A}_2$, das rechte Dreieck nach Definition der Hintereinanderschaltung von Abbildungen auch. Es bleibt noch zu zeigen, dass auch das äußere Reckteck kommutiert (das gilt nicht automatisch!). Mithilfe der bereits als kommutativ verifizierten Teildiagramme folgt aber



Die Kommutativität des äußeren Rechtecks ergibt nun

$$\Phi_D \circ \widetilde{A_1 A_2} = (f \circ g) \circ \Phi_B.$$

$$\Rightarrow \widetilde{A_1 A_2} = \Phi_D^{-1} \circ (f \circ g) \circ \Phi_B = \widetilde{M_D^B(f \circ g)} \Rightarrow M_D^C(f) \cdot M_C^B(g) = A_1 A_2 = M_D^B(f \circ g). \quad \square$$

Bemerkung. *Der Beweis muss nicht mithilfe des kommutativen Diagramms geführt werden. Alternativ kann man in jedem Schritt die Gleichheit der Formeln einsetzen.*

Bemerkung 4.46. *Insbesondere ergibt Satz 4.45 für* $f, g \in \mathrm{End}_K(V)$

$$M_B(f \circ g) = M_B(f) M_B(g)$$

und damit analog zu Satz 4.40 (i), dass

$$M_B : \mathrm{End}_K(V) \to M(n \times n, K), \;\; f \mapsto M_B(f)$$

ein Ringisomorphismus ist.

Definition 4.47. *Seien* B *und* B' *Basen von* V *sowie* $n = \dim V$.

$$T^B_{B'} := M^B_{B'}(id_V) \in M(n \times n, K)$$

heißt Transformationsmatrix des Basiswechsels von B *nach* B'.

Lemma 4.48. *Seien* B, B', B'' *Basen von* V. *Dann gelten folgende Aussagen:*

(i) $T^B_{B'} \in GL(n, K)$

(ii) $\widetilde{T^B_{B'}} = \Phi^{-1}_{B'} \circ \Phi_B$

(iii) $T^B_{B'} = \left(T^{B'}_B\right)^{-1}$

(iv) $T^B_{B''} = T^{B'}_{B''} \cdot T^B_{B'}$

(v) Ist $(v_1, \ldots, v_n)$ *eine Basis von* $V = K^n$, *so folgt* $T^{(v_1,\ldots,v_n)}_{(e_1,\ldots,e_n)} = \underbrace{(v_1 \ldots v_n)}_{\substack{\text{Matrix mit} \\ \text{Spalten } v_j}}$.

Bemerkung. *(ii) zeigt, dass* $\widetilde{T^B_{B'}}$ *die Koordinaten eines beliebigen Vektors* $v \in V$ *bzgl.* B *in die Koordinaten von* v *bzgl.* B' *überführt.*

Beweis. (ii) Wir wenden (4.2) auf $f = id$ an und erhalten

$$id_V = \Phi_{B'} \circ \widetilde{T^B_{B'}} \circ \Phi^{-1}_B.$$

Da Φ_B und $\Phi_{B'}$ Isomorphismen sind, folgt

$$\Phi^{-1}_{B'} = \widetilde{T^B_{B'}} \circ \Phi^{-1}_B \text{ und daraus } \Phi^{-1}_{B'} \circ \Phi_B = \widetilde{T^B_{B'}}.$$

(i) Da Φ_B und $\Phi_{B'}$ Isomorphismen sind, ist auch $\widetilde{T^B_{B'}} = \Phi^{-1}_{B'} \circ \Phi_B$ ein Isomorphismus. Nach Satz 4.39 ist $T^B_{B'}$ damit invertierbar, d.h. $T^B_{B'} \in GL(n, K)$.

(iii) Die Aussage ist heuristisch klar. Formal: Es gilt $\widetilde{A_1 A_2} = \tilde{A}_1 \circ \tilde{A}_2$, d.h.

$$\widetilde{T^B_{B'} \cdot T^{B'}_B} = \widetilde{T^B_{B'}} \circ \widetilde{T^{B'}_B} = \Phi^{-1}_{B'} \circ \underbrace{\Phi_B \circ \Phi^{-1}_B}_{id_V} \circ \Phi_{B'} = id_{K^n} = \tilde{E}_n.$$

$\overset{\text{Satz } 4.40(i)}{\Rightarrow} T^B_{B'} \cdot T^{B'}_B = E_n \Rightarrow T^B_{B'} = \left(T^{B'}_B\right)^{-1}$.

(iv) folgt aus Satz 4.45 für $f = g = id_V$.

(v) folgt aus der Definition 4.36 der Darstellungsmatrix. □

Beispiele 4.49. *Zur Erinnerung: Sind $B = (v_1, \ldots, v_n)$ und $B' = (v'_1, \ldots, v'_n)$ Basen eines K-VRs V ($\dim_K V = n$) und $(t_{ij}) = T^B_{B'} = M^B_{B'}(id_V)$, so gilt gemäß Definition 4.36*

$$v_j = id_V(v_j) = \sum_{i=1}^{n} t_{ij} v'_i.$$

Wir betrachten nun die Beispiele 4.34 und 4.35.

Beispiel 4.34: (gedrehtes Koordinatensystem)

Mit

$$B = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right), \quad B' = \left(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\right)$$

gelten

$$\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}-1\\1\end{pmatrix} \quad \textit{und} \quad \begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}-1\\1\end{pmatrix}.$$

$\Rightarrow$

$$T^B_{B'} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}.$$

Beispiel 4.35: (Polynomraum)

Seien $B = (P_0, P_1, \ldots, P_n)$ und $B' = (P_0, P_0 + P_1, \ldots, \sum_{k=0}^{n} P_k)$ mit dem Nullpolynom P_0 und $P_j(t) = t^j$ für $j = 1, \ldots, n$. Dann ist $P_j = \sum_{k=0}^{j} P_k - \sum_{k=0}^{j-1} P_k$, womit

$$T^B_{B'} = \begin{pmatrix} 1 & -1 & & 0 \\ & & \ddots & \\ & & \ddots & -1 \\ 0 & & & 1 \end{pmatrix}.$$

Das Polynom $P_1 - P_0$ hat bezüglich B die Koordinaten $(-1, 1, 0, \ldots, 0)'$ und als Koordinaten in B'

$$T^B_{B'} \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \textit{(Es gilt: } P_1 - P_0 = -2P_0 + 1 \cdot (P_0 + P_1)\textit{.)}.$$

Satz 4.50 (Transformationsformel)**.** *Seien $f : V \to W$ eine lineare Abbildung, B, B' Basen von V sowie C, C' Basen von W. Dann gilt*

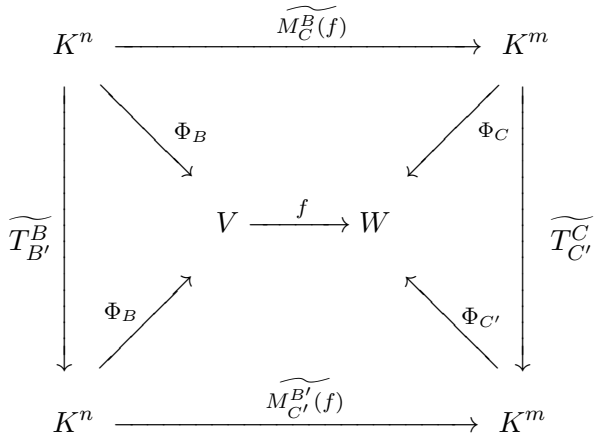
$$M_{C'}^{B'}(f) = T_{C'}^{C} M_C^B(f) T_B^{B'}.$$

Beweis. Nach (4.2) gilt

$$\widetilde{M_{C'}^{B'}(f)} = \Phi_{C'}^{-1} \circ f \circ \Phi_{B'} = \underbrace{\Phi_{C'}^{-1} \circ \Phi_C}_{=\widetilde{T_{C'}^{C}}} \circ \underbrace{\Phi_C^{-1} \circ f \circ \Phi_B}_{=\widetilde{M_C^B(f)}} \circ \underbrace{\Phi_B^{-1} \circ \Phi_{B'}}_{=\widetilde{T_B^{B'}}} = \widetilde{T_{C'}^{C} M_C^B(f) T_B^{B'}}.$$

$\Rightarrow$ Behauptung. □

Bemerkung. *Man kann den Beweis auch mit kommutativen Diagrammen zeigen (was letztlich dasselbe ist).*



Das Diagramm sieht ähnlich aus wie das im Beweis von Satz 4.45, allerdings zweimal mit anderen Richtungen im unteren Trapez ($\nearrow$ statt $\swarrow$ und $\nwarrow$ statt $\searrow$). Da $\Phi_{B'}$ und $\Phi_{C'}$ Isomorphismen sind, kann man diese Richtungen aber umdrehen und erhält wie im Beweis von Satz 4.45:

$$\underbrace{\widetilde{M_{C'}^{B'}(f)} \circ \widetilde{T_{B'}^{B}}}_{\| \; \widetilde{M_{C'}^{B'}(f) T_{B'}^{B}}} = \underbrace{\widetilde{T_{C'}^{C}} \circ \widetilde{M_C^B(f)}}_{\| \; \widetilde{T_{C'}^{C} M_C^B(f)}}$$

$\Rightarrow M_{C'}^{B'}(f) T_{B'}^{B} = T_{C'}^{C} M_C^B(f)$
$\Rightarrow M_{C'}^{B'}(f) = T_{C'}^{C} M_C^B(f) (T_{B'}^{B})^{-1} = T_{C'}^{C} M_C^B(f) T_B^{B'}$ *nach Lemma 4.48 (iii).*

Setzen wir $A_1 := M_C^B(f)$, $A_2 = M_{C'}^{B'}(f)$, $S := T_{C'}^{C}$ und $T := T_{B'}^{B}$, so gilt:

$$S \in GL(m, K), \; T \in GL(n, k) \text{ und } A_2 = S A_1 T^{-1}.$$

Im Spezialfall $W = V$ ist insbesondere der Fall $C = B$ und $C' = B'$ interessant. Hier folgt

$$M_{B'}^{B'}(f) = T_{B'}^{B} M_B^B(f) T_B^{B'},$$

d.h. mit $S := T_{B'}^B \in GL(n, K)$ ist

$$A_2 = S A_1 S^{-1}.$$

Definition 4.51.

(i) Zwei Matrizen $A_1, A_2 \in M(m \times n, K)$ heißen äquivalent ($A_1 \sim A_2$), falls es Matrizen $S \in GL(m, K)$ und $T \in GL(n, K)$ gibt mit $A_2 = S A_1 T^{-1}$.

(ii) Matrizen $A_1, A_2 \in M(n \times n, K)$ heißen ähnlich, falls es eine Matrix $S \in GL(n, K)$ gibt mit $A_2 = S A_1 S^{-1}$.

Die Relation $\sim$ ist reflexiv ($A \sim A$), symmetrisch ($A_1 \sim A_2 \Rightarrow A_2 \sim A_1$) und transitiv ($A_1 \sim A_2,\ A_2 \sim A_3 \Rightarrow A_1 \sim A_3$) und definiert somit eine Äquivalenzrelation auf $M(m \times n, K)$. Analog definiert Ähnlichkeit eine Äquivalenzrelation auf $M(n \times n, K)$.

Satz 4.52. *Seien $A_1, A_2 \in M(m \times n, K)$, B Basis von K^n, C Basis von K^m, desweiteren $f : K^n \to K^m$ eine lineare Abbildung und $M_C^B(f) = A_1$. Dann sind äquivalent:*

(i) $A_1 \sim A_2$.

(ii) Es gibt Basen B' von K^n und C' von K^m mit $M_{C'}^{B'}(f) = A_2$.

(iii) $\operatorname{Rang} A_1 = \operatorname{Rang} A_2$.

Insbesondere ist jede Matrix $A \in M(m \times n, K)$ vom Rang r äquivalent zu $\left(\begin{smallmatrix} E_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

Beweis. [Wir hatten oben (ii)⇒(i) gezeigt.]

(i)⇒(ii): Sei $A_1 \sim A_2$, d.h. es existieren $S \in GL(m, K)$, $T \in GL(n, K)$ mit $A_2 = S A_1 T^{-1}$. Sei $B = (v_1, \ldots, v_n)$, $T^{-1} = (t_{ij})$. Setze

$$v_j' := t_{1j} v_1 + \cdots + t_{nj} v_n \text{ für } j = 1, \ldots, n, \quad B' := (v_1', \ldots, v_n').$$

Es gilt: $T^{-1} \in GL(n, k) \overset{\text{Satz 4.39}}{\Rightarrow} \widetilde{T^{-1}}$ ist Isomorphismus $\overset{\text{Lemma 4.32}}{\Rightarrow} B'$ Basis. Nach Konstruktion ist

$$T^{-1} = M_B^{B'}(id_{K^n}) = T_B^{B'}.$$

Analog erhalten wir eine Basis C' mit $S^{-1} = T_C^{C'} \Leftrightarrow S = T_{C'}^C$. Es folgt

$$A_2 = S A_1 T^{-1} = T_{C'}^C M_C^B(f) T_B^{B'} = M_{C'}^{B'}(f).$$

(ii)⇒(iii): Es gilt:

$$\begin{aligned}\operatorname{Rang} A_1 = \operatorname{Rang} \tilde{A}_1 = \operatorname{Rang} \widetilde{M_C^B(f)} &= \operatorname{Rang}\left(\Phi_C^{-1} \circ f \circ \Phi_B\right) \\ &\overset{\Phi_B \text{ Isom.}}{=} \dim\left((\Phi_C^{-1} \circ f)(K^n)\right) \\ &\overset{\Phi_C^{-1} \text{ Isom.}}{=} \dim f(K^n) = \dim \operatorname{Bild}(f) = \operatorname{Rang} f.\end{aligned}$$

Analog zeigt man $\operatorname{Rang} A_2 = \operatorname{Rang} f \Rightarrow \operatorname{Rang} A_1 = \operatorname{Rang} A_2$.

(iii)⇒(i): Gelte $\operatorname{Rang} A_1 = \operatorname{Rang} A_2$ $(= r)$. Nach Lemma 4.43 gibt es Basen B von K^n und C von K^m mit

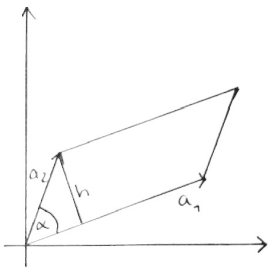
$$\begin{aligned}\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} &= M_C^B(\tilde{A}_1) \\ &= T_C^{(e_1,\ldots,e_n)} M_{(e_1,\ldots,e_n)}^{(e_1,\ldots,e_n)}(\tilde{A}_1) T_{(e_1,\ldots,e_n)}^B =: S A_1 T^{-1}.\end{aligned}$$

$\Rightarrow A_1 \sim \left(\begin{smallmatrix} E_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Analog zeigt man $\left(\begin{smallmatrix} E_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \sim A_2 \Rightarrow A_1 \sim A_2$. □

5 Determinanten

5.1 Axiomatische Definition nach Weierstraß und Leibniz-Formel

Motivation 5.1. *Sei $A = (a_{ij}) \in M(n \times n, \mathbb{R})$. Wir möchten das Volumen des von den Zeilenvektoren aufgespannten "Parallelotops" berechnen (im $\mathbb{R}^2$ des Parallelogramms). Bezeichnet a_i den Vektor der i-ten Zeile, so ergibt sich im $\mathbb{R}^2$:*



$$\begin{aligned}\textit{Fläche} &= |a_1| \cdot |a_2| \cdot \sin\alpha \\ &= \cdots = |a_{11}a_{22} - a_{12}a_{21}|.\end{aligned}$$

Man erhält dies auch unmittelbar aus dem Kreuzprodukt im $\mathbb{R}^3$, wenn man die z-Komponente 0 hinzufügt:

$$\begin{pmatrix} a_{11} \\ a_{12} \\ 0 \end{pmatrix} \times \begin{pmatrix} a_{21} \\ a_{22} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a_{11}a_{22} - a_{12}a_{21} \end{pmatrix}.$$

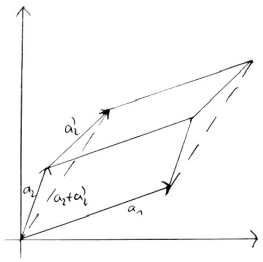
Man setzt nun

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

und kann zeigen:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{cases} > 0 & \textit{falls } \alpha \in (0,\pi) \\ < 0 & \textit{falls } \alpha \in (\pi, 2\pi) \\ = 0 & \textit{falls } \alpha = 0, \pi \textit{ (d.h. wenn } a_1 \textit{ und } a_2 \textit{ linear abhängig sind).} \end{cases}$$

Eine geschlossene Definition der Determinante im $\mathbb{R}^n$ ist schwierig und kompliziert. Deshalb wählen wir hier einen axiomatischen Zugang und halten dafür noch folgende Eigenschaft fest: Die Fläche des Parallelogramms von $(a_2 + a_2')$ und a_1 ist die Summe der Flächen der Parallelogramme von a_2 und a_1 plus a_2' und a_1.



Definition 5.2. *Eine Abbildung* $\det : M(n\times n, K) \to K$, $A \mapsto \det A$, *heißt* <u>Determinante</u>, *falls folgende Bedingungen erfüllt sind:*

(D1) $\det$ *ist linear in jeder Zeile, d.h. ist* $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ *mit Zeilen* $a_1, \ldots, a_n$, *dann gilt für alle* $i \in \{1, \ldots, n\}$

(D1a) $$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + \tilde{a}_i \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ \tilde{a}_i \\ \vdots \\ a_n \end{pmatrix}$$

(D1b) $$\det\begin{pmatrix} a_1 \\ \vdots \\ \lambda \cdot a_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \cdot \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} \textit{ für alle } \lambda \in K.$$

(D2) $\det$ *ist alternierend, d.h. hat A zwei gleiche Zeilen, so gilt* $\det A = 0$.

[Der Begriff "alternierend" beschreibt eigentlich (D6) unten. Falls (D1) gilt, folgt aber (D2)⇔(D6).]

(D3) $\det$ *ist normiert, d.h.* $\det E_n = 1$.

Bemerkung 5.3. *Es ist weder klar, ob* det *damit eindeutig festgelegt ist (Beweis folgt später), noch, ob überhaupt so eine Abbildung existiert.*

Zur Erinnerung:

- 3 Typen elementarer Zeilenumformungen
- jeweils realisierbar durch Multiplikation von A mit Matrizen Z ($ZM(i,\lambda)$ mit $\lambda \neq 0$, $ZA(i,j,\lambda)$, $ZV(i,j)$) von links
- Umformung von A damit auf ZDSF und SZSF möglich
- der Rang von A bleibt dabei erhalten.

Wir werde nachfolgend sehen, dass $\det A$ invariant gegenüber Umformungen von Typ II ist, bei Typ III kommt ein Faktor (-1) hinzu und bei Typ I der Faktor λ.

Satz 5.4. *Seien* det *eine Determinante und* $A \in M(n \times n, K)$ *mit Zeilenvektoren* $a_1, \ldots, a_n$. *Dann gilt:*

(D4) $\det(\lambda A) = \lambda^n \det(A)$.

(D5) Ist eine Zeile von A gleich 0, *so gilt* $\det(A) = 0$.

(D6) Zeilenumformung vom Typ III (Vertauschen von zwei Zeilen):

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = -\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$$

für $i \neq j$.

(D7) Zeilenumformung vom Typ II (Addieren des λ*-fachen der* j*-ten Zeile zur* i*-ten):*

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}$$

für $i \neq j$ *und* $\lambda \in K$.

(D8) Ist A eine obere Dreiecksmatrix, d.h. $A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, *so gilt* $\det A = \prod_{i=1}^{n} \lambda_i$.

(D9) Sind $n \geq 2$ und A von der Gestalt $A = \left(\begin{smallmatrix} A_1 & C \\ 0 & A_2 \end{smallmatrix}\right)$, *so gilt* $\det A = \det(A_1) \cdot \det(A_2)$. *Dasselbe gilt für* $A = \left(\begin{smallmatrix} A_1 & 0 \\ C & A_2 \end{smallmatrix}\right)$.

(D10) $\det A = 0 \Leftrightarrow \operatorname{Rang} A < n$ *(d.h.* $\det A \neq 0 \Leftrightarrow A \in GL(n, K)$*).*

(D11) $\det(A \cdot B) = \det(A) \cdot \det(B)$ *für* $A, B \in M(n \times n, K)$. *Insbesondere ist* $\det(A^{-1}) = (\det A)^{-1}$ *für* $A \in GL(n, K)$.

Beweis. (D4) Es gilt

$$\det(\lambda A) \overset{(D1b)}{=} \lambda \det \begin{pmatrix} a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix} = \cdots = \lambda^n \det(A).$$

$$\text{(D5) } \det \begin{pmatrix} a_1 \\ \vdots \\ 0 \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ 0_K \cdot 0 \\ \vdots \\ a_n \end{pmatrix} \overset{(D1b)}{=} 0_K \cdot \det \begin{pmatrix} a_1 \\ \vdots \\ 0 \\ \vdots \\ a_n \end{pmatrix} = 0.$$

(D6) Es gilt

$$0 \overset{(D2)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \\ a_n \end{pmatrix} = \underbrace{\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}}_{\overset{(D2)}{=} 0} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \underbrace{\det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}}_{\overset{(D2)}{=} 0}.$$

$\Rightarrow$ Behauptung.

(D7) Nach (D1) ist

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \underbrace{\lambda \det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}}_{\overset{(D2)}{=} 0} .$$

(a_j kommt $2x$ vor)

(D8) 1. Fall: $\lambda_i \neq 0 \ \forall\, i \in \{1, \ldots, n\}$.

Mit Zeilenumformungen vom Typ II kann man A auf die Form $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ bringen.

Es folgt $\det A \overset{(D7)}{=} \det\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \overset{(D1b)}{=} \lambda_1 \cdot \ldots \cdot \lambda_n \cdot \det(E_n) \overset{(D3)}{=} \prod_{i=1}^{n} \lambda_i$.

2. Fall: Es existiert $i \in \{1, \ldots, n\}$ mit $\lambda_i = 0$.

Sei $j := \max\{i \in \{1, \ldots, n\} | \lambda_i = 0\}$. Dann ist

$$A = \begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & * & & & \\ & & \lambda_{j-1} & & & & & \\ & & & 0 & \bullet & \ldots & \bullet \\ & & & & \lambda_{j+1} & * & & \\ & & & 0 & & \ddots & & \\ & & & & & & \lambda_n \end{pmatrix}$$

mit $\lambda_{j+1} \neq 0, \ldots, \lambda_n \neq 0$. Durch Zeilenumformungen vom Typ II kann man die Elemente $\bullet$ ... $\bullet$ zu 0 machen, also die j-te Zeile zu 0 transformieren, d.h. $\det A \overset{(D2)}{=} 0 = \prod_{i=1}^{n} \lambda_i$.

(D9) Sei $A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$. Man kann die Teilmatrix $(A_1 \ C)$ durch elementare Zeilenumformungen von Typ II und III auf Zeilenstufenform und damit auch A_1 auf obere Dreiecksgestalt B_1 bringen: $(A_1 \ C) \curvearrowright (B_1 \ C')$. Sei k die Anzahl der Zeilenvertauschungen (Typ III). Dann gilt $\det(B_1) = (-1)^k \det(A_1)$. Analog: $A_2 \curvearrowright B_2$ (obere Dreiecksmatrix) mit l Zeilenvertauschungen, d.h. $\det(B_2) = (-1)^l \det(A_2)$. Sei $B = \begin{pmatrix} B_1 & C' \\ 0 & B_2 \end{pmatrix}$. Dann folgt:

$$\det A \overset{(D6)+(D7)}{=} (-1)^{k+l} \det B \overset{(D8)}{=} (-1)^{k+l} \det(B_1) \cdot \det(B_2) = \det(A_1) \cdot \det(A_2).$$

Analog zeigt man die entsprechende Aussage für $A = \begin{pmatrix} A_1 & 0 \\ C & A_2 \end{pmatrix}$.

(D10) Wir bringen A mit elementaren Zeilenumformungen auf ZSF B, d.h.

$$\begin{aligned}\det A = \pm \det B \neq 0 \;&\Leftrightarrow\; \text{alle Diagonalelemente von } B \text{ sind ungleich } 0\\ &\Leftrightarrow \operatorname{Rang} B = n\\ &\Leftrightarrow \operatorname{Rang} A = n.\end{aligned}$$

(D11) 1. Fall: $\operatorname{Rang} A < n$.

$\Rightarrow \dim \operatorname{Bild} \tilde{A} < n$
$\Rightarrow \dim \operatorname{Bild}(\tilde{A} \circ \tilde{B}) < n$
$\Rightarrow \operatorname{Rang}(AB) < n$
$\Rightarrow \det(AB) \overset{(D10)}{=} 0 = 0 \cdot \det(B) \overset{(D10)}{=} \det(A) \cdot \det(B)$.

2. Fall: $\operatorname{Rang} A = n$.

A lässt sich mit elementaren Zeilenumformungen vom Typ II und III in ZSF bringen, die wegen $\operatorname{Rang} A = n$ von der Gestalt $\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ mit $\lambda_1 \neq 0, \ldots, \lambda_n \neq 0$ ist. Durch Zeilenumformungen vom Typ II kann man diese dann auf die Form

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} ZM(i, \lambda_i)$$

bringen (mit $ZM(i,\lambda)$ aus Lemma 3.54). Es gilt für alle $B \in M(n \times n, k)$

$$\det\big(ZM(i,\lambda)B\big) \overset{(D1b)}{=} \lambda \det B = \det\big(ZM(i,\lambda)\big) \cdot \det(B).$$

Analog gilt für die anderen elementaren Zeilenumformungen nach (D6) und (D7):

$$\det\big(ZA(i,j,\lambda) \cdot B\big) = \det B = \det\big(ZA(i,j,\lambda)\big) \cdot \det(B)$$

und

$$\det\big(ZA(i,j) \cdot B\big) = -\det B = \det\big(ZA(i,j)\big) \cdot \det(B).$$

Damit ist A ein Produkt $\prod_{i=1}^{m} Z_i$ mit Matrizen Z_i vom Typ $ZM(i,\lambda)$, $ZA(i,j,\lambda)$ oder $ZV(i,j)$, und es gilt

$$\det A = (-1)^k \prod_{i=1}^{n} \lambda_i,$$

wobei k die Anzahl der Zeilenvertauschungen = Anzahl der Matrizen $ZV(i,j)$ ist. Es

folgt

$$\det(AB) = \det\left(\Big(\prod_{i=1}^{m} Z_i\Big)B\right) = \det(Z_1)\det\left(\Big(\prod_{i=2}^{m} Z_i\Big)B\right)$$
$$= \cdots = \left(\prod_{i=1}^{m}\det(Z_i)\right)\det(B) = \det(A)\cdot\det(B).$$

Für $A \in GL(n,K)$ ist demnach $\det(A^{-1})\cdot\det(A) = \det\left(A^{-1}A\right) = \det E_n \overset{(D3)}{=} 1$, also $\det\left(A^{-1}\right) = \left(\det A\right)^{-1}$. □

Bemerkung 5.5. *Das nächste Ziel ist, die Determinante explizit auszurechnen. Sei*

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \text{ mit Zeilen } a_1,\ldots,a_n \in K^n.$$

Mit den Einheitsvektoren e_j als Zeilenvektoren und der Darstellung $a_i = \sum_{j=1}^{n} a_{ij}e_j$ folgt wegen Linearität der Determinante in jeder Zeile

$$\det A = \det\begin{pmatrix} \sum_{j=1}^{n} a_{1j}e_j \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \sum_{j_1=1}^{n} a_{1j_1}\det\begin{pmatrix} e_{j_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$
$$= \cdots = \sum_{j_1,\ldots,j_n=1}^{n} a_{1j_1}\cdot\ldots\cdot a_{nj_n}\cdot\det\begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix}.$$

Weiter gilt

$$\det\begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} = \begin{cases} 0 & \text{falls es } l \neq m \text{ gibt mit } j_l = j_m \\ \pm 1 & \text{falls alle } j_l \ (l = 1,\ldots,n) \text{ verschieden sind.} \end{cases}$$

Der letzte Fall bedeutet, dass die Indizes $j_1,\ldots,j_n$ eine Permutation der Zahlen $1,\ldots,n$ sind, d.h. $j_i = \sigma(i)$ für $\sigma \in S_n$. Der Faktor ± 1 hängt davon ab, ob man eine gerade oder ungerade Anzahl $A(\sigma)$ von Zeilenvertauschungen benötigt, um die $(e_{j_1},\ldots,r_{j_n}) = (e_{\sigma(1)},\ldots,e_{\sigma(n)})$ in die richtige Reihenfolge $(e_1,\ldots,e_n)$ zu bringen. Insgesamt erhält man die Leibniz-Formel

$$\det(A) = \sum_{\sigma\in S_n}(-1)^{A(\sigma)}a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)}$$

Das Problem ist, dass $A(\sigma)$ nicht eindeutig und somit nicht wohldefiniert ist. Man kann aber zeigen, dass für jedes $\sigma \in S_n$ die Anzahl $A(\sigma)$ entweder immer gerade oder immer

ungerade ist, d.h. das Vorzeichen $(-1)^{A(\sigma)}$ wäre eindeutig. Der Nachweis hiervon führt auf die Frage, wie man eine beliebige Permutation $\sigma \in S_n$ durch mehrfache Vertauschung von zwei Elementen (Transpositionen) in die Identität überführen kann.

Beispiele 5.6. *n=2:*

$$\begin{array}{rcc} \textit{Permutationen}: & \begin{pmatrix}1&2\\1&2\end{pmatrix} & \begin{pmatrix}1&2\\2&1\end{pmatrix} \\ \text{sign}: & +1 & -1 \end{array}$$

$\Rightarrow \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$. *Merkregel:* $\begin{matrix} a_{11} & & a_{12} \\ & \diagdown\!\diagup & \\ a_{21} & & a_{22} \end{matrix}$ *[gestrichelt mit Minus]*

n=3:

$$\begin{array}{rcccccc} \textit{Permutationen}: & \begin{pmatrix}1&2&3\\1&2&3\end{pmatrix} & \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix} & \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix} & \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix} & \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix} & \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} \\ \text{sign}: & +1 & -1 & -1 & -1 & +1 & +1 \end{array}$$

$$\Rightarrow \qquad \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31}.$$

Merkregel (Regel von Sarrus):

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

[durchgezogene Linie mit Faktor +1, gestrichelte Linie mit Faktor -1]

Definition 5.7. *(i) Ist $\sigma \in S_n$, so nennt man jedes Paar $(i,j) \in \{1,\dots,n\}^2$ mit $i<j$ und $\sigma(i) > \sigma(j)$ einen Fehlstand von σ [die Anzahl der Fehlstände bei festem σ ist eindeutig!].*

(ii) Wir definieren das Vorzeichen (Signum) von σ als

$$\text{sign}(\sigma) := \begin{cases} +1 & \textit{falls } \sigma \textit{ eine gerade Anzahl von Fehlständen hat} \\ -1 & \textit{falls } \sigma \textit{ eine ungerade Anzahl von Fehlständen hat.} \end{cases}$$

σ heißt gerade, wenn $\text{sign}(\sigma) = 1$, und ungerade, wenn $\text{sign}(\sigma) = -1$.

Lemma 5.8. *Für jedes $\sigma \in S_n$ gilt*

$$\operatorname{sign}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Beweis. Seien $\sigma \in S_n$ und m die Anzahl der Fehlstände von σ. Dann gilt

$$\begin{aligned}
\prod_{i<j} \big(\sigma(j) - \sigma(i)\big) &= \Bigg(\prod_{\substack{i<j: \\ \sigma(i)<\sigma(j)}} \big(\sigma(j) - \sigma(i)\big) \Bigg) (-1)^m \prod_{\substack{i<j: \\ \sigma(i)>\sigma(j)}} \big|\, \sigma(j) - \sigma(i) \,\big| \\
&= \underbrace{(-1)^m}_{=\operatorname{sign}(\sigma)} \prod_{i<j} \big|\, \sigma(j) - \sigma(i) \,\big| \\
&= \operatorname{sign}(\sigma) \prod_{i<j} (j - i),
\end{aligned}$$

wobei wir beim letzten Gleichheitszeichen ausgenutzt haben, dass $\sigma \in S_n$ bijektiv ist. □

Lemma 5.9. *Für alle $\sigma, \rho \in S_n$ gilt* $\operatorname{sign}(\sigma \circ \rho) = \operatorname{sign}(\sigma)\operatorname{sign}(\rho)$ *und insbesondere* $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$. *Damit ist*

$$\operatorname{sign} : (S_n, \circ) \to (\{-1, 1\}, \cdot)$$

ein Gruppenhomomorphismus.

Beweis. Wegen

$$\operatorname{sign}(\sigma \circ \rho) \overset{\text{Lemma 5.8}}{=} \prod_{i<j} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{j - i} = \Bigg(\prod_{i<j} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} \Bigg) \cdot \underbrace{\prod_{i<j} \frac{\rho(j) - \rho(i)}{j - i}}_{=\operatorname{sign}(\rho)}$$

bleibt zu zeigen, dass das erste Produkt im Ausdruck nach dem letzten Gleichheitszeichen mit $\operatorname{sign}(\sigma)$ übereinstimmt:

$$\begin{aligned}
\prod_{i<j} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} &= \Bigg(\prod_{\substack{i<j: \\ \rho(i)<\rho(j)}} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} \Bigg) \prod_{\substack{i<j: \\ \rho(i)>\rho(j)}} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} \\
&= \Bigg(\prod_{\substack{i<j: \\ \rho(i)<\rho(j)}} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} \Bigg) \prod_{\substack{i>j: \\ \rho(i)<\rho(j)}} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} \\
&= \prod_{\rho(i)<\rho(j)} \frac{\sigma\big(\rho(j)\big) - \sigma\big(\rho(i)\big)}{\rho(j) - \rho(i)} = \operatorname{sign}(\sigma)
\end{aligned}$$

nach Bijektivität von ρ und Lemma 5.8. □

Für den endgültigen Nachweis der Leibniz-Formel müssen noch die Transpositionen (Vertauschung von zwei Elementen) innerhalb der symmetrischen Gruppe untersucht werden.

Definition 5.10. *$\tau \in S_n$ heißt Transposition, wenn es $a, b \in \{1, \dots, n\}$ gibt mit $a \neq b$ und $\tau(a) = b$, $\tau(b) = a$ und $\tau(c) = c$ für alle $c \in \{1, \dots, n\} \setminus \{a, b\}$.*

Bemerkung. *Offenbar gilt $\tau^{-1} = \tau$.*

Lemma 5.11. *Ist $\tau \in S_n$ eine Transposition, so existiert ein $\sigma \in S_n$ mit $\tau = \sigma \circ \rho \circ \sigma^{-1}$, wobei $\rho = \binom{1\ 2\ 3\ \dots\ n}{2\ 1\ 3\ \dots\ n}$. Insbesondere gilt* $\operatorname{sign}(\tau) = \operatorname{sign}(\rho) = -1$.

Beweis. Sei $\tau = \binom{1\ \dots\ k\ \dots\ l\dots\ n}{1\ \dots\ l\ \dots\ k\dots\ n}$ die Transposition, welche die Elemente k und l vertauscht ($k \neq l$). Wir setzen $\sigma = \binom{1\ 2\ 3\ \dots\ n}{k\ l\ *\ \dots\ *} \in S_n$ ($* =$ beliebig). Damit folgt

$$
\begin{aligned}
&(\sigma \circ \rho \circ \sigma^{-1})(k) = l,\\
&(\sigma \circ \rho \circ \sigma^{-1})(l) = k\\
&\text{und für } i \notin \{k, l\}\ (\sigma \circ \rho \circ \sigma^{-1})(i) = (\sigma \circ \rho)\big(\underbrace{\sigma^{-1}(i)}_{\neq 1,2}\big) = \sigma\big(\sigma^{-1}(i)\big) = i,
\end{aligned}
$$

also $\sigma \circ \rho \circ \sigma^{-1} = \tau$. □

Wir können nun das folgende Resultat beweisen, welches dann zu einer sauberen Formulierung der Leibniz-Formel für die Determinante führt.

Proposition 5.12. *(i) Ist $\sigma \in S_n$, so existieren Transpositionen $\tau_1, \dots, \tau_k$ mit $\sigma = \tau_1 \circ \dots \circ \tau_k$ und* $\operatorname{sign}(\sigma) = (-1)^k$.

(ii) Für jede Permutation $\sigma \in S_n$ gilt $\det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = \operatorname{sign}(\sigma)$.

Beweis. (i) Ist $\sigma = \mathrm{Id}$, so gilt für eine beliebige Transposition τ

$$\sigma = \tau \circ \tau^{-1} = \tau \circ \tau.$$

Ist $\sigma \neq \mathrm{Id}$, so gibt es ein i_1 mit

$$\sigma(i) = i \text{ für } i = 1, \dots, i_1 - 1 \text{ und } \sigma(i_1) \neq i_1, \text{ d.h. } \sigma(i_1) > i_1.$$

Sei τ die Transposition, die i_1 mit $\sigma(i_1)$ vertauscht und $\sigma_1 := \tau_1 \circ \sigma$. Dann gilt $\sigma_1(i) = i$ für $i = 1, \dots, i_1$. Entweder ist nun $\sigma_1 = \mathrm{Id}$ oder es gibt ein $i_2 > i_1$ mit

$$\sigma_1(i) = i \text{ für } i = 1, \dots, i_2 - 1 \text{ und } \sigma_1(i_2) > i_2.$$

Sei analog nun τ_2 die Transposition, die i_2 und $\sigma_1(i_2)$ vertauscht, und $\sigma_2 := \tau_2 \circ \sigma_1 = \tau_2 \circ \tau_1 \circ \sigma$. Rekursiv erhält man so Transpositionen $\tau_1, \ldots, \tau_k$ mit

$$\tau_k \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{Id}\,. \quad \Longrightarrow \quad \sigma = \tau_1^{-1} \circ \cdots \circ \tau_k^{-1} = \tau_1 \circ \cdots \circ \tau_k.$$

Aus Lemma 5.9 und Lemma 5.11 folgt daraus

$$\mathrm{sign}(\sigma) = \prod_{j=1}^{k} \mathrm{sign}(\tau_j) = (-1)^k.$$

(ii) Ist $\sigma = \tau_1 \circ \cdots \circ \tau_k$, so kann man die Matrix $\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$ durch k Zeilenvertauschungen in die Einheitsmatrix E_n überführen, d.h.

$$\det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = (-1)^k \det(E_n) = (-1)^k = \mathrm{sign}(\sigma).$$

□

Bemerkung 5.13. *Die Zerlegung von σ in Transpositionen ist nicht eindeutig. Hat man eine andere Zerlegung $\sigma = \tau_1' \circ \cdots \circ \tau_l'$, so folgt $(-1)^l = \mathrm{sign}(\sigma) = (-1)^k$, d.h. k und l müssen entweder beide gerade oder beide ungerade sein. Damit ist die Eindeutigkeit einer Determinantenfunktion* $\det : M(n \times n, K)$ *mit (D1)–(D3) bewiesen, denn wenn es eine Funktion* det *mit (D1)–(D3) gibt, dann muss sie so aussehen wie in der Leibniz-Formel angegeben. Es fehlt noch die Existenz von* det*. Diese aber können wir beweisen, indem wir* det *über die Leibniz-Formel definieren und dann (D1)–(D3) nachrechnen.*

Satz 5.14. *Es gibt genau eine Determinante* $\det : M(n \times n, K) \to K$*. Diese ist für* $A = (a_{ij}) \in M(n \times n, K)$ *durch*

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(m)} \quad \textit{(Leibniz-Formel)} \tag{5.1}$$

gegeben.

