Vorlesung Lineare Algebra

Dirk Kussin

Institut für Mathematik, Universität Paderborn
E-mail address: dirk@math.upb.de

Hinweis. Für Druckfehler wird keine Haftung übernommen.

Inhaltsverzeichnis

KAPITEL 1

Grundbegriffe

1. Mengen und Abbildungen

1.1 (Mengen). Wir werden hier keine axiomatische Mengentheorie betreiben. Deshalb nehmen wir hier den naiven Standpunkt ein und verstehen unter einer *Menge* eine Zusammenfassung gewisser Objekte, den sogenannten *Elementen* der Menge. Eine Menge X ist also eindeutig durch ihre Elemente festgelegt, und man schreibt $x \in X$, falls x ein Element ist in X, und $x \notin X$, falls x kein Element von X ist.

Beispiele:

(1) Die *leere Menge* ist *die* Menge, die kein Element enthält. Sie wird mit $\emptyset$ bezeichnet.
(2) Die Menge aller Buchstaben in `OTTO` ist

$$X = \{\mathtt{O}, \mathtt{T}, \mathtt{T}, \mathtt{O}\} = \{\mathtt{T}, \mathtt{O}, \mathtt{T}, \mathtt{O}\} = \{\mathtt{O}, \mathtt{T}\}.$$

(3) $\mathbb{N} = \{1, 2, 3, 4, \dots\} = \{x \mid x \text{ ist eine natürliche Zahl}\}$.
(4) $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\} = \{x \mid x = 0 \text{ oder } x \text{ ist eine natürliche Zahl}\}$.
(5) $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} = \{0, \pm 1, \pm 2\} = \{x \mid x \text{ ist eine ganze Zahl}\}$.
(6) $\mathbb{Q} = \{p/q \mid p \in \mathbb{Z},\, q \in \mathbb{N}\} = \{x \mid x \text{ ist eine rationale Zahl}\}$.
(7) $\mathbb{R} = \{x \mid x \text{ ist eine reelle Zahl}\}$.

1.2 (Bildungen von Mengen). Die Mengenaxiome (die wir hier nicht behandeln) lassen bestimmte Prozesse zu, um aus gegeben Mengen weitere zu bilden. Dazu gehören:

(1) Teilmengen. Sei X eine Menge. Für jedes $x \in X$ sei $P(x)$ eine Aussage, von der man überprüfen kann, ob sie für x wahr oder falsch ist. Es ist $Y = \{x \in X \mid P(x) \text{ ist wahr}\}$ eine *Teilmenge* von X. Sie besteht aus denjenigen Elementen x von X, für die $P(x)$ zutrifft. Man schreibt $Y \subseteq X$. Zum Beispiel ist die Menge $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$ eine Teilmenge von $\mathbb{R}$.

Eine Menge Y ist genau dann Teilmenge einer Menge X, wenn für jedes $y \in Y$ gilt, dass $y \in X$ ist.

Zwei Mengen X und Y sind offenbar genau dann gleich, wenn $X \subseteq Y$ und $Y \subseteq X$ gilt.

(2) Vereinigung. Seien A und B Mengen. Dann heißt die Menge

$$A \cup B \overset{def}{=} \{x \mid x \in A \text{ oder } x \in B\}$$

(sprich "A vereinigt B") die Vereinigung oder Vereinigungsmenge von A und B.

Man kann folgende allgemeinere Konstruktion durchführen: Sei X eine Menge und I eine Indexmenge. Für jedes $i \in I$ sei eine Teilmenge $X_i \subseteq X$ gegeben. Dann ist

$$\bigcup_{i \in I} X_i = \{x \in X \mid \text{es gibt ein } i \in I \text{ mit } x \in X_i\}$$

eine Menge (eine Teilmenge von X), die *Vereinigung* der Mengen X_i $(i \in I)$.

(3) Durchschnitt. Seien A und B Mengen. Dann heißt die Menge

$$A \cap B \overset{def}{=} \{x \mid x \in A \text{ und } x \in A\}$$

(sprich "A Durchschnitt B") der Durchschnitt oder die Schnittmenge von A und B.

Allgemeiner: Sei X eine Menge und I eine Indexmenge. Für jedes $i \in I$ sei eine Teilmenge $X_i \subseteq X$ gegeben. Dann ist

$$\bigcap_{i \in I} X_i = \{x \in X \mid \text{für alle } i \in I \text{ gilt } x \in X_i\}$$

eine Menge (eine Teilmenge von X), der *Durchschnitt* der Mengen X_i $(i \in I)$.

Ist $I = \{1, 2, \dots, n\}$ endlich, so schreibt man auch $\bigcup_{i=1}^n X_i = X_1 \cup X_2 \cup \dots \cup X_n$ bzw. $\bigcap_{i=1}^n X_i = X_1 \cap X_2 \cap \dots \cap X_n$.

(4) Differenz. Seien X und Y Mengen. Dann heißt die Menge

$$X \setminus Y = \{x \in X \mid x \notin Y\}$$

die *Differenz* von X und Y. Gilt $Y \subseteq X$, so nennt man $X \setminus Y$ auch das *Komplement* von Y in X.

(5) Kartesisches Produkt. Seien $X_1, \dots, X_n$ Mengen. Dann heißt

$$\prod_{i=1}^n X_i = X_1 \times X_2 \times \dots \times X_n = \{(x_1, \dots, x_n) \mid x_i \in X_i \text{ für jedes } i = 1, \dots, n\}$$

das *(kartesische) Produkt* der Mengen $X_1, \dots, X_n$. Die Elemente $(x_1, \dots, x_n)$ nennt man auch n-Tupel (für $n = 2$ (geordnete) Paare und für $n = 3$ (geordnete) Tripel). Zwei n-Tupel $(x_1, \dots, x_n)$ und $(y_1, \dots, y_n)$ sind genau dann gleich, wenn $x_i = y_i$ für alle $i = 1, \dots, n$ gilt. Sei I eine Indexmenge und X_i eine Menge für jedes $i \in I$. Dann definiert man analog allgemeiner $\prod_{i \in I} X_i = \{(x_i)_{i \in I} \mid x_i \in X_i \text{ für alle } i \in I\}$. Gilt $X_i = X$ für alle $i \in I$, so schreibt man auch X^I statt $\prod_{i \in I} X$.

(6) Potenzmenge. Sei X eine Menge. Dann ist

$$2^X \overset{def}{=} \{Y \mid Y \subseteq X\}$$

eine Menge, die Menge aller Teilmengen von X. Sie heißt die *Potenzmenge* von X.

Bemerkung 1.3. Das *Auswahlaxiom* besagt: Ist I eine nichtleere Indexmenge, und ist X_i eine nichtleere Menge für jedes $i \in I$, so ist das kartesische Produkt $\prod_{i \in I} X_i$ eine *nichtleere* Menge.

1.4 (Abbildungen). Seien X und Y Mengen. Eine *Abbildung* $f\colon X \to Y$ zwischen X und Y besteht aus

(i) Dem *Definitionsbereich* X;
(ii) dem *Wertebereich* Y;
(iii) einer *Zuordnungsvorschrift* $x \mapsto f(x)$, die jedem $x \in X$ ein eindeutig bestimmtes Element $f(x) \in Y$ zuordnet.

Zwei Abbildungen $f\colon X \to Y$ und $f'\colon X' \to Y'$ sind also genau dann gleich, wenn $X = X'$, $Y = Y'$ *und* $f(x) = f'(x)$ gilt für alle $x \in X$. Beispiele:

(1) Sei X eine Menge. Dann heißt die Abbildung $1_X = id_x\colon X \to X$ mit $id_X(x) = x$ für jedes $x \in X$ die *identische Abbildung* auf X.
(2) Ordnet man jeder reellen Zahl x die reelle Zahl $x^2 + 5$ zu, so erhält man die Abbildung (Funktion) $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 + 5$. Da für jedes $x \in \mathbb{R}$ die Zahl $x^2 + 5$ immer positiv (> 0) ist, kann man auch die Abbildung $g\colon \mathbb{R} \to \mathbb{R}_{>0}$, $x \mapsto x^2 + 5$ betrachten. Diese hat denselben Definitionsbereich und dieselbe Zuordnungsvorschrift wie f, aber einen anderen Wertebereich, und daher sind dies zwei verschiedene Abbildungen!

(3) Ordnet man man jedem geordnetem Paar $(x, y) \in \mathbb{R} \times \mathbb{R}$ deren Produkt $xy \in \mathbb{R}$ zu, so erhält man die Abbildung $p\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto xy$ $(= p(x, y))$.
(4) Sei $f\colon X \to Y$ eine Abbildung. Ist $A \subseteq X$, so betrachtet man häufig die *Einschränkung* $f|_A\colon A \to Y$, $a \mapsto f(a)$ für alle $a \in A$. Analog werden auch Einschränkungen des Bildbereichs auf $B \subseteq Y$ mit $f(A) \subseteq B$ betrachtet, vgl. Beispiel oben.

1.5 (Komposition von Abbildungen). Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen (X, Y und Z Mengen). Dann kann man die folgende Abbildung betrachten:

$$g \circ f\colon X \to Z,\ g \circ f(x) \overset{def}{=} g(f(x)) \text{ für alle } x \in X.$$

(Sprich "g nach f".) Man nennt $g \circ f$ die Komposition (Verknüpfung, Verkettung, Hintereinanderschaltung) der Abbildungen g und f.

Bemerkung 1.6. (1) Seien $f\colon X \to Y$, $g\colon Y \to Z$ und $h\colon Z \to W$ Abbildungen. Dann gilt

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

(2) Sei $f\colon X \to Y$ eine Abbildung. Dann gilt $1_Y \circ f = f$ und $f \circ 1_X = f$.

Beweis. (1) Für jedes $x \in X$ gilt:

$$\begin{aligned}[(h \circ g) \circ f](x) &= (h \circ g)(f(x)) \\ &= h(g(f(x))) \\ &= h(g \circ f(x)) \\ &= [h \circ (g \circ f)](x).\end{aligned}$$

(2) Für jedes $x \in X$ gilt $1_Y \circ f(x) = 1_Y(f(x)) = f(x)$ und $f \circ 1_X(x) = f(1_X(x)) = f(x)$. □

1.7 (Bild und Urbild). Sei $f\colon X \to Y$ eine Abbildung. Seien $A \subseteq X$ und $B \subseteq Y$ Teilmengen.

(1) Man nennt

$$f(A) \overset{def}{=} \{f(a) \mid a \in A\} = \{y \in Y \mid \text{es gibt ein } a \in A \text{ mit } y = f(a)\} \subseteq Y$$

das *Bild* von A (unter f).
(2) Man nennt

$$f^{-1}(B) \overset{def}{=} \{x \in X \mid f(x) \in B\} \subseteq X$$

das *Urbild* von B (unter f). Ist $B = \{y\}$ einelementig, so schreibt man auch $f^{-1}(y)$ statt $f^{-1}(\{y\})$.

1.8 (Injektiv, surjektiv, bijektiv). Sei $f\colon X \to Y$ eine Abbildung.

(1) f heißt *injektiv*, wenn für alle x, $x' \in X$ aus $f(x) = f(x')$ stets $x = x'$ folgt. (Oder äquivalent dazu: Für alle x, $x' \in X$ mit $x \neq x'$ gilt $f(x) \neq f(x')$.)
(2) f heißt *surjektiv*, wenn zu jedem $y \in Y$ ein $x \in X$ existiert mit $y = f(x)$.
(3) f heißt *bijektiv*, wenn f injektiv und surjektiv ist.

1.9. Sei $f\colon X \to Y$ eine Abbildung. Eine Abbildung $g\colon Y \to X$ heißt *Umkehrabbildung* von f, falls $g \circ f = 1_X$ und $f \circ g = 1_Y$ gilt.

Bemerkung 1.10. Sei $f\colon X \to Y$ eine Abbildung. Eine Umkehrabbildung zu f ist (falls sie existiert) eindeutig bestimmt.

Beweis. Es seien g, $g'\colon Y \to X$ zwei Umkehrabbildungen zu f. Sei $y \in Y$. Dann gilt

$$\begin{aligned}g(y) &= g(1_Y(y)) = g(f \circ g'(y)) = [g \circ (f \circ g')](y) \\ &= [(g \circ f) \circ g'](y) = 1_X \circ g'(y) \\ &= g'(y).\end{aligned}$$

Es folgt $g' = g$. $\square$

Satz 1.11. *Sei $f\colon X \to Y$ eine Abbildung. Folgende beiden Aussagen sind äquivalent:*

(1) *f ist bijektiv.*
(2) *Zu f existiert eine Umkehrabbildung $g\colon Y \to X$.*

Die Umkehrabbildung zu einer bijektiven Abbildung f wird meist mit f^{-1} bezeichnet.

Beweis. "(1)⇒(2)" Es gelte (1), d. h. f sei bijektiv. Sei $y \in Y$. Da f surjektiv ist, gibt es $x \in X$ mit $y = f(x)$. Ein solches Urbild x von y ist eindeutig bestimmt, denn ist auch $x' \in X$ ein Urbild, also mit $y = f(x')$ und damit $f(x) = f(x')$, so folgt $x = x'$ aus der Injektivität von f. Definiere eine Abbildung $g\colon Y \to X$ durch $g(y) \overset{def}{=} x$. Dann gilt $f \circ g(y) = f(x) = y = 1_Y(y)$, also $f \circ g = 1_Y$ (y war beliebig in Y). Sei umgekehrt $x \in X$. Setze $y \overset{def}{=} f(x)$. Dann ist x ein (und damit das) Urbild von y. Nach Definition von g gilt also $g(y) = x$, und es folgt $g \circ f(x) = g(f(x)) = g(y) = x = 1_X(x)$; es folgt also auch $g \circ f = 1_X$.

"(2)⇒(1)" Es gelte (2), d. h. es gebe zu f eine Umkehrabbildung $g\colon Y \to X$.

f ist injektiv: Seien $x, x' \in X$ mit $f(x) = f(x')$. Dann folgt $x = g(f(x)) = g(f(x')) = x'$.

f ist surjektiv: Sei $y \in Y$. Setze $x \overset{def}{=} g(y)$. Dann gilt $f(x) = f(g(y)) = y$. $\square$

2. Gruppen

2.1. Sei X eine Menge. Eine Abbildung $f\colon X \times X \to X$, $(x, x') \mapsto f(x, x')$ heißt *(innere) Verknüpfung*. Statt $f(x, x')$ benutzt man meist die suggestivere Schreibweise $x * x' = f(x, x')$, mit einem Verknüpfungssymbol $*$. Weitere Symbole sind je nach Kontext in Gebrauch, häufig auch $\cdot$ und $+$; letzteres häufig, wenn die Verknüpfung *kommutativ* ist, d. h. wenn $f(x, x') = f(x', x)$ gilt für alle x, $x' \in X$. Das Paar (X, f) bzw. $(X, *)$ nennt man auch eine *algebraische Struktur*.

Definition 2.2. Eine Menge G mit einer Verknüpfung $G \times G \to G$, $(a, b) \mapsto a * b$ heißt eine *Gruppe*, wenn folgende Eigenschaften erfüllt sind:

(G1) Für alle a, b, $c \in G$ gilt $(a*b)*c = a*(b*c)$ (d. h. die Verknüpfung ist *assoziativ*).
(G2) Es gibt ein Element $e \in G$ mit folgender Eigenschaft: Für jedes $a \in G$ gilt $a * e = a$ und $e * a = a$. (Man nennt e ein *neutrales Element.*)
(G3) Zu jedem $a \in G$ gibt es ein $b \in G$ mit $b * a = e$ und $a * b = e$. (Man nennt b ein zu a *inverses Element.*

Bemerkung 2.3. Sei $(G, *)$ eine Gruppe.

(1) Das neutrale Element e ist eindeutig bestimmt.
(2) Zu jedem $a \in G$ ist das inverse Element b eindeutig bestimmt. (Man schreibt dafür $b = a^{-1}$.)

Beweis. (1) Seien e und e' zwei neutrale Elemente. Dann gilt

$$e' = e' * e = e$$

(die erste Gleichheit, da e neutral ist, die zweite, da e' neutral ist).

(2) Sei $a \in G$, und seien b, $b' \in G$ zwei zu a inverse Elemente. Dann gilt

$$b = b * e = b * (a * b') = (b * a) * b' = e * b' = b'.$$

$\square$

Wird die Verknüpfung mit $+$ bezeichnet, so schreibt man $-a$ statt a^{-1}. ("Additive Schreibweise versus multiplikative Schreibweise.")

Definition 2.4. Sei $(G, *)$ eine Gruppe. Ist $*$ kommutativ, d. h. gilt $a * b = b * a$ für alle $a, b \in G$, so heißt die Gruppe G *abelsch*. (Nach dem norwegischen Mathematiker Niels Henrik Abel (1802–1829).)

Beispiele 2.5. (1) $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ sind abelsche Gruppen.
(2) $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ sind abelsche Gruppen.
(3) Die algebraischen Strukturen $(\mathbb{Q}, \cdot)$ und $(\mathbb{R}, \cdot)$ sind keine Gruppen, denn die Null 0 ist nicht invertierbar: für jedes $x \in \mathbb{R}$ gilt $x \cdot 0 = 0 \neq 1$.
(4) Die algebraische Struktur $(\mathbb{N}, +)$ ist keine Gruppe, denn es gibt kein neutrales Element. (Für jedes $x \in \mathbb{N}$ gilt z. B. $x+2 \neq x$.) Die algebraische Struktur $(\mathbb{N}_0, +)$ hat zwar ein neutrales Element (nämlich 0), ist aber auch keine Gruppe, denn etwa hat $x = 2$ kein inverses Element: Für jedes $y \in \mathbb{N}_0$ gilt $2 + y \neq 0$.
(5) $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe, denn z. B. ist $a = 2$ nicht invertierbar.

Satz 2.6. *Sei X eine Menge. Dann ist $S(X) \stackrel{def}{=} \mathrm{Bij}(X, X)$ die Menge aller bijektiven Abbildungen von X nach X eine Gruppe; die Verknüpfung ist durch Komposition "$\circ$" von Abbildungen gegeben, neutrales Element ist die identische Abbildung 1_X auf X.*

Die Elemente von $S(X)$ heißen auch die *Permutationen* von X. Ist $X = \{1, 2, \dots, n\}$, so schreibt man auch $S_n = S(X)$.

Beweis. Da hier Abbildungen von X in sich selbst betrachtet werden, und da die Komposition bijektiver Abbildungen wieder bijektiv ist, ist die Komposition eine wohldefinierte Verknüpfung auf $S(X)$. Nach Bemerkung 1.6 ist Komposition von Abbildungen assoziativ und 1_X verhält sich als neutrales Element. Aus Satz 1.11 folgt die Existenz von inversen Elementen. □

Bemerkung 2.7. (1) Ist X eine Menge mit mindestens drei Elementen, so ist die Gruppe $S(X)$ nicht abelsch.
(2) Die Gruppe S_n hat $n! = 1 \cdot 2 \cdot \ldots \cdot n$ Elemente.

Im folgenden (und in Zukunft) werden wir die Verknüpfung oft als "Multiplikation $\cdot$" schreiben. Statt $a \cdot b$ werden wir dann häufig einfach ab schreiben. Für "Additionen $+$" und "Komposition $\circ$" gilt diese Regelung nicht.

Satz 2.8. *Sei G eine Gruppe, seien $a, b \in G$.*
(1) *(Kürzungsregel) Wenn es ein $c \in G$ gibt mit $ca = cb$ oder mit $ac = bc$, so ist $a = b$.*
(2) *Es gibt genau ein $x \in G$ mit $ax = b$, nämlich $x = a^{-1}b$. Es gibt genau ein $y \in G$ mit $ya = b$, nämlich $y = ba^{-1}$.*
(3) *Es gilt $(a^{-1})^{-1} = a$.*
(4) *Es gilt $(ab)^{-1} = b^{-1}a^{-1}$.*

Beweis. (1) Aus $ca = cb$ folgt $a = ea = c^{-1}ca = c^{-1}cb = eb = b$, aus $ac = bc$ folgt $a = ae = acc^{-1} = bcc^{-1} = be = b$.
(2) Für $x \stackrel{def}{=} a^{-1}b$ gilt $ax = aa^{-1}b = eb = b$. Ist auch $x' \in G$ mit $ax' = b$, so folgt $x' = ex' = a^{-1}ax' = a^{-1}b$. Die andere Beziehung mit y folgt analog.
(3) Es ist $a^{-1} \cdot (a^{-1})^{-1} = e = a^{-1}a$. Aus (1) (kürzen) folgt $(a^{-1})^{-1} = a$.
(4) Es ist $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}eb = b^{-1}b = e = (ab)^{-1}(ab)$. Aus (2) folgt $b^{-1}a^{-1} = (ab)^{-1}$. □

Bemerkung 2.9. (1) Satz 2.8 (3) und (4) besagen für $(\mathbb{R}, +)$: Für jedes $a \in \mathbb{R}$ ist $-(-a) = a$; für alle $a, b \in \mathbb{R}$ gilt $-(a+b) = (-b) + (-a) = (-a) + (-b)$. Statt $a + (-b)$ schreibt man einfacher $a - b$.
(2) Satz 2.8 (3) und (4) besagen für $(\mathbb{R} \setminus \{0\}, \cdot)$: Für alle $a \in \mathbb{R} \setminus \{0\}$ ist $\frac{1}{1/a} = a$. Für alle $a, b \in \mathbb{R} \setminus \{0\}$ ist $\frac{1}{ab} = \frac{1}{b} \cdot \frac{1}{a} = \frac{1}{a} \cdot \frac{1}{b}$.

(3) Es sei X eine Menge. Für $S(X)$ besagt Satz 2.8 (3) und (4): Für jedes $f \in S(X)$ ist $(f^{-1})^{-1} = f$; für alle $f, g \in S(X)$ ist $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.
(4) Ist G eine nichtabelsche Gruppe, so gibt es $a, b \in G$ mit $ab \neq ba$. Für solche Elemente gilt dann $(ab)^{-1} \neq a^{-1}b^{-1}$. (Übung.)

Definition 2.10. Sei $(G, \cdot)$ eine Gruppe mit neutralem Element e. Eine Teilmenge $U \subseteq G$ heißt *Untergruppe*, falls folgendes gilt:
(U1) $e \in U$.
(U2) Für alle $a, b \in U$ gilt $ab \in U$.
(U3) Für alle $a \in U$ gilt $a^{-1} \in U$.

Es ist offensichtlich, dass U mit der auf U eingeschränkten Verknüpfung "$\cdot$" selbst wieder eine Gruppe ist. Das Assoziativgesetz gilt in G, also dann erst recht in U.

Satz 2.11. *Sei $(G, \cdot)$ eine Gruppe. Eine Teilmenge $U \subseteq G$ ist genau dann eine Untergruppe, wenn folgendes gilt:*
(1) $U \neq \emptyset$.
(2) *Für alle $a, b \in U$ gilt $ab^{-1} \in U$.*

Beweis. "$\Rightarrow$" Sei U eine Untergruppe. Dann ist $e \in U$ nach (U1), also $U \neq \emptyset$. Sind $a, b \in U$, so ist $b^{-1} \in U$ wegen (U3) und dann $ab^{-1} \in U$ wegen (U2).