Beweis. Eindeutigkeit: Diese folgt aus Bemerkung 5.5 und Proposition 5.12 (ii).

Existenz: Wir zeigen, dass det, definiert durch (5.1), die Eigenschaften (D1), (D3) und (D6) (Vertauschung von zwei Zeilen) erfüllt und dass daraus im Falle eines Körpers K mit $-1 \neq 1$ (D2) folgt. Der Beweis von (D2) für K mit $-1 = 1$ ist Aufgabe 1 auf Übungsblatt 2.

(D1a) Mit $a_i = a_i' + a_i''$ gilt

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \dots a_{n\sigma(m)} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix},$$

(D1b) verifiziert man analog.

(D3) Mit

$$\delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{andernfalls} \end{cases} \qquad (\text{“Kronecker-delta”})$$

ist $E_n = (\delta_{ij})$ und

$$\delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)} \begin{cases} 1 & \text{falls } \sigma = \mathrm{Id} \\ 0 & \text{andernfalls.} \end{cases}$$

$$\Longrightarrow \qquad \det(E_n) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)} = 1.$$

(D6) Sei τ die Transposition, welche i und j vertauscht. Dann gilt $\operatorname{sign}(\tau) = -1$ und

$$\begin{aligned} \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_{\tau(1)} \\ \vdots \\ a_{\tau(i)} \\ \vdots \\ a_{\tau(j)} \\ \vdots \\ a_{\tau(n)} \end{pmatrix} &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma\circ\tau(1)} \cdot \ldots \cdot a_{n\sigma\circ\tau(n)} \\ &= \sum_{\tilde\sigma \in S_n} \underbrace{\operatorname{sign}(\tilde\sigma \circ \tau^{-1})}_{=\operatorname{sign}(\tilde\sigma)\operatorname{sign}(\tau)} a_{1\tilde\sigma(1)} \cdot \ldots \cdot a_{n\tilde\sigma(n)} \\ &= (-1) \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} = -\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}. \end{aligned}$$

(D2) Hat A zwei gleiche Zeilen, so verändert sich A beim Vertauschen dieser Zeilen nicht und mit (D6) folgt $\det A = -\det A$. Für K mit $1 \neq -1$ impliziert dies $\det A = 0$. $\square$

Satz 5.15. *Sei $A \in M(n \times n, K)$. Dann gilt* $\det(A^t) = \det(A)$.

Beweis. Mit $A = (a_{ij})$ ist

$$\begin{aligned}\det(A^t) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)}.\end{aligned}$$

Nun ist aber die Abbildung $\Psi : S_n \to S_n$, $\sigma \mapsto \sigma^{-1}$, bijektiv, denn $\Psi \circ \Psi = \mathrm{Id}_{S_n}$.
$\Longrightarrow \det(A^t) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} = \det(A)$. □

Bemerkung 5.16 (Algorithmus zum Berechnen der Determinante)**.**

Eingabe: $A \in M(n \times n, K)$ *Ziel: Ausgabe von* $\det A$.

Durchführung:

(i) Bringe A durch elementare Zeilenumformungen und/oder elementare Spaltenumformungen (jeweils vom Typ II oder III) auf obere Dreiecksgestalt

$$B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

(ii) Ist k die Anzahl der Vertauschungen von Zeilen oder Spalten, so folgt

$$\det A = (-1)^k \det B = (-1)^k \prod_{i=1}^{n} \lambda_i.$$

Der Beweis für die Gültigkeit dieses Algorithmus folgt aus (D6), (D7) und (D8) sowie Satz 5.15 (bei Spaltenumformungen).

5.2 Laplace-Entwicklungssatz und Cramersche Regel

Definition 5.17. *Sei $A \in M(n \times n, K)$. Wir setzen*

$$A'_{ij} := \left(\begin{array}{ccc|c|ccc} a_{11} & \ldots & a_{1j-1} & a_{1j} & a_{1j+1} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \hline a_{i1} & \ldots & a_{ij-1} & a_{ij} & a_{ij+1} & \ldots & a_{in} \\ \hline \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & a_{nj-1} & a_{nj} & a_{nj+1} & \ldots & a_{nn} \end{array}\right) \in M\big((n-1) \times (n-1), K\big)$$

("Streichmatrix", die durch Streichen der i-ten Zeile und der j-ten Spalte entsteht)

und

$$A_{ij} := \begin{pmatrix} a_{11} & \dots & a_{1(j-1)} & 0 & a_{1(j+1)} & \dots & a_{1n} \\ \vdots & & & \vdots & & & \vdots \\ a_{(i-1)1} & & & 0 & & & a_{(i-1)n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{(i+1)1} & & & 0 & & & a_{(i+1)n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & 0 & a_{n(i+1)} & \dots & a_{nn} \end{pmatrix} \in M(n \times n, K)$$

sowie

$$\begin{aligned} a_{ij}^{\#} &:= \det(A_{ji}) \\ A^{\#} &:= (a_{ij}^{\#}) \in M(n \times n, K) = \big(\det(A_{ij})\big)^t. \end{aligned}$$

$A^{\#}$ heißt die zu A komplementäre Matrix.

Beispiel 5.18. *Für $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ ergeben sich*

$$A_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix},\ A_{12} = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix},\ A_{21} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \text{ und } A_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

womit $A^{\#} = \left(\begin{smallmatrix} 4 & -3 \\ -2 & 1 \end{smallmatrix}\right)^t = \left(\begin{smallmatrix} 4 & -2 \\ -3 & 1 \end{smallmatrix}\right)$.

Lemma 5.19. *Sei $A \in M(n \times n, K)$. Dann gilt*

$$\det(A_{ij}) = (-1)^{i+j} \det(A'_{ij}) \quad \textit{für alle } i, j \in \{1, \dots, n\}.$$

Beweis. Durch $i-1$ Vertauschungen benachbarter Zeilen und $j-1$ Vertauschungen benachbarter Spalten kann man A_{ij} auf die Form $\left(\begin{smallmatrix} 1 & 0 \\ 0 & A'_{ij} \end{smallmatrix}\right)$ bringen.

$$\overset{\text{(D6) und Satz 5.15}}{\Longrightarrow} (-1)^{i+j-2} \det(A'_{ij}) = (-1)^{i+j} \det(A'_{ij}).$$

□

Lemma 5.20. *Seien $A = (a^1, \dots, a^n) \in M(n \times n, K)$ mit Spaltenvektoren $a^1, \dots, a^n$ und e^i der i-te Spalteneinheitsvektor im K^n. Dann gilt*

$$\det(A_{ij}) = \det\big((a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n)\big).$$

Beweis. Durch Addition geeigneter Vielfache der j-ten Spalte führen wir die Matrix $(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n)$ in A_{ij} über. Mit (D7) folgt dann die Behauptung. □

Satz 5.21. *Sei $A \in M(n \times n, K)$. Dann gelten folgende Aussagen:*

(i) $A \cdot A^{\#} = A^{\#} \cdot A = \det(A) \cdot E_n$.

(ii) Im Falle $\det(A) \neq 0$ *existiert* A^{-1} *und besitzt die Darstellung*

$$A^{-1} = \frac{1}{\det(A)} A^{\#} = \frac{1}{\det(A)} \big((-1)^{i+j} \det(A'_{ij})\big)^t.$$

Beweis. (i) Seien $A = (a^1, \ldots, a^n) = (a_{ij})$ und $A^{\#} \cdot A = (b_{jk})$. Dann gilt

$$\begin{aligned}
b_{ik} &= \sum_{j=1}^{n} a_{ij}^{\#} a_{jk} \\
&= \sum_{j=1}^{n} a_{jk} \det(A_{ji} \\
&\overset{\text{Lemma 5.20}}{=} \sum_{j=1}^{n} a_{jk} \det\big((a^1, \ldots, a^{i-1}, e^j, a^{i+1}, \ldots, a^n)\big) \\
&\overset{\text{(D1)}}{=} \det\Big(\Big(a^1, \ldots, a^{i-1}, \underbrace{\sum_{j=1}^{n} a_{jk} e^j}_{=a^k}, a^{i+1}, \ldots, a^n\Big)\Big) \\
&\overset{\text{(D2)}}{=} \delta_{jk} \det(A).
\end{aligned}$$

Damit ist $A^{\#} \cdot A = \det(A) \cdot E_n$. Analog folgt $A \cdot A^{\#} = \det(A) \cdot E_n$.

(ii) ist eine unmittelbare Konsequenz aus (i) und Lemma 5.19. □

Beispiel 5.22. *Sei* $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, K)$. *Dann ist nach Satz 5.21 (ii)*

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

Satz 5.23 (Laplacescher Entwicklungssatz)**.** *Seien $n \geq 2$ und $A \in M(n \times n, k)$. Dann gilt: Für jedes $i \in \{1, \ldots, n\}$ ist*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \quad \textit{(Entwicklung nach der } i\textit{-ten Zeile)}$$

und für jedes $j \in \{1, \ldots, n\}$ ist

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \quad \textit{(Entwicklung nach der } j\textit{-ten Spalte).}$$

Beweis. Nach Satz 5.21 (i) gilt $A \cdot A^{\#} = \det(A) \cdot E_n$, d.h. für die Diagonalelemente $\sum_{j=1}^n a_{ij} a_{ji}^{\#}$, $i = 1, \ldots, n$, ist

$$\det(A) = \sum_{j=1}^{n} a_{ij} a_{ji}^{\#} = \sum_{j=1}^{n} a_{ij} \det(A_{ij}) \overset{\text{Lemma 5.19}}{=} \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}).$$

Analog zeigt man die Entwicklung nach der j-ten Spalte über $A^{\#} \cdot A = \det(A) \cdot E_n$. □

Beispiel 5.24. *Wir berechnen* $\det \begin{pmatrix} -2 & 2 & 3 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$ *durch Entwicklung nach der 2. Spalte:*

$$\begin{pmatrix} -2 & 2 & 3 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} = (-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} -1 & 1 \\ -1 & 1 \end{vmatrix} + (-1)^{2+2} \cdot 1 \cdot \begin{vmatrix} -2 & 3 \\ -1 & 1 \end{vmatrix} + (-1)^{3+2} \cdot 0 \cdot \begin{vmatrix} -2 & 3 \\ -1 & 1 \end{vmatrix}$$
$$= 0 + 1 + 0 = 1.$$

Bemerkung. *Beim Entwickeln der Determinante nach einer Zeile oder Spalte versucht man diese so auszuwählen, dass sie möglichst viele Nullen enthält.*

Satz 5.25 (Cramersche Regel)**.** *Seien* $A = (a^1, \ldots, a^n) \in GL(nK)$, $b \in K^n$ *und*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := A^{-1} b$$

die eindeutig bestimmte Lösung des linearen Gleichungssystems $Ax = b$. *Dann gilt*

$$x_i = \frac{\det\left((a^1, \ldots, a^{i-1}, b, a^{i+1}, \ldots, a^n)\right)}{\det A} \qquad \text{für alle } i = 1, \ldots, n.$$

Beweis. Nach Satz 5.21 (ii) gilt $A^{-1} = (d_{ij})$ mit $d_{ij} = \frac{1}{\det A} a_{ij}^{\#}$, d.h.

$$d_{ij} = \frac{1}{\det A} \det(A_{ji}) \overset{\text{Lemma 5.20}}{=} \frac{\det\left((a^1, \ldots, a^{i-1}, e^j, a^{i+1}, \ldots, a^n)\right)}{\det A}.$$

$\Longrightarrow$

$$x_i = (A^{-1}b)_i = \sum_{j=1}^{n} d_{ij} b_j \overset{\text{(D1)}}{=} \frac{\det\left((a^1, \ldots, a^{i-1}, \sum_{j=1}^n b_j e^j, a^{i+1}, \ldots, a^n)\right)}{\det A}$$
$$= \frac{\det\left((a^1, \ldots, a^{i-1}, b, a^{i+1}, \ldots, a^n)\right)}{\det A}.$$

□

Beispiel 5.26. *Wir betrachten das LGS*

$$\underbrace{\begin{pmatrix} -2 & 2 & 3 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}}_{=:A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}_{=:b}.$$

Nach Beispiel 5.24 ist $\det A = 1$. *Anwendung der Cramerschen Regel ergibt nun*

$$x_1 = \det \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \overset{\text{Entw. 3. Zeile}}{=} 1 \cdot \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3$$

$$x_2 = \det \begin{pmatrix} -2 & 1 & 3 \\ -1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \overset{\text{Entw. 3. Zeile}}{=} (-1) \cdot \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} -2 & 1 \\ -1 & 2 \end{vmatrix} = 5 + (-3) = 2$$

und

$$x_3 = \det \begin{pmatrix} -2 & 2 & 1 \\ -1 & 1 & 2 \\ -1 & 0 & 0 \end{pmatrix} \overset{\text{Entw. 3. Zeile}}{=} (-1) \cdot \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = -3.$$

Damit ist $\text{Lös}(A, b) = \left\{ \begin{pmatrix} -3 \\ 2 \\ -3 \end{pmatrix} \right\}$.

6 Eigenwerte von Endomorphismen

6.1 Eigenwerte, Eigenvektoren und Diagonalisierbarkeit

Wie bislang seien auch nachfolgend K ein Körper und V ein K-Vektorraum.

Definition 6.1. *Seien* $\lambda \in K$, $A \in M(n \times n, K)$ *und* $\varphi \in \text{End}_K(V)$.

(i) $\lambda \in K$ *heißt* Eigenwert *von* A, *falls es ein* $v \in K^n \setminus \{0\}$ *gibt mit* $Av = \lambda v$. v *heißt* Eigenvektor *von* A *zum Eigenwert* λ.

(ii) λ *heißt* Eigenwert *von* φ, *falls es ein* $v \in V \setminus \{0\}$ *gibt mit* $\varphi(v) = \lambda v$. v *heißt* Eigenvektor *von* φ *zum Eigenwert* λ.

Es sei darauf hingewiesen, dass $0 \in V$ als Eigenvektor ausgeschlossen ist, $0 \in K$ als Eigenwert jedoch nicht!

Definition 6.2 (Diagonalisierbarkeit). *(i) $A \in M(n \times n, K)$ heißt diagonalisierbar, falls es ein $S \in GL(n, K)$ und $\lambda_1, \ldots, \lambda_n \in K$ gibt mit*

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

(ii) Ist V endlichdimensional, so heißt $\varphi \in \mathrm{End}_K(V)$ diagonalisierbar, falls es eine Basis von V gibt bzgl. welcher die Darstellungsmatrix von φ eine Diagonalmatrix ist.

Lemma 6.3. *Für $A \in M(n \times n, K)$ sind äquivalent:*

(i) A ist diagonalisierbar.

(ii) K^n besitzt eine Basis $B = (v_1, \ldots, v_n)$ aus Eigenvektoren von A.

In diesem Fall gilt

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei die λ_i die Eigenwerte von A sind und $S = T_B^{B'}$, $S^{-1} = T_{B'}^{B}$ mit $B' = (e_1, \ldots, e_n)$.

Beweis. (ii)⇒(i): Wir setzen $S^{-1} = (v_1, \ldots, v_n) \in GL(n, K)$. Dann gelten $S^{-1} = T_{B'}^{B}$ sowie $S = T_B^{B'}$ nach Lemma 4.48, insbesondere ist $v_j = S^{-1}e_j$.
⇒ $(AS^{-1})e_j = A(S^{-1}e_j) = Av_j = \lambda_j v_j = \lambda_j S^{-1}e_j = S^{-1}\lambda_j e_j$ für alle $j \in \{1, \ldots, n\}$
⇒ $(SAS^{-1})e_j = \lambda_j e_j$ für alle $j \in \{1, \ldots, n\}$
Da $(e_1, \ldots, e_n)$ Basis von K^n ist, ist die lineare Abbildung $\widetilde{SAS^{-1}}$ damit eindeutig festgelegt, d.h.

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

und (i) ist gezeigt.

(i)⇒(ii): Aus $SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ folgt durch Multiplikation mit S^{-1} von links

$$A(S^{-1}e_j) = S^{-1}\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} e_j = \lambda_j S^{-1}e_j.$$

Wegen $S^{-1} \in GL(n, K)$ ist $S^{-1}e_j \neq 0$, d.h. λ_j ist Eigenwert mit Eigenvektor $v_j = S^{-1}e_j$.

Wegen $S^{-1} \in GL(n, K)$ hat die Matrix

$$(v_1, \ldots, v_n) = (S^{-1}e_1, \ldots, S^{-1}e_n) = S^{-1}(e_1, \ldots, e_n) = S^{-1}E_n = S^{-1}$$

vollen Rang, also ist $B = (v_1, \ldots, v_n)$ ist Basis von K^n. Nach Lemma 4.48 (v) ist zudem $S^{-1} = T_{B'}^{B}$. □

Beispiel 6.4. *Seien $K = \mathbb{R}$ und $V = \mathbb{R}^2$.*

(i) Sei

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2, \ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{=:A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Diese Abbildung beschreibt die Spiegelung an der Geraden $\{(x_1, x_2) \in \mathbb{R}^2 | x_1 = x_2\}$.

Heuristisch klar: Die beiden Vektoren $\binom{1}{1}$ und $\binom{1}{-1}$ auf und senkrecht der Geraden sollten Eigenvektoren sein. Die zugehörigen Eigenwerte wären dann 1 und -1.

Formal: $A\binom{1}{1} = \binom{1}{1}$ und $A\binom{1}{-1} = \binom{-1}{1} = (-1)\binom{1}{-1}$. $\left(\binom{1}{1}, \binom{1}{-1}\right)$ ist Basis von $\mathbb{R}^2$. Mit $S^{-1} := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ ist dann nach Lemma 6.3 $SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(ii) Nicht jede lineare Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ besitzt einen Eigenwert. Ein Beispiel ist die Drehung um $\pi/2$:

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2, \ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Anschaulich klar – Beweis später!

Lemma 6.5. *Seien V endlichdimensional und $\varphi \in \mathrm{End}_K(V)$. Dann sind äquivalent:*

(i) φ ist diagonalisierbar.

(ii) V besitzt eine Basis, die nur aus Eigenvektoren von φ besteht.

Beweis. Per definitionem ist φ diagonalisierbar genau dann, wenn eine Basis $B = (v_1, \ldots, v_n)$ existiert, bezüglich welcher die darstellende Matrix $M_B(\varphi)$ von φ die Gestalt

$$M_B(\varphi) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

besitzt. Aber nach Definition der darstellenden Matrix bedeutet das gerade $\varphi(v_i) = \lambda_i v_i$ für alle $i \in \{1, \ldots, n\}$. □

Die Charakterisierungen der Diagonalisierbarkeit aus Lemma 6.3 (ii) und Lemma 6.5 (ii) werden nachfolgend häufig verwendet.

Lemma 6.6. *Sei $\varphi \in \mathrm{End}_K(V)$. Seien $v_1, \ldots, v_m$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_m \in K$ von φ. Dann ist $(v_1, \ldots, v_m)$ linear unabhängig und damit gilt $m \leq \dim(V)$. Insbesondere gilt: Ist V endlichdimensional, so besitzt φ höchstens $\dim(V)$ Eigenwerte.*

Beweis. Wir beweisen die Aussage durch vollständige Induktion nach m.

Induktionsanfang: $m = 1$: Da v_1 Eigenvektor ist, gilt $v_1 \neq 0 \Rightarrow (v_1)$ linear unabhängig.

Induktionsschritt: Sei $m \geq 2$ und die Aussage für $m - 1$ bereits bewiesen.

Seien $\alpha_1, \ldots, \alpha_m \in K$ mit $\sum_{i=1}^m \alpha_i v_i = 0$. (Zu zeigen: $\alpha_1 = \cdots = \alpha_m = 0$.) Dann folgen

$$0 = \varphi(0) = \varphi\left(\sum_{i=1}^m \alpha_i v_i\right) = \sum_{i=1}^m \alpha_i \varphi(v_i) = \sum_{i=1}^m \alpha_i \lambda_i v_i.$$

sowie $\lambda_1 \sum_{i=1}^m \alpha_i v_i = 0$ und damit

$$\sum_{i=2}^m \alpha_i (\lambda_i - \lambda_1) v_i = 0.$$

$\Rightarrow \alpha_i(\lambda_i - \lambda_1) = 0 \; \forall\, i \in \{2, \ldots, m\}$ (Induktionsvoraussetzung).
$\Rightarrow \alpha_i = 0 \; \forall\, i \in \{2, \ldots, m\}$, da λ_i paarweise verschieden nach Voraussetzung.
$\overset{\sum_{i=1}^m \alpha_i v_i = 0}{\Longrightarrow} \alpha_1 v_1 = 0 \Rightarrow \alpha_1 = 0$, da v_1 Eigenvektor und somit $v_1 \neq 0$.
$\Rightarrow (v_1, \ldots, v_m)$ linear unabhängig. □

Korollar 6.7. *Seien $\dim(V) = n$ und $\varphi \in \mathrm{End}_K(V)$ habe n verschiedene Eigenwerte. Dann ist φ diagonalisierbar.*

Beweis. Seien $v_1, \ldots, v_n$ Eigenvektoren zu $\lambda_1, \ldots, \lambda_n$. Nach Lemma 6.6 sind sie linear unabhängig und wegen $\dim(V) = n$ damit bereits eine Basis von V. Nach Lemma 6.5 (ii) ist φ dann diagonalisierbar. □

Definition 6.8. *Seien $\varphi \in \mathrm{End}_K(V)$ und $\lambda \in K$. Dann heißt*

$$\mathrm{Eig}(\varphi, \lambda) := \left\{ v \in V \;\middle|\; \varphi(v) = \lambda v \right\}$$

Eigenraum von φ bezüglich λ. Für $A \in M(n \times n, K)$ setzen wir $\mathrm{Eig}(A, \lambda) := \mathrm{Eig}(\tilde{A}, \lambda)$.

Lemma 6.9. *Seien $\varphi \in \mathrm{End}_K(V)$ und $\lambda \in K$. Dann gelten folgende Aussagen:*

(i) $\mathrm{Eig}(\varphi, \lambda)$ *ist ein UVR von V.*

(ii) *λ ist Eigenwert von $\varphi \Leftrightarrow \mathrm{Eig}(\varphi, \lambda) \neq \{0\}$.*
In diesem Falle ist $\mathrm{Eig}(\varphi, \lambda) \setminus \{0\}$ die Menge der Eigenvektoren von φ zu λ.

(iii) $\mathrm{Eig}(\varphi, \lambda) = \mathrm{Kern}(\lambda \,\mathrm{Id}_V - \varphi)$, *insbesondere ist*

$$\mathrm{Eig}(A, \lambda) = \mathrm{Kern}\left(\lambda \,\mathrm{Id}_{K^n} - \tilde{A}\right) = \mathrm{L\ddot{o}s}(\lambda E_n - A, 0).$$

(iv) $\lambda_1, \lambda_2 \in K$ *mit* $\lambda_1 \neq \lambda_2 \Rightarrow \mathrm{Eig}(\varphi, \lambda_1) \cap \mathrm{Eig}(\varphi, \lambda_2) = \{0\}$.

Beweis. (iii) Es gelten die Äquivalenzen
$v \in \mathrm{Eig}(\varphi, \lambda) \Leftrightarrow \varphi(v) = \lambda v \Leftrightarrow \lambda v - \varphi(v) = 0 \Leftrightarrow \left(\lambda \,\mathrm{Id}_V - \varphi\right)(v) = 0 \Leftrightarrow v \in \mathrm{Kern}\left(\lambda \,\mathrm{Id}_V - \varphi\right)$.

(ii) gilt per definitionem von Eigenwert und Eigenraum.

(i) folgt aus (iii) mit Satz 4.10 (i).

(iv) $v \in \mathrm{Eig}(\varphi, \lambda_1) \cap \mathrm{Eig}(\varphi, \lambda_2) \Leftrightarrow \lambda_1 v = \varphi(v) = \lambda_2 v \overset{\lambda_1 \neq \lambda_2}{\Rightarrow} v = 0$. □

Definition 6.10. *Seien* $\varphi \in \mathrm{End}_K(V)$ *und* $\lambda \in K$. *Wir nennen* $\mu_{geo} := \dim \mathrm{Eig}(\varphi, \lambda)$ *geometrische Vielfachheit von* λ. *Für* $A \in M(n \times n, K)$ *setzen wir* $\mu_{geo}(A, \lambda) := \mu_{geo}(\tilde{A}, \lambda)$.

Lemma und Definition 6.11. *Seien* V *endlichdimensional,* $\varphi \in \mathrm{End}_K(V)$ *und* B *eine Basis von* V. *Setze*

$$\det(\varphi) := \det\left(M_B(\varphi)\right).$$

Dann gilt:

(i) $\det(\varphi)$ *ist wohldefiniert, d.h.* $\det\left(M_B(\varphi)\right)$ *ist unabhängig von* B.

(ii) φ *ist ein Isomorphismus* $\Leftrightarrow \det(\varphi) \neq 0$.

Beweis. Übungsblatt 3, Aufgabe 1 (i). □

Lemma 6.12. *Seien* V *endlichdimensional und* $\lambda \in K$. *Dann sind äquivalent:*

(i) λ *ist Eigenwert von* φ.

(ii) $\det\left(\lambda \,\mathrm{Id}_V - \varphi\right) = 0$.

Beweis. Es gilt (i) $\Leftrightarrow \mathrm{Eig}(\varphi, \lambda) \neq \{0\} \Leftrightarrow \mathrm{Kern}\left(\lambda \,\mathrm{Id}_V - \varphi\right) \neq \{0\} \Leftrightarrow \lambda \,\mathrm{Id}_V - \varphi$ ist kein Isomorphismus $\overset{\text{Lemma 6.11}}{\Leftrightarrow} \det\left(\lambda \,\mathrm{Id}_V - \varphi\right) = 0$, also Aussage (ii). □

6.2 Charakteristisches Polynom

Definition 6.13. *Sei* $A = (a_{ij}) \in M(n \times n, K)$.

$$\chi_A := \det(tE_n - A) = \det \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & \dots & & t - a_{nn} \end{pmatrix} \in K[t]$$

heißt charakteristisches Polynom von A.

Beispiel 6.14. $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \Rightarrow \chi_A(t) = \det \left(\begin{smallmatrix} t-1 & -2 \\ -3 & t-4 \end{smallmatrix}\right) = t^2 - 5t - 2$. *Die Eigenwerte von A sind nach Lemma 6.12 die Nullstellen von χ_A, d.h.* $\lambda_{1/2} = \frac{5}{2} \pm \sqrt{\frac{25}{4} + 2}$.

Lemma 6.15. *Seien $A, B \in M(n \times n, K)$. Sind A und B ähnlich im Sinne von Definition 4.51 (ii) (d.h. $\exists\, S \in GL(n, K) :\ B = SAS^{-1}$), so gilt $\chi_A = \chi_B$.*

Beweis. Nach Voraussetzung ist

$$\begin{aligned} tE_n - B &= tE_n - SAS^{-1} \\ &= StE_nS^{-1} - SAS^{-1} \\ &= S\big(tE_n - A\big)S^{-1}. \end{aligned}$$

Somit gilt

$$\begin{aligned} \chi_B = \det(tE_n - B) &= \det\left(S(tE_n - A)S^{-1}\right) \\ &\overset{\text{Satz 5.4 (D11)}}{=} \det(S)\det(tE_n - A)\underbrace{\det(S^{-1})}_{=\det(S)^{-1}} = \chi_A. \end{aligned}$$

□

Lemma und Definition 6.16. *Seien $\dim(V) = n \in \mathbb{N}$ und $\varphi \in \mathrm{End}_K(V)$. Dann heißt*

$$\chi_\varphi := \det(t\,\mathrm{Id}_V - \varphi)$$

charakteristisches Polynom von φ. Ist B eine Basis von V und $A = M_B(\varphi)$, so gilt

$$\chi_\varphi = \chi_A.$$

Beweis. Wegen $M_B(t\,\mathrm{Id}_V - \varphi) = tM_B(\mathrm{Id}_V) - M_B(\varphi) = tE_n - A$ folgt die Aussage $\chi_\varphi = \chi_A$ unmittelbar aus Lemma 6.11 (i). □

Lemma 6.17 (Kästchenformel)**.** *Sei $A \in M(n \times n, K)$ von der Gestalt $A = \left(\begin{smallmatrix} A_1 & C \\ 0 & A_2 \end{smallmatrix}\right)$ mit quadratischen Matrizen $A_1 \in M(r \times r, K)$, $A_2 \in M(s \times s, K)$ mit $r + s = n$. Dann gilt*

$$\chi_A = \chi_{A_1} \cdot \chi_{A_2}.$$

Beweis.

$$\begin{aligned} \chi_A = \det(tE_n - A) &= \det\begin{pmatrix} tE_r - A_1 & C \\ 0 & tE_s - A_2 \end{pmatrix} \\ &\overset{\text{Satz 5.4 (D9)}}{=} \det(tE_r - A_1) \cdot \det(tE_s - A_2) = \chi_{A_1} \cdot \chi_{A_2}. \end{aligned}$$

□

Lemma und Definition 6.18. *Sei $A = (a_{ij}) \in M(n \times n, K)$. Dann wird die Spur von A definiert durch*

$$\mathrm{Spur}(A) := \sum_{i=1}^{n} a_{ii}.$$

Sind $\varphi \in \mathrm{End}_K(V)$ mit $n = \dim(V) < \infty$ und B eine Basis von V, so setzt man

$$\mathrm{Spur}(\varphi) := \mathrm{Spur}\left(M_B(\varphi)\right).$$

Es gelten folgende Aussagen:

(i) $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$, *insbesondere ist* $\mathrm{Spur}(S^{-1}AS) = \mathrm{Spur}(A)$.

(ii) $\mathrm{Spur}(\varphi)$ *ist wohldefiniert.*

(iii) $\mathrm{Spur}(\varphi \circ \psi) = \mathrm{Spur}(\psi \circ \varphi)$ *für* $\varphi, \psi \in \mathrm{End}_K(V)$.

Beweis. Übungsblatt 3, Aufgabe 2. □

Satz 6.19. *Seien $n = \dim(V) < \infty$ und $\varphi \in \mathrm{End}_K(V)$. Dann gilt:*

(i) χ_φ ist ein normiertes Polynom vom Grad n (d.h. der $\deg(\chi_\varphi)$-te Koeffizient ist 1)

$$\chi_\varphi = t^n + c_{n-1}t^{n-1} + \cdots + c_0$$

mit $c_0 = \pm\det(\varphi)$ und $c_{n-1} = -\mathrm{Spur}(\varphi)$.

(ii) Die Nullstellen von χ_φ sind genau die Eigenwerte von φ:

$$\lambda \in K \text{ ist Eigenwert von } \varphi \Leftrightarrow \chi_\varphi(\lambda) = 0.$$

Beweis. (i) Seien B eine Basis von V sowie $A := M_B(\varphi) \in M(n \times n, K)$. Dann gilt

$$\begin{aligned}\chi_\varphi = \chi_A &= \det(\underbrace{tE_n - A}_{=:B=(b_{ij})}) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) b_{1\sigma(1)} \dots b_{n\sigma(n)} \\ &= (t - a_{11}) \dots (t - a_{nn}) + \underbrace{\sum_{\sigma \in S_n \setminus \{\mathrm{Id}_{S_n}\}} \mathrm{sign}(\sigma) b_{1\sigma(1)} \dots b_{n\sigma(n)}}_{=:Q}\end{aligned}$$

Für $\sigma \in S_n \setminus \{\mathrm{Id}_{S_n}\}$ treten im Produkt $b_{1\sigma(1)} \dots b_{n\sigma(n)}$ höchstens $n-2$ Diagonalelemente auf, womit $\deg(Q) \leq n-2$ falls $Q \neq 0$.

$\Rightarrow \chi_\varphi = t^n - (a_{11} + \cdots + a_{nn})t^{n-1} +$ Terme kleinerer Ordnung

$\Rightarrow c_{n-1} = -\mathrm{Spur}(A) = -\mathrm{Spur}(\varphi)$.

Ferner gilt $c_0 = \chi_\varphi(0) = \det(0 \cdot E_n - A) = \det(-A) = (-1)^n \det(A)$.

(ii) ist mit der Definition des charakteristischen Polynoms Aussage von Lemma 6.12. □

Definition 6.20. *(i) Seien $P \in K[t]$, $P \neq 0$, $\lambda \in K$.*

$$\mu(P,\lambda) := \max\left\{ l \in \mathbb{N}_0 \,\middle|\, \exists\, Q \in K[t] \text{ mit } P = (t-\lambda)^l Q \right\}$$

heißt Vielfachheit der Nullstelle λ von P.

(ii) Seien V endlichdimensional und $\varphi \in \mathrm{End}_K(V)$. Dann nennt man $\mu_{alg}(\varphi,\lambda) := \mu(\chi_\varphi,\lambda)$ algebraische Vielfachheit von λ beim Endomorphismus φ.

Lemma 6.21. *Seien $n = \dim(V) < \infty$ und $\varphi \in \mathrm{End}_K(V)$. Dann gilt*

$$\mu_{geo}(\varphi,\lambda) \leq \mu_{alg}(\varphi,\lambda),$$

d.h. die Dimension des zu λ gehörenden Eigenraums ist höchstens gleich der Vielfachheit der Nullstelle λ von χ_φ.

Beweis. Im Falle $\mu_{geo}(\varphi,\lambda) = \dim\big(\mathrm{Eig}(\varphi,\lambda)\big) = 0$ ist nichts zu zeigen. Seien also $\dim\big(\mathrm{Eig}(\varphi,\lambda)\big) > 0$, ferner $b_1,\ldots,b_r$ eine Basis von $\mathrm{Eig}(\varphi,\lambda)$. Wir ergänzen diese zu einer Basis $B = (b_1,\ldots,b_n)$ von V. Bezüglich B ist die Darstellungsmatrix $M_B(\varphi)$ von der Gestalt

$$M_B(\varphi) = \left(\begin{array}{c|c} \lambda E_r & C \\ \hline 0 & D \end{array}\right).$$

Nach der Kästchenformel folgt $\chi_\varphi = \det\big((t-\lambda)E_r\big)\det(TE_n - D) = (t-\lambda)^r \det(tE_n - D)$, also die Behauptung. □

Beispiele 6.22. Frage: *Wann ist φ diagonalisierbar? Kann man das an χ_φ erkennen?*

(i) Wir betrachten noch einmal φ aus Beispiel 6.4 (i):

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{=:A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}.$$

Es gilt $\chi_\varphi = \chi_A = \det(tE_2 - A) = \det\left(\begin{smallmatrix} t & -1 \\ -1 & t \end{smallmatrix}\right) = t^2 - 1 = (t-1)(t+1)$. Nach Lemma 6.12 sind -1 und 1 die Eigenwerte von φ und damit ist φ nach Korollar 6.7 diagonalisierbar. Es sind

$$\mathrm{Eig}(\varphi,1) = \mathrm{Eig}(A,1) = \mathrm{L\ddot{o}s}(E_2 - A, 0) = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \text{ und}$$

$$\mathrm{Eig}(\varphi,-1) = \cdots = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right).$$

(ii) Wir betrachten nun φ aus Beispiel 6.4 (ii) (Drehung um $\pi/2$):

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{=:A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

Hier gilt $\chi_\varphi = \chi_A = \det(tE_2 - A) = \det \begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1$. Dieses hat in $\mathbb{R}$ keine Nullstelle! Nach Lemma 6.12 hat φ somit keinen Eigenwert und ist damit nach Lemma 6.5 nicht diagonalisierbar.

(iii) Sei schließlich

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Es gilt $\chi_\varphi = \chi_A = \det \begin{pmatrix} t-1 & -1 \\ 0 & t-1 \end{pmatrix} = (t-1)^2$, d.h. $\mu_{alg}(\varphi, 1) = 2$. Weiter ist

$$\mathrm{Eig}(\varphi, 1) = \mathrm{Eig}(A, 1) = \mathrm{L\ddot{o}s}(E_2 - A, 0) = \mathrm{L\ddot{o}s}\left(\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, 0 \right) = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right),$$

d.h. $\mu_{geo}(\varphi, 1) = 1$. Damit besitzt $\mathbb{R}^2$ keine Basis aus Eigenvektoren von φ und φ ist gemäß Lemma 6.5 nicht diagonalisierbar.

Lemma 6.23. *Seien $n = \dim V < \infty$ und $\varphi \in \mathrm{End}_K(V)$.*

(i) Ist φ diagonalisierbar, so zerfällt das charakteristische Polynom in Linearfaktoren, d.h. $\chi_\varphi = (t - \lambda_1) \dots (t - \lambda_n)$ mit nicht notwendig verschiedenen $\lambda_1, \dots, \lambda_n \in K$.

(ii) Gilt $\chi_\varphi = (t - \lambda_1) \dots (t - \lambda_n)$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_n \in K$, dann ist φ diagonalisierbar.

Beweis. (i) Ist φ diagonalisierbar, so gibt es gemäß Lemma 6.5 einen Basis von V aus Eigenvektoren zu Eigenwerten $\lambda_i \in K$ mit

$$M_B(\varphi) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

$$\Rightarrow \chi_\varphi = \det(tE_n - M_B(\varphi)) = \det \begin{pmatrix} t - \lambda_1 & & 0 \\ & \ddots & \\ 0 & & t - \lambda_n \end{pmatrix} \overset{(D8)}{=} (t - \lambda_1) \dots (t - \lambda_n).$$

(ii) Aus $\chi_\varphi = (t - \lambda_1) \dots (t - \lambda_n)$ folgt, dass $\lambda_1, \dots, \lambda_n$ Eigenwerte von φ sind und nach Voraussetzung sind diese paarweise verschieden. Nach Korollar 6.7 ist φ dann diagonalisierbar. □

Lemma 6.24. *Seien $\varphi \in \mathrm{End}_K(V)$ und $\lambda_1, \dots, \lambda_r \in K$ paarweise verschiedene Eigenwerte von φ. Dann gilt für alle $i \in \{1, \dots, r\}$*

$$\mathrm{Eig}(\varphi, \lambda_i) \cap \sum_{\substack{j=1: \\ j \neq i}}^{r} \mathrm{Eig}(\varphi, \lambda_j) = \{0\}, \qquad \textit{also} \qquad \sum_{j=1}^{r} \mathrm{Eig}(\varphi, \lambda_j) = \bigoplus_{j=1}^{r} \mathrm{Eig}(\varphi, \lambda_j).$$

Beweis. Angenommen, es existiere $v_i \neq 0$ mit $v_i \in \mathrm{Eig}(\varphi, \lambda_i) \cap \sum_{\substack{j=1: \\ j \neq i}}^{r} \mathrm{Eig}(\varphi, \lambda_j)$.