"$\Leftarrow$" Es gelten die beiden oben genannten Bedingungen. Da $U \neq \emptyset$ gibt es ein $c \in G$. Nach der zweiten Bedingung ist dann aber $e = cc^{-1} \in U$, also gilt (U1). Für jedes $b \in U$ gilt nach der zweiten Bedingung $b^{-1} = eb^{-1} \in U$, also (U3). Sind nun $a, b \in U$ beliebig, so ist dann $ab = a(b^{-1})^{-1} \in U$, also gilt (U2). □

Beispiele 2.12. (1) für jede Gruppe G (mit neutralem Element e) gilt: $\{e\}$ und G sind Untergruppen von G.
(2) Es ist $\mathbb{Z}$ eine Untergruppe von $(\mathbb{Q}, +)$ und auch von $(\mathbb{R}, +)$.
(3) Es ist $(\mathbb{Q} \setminus \{0\}, \cdot)$ eine Untergruppe von $(\mathbb{R} \setminus \{0\}, \cdot)$.

3. Körper

Wir betrachten nun eine weitere algebraische Strukur. Als Vorbild dient uns dabei die Menge der reellen Zahlen mit ihrer Addition und Multiplikation und den entsprechenden Rechenregeln.

Definition 3.1. Es sei K eine Menge, auf der zwei Verknüpfungen gegeben sind, eine "Addition" $+\colon K \times K \to K$ und eine "Multiplikation" $\cdot\colon K \times K \to K$, für die die folgenden Eigenschaften gelten:
(K1) $(K, +)$ ist eine abelsche Gruppe; das neutrale Elemente wird mit 0 ("null") bezeichnet.
(K2) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe; das neutrale Element wird mit 1 ("eins") bezeichnet.
(K3) In K gelten die Distributivgesetz: Für alle $a, b, c \in K$ gilt $a \cdot (b + c) = a \cdot b + a \cdot c$ und $(a + b) \cdot c = a \cdot c + b \cdot c$.

Dann heißt K ein *Körper*.

Bemerkung 3.2. Sei K ein Körper.
(1) Wir schreiben auch hier häufig kürzer ab statt $a \cdot b$.
(2) Um Klammern zu sparen, verwenden bei der Schreibweise die Konvention "Punktrechnung vor Strichrechnung", d. h. $\cdot$ bindet stärker als $+$. Diese Konvention haben wir schon in (K3) benutzt, denn sonst hätten wir dort z. B. $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ schreiben müssen.
(3) Für jedes $a \in K$ gilt $a \cdot 0 = 0 = 0 \cdot a$. (Beweis: $a \cdot 0 = a \cdot (0 + 0) \overset{(K3)}{=} a \cdot 0 + a \cdot 0$, und Kürzen (in der Gruppe $(K, +)$) ergibt $a \cdot 0 = 0$; analog $0 \cdot a = 0$.)

(4) Für alle $a, b \in K$ gilt $a \cdot b = b \cdot a$. (Beweis: Für $a, b \neq 0$ gilt dies nach Defnition (K2); ist $a = 0$ oder $b = 0$, so folgt dies aus der vorherigen Bemerkung.)
(5) Ebenso gilt $(ab)c = a(bc)$ für alle $a, b, c \in K$.
(6) Sind $a, b \in K$, so schreibt man $a - b$ statt $a + (-b)$.
(7) (Die folgende Aussage ist schon implizit in der Formulierung des Axioms (K2) enthalten. Sie würde aber auch unter einer abgeschwächten leicht Annahme mit folgendem Argument gelten.) Für alle $a, b \in K$ gilt: Sind $a \neq 0$ und $b \neq 0$, so ist $ab \neq 0$. (Beweis: Denn aus $ab = 0$ folgt $a = a(bb^{-1}) = (ab)b^{-1} = 0 \cdot b^{-1} \overset{s.o.}{=} 0$, Widerspruch.)

Beispiele 3.3. (1) $\mathbb{Q}$ und $\mathbb{R}$ sind Körper (mit der üblichen Addition und Multiplikation).
(2) $\mathbb{Z}$ ist kein Körper.
(3) Es sei $\mathbb{F}_2$ eine Menge mit zwei Elementen, die mit 0 und 1 bezeichnet seien. Auf $\mathbb{F}_2$ definieren wir zwei Verknüpfungen $+$ und $\cdot$ durch folgende Wertetafeln:

+	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Man prüft leicht nach, dass damit $\mathbb{F}_2$ zu einem Körper wird, mit Nullelement 0 und Einselement 1.

Definition 3.4. Es sei L ein Körper. Eine Teilmenge $K \subseteq L$ heißt *Unterkörper* (oder *Teilkörper*) von L, und L heißt dann auch *Oberkörper* von K, wenn folgendes gilt:
(1) K ist Untergruppe von $(L, +)$.
(2) $K \setminus \{0\}$ ist Untergruppe von $(L \setminus \{0\}, \cdot)$.

Es ist klar, dass ein Unterkörper K von L mit den von L vererbten Verknüpfungen selbst wieder ein Körper ist.

Beispiele 3.5. (1) Es ist $\mathbb{Q}$ ein Unterkörper von $\mathbb{R}$.
(2) Definiere $K \overset{def}{=} \{a + b \cdot \sqrt{2} \mid a, b \in \mathbb{Q}\} \subseteq \mathbb{R}$.
(a) Es gilt $\mathbb{Q} \subseteq K$. Für alle $a, a', b, b' \in \mathbb{Q}$ gilt

$$(a + b \cdot \sqrt{2}) + (a' + b' \cdot \sqrt{2}) = (a + a') + (b + b')\sqrt{2} \in K$$

und

$$-(a + b \cdot \sqrt{2}) = (-a) + (-b) \cdot \sqrt{2} \in K$$

. Also ist K eine Untergruppe von $(\mathbb{R}, +)$.
(b) Für alle $a, a', b, b' \in \mathbb{Q}$ gilt

$$(a + b \cdot \sqrt{2}) \cdot (a' + b' \cdot \sqrt{2}) = (aa' + 2bb') + (ab' + a'b)\sqrt{2} \in K.$$

Sei $x \in K$ mit $x \neq 0$. Dann existieren $a, b \in \mathbb{Q}$ mit $x = a + b\sqrt{2}$. Es ist $a - b\sqrt{2} \neq 0$ (denn andernfalls ist $a = b\sqrt{2}$; ist $b = 0$, so folgt auch $a = 0$ und daher $x = 0$, im Widerspruch zur Annahme $x \neq 0$; also $b \neq 0$, aber dann folgt $\sqrt{2} = \frac{a}{b} \in \mathbb{Q}$ im Widerspruch zu der Tatsache $\sqrt{2} \notin \mathbb{Q}$.)
Es folgt $a^2 - 2b^2 = (a + b\sqrt{2}) \cdot (a - b\sqrt{2}) \neq 0$ und

$$\begin{aligned}\frac{1}{x} &= \frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{(a + b\sqrt{2}) \cdot (a - b\sqrt{2})} \\ &= \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2}\sqrt{2} \in K.\end{aligned}$$

(c) Aus (a) und (b) folgt, dass K ein Unterkörper von $\mathbb{R}$ ist. Es ist K auch ein Oberkörper von $\mathbb{Q}$.

Lemma 3.6. $\sqrt{2} \notin \mathbb{Q}$.

Beweis. Wir führen einen Beweis durch Widerspruch und nehmen an, dass $\sqrt{2} \in \mathbb{Q}$ gilt. Weil außerdem $\sqrt{2} > 0$ ist, gibt es dann a, $b \in \mathbb{N}$ mit $\sqrt{2} = \frac{a}{b}$. Dabei können wir ohne Einschränkung annehmen, dass der Bruch $\frac{a}{b}$ gekürzt vorliegt, d. h. a und b haben außer 1 keinen gemeinsamen Teiler. Obige Gleichung bedeutet $\sqrt{2}b = a$, und Quadrieren ergibt

$$2b^2 = a^2.$$

Also ist a^2 eine gerade Zahl. Dann ist aber a selbst eine gerade Zahl (wäre a ungerade, von der Form $a = 2c + 1$, so wäre $a^2 = 4c^2 + 4c + 1$ ungerade). Wir könnenalso schreiben $a = 2c$ für ein $c \in \mathbb{N}$. Wir erhalten $2b^2 = a^2 = 4c^2$, also $b^2 = 2c^2$. Also ist b^2 und damit auch b eine gerade Zahl. Also haben sowohl a wie auch b die Zahl 2 als gemeinsamen Teiler, Widerspruch zur Gekürztheit des Bruches! Also war die Annahme $\sqrt{2} \in \mathbb{Q}$ falsch, und wir schließen $\sqrt{2} \notin \mathbb{Q}$. $\square$

3.7 (Allgemeines Distributivgesetz). Sei K ein Körper, und seien $a_1, \dots, a_m, b_1, \dots, b_n \in K$. Dann gilt

$$\left(\sum_{i=1}^{m} a_i\right) \cdot \left(\sum_{j=1}^{n} b_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j = \sum_{j=1}^{n}\sum_{i=1}^{m} a_i b_j.$$

4. Die komplexen Zahlen

4.1. Auf $\mathbb{R} \times \mathbb{R} = \{(a, b) \mid a, b \in \mathbb{R}\}$ werden zwei Verknüpfungen definiert: Für alle (a, b), $(a', b') \in \mathbb{R} \times \mathbb{R}$ definiere

$$(a, b) + (a', b') \stackrel{def}{=} (a + a', b + b')$$
$$(a, b) \cdot (a', b') \stackrel{def}{=} (aa' - bb', ab' + a'b)$$

(1) Man rechnet nach, dass $+$ und $\cdot$ innere Verknüpfungen sind, für die jeweils das Assoziativgesetz und das Kommutativitätsgesetz gelten, und das Distributivgesetz (dies folgt, weil die entsprechenden Gesetze in $\mathbb{R}$ gelten). Ferner sieht man, dass $(0, 0)$ neutrales Element bzgl. $+$ und $(1, 0)$ neutrales Element bzgl. $\cdot$ ist. Für jedes $(a, b) \in \mathbb{R} \times \mathbb{R}$ ist $(-a, -b)$ bzgl. $+$ invers zu (a, b). Es ist also $(\mathbb{R} \times \mathbb{R}, +)$ eine abelsche Gruppe.

(2) Sei $(a, b) \in \mathbb{R} \times \mathbb{R}$ mit $(a, b) \neq 0$. Dann gilt $a \neq 0$ oder $b \neq 0$, also ist die reelle Zahl $a^2 + b^2 > 0$. Es ist

$$\left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2}\right) \in \mathbb{R} \times \mathbb{R}$$

und

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2}\right) = (1, 0),$$

d. h. (a, b) ist invertierbar mit

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

Es folgt, dass $\mathbb{R} \times \mathbb{R}$ mit diesen Verknüpfungen ein Körper ist. Wir bezeichnen diesen Körper mit $\mathbb{C}$ und nennen ihn den *Körper der komplexen Zahlen*. Die Elemente von $\mathbb{C}$ heißen *komplexe Zahlen.*

(3) Für jedes $a \in \mathbb{R}$ sei $\phi \colon \mathbb{R} \to \mathbb{C}$ definiert durch $\phi(a) \stackrel{def}{=} (a, 0)$. Offenbar ist ϕ injektiv, und für alle a, $b \in \mathbb{R}$ gilt $\phi(a + b) = \phi(a) + \phi(b)$ und $\phi(ab) = \phi(a)\phi(b)$, ferner $\phi(1) = (1, 0)$.

Wir identifizieren jede reelle Zahl a mit ihrem Bild $\phi(a) = (a, 0) \in \mathbb{C}$. Auf diese Weise können wir also $\mathbb{R}$ mit Hilfe von ϕ als einen Teilkörper von $\mathbb{C}$ auffassen; statt $(a, 0)$ schreiben wir nur noch a.

(4) Wir setzen zur Abkürzung $i \overset{def}{=} (0,1)$. Jedes $z \in \mathbb{C}$ hat eine eindeutige Darstellung der Form
$$a + bi \quad \text{mit } a,\, b \in \mathbb{R}.$$
Beweis: Sei $z = (a,b)$ mit (eindeutigen) a, $b \in \mathbb{R}$. Wir haben nur zu zeigen, dass
$$(a,b) = a + bi$$
gilt. Die rechte Seite ist aber nur eine Abkürzung für
$$(a,0) + (b,0) \cdot (0,1),$$
und nach Definition von Addition und Multiplikation in $\mathbb{C}$ ist dies $(a,0)+(0,b) = (a,b)$.

(5) Es gilt
$$i^2 = (0,1) \cdot (0,1) = (-1,0) = \phi(-1) \overset{(3)}{=} -1.$$
Dies kann man zum Rechnen in $\mathbb{C}$ benutzen: Ist $z = a + bi$ und $z' = a' + b'i$ mit a, a', b, $b' \in \mathbb{R}$, so gilt
$$z + z' = (a + bi) + (a' + b'i) = (a + a') + (b + b')i$$
$$z \cdot z' = (a + bi) \cdot (a' + b'i) = (aa' - bb') + (ab' + a'b)i$$
Ist $z = a + bi \neq 0$, so ist auch $a - bi \neq 0$ und es gilt
$$\frac{1}{z} = \frac{1}{a+bi} = \frac{a-bi}{(a+bi)\cdot(a-bi)} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2} \cdot i.$$

4.2 (Die Gaußsche Ebene (Carl Friedrich Gauß, 1777–1855)). ZEICHNUNG

Definition 4.3. Es sei $z \in \mathbb{C}$, es seien a, $b \in \mathbb{R}$ mit $z = a + bi$.

(1) $\mathrm{Re}(z) \overset{def}{=} a \in \mathbb{R}$ heißt der *Realteil* von z.
(2) $\mathrm{Im}(z) \overset{def}{=} b \in \mathbb{R}$ heißt der *Imaginärteil* von z.
(3) $\overline{z} \overset{def}{=} a - bi$ heißt die zu z *konjugierte komplexe* Zahl.
(4) $|z| \overset{def}{=} \sqrt{a^2 + b^2} \in \mathbb{R}_{\geq 0}$ heißt der *Betrag* von z.

Bemerkung 4.4. Es seien x, $y \in \mathbb{C}$. Es gilt

(1) $\mathrm{Re}(x) = \frac{1}{2}(x + \overline{x})$, $\mathrm{Im}(x) = \frac{1}{2i}(x - \overline{x})$.
(2) $\overline{x} = x \quad \Leftrightarrow \quad \mathrm{Im}(x) = 0 \quad \Leftrightarrow \quad x \in \mathbb{R}$.
(3) $\overline{x} = -x \quad \Leftrightarrow \quad \mathrm{Re}(x) = 0 \quad \Leftrightarrow \quad$ es gibt ein $b \in \mathbb{R}$ mit $x = bi$.
(4) $|x|^2 = x \cdot \overline{x}$, $|\overline{x}| = |x|$, $\overline{(\overline{x})} = x$.
(5) $|x| \geq 0$, und es gilt $|x| = 0 \quad \Leftrightarrow \quad x = 0$.
(6) $\overline{x + y} = \overline{x} + \overline{y}$, $\overline{x \cdot y} = \overline{x} \cdot \overline{y}$.
(7) $|x \cdot y| = |x| \cdot |y|$.

Beweis. Nachrechnen! □

Satz 4.5 (Dreiecksungleichung). *Für alle x, $y \in \mathbb{C}$ gilt*
$$|x + y| \leq |x| + |y|.$$

Beweis. Setze $z \overset{def}{=} x\overline{y} + \overline{x}y$. Es gilt $\overline{z} = z$, also $z \in \mathbb{R}$. Setze $w \overset{def}{=} x\overline{y} - \overline{x}y$. Es gilt $\overline{w} = -w$, also $w^2 = -w\overline{w} = -|w|^2 \leq 0$. Außerdem
$$z^2 = (x\overline{y} + \overline{x}y)^2 = (x\overline{y} - \overline{x}y)^2 + 4x\overline{y})(\overline{x}y) = w^2 + 4|x|^2|y|^2,$$
und damit
$$z \leq |z| = \sqrt{z^2} \leq \sqrt{4|x|^2|y|^2} = 2|x||y|.$$
Es folgt
$$\begin{aligned} |x+y|^2 &= (x+y)(\overline{x} + \overline{y}) = x\overline{x} + (x\overline{y} + \overline{x}y) + y\overline{y} \\ &= |x|^2 + z + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2. \end{aligned}$$

□

Aus den (Ordnungs-) Axiomen für die reellen Zahlen folgt, dass eine reelle Zahl x mit $x < 0$ keine reelle Quadratwurzel besitzt. (Ist $y \in \mathbb{R}$ so ist stets $y^2 \geq 0$.) Dieses Manko wird in $\mathbb{C}$ behoben, Quadratwurzelziehen ist in $\mathbb{C}$ stets möglich.

Satz 4.6 (Existenz von Quadratwurzeln). *Sei $z \in \mathbb{C}$. Dann existiert ein $w \in \mathbb{C}$ mit $w^2 = z$.*

Beweis. Sei $z = a + bi$ mit $a, b \in \mathbb{R}$.

(1) Ist $z = 0$, so setze $w \overset{def}{=} 0$.
(2) Sei $z \neq 0$ und $a > 0$. Dann ist

$$c \overset{def}{=} \sqrt{\frac{1}{2}\left(a + \sqrt{a^2 + b^2}\right)} \in \mathbb{R}$$

und $c > 0$. Setze

$$w \overset{def}{=} c + \frac{b}{2c} i.$$

Dann gilt

$$\begin{aligned} w^2 &= c^2 + bi - \frac{b^2}{4c^2} = \frac{1}{2}(a + \sqrt{a^2 + b^2}) - \frac{b^2}{2(a + \sqrt{a^2 + b^2})} + bi \\ &= \frac{2a^2 + 2a\sqrt{a^2 + b^2}}{2(a + \sqrt{a^2 + b^2})} + bi = a + bi = z. \end{aligned}$$

(3) Sei $z \neq 0$ und $a \leq 0$. Dann ist

$$d \overset{def}{=} \sqrt{\frac{1}{2}\left(-a + \sqrt{a^2 + b^2}\right)} \in \mathbb{R}$$

und $d > 0$. Setze

$$w \overset{def}{=} \frac{b}{2d} + di.$$

Dann verifiziert man analog $w^2 = z$.

□

Wir erwähnen hier nur ein viel besseres Ergebnis:

Satz 4.7 (Fundamentalsatz der Algebra). *Jedes nicht-konstante Polynom mit komplexen Koeffizienten hat eine Nullstelle in $\mathbb{C}$.*

Dieser Satz wird am elegantesten in der Funktionentheorie bewiesen.

5. Der binomische Lehrsatz

Wir wollen im folgenden eine allgemeine Rechenregel beweisen. Diese ist eine gute Illustration für einen Beweis durch vollständige Induktion.

Definition 5.1. (1) Sei $n \in \mathbb{N}_0$. Definiere $n! \overset{def}{=} 1 \cdot 2 \cdot \ldots \cdot n$ ("n Fakultät"). Man setzt $0! = 1$. (Leeres Produkt.)
(2) Seien $n, k \in \mathbb{N}_0$ mit $0 \leq k \leq n$. Man definiert den *Binomialkoeffizienten*

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}.$$

("n über k")

Lemma 5.2. *Sei $n \in \mathbb{N}_0$.*
(1) *Es gilt $\binom{n}{0} = 1 = \binom{n}{n}$.*

(2) *Für jedes* $k \in \mathbb{N}_0$ *mit* $0 \le k \le n$ *gilt* $\binom{n}{k} = \binom{n}{n-k}$.
(3) *Seien* $n, k \in \mathbb{N}_0$ *mit* $1 \le k \le n$. *Dann gilt*

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

Beweis. (1) und (2) sind unmittelbar klar.
(3) Es gilt

$$\begin{aligned}\binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!\cdot(n-k)!} + \frac{n!}{(k-1)!\cdot(n-k+1)!}\\ &= \frac{n!(n-k+1)}{k!\cdot(n-k+1)!} + \frac{n!k}{k!\cdot(n-k+1)!}\\ &= \frac{n!\cdot(n+1)}{k!\cdot(n-k+1)!} = \frac{(n+1)!}{k!\cdot(n-k+1)!}\\ &= \binom{n+1}{k}.\end{aligned}$$

□

5.3 (Pascalsches Dreieck (Blaise Pascal, 1623–1662, französischer Mathematiker, Physiker, Literat und Philosoph.)). ZEICHNUNG

Satz 5.4 (Binomischer Lehrsatz). *Sei* K *ein Körper, und seien* $a, b \in K$. *Für jedes* $n \in \mathbb{N}_0$ *gilt*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Beweis. Wir beweisen diese Aussage durch **vollständige Induktion** nach n.
Induktionsanfang: Für $n = 0$ ist die Aussage richtig, denn

$$(a+b)^0 = 1 = \binom{0}{0} a^0 b^0 = \sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^0.$$

[Wem diese Argumentation wegen des hohen Grades an Trivialität suspekt erscheint, kann den Induktionsanfang auch für $n = 1$ machen, allerdings wird die Aussage dann nur für $n \ge 1$ bewiesen:

$$(a+b)^1 = a + b = \binom{1}{0} a^1 b^0 + \binom{1}{1} a^0 b^1 = \sum_{k=0}^{1} \binom{1}{k} a^{1-k} b^k.$$

(Auch nicht wesentlich schwieriger.)]
Induktionsvoraussetzung: (IV) Sei $n \in \mathbb{N}_0$, für dass die Aussagen bereits bewiesen ist, so das also (für dieses n) gilt

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Induktionsschluß: Wir haben zu zeigen, dass aus der IV folgt, dass die Aussage auch für $n+1$ statt n gilt, d. h. die Aussage "vererbt" sich von n auf $n+1$. In der Tat:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)\cdot(a+b)^n \\
&\overset{IV}{=} (a+b)\cdot\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=0}^{n-1}\binom{n}{k}a^{n-k}b^{k+1} + b^{n+1} \\
&\overset{(*)}{=} a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=1}^{n}\binom{n}{k-1}a^{n-k+1}b^k + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\left[\binom{n}{k}+\binom{n}{k-1}\right]a^{n-k+1}b^k + b^{n+1} \\
&\overset{5.2}{=} a^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}a^{n-k+1}b^k + b^{n+1} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^{(n+1)-k}b^k,
\end{aligned}$$

und dies ist die behauptete Formel für $n+1$. Bei $(*)$ wurde eine Indexverschiebung durchgeführt: k wurde durch $k-1$ ersetzt. □

KAPITEL 2

Matrizen

6. Das Rechnen mit Matrizen

Im folgenden werde mit K immer ein Körper bezeichnet.

Definition 6.1. Seien m, $n \in \mathbb{N}$.

(1) Für jedes $i \in \{1, \dots, m\}$ und jedes $j \in \{1, \dots, n\}$ sei ein Element $\alpha_{ij} \in K$ gegeben. Das rechteckige Schema

$$\begin{aligned} A &= \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \\ &= (\alpha_{ij})_{1 \le i \le m, 1 \le j \le n} = (\alpha_{ij}) \end{aligned}$$

heißt eine *Matrix* über K mit m Zeilen und n Spalten, oder auch eine $m \times n$-Matrix über K.

(2) Die Menge aller $m \times n$-Matrizen über K bezeichnen wir mit $\mathrm{M}(m, n; K)$.