$\Rightarrow v_i = v_1 + \cdots + v_{i-1} + v_{i+1} + \cdots + v_r$ mit $v_j \in \mathrm{Eig}(\varphi, \lambda_j)$ für alle $j \in \{1, \dots, n\}$.

$\Rightarrow (v_j : v_j \neq 0)$ ist linear abhängig. $\lightning$ (zu Lemma 6.6)

Lemma 3.76 impliziert dann, dass die Summe direkt ist. □

Satz 6.25 (Eigenraumzerlegung)**.** *Seien V endlichdimensional und $\varphi \in \mathrm{End}_K(V)$. Dann sind äquivalent:*

(i) *φ ist diagonalisierbar.*

(ii) *χ_φ zerfällt in Linearfaktoren und es gilt $\mu_{alg}(\varphi, \lambda) = \mu_{geo}(\varphi, \lambda)$ für alle Eigenwerte λ von φ.*

(iii) *Sind $\lambda_1, \dots, \lambda_k \in K$ die paarweise verschiedenen Eigenwerte von φ, so gilt*

$$V = \bigoplus_{j=1}^{k} \mathrm{Eig}(\varphi, \lambda_j).$$

Beweis. (i)⇒(ii): Sei φ diagonalisierbar. Nach Lemma 6.5 besitzt V dann eine Basis B aus Eigenvektoren von φ. Wir ordnen diese Eigenvektoren den verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ von φ zu: Seien $v_1^{(i)}, \dots, v_{s_i}^{(i)} \in \mathrm{Eig}(\varphi, \lambda_i)$, $i \in \{1, \dots, k\}$, und $B_i := (v_1^{(i)}, \dots, v_{s_i}^{(i)})$.

(1) Wir zeigen: B_i ist eine Basis von $\mathrm{Eig}(\varphi, \lambda_i)$.
Da B_i Teilfamilie der Basis B ist, ist auch B_i linear unabhängig. Bleibt der Nachweis, dass B_i Erzeugendensystem von $\mathrm{Eig}(\varphi, \lambda_i)$ ist. Sei dazu $v \in \mathrm{Eig}(\varphi, \lambda_i)$. Da B Basis ist, besitzt v eine Darstellung als Linearkombination aus B, d.h. es existieren $\lambda_m^{(j)} \in K$ mit

$$v = \sum_{j=1}^{k} \left(\lambda_1^{(j)} v_1^{(j)} + \cdots + \lambda_{s_j}^{(j)} v_{s_j}^{(j)} \right).$$

$$\Longrightarrow \qquad \underbrace{v - \left(\lambda_1^{(i)} v_1^{(i)} + \cdots + \lambda_{s_i}^{(i)} v_{s_i}^{(i)} \right)}_{\in \mathrm{Eig}(\varphi, \lambda_i)} = \underbrace{\sum_{j \neq i} \left(\lambda_1^{(j)} v_1^{(j)} + \cdots + \lambda_{s_j}^{(j)} v_{s_j}^{(j)} \right)}_{\in \sum_{j \neq i} \mathrm{Eig}(\varphi, \lambda_j)}.$$

Nach Lemma 6.24 ist der Nullvektor aber der einzige Vektor, der diese Identität erfüllen kann, womit $v = \lambda_1^{(i)} v_1^{(i)} + \cdots + \lambda_{s_i}^{(i)} v_{s_i}^{(i)}$.

(2) Nach (1) gilt

$$\mu_{geo}(\varphi, \lambda_1) + \cdots + \mu_{geo}(\varphi, \lambda_k) = s_1 + \cdots + s_k = \dim V.$$

Nach Lemma 6.23 (i) zerfällt χ_φ in Linearfaktoren, d.h.

$$\mu_{alg}(\varphi, \lambda_1) + \cdots + \mu_{alg}(\varphi, \lambda_i) = \deg(\chi_\varphi) = \dim V.$$

Wegen $\mu_{geo}(\varphi, \lambda_i) \leq \mu_{alg}(\varphi, \lambda_i)\ \forall\, i$ nach Lemma 6.21 folgt daraus $\mu_{geo}(\varphi, \lambda_i) = \mu_{alg}(\varphi, \lambda_i)\ \forall\, i$.

(ii)⇒(iii): Es gelte (ii). Seien $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von φ und

$$W = \bigoplus_{j=1}^{k} \mathrm{Eig}(\varphi, \lambda_j) \quad \text{(die Summe ist direkt nach Lemma 6.24).}$$

Es folgt

$$\dim W = \sum_{j=1}^{k} \mu_{geo}(\varphi, \lambda_j) \overset{\text{Voraussetzung}}{=} \sum_{j=1}^{k} \mu_{alg}(\varphi, \lambda_j) = \deg(\chi_\varphi) = \dim V,$$

was (wegen $W \subset V$) nach Korollar 3.34 (iii) $W = V$ impliziert.

(iii)⇒(i): Für $i = 1, \ldots, k$ sei $B_i := (v_1^{(i)}, \ldots, v_{s_i}^{(i)})$ eine Basis von $\mathrm{Eig}(\varphi, \lambda_i)$. Nach Satz 3.75 (ii) ist

$$B = \left(v_1^{(1)}, \ldots, v_{s_1}^{(1)}, v_1^{(2)}, \ldots, v_{s_2}^{(2)}, \ldots, v_1^{(k)}, \ldots, v_{s_k}^{(k)}\right)$$

eine Basis von V, die nur aus Eigenvektoren von φ besteht. Gemäß Lemma 6.5 ist φ damit diagonalisierbar. □

Beispiele 6.26. *(i) Wir betrachten nochmal die Matrix aus Beispiel 6.22:* $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. *Dort hatten wir bereits ermittelt:*

$$\chi_A = \det \begin{pmatrix} t-1 & -1 \\ 0 & t-1 \end{pmatrix} = (t-1)^2,\ \mu_{alg}(A, 1) = 2,\ \mu_{geo}(A, 1) = 1.$$

Nach Satz 6.25 (ii) ist A nicht diagonalisierbar.

(ii) Sei

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -6 & 1 & 2 \\ 3 & -1 & -2 \end{pmatrix}.$$

Dann ist

$$\chi_A = \det A = \begin{pmatrix} t-2 & 1 & 1 \\ 6 & t-1 & -2 \\ -3 & 1 & t+2 \end{pmatrix} = \cdots = t^3 - t^2 - 5t - 3 \overset{(*)}{=} (t+1)^2(t-3).$$

$\Rightarrow$ *Eigenwerte von* A*:* -1 *und* 3*,* $\mu_{alg}(A,-1) = 2$*,* $\mu_{alg}(A,3) = 1$*.*

Weiter sind

$$\mathrm{Eig}(A,3) = \mathrm{L\ddot{o}s}(3E_3 - A, 0) = \cdots = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right),$$

$$\mathrm{Eig}(A,-1) = \mathrm{L\ddot{o}s}(-E_3 - A, 0) = \cdots = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right).$$

$\Rightarrow$ $\mu_{geo}(A,-1) = 2 = \mu_{alg}(A,-1)$ *und* $\mu_{geo}(A,3) = 1 = \mu_{alg}(A,3)$*.*

Nach Satz 6.25 (ii) ist A *diagonalisierbar und*

$$B = \left(\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right)$$

ist eine Basis des $\mathbb{R}^3$ *aus Eigenvektoren von* A*.*

Bemerkung 6.27. *Im Allgemeinen ist der Schritt* $(*)$ *schwierig, weil es kein Verfahren gibt, um die Nullstellen eines Polynoms vom Grad* n *zu bestimmen. In diesem Fall benötigt man einen numerischen Algorithmus.*

6.3 Satz von Cayley-Hamilton und Minimalpolynom

Motivation 6.28. *Für eine diagonalisierbare Matrix* $A \in M(n \times n, K)$ *gilt*

$$A = S^{-1} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} S, \quad \text{d.h. } A^j = \underbrace{A \cdot \cdots \cdot A}_{j\text{-mal}} = S^{-1} \begin{pmatrix} \lambda_1^j & & 0 \\ & \ddots & \\ 0 & & \lambda_n^j \end{pmatrix} S.$$

Gilt $\chi_A(t) = \sum_{j=0}^n c_j t^j$ *mit* $c_n = 1$*, so folgt daraus*

$$\chi_A(A) = \sum_{j=0}^n c_j S^{-1} \begin{pmatrix} \lambda_1^j & & 0 \\ & \ddots & \\ 0 & & \lambda_n^j \end{pmatrix} S = S^{-1} \begin{pmatrix} \chi_A(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \chi_A(\lambda_n) \end{pmatrix} S = 0.$$

Wir werden zeigen, dass dieses Ergebnis auch für beliebige Matrizen $A \in M(n \times n, K)$ *und Endomorphismen* $\varphi \in \mathrm{End}_K(V)$ *(*$\dim V < \infty$*) gilt.*

Bemerkung 6.29. *Gemäß Lemma 4.6 (ii) ist* $(\mathrm{End}_K(V), +, \circ)$ *ein Ring mit Einselement* Id_V*. Für* $f \in K[t]$*,* $f = a_m t^m + \cdots + a_1 t + a_0 = a_m t^m + \cdots + a_1 t + a_0 t^0$*, erhalten wir mit* $\varphi \in \mathrm{End}_K(V)$ *als Unbestimmte*

$$f(\varphi) = a_m \varphi^m + \cdots + a_1 \varphi + a_0 \varphi^0 \in \mathrm{End}_K(V),$$

wobei $\varphi^k = \underbrace{\varphi \circ \cdots \circ \varphi}_{k\text{-mal}}$ *und* $\varphi^0 := \mathrm{Id}_V$ *(d.h. wir ersetzen* $a_0 = a_0 \cdot 1$ *durch* $a_o \cdot 1_{\mathrm{End}_K(V)}$*).*
Es gilt für $f, g \in K[t]$ *und* $\varphi \in \mathrm{End}_K(V)$ *(Blatt 5, Aufgabe 2 (ii)):*

$$f(\varphi) \circ g(\varphi) = (f \cdot g)(\varphi) = (g \cdot f)(\varphi) = g(\varphi) \circ f(\varphi).$$

Vorbetrachtung 6.30. *Seien* $n = \dim V < \infty$ *und* $\varphi \in \mathrm{End}_K(V)$*. Wir betrachten für festes* $v \neq 0$ *die Vektoren*

$$v, \varphi(v), \varphi^2(v), \ldots$$

Wegen $n = \dim V$ *hat eine Familie unabhängiger Vektoren höchstens* n *Elemente, womit* $(v, \varphi(v), \ldots, \varphi^n(v))$ *linear abhängig ist. Also gibt es ein* $r \in \mathbb{N}$*,* $r \leq n$*, so dass*

$$\begin{aligned} &(v, \varphi(v), \ldots, \varphi^{r-1}(v)) \text{ linear unabhängig, aber} \\ &(v, \varphi(v), \ldots, \varphi^{r}(v)) \text{ linear abhängig sind.} \end{aligned} \tag{6.1}$$

Der Vektor $\varphi^r(v)$ *ist folglich Linearkombination aus* $v, \varphi(v), \ldots, \varphi^{r-1}(v)$*, d.h. es existieren* $c_0, \ldots, c_{r-1} \in K$ *mit*

$$\varphi^r(v) = \sum_{j=0}^{r-1} c_j \varphi^j(v).$$

Sei nun $U = \mathrm{Lin}(v, \varphi(v), \ldots, \varphi^{r-1}(v))$*. Wegen (6.1) ist die Familie* $B' = (b_1, \ldots, b_r)$ *mit* $b_1 = v, b_2 = \varphi(v), \ldots, b_r = \varphi^{r-1}(v)$ *eine Basis von* U *und es gelten* $\varphi(b_i) = b_{i+1}$ *für* $1 \leq i \leq r$ *sowie*

$$\varphi(b_r) = \varphi^r(v) = \sum_{j=0}^{r-1} c_j \varphi^j(v) = \sum_{j=0}^{r-1} c_j b^{j+1}. \tag{6.2}$$

Damit ist $\varphi(U) \subset U$ *und* $\varphi_U := \varphi_{|U} \in \mathrm{End}_K(U)$*.* φ_U *hat dann bezüglich der Basis* B' *von* U *die Darstellungsmatrix*

$$A_U = M_{B'}(\varphi_U) = \begin{pmatrix} 0 & \ldots & 0 & c_0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & \vdots \\ 0 & & 1 & c_{r-1} \end{pmatrix}. \tag{6.3}$$

Es folgt

$$\chi_{\varphi_U} = \det(tE_r - A_U) = \det \begin{pmatrix} t & \dots & 0 & -c_0 \\ -1 & \ddots & \vdots & \vdots \\ & \ddots & t & -c_{r-2} \\ 0 & & -1 & t - c_{r-1} \end{pmatrix} \overset{\substack{\text{Entwicklung} \\ \text{nach der} \\ \text{letzten Spalte}}}{=} t^r - \sum_{j=0}^{r-1} c_j t^j$$

und zusammen mit (6.2) $\chi_{\varphi_U}(\varphi_U)(v) = 0$. *Es folgt sogar* $\chi_{\varphi_U}(\varphi_U) = 0 \in \mathrm{End}_K(U)$, *denn ist* $w = \sum_{i=1}^r \lambda_i w_i \in U$ *mit* $\lambda_1, \dots, \lambda_r \in K$, *so ergibt sich*

$$\begin{aligned} \chi_{\varphi_U}(\varphi_U)(w) &= \varphi_U^r(w) - \sum_{j=0}^{r-1} c_j \varphi_U^j(w) \\ &= \sum_{i=1}^r \lambda_i \Big(\varphi_U^r(b_i) - \sum_{j=0}^{r-1} c_j \varphi_U^j(b_i) \Big) \\ &= \sum_{i=1}^r \lambda_i \Big(\varphi_U^{r+i-1}(v) - \sum_{j=0}^{r-1} c_j \varphi_U^{j+i-1}(v) \Big) \\ &= \sum_{i=1}^r \lambda_i \varphi_U^{i-1} \Big(\underbrace{\varphi_U^r(v) - \sum_{j=0}^{r-1} c_j \varphi_U^j(v)}_{=0} \Big) = 0. \end{aligned}$$

Man beachte, dass hier U *und* r *von* v *abhängen.*

Satz 6.31 (Cayley-Hamilton)**.** *Seien* $n = \dim V < \infty$ *und* $\varphi \in \mathrm{End}_K(V)$. *Dann gilt*

$$\chi_\varphi(\varphi) = 0.$$

Insobesondere gilt $\chi_A(A) = 0$ *für alle* $A \in M(n \times n, K)$.

Beweis. Wir zeigen $\chi_\varphi(\varphi)(v) = 0$ für alle $v \in V$. Für $v = 0$ ist die Aussage korrekt, denn $\chi_\varphi(\varphi)$ ist K-linear. Sei also $v \neq 0$. Sei U der zu diesem v in der Vorbetrachtung definierte Teilraum mit der Basis $(b_1, \dots, b_r)$ wie oben. Ergänzen wir Letztere zu einer Basis $B = (b_1, \dots, b_n)$ von V, so besitzt die Darstellungsmatrix $M_B(\varphi)$ die Blockgestalt

$$M_B(\varphi) = \begin{pmatrix} A_U & C \\ 0 & A' \end{pmatrix} \text{ mit } A_U \text{ aus (6.3)}.$$

Es folgt $\chi_\varphi \overset{\substack{\text{Kästchen-} \\ \text{formel}}}{=} \chi_{A_U} \cdot \chi_{A'} = \chi_{A'} \cdot \chi_{A_U}$ und damit aus Vorbetrachtung 6.30

$$\chi_\varphi(\varphi)(v) = \Big(\chi_{A'}(\varphi) \circ \chi_{A_U}(\varphi) \Big)(v) = \chi_{A'}(\varphi)\Big(\chi_{A_U}(\varphi)(v) \Big) = \chi_{A'}(\varphi)\Big(\underbrace{\chi_{\varphi_U}(\varphi)(v)}_{\substack{=\chi_{\varphi_U}(\varphi_U)(v) \\ =0}} \Big) = 0.$$

Für $A \in M(n \times n, K)$ gilt mit der zugehörigen linearen Abbildung $\tilde{A} \in \mathrm{End}_K(K^n)$ und jedes $x \in K^n$:

$$\chi_A(A) \cdot x = \chi_{\tilde{A}}(\tilde{A})(x) = 0.$$

□

Bemerkung 6.32. *Der "Beweis"*

$$\chi_A(A) = \big(\det(tE_n - A)\big)(A) = \det(AE_n - A) = \det(A - A) = \det(0) = 0$$

ist ***falsch****, denn*

$$\underbrace{\underbrace{\det(tE_n - A)}_{\in K[t]}(A)}_{\in M(n\times n,K)} \neq \underbrace{\det(\underbrace{AE_n - A}_{\in (M(n\times n,K)}) .}_{\in K}$$

Satz und Definition 6.33. *Seien $n = \dim V < \infty$, $\varphi \in \mathrm{End}_K(V)$ und*

$$G = \{g \in K[t] |\, g(\varphi) = 0\}.$$

Dann gibt es ein eindeutig bestimmtes normiertes Polynom μ_φ kleinsten Grades mit $\mu_\varphi(\varphi) = 0$. μ_φ heißt Minimalpolynom von φ. Es gilt $G = \mu_\varphi K[t] := \{\mu_\varphi Q |\, Q \in K[t]\}$. μ_φ ist auch das eindeutig bestimmte normierte Polynom, mit welchem $G = \mu_\varphi K[t]$ gilt.

Für $A \in M(n \times n, K)$ definiert man das Minimalpolynom μ_A durch $\mu_A := \mu_{\tilde{A}}$. μ_A hat dann die zu μ_φ analogen Eigenschaften.

Zur Erinnerung: Wir hatten deg(Nullpolynom) $= \infty$ gesetzt.

Beweis. Sei $\mu_\varphi \in G$ ein Polynom minimalen Grades. Nach dem Satz von Cayley-Hamilton ist $\chi_\varphi \in G$, also insbesondere $G \neq \emptyset$ und $\deg(\mu_\varphi) \leq n$. Ohne Einschränkung kann man annehmen, dass μ_φ normiert ist, d.h. $\mu_\varphi = \sum_{k=0}^{d} c_k t^k$ mit $c_d = 1$ für ein $k \in \mathbb{N}$. Sei $g \in G$ beliebig. Nach Satz 2.34 (ii) ergibt Division mit Rest

$$g = q\mu_\varphi + r \quad \text{mit} \quad \deg(r) < \deg(\mu_\varphi) \text{ falls } r \neq 0$$

und wegen $g \in G$ ist

$$0 = g(\varphi) = \underbrace{q(\varphi) \circ \underbrace{\mu_\varphi(\varphi)}_{=0}}_{=0} + r(\varphi),$$

womit $r(\varphi) = 0 \in \mathrm{End}_K(V)$, also $r \in G$ und $\deg(r) < \deg(\mu_\varphi)$ falls $r \neq 0$. Da $\mu_\varphi \in G$ minimalen Grad hat, muss $r = 0 \in K[t]$ gelten. Folglich ist $g = q\mu_\varphi$, also entweder $g = \lambda\mu_\varphi$ mit $\lambda \in K \setminus \{0\}$ oder $\deg(g) > \deg(\mu_\varphi)$, womit μ_φ als normiertes Polynom

kleinsten Grades eindeutig ist und $G \subset \{\mu_\varphi q \mid q \in K[t]\}$ gezeigt ist. Die umgekehrte Inklusion gilt auch, weil $(\mu_\varphi q)(\varphi) = (q\mu_\varphi)(\varphi) = q(\varphi) \circ \mu_\varphi(\varphi) = 0 \in \mathrm{End}_K(V)$.

Bleibt zu zeigen, dass μ_φ das eindeutig bestimmte normierte Polynom ist, mit welchem $G = \mu_\varphi K[t]$ gilt. Angenommen, es gibt ein weiteres normiertes Polynom $h \in K[t]$, mit welchem $hK[t] = G = \mu_\varphi K[t]$ gilt. Wegen $\mu_\varphi \in G$ existiert ein $P \in K[t]$ mit $\mu_\varphi = hP$, weiter ist wegen $h = h \cdot 1$ auch $h \in G$, womit $h = \mu_\varphi Q$ für ein $Q \in K[t]$ ist. Also gilt

$$\mu_\varphi = hP = \mu_\varphi QP,$$

womit $PQ = 1$ und damit $P, Q \in K$. Da aber beide, h und μ_φ, normiert sind, folgt $P = Q = 1$, also $h = \mu_\varphi$. □

Lemma 6.34. *Seien V endlichdimensional, $\varphi \in \mathrm{End}_K(V)$ und $\lambda \in K$. Dann gilt:*

$$\chi_\varphi(\lambda) = 0 \;\Leftrightarrow\; \mu_\varphi(\lambda) = 0,$$

d.h. χ_φ und μ_φ haben dieselben Nullstellen.

Beweis. "⇐": Sei $\mu_\varphi(\lambda) = 0$. Nach Satz 6.33 existiert $Q \in K[t]$ mit $\chi_\varphi = \mu_\varphi \cdot Q$. Es folgt $\chi_\varphi(\lambda) = \underbrace{\mu_\varphi(\lambda)}_{=0} \cdot Q(\lambda) = 0$.

"⇒": Sei $\chi_\varphi(\lambda) = 0$. Dann ist λ Eigenwert von φ (Satz 6.19 (ii)) mit einem Eigenvektor $v \in V$ $(v \neq 0)$. Sei $\mu_\varphi = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$. Dann gilt

$$\begin{aligned}\mu_\varphi(\varphi)(v) &= \big(\varphi^r + a_{r-1}\varphi^{r-1} + \cdots + a_1\varphi + a_0 \,\mathrm{Id}_V\big)(v)\\ &= \big(\lambda^r + a_{r-1}\lambda^{r-1} + \cdots + a_1\lambda + a_0\big)v = \mu_\varphi(\lambda)v.\end{aligned}$$

Wegen $v \neq 0$ und $\mu_\varphi(\varphi) = 0$ folgt $\mu_\varphi(\lambda) = 0$. □

Beispiele 6.35. *(i)* $A = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right) \in M(2\times 2, \mathbb{R}) \Rightarrow \chi_A = (t-1)^2$. *Es gelten:*

- *μ_A ist normiert (gemäß Definition),*
- *$\mu_A \mid \chi_A$ ("μ_A teilt χ_A") (nach Satz 6.33)*
- *$\mu_A(1) = 0$ (nach Lemma 6.34).*

$\Rightarrow \mu_A \in \{(t-1), (t-1)^2\}$.

Einsetzen von A in $(t-1)$ zeigt: $(t-1)(A) = A - E_2 = 0$, *d.h.* $\mu_A = (t-1)$.

(ii) $A = \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right) \in M(2\times 2, \mathbb{R}) \Rightarrow \chi_A = t^2 - 1 = (t-1)(t+1)$. *Nach Satz 6.33 und Lemma 6.34 ist* $\mu_A = (t-1)(t+1)$.

(iii) Sei

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -8 & 1 & 4 \\ 2 & -1 & -1 \end{pmatrix} \implies \chi_A = \cdots = (t+1)^2(t-3).$$

Nach Satz 6.33 und Lemma 6.34 folgt $\mu_A \in \{(t+1)(t-3), (t+1)^2(t-3)\}$. *Einsetzen von* A *in* $(t+1)(t-3)$ *zeigt:* $(A+E_3)(A-3E_3) = \cdots \neq 0$, *womit* $\mu_A = (t+1)^2(t-3)$.

(iv) Sei

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -6 & 1 & 2 \\ 3 & -1 & -2 \end{pmatrix} \overset{\text{Bsp. 6.26(ii)}}{\implies} \chi_A = \cdots = (t+1)^2(t-3).$$

Nach Satz 6.33 und Lemma 6.34 folgt $\mu_A \in \{(t+1)(t-3), (t+1)^2(t-3)\}$. *Einsetzen von* A *in* $(t+1)(t-3)$ *zeigt:* $(A+E_3)(A-3E_3) = \cdots = 0$, *womit* $\mu_A = (t+1)(t-3)$.

Satz 6.36. *Seien* V *endlichdimensional und* $\varphi \in \mathrm{End}_K(V)$. *Dann sind äquivalent:*

(i) φ *ist diagonalisierbar.*

(ii) Das Minimalpolynom μ_φ *zerfällt in Linearfaktoren und besitzt nur einfache Nullstellen, d.h.* $\mu_\varphi = (t-\lambda_1)\dots(t-\lambda_r)$ *mit paarweise verschiedenen* $\lambda_1, \dots, \lambda_r \in K$.

Beweis. (i)⇒(ii): Sei $\{\lambda_1, \dots, \lambda_r\}$ mit paarweise verschiedenen λ_i die Menge der Eigenwerte von φ.

$$\varphi \text{ diagonalisierbar} \overset{\text{Satz 6.25}}{\implies} V = \bigoplus_{i=1}^{r} \mathrm{Eig}(\varphi, \lambda_i),$$

d.h. jedes $v \in V$ besitzt eine (eindeutige) Darstellung $v = v_1 + \cdots + v_r$ mit $v_i \in \mathrm{Eig}(\varphi, \lambda_i)$. Wir definieren nun $h \in \mathrm{End}_K(V)$ durch Einsetzen von φ als Unbestimmte in das Polynom $(t-\lambda_1)\dots(t-\lambda_r)$:

$$h := (\varphi - \lambda_1 \mathrm{Id}_V) \circ (\varphi - \lambda_2 \mathrm{Id}_V) \circ \cdots \circ (\varphi - \lambda_r \mathrm{Id}_V)$$

Mit $h_{-j} := (\varphi - \lambda_1 \mathrm{Id}_V) \circ \cdots \circ (\varphi - \lambda_{j-1} \mathrm{Id}_V) \circ (\varphi - \lambda_{j+1} \mathrm{Id}_V) \cdots \circ (\varphi - \lambda_r \mathrm{Id}_V)$, $1 \le j \le r$, folgt für $v = v_1 + \cdots + v_r$ mit $v_i \in \mathrm{Eig}(\varphi, \lambda_i)$

$$\begin{aligned} h(v) &= \Big((\varphi - \lambda_1 \mathrm{Id}_V) \circ (\varphi - \lambda_2 \mathrm{Id}_V) \circ \cdots \circ (\varphi - \lambda_r \mathrm{Id}_V)\Big)(v_1 + \cdots + v_r) \\ &= \sum_{j=1}^{r} \Big((\varphi - \lambda_1 \mathrm{Id}_V) \circ (\varphi - \lambda_2 \mathrm{Id}_V) \circ \cdots \circ (\varphi - \lambda_r \mathrm{Id}_V)\Big)(v_j) \\ &\overset{\text{Bem.6.29}}{=} \sum_{j=1}^{r} h_{-j} \circ (\varphi - \lambda_j \mathrm{Id}_V)(v_j) = \sum_{j=1}^{r} h_{-j}\Big(\underbrace{(\varphi - \lambda_j \mathrm{Id}_V)(v_j)}_{=0}\Big) = 0. \end{aligned}$$

Also gilt $h = 0$ und damit $\mu_\varphi | (t-\lambda_1)\dots(t-\lambda_r)$ nach Satz 6.33. Nach Satz 6.19 (ii) und Lemma 6.34 folgt $\mu_\varphi(\lambda_j) = 0$ für alle $j \in \{1,\dots,r\}$, womit $\mu_\varphi = (t-\lambda_1)\dots(t-\lambda_r)$.

(ii)⇒(i): Nach (ii) gilt $\mu_\varphi = (t-\lambda_1)\dots(t-\lambda_r)$ mit paarweise verschiedenen $\lambda_j \in K$. Wir folgern daraus (i) mittels vollständiger Induktion nach $n = \dim V$.

Induktionsanfang: Für $n = 1$ ist nichts zu zeigen, denn jedes φ ist diagonalisierbar, da jedes $v \in V \setminus \{0\}$ Eigenvektor von jedem $\varphi \in \mathrm{End}_K(V)$ ist.

Für den Induktionsschritt beweisen wir zunächst

$$V = \mathrm{Kern}(\varphi - \lambda_1 \mathrm{Id}_V) \oplus \mathrm{Bild}(\varphi - \lambda_1 \mathrm{Id}_V). \tag{6.4}$$

Nach Satz 2.34 (ii) existieren $q, s \in K[t]$ mit

$$(t-\lambda_2)\dots(t-\lambda_r) = q(t-\lambda_1) + s \quad \text{mit} \quad \deg(s) < \underbrace{\deg(t-\lambda_1)}_{=1} \text{ falls } s \neq 0.$$

Wegen

$$s(\lambda_1) = \prod_{j=2}^{r}(\lambda_1 - \lambda_j) \; - q(\lambda_1)\underbrace{(\lambda_1 - \lambda_1)}_{=0} \neq 0$$

folgt $s \neq 0$. Folglich gilt $\deg(s) = 0$ und s ist von der Form $s = s_0 t^0$ mit $s_0 \in K \setminus \{0\}$. Einsetzen von φ in s ergibt:

$$v = \underbrace{\frac{1}{s_0}\Big((\varphi - \lambda_2 \mathrm{Id}_V) \circ \dots \circ (\varphi - \lambda_r \mathrm{Id}_V)\Big)(v)}_{=:u} - \underbrace{\frac{1}{s_0} q(\varphi) \circ (\varphi - \lambda_1 \mathrm{Id}_V)(v)}_{=:w}.$$

Es gilt

$$(\varphi - \lambda_1 \mathrm{Id}_V)(u) = \frac{1}{s_0}\underbrace{\mu_\varphi(\varphi)}_{=0}(v) = 0,$$

womit $u \in \mathrm{Kern}(\varphi - \lambda_1 \mathrm{Id}_V)$. Außerdem ist

$$w \overset{\text{Bem. 6.29}}{=} \frac{1}{s_0}(\varphi - \lambda_1 \mathrm{Id}_V) \circ q(\varphi)(v) = \underbrace{(\varphi - \lambda_1 \mathrm{Id}_V)\Big(\underbrace{q(\varphi)(v)}_{\in V}\Big)}_{\in \mathrm{Bild}(\varphi - \lambda_1 \mathrm{Id}_V)}.$$

$\Longrightarrow V = \mathrm{Kern}(\varphi - \lambda_1 \mathrm{Id}_V) + \mathrm{Bild}(\varphi - \lambda_1 \mathrm{Id}_V)$. Nach der Dimensionsformel (Satz 4.13) gilt aber $\dim V = \dim \mathrm{Kern}(\varphi - \lambda_1 \mathrm{Id}_V) + \dim \mathrm{Bild}(\varphi - \lambda_1 \mathrm{Id}_V)$, woraus nach Satz 3.70 (iv) die Summe sogar direkt ist, d.h. Identität (6.4) ist verifiziert.

Induktionsschritt: Sei die Diagonalisierbarkeit unter Voraussetzung (ii) für $\dim V \leq n-1$ bewiesen. Da $\mathrm{Kern}(\varphi - \lambda_1 \mathrm{Id}) = \mathrm{Eig}(\varphi, \lambda_1) \neq \{0\}$ ist, folgt für $W = \mathrm{Bild}(\varphi - \lambda_1 \mathrm{Id}_V)$ aus der Dimensionsformel: $\dim W \leq n-1$. Wegen

$$\varphi(W) = \varphi \circ (\varphi - \lambda_1 \operatorname{Id}_V)(V) \overset{\text{Bem. } 6.29}{=} (\varphi - \lambda_1 \operatorname{Id}_V) \circ \varphi(V) = (\varphi - \lambda_1 \operatorname{Id}_V)\big(\overbrace{\varphi(V)}^{\subset V}\big) \subset W$$

ist die Einschränkung $\varphi_{|W} \in \operatorname{End}_K(W)$.

$\Longrightarrow$ $\mu_\varphi(\varphi_{|W})(W) = \overbrace{\mu_\varphi(\varphi)}^{=0}(W) = 0.$

$\Longrightarrow$ $\mu_{\varphi_{|W}} | \mu_\varphi = (t - \lambda_1) \dots (t - \lambda_r)$ nach Satz 6.33.

$\Longrightarrow$ $\mu_{\varphi_{|W}}$ zerfällt in Linearfaktoren und besitzt nur einfache Nullstellen.

Nach Induktionsvoraussetzung ist $\varphi_{|W}$ diagonalisierbar, d.h. es gibt eine Basis von W, die nur aus Eigenvektoren von $\varphi_{|W}$ besteht. Wegen $V = \operatorname{Eig}(\varphi, \lambda_1) \oplus W$ nach (6.4) existiert damit auch eine Basis von V aus Eigenvektoren von φ, d.h. φ ist diagonalisierbar. □

Beispiele 6.37. *Wir betrachten nochmal die Matrizen aus Beispielen 6.35 (iii) und (iv).*

(i) Sei

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -8 & 1 & 4 \\ 2 & -1 & -1 \end{pmatrix} \in M(3 \times 3, \mathbb{R}).$$

Wir hatten gezeigt: $\mu_A = (t+1)^2(t-3)$. Nach Satz 6.36 ist A nicht diagonalisierbar.

(ii) Sei

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -6 & 1 & 2 \\ 3 & -1 & -2 \end{pmatrix} \in M(3 \times 3, \mathbb{R}).$$

Wir hatten gezeigt: $\mu_A = (t+1)(t-3)$. Nach Satz 6.36 ist A diagonalisierbar.

Bemerkung 6.38. *Ist $\varphi \in \operatorname{End}_K(V)$ mit $n = \dim V < \infty$, so ist das Minimalpolynom μ_φ nach Satz 6.33 Teiler von χ_φ. Man kann zeigen, dass χ_φ im Gegenzug Teiler von μ_φ^n ist. Im Falle $K = \mathbb{C}$ ist das unmittelbar ersichtlich, weil χ_φ dann in Linearfaktoren zerfällt und dieselben Nullstellen wie μ_φ besitzt (Blatt 6, Aufgabe 3). Im Allgemeinen benötigt man hierfür den Zerfällungskörper für φ (siehe Definition 6.41). Der Beweis von $\chi_\varphi | \mu_\varphi^n$ folgt aber auch später aus der Frobenius-Normalform-Darstellung (Satz 10.11).*

6.4 Trigonalisierbarkeit und nilpotente Abbildungen

Seien $\dim V < \infty$ und $\varphi \in \operatorname{End}_K(V)$. Nach Satz 6.25 gilt:

$$\varphi \text{ diagonalisierbar}$$
$$\Updownarrow$$
χ_φ zerfällt in Linearfaktoren und $\mu_{geo}(\varphi, \lambda) = \mu_{alg}(\varphi, \lambda)$ für alle Eigenwerte λ von φ.

D.h. es kann zwei Probleme geben:

(i) χ_φ zerfällt nicht in Linearfaktoren.
Lösung: Übergang zu einem größeren Körper, bspw. von $\mathbb{R}$ nach $\mathbb{C}$.

(ii) $\mu_{geo}(\varphi, \lambda) < \mu_{alg}(\varphi, \lambda)$ für einen Eigenwert λ von φ.
Hier kann man nichts machen, d.h. φ ist dann wirklich nicht diagonalisierbar.

Wir benötigen aber "vereinfachende Strukturen" für viele Anwendungen. Deshalb betrachten wir statt dessen die sogenannte *Trigonalisierbarkeit* und später dann die Jordansche Normalform als eine spezielle Form hiervon.

Satz 6.39. *Seien $n = \dim V < \infty$ und $\varphi \in \mathrm{End}_K(V)$. Dann sind äquivalent:*

(i) Das charakteristische Polynom χ_φ zerfällt in Linearfaktoren, d.h.

$$\chi_\varphi = (t - \lambda_1) \dots (t - \lambda_n)$$

(wobei die Eigenwerte λ_i hier nicht notwendig verschieden sind).

(ii) Es gibt eine Basis B von V mit

$$M_B(\varphi) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Beweis. (ii)⇒(i): Klar nach (D8) aus Satz 5.4.

(i)⇒(ii): Wir beweisen die Aussage mittels vollsändiger Induktion nach n.

Induktionsanfang: $n = 1$ – offensichtlich korrekt.

Induktionsschritt: Sei die Aussage $\forall\, \varphi \in \mathrm{End}_K(V)$ und $\dim V \leq n - 1$ bereits bewiesen. Sei $b_1 \neq 0$ ein Eigenvektor zum Eigenwert λ_1. Wir ergänzen b_1 zu einer Basis $B' = (b_1, b_2, \dots, b_n)$ von V. Es gilt

$$M_{B'}(\varphi) = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & C \end{array}\right) \quad \text{mit } C \in M\big((n-1) \times (n-1), K\big).$$

Zu zeigen: C ist bei geeigneter Wahl von $b_2, \dots, b_n$ eine obere Dreiecksmatrix. Sei

$$W := \mathrm{Lin}(b_2, \dots, b_n).$$

Für $w \in W$ existieren $a_1(w), \dots, a_n(w) \in K$ mit $\varphi(w) = a_1(w)b_1 + \underbrace{\sum_{i=2}^{n} a_i(w)b_i}_{=:\,\psi(w)}$.

Dann ist $\psi \in \mathrm{End}_K(W)$, der bezüglich der Basis $(b_2, \ldots, b_n)$ von W genau die Matrix C als Darstellungsmatrix besitzt. Aus der Kästchenformel folgt

$$\chi_\varphi = (t - \lambda_1) \cdot \chi_C = (t - \lambda_1) \cdot \chi_\psi$$

und damit (nach (i)) $\chi_\psi = (t - \lambda_2) \ldots (t - \lambda_n)$. Nach Induktionsvoraussetzung kann man deshalb die Basis $(b_2, \ldots, b_n)$ von W so wählen, dass C eine obere Dreiecksmatrix ist. □

Bemerkung 6.40. *Für Matrizen $A \in M(n \times n, K)$ liefert Satz 6.39 die Äquivalenz von*

(i) χ_A *zerfällt in Linearfaktoren;*

(ii) $\tilde{A} : K^n \to K^n$ *ist trigonalisierbar;*

(iii) A *ist ähnlich zu einer oberen Dreiecksmatrix, d.h. es existiert* $S \in GL(n, K)$ *mit*

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Bemerkung und Definition 6.41. *Seien* $\dim V < \infty$ *und* $\varphi \in \mathrm{End}_K(V)$. *Der Körper K heißt Zerfällungskörper für φ, wenn sein charakteristisches Polynom χ_φ über K in Linearfaktoren zerfällt.* $\mathbb{C}$ *ist nach dem Fundamentalsatz der Algebra (Satz 2.38) Zerfällungskörper für jedes* $\varphi \in \mathrm{End}_{\mathbb{C}}(V)$. $\mathbb{R}$ *ist beispielweise kein Zerfällungskörper für* $\tilde{A} : \mathbb{R}^2 \to \mathbb{R}^2$ *mit* $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, *da* $\chi_A = t^2 + 1$. *In der Vorlesung "*Algebra und Zahlentheorie*" konstruiert man zu jedem Körper K seinen algebraischen Abschluss $\bar{K}$, in dem jedes Polynom aus $K[t]$ in Linearfaktoren zerfällt. Insbesondere ist $\bar{K}$ ein Zerfällungskörper.*

Beispiel 6.42. *Sei*

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 2 & 2 & 3 \end{pmatrix} \in M(3 \times 3, \mathbb{R}).$$

Es gilt $\chi_A = \cdots = (t-1)(t-3)^2$, *d.h. die Eigenwerte von A sind* 1 *und* 3 *und*

$$\mu_A \in \left\{(t-1)(t-3), (t-1)(t-3)^2\right\}.$$

Es gilt $(A - E_3)(A - 3E_3) = \cdots \neq 0$, *womit* $\mu_A = (t-1)(t-3)^2$. $\overset{\text{Satz 6.36}}{\Longrightarrow}$ *A ist nicht diagonalisierbar.*

$$\mathrm{Eig}(A, 1) = \mathrm{Lös}(E_3 - A, 0) = \cdots = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right).$$

Mit $b_1 := (1,0,-1)^t$ *ist* $B' = (b_1, e_2, e_3)$ *eine Basis des* $\mathbb{R}^3$*, bezüglich welcher* $\tilde{A}$ *die Darstellungsmatrix*

$$M_{B'}(\tilde{A}) = \left(\begin{array}{c|cc} 1 & -1 & 1 \\ \hline 0 & 2 & -1 \\ 0 & 1 & 4 \end{array}\right).$$

besitzt. Wir müssen nun die Matrix $C = \left(\begin{smallmatrix} 2 & -1 \\ 1 & 4 \end{smallmatrix}\right)$ *trigonalisieren. Es gelten* $\chi_C = \cdots = (t-3)^2$ *und*

$$\mathrm{Eig}(C,3) = \mathrm{L\ddot{o}s}(3E_2 - C, 0) = \mathrm{L\ddot{o}s}\left(\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, 0\right) = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right).$$

Bezüglich der Basis (b_1, b_2, b_3) *mit* $b_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ *und* $b_3 = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *ist*

$$M_B(\tilde{A}) = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 3 \end{pmatrix}. \quad \textit{[Man beachte, dass sich auch die erste Zeile verändert hat!]}$$

Im späteren Kapitel über "*Jordansche Normalformen*" werden wir die Einträge der Trigonalmatrix genauer untersuchen bzw. eine Basis in geeignetem Sinne optimal aussuchen. Dazu wird der Begriff der nilpotenten Abbildung benötigt.

Definition 6.43. $\varphi \in \mathrm{End}_K(V)$ *heißt nilpotent, falls es ein* $k \in \mathbb{N}$ *gibt mit* $\phi^k = 0$.

Lemma 6.44. *Sei* $n = \dim V < \infty$*. Folgende Aussagen sind äquivalent:*

(i) φ *ist nilpotent.*

(ii) $\chi_\varphi = t^n$.

(iii) Es gibt eine Basis B *mit* $M_B(\varphi) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$.

(iv) $\varphi^n = 0$.

Beweis. (i)⇒(ii): $\exists\, k \in \mathbb{N}$ mit $\varphi^k = 0 \Rightarrow \mu_\varphi | t^k \Rightarrow \mu_\varphi = t^l$ mit $l \le k \overset{\text{Bem. } 6.38}{\Rightarrow} \chi_\varphi = t^n$.

(ii)⇒(iii) ist unmittelbare Konsequenz aus Satz 6.39.

(iii)⇒(ii) klar.

(ii)⇒(iv) folgt aus dem Satz von Cayley-Hamilton.

(iv)⇒(i): Definition von "nilpotent". □

Bemerkung 6.45. *Für die Dreiecksmatrix aus Satz 6.39 gilt*

$$\begin{aligned} M_B(\varphi) &= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} + \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} \\ &=: \quad M_B(\varphi_1) \quad + \quad M_B(\varphi_2), \end{aligned}$$

d.h. wir haben $\varphi = \varphi_1 + \varphi_2$ bezüglich B in eine diagnoalisierbare Abbildung φ und eine nilpotente Abbildung φ_2 zerlegt.

7 Euklidische und unitäre Vektorräume

In diesem Kapitel seien K ein Körper mit $1+1 \neq 0$ (Charakteristik$(K) \neq 2$) und V ein K-Vektorraum.

7.1 Symmetrische Bilinearformen

Motivation 7.1. *Wir möchten nun zusätzlich ein Konzept für die Länge von Vektoren und bspw. für die Abstände zwischen Vektoren entwickeln. Zum Beispiel ist im $\mathbb{R}^2$ der euklidische Abstand zwischen $P_1 = \binom{x_1}{x_2}$ und $P_2 = \binom{y_1}{y_2}$ gegeben durch*

$$\|P_1 - P_2\| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2} \quad \textit{mit} \quad \|P_1\|^2 = x_1^2 + x_2^2 = (x_1, x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Eng damit verbunden ist das Skalarprodukt, welches wir daraus durch

$$\begin{aligned} \langle P_1, P_2 \rangle :&= \frac{1}{2}\Big(\|P_1+P_2\|^2 - \|P_1\|^2 - \|P_2\|^2\Big) \\ &= \frac{1}{2}\Big((x_1+y_1)^2 + (x_2+y_2)^2 - (x_1^2+x_2^2) - (y_1^2+y_2^2)\Big) \\ &= x_1y_1 + x_2y_2 \\ &= (x_1, x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned}$$

definieren können. Will man die Komponenten gewichten, bspw. weil x_1 und x_2 in verschiedenen Einheiten gemessen werden (wie cm und m), wählt man bspw. $\|P_1\|^2 = (x_1, x_2)\left(\begin{smallmatrix} 1 & 0 \\ 0 & 100^2 \end{smallmatrix}\right)\binom{x_1}{x_2}$ oder man macht eventuell vorher noch eine Basistransformation

$$\|P_1\|^2 = (x_1, x_2)\underbrace{S^t\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}S}_{=:A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Das zugehörige Skalarprodukt $\langle P_1, P_2\rangle = (x_1, x_2)A\binom{y_1}{y_2}$ *definiert eine symmetrische Bilinearform.*

Definition 7.2. *Eine Abbildung* $\gamma : V \times V \to K$ *heißt* Bilinearform *auf* V*, falls* $v \mapsto \gamma(v, w)$ *bei festem* w *linear in* v *und* $w \mapsto \gamma(v, w)$ *bei festem* v *linear in* w *ist, d.h.*

(i) $\gamma(v_1 + v_2, w) = \gamma(v_1, w) + \gamma(v_2, w)$ *für alle* $v_1, v_2, w \in V$ *und*
$\gamma(\lambda v, w) = \lambda\gamma(v, w)$ *für alle* $\lambda \in K$ *und alle* $v, w \in V$*;*

(ii) $\gamma(v, w_1 + w_2) = \gamma(v, w_1) + \gamma(v, w_2)$ *für alle* $v, w_1, w_2 \in V$ *und*
$\gamma(v, \lambda w) = \lambda\gamma(v, w)$ *für alle* $\lambda \in K$ *und alle* $v, w \in V$*.*

Eine Bilinearform γ *heißt* symmetrisch*, falls* $\gamma(v, w) = \gamma(w, v)$ *für alle* $v, w \in V$ *gilt.*

Beispiele 7.3. *(i) Seien* $V = \mathbb{R}^n$ *und* $A \in M(n \times n, K)$ *mit* $A = (a_{ij}) = (a_{ji}) = A^t$ *(d.h.* A *ist symmetrisch),* $x, y \in \mathbb{R}^n$*. Dann ist* $\gamma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *mit*

$$\gamma(x, y) = x^t Ay = \sum_{i,j=1}^{n} a_{ij}x_iy_j$$

eine symmetrische Bilinearform.

(ii) Seien $K = \mathbb{R}$ *und* $V = \mathcal{C}[0, 1] := \{f : [0, 1] \to \mathbb{R} \mid f \text{ stetig}\}$*. Dann ist die Abbildung* $\gamma : \mathcal{C}[0, 1] \times \mathcal{C}[0, 1] \to \mathbb{R}$ *mit*

$$\gamma(f, g) := \int_0^1 f(t)g(t)dt$$

eine symmetrische Bilinearform auf $\mathcal{C}[0, 1]$*.*

Definition 7.4. *Seien* $n = \dim(V) < \infty$*,* $B = (v_1, \ldots, v_n)$ *eine Basis von* V *und* γ *eine Bilinearform auf* V*. Die Matrix*

$$M_B^*(\gamma) := (\gamma(v_i, v_j)) \in M(n \times n, K)$$

heißt Darstellungsmatrix *von* γ *bezüglich* B*. [Wir verwenden " $*$ " zur Unterscheidung von* $M_B(\varphi)$ *für* $\varphi \in \mathrm{End}_K(V)$*.] Wir setzen* $\mathrm{Rang}(\gamma) := \mathrm{Rang}(M_B^*(\gamma))$ *(die Wohldefiniertheit von* $\mathrm{Rang}(\gamma)$ *ist unmittelbare Konsequenz des nachfolgenden Lemmas 7.6 (ii)).*

Für $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{j=1}^n \mu_j v_j$ folgt

$$\gamma(v, w) = \gamma\Big(\sum_{i=1}^{n} \lambda_i v_i, \sum_{j=1}^{n} \mu_j v_j\Big) = \sum_{i,j=1}^{n} \lambda_i\mu_j\gamma(v_i, v_j) = (\lambda_1, \ldots, \lambda_n)M_B^*(\gamma)\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Beispiel 7.5. *Sei* $V = \operatorname{Lin}\underbrace{(t^0, \ldots, t^{n-1})}_{=:B} \subset K[t]$. *Dann gilt für* γ *aus Beispiel 7.3 (ii)*

$$\gamma(t^i, t^j) = \int_0^1 t^{i+j} dt = \frac{1}{i+j+1} \quad \text{für } i, j = 0, \ldots, n-1,$$

d.h.

$$M_B^*(\gamma) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & & & & \frac{1}{n+1} \\ \vdots & & & & \vdots \\ \frac{1}{n} & & \cdots & & \frac{1}{2n-1} \end{pmatrix}.$$

Lemma 7.6. *Seien* $n = \dim V < \infty$, B, C *Basen von* V *mit Koordinatensystemen* $\Phi_B, \Phi_C : K^n \to V$. *Seien ferner* γ *eine Bilinearform auf* V *und* $A := M_B^*(\gamma)$. *Dann gilt:*

(i) Sind $B = (b_1, \ldots, b_n)$, $v = \sum_{i=1}^n x_i b_i$ *und* $w = \sum_{i=1}^n y_i b_i$, *so ist*

$$\gamma(v, w) = \Phi_B^{-1}(v)^t M_B^*(\gamma) \Phi_B^{-1}(w) = (x_1, \ldots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

(ii) $M_C^*(\gamma) = (T_B^C)^t M_B^*(\gamma) T_B^C$ *mit* $\widetilde{T_B^C} = \Phi_B^{-1} \circ \Phi_C$.

(iii) Ist umgekehrt $A \in M(n \times n, K)$, *so ist* $\gamma_A^B : V \times V \to K$ *mit*

$$\gamma_A^B(v, w) := \Phi_B^{-1}(v)^t A \Phi_B^{-1}(w)$$

eine Bilinearform. Für $V = K^n$ *und* $B = (e_1, \ldots, e_n)$ *gilt*

$$\gamma_A(v, w) := \gamma_A^B(v, w) = v^t A w.$$

Beweis. (i) Sei $B = (b_1, \ldots, b_n)$. Per definitionem ist $\Phi_B(e_i) = b_i$. Für $v, w \in V$ mit $v = \sum_{i=1}^n x_i b_i$ und $w = \sum_{i=1}^n y_i b_i$ gelten

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \Phi_B^{-1}(v), \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \Phi_B^{-1}(w)$$

und damit

$$\gamma(v, w) = \sum_{i,j=1}^n x_i y_j \gamma(b_i, b_j) = (x_1, \ldots, x_n) \underbrace{M_B^*(\gamma)}_{=A} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \Phi_B^{-1}(v)^t M_B^*(\gamma) \Phi_B^{-1}(w).$$

(ii) Nach (i) gilt für alle $v, w \in V$

$$\Phi_B^{-1}(v)^t M_B^*(\gamma)\Phi_B^{-1}(w) = \gamma(v,w) = \Phi_C^{-1}(v)^t M_C^*(\gamma)\Phi_C^{-1}(w).$$

Mit $\Phi_B^{-1}(w) = T_B^C \Phi_C^{-1}(w)$ (hier ohne $\sim$, denn $\Phi_C^{-1}(w) \in K^n$) erhält man für die linke Seite

$$\Phi_C^{-1}(v)^t (T_B^C)^t M_B^*(\gamma) T_B^C \Phi_C^{-1}(w).$$

Ist $C = (c_1, \ldots, c_n)$, so folgt daraus mit $v = c_i$ und $w = c_j$

$$e_i^t M_C^*(\gamma) e_j = e_i^t (T_B^C)^t M_B^*(\gamma) T_B^C e_j, \quad 1 \le i, j \le n.$$

Also stimmen die Einträge von $M_C^*(\gamma)$ allesamt mit denen von $(T_B^C)^t M_B^*(\gamma) T_B^C$ überein. [(ii) ist übrigens auch heuristisch sofort klar, da T_B^C die Koordinaten eines Vektors v bzgl. C in die bzgl. B überführt.]

(iii) Dass γ_A^B eine Bilinearform ist, zeigt man durch Nachrechnen. Für $V = K^n$ und $B = (e_1, \ldots, e_n)$ ist $\Phi_B(e_i) = e_i$, d.h. $\Phi_B = \mathrm{Id}_V$ und es folgt die Behauptung. □

Lemma 7.7. *(i)* $\mathrm{Bil}(V) := \{\gamma : V \times V \to K \mid \gamma$ *ist Bilinearform*$\}$ *ist ein* K*-Vektorraum (ein UVR des* K*-Vektorraums der Abbildungen von* $V \times V$ *nach* K*).*

(ii) Seien $n = \dim V < \infty$ *und* $B = (v_1, \ldots, v_n)$ *eine Basis von* V*. Dann ist die Abbildung*

$$M_B^* : \mathrm{Bil}(V) \to M(n \times n, K)$$

ein Isomorphismus von K*-Vektorräumen mit Umkehrabbildung*

$$\begin{aligned} \Gamma^B : M(n \times n) &\to \mathrm{Bil}(V) \\ A &\mapsto \gamma_A^B. \end{aligned}$$

Beweis. (i) Nachrechnen.

(ii) Dass M_B^* linear ist, kann man ebenfalls einfach nachrechnen. Für $\gamma \in \mathrm{Bil}(V)$ und alle $1 \le i, j \le n$ gilt weiter

$$\begin{aligned} \Big(\big(\Gamma^B \circ M_B^*\big)(\gamma)\Big)(v_i, v_j) &= \gamma^B_{M_B^*(\gamma)}(v_i, v_j) \\ &\overset{\text{Lemma 7.6}(iii)}{=} \Phi_B^{-1}(v_i)^t M_B^*(\gamma) \Phi_B^{-1}(v_j) \\ &= e_i^t M_B^*(\gamma) e_j = \gamma(v_i, v_j), \end{aligned}$$

also $\Gamma^B \circ M_B^* = \mathrm{Id}_{\mathrm{Bil}(V)}$. Schließlich ist auch $M_B^* \circ \Gamma^B = \mathrm{Id}_{M(n\times n, K)}$, denn es gilt für alle

$A = (A_{ij}) \in M(n \times n, K)$

$$\begin{aligned}\Big(\big(M_B^* \circ \Gamma^B\big)(A)\Big)_{ij} &= \Big(M_B^*\big(\gamma_A^B\big)\Big)_{ij} \\ &= \gamma_A^B(v_i, v_j) \\ &\overset{\text{Lemma 7.6}(iii)}{=} \Phi_B^{-1}(v_i)^t A \Phi_B^{-1}(v_j) \\ &= e_i^t A e_j = A_{ij}.\end{aligned}$$

□

Definition 7.8.

- *Ein quadratischer Raum ist ein Paar (V, γ) bestehend aus einem endlichdimensionalen K-VR V und einer symmetrischen Bilinearform γ auf V.*
- *$v, w \in V$ heißen orthogonal bzgl. γ (Notation $v \perp w$), falls $\gamma(v, w) = 0$. Eine Familie $(v_i)_{i \in I}$ von Vektoren heißt orthogonal bzgl. γ, falls $\gamma(v_i, v_j) = 0$ für alle $i, j \in I$ mit $i \neq j$.*
- *Eine Familie $(v_1, \ldots, v_n)$ heißt Orthogonalbasis von (V, γ), falls $(v_1, \ldots, v_n)$ eine Basis von V und orthogonal bzgl. γ ist.*

Bemerkung. *Für eine Basis B von V gilt:*

$$B \text{ ist Orthogonalbasis von } (V, \gamma) \iff M_B^*(\gamma) \text{ ist diagonal.}$$

Lemma 7.9. *Seien $n = \dim V < \infty$, (V, γ) ein quadratischer Raum und $v_1 \in V$ ein Vektor mit $\gamma(v_1, v_1) \neq 0$ sowie $H := \{w \in V \,|\, \gamma(v_1, w) = 0\}$. Dann ist H ein UVR von V mit $\dim H = n - 1$, und es gilt*

$$V = \operatorname{Lin}(v_1) \oplus H.$$

Beweis. Sei $\delta_{v_1} : V \to K$, $w \mapsto \delta_{v_1}(w) = \gamma(v_1, w)$. Dann gilt: δ_{v_1} ist linear mit $H = \operatorname{Kern} \delta_{v_1}$ und

$$\begin{aligned}\dim H &= \underbrace{\dim V}_{=n} - \underbrace{\dim \underbrace{\operatorname{Bild} \delta_{v_1}}_{\subset K}}_{\substack{\text{also } =0 \text{ oder } 1 \\ \text{d.h. } = 1 \text{ wegen } \gamma(v_1, v_1) \neq 0}} \\ &= n - 1.\end{aligned}$$

Es gilt für alle $v \in V$:

$$v = \underbrace{\frac{\gamma(v_1, v)}{\gamma(v_1, v_1)} v_1}_{\in \mathrm{Lin}(v_1)} + \underbrace{\left(v - \frac{\gamma(v_1, v)}{\gamma(v_1, v_1)} v_1\right)}_{\substack{\in H, \\ \text{da } \gamma\left(v_1, v - \frac{\gamma(v_1, v)}{\gamma(v_1, v_1)} v_1\right) = 0}} .$$

Es folgt $V = \mathrm{Lin}(v_1) + H$ und wegen $\dim V = \dim \mathrm{Lin}(v_1) + \dim H$ aus Satz 3.70 (iv) auch $V = \mathrm{Lin}(v_1) \oplus H$. □

Satz 7.10. *Seien $n = \dim V < \infty$ und (V, γ) ein quadratischer Raum. Dann besitzt (V, γ) eine Orthogonalbasis $(v_1, \ldots, v_n)$ (bzgl. γ), und es gilt*

$$V = \mathrm{Lin}(v_1) \oplus \cdots \oplus \mathrm{Lin}(v_n).$$

Beweis. Gilt $\gamma(v, w) = 0 \ \forall\, v, w \in V$ (man schreibt dafür auch $\gamma \equiv 0$), so ist nichts zu zeigen (jede Basis von V ist dann orthogonal). Sei also $\gamma \not\equiv 0$. Wir führen den Beweis mittels vollständiger Induktion nach n.

Induktionsanfang: Für $n = 1$ ist die Aussage klar.

Induktionsschritt: Sei $n \geq 2$ und die Aussage für $n - 1$ bereits bewiesen. Da $\gamma \not\equiv 0$ ist, gibt es ein $v_1 \in V$ mit $\gamma(v_1, v_1) \neq 0$, denn wäre $\gamma(v, v) = 0$ für alle $v \in V$, dann wäre auch

$$\gamma(v + w, v + w) - \gamma(v, v) - \gamma(w, w) = 2\gamma(v, w) = 0$$

für alle v, w, im Widerspruch zu Charakteristik$(K) \neq 2$ (siehe Voraussetzung zu Beginn des Kapitels!). Nach Lemma 7.9 gilt mit $H := \{w \in V \,|\, \gamma(v_1, w) = 0\}$

$$V = \mathrm{Lin}(v_1) \oplus H \ \text{ mit } \ \dim H = n - 1.$$

$(H, \gamma_{|H \times H})$ ist aber auch ein quadratischer Raum. Nach Induktionsvoraussetzung existiert eine Orthogonalbasis $(v_2, \ldots, v_n)$ von H mit $H = \mathrm{Lin}(v_2) \oplus \cdots \oplus \mathrm{Lin}(v_n)$. Nach Satz 3.70 (ii) ist $(v_1, \ldots, v_n)$ Orthogonalbasis von V mit

$$V = \mathrm{Lin}(v_1) \oplus H \overset{\text{Bem. } 3.74(ii)}{=} \mathrm{Lin}(v_1) \oplus \cdots \oplus \mathrm{Lin}(v_n).$$

□

Beispiel 7.11. *Sei $V = \mathbb{R}^n$ mit dem Standardskalarprodukt $\gamma(v, w) = \langle v, w \rangle = \sum_{i=1}^n v_i w_i$ als symmetrischer Bilinearform. Dann ist $(e_1, \ldots, e_n)$ eine Orthogonalbasis von $(\mathbb{R}^n, \langle ., . \rangle)$. In $\mathbb{R}^2$ definiert aber auch das bspw. um $45°$ gedrehte Koordinatensystem eine Orthogonalbasis, d.h. die Orthogonalbasis ist nicht eindeutig.*

Korollar 7.12. *Sei $A \in M(n \times n, K)$ symmetrisch. Dann gibt es $T \in GL(n, K)$, so dass $T^t A t$ eine Diagonalmatrix ist, d.h.*

$$T^t A T = \left(\begin{array}{c|c} \Lambda_r & 0 \\ \hline 0 & 0 \end{array}\right) \quad \text{mit } \Lambda_r = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}$$

mit $\lambda_i \neq 0$ und $r = \mathrm{Rang}(A)$. Die Spalten von T bilden eine Orthogonalbasis von K^n (d.h. sie sind orthogonal bzgl. γ_A).

Beweis. Seien $V = K^n$, $B = (e_1, \ldots, e_n)$ und $\gamma(v, w) := \gamma_A^B(v, w) \overset{\text{Lemma 7.6}(iii)}{=} v^t A w$, für $v, w \in K^n$. (V, γ) ist dann ein quadratischer Raum. Nach Satz 7.10 existiert eine Orthogonalbasis $C = (v_1, \ldots, v_n)$ und Lemma 7.6 (ii) ergibt

$$(T_B^C)^t \underbrace{M_B^*(\gamma)}_{=A} T_B^C = M_C^*(\gamma) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Wegen $B = (e_1, \ldots, e_n)$ ist $T := T_B^C = (v_1, \ldots, v_n)$ (Matrix mit den Spalten $v_1, \ldots, v_n$). Man kann die Reihenfolge der v_i's nun so ändern, dass obige Form entsteht. Wegen $T \in GL(n, K)$ gilt $\mathrm{Rang}(A) = r$. □

Bemerkung. *Die $\lambda_1, \ldots, \lambda_r$ sind hier nicht eindeutig bestimmt.*

Korollar 7.12 können wir im Falle $K = \mathbb{R}$ noch weiter vertiefen.

Satz 7.13 (Sylvesterscher Trägheitssatz)**.** *Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Dann gibt es $T \in GL(n, K)$ sowie $p, q \in \{0, \ldots, n\}$ mit*

$$T^t A T = \begin{pmatrix} E_p & & 0 \\ & -E_q & \\ 0 & & 0 \end{pmatrix}$$

Die Zahlen p, q sind durch A eindeutig bestimmt (d.h. unabhängig von T).

Beweis. Aus Korollar 7.12 folgt die Existenz von T mit

$$\hat{T}^t A \hat{T} = \left(\begin{array}{c|c} \Lambda_r & 0 \\ \hline 0 & 0 \end{array}\right) \quad \text{mit } \Lambda_r = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix},$$

wobei $\hat{T} = (\hat{v}_1, \ldots, \hat{v}_n)$ mit einer Orthogonalbasis $C = (\hat{v}_1, \ldots, \hat{v}_n)$ gilt. Sei weiter wie in

Korollar 7.12 $\gamma(v,w) = \gamma_A^B(v,w) = v^t A w$ mit $B = (e_1, \ldots, e_n)$. Wir setzen

$$\tilde{v}_i := \begin{cases} \frac{1}{\sqrt{|\gamma(\hat{v}_i, \hat{v}_j)|}} \hat{v}_i & \text{falls } \gamma(\hat{v}_i, \hat{v}_j) \neq 0 \\ \hat{v}_i & \text{falls } \gamma(\hat{v}_i, \hat{v}_j) = 0 \end{cases}$$

und erhalten damit

$$\gamma(\tilde{v}_i, \tilde{v}_j) = \begin{cases} 0 & \text{falls } i \neq j \\ 0 & \text{falls } i = j > r \\ \frac{\gamma(\hat{v}_i, \hat{v}_j)}{|\gamma(\hat{v}_i, \hat{v}_j)|} = \pm 1 & \text{falls } i = j \leq r. \end{cases}$$

Wir sortieren jetzt die $\tilde{v}_i$ aufsteigend nach $\gamma(\tilde{v}_i, \tilde{v}_i) = +1, -1, 0$ und erhalten die neue Basis $D = (v_1, \ldots, v_n)$. Mit $T = (v_1, \ldots, v_n)$ gilt dann $T = T_B^D$ und damit

$$T^t A T = \begin{pmatrix} E_p & & 0 \\ & -E_q & \\ 0 & & 0 \end{pmatrix}.$$

Es bleibt der Nachweis der Eindeutigkeit von p und q. Sei dazu $(v_1^*, \ldots, v_n^*)$ eine andere Orthogonalbasis mit

$$\gamma(v_i^*, v_i^*) = \begin{cases} 1 & \text{falls } 1 \leq i \leq p* \\ -1 & \text{falls } p^* + 1 \leq i \leq p^* + q^* \\ 0 & \text{falls } i > p^* + q^*. \end{cases}$$

Zu zeigen: $p = p^*$ ($\Rightarrow q = q^*$, da $p + q = p^* + q^* = \text{Rang}(A)$). Wir beweisen dafür unten, dass

$$v_1, \ldots, v_p, v_{p^*+1}^*, \ldots, v_n^* \tag{7.1}$$

und

$$v_1^*, \ldots, v_{p^*}^*, v_{p+1}, \ldots, v_n \tag{7.2}$$

linear unabhängig sind. Denn dies impliziert

$$p + (n - p^*) \leq n, \quad p^* + (n - p) \leq n \quad \Longrightarrow \quad p \leq p^*, \; p^* \leq p \quad \Longrightarrow \quad p = p^*.$$

Angenommen, die Familie der Vektoren in (7.1) sei nicht linear unabhängig, d.h. es existieren $\lambda_1, \ldots, \lambda_p, \mu_{p^*+1}, \ldots, \mu_n \in \mathbb{R}$ mit

$$v = \lambda_1 v_1 + \cdots + \lambda_p v_p = \mu_{p^*+1} v_{p^*+1}^* + \cdots + \mu_n v_n^*.$$

Wir berechnen (auf beiden Seiten) $\gamma(v,v)$:

$$\gamma(v,v) = \sum_{i=1}^{p} \lambda_i^2 \underbrace{\gamma(v_i,v_i)}_{=1} = \sum_{i=p^*+1}^{n} \mu_i^2 \underbrace{\gamma(v_i^*,v_i^*)}_{=-1 \text{ oder } 0} \leq 0.$$

$\Rightarrow \lambda_1 = \cdots = \lambda_p = 0 \Rightarrow \mu_{p^*+1} = \cdots = \mu_n = 0$, also (7.1). (7.2) zeigt man analog. □

Lemma und Definition 7.14. *Für eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ definieren wir Signatur von A als Signatur $(A) := (p,q)$ mit p,q aus Satz 7.13. Ist nun $S \in GL(n,\mathbb{R})$, dann haben A und S^tAS dieselbe Signatur.*

Beweis. S^tAS ist auch symmetrisch. Sei $\tilde{T} \in GL(n,K)$ mit

$$\begin{pmatrix} E_{\tilde{p}} & & 0 \\ & -E_{\tilde{q}} & \\ 0 & & 0 \end{pmatrix} = \tilde{T}^t S^t A S \tilde{T} = (S\tilde{T})^t A (S\tilde{T}).$$

Damit haben wir auch eine "Sylvester"-Darstellung für A gefunden. Diese ist nach Satz 7.13 aber eindeutig, womit $\tilde{p} = p$ und $\tilde{q} = q$, d.h. Signatur(S^tAS) = Signatur(A). □

7.2 Euklidische Räume und Orthogonalität

In diesem gesamten Abschnitt seien $K = \mathbb{R}$ und entsprechend V ein $\mathbb{R}$-Vektorraum.

Definition 7.15. *(i) Eine symmetrische Bilinearform $\gamma : V \times V \to \mathbb{R}$ heißt*

positiv definit, falls $\gamma(v,v) > 0$ für alle $v \in V \setminus \{0\}$

positiv semidefinit, falls $\gamma(v,v) \geq 0$ für alle $v \in V \setminus \{0\}$.

Eine positiv definite symmetrische Bilinearform nennt man auch Skalarprodukt.

(ii) Eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ heißt

positiv definit, falls $x^tAx > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ (also γ_A positiv definit)

positiv semidefinit, falls $x^tAx \geq 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ (also γ_A positiv semidefinit).

Beispiel und Definition 7.16. *(i) Das Standardskalarprodukt auf $\mathbb{R}^n$ ist*

$$\langle .,. \rangle_{E_n} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \ \langle x,y \rangle_{E_n} = x^t E_n y = \sum_{i=1}^{n} x_i y_i.$$

Wegen $\langle x,x \rangle_{E_n} = \sum_{i=1}^{n} x_i^2 > 0$ für $x \neq 0$ ist $\langle .,. \rangle_{E_n}$ positiv definit.

(ii) *Für* $A = \left(\begin{smallmatrix} 1 & -2 \\ -2 & 1 \end{smallmatrix}\right)$ *gilt*

$$e_1^t A e_1 = 1,\ e_2^t A e_2 = 1,\ \textit{aber}\ (1,1)A\begin{pmatrix}1\\1\end{pmatrix} = -2 < 0,$$

d.h. es reicht nicht, $\gamma(v,v) > 0$ *für Basisvektoren zu überprüfen.*

(iii) *Entsprechend gibt es die Begriffe* negativ definit ($\gamma(v,v) < 0\ \forall\, v \neq 0$), negativ semidefinit *und* indefinit *(Bsp. (ii)).*

(iv) *Sei* $V = \mathcal{C}[0,1]$. *Dann ist*

$$\begin{aligned} \gamma : V \times V &\to \mathbb{R} \\ (f,g) &\mapsto \int_0^1 f(t)g(t)dt \end{aligned}$$

ein Skalarprodukt (d.h. die symmetrische Bilinearform ist auch positiv definit!).

Definition 7.17. *Ein* euklidischer Raum *ist ein Paar* (V,γ) *bestehend aus einem* $\mathbb{R}$*-Vektorraum* V *und einem Skalarprodukt* $\gamma : V \times V \to \mathbb{R}$.

Im Rest dieses Abschnittes sei (V,γ) immer ein euklidischer Raum. Wir verwenden die Bezeichnungen $\langle v,w\rangle_\gamma := \gamma(v,w)$ oder $\langle v,w\rangle_A := v^t A w$. Wenn klar ist, welches Skalarprodukt gemeint ist, schreibt man auch einfach $\langle v,w\rangle$ anstelle von $\langle v,w\rangle_\gamma$ oder $\langle v,w\rangle_A$. Außerdem lässt man in der Angabe des euklidischen Raumes auch einfach das Skalarprodukt weg ("Sei V ein euklidischer Raum"), wenn klar ist, welches gemeint ist.

Definition 7.18. *Sei* (V,γ) *ein euklidischer Raum.*

- *Die Abbildung* $\|.\| : V \to \mathbb{R}$ *mit* $\|v\| := \sqrt{\langle v,v\rangle_\gamma}$ *heißt die* Norm *von* (V,γ).
- *Eine Familie von Vektoren* $(v_i)_{i\in I}$ *aus* V *heißt* orthonormal, *falls* $(v_i)_{i\in I}$ *orthogonal ist mit* $\|v_i\| = 1$ *für alle* $i \in I$.
- *Eine Familie* B *von Vektoren heißt* Orthonormalbasis *(kurz ONB) von* V, *falls* B *Basis sowie orthonormal ist.*

Lemma 7.19. *Seien* V *ein euklidischer Raum und* $(v_1,\dots,v_n)$ *eine orthogonale Famile von Vektoren aus* $V \setminus \{0\}$. *Dann gilt:*

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

(ii) $\left(\frac{v_1}{\|v_1\|},\dots,\frac{v_n}{\|v_n\|}\right)$ *ist eine orthonormale Familie. Damit besitzt jder UVR* $U \neq \{0\}$ *von* V *eine ONB.*

Beweis. (i) Gelte $\sum_{j=1}^n \lambda_j v_j = 0$ für $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Anwendung von $\langle ., v_i \rangle$ auf beiden Seiten ergibt

$$0 = \langle 0, v_i \rangle = \sum_{j=1}^{n} \lambda_j \langle v_j, v_i \rangle = \lambda_i \underbrace{\langle v_i, v_i \rangle}_{>0} \quad \Rightarrow \ \lambda_i = 0$$

für alle $i = 1, \ldots, n$.

(ii) Die Orthomormalität ist eine unmittelbare Konsequenz aus der Bilinearität von γ. Da nach Satz 7.10 zu jedem UVR eine Orthogonalbasis existiert, besitzt damit jeder UVR auch eine Orthonormalbasis. □

Lemma 7.20. *Seien (V, γ) ein euklidischer Raum und $B = (v_1, \ldots, v_n)$ eine Orthonormalbasis. Für $v = \sum_{j=1}^n \lambda_j v_j \in V$ folgt dann $\lambda_i = \langle v, v_i \rangle$.*

Beweis. Der Beweis folgt wie oben unter (i): Für alle $i = 1, \ldots n$ ist

$$v = \sum_{j=1}^{n} \lambda_j v_j \quad \Rightarrow \ \langle v, v_i \rangle = \sum_{j=1}^{n} \lambda_j \underbrace{\langle v_j, v_i \rangle}_{=\delta_{ij}} = \lambda_i$$

mit dem Kronecker-Symbol ("Kronecker-Delta") $\delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$ □

Satz 7.21 (Cauchy-Schwarz-Ungleichung)**.** *Sei $(V, \langle ., . \rangle)$ ein euklidischer Raum. Dann gilt für $v, w \in V$*

$$\big|\, \langle v, w \rangle \,\big| \ \leq \|v\| \cdot \|w\|.$$

Gleichheit gilt genau dann, wenn v und w linear abhängig sind (d.h. $v = \alpha w$ für $\lambda \in \mathbb{R}$).

Beweis. Für $w = 0$ ist die Aussage offenbar korrekt. Sei also $w \neq 0$. Für $\lambda, \mu \in \mathbb{R}$ gilt

$$0 \leq \langle \lambda v + \mu w, \lambda v + \mu w \rangle = \lambda^2 \langle v, v \rangle + 2\lambda\mu \langle v, w \rangle + \mu^2 \langle w, w \rangle.$$

Setze nun $\lambda := \langle w, w \rangle > 0$ und dividiere durch λ:

$$\Longrightarrow \quad 0 \leq \langle w, w \rangle \langle v, v \rangle + 2\mu \langle v, w \rangle + \mu^2.$$

Einsetzen von $\mu := -\langle v, w \rangle$:

$$\Longrightarrow \quad 0 \leq \langle w, w \rangle \langle v, v \rangle - 2\langle v, w \rangle \langle v, w \rangle + \langle v, w \rangle^2 = \langle w, w \rangle \langle v, v \rangle - \langle v, w \rangle \langle v, w \rangle,$$

also $\langle v, w \rangle^2 \leq \langle v, v \rangle \langle v, w \rangle$.

Bleibt zu zeigen, dass Gleichheit genau dann gilt, wenn V und w linear abhängig sind. Für $w = 0$: v, w sind linear abhängig und es gilt Gleichheit ($0 = 0$). Sei also $w \neq 0$.

"$\Leftarrow$": Seien v, w linear abhängig, d.h. $v = \alpha w$ für ein $\alpha \in \mathbb{R}$. Dann gilt

$$\big|\, \langle v, w\rangle \,\big| \;=\; \big|\, \alpha\langle w, w\rangle \,\big| \;= \|\alpha w\| \cdot \|w\|.$$

"$\Rightarrow$": Gelte "=". Führe die Berechnung der Ungleichung rückwärts durch (überall kann man $\Rightarrow$ durch $\Leftrightarrow$ ersetzen). Dann folgt für $\mu = -\langle v, w\rangle$ und $\lambda = \langle w, w\rangle > 0$ (da $w \neq 0$)

$$0 = \langle \lambda v + \mu w, \lambda v + \mu w\rangle = 0.$$

$\Rightarrow \lambda v + \mu w = 0 \overset{\lambda \neq 0}{\Rightarrow} v$ und w sind linear abhängig. □

Die Cauchy-Schwarz-Ungleichung hat aufgrund des allgemeinen Skalarproduktes viele Anwendungen.

Beispiele 7.22. *(i) Sei $A \in M(n \times n, \mathbb{R})$ positiv definit. Dann gilt für alle $x \in \mathbb{R}^n$*

$$\begin{aligned}(x^t x)^2 = \left(x^t A A^{-1} x\right)^2 &= \langle x, A^{-1}x\rangle_A^2 \\ &\leq \langle x, x\rangle_A \langle A^{-1}x, A^{-1}x\rangle_A \\ &= (x^t A x)(x^t A^{-1} A A^{-1} x) = (x^t A x)(x^t A^{-1} x).\end{aligned}$$

(ii) Für $f, g \in \mathcal{C}[0,1]$ gilt

$$\left(\int_0^1 f(t)g(t)dt\right)^2 \leq \int_0^1 f(t)^2 dt \;\cdot\; \int_0^1 g(t)^2 dt.$$

(iii) Sei $h \in \mathcal{C}[0,1]$, $h(.) > 0$. Dann definiert $\gamma(f,g)_h := \int_0^1 f(t)g(t)h(t)dt$ ein Skalarprodukt, womit nach der Cauchy-Schwarz-Ungleichung

$$\left(\int_0^1 f(t)g(t)h(t)dt\right)^2 \leq \int_0^1 f(t)^2 h(t) dt \;\cdot\; \int_0^1 g(t)^2 h(t) dt.$$

(Das folgt aber auch bereits aus (ii).)