(3) Eine $n \times n$-Matrix nennt man quadratisch. Die Menge aller $n \times n$-Matrizen über K bezeichnen wir mit $\mathrm{M}(n; K)$.

Man beachte: Zwei Matrizen $A = (\alpha_{ij})$ und $B = (\beta_{ij})$ über K sind genau dann gleich, wenn sie dasselbe Format haben und wenn $\alpha_{ij} = \beta_{ij}$ an jeder Stelle i, j gilt.

6.2 (Weitere Bezeichnungen). (1) Sei $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$. Wir schreiben manchmal auch $A[i, j] \stackrel{def}{=} \alpha_{ij}$. Für jedes $i \in \{1, \dots, m\}$ heißt

$$A_{i\bullet} \stackrel{def}{=} (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}) \in \mathrm{M}(1, n; K)$$

die i-te Zeile von A. Für jedes $j \in \{1, \dots, n\}$ sei

$$A_{\bullet j} \stackrel{def}{=} \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$$

die j-te Spalte von A.

(2) Seien

$$b_1 = \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{m1} \end{pmatrix}, \quad \begin{pmatrix} \beta_{12} \\ \beta_{22} \\ \vdots \\ \beta_{m2} \end{pmatrix}, \dots, \begin{pmatrix} \beta_{1n} \\ \beta_{2n} \\ \vdots \\ \beta_{mn} \end{pmatrix} \in \mathrm{M}(m, 1; K).$$

Wir schreiben $[b_1, b_2, \dots, b_n] \stackrel{def}{=} (\beta_{ij}) \in \mathrm{M}(m, n; K)$ für die Matrix mit den Spalten b_1, $b_2, \dots, b_n$.

(3) Es seien $c_1 = (\gamma_{11}, \gamma_{12}, \dots, \gamma_{1n})$, $c_2 = (\gamma_{21}, \gamma_{22}, \dots, \gamma_{2n}), \dots, c_m = (\gamma_{m1}, \gamma_{m2}, \dots, \gamma_{mn}) \in \mathrm{M}(1, n; K)$. Wir schreiben

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \stackrel{def}{=} (\gamma_{ij}) \in \mathrm{M}(m, n; K)$$

für die Matrix mit den Zeilen $c_1, c_2, \dots, c_m$.

6.3. Es seien $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ und $B = (\beta_{ij}) \in \mathrm{M}(m, n; K)$. Es sei $\lambda \in K$.

(1) Man setzt

$$A + B \stackrel{def}{=} (\alpha_{ij} + \beta_{ij}) \in \mathrm{M}(m, n; K).$$

Mit anderen Worten, für jedes $i \in \{1, \dots, m\}$ und jedes $j \in \{1, \dots, n\}$ ist

$$(A + B)[i, j] = A[i, j] + B[i, j].$$

(2) Man setzt

$$\lambda A = \lambda \cdot A \stackrel{def}{=} (\lambda \alpha_{ij}) \in \mathrm{M}(m, n; K).$$

Satz 6.4. (1) *Mit der oben definierten Addition von Matrizen ist $M(m, n; K)$ eine abelsche Gruppe. Ihr neutrales Element ist die Nullmatrix*

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

und zu jedem $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ ist die Matrix $-A \stackrel{def}{=} (-\alpha_{ij}) \in \mathrm{M}(m, n; K)$ invers.

(2) *Für alle $\lambda, \mu \in K$ und alle $A, B \in \mathrm{M}(m, n; K)$ gilt*
  (a) $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$.
  (b) $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$.
  (c) $(\lambda \cdot \mu) \cdot A = \lambda \cdot (\mu \cdot A)$.
  (d) $1 \cdot A = A$.

Beweis. Der Beweis dieser Aussagen ist einfach; man erhält sie durch Rechnen im Körper K. □

Wir definieren jetzt auch die Multiplikation von Matrizen geeigneten Formats.

Definition 6.5. Es seien $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ und $B = (\beta_{ij}) \in \mathrm{M}(n, p; K)$. Man setzt

$$AB = A \cdot B \stackrel{def}{=} \Big(\sum_{k=1}^{n} \alpha_{ik} \cdot \beta_{kj} \Big)_{1 \le i \le m, 1 \le j \le p} \in \mathrm{M}(m, p; K).$$

Mit anderen Worten, für jedes $i \in \{1, \dots, m\}$ und jedes $j \in \{1, \dots, p\}$ ist

$$AB[i, j] = \sum_{k=1}^{n} A[i, k] \cdot B[k, j].$$

Beispiel: (vorgeführt)

Satz 6.6. (1) *Für alle $A \in \mathrm{M}(m, n; K)$, $B \in \mathrm{M}(n, p; K)$ und $C \in \mathrm{M}(p, q; K)$ gilt*

$$(AB)C = A(BC).$$

(Diese Gleichung gilt in $\mathrm{M}(m, q; K)$.)

(2) *Für alle* $A \in \mathrm{M}(m,n;K)$ *und* B, $C \in \mathrm{M}(n,p;K)$ *und gilt*

$$A(B+C) = AB + AC.$$

(3) *Für alle* $B\,C \in \mathrm{M}(m,n;K)$ *und* $A \in \mathrm{M}(n,p;K)$ *gilt*

$$(B+C)A = BA + CA.$$

(4) *Für alle* $\lambda \in K$ *und* $A \in \mathrm{M}(m,n;K)$ *und* $B \in \mathrm{M}(n,p;K)$ *gilt*

$$(\lambda A)B = A(\lambda B) = \lambda(AB).$$

Beweis. (1) Seien $i \in \{1,\dots,m\}$ und $j \in \{1,\dots,q\}$. Dann gilt

$$\begin{aligned}(AB)C[i,j] &= \sum_{k=1}^{p}(AB)[i,k]\cdot C[k,j]\\ &= \sum_{k=1}^{p}\left(\sum_{\ell=1}^{n}A[i,\ell]\cdot B[\ell,k]\right)\cdot C[k,j]\\ &\overset{3.7}{=} \sum_{\ell=1}^{n}A[i,\ell]\cdot\left(\sum_{k=1}^{p}B[\ell,k]\cdot C[k,j]\right)\\ &= \sum_{\ell=1}^{n}A[i,\ell]\cdot(BC)[\ell,j]\\ &= A(BC)[i,j].\end{aligned}$$

(2) Seien $i \in \{1,\dots,m\}$ und $j \in \{1,\dots,p\}$. Dann gilt

$$\begin{aligned}A\cdot(B+C)[i,j] &= \sum_{k=1}^{n}A[i,k]\cdot\big(B[k,j]+C[k,j]\big)\\ &= \sum_{k=1}^{n}A[i,k]\cdot B[k,j] + \sum_{k=1}^{n}A[i,k]\cdot C[k,j]\\ &= AB[i,j] + AC[i,j] = (AB+AC)[i,j]\end{aligned}$$

□

Definition 6.7. Es sei $A = (\alpha_{ij} \in \mathrm{M}(m,n;K)$. Dann heißt die Matrix

$${}^tA \overset{def}{=} \begin{pmatrix}\alpha_{11} & \alpha_{21} & \dots & \alpha_{m1}\\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{m2}\\ \vdots & \vdots & & \vdots\\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{mn}\end{pmatrix} \in \mathrm{M}(n,m;K)$$

die zu A *transponierte Matrix*. Für alle $i \in \{1,\dots,n\}$ und $j \in \{1,\dots,m\}$ ist also ${}^tA[i,j] = A[j,i]$.

Beispiel: (angegeben)

Satz 6.8. (1) *Für jedes* $A \in \mathrm{M}(m,n;K)$ *gilt* ${}^t({}^tA) = A$.
(2) *Für alle* A, $B \in \mathrm{M}(m,n;K)$ *gilt* ${}^t(A+B) = {}^tA + {}^tB$.
(3) *Für alle* $\lambda \in K$ *und* $A \in \mathrm{M}(m,n;K)$ *gilt* ${}^t(\lambda\cdot A) = \lambda\cdot{}^tA$.
(4) *Für alle* $A \in \mathrm{M}(m,n;K)$ *und* $B \in \mathrm{M}(n,p;K)$ *gilt*

$${}^t(A\cdot B) = {}^tB\cdot{}^tA.$$

Beweis. (1) bis (3) sind klar.

(4) Offenbar sind sowohl ${}^t(A \cdot B)$ als auch ${}^tB \cdot {}^tA$ in $\mathrm{M}(p, n; K)$. Für alle $i \in \{1, \dots, p\}$ und $j \in \{1, \dots, m\}$ gilt

$$\begin{aligned} {}^t(A \cdot B)[i, j] &= AB[j, i] = \sum_{k=1}^{n} A[j, k] \cdot B[k, i] \\ &= \sum_{k=1}^{n} {}^tA[k, j] \cdot {}^tB[i, k] = \sum_{k=1}^{n} {}^tB[i, k] \cdot {}^tA[k, j] \\ &= ({}^tB \cdot {}^tA)[i, j]. \end{aligned}$$

□

6.9 (Kronecker-Symbol). (1) Für ganze Zahlen i und j setzt man

$$\delta_{ij} \overset{def}{=} \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Dies heißt das *Kronecker-Symbol* (nach dem Mathematiker Leopold Kronecker, 1823–1891).

(2) Seien $k \in \{1, \dots, m\}$ und $\ell \in \{1, \dots, n\}$. Definiere

$$E_{k\ell} \overset{def}{=} \big(\delta_{ik} \cdot \delta_{j\ell}\big)_{1 \le i \le m,\, 1 \le j \le n} \in \mathrm{M}(m, n; K).$$

Für jedes $i \in \{1, \dots, m\}$ und jedes $j \in \{1, \dots, n\}$ gilt also

$$E_{kl}[i, j] = \begin{cases} 1 & \text{falls } i = k \text{ und } j = \ell \text{ ist}, \\ 0 & \text{sonst}. \end{cases}$$

Die Matrizen E_{11}, $E_{12}, \dots, E_{mn} \in \mathrm{M}(m, n; K)$ heißen die *Basismatrizen* in $M(m, n; K)$.

(3) Die Matrix

$$E_n \overset{def}{=} (\delta_{ij}) \in \mathrm{M}(n; K)$$

heißt die (n-te) *Einheitsmatrix* (über K).

Bemerkung 6.10. (1) Für jedes $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ gilt

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} E_{ij}.$$

(2) Für jedes $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ gilt

$$E_m \cdot A = \big(\sum_{k=1}^{m} \delta_{ik} \cdot \alpha_{kj}\big) = (\alpha_{ij}) = A$$

und

$$A \cdot E_n = \big(\sum_{k=1}^{n} \alpha_{ik} \cdot \delta_{kj}\big) = (\alpha_{ij}) = A.$$

Bemerkung 6.11. Für k, ℓ, r, $s \in \{1, \dots, n\}$ gilt in $\mathrm{M}(n; K)$

$$E_{k\ell} \cdot E_{rs} = \delta_{\ell r} \cdot E_{ks} = \begin{cases} E_{ks} & \ell = r, \\ 0 & \text{sonst}. \end{cases}$$

Beweis. Für $i,\, j \in \{1, \ldots, n\}$ gilt

$$\begin{aligned} E_{k\ell} \cdot E_{rs}[i,j] &= \sum_{p=1}^{n} E_{k\ell}[i,p] \cdot E_{rs}[p,j] \\ &= \sum_{p=1}^{n} \delta_{ki} \cdot \delta_{\ell p} \cdot \delta_{rp} \cdot \delta_{sj} \\ &= \delta_{ki} \cdot \left(\sum_{p=1}^{n} \delta_{\ell p} \cdot \delta_{rp} \right) \cdot \delta_{sj} \\ &= \delta_{ki} \cdot \delta_{\ell r} \cdot \delta_{sj} \\ &= \delta_{\ell r} \cdot E_{ks}[i,j]. \end{aligned}$$

□

Definition 6.12. Eine Matrix $A \in \mathrm{M}(n;K)$ heißt *invertierbar*, wenn es eine Matrix $B \in \mathrm{M}(n;K)$ gibt mit

$$A \cdot B = E_n \quad \text{und} \quad B \cdot A = E_n. \tag{6.1}$$

Bemerkung 6.13. (1) Man zeigt (etwa wie in 1.10 oder in 2.3), dass für jedes $A \in \mathrm{M}(n;K)$ ein $B \in \mathrm{M}(n;K)$, für das (6.1) gilt, eindeutig ist. Man bezeichnet B meist mit A^{-1}, und nennt A^{-1} die zu A *inverse* Matrix.