Aus der Cauchy-Schwarz-Ungleichung folgt auch die Dreiecksungleichung. Wir formulieren das etwas umfassender.

Lemma 7.23 (Eigenschaften einer Norm)**.** *Sei $(V, \langle ., .\rangle)$ ein euklidischer Raum. Dann gilt für alle $v, w \in V$ sowie $\lambda \in \mathbb{R}$:*

(i) $\|v\| = 0 \Leftrightarrow v = 0$

(ii) $\|\lambda v\| = |\lambda| \cdot \|v\|$

(iii) $\|v + w\| \leq \|v\| + \|w\|$ *(Dreiecksungleichung).*

Beweis. (i) gilt wegen der positiven Definitheit und (ii) wegen Bilinearität des Skalarproduktes. Schließlich gilt (iii) wegen

$$\begin{aligned} \|v + w\|^2 = \langle v + w, v + w\rangle &= \|v\|^2 + \|w\|^2 + 2\langle v, w\rangle \\ &\overset{C.S.}{\leq} \|v\|^2 + \|w\|^2 + 2\|v\| \cdot \|w\| = \big(\|v\| + \|w\|\big)^2. \end{aligned} \tag{7.3}$$

□

Lemma 7.24. *Seien $(V, \langle .,. \rangle)$ ein euklidischer Raum und $v, w \in V$. Dann gilt:*

(i) $\|v + w\| = \|v\| + \|w\| \;\Leftrightarrow\; \langle v, w\rangle = 0$ *(Satz von Pythagoras)*

(ii) $\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$ *(Parallelogrammgleichung).*

Beweis. (i) folgt aus (7.3) oben. (ii) rechnet man unmittelbar nach:

$$\begin{aligned} \|v + w\|^2 + \|v - w\|^2 &= \|v\|^2 + \|w\|^2 + 2\langle v, w\rangle + \|v\|^2 + \|w\|^2 - 2\langle v, w\rangle \\ &= 2\|v\|^2 + 2\|w\|^2. \end{aligned}$$

□

Bemerkung 7.25. *Mit dem letzten Lemma haben wir zwei Aussagen der elementaren Geometrie auf euklidische Räume übertragen. Eine weitere Möglichkeit: Definiere für $v, w \in V \setminus \{0\}$ den Winkel α zwischen v und w durch*

$$\cos\alpha := \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \quad \textit{(liegt wegen der Cauchy-Schwarz-Ungl. in } [-1, 1]\textit{).}$$

$\langle .,. \rangle$ ist hier ein beliebiges Skalarprodukt. Dann gilt der Kosinussatz:

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\| \cdot \|w\| \cdot \cos\alpha.$$

7.3 Orthogonale Projektionen und Gram-Schmidt-Orthogonalisierung

Motivation 7.26. *Seien V ein euklidischer Raum und $v, v_1, v_2 \in V$. Wir möchten v auf $U = \mathrm{Lin}(v_1, v_2)$ "orthogonal projezieren". Dabei wird orthogonal definiert durch $v - p_u(v) \perp U$, wobei $p_U : V \to V$ die zugehörige Projektionsabbildung bezeichnet. Insbesondere hängt diese vom verwendeten Skalarprodukt ab. Offenbar gelten $p_U^2 = p_U$ sowie $p_{U|U} = \mathrm{Id}_U$.*

Lemma und Definition 7.27. *Sei V ein euklidischer Raum. Eine lineare Abbildung $p : V \to V$ heißt Projektion, falls $p^2 = p$ gilt (p heißt dann auch idempotent). Es gilt:*

(i) Die Eigenwerte von p sind 0 mit $\mathrm{Eig}(p,0) = \mathrm{Kern}(p)$ und 1 mit $\mathrm{Eig}(p,1) = \mathrm{Bild}(p)$.

(ii) $V = \mathrm{Kern}(p) \oplus \mathrm{Bild}(p)$

(iii) Ist p eine Projektion, so ist auch $\mathrm{Id}_V - p$ eine Projektion mit $\mathrm{Kern}(\mathrm{Id}_V - p) = \mathrm{Bild}(p)$ und $\mathrm{Bild}(\mathrm{Id}_V - p) = \mathrm{Kern}(p)$.

Beweis. (i) Seien λ Eigenwert von p und v ($\neq 0$) ein zugehöriger Eigenvektor. Dann gilt

$$\lambda v = p(v) = p^2(v) = p\big(p(v)\big) = p(\lambda v) = \lambda p(v) = \lambda^2 v$$

$\Rightarrow \lambda^2 = \lambda$, d.h. $\lambda = 0$ oder $\lambda = 1$. Weiter ist $\mathrm{Eig}(p,0) = \{v \,|\, p(v) = 0\} = \mathrm{Kern}(p)$ $\surd$.

Wir zeigen: $\{v \in V \,|\, p(v) = v\} = \mathrm{Eig}(p,1) \overset{!}{=} \mathrm{Bild}(p) = \{p(v) \,|\, v \in V\}$. ["!" über einem Gleichheitszeichen bedeutet, dass diese Gleichheit zu zeigen ist.]

"$\subset$": $\{v \in V \,|\, p(v) = v\} \subset \{p(v) \,|\, v \in V\}$ ist klar $\surd$

"$\supset$": Für $p(v)$ gilt $p\big(p(v)\big) = p^2(v) = p(v)$ $\surd$.

(ii) und (iii) wurden auf Übungsblatt 12 (Aufgabe 1) zur Linearen Algebra 1 bewiesen. Wir geben den Beweis hier der Vollständigkeit halber noch einmal an.

(ii) Für $v \in V$ gilt

$$v = \underbrace{v - p(v)}_{\substack{\in \mathrm{Kern}(p),\ \text{da} \\ p(v-p(v)) \\ = p(v) - p^2(v) = 0}} + \underbrace{p(v)}_{\in \mathrm{Bild}(p)},$$

d.h. $V = \mathrm{Kern}(p) + \mathrm{Bild}(p)$. Wegen $\dim V = \dim\big(\mathrm{Kern}(p)\big) + \dim\big(\mathrm{Bild}(p)\big)$ nach der Dimensionsformel folgt aus Satz 3.70 (iv) die Behauptung.

(iii) Es gilt $(\mathrm{Id}_V - p)^2 = \mathrm{Id}_V - 2p + p^2 = \mathrm{Id}_V - p$, d.h. $\mathrm{Id}_V - p$ ist Projektion. Nach (i) hat $\mathrm{Id}_V - p$ wieder die Eigenwerte 0 und 1, und es gilt

$$\mathrm{Kern}(\mathrm{Id}_V - p) = \mathrm{Eig}(\mathrm{Id}_V - p, 0) = \mathrm{Eig}(p,1) = \mathrm{Bild}(p),$$
$$\mathrm{Bild}(\mathrm{Id}_V - p) = \mathrm{Eig}(\mathrm{Id}_V - p, 1) = \mathrm{Eig}(p,0) = \mathrm{Kern}(p).$$

□

Definition 7.28. *Sei V ein euklidischer Raum.*

(i) Zwei Mengen $M, N \subset V$ heißen orthogonal (Notation $M \perp N$), falls $\langle v, w\rangle = 0$ $\forall\, v \in M, \forall\, w \in N$ gilt. [Bei einelementigen Mengen $\{v\}$ lässt man die Mengenklammer weg, schreibt also z. Bsp. $v \perp N$ oder $v \perp w$.] Für $M \subset V$ definiert

man

$$M^\perp := \{v \in V \,|\, v \perp w \;\forall\, w \in M\}.$$

Für einen UVR $W \subset V$ heißt $W^\perp$ das orthogonale Komplement von W in V.

(ii) Sei $U \subset V$ UVR. Dann heißt eine lineare Abbildung $p_U : V \to V$ orthogonale Projektion von V auf U, falls

- *p_U eine Projektion mit $\operatorname{Bild} p_U = U$ ist und*
- *$v - p_U(v) \perp U$ für alle $v \in V$ gilt, d.h. falls der Projektionsrest $v - p_U(v)$ immer orthogonal zu U ist.*

Die Existenz von p_U ist zunächst nicht klar. Der mathematisch elegante Beweis wäre, zunächst eine Orthonormalbasis für U zu konstruieren und dann p_U anzugeben. Für viele Anwendungen benötigt man aber die Darstellung bzgl. einer beliebigen Basis von U, die wir zunächst präsentieren wollen. Diese Konstruktion führt auch zu einem tieferen Verständnis von Projektionen. Wir zeigen dann unten die Existenz von p_U, indem wir p_U einfach angeben. Für die Eindeutigkeit benötigen wir das folgende Lemma.

Lemma 7.29. *Das Skalarprodukt $\langle .,. \rangle$ auf einem euklidischen Raum ist nicht ausgeartet, d.h. es gilt*

$$\langle v, w\rangle = 0 \;\forall\, w \in V \quad \Longleftrightarrow \quad v = 0.$$

Beweis. "$\Rightarrow$": $\langle v, w\rangle = 0 \;\forall\, w \in V \Rightarrow \langle v, v\rangle = 0 \overset{\text{pos. def.}}{\overset{\langle .,. \rangle}{\Longrightarrow}} v = 0$

"$\Leftarrow$": Es gilt für $v = 0$:

$$\begin{aligned}\langle v, w\rangle &= \frac{1}{2}\Big(\langle v + w, v + w\rangle - \langle v, v\rangle - \langle w, w\rangle\Big)\\ &= \frac{1}{2}\Big(\langle w, w\rangle - \langle w, w\rangle\Big) = 0.\end{aligned}$$

□

Satz 7.30. *Seien V ein euklidischer Raum und U UVR von V. Dann existiert eine eindeutig bestimmte Orthogonalprojektion p_u, die folgendermaßen gegeben ist.*

(i) Im Falle $V = \mathbb{R}^n$: Sei $E = (e_1, \ldots, e_n)$ die Einheitsbasis und $A = M_E^*(\langle .,. \rangle)$ die Darstellungsmatrix von $\langle .,. \rangle$ bzgl. E. Sei $(x_1, \ldots, x_r)$ eine beliebige Basis von U. Dann gilt mit $X = (x_1, \ldots, x_r) \in M(n \times r, \mathbb{R})$

$$P_U = M_E(p_u) = X(X^t A X)^{-1} X^t A.$$

(ii) Für allgemeines V: Ist $B = (v_1, \dots, v_r)$ eine Basis von U, so gilt für alle $v \in V$

$$p_U(v) = (v_1, \dots, v_r)\big(M_B^*(\langle ., . \rangle)\big)^{-1} \begin{pmatrix} \langle v_1, v \rangle \\ \vdots \\ \langle v_r, v \rangle \end{pmatrix},$$

wobei $(v_1, \dots, v_r)x := \sum_{k=1}^r x_k v_k$ *für* $x \in \mathbb{R}^r$.

Beweis. (i) Da $(x_1, \dots, x_r)$ eine Basis ist, ist $\operatorname{Rang}(X) = r$. Weiter folgt $\operatorname{Rang}(A) = n$ aus der positiven Definitheit von $\langle ., . \rangle$ und damit nach Aufgabe 4 (a) auf Übungsblatt 8 $\operatorname{Rang}(X^t A X) = r$. Es folgt

$$P_U^2 = X \underbrace{(X^t A X)^{-1} X^t A X}_{=E_r} (X^t A X)^{-1} X^t A = P_U,$$

d.h. die Matrix P_U ist idempotent. Wegen $x_j = Xe_j$ gilt

$$P_U x_j = X(X^t A X)^{-1} X^t A X e_j = X e_j = x_j, \tag{7.4}$$

womit $U \subset \operatorname{Bild} \tilde{P}_U$. Da andererseits jedes $y \in U$ Linearkombination der $x_1, \dots, x_r$ ist, folgt dies nach Linearität wegen (7.4) auch für $P_U y$, womit $\operatorname{Bild} \tilde{P}_U \subset U$ und damit

$$U = \operatorname{Bild} \tilde{P}_U.$$

Wir zeigen nun $y - P_U y \perp U$ für alle $y \in \mathbb{R}^n$ ist, also $\langle y - P_U y, x_j \rangle = 0$ für $j = 1, \dots, r$. Es gilt

$$\begin{aligned} \langle y - P_U y, x_j \rangle &= \big((E_n - P_U)y\big)^t A X e_j \\ &= y^t \Big(E_n - AX(X^t A X)^{-1} X^t\Big) A X e_j \\ &= y^t A X e_j - y^t A X e_j = 0, \end{aligned}$$

d.h. p_U mit $P_U = M_E(p_U)$ ist eine Orthogonalprojektion auf U. Es bleibt der Nachweis der Eindeutigkeit. Ist $q_U : V \to V$ eine weitere Orthogonalprojektion auf U, so folgt für jedes $v \in V$ aus $\langle v - p_U(v), u \rangle = 0$ und $\langle v - q_U(v), u \rangle = 0$, dass

$$\langle \underbrace{p_U(v) - q_U(v)}_{\in U}, u \rangle = 0.$$

Da auch $(U, \langle ., . \rangle_{|U})$ ein euklidischer Raum ist, impliziert Lemma 7.29 $p_U(v) - q_U(v) = 0$ für alle $V \in V \Rightarrow p_U = q_U$.

(ii) folgt analog → Übungsblatt 8, Aufgabe 4 (b). □

Bemerkung. *Ist im Falle von Satz 7.30 (i) $(x_1, \ldots, x_r)$ nicht linear unabhängig und $U = \mathrm{Lin}(x_1, \ldots, x_r)$, so gilt mit $X = (x_1, \ldots, x_r) \in M(n \times r, \mathbb{R})$ und $y \in \mathbb{R}^n$*

$$P_U y = M_E(p_U) y = X\beta,$$

wobei β eine Lösng der "Normalengleichungen" $(X^t A X)\beta = X^t A y$ ist. β ist dann nicht eindeutig, aber $X\beta$ ist eindeutig bestimmt. In den meisten Anwendungen ist eine solche Modellierung wegen der Redundanz der x_j aber nicht sinnvoll.

Beispiel 7.31. *Ein Bekleidungsunternehmen mit 10 Filialen in Kleinstädten möchte die Einflussfaktoren für den Jahresgewinn ermitteln. Gesucht wird eine Funktion der Form*

$$\underset{y}{\textit{Gewinn}} = \beta_1 \cdot \underset{x_1}{\textit{Verkaufsfläche}} + \beta_2 \cdot \underset{x_2}{\textit{Anzahl Produkte}} + \beta_3 \cdot \underset{x_3}{\textit{Werbeetat}} + \beta_4,$$

d.h. gegeben sind vier Vektoren y, x_1, x_2, x_3 der Länge 10. Um β_4 bestimmen zu können, setzen wir $x_4 = (1, \ldots, 1)^t$. Mit $U = \mathrm{Lin}(x_1, \ldots, x_4)$ und $A = E_{10}$ berechnen wir die Orthogonalprojektion $P_U y$ nach Satz 7.30 (i). Falls $(x_1, \ldots, x_4)$ linear unabhängig sind, gilt

$$X \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_4 \end{pmatrix} = X(X^t X)^{-1} X^t y \quad \textit{mit } X = (x_1, \ldots, x_4),$$

also $(\beta_1, \ldots, \beta_4)^t = (X^t X)^{-1} X^t y$. Für die Beurteilung der Ergebnisse (bspw. "Sind bestimmte Koeffizienten β_i "signifikant" von Null verschieden?") braucht man zusätzlich ein statistisches Modell bspw. der Form $y = X\beta + \varepsilon$ mit einem "zufälligen Fehler" ε.

Lemma 7.32. *Seien V ein euklidischer Raum, U UVR von V und $v \in V$. Dann ist die orthogonale Projektion $p_U(v)$ die Bestapproximation bzgl. $\|\cdot\|$ von v durch ein Element aus U, d.h.*

$$\inf_{u \in U} \|v - u\| = \|v - p_U(v)\| = \left(\|v\|^2 - \|p_U(v)\|^2\right)^{1/2}.$$

Anders ausgedrückt: Der Abstand von v zu einem $u \in U$ ist an der Stelle u am kleinsten, wo $v - u \perp U$.

Beweis. Es gilt

$$\begin{aligned} \|v - u\|^2 &= \| \underbrace{v - p_U(v)}_{\perp U} + \underbrace{p_U(v) - u}_{\in U} \|^2 \\ &= \|v - p_U(v)\|^2 + \|p_U(v) - u\|^2 \quad \text{(Pythagoras)}, \end{aligned}$$

d.h. $\|v-u\|$ wird minimal für $u = p_U(v)$. Die zweite behauptete Gleichung folgt erneut mit Pythagoras: Wegen $v - p_U(v) \perp p_U(v)$ gilt $\|v\|^2 = \|v - p_U(v)\|^2 + \|p_U(v)\|^2$. □

Bemerkung 7.33. *Ein wichtiger Spezialfall der Orthogonalprojektion ist gegeben, wenn die Basis $(v_1, \ldots, v_r)$ in Satz 7.30 eine Orthonormalbasis ist. Dann erhält man (mit der dortigen Notation)*

$$p_U = (v_1, \ldots, v_r) \begin{pmatrix} \langle v_1, v\rangle \\ \ddots \\ \langle v_r, v\rangle \end{pmatrix}, \quad \textit{d.h. } p_U(v) = \sum_{j=1}^{r} \langle v_j, v\rangle v_j.$$

Man kann diese Beziehung aber auch einfach direkt einsehen: Sei $p_U(v) = \sum_{j=1}^r \lambda_j v_j$. Dann muss ja gelten $\langle v - p_U(v), v_i\rangle = 0$, d.h.

$$\langle v, v_i\rangle - \sum_{j=1}^{r} \lambda_j \underbrace{\langle v_j, v_i\rangle}_{\delta_{ij}} = 0 \quad \Rightarrow \quad \lambda_i = \langle v, v_i\rangle.$$

Wir zeigen nun, wie man eine beliebige Basis von V in eine Orthonormalbasis bzgl. eines allgemeinen Skalarproduktes $\langle .,. \rangle$ überführen kann.

Gram-Schmidt-Orthogonalisierung

Ausgangssituation $B = (v_1, \ldots, v_n)$ Basis von V. Sei

$$U_k = \mathrm{Lin}(v_1, \ldots, v_k) = \bigoplus_{j=1}^{k} \mathrm{Lin}(v_j), \quad \text{insbesondere } U_n = V.$$

Algorithmus

(i) Setze $\tilde{w}_1 := v_1$ und $\tilde{w}_1 := \frac{\tilde{w}_1}{\|\tilde{w}_1\|}$.

(ii) Projeziere (orthogonal) für $k = 2, \ldots, n$ den Vektor v_k auf U_{k-1} und verwende den normierten Projektionsrest als neuen Basisvektor w_k. In Formeln:

$$\tilde{w}_k := v_k - \sum_{j=1}^{k-1} \langle v_k, w_j\rangle w_j \quad \text{and} \quad w_k := \frac{\tilde{w}_k}{\|\tilde{w}_k\|}.$$

Satz 7.34 (Gram-Schmidt Orthogonalisierung)**.** *(i) $(w_1, \ldots, w_k)$ ist für alle k eine Orthonormalbasis von $(U_k, \langle .,. \rangle_{|U_k})$ und es gilt*

$$\bigoplus_{j=1}^{k} \mathrm{Lin}(w_j) = U_k = \bigoplus_{j=1}^{k} \mathrm{Lin}(v_j).$$

(ii) Die Transformationsmatrix $T_k := T^{(v_1,\dots,v_k)}_{(w_1,\dots,w_k)}$ *ist für alle* $k \le n$ *eine obere Dreiecksmatrix mit positiven Diagonalelementen* $\|\tilde{w}_j\|$, $j = 1,\dots,k$, *und*

$$\det(T_k) = \prod_{j=1}^{k} \|w_j\| > 0.$$

Analog ist $T_k^{-1} = T^{(w_1,\dots,w_k)}_{(v_1,\dots,v_k)}$ *obere Dreiecksmatrix mit Diagonalelementen* $\|\tilde{w}_j\|^{-1}$, $j = 1,\dots,k$.

Beweis. Wir zeigen die Aussage mittels Induktion nach k.

Induktionsanfang: $k = 1$ folgt direkt aus der Definition von w_1.

Induktionsschritt: Sei die Aussage für $k-1$ bereits bewiesen. Da B Basis ist, gilt

$$v_k \notin U_{k-1} = \bigoplus_{j=1}^{k-1} \mathrm{Lin}(v_j) \overset{\text{Ind.-vorauss.}}{=} \bigoplus_{j=1}^{k-1} \mathrm{Lin}(w_j),$$

d.h.

$$\tilde{w}_k := v_k - p_{U_{k-1}}(v_k) = v_k - \sum_{j=1}^{k-1} \langle v_k, w_j\rangle w_j.$$

Weiter ist $\tilde{w}_k \neq 0$ und damit auch $w_k = \frac{\tilde{w}_k}{\|\tilde{w}_k\|} \neq 0$. Es gilt $w_k \perp U_{k-1}$ und $(w_1,\dots,w_k)$ ist damit eine Orthonormalbasis. Nach Lemma 7.9 folgt

$$U_k = \mathrm{Lin}(w_k) \oplus U_{k-1} \overset{\text{Bem. 3.74}}{=} \bigoplus_{j=1}^{k} \mathrm{Lin}(w_j).$$

Insbesondere ist $(w_1,\dots,w_k)$ eine Orthomnormalbasis von $(U_k, \gamma_{|U_k})$. Nach Konstruktion von w_k gilt

$$w_k = \frac{1}{\|\tilde{w}_k\|} v_k + y \quad \text{mit einem } y \in U_{k-1},$$

d.h. $v_k = \alpha w_k - y$ mit $\alpha = \|\tilde{w}_k\|$ und $y \in U_{k-1}$. Aber damit ist

$$T_k = T^{(v_1,\dots,v_k)}_{(w_1,\dots,w_k)} = \begin{pmatrix} T^{(v_1,\dots,v_{k-1})}_{(w_1,\dots,w_{k-1})} & * \\ 0 & \alpha \end{pmatrix} \quad \text{mit } \det(T_k) = \det(T_{k-1})\alpha \overset{\text{Ind.-vorauss.}}{>} 0.$$

Analog zeigt man die behauptete Aussage auch für T_k^{-1}. [Die Gestalt als obere Dreiecksmatrix folgt auch allgemein, weil die Inverse einer oberen Dreiecksmatrix wieder eine obere Dreiecksmatrix ist.] □

Beispiel 7.35. *Im* $(\mathbb{R}^3, \langle .,. \rangle_{E_3})$ *sei* $U = \mathrm{Lin}(v_1, v_2)$ *mit*

$$v_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \quad \textit{und} \quad v_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Gesucht ist eine Orthonormalbasis von U*. Wir verwenden das Gram-Schmidt-Verfahren.*

$$w_1 := \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

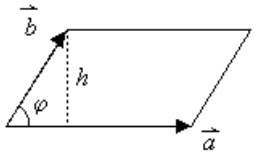
Weiter ist

$$\tilde{w}_2 := v_2 - \langle v_2, w_1 \rangle w_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \underbrace{\left\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\rangle}_{\frac{-2}{\sqrt{5}}} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{5}(-2) \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix},$$

womit schließlich $w_2 = \frac{\tilde{w}_2}{\|\tilde{w}_2\|} = \frac{1}{\sqrt{30}} \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}$.

Beispiel 7.36. *Als Anwendung des Gram-Schmidt-Orthogonalisierungsverfahrens bestimmen wir das Volumen eines Parallelotops (oder auch Parallelepipeds) im* $\mathbb{R}^n$*, welches von den Vektoren* $x_1, \ldots, x_n$ *aufgespannt wird. Motivation im* $\mathbb{R}^2$ *(mit* $x_1 = \vec{a}, x_2 = \vec{b}$*):*



Die Fläche beträgt

$$\|x_1\| \cdot h = \|x_1\| \cdot \| \underbrace{x_2 - p_{\mathrm{Lin}(x_1)}(x_2)}_{\perp x_1} . \|$$

Analog kommt im $\mathbb{R}^3$ *die Komponente* $\|x_3 - p_{\mathrm{Lin}(x_1,x_2)}(x_3)\|$ *hinzu.*

Sei allgemein $U_k = \mathrm{Lin}(x_1, \ldots, x_k)$ *und das Volumen definiert durch*

$$\mathrm{Vol}(x_1, \ldots, x_n) := \|x_1\| \prod_{j=2}^{n} \|x_j - p_{\mathrm{Lin}(x_1,\ldots,x_{j-1})}(x_j)\|.$$

Definition 7.37. *Für einen euklidischen Vektorraum V und Vektoren $v_1, \dots, v_r \in V$ definieren wir die Gramsche Determinante durch*

$$G(v_1, \dots, v_r) = \det\Big(\big(\langle v_i, v_j\rangle\big)_{i,j=1,\dots,n}\Big).$$

Gelten $V = \mathbb{R}^n$ und $\langle v_i, v_j\rangle = v_i^t A v_j$, so gilt

$$G(v_1, \dots, v_n) = \det\Big((v_1, \dots, v_r)^t A (v_1, \dots, v_r)\Big).$$

In der Literatur wird die Gramsche Determinante meist nur für das Standardskalarprodukt (d.h. $A = E_n$) definiert.

Satz 7.38. *Seien $V = \mathbb{R}^n$ mit Skalarprodukt $\langle .,. \rangle_A$ und $x_1, \dots, x_n \in V$. Dann gilt mit T_n aus Satz 7.34*

$$\mathrm{Vol}(x_1, \dots, x_n) = \sqrt{G(x_1, \dots, x_n)} = \det(T_n).$$

Beweis. Sei die Famile $(x_1, \dots, x_n)$ zunächst linear unabhängig, also eine Basis von V. Die Gram-Schmidt-Orthogonalisierung ergibt dann eine Orthonormalbasis $(w_1, \dots, w_n)$, die der Matrizenidentität $\underbrace{(w_1, \dots, w_n)}_{T^{(w_1,\dots,w_n)}_{(e_1,\dots,e_n)}} T_n = \underbrace{(x_1, \dots, x_n)}_{T^{(x_1,\dots,x_n)}_{(e_1,\dots,e_n)}}$ genügt, womit

$$\begin{aligned} G(x_1, \dots, x_n) &= \det\Big((x_1, \dots, x_n)^t A (x_1, \dots, x_n)\Big) \\ &= \det\Big(T_n^t \underbrace{(w_1, \dots, w_n)^t A (w_1, \dots, w_n)}_{=E_n} T_n\Big) \\ &= \big(\det T_n\big)^2 \\ &= \Big(\prod_{j=1}^{n} \|\tilde{w}_j\|\Big)^2 \end{aligned}$$

mit $\tilde{w}_j$ aus Satz 7.34. Nach Konstruktion sind $\tilde{w}_1 = x_1$ sowie $\tilde{w}_j = x_j - p_{\mathrm{Lin}(x_1,\dots,x_{j-1})}(x_j)$, also $G(x_1, \dots, x_n) = \big(\mathrm{Vol}(x_1, \dots, x_n)\big)^2$. Ist die Familie $(x_1, \dots, x_n)$ linear abhängig, sind beide Seiten gleich Null. □

7.4 Die orthogonale Gruppe

Definition 7.39. *Seien $(V, \langle .,. \rangle_V)$, $(W, \langle .,. \rangle_W)$ endlichdimensionale euklidische Räume und $\varphi : V \to W$ linear. φ heißt orthogonal oder Isometrie, falls*

$$\langle \varphi(v_1), \varphi(v_2)\rangle_W = \langle v_1, v_2\rangle_V \quad \text{für alle } v_1, v_2 \in V.$$

Bemerkung 7.40. *Eine Isometrie wird meistens als normerhaltende lineare Abbildung definiert, d.h. als* $\varphi \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ *mit*

$$\|\varphi(v)\|_W = \|v\|_V \quad \textit{für alle } v \in V.$$

Wegen

$$\langle v, w\rangle = \frac{1}{2}\Big(\langle v+w, v+w\rangle - \langle v, v\rangle - \langle w, w\rangle\Big)$$

ist das aber identisch zu Definition 7.39.

Lemma 7.41. *Seien* $(V, \langle .,. \rangle_V)$, $(W, \langle .,. \rangle_W)$ *endlichdimensionale euklidische Räume und* $\varphi : V \to W$ *eine Isometrie. Dann gilt:*

(i) $\|\varphi(v)\|_W = \|v\|_V$ *für alle* $v \in V$.

(ii) $v_1 \perp v_2 \Leftrightarrow \varphi(v_1) \perp \varphi(v_2)$.

(iii) φ *ist injektiv, d.h.* φ *ist ein Isomorphismus (da* $\dim V = \dim W$*).*

(iv) φ^{-1} *ist ebenfalls eine Isometrie.*

(v) Sind $V = W$ *und* λ *Eigenwert von* φ*, so ist* $|\lambda| = 1$*, d.h.* $\lambda \in \{-1, 1\}$.

Beweis. Aus Definition 7.39 folgen (i) mit $v = w$ und (ii) mit $v \perp w$.

(iii) $v \in V$ mit $\varphi(v) = 0 \Rightarrow \|\varphi(v)\|_W = 0 \overset{(i)}{\Rightarrow} \|v\|_V = 0 \Rightarrow v = 0$.

(iv) $w_1, w_2 \in W \Rightarrow \langle \varphi^{-1}(w_1), \varphi^{-1}(w_2)\rangle_V \overset{\varphi \text{ Isom.}}{=} \langle \varphi(\varphi^{-1}(w_1)), \varphi(\varphi^{-1}(w_2))\rangle_W = \langle w_1, w_2\rangle_W$

(v) Ist $v \in V$ Eigenvektor zum Eigenwert λ, so folgt

$$\|v\|_V = \|\varphi(v)\|_V = \|\lambda v\|_V = |\lambda| \cdot \|v\|_V \overset{v \neq 0}{\Rightarrow} |\lambda| = 1.$$

□

Lemma 7.42. *Seien* $(V, \langle .,. \rangle_V)$ *euklidischer Raum,* $n = \dim V < \infty$ *und* B *Orthonormalbasis von* $(V, \langle .,. \rangle_V)$*. Dann ist das Koordinatensystem* $\Phi_B : (\mathbb{R}^n, \langle .,. \rangle_{E_n}) \to (V, \langle .,. \rangle_V)$ *eine Isometrie.*

Beweis. Es gilt $\langle \Phi_B(e_i), \Phi_B(e_j)\rangle_V = \langle v_i, v_j\rangle_V = \delta_{ij} = \langle e_i, e_j\rangle_{E_n}$. □

Lemma 7.43. *Seien* $(U, \langle .,. \rangle_U)$, $(V, \langle .,. \rangle_V)$ *und* $(W, \langle .,. \rangle_W)$ *euklidische Vektorräume mit* $\dim(U) = \dim(V) = \dim(W)$*. Seien* $\varphi : U \to V$ *und* $\psi : V \to W$ *Isometrien. Dann ist auch* $\psi \circ \varphi$ *eine Isometrie.*

Beweis. Übungsaufgabe. □

Lemma 7.44. *Seien $n = \dim(V)$, B eine Orthonormalbasis von V, $\varphi \in \mathrm{End}_K(V)$ und $Q = M_B(\varphi)$. Dann sind äquivalent:*

(i) φ ist eine Isometrie.

(ii) $Q^tQ = E_n$, d.h. Q ist orthogonal (im Sinne der nachfolgenden Definition).

Insbesondere gilt dann:

$$Q^{-1} = Q^t \text{ und } M_B(\varphi^{-1}) = Q^{-1} = Q^t$$

Beweis. Aus Lemmata 7.42 und 7.43 folgt:

$$\varphi \text{ Isometrie} \Leftrightarrow \widetilde{Q} = \Phi_B^{-1} \circ \varphi \circ \Phi_B \text{ Isometrie} \Leftrightarrow \forall x, y \in \mathbb{R}^n : \underbrace{\langle Qx, Qy \rangle_E}_{=x^tQ^tQy} = \langle x, y \rangle \Leftrightarrow Q^tQ = E_n$$

Dass daraus $Q^{-1} = Q^t$ folgt, ist klar. Da $M_B : \mathrm{End}_K(V) \to M(n \times n, K)$ nach Bemerkung 4.46 ein Ringisomorphismus ist, folgt auch $M_B(\varphi^{-1}) = Q^{-1}$. □

Lemma und Definition 7.45. *Eine Matrix $Q \in M(n \times n, \mathbb{R})$ heißt orthogonal, falls*

$$Q^tQ = E_n.$$

[Sinngemäß müsste man solche Q orthonormal nennen (da Q^tQ ja kein beliebiges Vielfaches von E_n ist, sondern genau E_n), aber wir nennen es trotzdem orthogonal.] Wir setzen

$$O(n) := \{Q \in M(n \times n, \mathbb{R}) \mid Q \text{ ist orthogonal}\}.$$

$O(n)$ ist bzgl. der Matrixmultiplikation eine Gruppe, die orthogonale Gruppe vom Rang n genannt wird. Es gilt:

$$\det(Q) = \pm 1 \ \forall Q \in O(n)$$

Beweis. Abgeschlossenheit von $O(n)$ bzgl. Matrixmultiplikation:

$$Q, R \in O(n) \Rightarrow (QR)^t(QR = R^t \underbrace{Q^tQ}_{=E_n} R) = R^tR = E_n \Rightarrow QR \in O(n)$$

Neutrales Element: $E_n^t \cdot E_n = E_n \cdot E_n = E_n \Rightarrow E_n \in O(n)$.

Inverse Elemente:

$$Q \in O(n) \Rightarrow \left(Q^{-1}\right)^t Q^{-1} = QQ^{-1} = E_n \Rightarrow Q^{-1} \in O(n)$$

Assoziativität: klar.

Wegen $Q^tQ = E_n$ folgt $= \det(E_n) = \det(Q^tQ) = (\det(Q))^2$. □

Lemma 7.46. *Sei $Q \in M(n \times n, \mathbb{R})$. Dann sind äquivalent:*

(i) $Q \in O(n)$

(ii) $Q^t Q = E_n$

(iii) Die Spalten von Q bilden eine Orthonormalbasis von $(\mathbb{R}^n, \langle ., . \rangle_{E_n})$

(iv) $QQ^t = E_n$

(v) Die Zeilen von Q bilden eine Orthonormalbasisvon $(\mathbb{R}^n, \langle ., . \rangle_{E_n})$.

Lemma und Definition 7.47. *Die Menge*

$$SO(n) := \{Q \in O(n) \mid \det(Q) = 1\}$$

ist eine Untergruppe von $O(n)$. Sie heißt spezielle orthogonale Gruppe.

Beweis. Übungsaufgabe. □

Satz 7.48 (Hauptachsentransformation)**.** *Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Dann existiert eine orthogonale Matrix Q, so dass*

$$Q^t A Q = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

für $\lambda_1, ..., \lambda_n \in \mathbb{R}$. Man kann Q sogar so wählen, dass $\det(Q) = 1$ ist.

Wir brauchen für den Beweis folgendes

Hilfslemma: Sei S symmetrisch und positiv definit. Gilt $u^t S u = 0$ für ein $u \in \mathbb{R}^n$, so ist $Su = 0$.

Beweis des Hilfslemmas. Für alle $x \in \mathbb{R}^n, \lambda \in \mathbb{R}$ gilt

$$0 \leq (\lambda u + x)^t S(\lambda u + x) = 2\lambda x^t S u + \underbrace{x^t S x}_{\geq 0}.$$

Da man λ beliebig groß/klein auswählen kann, folgt $x^t S u = 0 \; \forall x \in \mathbb{R}^n \Rightarrow Su = 0$. □

Beweis von Satz 7.48. wir führen den Beweis mittels vollständiger Induktion nach n.

Induktionsanfang: $n = 1$ ist klar.

Induktionsschritt: Sei $n \geq 2$ und die Aussage für $n - 1$ bereits bewiesen. Die Einheitssphäre $S^{n-1} := \{x \in \mathbb{R}^n \mid ||x|| = 1\}$ ist eine kompakte Menge im $\mathbb{R}^n$ (da abgeschlossen

und beschränkt). Die stetige, reellwertige Funktion $x \mapsto x^t Ax$ nimmt damit ihr Maximum $\mu \in S^1$ an, z.B. an der Stelle $u \in S^{n-1}$ (u ist nicht notwendigerweise eindeutig). Formal:

$$u := \underset{x \in S^{n-1}}{\operatorname{argmax}}\, x^t Ax,\ \mu := u^t Au.$$

Dann gilt $\mu = u^t Au \geq x^t Ax\ \forall x \in S^1$, d.h. $\forall x \in \mathbb{R}^n$ gilt:

$$x^t Ax = ||x||^2 \left(\frac{x}{||x||}\right)^t A \left(\frac{x}{||x||}\right) \leq \mu ||x||^2 = \mu(x^t x)$$

$$\Rightarrow x^t(\mu E_n - A)x \geq 0\ \forall x \in \mathbb{R}^n$$

Folglich ist $\mu E_n - A$ symmetrisch und positiv semidefinit mit

$$u^t(\mu E_n - A)u = u^t \mu E_n u - \underbrace{u^t Au}_{\mu} = 0$$

Nach dem Hilfslemma ist damit $(\mu E_n - A)u = 0 \Leftrightarrow Au = \mu u$, d.h. μ ist Eigenwert von A mit Eigenvektor u. Da $||u|| = 1$ ist, kann man u zu einer Orthonormalbasis des $\mathbb{R}^n$ ergänzen.

$$\overset{\text{Lemmata 7.42+7.44}}{\Longrightarrow} \exists K \in M(n \times n, \mathbb{R}), K \text{ orthogonal}, Ke_1 = u.$$

$$\Rightarrow AKe_1 = Au = \mu u = \mu Ke_1 = K(\mu e_1)\ \Rightarrow\ (K^{-1}AK)e_1 = \mu e_1$$

$$\Rightarrow K^t AK = K^{-1}AK = \begin{pmatrix} \mu & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} \text{ mit } B^t = B$$

Nach Induktionsvoraussetzung existiert $L \in M((n-1) \times (n-1), \mathbb{R}) \cap O(n-1)$, so dass

$$L^t BL = \begin{pmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Nun definieren wir die $n \times n$-Matrix Q durch

$$Q := K \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L & \\ 0 & & & \end{pmatrix}.$$

Wegen $L^tL = E_n$ gilt

$$Q^tQ = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L^t & \\ 0 & & & \end{pmatrix} \underbrace{K^tK}_{=E_n} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L^tL & \\ 0 & & & \end{pmatrix} = E_n.$$

Mit $\lambda_1 := \mu$ gilt weiter

$$Q^tAQ = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L^t & \\ 0 & & & \end{pmatrix} \underbrace{K^tAK}_{=\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L & \\ 0 & & & \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L^tBL & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Will man eine Matrix Q mit $\det(Q) = 1$ erhalten, so ersetzt man einfach Q durch:

$$Q \cdot \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$$

was die Rechnung nicht ändert. □

Bemerkung 7.49. *Man kann als Anwendung der Hauptachsentransformation für symmetrische, positiv semidefinite Matrizen A eine Wurzel $A^{1/2}$ definieren. Nach Satz 7.48 gilt*

$$A = Q \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} Q^t \quad \text{mit } \lambda_1, \ldots, \lambda_n \geq 0.$$

Wir definieren nun

$$A^{1/2} := Q \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} Q^t.$$

Offenbar gilt dann $A^{1/2} \cdot A^{1/2} = A$, *denn*

$$Q \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \underbrace{Q^t Q}_{E_n} \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} Q^t = Q \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} Q^t = A.$$

7.5 Unitäre Räume

Im letzten Abschnitt dieses Kapitels sei nun V ein $\mathbb{C}$-Vektorraum. Für $x = x_1 + ix_2 \in \mathbb{C}$ setzen wir $\bar{x} := x_1 - ix_2$ (komplex Konjugiertes von x).

Definition 7.50. $h : V \times V \to \mathbb{C}$ *heißt* Sesquilinearform *auf* V, *falls für alle* $\lambda \in \mathbb{C}$ *und alle* $v, v_1, v_2, w, w_1, w_2 \in V$ *gilt:*

(i) $h(v_1 + v_2, w) = h(v_1, w) + h(v_2, w)$, $h(\lambda v_1, v_2) = \lambda h(v_1, v_2)$ *und*

(ii) $h(v, w_1 + w_2) = h(v, w_1) + h(v, w_2)$, $h(v, \lambda w) = \bar{\lambda} h(v, w)$.

Bezeichnung: $SeLF := \{h : V \times V \to \mathbb{C} |\, h \text{ ist Sesquilinearform}\}$.

Beispiel 7.51. *Auf* $V = \mathbb{C}^n$ *definiert* $h(x, y) = x^t \bar{y} = \sum_{i=1}^n x_i \bar{y}_i$ *mit* $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_n)^t$ *eine Sesquilinearform aber keine Bilinearform (denn* $h(x, \lambda y) = \bar{\lambda} h(x - y)$*!).*

Definition 7.52. *Seien* $h \in SeLF(V)$, $B = (v_1, \ldots, v_n)$ *Basis von* V. *Dann heißt*

$$M_B^*(h) = \big(h(v_i, v_j)\big) \in M(n \times n, \mathbb{C})$$

Darstellungsmatrix *von* h *bezüglich* B.

Definition 7.53. *Eine Sesquilinearform* $h \in SeLF(V)$ *heißt* hermitesch, *falls* $h(v, w) = \overline{h(w, v)}$ *für alle* $v, w \in V$ *gilt. Insbesondere ist* $h(v, v) = \overline{h(v, v)}$, *d.h.* $h(v, v) \in \mathbb{R}\ \forall\, v \in V$.

Darstellende Matrizen $A = (a_{ij})$ hermitescher Sesquilinearformen genügen der Identität $\bar{A}^t = A$, wobei $\bar{A} = (\bar{a}_{ij})$. Entsprechend bezeichnet man $A \in M(n \times n, \mathbb{C})$ als hermitesch, falls $\bar{A}^t = A$ gilt.

Definition 7.54. *Sei* h *eine hermitesche Sesquilinearform auf* V.

(i) h *heißt* positiv definit, *falls* $h(v, v) > 0\ \forall\, v \in V \setminus \{0\}$ *gilt. Eine positiv definite hermitesche Sesquilinearform heißt* Skalarprodukt.

(ii) Ein unitärer Raum *ist ein Paar* (V, h) *bestehend aus einem* C*-Vektorraum* V *und einem Skalarprodukt* h *auf* V.

Beispiel 7.55. *Auf $V = \mathbb{C}^n$ ist $\langle .,. \rangle_{\mathbb{C}}$, gegeben durch $\langle x, y \rangle_{\mathbb{C}} := x^t \bar{y}$ für $x, y \in \mathbb{C}^n$, ein Skalarprodukt.*

- *$\langle .,. \rangle_{\mathbb{C}}$ ist hermitesch: $\langle x, y \rangle_{\mathbb{C}} = x^t \bar{y} = \bar{x}^t y = \overline{x^t \bar{y}} = \overline{\langle x, y \rangle_{\mathbb{C}}}$.*
- *$\langle .,. \rangle_{\mathbb{C}}$ ist positiv definit: $\langle x, x \rangle_{\mathbb{C}} := x^t \bar{x} = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n |x_i|^2 > 0$ für $x \neq 0$. [Für $z = a + ib \in \mathbb{C}$ mit $a, b \in \mathbb{R}$ setzt man $|z| := \sqrt{a^2 + b^2}$.]*

Die mathematische Behandlung von unitären Räumen ist weitgehend analog zu der von euklidischen, indem man systematisch folgende Begriffe ersetzt:

$\mathbb{R}$	$\mathbb{C}$
euklidischer Raum	unitärer Raum
Bilinearform	Sesquilinearform
symmetrische Bilinearform	hermitesche Sesquilinearform
transponierte Matrix A^t	adjungierte Matrix $A^* := \bar{A}^t$
orthogonal	unitär.

Wir verzichten daher an dieser Stelle auf die Ausführung und verweisen auf die Literatur.

8 Duale und adjungierte Abbildungen

8.1 Dualräume

Einführung 8.1. *Im $\mathbb{R}^n$ kann man z.B. das Standardskalarprodukt*

$$\langle x, y \rangle = (x_1, ..., x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i$$

für festes $x = (x_1, ..., x_n)^t$ als lineare Abbildung vom $\mathbb{R}^n$ nach $\mathbb{R}$ betrachten. Die Spaltenvektoren sind dann Elemente des $\mathbb{R}^n$, die Zeilenvektoren Elemente von

$$(\mathbb{R}^n)^* := \{f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ ist linear}\}.$$

Es gibt eine "kanonische Abbildung" $\psi : \mathbb{R}^n \to (\mathbb{R}^n)^*$, $x \mapsto x^t$.

Definition 8.2. *Seien K ein Körper und V ein K-Vektorraum. Die Menge*

$$V^* := \mathrm{Hom}_K(V, K) = \{\varphi : V \to K \mid \varphi \text{ ist } K\text{-linear}\}$$

heißt Dualraum von V. Die Elemente von V^* heißen Linearformen auf V.

Beispiele 8.3. *(i) Seien $K = \mathbb{R}$ und $V = \mathbb{R}^n$. Die Funktionen*

$$\varphi : \mathbb{R}^n \to \mathbb{R}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{j=1}^{n} x_j$$

und

$$\psi : \mathbb{R}^n \to \mathbb{R}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1$$

sind Linearformen auf $\mathbb{R}^n$.

(ii) Seien $K = \mathbb{R}$ und $V = \mathcal{C}([0,1]) = \{f : [0,1] \to \mathbb{R} | \, f \text{ stetig}\}$. Die Funktion

$$\varphi : \mathcal{C}([0,1]) \to \mathbb{R}, f \mapsto \int_0^1 f(t) \, dt$$

eine Linearform auf $\mathcal{C}([0,1])$.

8.2 Duale Basen und duale Abbildungen

Lemma und Definition 8.4. *Seien V K-Vektorraum und $B = (v_1, ..., v_n)$ eine Basis ($\dim V = n \in \mathbb{N}$). Wir definieren für $i \in \{1, ..., n\}$ die linearen Abbildungen*

$$v_i^* : V \to K, \; v_j \mapsto \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases}$$

d.h. es gilt

$$v = \sum_{j=1}^{n} \lambda_j v_j \mapsto \sum_{j=1}^{n} \lambda_j \delta_{ij} = \lambda_i.$$

Dann ist $B^* := (v_1^*, ..., v_n^*)$ eine Basis von V^*, die sogenannte duale Basis zu B.

Beweis. Lineare Unabhängigkeit: Seien $\lambda_1, ..., \lambda_n \in K$ mit $\sum_{j=1}^{n} \lambda_j v_j^* = 0$. Dann folgt

$$0 = \sum_{j=1}^{n} \lambda_j v_j^*(v_i) = \lambda_i \; \forall \, i \in \{1, ..., n\}.$$

Erzeugendensystem: Sei $\varphi \in V^*$. Setze $\lambda_i := \varphi(v_i)$ für $i \in \{1, ..., n\}$. Damit gilt

$$\varphi(v_i) = \lambda_i = \left(\sum_{j=1}^{n} \lambda_j v_j^* \right)(v_i) \;\Rightarrow\; \varphi(v) = \left(\sum_{j=1}^{n} \lambda_j v_j^* \right)(v) \; \forall \, v \in V.$$

□

Satz 8.5. *Sei $n = \dim(V) < \infty$, $B = (v_1, ..., v_n)$ eine Basis von V, und $B^* = (v_1^*, ..., v_n^*)$ die duale Basis zu B. Dann existiert ein Isomorphismus*

$$\Psi_B : V \to V^* \text{ mit } v_i \mapsto v_i^* \ \forall i \in \{1, ..., n\}.$$

Insbesondere gilt $\dim(V) = \dim(V^*)$ und Ψ_B ist injektiv, d.h. $\forall \varphi \in V^*$ existieren $\lambda_1, ..., \lambda_n \in K$, so dass $\varphi = \sum_{j=1}^n \lambda_j v_j^*$.

Beweis. Definiere $\Psi_B : V \to V^*$ durch

$$v = \sum_{j=1}^n \lambda_j v_j \mapsto \sum_{j=1}^n \lambda_j v_j^*.$$

Ψ_B ist offensichtlich linear. Da B^* nach Lemma 8.4 eine Basis ist, existieren zu allen $\varphi \in V^*$ eindeutige Skalare $\lambda_1, ..., \lambda_n \in K$, so dass $\varphi = \sum_{j=1}^n \lambda_j v_j^*$ ist. Damit ist Ψ_B injektiv. Aus Lemma 8.4 folgt auch $\dim(V) = \dim(V^*)$, da die Basen B und B^* gleich lang sind. Nach Korollar 4.16 ist damit Ψ_B bijektiv, also ein Isomorphismus. □

Bemerkung 8.6. *Man beachte, dass v_i^* nicht nur von v_i, sondern auch von den anderen Basisvektoren $v_j, j \neq i$, abhängt. In diesem Sinne ist $B \mapsto B^*$ zwar eine wohldefinierte Zuordnung, aber man müsste eigentlich $v_i^*(B)$ anstelle von v_i^* schreiben. Insbesondere gibt es keine kanonische Abbildung mit $v \mapsto v^*$ (wie im $\mathbb{R}^n$). Im euklidischen Raum könnte man ein solches $^* : V \to V^*$ durch $v \mapsto \langle v, \cdot \rangle$ definieren, benötigt dafür aber eben ein Skalarprodukt. Analog hängt auch der Isomorphismus Ψ_B in Satz 8.5 von B ab.*

Beispiele 8.7. *(i) Seien $V = K^n$, $B = (e_1, ..., e_n)$. Für die duale Basis $B^* = (e_1^*, ..., e_n^*)$ gilt*

$$M_{e_1}^{(e_1,...,e_n)}(e_i^*) = (0, ..., 0, \overbrace{1}^{\substack{i\text{-te} \\ \text{Stelle}}}, 0, ..., 0) = e_i^t.$$

(ii) Seien $K = \mathbb{R}$, $V = \mathbb{R}^2$ und $C = (v_1, v_2)$ mit $v_1 = \binom{1}{0} = e_1$, $v_2 = \binom{1}{1}$. Es gelten

$$\begin{aligned} M_{(e_1)}^{(e_1,e_2)}(v_1^*) &= M_{(e_1)}^{(v_1,v_2)}(v_1^*) \cdot T_{(v_1,v_2)}^{(e_1,e_2)} \\ &= M_{(e_1)}^{(v_1,v_2)}(v_1^*) \cdot \left(T_{(e_1,e_2)}^{(v_1,v_2)}\right)^{-1} \\ &= (1,0) \cdot \left(T_{(e_1,e_2)}^{(v_1,v_2)}\right)^{-1} \quad (\text{da } v_1^*(v_1) = 1 \text{ und } v_1^*(v_2) = 0 \text{ ist}) \\ &= (1,0) \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = (1,0) \cdot \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = (1,-1) \end{aligned}$$

sowie

$$\begin{aligned} M_{(e_1)}^{(e_1,e_2)}(v_2^*) &= M_{(e_1)}^{(v_1,v_2)}(v_2^*) \cdot T_{(v_1,v_2)}^{(e_1,e_2)} \\ &= M_{(e_1)}^{(v_1,v_2)}(v_2^*) \cdot \left(T_{(e_1,e_2)}^{(v_1,v_2)}\right)^{-1} \\ &= (0,1) \cdot \left(T_{(e_1,e_2)}^{(v_1,v_2)}\right)^{-1} \text{ (da } v_2^*(v_1) = 0 \text{ und } v_2^*(v_2) = 1 \text{ ist)} \\ &= (0,1) \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = (0,1) \cdot \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = (0,1), \end{aligned}$$

d.h. für $v_1 = \binom{1}{0}$ *sind die* v_1^* *in* (i) *und* (ii) *unterschiedlich.*

Korollar 8.8. *Seien* V *ein* K*-Vektorraum,* $n := \dim(V) < \infty$ *und* $v \in V \setminus \{0\}$*. Dann existiert* $\varphi \in V^*$ *mit* $\varphi(v) \neq 0$*.*

Beweis. Wir ergänzen zunächst (v) zu einer Basis $(v, v_2, ..., v_n)$ von V. Wir betrachten nun die dazu duale Basis $(v^*, v_2^*, ..., v_n^*)$. Dann gilt $v^*(v) = 1 \neq 0$. □

Lemma und Definition 8.9. *Ist* $U \subseteq V$ *ein Untervektorraum von* V*, so heißt*

$$U^0 := \{\varphi \in V^* \mid \varphi(u) = 0 \ \forall u \in U\} \subseteq V^*$$

der Annulator von U*.* U^0 *ist ein Untervektorraum von* V^**.*

Beweis. Nachrechnen. □

Satz 8.10. *Seien* $U \subset V$ *ein Untervektorraum und* $(u_1, ..., u_k)$ *eine Basis von* U*. Sind* $B = (u_1, ..., u_k, v_1, ..., v_r)$ *eine Basis von* V *und* $B^* = (u_1^*, ..., u_k^*, v_1^*, ..., v_r^*)$ *die dazu duale Basis, so ist* $(v_1^*, ..., v_r^*)$ *eine Basis von* U^0*. Insbesondere gilt*

$$\dim(U^0) = \dim(V) - \dim(U).$$

Beweis. $(v_1^*, ..., v_r^*)$ ist linear unabhängig, da diese Familie eine Teilfamilie der Basis B^* ist. Wir müssen also noch zeigen, dass $\mathrm{Lin}(v_1^*, ..., v_r^*) = U^0$ ist.

"$\subset$": Sei $\varphi \in \mathrm{Lin}(v_1^*, ..., v_r^*)$, d.h. $\exists \lambda_1, ..., \lambda_r \in K$ mit $\varphi = \sum_{j=1}^r \lambda_j v_j^*$. Für $i \in \{1, ..., k\}$ gilt

$$\varphi(u_i) = \sum_{j=1}^{r} \lambda_j \underbrace{v_j^*(u_i)}_{=0} = 0 \ \Rightarrow \ \varphi(u) = 0 \ \forall u \in U \ \Rightarrow \ \varphi \in U^0.$$

"$\supset$": Sei $\varphi \in U^0$. Da B^* eine Basis von $V^* \supset U^0$ ist, gibt es $\lambda_1, ..., \lambda_r, \mu_1, ..., \mu_k \in K$ mit

$$\varphi = \sum_{j=1}^{k} \mu_j u_j^* + \sum_{j=1}^{r} \lambda_j v_j^*.$$

Für $i \in \{1, ..., k\}$ ist $0 = \varphi(u_i) = \mu_i \cdot \underbrace{u_i^*(u_i)}_{=1} = \mu_i$ (da $u_j^*(u_i) = 0$ für $i \neq j$ und $v_j^*(u_i) = 0$).
Dies gilt für alle $\mu_1, ..., \mu_k$, womit

$$\varphi = \sum_{j=1}^{r} \lambda_j v_j^* \;\Rightarrow\; \varphi \in \mathrm{Lin}(v_1^*, ..., v_r^*).$$

□

Lemma und Definition 8.11. *Seien V und W K-Vektorräume und $f : V \to W$ eine lineare Abbildung. Wir definieren die zu f duale Abbildung f^* durch*

$$f^* : W^* \to V^*,\; \psi \mapsto \psi \circ f.$$

f^* ist ebenfalls linear.

Beweis. Für $\psi, \psi_1, \psi_2 \in W^*$ und $\lambda \in K$ folgen

$$f^*(\psi_1 + \psi_2) = (\psi_1 + \psi_2) \circ f = \psi_1 \circ f + \psi_2 \circ f = f^*(\psi_1) + f^*(\psi_2)$$

sowie

$$f^*(\lambda\psi) = (\lambda\psi) \circ f = \lambda f^*(\psi).$$

□

Lemma 8.12. *Seien V und W endlichdimensionale K-Vektorräume. Dann ist die Abbildung*

$$^* : \mathrm{Hom}_K(V, W) \to \mathrm{Hom}_K(W^*, V^*)$$

ein Isomorphismus von K-Vektorräumen.

Beweis. Die Linearität kann man einfach nachrechnen. Bleibt die Isomorphie zu zeigen.

Injektivität: Sei $f \in \mathrm{Hom}_K(V, W)$ mit $f^* = 0$. $\Rightarrow\; \psi \circ f = 0\ \forall\, \psi \in W^*$. Angenommen, $f \neq 0$, d.h. $\exists\, v \in V$ mit $f(v) \neq 0$.

$$\overset{\text{Kor. } 8.8}{\Rightarrow} \exists \psi \in W^* : \psi(f(v)) \neq 0 \Rightarrow \psi \circ f \neq 0. \;\lightning$$

Damit ist $f = 0$, also ist * injektiv (da * einen trivialen Kern hat).

Nach Korollar 4.16 ist * damit aber bereits bijektiv, denn

$$\dim_K(\mathrm{Hom}_K(V,W)) \overset{\text{Satz 4.36 } (iii)}{=} \dim_K(V)\cdot\dim_K(W)$$
$$\overset{\text{Satz 8.5}}{=} \dim_K(V^*)\cdot\dim_K(W^*) = \dim(\mathrm{Hom}_K(W^*,V^*)). \qquad \square$$

Der nächste Satz zeigt, dass die duale Abbildung f^* bezüglich der dualen Basen durch die Transponierte der Darstellungsmatrix von f beschrieben wird.

Satz 8.13. *Seien V und W endlichdimensionale K-Vektorräume, B und C Basen von V bzw. W und $f: V \to W$ eine lineare Abbildung. Dann gilt*

$$M_{B^*}^{C^*}(f^*) = M_C^B(f)^t.$$

Beweis. Seien $B = (v_1, ..., v_n)$, $C = (w_1, ..., w_m)$ und $A = (a_{kj}) := M_C^B(f)$. Dann gilt

$$f(v_j) = \sum_{k=1}^{m} a_{kj} w_k.$$

Anwendung von w_i^* ergibt $w_i^*(f(v_j)) = w_i^*\left(\sum_{k=1}^{m} a_{kj} w_k\right) = \sum_{k=1}^{m} a_{kj}\delta_{ik} = a_{ij}$, also

$$a_{ij} = w_i^*(f(v_j)) = (w_i^* \circ f)(v_j) = f^*(w_i^*)(v_j). \tag{8.1}$$

Andererseits gilt mit $D = (d_{kj}) := M_{B^*}^{C^*}(f^*)$:

$$f^*(w_i^*) = \sum_{k=1}^{n} d_{ki} v_k^*.$$

Durch Einsetzen von v_j ergibt sich

$$d_{ji} = \sum_{k=1}^{n} d_{ki} v_k^*(v_j) = f^*(w_i^*)(v_j) \overset{(8.1)}{=} a_{ij},$$

also $M_{B^*}^{C^*}(f^*) = M_C^B(f)^t$. $\square$

Satz 8.14. *Seien V und W endlichdimensionale K-Vektorräume und $f: V \to W$ eine lineare Abbildung. Dann gilt*

$$(i)\ \ \mathrm{Bild}(f^*) = \mathrm{Kern}(f)^0 \quad \textit{und} \quad (ii)\ \ \mathrm{Kern}(f^*) = \mathrm{Bild}(f)^0.$$

Beweis. (ii) Es gilt

$$\varphi \in \mathrm{Kern}(f^*) \Leftrightarrow f^*(\varphi) = 0_{V^*} \Leftrightarrow \varphi \circ f = 0_{V^*} \Leftrightarrow \varphi_{|\,\mathrm{Bild}(f)} = 0_{\mathrm{Bild}(f)^*} \Leftrightarrow \varphi \in \mathrm{Bild}(f)^0.$$

(i) Sei $\varphi \in \mathrm{Bild}(f^*)$, d.h. $\exists\, \psi \in W^* : \varphi = f^*(\psi) = \psi \circ f$. Dann ist

$$\varphi|_{\mathrm{Kern}(f)} = 0 \;\Rightarrow\; \varphi \in \mathrm{Kern}(f)^0,$$

womit $\mathrm{Bild}(f^*) \subset \mathrm{Kern}(f)^0$. Wegen

$$\begin{aligned}
\dim(\mathrm{Bild}(f^*)) &\overset{\text{Dim.-Formel}}{=} \dim(W^*) - \dim(\mathrm{Kern}(f^*)) \\
&\overset{(ii)}{=} \dim(W^*) - \dim(\mathrm{Bild}(f)^0) \\
&\overset{\text{Satz 8.10}}{=} \dim(W^*) - \big(\dim(W) - \dim(\mathrm{Bild}(f))\big) \\
&\overset{\text{Satz 8.5}}{=} \dim(\mathrm{Bild}(f)) \\
&\overset{\text{Dim.-Formel}}{=} \dim(V) - \dim(\mathrm{Kern}(f)) \\
&\overset{\text{Satz 8.10}}{=} \dim(\mathrm{Kern}(f)^0),
\end{aligned}$$

folgt (i) damit aus Korollar 4.14. □

Korollar 8.15. *Seien V und W endlichdimensionale K-Vektorräume und $f : V \to W$ eine lineare Abbildung. Dann gilt*

$$\mathrm{Rang}(f^*) = \mathrm{Rang}(f).$$

Beweis.

$$\begin{aligned}
\mathrm{Rang}(f^*) = \dim(\mathrm{Bild}(f^*)) &= \dim(\mathrm{Kern}(f)^0) \\
&\overset{\text{Satz 8.10}}{=} \dim(V) - \dim(\mathrm{Kern}(f)) = \dim(\mathrm{Bild}(f)) = \mathrm{Rang}(f).
\end{aligned}$$

□

Aus Korollar 8.15 folgt ein alternativer Beweis von Satz 4.30.

Korollar 8.16 (Satz 4.30)**.** *Sei $A \in M(m \times n, K)$. Dann gilt*

$$\textit{Zeilenrang}(A) = \textit{Spaltenrang}(A).$$

Beweis. Nach Satz 8.13 gelten

$$A = M^{(e_1,\ldots,e_n)}_{(e_1,\ldots,e_m)}(\widetilde{A}) \quad \text{und} \quad A^t = M^{(e_1^*,\ldots,e_m^*)}_{(e_1,\ldots,e_n)}(\widetilde{A}^*)$$

und damit

$$\begin{aligned}\text{Spaltenrang}(A) = \dim(\text{Bild}(\widetilde{A})) &= \text{Rang}(\widetilde{A}) \\ &\overset{\text{Kor. 8.15}}{=} \text{Rang}(\widetilde{A}^*) = \text{Spaltenrang}(A^t) = \text{Zeilenrang}(A).\end{aligned}$$

□

Definition 8.17. *Der Vektorraum*

$$V^{**} := (V^*)^* = \text{Hom}_K(V^*, K)$$

heißt Bidualraum von V.

Der nächste Satz zeigt, dass im Gegensatz zum Isomorphismus $\Psi_B : V \to V^*$ aus Satz 8.5 für einen endlichdimensionalen Vektorraum V ein Basis-unabhängiger Isomorphismus von V nach V^{**} existiert.

Satz 8.18. *Sei V ein K-Vektorraum mit $\dim(V) < \infty$. Dann existiert ein kanonischer (d.h. von einer Basis in V unabhängiger) Isomorphismus*

$$i : V \to V^{**},\ v \mapsto i_v \quad \text{mit} \quad i_v : V^* \to K,\ \varphi \mapsto \varphi(v).$$

Beweis. (i) Die Abbildung ist wohldefiniert und linear (Beweis durch Nachrechnen).

(ii) *Injektivität von i:* Sei $v \in \text{Kern}(i)$, d.h. $i_v = 0$. Daraus folgt:

$$\forall \varphi \in V^* : i_v(\varphi) = \varphi(v) = 0 \overset{\text{Kor. 8.8}}{\Longrightarrow} v = 0.$$

(iii) *Surjektivität von i:* Es gilt $\dim(V^{**}) = \dim(V^*) = \dim(V)$ nach Satz 8.5. Damit ist i nach Korollar 4.16 auch surjektiv und folglich ein Isomorphismus. □

8.3 Adjungierte Abbildungen und Spektralsatz

Lemma und Definition 8.19. *Eine Bilinearform $\gamma : V \times V \to K$ auf einem K-Vektorraum V induziert folgende lineare Abbildungen:*

$$\Gamma_l : V \to V^*, w \mapsto \gamma(\cdot, w)$$
$$\Gamma_r : V \to V^*, v \mapsto \gamma(v, \cdot)$$

(l steht für links, r für rechts), wobei

$$\gamma(\cdot, w) : V \to K,\ \gamma(\cdot, w)(v) = \gamma(v, w)$$
$$\gamma(v, \cdot) : V \to K,\ \gamma(v, \cdot)(w) = \gamma(v, w).$$

Die Bilinearform γ heißt nicht ausgeartet, wenn Γ_l und Γ_r injektiv sind. Ist γ symmetrisch, so gilt $\Gamma_l = \Gamma_r$.

Beweis. Klar nach Definition einer (symmetrischen) Bilinearform. □

Lemma 8.20. *Sei V endlichdimensionaler K-Vektorraum und γ eine symmetrische Bilinearform auf V. Ist γ nicht ausgeartet, so sind Γ_l und Γ_r Isomorphismen.*

Beweis. Wegen $\dim(V) = \dim(V^*)$ nach Satz 8.5 folgt aus der Injektivität von Γ_l und Γ_r nach Korollar 4.16 auch die Surjektivität. □

Bemerkung und Definition 8.21. *Im Gegensatz zum Isomorphismus $\Psi_B : V \to V^*$ aus Satz 8.5 (der von der Basis B abhängt) ist der Isomorphismus Γ_r (und auch Γ_l) kanonisch, wenn ein Skalarprodukt vorgegeben ist. Im Spezialfall $V = \mathbb{R}^n$ mit dem Standardskalarprodukt gilt mit der Einheitsbasis $B = (e_1, ..., e_n)$, dass $\Psi_B = \Gamma_r$. Insbesondere sind für einen endlichdimensionalen euklidischen Raum V die Vektorräume V und V^* kanonisch isomorph. Deshalb bezeichnet man auch $\Psi = \Gamma_r$ als kanonischen Isomorphismus, d.h.*

$$\Psi : V \to V^*, \ v \mapsto \langle v, \cdot \rangle.$$

Satz 8.22. *Seien V ein endlichdimensionaler euklidischer Raum und Ψ der kanonische Isomorphismus. Dann gilt:*

(i) Für jeden Untervektorraum $U \subset V$ gilt $\Psi(U^\perp) = U^0$.

(ii) Sind $B = (v_1, ..., v_n)$ eine Orthonormalbasis von V und $B^* = (v_1^*, ..., v_n^*)$ die dazu duale Basis, so ist $\Psi = \Psi_B$, d.h. $\Psi(v_i) = v_i^* \ \forall\, i \in \{1, ..., n\}$.

Beweis. (i) Es gilt $\Psi(U^\perp) \subset U^0$, denn $\forall v \in U^\perp$ ist $\Psi(v)(u) = \langle v, u \rangle = 0 \ \forall u \in U$, womit $\Psi(v) \in U^0$. Wegen

$$\dim(\Psi(U^\perp)) = \dim(U^\perp) = \dim(V) - \dim(U) = \dim(U^0)$$

(die erste Gleichheit gilt, da Ψ ein Isomorphismus ist) folgt, dass $\Psi(U^\perp) = U^0$ sein muss (da beide Mengen Untervektorräume von V^* sind und $\Psi(U^\perp) \subset U^0$ ist).

(*ii*) Es gilt $\Psi(v_i)(v_j) = \langle v_i, v_j \rangle = \delta_{ij} = v_i^*(v_j) \ \forall\, i, j \in \{1, ..., n\}$. □

Lemma und Definition 8.23. *Seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ endlichdimensionale euklidische Räume sowie $\varphi : V \to W$ eine lineare Abbildung. Dann existiert genau eine lineare Abbildung*

$$\varphi^{ad} : W \to V$$

mit der Eigenschaft

$$\langle \varphi(v), w\rangle_W = \langle v, \varphi^{ad}(w)\rangle_V \ \forall v \in V, w \in W.$$

Die Funktion φ^{ad} heißt die zu φ adjungierte Abbildung. Sind Φ und Ψ die kanonischen Isomorphismen auf $(V, \langle\cdot,\cdot\rangle_V)$ bzw. $(W, \langle\cdot,\cdot\rangle_W)$, so gilt mit der dualen Abbildung φ^*:

$$\varphi^{ad} = \Phi^{-1} \circ \varphi^* \circ \Psi,$$

d.h. das folgende Diagramm kommutiert:

$$\begin{array}{ccc} V & \xleftarrow{\varphi^{ad}} & W \\ {\scriptstyle\Phi}\downarrow & & \downarrow{\scriptstyle\Psi} \\ V^* & \xleftarrow{\varphi^*} & W^* \end{array}$$

Insbesondere ist die Abbildung $\varphi \mapsto \varphi^{ad}$ ein Isomorphismus.

Beweis. Wir definieren einfach

$$\varphi^{ad} = \Phi^{-1} \circ \varphi^* \circ \Psi \iff \Phi \circ \varphi^{ad} = \varphi^* \circ \Psi. \tag{8.2}$$

Wegen $\varphi^*(h) = h \circ \varphi$ und $\Psi(w) = \langle w, \cdot\rangle_W$ gilt damit

$$\langle \varphi(\cdot), w\rangle_W = \Psi(w) \circ \varphi = \varphi^*(\Psi(w)) \overset{(8.2)}{=} \Phi(\varphi^{ad}(w)) = \langle \cdot, \varphi^{ad}(w)\rangle_V \quad \forall\, w \in W.$$

Umgekehrt folgt aus

$$\langle \varphi(\cdot), w\rangle_W = \langle \cdot, \varphi^{ad}(w)\rangle_V$$

auch (8.2), und damit die Eindeutigkeit. Dass $\varphi \mapsto \varphi^{ad}$ ein Isomorphismus ist, gilt, da $\varphi \mapsto \varphi^*$ nach Lemma 8.12 ein Isomorphismus ist (und die Verkettung von Isomorphismen ebenfalls ein Isomorphismus ist). □

Lemma 8.24. *Für endlichdimensionale euklidische Vektorräume V, W und orthogonales $\varphi : V \to W$ gilt $\varphi^{ad} = \varphi^{-1}$.*

Beweis. $\langle \varphi(v), w\rangle_W = \langle \varphi(v), \varphi(\varphi^{-1}(w))\rangle_W = \langle v, \varphi^{-1}(w)\rangle_V$. □

Lemma 8.25. *Seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ endlichdimensionale euklidische Räume, B eine Orthonormalbasis von V, C eine Orthonormalbasis von W und $\varphi : V \to W$ eine lineare Abbildung. Dann gilt*

$$M_B^C(\varphi^{ad}) = M_C^B(\varphi)^t$$

und damit insbesondere $(\varphi^{ad})^{ad} = \varphi$.

Beweis. Übungsblatt 11, Aufgabe 1 (a). □

Satz 8.26. *Seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ endlichdimensionale euklidische Räume. Ist $\varphi : V \to W$ linear, so gilt*

$$(i)\ \operatorname{Kern}(\varphi^{ad}) = (\operatorname{Bild}(\varphi))^{\perp} \quad \textit{und} \quad (ii)\ \operatorname{Bild}(\varphi^{ad}) = (\operatorname{Kern}(\varphi))^{\perp}.$$

Beweis. Übungsblatt 11, Aufgabe 1 (b) und (c). □

Korollar 8.27. *Seien V endlichdimensionaler euklidischer Vektorraum und $\varphi \in \operatorname{End}_{\mathbb{R}}(V)$. Dann gilt*

$$V = \operatorname{Kern}(\varphi) \oplus \operatorname{Bild}(\varphi^{ad}),$$

wobei das Symbol $\oplus$ die direkte Summe von zueinander orthogonalen Untervektorräumen bezeichnet. Es gilt ferner

$$V = \operatorname{Kern}(\varphi) \oplus (\operatorname{Kern}(\varphi))^{\perp} = \operatorname{Kern}(\varphi) \oplus \operatorname{Bild}(\varphi^{ad}).$$

Beweis. $V = \operatorname{Kern}(\varphi) \oplus (\operatorname{Kern}(\varphi))^{\perp} = \operatorname{Kern}(\varphi) \oplus \operatorname{Bild}(\varphi^{ad})$. □

Definition 8.28. *Sei V ein endlichdimensionaler euklidischer Vektorraum. $\varphi \in \operatorname{End}_{\mathbb{R}}(V)$ heißt selbstadjungiert, falls $\varphi = \varphi^{ad}$ ist.*

Lemma 8.29. *Seien $(V, \langle\cdot,\cdot\rangle_V)$ ein endlichdimensionaler euklidischer Vektorraum und B eine Orthonormalbasis. Dann sind äquivalent:*

(i) φ ist selbstadjungiert.

(ii) $M_B(\varphi)$ ist symmetrisch.

In diesem Fall gilt $V = \operatorname{Kern}(\varphi) \oplus \operatorname{Bild}(\varphi)$.

Beweis. Es gilt

φ selbstadjungiert $\Leftrightarrow \varphi = \varphi^{ad} \Leftrightarrow M_B(\varphi) = M_B(\varphi^{ad}) = M_B(\varphi)^t \Leftrightarrow M_B(\varphi)$ symmetrisch.

Sind die beiden Aussagen (i) und (ii) erfüllt, so ist

$$V = \operatorname{Kern}(\varphi) \oplus \operatorname{Bild}(\varphi^{ad}) = \operatorname{Kern}(\varphi) \oplus \operatorname{Bild}(\varphi)$$

nach Satz 8.26. □

Lemma 8.30. *Seien $(V, \langle\cdot,\cdot\rangle_V)$ ein endlichdimensionaler euklidischer Raum, $U \subset V$ ein Untervektorraum von V und $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ eine selbstadjungierte Abbildung mit $\varphi(U) \subset U$. Dann gilt*

$$\varphi(U^{\perp}) \subseteq U^{\perp}.$$

Beweis. Sei $x \in \varphi(U^{\perp})$, d.h. $\exists\, v \in U^{\perp}$ mit $\varphi(v) = x$. Dann gilt $\forall\, u \in U$

$$\langle u, \varphi(v)\rangle_V = \langle \varphi^{ad}(u), v\rangle_V = \langle \underset{\in U}{\varphi(u)}\,,\ \underset{\in U^{\perp}}{v}\ \rangle_V = 0.$$

Damit ist $x \in U^{\perp}$. □

Lemma 8.31. *Seien $(V, \langle\cdot,\cdot\rangle_V)$ ein endlichdimensionaler euklidischer Vektorraum und $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ selbstadjungiert. Dann zerfällt χ_φ über $\mathbb{R}$ in Linearfaktoren. Insbesondere sind alle Eigenwerte von φ reell.*

Beweis. Sei B eine Orthonormalbasis von V. Wir definieren $A := M_B(\varphi)$, d.h. es gilt $\chi_\varphi = \chi_A$ und $A = A^t$. Wir setzen ferner

$$\widetilde{A}_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n,\ z \mapsto Az.$$

Nach Satz 2.38 zerfällt χ_A über $\mathbb{C}$ in Linearfaktoren, d.h. es gilt

$$\chi_A = \chi_{\widetilde{A}_{\mathbb{C}}} = \prod_{j=1}^{n}(t - \lambda_j),$$

wobei $\lambda_1, ..., \lambda_n \in \mathbb{C}$ die Eigenwerte von $\widetilde{A}_{\mathbb{C}}$ sind.

Zu zeigen: $\lambda_1, ..., \lambda_n \in \mathbb{R}$. Sei dazu $i \in \{1, ..., n\}$ beliebig. Sei $z \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert λ_i von $\widetilde{A}_{\mathbb{C}}$, d.h. es gilt

$$\lambda_i z^t\overline{z} = (\lambda_i z)^t\overline{z} = (Az)^t\overline{z} = z^tA^t\overline{z} = z^tA\overline{z} = z^t\overline{Az} = z^t\overline{\lambda_i z} = \overline{\lambda_i} z^t\overline{z}.$$

Da z als Eigenvektor ungleich 0 ist, gilt zudem $z^t\overline{z} = \sum_{i=1}^{n}|z_i|^2 > 0$. Es folgt $\lambda_i = \overline{\lambda_i}$, also liegt λ_i in $\mathbb{R}$. □

Satz 8.32 (Spektralsatz für selbstadjungierte Abbildungen)**.** *Seien $(V, \langle\cdot,\cdot\rangle_V)$ ein endlichdimensionaler euklidischer Raum und $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ selbstadjungiert. Dann existiert eine Orthonormalbasis von V, die nur aus Eigenvektoren von φ besteht (wobei die zugehörigen Eigenwerte nicht notwendigerweise verschieden sind). Sind $\lambda_1, ..., \lambda_r$ die paarweise verschiedenen Eigenwerte von φ, so gilt*

$$V = \mathrm{Eig}(\varphi, \lambda_1) \oplus ... \oplus \mathrm{Eig}(\varphi, \lambda_r).$$

Beweis. Wir führen diesen Beweis mittels vollständiger Induktion nach $n = \dim(V)$.

Induktionsanfang: Klar, weil jeder von 0 verschiedene Vektor im eindimensionalen Vektorraum V Eigenvektor ist und eine Basis von V bildet.