(2) Es ist

$$\mathrm{GL}(n;K) \stackrel{def}{=} \{A \in \mathrm{M}(n;K) \mid A \text{ ist invertierbar}\}$$

mit der Matrizenmultilikation $\cdot$ eine Gruppe; das neutrale Element ist die einheitsmatrix E_n, und für jedes $A \in \mathrm{GL}(n;K)$ gilt: invers zu A ist die inverse Matrix A^{-1}. (GL steht für "**g**eneral **l**inear group".) All dies ergibt sich aus dem folgenden ((a) und (b)):

(3) Es seien $A,\, B \in \mathrm{GL}(n;K)$. Dann gilt

(a) Es ist $A \cdot B \in \mathrm{GL}(n;K)$ und $(AB)^{-1} = B^{-1}A^{-1}$.
(b) Es ist $A^{-1} \in \mathrm{GL}(n;K)$ und $(A^{-1})^{-1} = A$.
(c) Es ist ${}^tA \in \mathrm{GL}(n;K)$ und $({}^tA)^{-1} = {}^t(A^{-1})$.

(4) Es gelte $n \geq 2$. Dann ist $\mathrm{GL}(n;K)$ nicht abelsch.

(5) Die Gruppe $\mathrm{GL}(1;K)$ ist (bis auf Schreibung ihrer Elemente) gerade die (abelsche) Gruppe $K \setminus \{0\}$.

Beweis. (3) Für (a) und (b) macht man dieselben Rechnungen wie im Beweis von Satz 2.8, die zeigen, dass $B^{-1}A^{-1}$ invers zu AB ist und A invers zu A^{-1}; es folgt dann insbesondere, dass AB und A^{-1} invertierbar sind.

(c) Es gilt

$${}^t(A^{-1}) \cdot {}^tA \stackrel{6.8}{=} {}^t(A \cdot A^{-1}) = {}^tE_n = E_n$$

und

$${}^tA \cdot {}^t(A^{-1}) = {}^t(A^{-1} \cdot A) = {}^tE_n = E_n.$$

(4) Übung. □

7. Der Algorithmus von Gauß

Im folgenden werde mit K immer ein Körper bezeichnet.

Definition 7.1. Eine (quadratische) Matrix $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$ heißt *Diagonalmatrix*, wenn $\alpha_{ij} = 0$ gilt für alle $i, j \in \{1, \dots, n\}$ mit $i \neq j$, d. h. es gilt

$$A = \begin{pmatrix} \alpha_{11} & & & \\ & \alpha_{22} & & \\ & & \ddots & \\ & & & \alpha_{nn} \end{pmatrix} \overset{def}{=} \mathrm{diag}(\alpha_{11}, \alpha_{22}, \dots, \alpha_{nn}).$$

(Hierbei sind die nicht sichtbaren Einträge $= 0$.)

Bemerkung 7.2. Seien $A = \mathrm{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$, $B = (\mathrm{diag}(\beta_1, \beta_2, \dots, \beta_n) \in \mathrm{M}(n;K)$.

(1) Es ist
$$\begin{aligned} A + B &= \mathrm{diag}(\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n), \\ A \cdot B &= \mathrm{diag}(\alpha_1 \cdot \beta_1, \alpha_2 \cdot \beta_2, \dots, \alpha_n \cdot \beta_n) \end{aligned}$$
und ${}^tA = A$.

(2) Es gilt $A \in \mathrm{GL}(n;K)$ genau dann, wenn $\alpha_i \neq 0$ gilt für alle $i \in \{1, \dots, n\}$. In dem Fall gilt
$$A^{-1} = \mathrm{diag}(\alpha_1^{-1}, \alpha_2^{-1}, \dots, \alpha_n^{-1}).$$

Beweis. (1) Für $i \neq j$ hat man

$$AB[i,j] = \sum_{k=1}^{n} A[i,k] \cdot B[k,j] = \sum_{k=1}^{n} 0 = 0,$$

denn es kann nicht sowohl $k = i$ und $k = j$ gelten, also ist in der Summe $A[i,k] = 0$ oder $B[k,j] = 0$. Außerdem ist

$$AB[i,i] = \sum_{k=1}^{n} A[i,k] \cdot B[k,i] = A[i,i] \cdot B[i,i].$$

Der Rest in (1) ist klar.

(2) Sind alle $\alpha_i \neq 0$, so folgt aus der gerade bewiesenen Rechenregel sofort, dass $\mathrm{diag}(\alpha_1^{-1}, \alpha_2^{-1}, \dots, \alpha_n^{-1})$ invers zu A ist; insbesondere ist dann A invertierbar. Umgekehrt gelte, dass $A = \mathrm{diag}(\alpha_1, \dots, \alpha_n)$ invertierbar ist. Sei $A^{-1} = (\widetilde{\alpha}_{ij})$. Dann gilt

$$1 = E_n[i,i] = AA^{-1}[i,i] = \sum_{k=1}^{n} \delta_{ik} \cdot \alpha_i \cdot \widetilde{\alpha}_{ki} = \alpha_i \cdot \widetilde{\alpha}_{ii},$$

und es folgt $\alpha_i \neq 0$ $(i = 1, \dots, n)$. □

Bemerkung 7.3. Sei $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \in \mathrm{M}(m;K)$, seien $A = (\alpha_{ij}) \in \mathrm{M}(m,n;K)$ und $B = (\beta_{ij}) \in \mathrm{M}(n,m;K)$. Dann gilt für alle i, j

$$D \cdot A[i,j] = \lambda_i \cdot A[i,j] \quad \text{und} \quad B \cdot D[i,j] = \lambda_j \cdot B[i,j],$$

d. h.

$$(D \cdot A)_{i\bullet} = \lambda_i \cdot A_{i\bullet} \quad \text{und} \quad (B \cdot D)_{\bullet j} = \lambda_j \cdot A_{\bullet j}.$$

Speziell: Sei $\lambda \in K \setminus \{0\}$. Bezeichne

$$D_k(\lambda) \overset{def}{=} \mathrm{diag}(1, \dots, 1, \lambda, 1, \dots, 1)$$

(mit λ an der k-ten Stelle). Dann ist $D_k(\lambda) \in \mathrm{GL}(m;K)$ mit $D_k(\lambda)^{-1} = D_k(\lambda^{-1})$, und $D_k \cdot A$ erhält man aus A durch Multiplikation der k-ten Zeile mit λ (bzw. $B \cdot D_k(\lambda)$ erhält man aus B durch Multiplikation der k-ten Spalte mit λ.

Bemerkung 7.4. Seien $E_{11}, E_{12}, \dots, E_{mm} \in \mathrm{M}(m;K)$ die Basismatrizen. Es seien k, $\ell \in \{1, \dots, m\}$.

(1) Es sei $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$. Dann ist

$$E_{k\ell} \cdot A[i, j] = \sum_{p=1}^{m} \delta_{ik} \cdot \delta_{p\ell} \cdot \alpha_{pj} = \delta_{ik} \cdot \alpha_{\ell j}.$$

Also ist $(E_{k\ell} \cdot A)_{k\bullet} = A_{\ell\bullet}$, und sonst stehen in $E_{k\ell} \cdot A$ nur Nullen:

$$E_{k\ell} \cdot A = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \\ \alpha_{\ell 1} & \alpha_{\ell 2} & \dots & \alpha_{\ell n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \tag{7.1}$$

wobei die $\alpha_{\ell j}$ in der k-ten Zeile stehen.

(2) Sei $B = (\beta_{ij}) \in \mathrm{M}(n, m; K)$. Analog gilt dann

$$B \cdot E_{k\ell} = \begin{pmatrix} 0 & \dots & 0 & \beta_{1k} & 0 & \dots & 0 \\ 0 & \dots & 0 & \beta_{2k} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \beta_{nk} & 0 & \dots & 0 \end{pmatrix}$$

wobei die β_{ik} in der ℓ-ten Spalte stehen.

7.5 (Vertauschungsmatrizen). Seien $k, \ell \in \{1, \dots, m\}$.

(1) Die Matrizen

$$V_{k\ell} \overset{def}{=} E_m - E_{kk} - E_{\ell\ell} + E_{k\ell} + E_{\ell k}$$

heißen *Vertauschungsmatrizen*.

(2) Sei $A \in \mathrm{M}(m, n; K)$. Aus (7.1) folgt:

$$V_{k\ell} \cdot A = A - E_{kk} \cdot A - E_{\ell\ell} \cdot A + E_{k\ell} \cdot A + E_{\ell k} \cdot A$$

entsteht aus A durch Vertauschen der k-ten und der ℓ-ten Zeile.

(3) Sei $B \in \mathrm{M}(n, m; K)$. Dann gilt:

$$B \cdot V_{k\ell} = B - B \cdot E_{kk} - B \cdot E_{\ell\ell} + B \cdot E_{k\ell} + B \cdot E_{\ell k}$$

entsteht aus B durch Vertauschen der k-ten und der ℓ-ten Spalte.

(4) Es gilt

$$V_{k\ell}^2 = V_{k\ell} \cdot V_{k\ell} = V_{k\ell} \cdot \big(V_{k\ell} \cdot E_m\big) = E_m.$$

Also gilt $V_{k\ell} \in \mathrm{GL}(m; K)$ mit $V_{k\ell}^{-1} = V_{k\ell}$.

7.6 (Additionsmatrizen). Seien $k, \ell \in \{1, \dots, m\}$ mit $k \neq \ell$, sei $\lambda \in K$.

(1) Die Matrizen

$$W_{k\ell}(\lambda) \overset{def}{=} E_m + \lambda E_{k\ell} \in \mathrm{M}(m; K)$$

heißen *Additionsmatrizen*.

(2) Es gilt

$$W_{k\ell}(\lambda) = W_{k\ell}(1) \cdot D_\ell(\lambda).$$

(3) Sei λ, $\mu \in K$. Es gilt

$$\begin{aligned} W_{k\ell}(\lambda) \cdot W_{k\ell}(\mu) &= (E_m + \lambda \cdot E_{k\ell}) \cdot (E_m + \mu \cdot E_{k\ell}) \\ &= E_m + \lambda \cdot E_{k\ell} + \mu \cdot E_{k\ell} + \lambda \cdot \mu \cdot E_{k\ell} \cdot E_{k\ell} \\ &= E_m + (\lambda + \mu) \cdot E_{k\ell} \\ &= W_{k\ell}(\lambda + \mu) = W_{k\ell}(\mu + \lambda) = \dots \\ &= W_{k\ell}(\mu) \cdot W_{k\ell}(\lambda). \end{aligned}$$

(4) Es folgt $W_{k\ell}(-\lambda) \cdot W_{k\ell}(\lambda) = W_{k\ell}(0) = E_m = W_{k\ell}(\lambda) \cdot W_{k\ell}(-\lambda)$, also ist $W_{k\ell}(\lambda) \in \mathrm{GL}(m; K)$ mit $W_{k\ell}(\lambda)^{-1} = W_{k\ell}(-\lambda)$.