Induktionsschritt (von $n-1$ nach n, $n \geq 2$): Zunächst existiert nach Satz 6.19 (ii) und Lemma 8.31 ein Eigenwert λ von φ. Sei w_1 ein Eigenvektor von φ zum Eigenwert λ. Setze

$$v_1 = \frac{w_1}{||w_1||}$$

und $U = \mathrm{Lin}(v_1)$. $\Rightarrow \ \varphi(U) \subset U \overset{\text{Lemma } 8.30}{\Longrightarrow} \ \varphi(U^\perp) \subset U^\perp$. Sei $\psi := \varphi_{|U^\perp} : U^\perp \to U^\perp$. Die Abbildung ψ ist selbstadjungiert, da

$$\langle \psi(x), y \rangle = \langle \varphi(x), y \rangle = \langle x, \varphi(y) \rangle = \langle x, \psi(y) \rangle \ \ \forall\, x, y \in U^\perp$$

gilt. Weiter ist

$$V = U \oplus U^\perp \ \Rightarrow \ \dim(U^\perp) = \dim(V) - \dim(U) = n - 1.$$

Nach Induktionsvoraussetzung existiert folglich eine Orthonormalbasis $(v_2, ..., v_n)$ von $U^\perp$, die nur aus Eigenvektoren von ψ besteht (welche alle auch Eigenvektoren von φ sind). Damit ist $(v_1, ..., v_n)$ eine Orthonormalbasis von V aus Eigenvektoren von φ, und es gilt mit der Induktionsvoraussetzung

$$V = \mathrm{Eig}(\varphi, \lambda_1) \oplus (\mathrm{Eig}(\varphi, \lambda_2) \oplus ... \oplus \mathrm{Eig}(\varphi, \lambda_r)) = \mathrm{Eig}(\varphi, \lambda_1) \oplus ... \oplus \mathrm{Eig}(\varphi, \lambda_r).$$

□

9 Ideale und euklidische Ringe

In diesem Kapitel seien R und S **immer** kommutative Ringe mit Einselement.

9.1 Ringhomomorphismen und Ideale

Ein Ringhomomorphismus ist eine strukturerhaltende Abbildung zwischen zwei Ringen. Für einen "Homomorphismus von Ringen mit Eins" wird zusätzlich gefordert, dass die Eins auf die Eins abgebildet wird.

Definition 9.1. *Eine Abbildung $\varphi : R \to S$ heißt Ringhomomorphismus von Ringen mit Eins, falls folgende Bedingungen erfüllt sind:*

(RH1) $\varphi(a + b) = \varphi(a) + \varphi(b)$,

(RH2) $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ *und*

(RH3) $\varphi(1_R) = 1_S$.

Eine zentrale Unterstruktur kommutativer Ringe sind Ideale. Sie sind wichtig, weil sie als Kerne von Ringhomomorphismen auftreten und die Definition von Faktorringen ermöglichen.

Definition 9.2. *$I \subset R$ heißt Ideal in R, falls folgende Bedingungen erfüllt sind:*

(I1) $0 \in I$,

(I2) $a, b \in I \Rightarrow a + b \in I$ *und*

(I3) $a \in I, r \in R \Rightarrow r \cdot a \in I$.

Beispiele 9.3.

(i) $(a) = \{ra \mid r \in R\}$ *für ein $a \in R$, $\{0\}$ und R sind Ideale in R.*

(ii) Für $n \in \mathbb{Z}$ ist $(n) = \{nr \mid r \in \mathbb{Z}\}$ ein Ideal in $\mathbb{Z}$.

Man schreibt in (i) auch aR statt (a) und in (ii) auch $n\mathbb{Z}$ statt (n).

Lemma und Definition 9.4. *Sei $\varphi : R \to S$ ein Ringhomomorphismus. Dann gilt:*

(i) Ist $I \subset S$ ein Ideal in S, so ist $\varphi^{-1}(I) \subset R$ ein Ideal in R.

(ii) $\mathrm{Kern}(\varphi) := \{a \in R \mid \varphi(a) = 0\}$ *ist ein Ideal in R.*

(iii) φ injektiv $\Leftrightarrow \mathrm{Kern}(\varphi) = \{0\}$

(iv) Ist $I \subset R$ ein Ideal in R und φ surjektiv, so ist $\varphi(I) \subset S$ ein Ideal in S.

(v) $\mathrm{Bild}(\varphi) := \varphi(R)$ *ist ein Unterring von S.*

Beweis. (i) $(I1)$ gilt, da: $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0) \;\Rightarrow\; \varphi(0) = 0$, d.h. $0 \in \varphi^{-1}(I)$. $(I2)$ und $(I3)$ ergeben sich durch Nachrechnen.

(ii) folgt aus (i), da $\mathrm{Kern}(\varphi) = \varphi^{-1}(\{0\})$ ist, und $\{0\}$ ein Ideal ist.

(iii)–(v): Bonusblatt 12, Aufgabe 1. □

Bemerkung. *(iv) wird falsch, wenn φ nicht surjektiv ist. Gegenbeispiel:*
Die kanonische Inklusion $\varphi : \mathbb{Z} \to \mathbb{Q},\ x \mapsto x$ ist ein Ringhomomorphismus. $\varphi(\mathbb{Z}) = \mathbb{Z}$ ist nach (v) ein Unterring von $\mathbb{Q}$. $\mathbb{Z}$ ist ein Ideal in $\mathbb{Z}$, aber kein Ideal in $\mathbb{Q}$, denn:

$$\underbrace{\frac{1}{3}}_{\in\mathbb{Q}} \cdot \underbrace{2}_{\in\mathbb{Z}} = \frac{2}{3} \notin \mathbb{Z}$$

Lemma und Definition 9.5. *Sei $I \subset R$ ein Ideal. Dann ist durch*

$$v_1 \sim v_2 \ :\Longleftrightarrow\ v_1 - v_2 \in I$$

eine Äquivalenzrelation auf R gegeben. Die Äquivalenzklasse von r ist

$$\overline{r} := r + I := \{r + a \mid a \in I\}$$

und heißt Restklasse von r modulo I. Die Menge der Restklassen bezeichnen wir mit

$$R/I := \{\overline{r} \mid r \in R\}.$$

Beweis. Reflexivität, Symmetrie und Transitivität der Relation $\sim$ verifiziert man durch einfaches Nachrechnen. □

Lemma und Definition 9.6. *Seien $I \subset R$ ein Ideal und $\sim$ die zugehörige Äquivalenzrelation aus Lemma 9.5. Dann gelten folgende Implikationen:*

$$r_1 \sim r_2,\ s_1 \sim s_2 \ \Rightarrow\ r_1 + s_1 \sim r_2 + s_2 \text{ und } r_1 \cdot s_1 \sim r_2 \sim s_2.$$

Damit sind die folgende Addition

$$+ : R/I \times R/I \to R/I,\ \overline{r} + \overline{s} := \overline{r + s}$$

und die folgende Multiplikation

$$\cdot : R/I \times R/I \to R/I,\ \overline{r} \cdot \overline{s} := \overline{r \cdot s}$$

wohldefiniert. Mit diesen Verknüpfungen wird R/I zu einem kommutativen Ring mit Eins, dem sogenannten Faktorring (oder Restklassenring) R/I. Die Abbildung

$$\pi : R \to R/I,\ r \mapsto \overline{r},$$

ist ein surjektiver Ringhomomorphismus mit $\operatorname{Kern}(\pi) = I$.

Beweis. Übungsblatt 11, Aufgabe 2. □

Satz 9.7 (Homomorphiesatz für Ringe)**.** *Sei $\varphi : R \to S$ ein Ringhomomorphismus. Dann existiert ein Ringisomorphismus*

$$\Phi : R/\operatorname{Kern}(\varphi) \to \operatorname{Bild}(\varphi),\ \overline{r} = r + \operatorname{Kern}(\varphi) \mapsto \varphi(r).$$

Bemerkung. $\mathrm{Kern}(\varphi)$ *ist nach Lemma 9.4 (ii) ein Ideal, folglich* $R/\mathrm{Kern}(\varphi)$ *kommutativer Ring mit* 1 *nach Lemma 9.6. Nach Lemma 9.4 (v) ist* $\mathrm{Bild}(\varphi)$ *Unterring von* S*.*

Beweis. Φ ist wohldefiniert: Seien $r_1, r_2 \in R$ mit $\overline{r_1} = \overline{r_2}$. Daraus folgt $r_1 - r_2 \in \mathrm{Kern}(\varphi)$ $\Rightarrow \varphi(r_1 - r_2) = 0 \Rightarrow \varphi(r_1) = \varphi(r_2)$.

Φ ist ein Homomorphismus:

$$\Phi(\overline{r_1} + \overline{r_2}) \overset{\text{Lemma } 9.6}{=} \Phi(\overline{r_1 + r_2}) = \varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2) = \Phi(\overline{r_1}) + \Phi(\overline{r_2}),$$

also gilt (RH1). Für (RH2) verfährt man analog; mit $\Phi(\overline{1}) = \varphi(1) = 1$ gilt auch (RH3).

Φ ist injektiv: Sei $\overline{r} \in R/\mathrm{Kern}(\varphi)$ mit $\Phi(\overline{r}) = 0$. Es folgt $\overline{r} = r + \mathrm{Kern}(\varphi)$ mit $\varphi(r) = 0$ $\Rightarrow r \in \mathrm{Kern}(\varphi) \Rightarrow \overline{r} = \overline{0} \Rightarrow \mathrm{Kern}(\Phi) = \{\overline{0}\} \Rightarrow \Phi$ ist injektiv nach Lemma 9.4 (iii).

Φ ist surjektiv: Klar nach Konstruktion. □

Beispiel 9.8. *Sei* K *ein Körper,* $R = K[t]$ *und* $\varphi : K[t] \to K,\ \varphi(f) = f(0)$*. Es gilt*

$$\mathrm{Bild}(\varphi) = K, \quad \mathrm{Kern}(\varphi) = \{f \in K[t] \mid f(0) = 0\} \overset{\text{Lemma } 2.35}{=} t \cdot K[t],$$

und man verifiziert leicht, dass φ *ein Ringhomomorphismus ist. Gemäß Satz 9.7 ist*

$$\Phi : K[t]/t \cdot K[t] \to K,\ f + t \cdot K[t] \mapsto f(0)$$

ein Ringisomorphismus.

Als Nächstes wollen wir Ergebnisse über die Eindeutigkeit von Primfaktorzerlegungen formulieren, beispielsweise in $R = K[t]$. Hier gilt zum Beispiel

$$2t^2 - 2 = 2(t-1)(t+1) = 2 - 3\left(\frac{t}{3} - \frac{1}{3}\right)(t+1),$$

d.h. wir können Eindeutigkeit höchstens bis auf "invertierbare" Faktoren (hier 3 und $\frac{1}{3}$) erwarten (solche nennen wir "Einheiten", die Elemente $(t-1)$ und $\left(\frac{t}{3} - \frac{1}{3}\right)$ bezeichnen wir als "assoziiert").

Lemma und Definition 9.9. *Ein Element* $x \in R$ *heißt* <u>*Einheit*</u>*, falls es ein* $y \in R$ *gibt mit* $xy = 1$*. Die Menge*

$$R^* := \{x \in R \mid x \text{ ist eine Einheit}\}$$

bildet eine Gruppe bzgl. der Ringmultiplikation "$\cdot$*".*

Beweis. Sind $x_1, x_2 \in R^*$, d.h. $\exists\, y_1, y_2 \in R$ mit $x_1 y_1 = x_2 y_2 = 1$. $\overset{R \text{ kom.}}{\Rightarrow}$ $(x_1 x_2)(y_1 y_2) = (x_1 y_1)(x_2 y_2) = 1$, d.h. $x_1 x_2 \in \mathbb{R}^*$. Weiter gilt $1 \in R^*$, da $1 \cdot 1 = 1$. Das inverse Element zu $x \in R^*$ ist gerade das $y \in R$ mit $xy = 1$. Kommutativität und Assoziativität gelten, da $R^* \subset R$ ist. □

Beispiele 9.10. *(i)* $\mathbb{Z}^* = \{-1, 1\}$, *da* $1 \cdot 1 = 1$ *und* $(-1) \cdot (-1) = 1$.

(ii) Ist K ein Körper, so ist $K^* = K \setminus \{0\}$.

Definition 9.11. *Seien $a_1, \ldots, a_n \in R$. Dann heißt*

$$(a_1, \ldots, a_n) := \left\{ \sum_{i=1}^{n} a_i r_i \;\middle|\; r_1, \ldots, r_n \in R \right\} = a_1 R + \ldots + a_n R$$

das von den $a_1, \ldots, a_n$ erzeugte Ideal.

Definition 9.12. *Sei $I \subset R$ ein Ideal. I heißt Hauptideal, falls es ein $a \in R$ gibt mit*

$$I = (a) = \{ra \mid r \in R\} = aR.$$

R heißt Hauptidealring, falls R nullteilerfrei ist und jedes Ideal I in R ein Hauptideal ist.

Beispiele 9.13. *(i)* $\mathbb{Z}$ *ist ein Hauptidealring.*

(ii) Für einen Körper K ist $K[t]$ ein Hauptidealring.

(iii) $\mathbb{Z}[t]$ *ist kein Hauptidealring.*
Beweis: Wir zeigen, dass es kein $f \in \mathbb{Z}[t]$ gibt mit $(f) = (2, t)$. Angenommen, es gibt ein solches $f \in \mathbb{Z}[t]$. Dann muss einerseits ein $h_1 \in \mathbb{Z}[t]$ existieren mit $2 = h_1 \cdot f$. $\Rightarrow$ $deg(h_1) = deg(f) = 0$ $\Rightarrow$ f ist konstant $\Rightarrow$ $\exists a \in \mathbb{Z} : f = a$. Andererseits wäre aber auch $t = h_2 \cdot f = h_2 \cdot a$ mit $h_2 \in \mathbb{Z}[t]$. Da t normiert ist, muss $a = \pm 1$ sein, d.h. $f = \pm 1$. Aber $\pm 1 \notin (2, t)$. ↯

Lemma und Definition 9.14. *$b \in R$ heißt Teiler von $a \in R$ (Notation: $b|a$), falls es ein $c \in R$ gibt mit $a = b \cdot c$. $a \in R$ und $b \in R$ heißen assoziiert (Notation: $a \widehat{=} b$), falls es eine Einheit $u \in R^*$ gibt mit $a = bu$. Es gilt:*

$$a \widehat{=} b \;\Leftrightarrow\; a|b \text{ und } b|a \;\Leftrightarrow\; (a) = (b).$$

Beweis. Bonusblatt 12, Aufgabe 2. □

Definition 9.15. *Seien R nullteilerfrei und $a_1, \ldots, a_n \in R$. $d \in R$ heißt größter gemeinsamer Teiler (kurz: ggT) von $a_1, \ldots, a_n$, falls gilt:*

(GGT1) $d|a_1, \ldots, d|a_n$. *(GGT2) $c \in R$ mit $c|a_1, \ldots, c|a_n \Rightarrow c|d$.*

Die Menge der größten gemeinsamen Teiler wird mit $\mathrm{GGT}(a_1, \ldots, a_n)$ bezeichnet.

Lemma 9.16. *Seien R nullteilerfrei und seien $a_1, \ldots, a_n \in R$. Dann gelten:*

(i) $d_1, d_2 \in \mathrm{GGT}(a_1, \ldots, a_n) \Leftrightarrow d_1 | d_2$ *und* $d_2 | d_1 \Leftrightarrow d_1 \widehat{=} d_2 \Leftrightarrow (d_1) = (d_2)$,

(ii) $\mathrm{GGT}(a_1, \ldots, a_n) = \mathrm{GGT}(a, a_n)$ *für jedes* $a \in \mathrm{GGT}(a_1, \ldots, a_{n-1})$.

Beweis. Bonusblatt 12, Aufgabe 3. □

Satz 9.17. *Sei R ein Hauptidealring und seien $a, b \in R$. Dann sind äquivalent:*

(i) $d \in \mathrm{GGT}(a, b)$;

(ii) $(d) = (a, b)$.

Beweis. Da R Hauptidealring ist, existiert $\widetilde{d}$ mit $(\widetilde{d}) = (a, b)$, womit insbesondere $\widetilde{d} | a$ und $\widetilde{d} | b$.

(i) $\Rightarrow$ (ii): Sei $d \in \mathrm{GGT}(a, b)$. Damit gilt $\widetilde{d} | d$. Wegen $\widetilde{d} \in (a, b)$ folgt $d | \widetilde{d} \Rightarrow (d) = (\widetilde{d}) = (a, b)$.

(ii) $\Rightarrow$ (i): Sei $(d) = (a, b)$. Dann gilt $d | a$ und $d | b$ (i.e. $(GGT1)$). Ferner folgt, dass es $u, v \in R$ gibt mit $d = au + bv$, d.h. aus $c | a$ und $c | b$ folgt $c | d$ (i.e. $(GGT2)$). □

Definition 9.18. *Seien R nullteilerfrei und $p \in R \setminus (R^* \cup \{0\})$.*

(i) p heißt Primelement, falls für beliebige $a, b \in R$ aus $p | ab$ stets $p | a$ oder $p | b$ folgt.

(ii) p heißt irreduzibel (oder unzerlegbar), falls für beliebige $a, b \in R$ aus $p = ab$ stets $a \in R^*$ oder $b \in R^*$ folgt.

Satz 9.19. *Sei R ein Hauptidealring. Dann ist $p \neq 0$ genau dann irreduzibel, wenn p ein Primelement ist.*

Beweis. Später. [Wir führen im nächsten Abschnitt einen Beweis der Richtung "$\Rightarrow$" für den Spezialfall euklidischer Ringe (Lemma 9.30).] □

Definition 9.20. *Sei R nullteilerfrei. R heißt faktoriell, falls jedes $a \in R \setminus (R^* \cup \{0\})$ eine bis auf Reihenfolge und Assoziiertheit eindeutige Zerlegung $a = p_1 \cdot \ldots \cdot p_r$ in irreduzible Faktoren $p_1, \ldots, p_r$ besitzt. [Die Eindeutigkeit bedeutet hier, dass aus $a = p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s$ stets folgt, dass $r = s$ ist und $p_i \widehat{=} q_i$ $\forall i \in \{1, \ldots, r\}$ (nach Umnummerierung).]*

Satz 9.21. *Jeder Hauptidealring ist faktoriell.*

Beweis. Später. Wir präsentieren an dieser Stelle lediglich eine Beweisskizze:

(i) Es genügt zu zeigen, dass für einen Hauptidealring R jedes Element in $R \setminus (R^* \cup \{0\})$ als Produkt irreduzibler Elemente darstellbar ist.

(ii) *Widerspruchsannahme:* Es gibt ein Element, das keine solche Darstellung besitzt.

(iii) Verwende dieses Element, um eine unendliche echt aufsteigende Kette von Idealen in R zu konstruieren.

(iv) Die Vereinigung aller Ideale in dieser Kette ist wieder ein Ideal I. Die Hauptidealeigenschaft von I ist *unvereinbar* mit der Existenz der unendlichen aufsteigenden Kette. □

9.2 Euklidische Ringe

Definition 9.22. *Sei R nullteilerfrei. R heißt euklidischer Ring, falls es eine Abbildung $\delta : R \setminus \{0\} \to \mathbb{N}_0$ gibt, so dass gilt: Für alle $f, g \in R$ mit $g \neq 0$ existieren $q, r \in R$ mit $f = q \cdot g + r$ und*

$$\delta(r) < \delta(g) \text{ oder } r = 0.$$

δ heißt Normabbildung auf R.

Beispiele 9.23. *(i) Sei $R = \mathbb{Z}$. Mit $\delta = |\cdot|$ wird R zu einem euklidischen Ring.*

(ii) Ist K ein Körper und $R = K[x]$, so ist R mit $\delta = \deg$ ein euklidischer Ring.

(iii) Sei $R = \mathbb{Z}(i) = \{a+bi \mid a, b \in \mathbb{Z}\} \subset \mathbb{C}$. Mit $\delta(x+iy) = x^2+y^2$ ist R ein euklidischer Ring. Dieser heißt Ring der ganzen gaußschen Zahlen.

Beweis: Übungsaufgabe (vgl. Aufgabe 4 auf Bonusblatt 12).

Satz 9.24. *Sei R ein euklidischer Ring. Dann ist R ein Hauptidealring.*

Beweis. Sei $I \subset R$ ein Ideal mit $I \neq 0$ (im Fall $I = 0$ wird I von der Null erzeugt, und ist damit ein Hauptideal). Es gilt: $\emptyset \neq \{\delta(a) \mid a \in I \setminus \{0\}\} \subseteq \mathbb{N}_0$. Wähle nun $a \in I \setminus \{0\}$, so dass $\delta(a)$ minimal ist. *Behauptung:* $I = (a)$.

"$\subset$": Wegen $a \in I$ gilt $(a) \subset I$.

"$\supset$": Sei $f \in I$. Dann existieren $q, r \in R$ mit $f = qa + r$ und $\delta(r) < \delta(a)$ oder $r = 0$. Damit gilt

$$\underbrace{f}_{\in I} - \underbrace{qa}_{\in I} = r \in I.$$

Da $\delta(a)$ aber minimal ist, muss $r = 0$ sein. Es folgt $f = qa \in (a)$. □

Korollar 9.25. *Sei R ein euklidischer Ring. Dann ist R faktoriell.*

Beweis. R ist euklidischer Ring $\overset{\text{Satz } 9.24}{\Longrightarrow}$ R ist Hauptidealring $\overset{\text{Satz } 9.21}{\Longrightarrow}$ R ist faktoriell. □

Korollar 9.26. *Sei K ein Körper. Dann ist $K[t]$ faktoriell, d.h. für jedes $f \in K[t]$ existiert eine (bis auf Reihenfolge und Assoziiertheit) eindeutige Zerlegung in irreduzible Polynome. Entsprechendes gilt für $\mathbb{Z}$.*

Beweis. $K[t]$ und $\mathbb{Z}$ sind euklidische Ringe (Beispiele 9.23) $\overset{\text{Kor. 9.25}}{\Longrightarrow}$ Behauptung. □

Der euklidische Algorithmus. Sei R ein euklidischer Ring, und seien $a, b \in R \setminus \{0\}$ mit $\delta(b) \leq \delta(a)$. Setze $c_1 := a$, $c_2 := b$, also $\delta(c_2) \leq \delta(c_1)$. Der euklidische Algorithmus ist eine Sequenz von Divisionen

$$\begin{aligned} c_1 &= q_1 c_2 + c_3 \text{ mit } \delta(c_3) < \delta(c_2) \\ c_2 &= q_2 c_3 + c_4 \text{ mit } \delta(c_4) < \delta(c_3) \\ &\vdots \\ c_k &= q_k c_{k+1} + c_{k+2} \text{ mit } \delta(c_{k+2}) < \delta(c_{k+1}), \end{aligned}$$

endend mit $c_{n-1} = q_{n-1} c_n + 0$. Dieser Algorithmus terminiert nach endlich vielen Schritten, da $\delta(c_k) \in \mathbb{N}_0$ in jedem Schritt strikt abnimmt. Die Existenz der Darstellungen in jedem Schritt folgt sofort aus der Definition eines euklidischen Rings.

Satz 9.27 (Euklidischer Algorithmus)**.** *Seien $c_1, \ldots, c_n$ wie oben und $d := c_n$. Dann gilt:*

(i) $d|a$ und $d|b$.

(ii) Gilt $g|a$ und $g|b$, so folgt $g|d$.

Damit ist $d \in GGT(a, b)$. Ferner gilt:

(iii) Es existieren $u, v \in R$ mit $d = ua + vb$ (Lemma von Bézout).

Beweis. (i) Wir steigen den Algorithmus (von unten) auf: d teilt c_n (wegen $d = c_n$) und $c_{n-1} = q_{n-1} c_n + 0$ (s. Algorithmus), d.h. es teilt auch $c_{n-2} = q_{n-2} c_{n-1} + c_n$. Induktiv folgt $\forall\, k \in \{0, \ldots, n-2\}$: $d|c_{k+1}$ und $d|c_{k+2}$, d.h. schließlich auch $d|a$ und $d|b$.

(ii) Wir steigen den Algorithmus (von oben) ab:

$$g|a \text{ und } g|b \overset{c_3 = c_1 - q_1 c_2}{\Longrightarrow} g|c_3 \text{ (und } g|c_2) \overset{c_4 = c_2 - q_2 c_3}{\Longrightarrow} g|c_4 \text{ (und } g|c_3) \Longrightarrow \ldots \Longrightarrow g|c_n.$$

(iii) Wir steigen den Algorithmus ab: $a = c_1$ und $b = c_2$ haben offenbar die Darstellung $c_i = u_i a + v_i b$, $i = 1, 2$ (da $a = 1 \cdot a + 0 \cdot b$ und $b = 0 \cdot a + 1 \cdot b$). Wegen $c_{k+2} = c_k - q_k c_{k+1}$ folgt induktiv daraus schließlich, dass auch c_{n-1} und $c_n = d$ solch eine Darstellung besitzen. □

Beispiel 9.28. *Seien $R = \mathbb{Z}$ mit $\delta = |\cdot|$, $a = 24$, $b = 15$. Es gilt gemäß dem Algorithmus:*

$$\begin{aligned}
\underset{c_1}{24} &= 1 \cdot \underset{c_2}{15} + \underset{c_3}{9} \\
\underset{c_2}{15} &= 1 \cdot \underset{c_3}{9} + \underset{c_4}{6} \\
\underset{c_3}{9} &= 1 \cdot \underset{c_4}{6} + \underset{c_5}{3} \\
\underset{c_4}{6} &= 2 \cdot \underset{c_5}{3} + 0
\end{aligned}$$

Folglich ist $d = 3$ ein größter gemeinsamer Teiler von 24 und 15.

Bemerkung 9.29. *Man kann für euklidische Ringe viele Aussagen, die allgemein für Hauptidealringe gelten, auch direkt beweisen – insbesondere, dass sie faktoriell sind. Beispielsweise folgt aus dem Lemma von Bézout unmittelbar Satz 9.17, und man erhält die Implikation "irreduzibel $\Rightarrow$ prim" aus Satz 9.19 recht einfach wie folgt.*

Lemma 9.30 (Euklidisches Lemma)**.** *Sei R ein euklidischer Ring. Ist $p \neq 0$ irreduzibel, so ist p auch Primelement.*

Beweis. Seien $a, b \in R$ mit $p|ab$ und $p \nmid a$. [Zu zeigen: $p|b$.] Es folgt $\mathrm{GGT}(p, a) = R^*$, denn gäbe es $d \notin R^*$ mit $p = dr$ und $a = ds$, so würde aus der Irreduzibilität von p folgen, dass $r \in R^*$, also $pr^{-1} = d \Rightarrow a = pr^{-1}s \Rightarrow p|a$. $\lightning$ Wegen $(d) = (a, b) \Leftrightarrow d \in \mathrm{GGT}(a, b)$ nach Satz 9.17 existieren damit $r, s \in R$ mit $1 = pr + as$. Also folgt $b = prb + abs \overset{p|ab}{\Longrightarrow} p|b$. $\square$

9.3 Elementarteilersatz und Determinantenteiler

Im Hinblick auf das nächste Kapitel zu Normalformen wenden wir uns nun $(n \times n)$-Matrizen mit Einträgen aus R zu, die Menge dieser wird mit $M(n \times n, R)$ bezeichnet. Auch für solche Matrizen kann man dann elementare Zeilen- bzw. Spaltenumformungen durchführen. Das Problem ist allerdings, dass der Gaußalgorithmus über Körpern K die Existenz der inversen Elemente für $K \setminus \{0\}$ bzgl. "$\cdot$" verwendet, die in Ringen nicht garantiert ist. Folglich müssen wir ihn modifizieren, um Matrizen mit Einträgen in R immer noch auf ZSF zu bringen.

Satz 9.31 (Gaußdiagonalisierung für euklidische Ringe)**.** *Seien R ein euklidischer Ring und $A \in M(n \times n, R)$. Dann lässt sich A durch elementare Zeilen- und Spaltenumformungen vom Typ II und III gemäß Definition 3.52 in eine Matrix der Gestalt*

$$\left(\begin{array}{ccc|c}
c_1 & & 0 & \\
& \ddots & & 0 \\
0 & & c_r & \\
\hline
& 0 & & 0 \\
& & &
\end{array}\right)$$

mit $c_1, \ldots, c_r \in R \setminus \{0\}$ und $c_1|c_2| \ldots |c_r$ bringen.

Beweis (Algorithmus zur Durchführung). Im Falle $A = 0$ sind wir bereits fertig, sei also $A \neq 0$. *Notation:* Wir verwenden nachfolgend für alle umgeformten Matrizen den Buchstaben A.

Schritt 1: Durch Zeilen- und Spaltenvertauschungen erreicht man $a_{11} \neq 0$ mit der Eigenschaft $\delta(a_{11}) \leq \delta(a_{ij})$ für alle i, j mit $a_{ij} \neq 0$.

Schritt 2: Bringe A auf die neue Form

$$A = \left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right). \tag{9.1}$$

Falls A diese Form bereits hat, ist nichts dafür zu tun. Andernfalls existiert in der ersten Zeile oder Spalte noch ein Element $\neq 0$. Sei ohne Einschränkung $a_{21} \neq 0$. Nach Schritt 1 ist $\delta(a_{11}) \leq \delta(a_{21}) \Rightarrow \exists\, q \in R$ mit $a_{21} = qa_{11}$ oder $\delta(a_{21} - qa_{11}) < \delta(a_{11})$. Addiere nun das $(-q)$-fache der ersten Zeile zur zweiten Zeile. Man erhält damit eine neue Matrix A mit $a_{21} = 0$ oder $\delta(a_{21}) < \delta(a_{11})$. Falls $\delta(a_{21}) < \delta(a_{11})$ gehe zurück zu Schritt 1. In diesem Falle ist a_{21} das Element mit dem kleinsten $\delta(a_{ij})$ unter allen von Null verschiedenen Einträgen von A und wird dann zum neuen a_{11}. Da δ nach unten durch Null beschränkt ist, liefert diese Iteration nach endlich vielen Schritten (und für alle außer a_{11} von Null verschiedenen Elemente der ersten und zweiten Zeile) A von der Form (9.1).

3. Schritt: Bringe A auf die neue Form

$$A = \left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right) \text{ mit der zusätzlichen Eigenschaft } a_{11} | a_{ij}\ \forall\, i, j. \tag{9.2}$$

Falls a_{11} bereits alle übrigen Einträge von A teilt, ist nichts zu tun. Andernfalls existiert ein Paar (i, j) mit $a_{11} \nmid a_{ij}$. $\Rightarrow \exists\, q \in R$ mit $a_{ij} = qa_{11} + r$ und $\delta(a_{ij} - qa_{11}) < \delta(a_{11})$. Addition der ersten zur i-ten Zeile ($i \neq 1$, da $a_{11} | 0$) ergibt

$$\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \\ a_{11} & a_{i2} & \ldots a_{ij} \ldots & a_{in} \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right).$$

Anschließende Subtraktion des q-fachen der ersten Spalte von der j-ten Spalte ergibt

$$\left(\begin{array}{c|ccc} a_{11} & 0 \ \dots \ 0 & -qa_{11} & 0 \ \dots \ 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \\ a_{11} & * & a_{ij} - qa_{11} & * \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right) \quad \text{mit } \delta(a_{ij} - qa_{11}) < \delta(a_{11}).$$

Wende nun die gesamten bisherigen Schritte auf diese neue Matrix an, beginnend damit, dass man gemäß Schritt 1 den Eintrag $a_{ij} - qa_{11}$ an die Stelle $(1,1)$ verschiebt. Dieses Prozedere iteriert man so lange, bis A von der Form (9.2) ist (was nach endlich vielen Iterationen passiert).

Schritt 4: Wir erhalten also eine Matrix der Gestalt

$$A = \left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & a_{11}B & \\ 0 & & & \end{array}\right) \quad \text{mit } B \in M\big((n-1) \times (n-1), R\big).$$

Wende unser bisheriges Verfahren nun auf B an. Iteration ergibt dann die Aussage des Satzes. □

Beispiele 9.32. *(i)* $R = \mathbb{Z}$ *mit* $\delta = |\cdot|$

$$A = \begin{pmatrix} 5 & 3 \\ 4 & 6 \end{pmatrix} \to \begin{pmatrix} 3 & 5 \\ 6 & 4 \end{pmatrix} \to \begin{pmatrix} 3 & 2 \\ 6 & -2 \end{pmatrix} \to \begin{pmatrix} 2 & 3 \\ -2 & 6 \end{pmatrix} \to \begin{pmatrix} 2 & 3 \\ 0 & 9 \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -18 \end{pmatrix} \leftarrow \begin{pmatrix} 1 & 0 \\ 9 & -18 \end{pmatrix} \leftarrow \begin{pmatrix} 1 & 2 \\ 9 & 0 \end{pmatrix} \leftarrow \begin{pmatrix} 2 & 1 \\ 0 & 9 \end{pmatrix}$$

(ii) $R = \mathbb{Q}[t]$ *mit* $\delta = \deg$

$$A = \begin{pmatrix} t-1 & 0 \\ -1 & t-1 \end{pmatrix} \to \begin{pmatrix} -1 & t-1 \\ t-1 & 0 \end{pmatrix} \to \begin{pmatrix} -1 & 0 \\ t-1 & (t-1)^2 \end{pmatrix} \to \begin{pmatrix} -1 & 0 \\ 0 & (t-1)^2 \end{pmatrix}$$

Wir erweitern nun die Äquivalenzrelationen aus Definition 4.51 von $M(n \times n, K)$-Matrizen für einen Körper K auf Matrizen aus $M(n \times n, R)$.

Definition 9.33. *Seien $A, B \in M(n \times n, R)$.*

(i) A heißt äquivalent zu B ($A \sim B$), falls es invertierbare $M(n \times n, R)$-Matrizen P, Q gibt mit $B = P^{-1}AQ$.

(ii) A heißt ähnlich zu B ($A \approx B$), falls es eine invertierbare Matrix $S \in M(n \times n, R)$ gibt mit $B = S^{-1}AS$.

Der für das nächste Kapitel zentrale Satz dieses letzten Abschnittes ist der folgende.

Satz 9.34 (Elementarteilersatz über euklidischen Ringen)**.** *Seien R ein euklidischer Ring und $A \in M(n \times n, R)$. Dann existieren $c_1, \ldots, c_r \in R \setminus \{0\}$ mit $c_1 \mid c_2 \mid \ldots \mid c_r$, so dass*

$$A \sim \left(\begin{array}{ccc|c} c_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & c_r & \\ \hline & 0 & & 0 \end{array}\right) =: C.$$

$r \in \mathbb{N}_0$ ist eindeutig bestimmt, und $c_1, \ldots, c_r$ sind bis auf Assoziiertheit eindeutig bestimmt. $c_1, \ldots, c_r$ heißen Elementarteiler von A.

Beweis der Existenz. Diese ergibt sich aus Satz 9.31, da die elementaren Umformungen von Typ II und III einer Multiplikation von links (für Zeilenumformungen) bzw. von rechts (für Spaltenumformungen) mit den jeweiligen auch in R invertierbaren Elementarmatrizen $ZM(i, j, \lambda)$ und $ZV(i, j)$ aus Lemma 3.54 entsprechen. [Die Eindeutigkeit zeigen wir im Anschluss an Satz 9.36 unten.] □

Um die im Elementarteilersatz behauptete Eindeutigkeit zu verifizieren, führen wir nun noch sogenannte Determinantenteiler ein. Dabei wird $\det A$ für $A \in M(n \times n, R)$ durch die Leibniz-Formel definiert und wir verstehen unter $(l \times l)$-Untermatrizen von A all diejenigen $M(l \times l, R)$-Matrizen, die durch Streichen $n - l$ beliebiger Zeilen und $n - l$ beliebiger Spalten aus A entstehen.

Definition 9.35. *Seien R euklidischer Ring und $A \in M(n \times n, R)$. Wir nennen*

$$\bar{d}_l(A) \in GGT\big\{ \det(B) \,\big|\, B \text{ ist } (l \times l)\text{-Untermatrix von } A \big\}$$

einen l-ten Determinantenteiler von A [dieser ist eindeutig bis auf Assoziiertheit gemäß Lemma 9.16 (i)].

Satz 9.36. *In der Situation von Satz 9.34 gilt mit* $c_k := 0$ *für* $k > r$

$$\bar{d}_l(A) \mathrel{\widehat{=}} c_1 \cdots c_l$$

für alle $l = 1, \ldots, n$*; insbesondere* $\det A \mathrel{\widehat{=}} c_1 \cdots c_n$*.*

Beweis. Der Beweis untergliedert sich in zwei Schritte:

(i) $\bar{d}_l(A)$ ist invariant unter Zeilen- und Spaltenumformungen von A vom Typ II/III.

(ii) $\bar{d}_l(C) \mathrel{\widehat{=}} c_1 \cdots c_l$ mit C aus Satz 9.34.

(i) Sei A' die Matrix, die durch Ersetzen der i-ten Zeile durch das λ-fache der j-ten Zeile + i-te Zeile ($\lambda \in R$) ensteht (Zeilenumformung vom Typ II). Zu zeigen: $\bar{d}_l(A') \mathrel{\widehat{=}} \bar{d}_l(A)$ $\forall\, l \in \{1, \ldots, n\}$, d.h. "$d$ teilt alle $(l \times l)$-Unterdeterminanten von A" $\Leftrightarrow$ "d teilt alle $(l \times l)$-Unterdeterminanten von A'". Sei nun U eine $(l \times l)$-Untermatrix von A und U' die entsprechende von A'. Es gibt drei Fälle:

(1) Die i-te Zeile (der ursprünglich n Zeilen) ist nicht in U und U' enthalten (d.h. sie wurde beim Bilden von U aus A bzw. von U' aus A' gestrichen). Dann gilt $U = U'$ und entsprechend $d \,|\, \det(U) \;\Leftrightarrow\; d \,|\, \det(U')$.

(2) Die (ursprünglich) i-te und j-te Zeile sind beide in U und U' enthalten (d.h. beide wurden beim Bilden von der $(l \times l)$-Untermatrix nicht gestrichen). Dann entsteht U' aus U durch Multiplikation von links mit der Matrix $ZA(i, j, \lambda)$ aus Lemma 3.54 und $\det(U) = \det(U')$, also ebenfalls $d \,|\, \det(U) \;\Leftrightarrow\; d \,|\, \det(U')$.