(5) Sei $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$. Dann gilt

$$\begin{aligned} W_{kl}(\lambda) \cdot A &= A + \lambda \cdot E_{k\ell} \cdot A \\ &\overset{(7.1)}{=} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{k-1,1} & \alpha_{k-1,2} & \dots & \alpha_{k-1,n} \\ \alpha_{k1} + \lambda \cdot \alpha_{\ell 1} & \alpha_{k2} + \lambda \cdot \alpha_{\ell 2} & \dots & \alpha_{kn} + \lambda \cdot \alpha_{\ell n} \\ \alpha_{k+1,1} & \alpha_{k+1,2} & \dots & \alpha_{k+1,n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}, \end{aligned}$$

d. h. außer der k-ten Zeile bleiben alle Zeilen von A unverändert, nur zur k-ten wird das λ-fache der ℓ-ten Zeile hinzuaddiert.

(6) Sei $B = (\beta_{ij}) \in \mathrm{M}(n, m; K)$. Dann geht $B \cdot W_{k\ell}(\lambda)$ aus B wie folgt hervor: alle Spalten außer der ℓ-ten Spalte von B bleiben unverändert, nur zur ℓ-ten Spalte wird das λ-fache der k-ten Spalte hinzuaddiert.

Definition 7.7. Eine Matrix $P \in \mathrm{M}(m; K)$ heißt *Elementarmatrix*, wenn sie eine der folgenden Matrizen ist:

(1) $D_k(\lambda)$ mit einem $k \in \{1, \dots, m\}$ und einem $\lambda \in K \setminus \{0\}$.
(2) $V_{k\ell}$ mit k, $\ell \in \{1, \dots, m\}$ (Vertauschungsmatrix).
(3) $W_{k\ell}(\lambda)$ mit verschiedenen k, $\ell \in \{1, \dots, m\}$ und mit einem $\lambda \in K$ (Additionsmatrix).

Die vorherige Diskussion hat u. a. folgendes gezeigt:

Satz 7.8. *Sei $P \in \mathrm{M}(m; K)$ eine Elementarmatrix. Dann gilt*

(1) $P \in \mathrm{GL}(m; K)$.
(2) P^{-1} *ist eine Elementarmatrix.*
(3) ${}^tP$ *ist eine Elementarmatrix.*

7.9 (Elementare Zeilenumformungen). (1) Sei A, $A' \in \mathrm{M}(m, n; K)$. Man sagt, dass A' aus A durch eine *elementare Zeilenumformung* hervorgeht, wenn es eine Elementarmatrix $P \in \mathrm{M}(m; K)$ gibt mit $A' = P \cdot A$. (Analog, durch Multiplikation von rechts, definiert man elementare Spaltenumformungen.)

(2) Den drei Typen von Elementarmatrizen entsprechend gibt es drei Typen von elementaren Zeilenumformungen. Schreibe dazu

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

mit den Zeilen $a_1, \dots, a_m \in \mathrm{M}(1, n; K)$.

I. Multiplikation einer k-ten Zeile mit $\lambda \neq 0$:

$$A = \begin{pmatrix} \vdots \\ a_k \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ \lambda \cdot a_k \\ \vdots \end{pmatrix} = A'.$$

II. Vertauschung zweier Zeilen:

$$A = \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_\ell \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_\ell \\ \vdots \\ a_k \\ \vdots \end{pmatrix} = A'.$$

III. Addition des λ-fachen der ℓ-ten Zeile zur k-ten Zeile ($k \neq \ell$, $\lambda \in K$):

$$A = \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_\ell \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_k + \lambda \cdot a_\ell \\ \vdots \\ a_\ell \\ \vdots \end{pmatrix} = A'.$$

(Analoges gilt für elementare Spaltenumformungen.)

Definition 7.10. Sei $T = (\tau_{ij}) \in \mathrm{M}(m, n; K)$. Es sei $r \in \mathbb{N}_0$ mit $r \leq \min\{m, n\}$, und es seien $j(1), j(2), \ldots, j(r) \in \{1, \ldots, n\}$ mit $j(1) < j(2) < \cdots < j(r)$.

(1) T heißt eine schwache (obere) Treppenmatrix, wenn folgendes gilt:
  (a) Für jedes $i \in \{1, \ldots, r\}$ ist
  $$\tau_{i1} = \tau_{i2} = \cdots = \tau_{i,j(i)-1} = 0 \quad \text{und} \quad \tau_{i,j(i)} = 1.$$
  (b) Für jedes $i \in \{r+1, \ldots, m\}$ gilt
  $$\tau_{i1} = \tau_{i2} = \cdots = \tau_{in} = 0.$$
(2) T heißt eine (obere) Treppenmatrix, wenn T eine schwache Treppenmatrix ist (also (a) und (b) gilt) und zusätzlich folgendes gilt:
  (c) Für jedes $i \in \{1, \ldots, r\}$ ist
  $$\tau_{k,j(i)} = 0 \quad \text{für alle} \quad k = 1, \ldots, i-1.$$
(3) Ist T eine schwache Treppenmatrix, so heißt r der *Rang* von T und $j(1), \ldots, j(r)$ heißen die *charakteristischen Spaltenindizes* von T.

Bemerkung 7.11. (1) $T \in \mathrm{M}(m, n; K)$ ist genau dann eine (schwache) Treppenmatrix vom Rang 0, wenn $T = 0$ ist.
(2) Eine schwache Treppenmatrix $T \in \mathrm{M}(5, 10; K)$ vom Rang 4 und mit den charakteristischen Spaltenindizes $j(1) = 2$, $j(2) = 3$, $j(3) = 6$ und $j(4) = 8$ hat die folgende Form:
$$T = \begin{pmatrix} 0 & 1 & * & * & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$
wobei $*$ für beliebige Elemente in K steht.

(3) Eine Treppenmatrix $T \in \mathrm{M}(5, 10; K)$ vom Rang 4 und mit den charakteristischen Spaltenindizes $j(1) = 2$, $j(2) = 3$, $j(3) = 6$ und $j(4) = 8$ hat die folgende Form:

$$T = \begin{pmatrix} 0 & 1 & 0 & * & * & 0 & * & 0 & * & * \\ 0 & 0 & 1 & * & * & 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Satz 7.12. *Sei $A \in \mathrm{M}(m, n; K)$. Dann gibt es ein $P \in \mathrm{M}(m; K)$ mit folgenden Eigenschaften:*

(1) *P ist ein Produkt von endlich vielen Elementarmatrizen; insbesondere $P \in \mathrm{GL}(m; K)$.*
(2) *$T \stackrel{def}{=} P \cdot A$ ist eine Treppenmatrix.*

Die Existenz solcher Matrizen P und T wird durch den sogenannten Gauß-Algorithmus weiter unten gezeigt.

Bemerkung 7.13. Die Aussagen des vorherigen Satzes kann man auch so lesen: Ist $P = P_s \cdot P_{s-1} \cdot \ldots \cdot P_1$ ein Produkt von Elementarmatrizen $P_1, \ldots, P_s \in \mathrm{M}(m; K)$, so ergibt sich die Treppenmatrix T aus A durch sukzessive Anwendung von elementaren Zeilenumformungen, die zu $P_1, P_2, \ldots, P_s$ (in der Reihenfolge!) wie in 7.9 korrespondieren.

Algorithmus A. Eingabe: $A \in \mathrm{M}(m, n; K)$.

Ausgabe: Eine Matrix $P \in \mathrm{M}(m; K)$, welche Produkt von Elementarmatrizen ist, und eine schwache Treppenmatrix $S \in \mathrm{M}(m, n; K)$ mit $S = P \cdot A$, sowie den Rang r von S und die charakteristischen Spaltenindizes $j(1), \ldots, j(r)$.

1. $j := 1$; $k := 1$; $P := E_m$; $S := A$; $r := 0$;
2. Falls $S[i, j] = 0$ für alle $i = k, \ldots, m$, so setze $j := j + 1$. Andernfalls gehe zu 4.
3. Falls $j \leq n$, so gehe zu 2. Andernfalls gebe P, S, r, $j(1), \ldots, j(r)$ aus und STOP.
4. $r := r + 1$; $j(k) := j$;
5. Wähle ein $p \in \{k, \ldots, m\}$ mit $S[p, j] \neq 0$.
6. $P := V_{kp} \cdot P$; $S := V_{kp} \cdot S$;
7. $P := D_k(S[k, j]^{-1}) \cdot P$; $S := D_k(S[k, j]^{-1}) \cdot S$;
8. Für $i = k + 1, \ldots, m$ tue $P := W_{ik}(-S[i, j]) \cdot P$; $S := W_{ik}(-S[i, j]) \cdot S$;
9. $k := k + 1$; $j := j + 1$; falls $j \leq n$, so gehe zu 2. Andernfalls gebe P, S, r, $j(1), \ldots, j(r)$ aus und STOP.

Es ist offensichtlich, dass am Ende P ein Produkt von Elementarmatrizen und S eine schwache Treppenmatrix ist, wobei $S = PA$ gilt.

Beispiel 7.14.

$$A = \begin{pmatrix} 0 & 1 & 2 & 9 & -1 \\ 3 & 4 & 5 & 9 & -1 \\ 6 & 7 & 8 & 9 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{M}(4, 5; \mathbb{R}).$$

Der Algoritmus A liefert der Reihe nach: $S_1 = A$,

$$S_2 = V_{14}S_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 4 & 5 & 9 & -1 \\ 6 & 7 & 8 & 9 & -1 \\ 0 & 1 & 2 & 9 & -1 \end{pmatrix}, \quad S_3 = W_{21}(-3)S_2 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 6 & 7 & 8 & 9 & -1 \\ 0 & 1 & 2 & 9 & -1 \end{pmatrix},$$

$$S_4 = W_{31}(-6)S_3 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 1 & 2 & 3 & -7 \\ 0 & 1 & 2 & 9 & -1 \end{pmatrix}, \quad S_5 = W_{41}(0)S_4 = E_4S_4 = S_4,$$

$$S_6 = W_{32}(-1)S_5 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & -3 & -3 \\ 0 & 1 & 2 & 9 & -1 \end{pmatrix}, \quad S_7 = W_{42}(-1) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & -3 & -3 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix},$$

$$S_8 = D_3(-1/3)S_7 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix}, \quad S_9 = W_{43}(-3) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Man erhält also die schwache Treppenmatrix

$$S = S_9 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

den Rang 3 und die charakteristischen Spaltenindizes $j(1) = 1$, $j(2) = 2$, $j(3) = 4$. Die Matrix

$$\begin{aligned} P &= W_{43}(-3)D_3(-1/3)W_{42}(-1)W_{32}(-1)W_{41}(0)W_{31}(-6)W_{21}(-3)V_{14} \\ &= \begin{pmatrix} 0 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & -3 \\ 0 & 1/3 & -1/3 & 1 \\ 1 & -2 & 1 & 0 \end{pmatrix} \end{aligned}$$

ist ein Produkt von Elementarmatrizen und es gilt $S = PA$.

Interessiert man sich nur für die schwache Treppenmatrix S (und nicht auch für P), so braucht man in den obigen Schritten die jeweiligen Elementarmatrizen (bzw. elementaren Zeilenumformungen) nicht jedesmal mitzunotieren.

Algorithmus B. Eingabe: Ein schwache Treppenmatrix $S \in \mathrm{M}(m, n; K)$ sowie deren Rang r und charakteritischen Spaltenindizes $j(1), \ldots, j(r)$.

Ausgabe: Eine Treppenmatrix $T \in \mathrm{M}(m, n; K)$ sowie ein $Q \in \mathrm{M}(m; K)$, welche Produkt von Elementarmatrizen ist und mit $T = QS$.