(3) Die (ursprünglich) i-te Zeile ist in U und U' enthalten, die (ursprünglich) j-te aber nicht. In diesem Falle sei V die Matrix, die aus U dadurch entsteht, dass die (ursprünglich) i-te durch die (ursprünglich) j-te Zeile ausgetauscht ist. Linearität der Determinante in jeder Zeile impliziert $\det U' = \det U + \lambda \det V$ und damit "$d \,|\, \det U$ und $d \,|\, \det V$" $\Leftrightarrow$ "$d \,|\, \det U'$ und $d \,|\, \det V$". V ist aber bis auf Zeilenvertauschungen ($\det \to -\det$) auch eine $(l \times l)$-Untermatrix, die bei der Zeilenumformung $A \to A'$ unverändert bleibt (also $V = V'$, wenn V' aus U' durch Ersetzen der i-ten durch die j-te Zeile von A' entsteht).

Insgesamt folgt aus (1)–(3), dass, falls A' aus A durch Zeilenumformungen vom Typ II entstanden ist, gilt: "d teilt alle $(l \times l)$-Unterdeterminanten von A" $\Leftrightarrow$ "d teilt alle $(l \times l)$-Unterdeterminanten von A'". Da man eine Zeilenvertauschung vom Typ III durch geeignete Hintereinanderausführungen von Umformungen des Typs I mit Multiplikation einer Zeile mit -1 und des Typs II darstellen kann, und die Multiplikation einer Zeile mit -1 dann und genau dann das Vorzeichen der Determinante ändert, wenn die betreffende Zeile in U und U' enthalten ist, folgt ebenfalls die Äquivalenz "d teilt alle

$(l \times l)$-Unterdeterminanten von A" $\Leftrightarrow$ "d teilt alle $(l \times l)$-Unterdeterminanten von A'". Dasselbe gilt für Spaltenumformungen.

(ii) Nach (i) können wir ohne Einschränkung annehmen, dass die Matrix Diagonalgestalt C gemäß Satz 9.34 hat. Angenommen, für eine Untermatrix U von C werden die Zeilen $i_1 < \cdots < i_l$ und die Spalten $j_1 < \cdots < j_l$ ausgewählt. Sind diese Zeilen und Spalten gleich ($i_k = j_k \; \forall\, k$), so gilt mit $c_1|c_2|\ldots|c_r$ aus Satz 9.31 und (i)

$$c_1 \cdots c_l \,|\, \underbrace{\det(U)}_{=c_{i_1}\cdots c_{i_l}} \quad \text{sowie} \quad \bar{d}_l(A)|c_1 \cdots c_l, \tag{9.3}$$

wobei $c_k := 0$ für $k > r$ (r aus Satz 9.31). Gilt $\{i_1, \ldots, i_l\} \neq \{j_1, \ldots, j_l\}$, so folgt aus der Implikation $U_{i_m j_k} \neq 0 \Rightarrow i_m = j_k$ (da C diagonal), dass es höchstens $(l-1)$-mal einen Nicht-Null-Eintrag in der Matrix U gibt, womit für mindestens eine Zeile alle Einträge Null sind $\Rightarrow \; \det(U) = 0$. Aus (9.3) folgt damit $\bar{d}_l(A) \mathrel{\widehat{=}} c_1 \cdots c_l$. □

Beweis der Eindeutigkeit in Satz 9.34. Wegen $c_1 \mathrel{\widehat{=}} \bar{d}_1(A)$ und $\bar{d}_l(A) \mathrel{\widehat{=}} c_l \cdot \bar{d}_{l-1}(A)$ sind r eindeutig und alle c_i eindeutig bis auf Assoziiertheit. □

Die Elementarteiler $c_1, \ldots, c_r$ einer $M(n \times n, R)$-Matrix können wir nun entweder mithilfe der Gaußdiagonalisierung für euklidische Ringe (oft recht aufwendig) oder mithilfe der Determinantenteiler bestimmen.

Beispiel 9.37. *Mit* $R = \mathbb{R}[t]$ *und* $A = \begin{pmatrix} t & -1 & -3 \\ -3 & t-1 & 4 \\ 2 & -1 & t-5 \end{pmatrix} \in M(3 \times 3, \mathbb{R}[t])$ *gilt*

- $\bar{d}_1(A) \in GGT(-1, \ldots)$, *d.h.* $\bar{d}_1(A) \mathrel{\widehat{=}} 1$,
- $\bar{d}_2(A) \in GGT(\underbrace{-4+3(t-1)}_{\det\left(\begin{smallmatrix} -1 & -3 \\ t-1 & 4 \end{smallmatrix}\right)}, \underbrace{3-2(t-1)}_{\det\left(\begin{smallmatrix} -3 & t-1 \\ 2 & -1 \end{smallmatrix}\right)}, \ldots) = GGT(\underbrace{2t-7, -2t+5}_{\textit{teilerfremd}}, \ldots)$, *d.h.* $\bar{d}_2(A) \mathrel{\widehat{=}} 1$,
- $\bar{d}_3(A) \mathrel{\widehat{=}} \det A = \cdots = (t-2)^3$.

$\overset{\text{Satz 9.36}}{\Longrightarrow}$ $c_1 \mathrel{\widehat{=}} 1$, $c_2 \mathrel{\widehat{=}} 1$, $c_3 \mathrel{\widehat{=}} (t-2)^2$.

10 Normalformen von Endomorphismen

10.1 Satz von Frobenius und Invariantenteilersatz

Vorbetrachtung 10.1. *Wir wollen schließlich Ähnlichkeit von Matrizen in* $M(n \times n, K)$ *untersuchen. Damit führen wir insbesondere die Diskussionen aus Kapitel 6 über die Diagonalisierbarkeit fort. Überraschend ist, dass wir dies auf die Äquivalenz von bestimmten Matrizen aus* $M(n \times n, K[t])$ *und damit auf den Elementarteilersatz zurückführen können!*

Zur Erinnerung: Demgegenüber war das Äquivalenzproblem in $M(n \times n, K)$ gelöst nach Satz 4.52:

$$A \sim B \iff \operatorname{Rang}(A) = \operatorname{Rang}(B).$$

Definition 10.2. *Sei $A \in M(n \times n, K)$. Wir bezeichnen $P_A := tE_n - A \in M(n \times n, K[t])$ als die charakteristische Matrix von A.*

Satz 10.3 (Frobenius 1878)**.** *Seien K ein Körper und $A, B \in M(n \times n, K)$. Dann gilt*

$$A \approx B \iff P_A \sim P_B \text{ (über } K[t]).$$

Damit haben wir das schwerere Problem der Ähnlichkeit auf das leichtere Problem der Äquivalenz zurückgeführt – allerdings von Matrizen über dem Polynomring $K[t]$.

Beweis. "$(i) \Rightarrow (ii)$": Sei $A \approx B$, d.h. es existiert $S \in \mathrm{GL}(n, K)$ mit $B = SAS^{-1}$.

$$\Longrightarrow \quad P_B = tE_n - B = \underset{S(tE_n)S^{-1}}{tE_n} - \underset{SAS^{-1}}{B} = S(tE_n - A)S^{-1} = SP_AS^{-1},$$

d.h. $P_A \approx P_B$ und insbesondere $P_A \sim P_B$.

"$(ii) \Rightarrow (i)$": Sei $P_A \sim P_B$, d.h. es existieren $S, T \in \mathrm{GL}(n, K[t])$ mit $P_A = SP_BT^{-1}$, also

$$S(tE_n - B) = (tE_n - A)T. \tag{10.1}$$

Weiter besitzen S und T eindeutige Darstellungen $S = \sum_{i=0}^{m} t^i S_i$, $T = \sum_{i=0}^{m} t^i T_i$ mit $T_i, S_i \in M(n \times n, K)$ für alle $i \in \{0, \ldots, m\}$ und ein geeignetes $m \in \mathbb{N}$. Einsetzen in (10.1) ergibt nun:

$$\sum_{i=0}^{m} \left(t^{i+1} S_i - t^i S_i B\right) = \sum_{i=0}^{m} \left(t^{i+1} T_i - t^i A T_i\right).$$

Da Polynome dann und genau dann identisch sind, wenn ihre Koeffizienten übereinstimmen, impliziert dies $S_{i-1} - S_iB = T_{i-1}AT_i$ für alle $i \in \{1, \ldots, m\}$, $S_m = T_m$ und $S_0B = AT_0$, d.h. mit $S_{-1} := 0$, $T_{-1} := 0$, $S_{m+1} := 0$ und $T_{m+1} := 0$ folgt

$$S_{i-1} - S_iB = T_{i-1} - AT_i \quad \forall\, i \in \{0, \ldots, m+1\}.$$

Multipliziert man nun diese Gleichung von links mit A^i und summiert, ergibt sich

$$\sum_{i=0}^{m+1} (A^i S_{i-1} - A^i SB) = \sum_{i=0}^{m+1} (A^i T_{i-1} - A^{i+1} T_i) \underset{\text{Teleskopsumme}}{=} A^0 T_{-1} - A^{m+2} T_{m+1} = 0.$$

Aber daraus folgt

$$A\sum_{i=0}^{m} A_i S_i = \left(\sum_{i=0}^{m} A^i S_i\right) B,$$

d.h. mit $R = \sum_{i=0}^{m} A^i S_i$ gilt $AR = RB$. Bleibt zu zeigen: R ist invertierbar (denn dann gilt $R^{-1}AR = B$, also $A \approx B$). Nach Voraussetzung ist $S = \sum_{i=0}^{m} t^i S_i$ invertierbar, also existiert $M = \sum_{i=0}^{m} t^i M_i$, $M_i \in M(n \times n, K)$ $\forall i \in \{1, \ldots, m\}$, mit $SM = E_n$ (gegebenenfalls nach Vergrößern kann man erreichen, dass m in der Darstellung von S, T und M dasselbe ist). Wir setzen jetzt

$$N := \sum_{j=0}^{m} B^j M_j \ \in M(n \times n, K)$$

und zeigen, dass $RN = E_n$ ist. Wegen $RB = AR$ gilt

$$RB^j = ARB^{j-1} \underset{\substack{\uparrow \\ RB^{j-1} = ARB^{j-2}}}{=} A^2 RB^{j-2} = \ldots = A^j R$$

und damit

$$\begin{aligned} RN = \sum_{j=0}^{m} RB^j M_j &= \sum_{j=0}^{m} A^j R M_j \\ &= \sum_{j=0}^{m} A^j \left(\sum_{i=0}^{m} A^i S_i\right) M_j \\ &= \sum_{i,j=0}^{m} A^{j+i} S_i M_j = S_0 M_0 + \sum_{k=1}^{2m} A^k \left(\sum_{i+j=k} S_i M_j\right). \end{aligned} \tag{10.2}$$

Wegen

$$E_n = \underbrace{\left(\sum_{i=0}^{m} t^i S_i\right)}_{S} \cdot \underbrace{\left(\sum_{j=0}^{m} t^j M_j\right)}_{M} = S_0 M_0 + \sum_{k=1}^{2m} \left(\sum_{i+j=k} S_i M_j\right) t^k$$

gilt aber

$$S_0 M_0 = E_n \text{ und } \sum_{i+j=k} S_i M_j = 0 \ \forall k \geq 1,$$

da E_n keine Unbestimmte t enthält. Einsetzen in (10.2) ergibt $RN = S_0 M_0 = E_n$, also $R \in GL(n, K)$ nach Satz 4.39 und $B = R^{-1}AR$. □

Wir fassen die bisherigen Ergebnisse in den beiden folgenden Sätzen noch einmal zusammen.

Satz 10.4. *Sei $A \in M(n \times n, K)$. Dann gilt:*

(i) *Es gibt eindeutig bestimmte, normierte Polynome $c_1(A), \ldots, c_n(A) \in K[t]$, so dass*

$$P_A \sim \begin{pmatrix} c_1(A) & & 0 \\ & \ddots & \\ 0 & & c_n(A) \end{pmatrix} \text{ und } c_1(A) \mid c_2(A) \mid \ldots \mid c_n(A).$$

Die Polynome $c_1(A), \ldots, c_n(A)$ heißen Invariantenteiler von A.

(ii) *Es gibt eindeutig bestimmte, normierte Polynome $d_1(A), \ldots, d_n(A) \in K[t]$ mit $d_l(A) \mathrel{\widehat{=}} \bar{d}_l(P_A)$. Es gelten $d_n(A) = \chi_A$ sowie*

$$d_l(A) = c_1(A) \cdots c_l(A) \quad \textit{für alle } l = 1, \ldots, n.$$

[Achtung: Im Sinne von Definition 9.35 sind $d_1(A), \ldots, d_n(A)$ Determinantenteiler von P_A, werden aber oft ebenfalls als Determinantenteiler von A bezeichnet.]

Beweis. (i) folgt unmittelbar aus Satz 9.34 – wenn man zusätzlich verlangt, dass die Elementarteiler von P_A normiert sind, sind sie eindeutig, und wegen $\bar{d}_n(P_A) \mathrel{\widehat{=}} d_n(A) = \chi_A \neq 0$ ist $r = n$ nach Satz 9.36. (ii) ist eine Konsequenz aus (i) und Satz 9.36. □

Satz 10.5 (Invariantenteilersatz)**.** *Seien $A, B \in M(n \times n, K)$. Dann sind äquivalent:*

(i) $A \approx B$;

(ii) $c_l(A) = c_l(B)$ *für alle* $l \in \{1, \ldots, n\}$;

(iii) $d_l(A) = d_l(B)$ *für alle* $l \in \{1, \ldots, n\}$.

Beweis. Die Aussage folgt aus Satz 10.3 (Satz von Frobenius) und Satz 10.4. □

Beispiel 10.6. *Sei* $A = \begin{pmatrix} 0 & 1 & 3 \\ 3 & 1 & -4 \\ -2 & 1 & 5 \end{pmatrix}$, *also P_A wie in Beispiel 9.37, womit*

- $d_1(A) = d_2(A) = 1$ *und* $d_3(A) = (t-2)^3$.

Sei

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ -1 & 1 & 4 \end{pmatrix} \quad \textit{d.h.} \quad P_B = \begin{pmatrix} t-1 & -1 & -2 \\ -1 & t-1 & 2 \\ 1 & -1 & t-4 \end{pmatrix}.$$

Da hier $d_2(B)$ schwieriger zu ermitteln ist, bestimmen wir die Invariantenteiler mittels Gaußdiagonalisierung gemäß Satz 9.31 (Übungsaufgabe). Wir erhalten

$$P_B \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & t-2 & 0 \\ 0 & 0 & (t-2)^2 \end{pmatrix} \implies c_1(B) = 1,\ c_2(B) = (t-2),\ c_3(B) = (t-2)^2,\ \textit{womit}$$

- $d_1(B) = 1$, $d_2(B) = (t-2)$, $d_3(B) = (t-2)^3$.

$\overset{\text{Satz 10.5}}{\implies} A \not\approx B$.

10.2 Frobenius-, Weierstraß- und Jordan-Normalformen

Wir suchen möglichst einfache Matrizen B, die zu gegebenem A, also zu gegebenen Invariantenteilern bzw. Determinantenteilern, passen. Wir erreichen dies in drei Schritten und werden dabei immer detaillierter. Die resultierenden Matrizen B heißen dann Normalformen.

Lemma und Definition 10.7. *Sei $g = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in K[t], n \geq 1$. Die Matrix*

$$B_g := \begin{pmatrix} 0 & \cdots & 0 & -a_0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & -a_{n-2} \\ 0 & & 1 & -a_{n-1} \end{pmatrix}$$

heißt Begleitmatrix zu g. Es gelten folgende Aussagen:

(i) $c_1(B_g) = \ldots = c_{n-1}(B_g) = 1$ *und* $c_n(B_g) = d_n(B_g) = \chi_{B_g} = g$ *d.h.*

$$P_{B_g} \sim \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ 0 & \cdots & 0 & g \end{pmatrix}.$$

(ii) $\mu_{B_g} = \chi_{B_g} (= g)$.

Beweis. (i) Es gilt

$$P_{B_g} = tE_n - B_g = \begin{pmatrix} t & & 0 & a_0 \\ -1 & \ddots & & \vdots \\ & \ddots & t & a_{n-2} \\ 0 & & -1 & t + a_{n-1} \end{pmatrix}.$$

Laplace-Entwicklung von $\det(P_{B_g})$ nach der letzten Spalte gemäß Satz 5.23 ergibt

$$\chi_{B_g} = d_n(B_g) = \det(P_{B_g}) = \sum_{i=1}^{n} (-1)^{i+n} a_{i-1} \det((B_g)'_{in}) = \ldots = g.$$

Streichen der ersten Zeile und letzten Spalte führt zur Matrix

$$C := \begin{pmatrix} -1 & t & & 0 \\ & \ddots & \ddots & \\ & & \ddots & t \\ 0 & & & -1 \end{pmatrix} \overset{\substack{\text{(D8)} \\ \text{Satz } 5.4}}{\Longrightarrow} \det(C) = (-1)^{n-1}.$$

Wegen $d_{n-1}(B_g) | \det(C)$ folgt $d_1(B_g) = \ldots = d_{n-1}(B_g) = 1$, womit (i) bewiesen ist.

(ii) Die Vektoren

$$e_1, \underbrace{B_g e_1}_{e_2}, \underbrace{B_g^2 e_1}_{e_3}, \ldots, \underbrace{B_g^{n-1} e_1}_{e_n}$$

sind linear unabhängig. Folglich existiert kein $h \in K[t]$ mit $\deg(h) \leq n-1$ und $h(B_g) = 0$. Damit ist $\mu_{B_g} = \chi_{B_g}$, da $\chi_{B_g}(B_g) = 0$ nach Satz 6.31 (Satz von Cayley-Hamilton). □

Lemma 10.8. *Seien $g_1, \ldots, g_r \in K[t]$ normierte Polynom mit $g_1 \mid g_2 \mid \ldots \mid g_r$, $\deg(g_i) \geq 1$ $\forall i \in \{1, \ldots, g_r\}$ und $\deg(g_1) + \ldots + \deg(g_r) = n$. Sei ferner*

$$B = B_{g_1,\ldots,g_r} := \begin{pmatrix} B_{g_1} & & 0 \\ & \ddots & \\ 0 & & B_{g_r} \end{pmatrix} \in M(n \times n, K).$$

Dann gilt

$$P_B \sim \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & g_1 & & \\ & 0 & & & \ddots & \\ & & & & & g_r \end{pmatrix},$$

d.h. B besitzt die Invariantenteiler $1, \ldots, 1, g_1, \ldots, g_r$.

Beweis. Es gilt

$$P_B = tE_n - B = \begin{pmatrix} \boxed{tE_{\deg(g_1)} - B_{g_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{tE_{\deg(g_r)} - B_{g_r}} \end{pmatrix}.$$

Nach Lemma 10.7 ist diese Matrix äquivalent zu

$$\begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} \end{pmatrix} \begin{matrix} g_1 \\ \\ \\ \\ g_r \end{matrix} \sim \begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ & & & g_1 & \\ 0 & & & & \ddots \\ & & & & & g_r \end{pmatrix},$$

wobei die Äquivalenz zur rechts stehenden Diagonalmatrix dann über Zeilen- und Spaltenvertauschungen erfolgt. □

Satz und Definition 10.9. *Sei $A \in M(n\times n, K)$. Dann existiert ein eindeutig bestimmtes $r \in \mathbb{N}$ sowie eindeutig bestimmte normierte Polynome $g_1, \dots, g_r$ mit $g_1 \mid g_2 \mid \dots \mid g_r$, $\deg(g_i) \geq 1$ $\forall i \in \{1, \dots, r\}$ und*

$$A \approx B_{g_1,\dots,g_r}.$$

Diese Polynome sind die nicht-konstanten Invariantenteiler von A. Die Matrix $B_{g_1,\dots,g_r}$ heißt Frobenius-Normalform von A.

Beweis. Es seien $1, \dots, 1, g_1, \dots, g_r$ die Invariantenteiler von A mit $\deg(g_i) \geq 1$ für alle $i \in \{1, \dots, r\}$. Die Aussage folgt dann aus Lemma 10.8 und Satz 10.5 (Invariantenteilersatz). □

Beispiel 10.10. *Wir betrachten noch einmal die Matrizen A und B aus Beispiel 10.6.*

(i) Sei

$$A = \begin{pmatrix} 0 & 1 & 3 \\ 3 & 1 & -4 \\ -2 & 1 & 5 \end{pmatrix}.$$

In Beispiel 9.37 hatten wir gezeigt, dass die Invariantenteiler von A die folgenden sind:

$$c_1(A) = c_2(A) = 1, \quad c_3(A) = (t-2)^3 = t^3 - 6t^2 + 12t - 8 =: g_1.$$

Damit folgt nun aus Satz 10.9:

$$A \approx B_{g_1} \overset{10.7}{=} \begin{pmatrix} 0 & 0 & 8 \\ 1 & 0 & -12 \\ 0 & 1 & 6 \end{pmatrix}.$$

(ii) Sei nun

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ -1 & 1 & 4 \end{pmatrix}$$

mit Invariantenteilern

$$c_1(B) = 1, \quad c_2(B) = (t-2) =: g_1, \quad c_3(B) = (t-2)^2 = t^2 - 4t + 4 =: g_2.$$

Per definitionem sind (2) *und* $\left(\begin{smallmatrix} 0 & -4 \\ 1 & 4 \end{smallmatrix}\right)$ *die Begleitmatrizen von* g_1 *und* g_2*, womit*

$$B \overset{\text{Satz 10.9}}{\approx} \left(\begin{array}{c|cc} 2 & 0 & 0 \\ \hline 0 & 0 & -4 \\ 0 & 1 & 4 \end{array}\right).$$

Als Korollar aus Satz 10.9 erhalten wir nun den noch ausstehenden Beweis von $\chi_\varphi | \mu_\varphi^n$ in Bemerkung 6.38!

Satz 10.11. *Sei $A \in M(n \times n, K)$. Dann ist das Minimalpolynom von A gleich dem n-ten Invariantenteiler von A, d.h. $\mu_A = c_n(A)$. Ferner gelten $\mu | \chi_A$ und $\chi_A | \mu_A^n$.*

Beweis. Es seien wie in Satz 10.9 $g_1, \ldots, g_r$ die nicht-konstanten Invariantenteiler von A sowie $B := B_{g_1,\ldots,g_r}$ die Frobenius-Normalform von A. Es gelten $\mu_A \overset{A \approx B}{=} \mu_B$ sowie $g_r = c_n(A)$. Für $g \in K[t]$ gilt weiter

$$g(B) = \begin{pmatrix} g(B_{g_1}) & & 0 \\ & \ddots & \\ 0 & & g(B_{g_r}) \end{pmatrix}. \tag{10.3}$$

Wegen $g_1 | g_2 | \ldots | g_r$ sowie $g_i(B_{g_i}) = 0$ nach Lemma 10.7 (i) und dem Satz von Cayley-Hamilton folgt aus (10.3), dass $g_r(B) = 0$. Gilt andererseits $g(B) = 0$ für ein $g \in K[t]$, so folgt aus (10.3) insbesondere $g(B_r) = 0$ und damit nach Lemma 10.7 (ii) und Satz 6.33 auch $\mu_{B_{g_r}} = g_r \mid g$, d.h. $\mu_A = \mu_B = g_r = c_n(A)$. Ferner gilt

$$\chi_A = \underbrace{\det(P_A)}_{=d_n(A)} = \underbrace{c_1(A) \cdots c_n(A)}_{\text{Invariantenteiler von } A} \mid c_n(A)^n = \mu_A^n.$$

□

Ein Nachteil der Frobenius-Normalform ist, dass bspw. die Diagonalmatrix $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$ keine Frobenius-Form mit Diagonalgestalt hat: Hier ist $c_2(A) = (t-1)(t-2) = t^2 - 3t + 2$ der einzige nicht-konstante Inavariantenteiler, womit $\left(\begin{smallmatrix} 0 & -2 \\ 1 & 3 \end{smallmatrix}\right)$ die zugehörige Frobenius-Form ist. Dieser Missstand motiviert die sogenannte Weierstraß-Normalform, in welcher jede Begleitmatrix $B_{g_1}, \ldots, B_{g_r}$ innerhalb der Frobenius-Normalform $B_{g_1,\ldots,g_r}$ weiter in Unterstrukturen zerlegt wird.

Lemma 10.12. *Ist $g = h_1 \cdots h_k$ ein Produkt von paarweise teilerfremden normierten Polynomen $h_i \in K[t]$ mit $\deg(g_i) \geq 1$ für $i = 1, \ldots, k$, so gilt*

$$B_g \approx \begin{pmatrix} B_{h_1} & & 0 \\ & \ddots & \\ 0 & & B_{h_k} \end{pmatrix} =: C.$$

Beweis. Wegen Satz 10.5 (Invariantenteilersatz) genügt es zu zeigen, dass die Invariantenteiler der beiden $M(n \times n, K)$-Matrizen B_g und C identisch sind. Die charakteristische Matrix

$$P_C = \begin{pmatrix} P_{B_{h_1}} & & 0 \\ & \ddots & \\ 0 & & P_{B_{h_k}} \end{pmatrix}$$

ist nach blockweiser Anwendung von Lemma 10.7 äquivalent zu

$$\begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & & h_1 \end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & & h_r \end{matrix}} \end{pmatrix} \sim \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & h_1 & & \\ & 0 & & & \ddots & \\ & & & & & h_r \end{pmatrix}, \tag{10.4}$$

wobei die Äquivalenz zur rechts stehenden Diagonalmatrix über Zeilen- und Spaltenvertauschungen erfolgt. Da $h_1, \ldots, h_k$ normiert sind, ist auch g normiert. Entsprechend ist nach Lemma 10.7

$$P_{B_g} \sim \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ & & 1 \\ 0 & & g \end{pmatrix}.$$

Es gilt

$$d_n(C) = \det(P_C) = h_1 \cdots h_k = g = \det(P_{B_g}) = d_n(B_g).$$

Ferner ist klar, dass $d_{n-1}(B_g) = 1$ gilt, d.h.

$$d_1(B_g) = \ldots = d_{n-1}(B_g) = 1.$$

Wir zeigen nun, dass $d_{n-1}(C) = 1$ ist und betrachten dafür die rechte Matrix in (10.4)

(Elementarteiler äquivalenter Matrizen stimmen bis auf Assoziiertheit überein). Für jedes $i \in \{1, \ldots, k\}$ taucht das Produkt $\prod_{j \neq i}^{k} h_j$ als $(n-1) \times (n-1)$-Unterdeterminante der rechten Matrix in (10.4) auf. Da die h_j teilerfremd sind und $\overline{d}_{n-1}(P_C)$ ein größter gemeinsame Teiler aller $(n-1)\times(n-1)$-Unterdeterminanten von P_C ist, folgt $\overline{d}_{n-1}(P_C) \mathrel{\widehat{=}} 1$, also $d_{n-1}(C) = 1$. □

Satz und Definition 10.13. *Sei $A \in M(n \times n, K)$. Dann existiert ein eindeutig bestimmtes $m \in \mathbb{N}$ und bis auf die Reihenfolge eindeutige Polynome $h_1, \ldots, h_m \in K[t]$, die Potenzen von irreduziblen normierten Polynomen sind, so dass*

$$A \approx B_{h_1,\ldots,h_m} = \begin{pmatrix} B_{h_1} & & 0 \\ & \ddots & \\ 0 & & B_{h_m} \end{pmatrix}.$$

Die Matrix $B_{h_1,\ldots,h_m}$ heißt Weierstraß-Normalform von A.

Beweis der Existenz. Es seien $g_1, \ldots, g_r$ die nicht-konstanten Invariantenteiler von A. Dann gilt

$$A \approx \begin{pmatrix} B_{g_1} & & 0 \\ & \ddots & \\ 0 & & B_{g_r} \end{pmatrix}$$

nach Satz 10.9. Jeder nicht-konstante Invariantententeiler $g_l \in K[t]$ $(l = 1, \ldots, r)$ besitzt nun eine bis auf Assoziiertheit eindeutige Primfaktorzerlegung $g_l = h_{l1} \cdot \ldots \cdot h_{ls_l}$ mit Potenzen h_{li} von paarweise teilerfremden Primpolynomen aus $K[t]$. Fordert man Normiertheit, so ist diese Zerlegung echt eindeutig. Wendet man nun Lemma 10.12 an, so erhält man die Existenz einer Weierstraß-Normalform. Für den Beweis der Eindeutigkeit (bis auf Reihenfolge) muss gezeigt werden, dass durch die Invariantententeiler von $B_{h_1,\ldots,h_m}$ die Primpolynompotenzen $h_1, \ldots, h_m$ bereits eindeutig festgelegt sind. Wir verweisen an dieser Stelle auf die Literatur. □

Beispiel 10.14. *Sei*

$$A = \begin{pmatrix} 4 & -1 & -2 & 3 \\ -1 & 5 & 2 & -4 \\ 0 & 1 & 3 & -1 \\ -1 & 2 & 2 & 1 \end{pmatrix}.$$

Die Gaußdiagonalisierung über euklidischen Ringen aus Satz 9.31 ergibt angewendet auf P_A die Invariantenteiler

- $c_1(A) = 1,\ c_2(A) = 1,\ c_3(A) = (t-3),\ c_4(A) = (t-3)^2(t-2)$, *d.h.*
- $h_1 = c_3(A) = t - 3,\ h_2 = (t-3)^2 = t^2 - 6t + 9,\ h_3 = t - 2$ *($c_4(A) = h_2 h_3$).*

Nach Satz 10.3 ist $A \approx B_{h_1,h_2,h_3}$ mit der Weierstraß-Normalform

$$B_{h_1,h_2,h_3} = \left(\begin{array}{c|ccc} 3 & 0 & 0 & 0 \\ \hline 0 & 0 & -9 & 0 \\ 0 & 1 & 6 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right).$$

Bemerkung 10.15. *Die nach der Weierstraß-Normalform zu $A \in M(n\times n, K)$ gehörenden Polynome $h_1, \ldots, h_m \in K[t]$ heißen Weierstraßsche Elementarteiler von A über K. Diese sind genau die Primpolynompotenzen, welche in der Primfaktorzerlegung der nicht-konstanten Invariantenteiler von A auftreten. Fasst man daher A als Matrix über einem Erweiterungskörper E von K auf, so besitzt A im Allgemeinen andere Weierstraß-Elementarteiler über E als über K, d.h. die Weierstraß-Normalform hängt von K ab. Demgegenüber ändern sich die Invariantenteiler und damit die Frobenius-Normalform nicht beim Übergang von K zu E.*

Wir wollen schließlich die Jordan-Normalform einer Matrix $A \in M(n\times n, K)$ herleiten. Diese gilt, falls χ_A in Linearfaktoren zerfällt, wenn also K ein Zerfällungskörper von χ_A ist, und insbesondere, wenn K algebraisch abgeschlossen ist. Im Falle eines algebraisch abgeschlossenen Körpers K kommen als Weierstraß-Elementarteiler einer $M(n \times n, K)$-Matrix nur Polynome der Form $(t-\lambda)^l$ $(1 \le l \le n)$ in Frage.

Lemma und Definition 10.16. *Seien $\lambda \in K$ und $g = (t-\lambda)^l \in K[t]$ mit $l \in \mathbb{N}$. Dann gilt*

$$B_g \approx \begin{pmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix} =: \mathcal{J}(\lambda, l) \in M(l \times l, K).$$

Die Matrix $\mathcal{J}(\lambda, l)$ heißt Jordan-Matrix über K.

Beweis. Für $\mathcal{J} = \mathcal{J}(\lambda, l)$ ist

$$P_{\mathcal{J}} = \begin{pmatrix} t-\lambda & & & 0 \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & -1 & t-\lambda \end{pmatrix} \in M(l \times l, K[t]).$$

Satz 5.4 (D8) und Satz 5.15 $\Longrightarrow$ $\overline{d}_l(P_{\mathcal{J}}) \mathrel{\widehat{=}} (t-\lambda)^l$, d.h. $d_l(\mathcal{J}) = (t-\lambda)^l$.

Ferner gilt für die $(l-1)\times(l-1)$-Streichmatrix, die durch Weglassen der ersten Zeile und letzten Spalte aus $P_{\mathcal{J}}$ hervorgeht,

$$\det\begin{pmatrix} -1 & t-\lambda & & 0 \\ & \ddots & \ddots & \\ & & \ddots & t-\lambda \\ 0 & & & -1 \end{pmatrix} \overset{\text{(D8)}}{\underset{\text{Satz 5.4}}{=}} (-1)^{l-1}.$$

$\Rightarrow \overline{d}_{l-1}(P_{\mathcal{J}}) \widehat{=} 1 \Rightarrow \overline{d}_1(P_{\mathcal{J}}) = \ldots = \overline{d}_{l-2}(P_{\mathcal{J}}) \widehat{=} 1$, d.h. $d_1(\mathcal{J}) = \cdots = d_{l-1}(\mathcal{J}) = 1$. Nach Lemma 10.7 gilt damit $d_k(B_g) = d_k(\mathcal{J})$ für alle $k \in \{1,\ldots,l\}$. Die Behauptung folgt nun aus dem Invariantenteilersatz. □

Satz und Definition 10.17. *Sei $A \in M(n\times n, K)$ eine Matrix, deren charakteristisches Polynom χ_A über K in Linearfaktoren zerfällt. Dann gibt es bis auf die Reihenfolge eindeutig bestimmte Jordan-Matrizen $\mathcal{J}_1 = \mathcal{J}(\lambda_1, l_1), \ldots, \mathcal{J}_m = \mathcal{J}(\lambda_m, l_m)$ über K, so dass*

$$A \approx \begin{pmatrix} \mathcal{J}_1 & & 0 \\ & \ddots & \\ 0 & & \mathcal{J}_m \end{pmatrix} =: \mathcal{J}.$$

Hierbei sind die $\lambda_1, \ldots, \lambda_m$ die (nicht notwendigerweise paarweise verschiedenen) Eigenwerte von A. Die Matrix $\mathcal{J}$ heißt Jordansche Normalform von A. Die Matrizen $\mathcal{J}_i$ heißen Jordanblöcke (oder Jordankästchen).

Beweis. Es gilt $\chi_A = d_n(A) = c_1(A)\cdots c_n(A)$. $\Rightarrow c_1(A), \ldots, c_n(A)$ zerfallen in Linearfaktoren. $\Rightarrow$ Alle Weierstraß-Elementarteiler $h_1, \ldots, h_m$ sind Potenzen von linearen Polynomen $h_i = (t-\lambda_i)^{l_i}$ für ein $\lambda_i \in K$ und ein $l_i \in \mathbb{N}$. Wegen $h_1 \cdots h_m = c_1(A)\cdots c_n(A) = d_n(A) = \chi_A$ sind nach Satz 6.19 (ii) die λ_i gerade die Eigenwerte von A. Setzen wir nun $\mathcal{J}_i := \mathcal{J}(\lambda_i, l_i)$, so git $B_{h_i} \approx \mathcal{J}_i$ für $i = 1, \ldots, m$ nach Lemma 10.16. Zusammen mit Satz 10.13 folgt

$$A \approx \begin{pmatrix} \mathcal{J}_1 & & 0 \\ & \ddots & \\ 0 & & \mathcal{J}_m \end{pmatrix}.$$

Die Eindeutigkeit von $\mathcal{J}_1, \ldots, \mathcal{J}_m$ folgt bis auf die Reihenfolge aus der Eindeutigkeit der Weierstraß-Elementarteiler $h_1, \ldots, h_m$ in der Weierstraß-Normalform. □

Bemerkung 10.18. *Die Jordan-Normalform von A ist eine untere Dreiecksmatrix, auf deren Hauptdiagonale nur Eigenwerte von A stehen. Es ist oft zweckmäßig, Jordanblöcke, welche zum selben Eigenwert gehören, jeweils entlang der Diagonale zu gruppieren. Da man für diese Umordnung dieselben Zeilen- und Spaltenvertauschungen durchführt und $(ZV(i,j))^{-1} = ZV(i,j)$, bleibt die umgeordnete Version ähnlich zur vorherigen.*

Algorithmus (Jordan-Normalform).

Eingabe: $A \in M(N \times n, K)$ mit χ_A, welches in Linearfaktoren zerfällt.

Ausgabe: Jordan-Normalform von A.

(1) Bestimme die nicht-konstanten Invariantenteiler $g_1, \ldots, g_r$ von A.

(2) Bestimme deren Primfaktorzerlegung

$$g_i = (t - \lambda_{i1})^{m_{i1}} (t - \lambda_{i2})^{m_{i2}} \ldots (t - \lambda_{ik_i})^{m_{ik_i}},$$

$i = 1, \ldots, r$.

(3) Erhalte

$$A \approx \begin{pmatrix} \mathcal{J}(\lambda_{11}, m_{11}) & & 0 \\ & \ddots & \\ 0 & & \mathcal{J}(\lambda_{rk_r}, m_{rk_r}) \end{pmatrix}.$$

(4) Gruppiere Jordan-Matrizen zu gleichen Eigenwerten zusammen.

Beispiele 10.19. *(i) Wir betrachten nochmal die (4×4)-Matrix A aus Beispiel 10.14, wo wir die Weierstraß-Elementarteiler $h_1 = t - 3$, $h_2 = (t-3)^2$ und $h_3 = t - 2$ ermittelt hatten. Damit gilt*

$$A \approx B_{h_1,h_2,h_3} \approx \begin{pmatrix} \mathcal{J}(3,1) & & 0 \\ & \mathcal{J}(3,2) & \\ 0 & & \mathcal{J}(2,1) \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

(ii) Für die (3×3)-Matrix A aus Beispiel 10.10 (i) hatten wir die Invariantenteiler $c_1(A) = c_2(A) = 1$ und $c_3(A) = (t-2)^3$ bestimmt. Damit ist $(t-2)^3$ der einzige Weierstraß-Elementarteiler und

$$A \approx \mathcal{J}(2,3) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

(iii) Für die (3×3)-Matrix B aus Beispiel 10.10 (ii) hatten wir die Invariantenteiler $c_1(A) = 1$, $c_2(A) = t - 2$, $c_3(A) = (t-2)^2$ hergeleitet. Damit ist

$$B \approx \begin{pmatrix} \mathcal{J}(2,1) & 0 \\ 0 & \mathcal{J}(2,2) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Bemerkung 10.20. *Eine Matrix $A \in M(n \times n, K)$ ist genau dann diagonalisierbar, wenn die zugehörige Jordan-Normalform existiert (also χ_A über K in Linearfaktoren zerfällt) und Diagonalgestalt hat. Denn es treten genau dann ausschließlich (1×1)-Jordanblöcke auf, wenn der n-te Invariantenteiler $c_n(A)$ nur einfache Nullstellen hat (wegen $c_1(A)|c_2(A)|\ldots|c_n(A)$ gilt dies dann auch für alle nicht-konstanten Invariantenteiler); da $c_n(A) = \mu_A$ nach Satz 10.11 folgt die Äquivalenzaussage dann aus Satz 6.36.*

Literatur

wird noch ergänzt, insbesondere verwendet wurden

Rainer Dahlhaus, Vorlesungsskript Lineare Algebra, Universität Heidelberg.

Gerd Fischer, Lineare Algebra, Springer.

Falko Lorenz, Lineare Algebra I und II, Spektrum.