1. $Q := E_m$; $T := S$;
2. Falls $r = 0$ oder $r = 1$, dann gebe Q, T aus und STOP. Andernfalls gehe zu 3.
3. $k := 2$;
4. Für $i = 1, \ldots, k-1$ tue: $Q := W_{ik}(-T[i, j(k)]) \cdot Q$; $T := W_{ik}(-T[i, j(k)]) \cdot T$;
5. $k := k + 1$. Falls $k \leq r$, gehe zu 4. Andernfalls gebe Q, T aus und STOP.

Beispiel 7.15. Sei

$$S = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es ist $r = 3$ und $j(1) = 1$, $j(2) = 2$, $j(3) = 4$. Der Algorithmus B liefert der Reihe nach $T_1 = S$,

$$T_2 = W_{12}(-1)T_1 = \begin{pmatrix} 1 & 0 & -1 & -5 & 5 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \ T_3 = W_{13}(5)T_2 = \begin{pmatrix} 1 & 0 & -1 & 0 & 10 \\ 0 & 1 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T_4 = W_{23}(-6) = \begin{pmatrix} 1 & 0 & -1 & 0 & 10 \\ 0 & 1 & 2 & 0 & -10 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = T.$$

T ist eine Treppenmatrix (mit selbem Rang und charakteristischen Spaltenindizes wie S).

Für

$$Q = W_{23}(-6)W_{13}(5)W_{12}(-1) = \begin{pmatrix} 1 & -1 & 5 & 0 \\ 0 & 1 & -6 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

gilt $T = SQ$.

7.16 (Der Algorithmus von Gauß). (Nach dem Mathematiker Carl Friedrich Gauß, 1777–1855.) Es sei $A \in \mathrm{M}(m, n; K)$. Der Algorithmus liefert eine Treppenmatrix $T \in \mathrm{M}(m, n; K)$ und eine Matrix $P \in \mathrm{M}(m; K)$, die ein Produkt von Elementarmatrizen ist, und für die $T = PA$ gilt. Ferner liefert er den Rang r und die charakteristischen Spaltenindizes $j(1), \ldots, j(r)$ von T.

1. Man wendet auf A den Algorithmus A an und erhält eine schwache Treppenmatrix $S \in \mathrm{M}(m, n; K)$ vom Rang r und mit den charakteristischen Spaltenindizes $j(1), \ldots, j(r)$ sowie eine Matrix $P_1 \in \mathrm{M}(m; K)$, die ein Produkt von Elementarmatrizen ist, und für die $S = P_1 A$ gilt.
2. Im zweiten Schritt wendet man auf S den Algorithmus B an und erhält eine Treppenmatrix T mit selben Rang und selben charakteristischen Spaltenindizes wie S, sowie eine Matrix $P_2 \in \mathrm{M}(m; K)$, die ein Produkt von Elementarmatrizen ist, und für die $T = P_2 S$ gilt. Dann ist $P \overset{def}{=} P_2 P_1$ ein Produkt von Elementarmatrizen, und es gilt $T = P_2 S = P_2 P_1 A = PA$.

Bemerkung 7.17. (1) Satz 7.12 ist damit bewiesen.

(2) Der Algorithmus von Gauß wird auch Eliminationsverfahren von Gauß genannt.

(3) Es gibt Varianten des Gauß-Algorithmus. Oftmals benötigt man nur eine zu A gehörige schwache Treppenmatrix bzw. oftmals nur deren Rang und/oder charakteristischen Spaltenindizes. Dann benötigt man nur den Algorithmus A, wobei man auch noch auf die Berechnung der Matrix P verzichten mag. Statt auf A erst Algorithmus A und dann auf das Ergebnis Algorithmus B anzuwenden, kann man auch Algorithmus B schon innerhalb eines modifizierten Algorithmus A abarbeiten. Man erweitert dafür in 8. den Bereich des Laufindex i, so dass dieser Schritt ersetzt wird durch

8.' Für $i = 1, \ldots, k-1, k+1, \ldots, m$ tue $P := W_{ik}(-S[i,j]) \cdot P$; $S := W_{ik}(-S[i,j]) \cdot S$;

Satz 7.18. *Seien $T, T' \in \mathrm{M}(m, n; K)$ Treppenmatrizen. Es gebe ein $P \in \mathrm{GL}(m; K)$ mit $T' = PT$. Dann gilt $T' = T$.*

Beweis. Induktion nach n. Induktionsanfang: $n = 1$. In $\mathrm{M}(m, 1; K)$ gibt es genau zwei Treppenmatrizen, nämlich $0 = {}^t(0, \ldots, 0)$ und $E_{11} = {}^t(1, 0, \ldots, 0)$. Wäre $T \neq T'$, dann wäre etwa $T = 0$ und $T' = E_{11}$. Aber aus der Beziehung $T' = PT$ mit $P \in \mathrm{GL}(m; K)$ folgt, dass $T = 0$ ist genau dann, wenn $T' = 0$ ist.

Induktionsschluß: Es sei $n \geq 2$, und es sei bereits gezeigt: Sind $S, S' \in \mathrm{M}(m, n-1; K)$ Treppenmatrizen und ist $Q \in \mathrm{GL}(m, K)$ mit $S' = QS$, so ist $S' = S$.

Seien nun $T = (\tau_{ij})$, $T' = (\tau'_{ij}) \in \mathrm{M}(m,n;K)$ Treppenmatrizen und $P \in \mathrm{GL}(m;K)$ mit $T' = PT$. Definiere

$$S \overset{def}{=} (\tau_{ij})_{1\leq i\leq m,\, 1\leq j\leq n-1} \quad \text{und} \quad S' \overset{def}{=} (\tau'_{ij})_{1\leq i\leq m,\, 1\leq j\leq n-1}.$$

Offenbar sind S und S' Treppenmatrizen in $M(m,n-1;K)$, und es gilt $S' = PS$. Nach Induktionsvorausstezung gilt deswegen $S' = S$. Damit gilt also

$$T'_{\bullet j} = T_{\bullet j} \qquad\qquad j = 1,\ldots,n-1.$$

Es ist noch $T'_{\bullet n} = T_{\bullet n}$ zu zeigen.

Sei r der Rang von T, sei r' der Rang von T'. Seien $j(1),\ldots,j(r)$ bzw. $j'(1),\ldots,j'(r')$ die charakteristischen Spaltenindizes von T bzw. T'.

1. Fall: Es gelte $j(r) \leq n-1$. Wegen $S = S' = PS$ gilt für jedes $i \in \{1,\ldots,m\}$ und jedes $k \in \{1,\ldots,r\}$

$$\begin{aligned}
\delta_{ik} &\overset{*}{=} T[i,j(k)] = S[i,j(k)] = PS[i,j(k)] \\
&= \sum_{p=1}^{m} P[i,p]\cdot S[p,j(k)] = \sum_{p=1}^{m} P[i,p]\cdot T[p,j(k)] \\
&\overset{*}{=} \sum_{p=1}^{m} P[i,p]\cdot \delta_{pk} \\
&= P[i,k],
\end{aligned}$$

und daher ergibt sich

$$\begin{aligned}
T'[i,n] &= PT[i,n] = \sum_{k=1}^{m} P[i,k]\cdot T[k,n] \\
&\overset{*}{=} \sum_{k=1}^{r} P[i,k]\cdot T[k,n] \\
&= \sum_{k=1}^{r} \delta_{ik}\cdot T[k,n] \\
&= T[i,n].
\end{aligned}$$

(Bei $*$ wurde jeweils die Definition einer Treppenmatrix benutzt.) Also gilt $T'_{\bullet n} = T_{\bullet n}$ und damit $T' = T$.

2. Fall: Es gelte $j'(r') \leq n-1$. Wegen $T = P^{-1}T'$ schließt man wie im 1. Fall, dass $T = T'$ gilt.

3. Fall: $j(r) = n = j'(r')$. Dann folgt, dass $r-1$ der Rang von S und $r'-1$ der Rang von S' ist. Da $S' = S$ gilt, folgt $r'-1 = r-1$, also $r' = r$. Dann folgt aber (nach Definition einer Treppenmatrix), dass sowohl $T'_{\bullet n} = E_{r1} \in \mathrm{M}(n,1;K)$ als auch $T_{\bullet n} = E_{r1} \in \mathrm{M}(n,1;K)$ gilt. Es folgt $T' = T$. □

Folgerung 7.19. *Sei $A \in \mathrm{M}(m,n;K)$. Seien P, $P' \in \mathrm{GL}(m;K)$, so dass $T \overset{def}{=} PA$ und $T' \overset{def}{=} P'A$ Treppenmatrizen sind. Dann gilt $T' = T$. Insbesondere ist die im Satz 7.12 vorkommende Treppenmatrix T durch A eindeutig bestimmt.*

Bezeichnung: T heißt *die zu A gehörige Treppenmatrix.*

Beweis. Es gilt $T' = P'A = P'P^{-1}PA = (P'P^{-1})T$. Da $P'P^{-1} \in \mathrm{GL}(m;K)$ ist, folgt aus dem vorherigen Satz $T' = T$. □

Definition 7.20. Sei $A \in \mathrm{M}(m,n;K)$. Sei T die zu A gehörige Treppenmatrix. Der *Rang* von A ist definiert als der Rang von T, also

$$\mathrm{rang}(A) = \#\{i \mid 1 \leq i \leq m \text{ mit } T_{i\bullet} \neq (0,0,\ldots,0)\}.$$

Bemerkung 7.21. Sei $A \in \mathrm{M}(m,n;K)$ mit zugehöriger Treppenmatrix T. Dann gilt:

(1) Es gilt $0 \le \mathrm{rang}(A) = \mathrm{rang}(T) \le \min\{m,n\}$.
(2) $\mathrm{rang}(A) = 0 \;\Leftrightarrow\; \mathrm{rang}(T) = 0 \;\Leftrightarrow\; T = 0 \;\Leftrightarrow\; A = 0$.
(3) Für jedes $B \in \mathrm{GL}(m;K)$ gilt $\mathrm{rang}(BA) = \mathrm{rang}(A)$.

Beweis. (3) Es gibt ein $P \in \mathrm{GL}(m;K)$ mit $T = PA$. Es ist $PB^{-1} \in \mathrm{GL}(m;K)$ und $T = (PB^{-1})(BA)$, also ist T auch die zu BA gehörige Treppenmatrix. Also $\mathrm{rang}(BA) = \mathrm{rang}(T) = \mathrm{rang}(A)$. □

Definition 7.22. Seien A, $B \in \mathrm{M}(m,n;K)$. Man nennt A und B *äquivalent*, wenn es $P \in \mathrm{GL}(m;K)$ und $Q \in \mathrm{GL}(n;K)$ gibt mit $B = PAQ$; man schreibt dann $A \sim B$. (Dies definiert eine Äquivalenzrelation auf der Menge $\mathrm{M}(m,n;K)$, vgl. Übungen.)

Satz 7.23. (1) *Sei $A \in \mathrm{M}(m,n;K)$ vom Rang $r = \mathrm{rang}(A)$. Dann gilt*

$$A \sim T_r \overset{def}{=} \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & & \\ & & & & & \end{pmatrix}.$$

(2) *Seien A, $B \in \mathrm{M}(m,n;K)$. Dann gilt*

$$A \sim B \quad \Leftrightarrow \quad \mathrm{rang}(A) = \mathrm{rang}(B).$$

Beweis. Mit den gezeigten Ergebnissen und Methoden dieses Abschnitts ergibt sich dies leicht. Siehe Übungen. □

Satz 7.24. *Sei $A \in \mathrm{M}(n;K)$. Folgende Aussagen sind äquivalent:*

(1) *A ist invertierbar.*
(2) $\mathrm{rang}(A) = n$.
(3) *Die zu A gehörige Treppenmatrix ist die Einheitsmatrix E_n.*
(4) *A ist Produkt von Elementarmatrizen.*

Beweis. (1)⇒(2): Sei $A \in \mathrm{GL}(n;K)$. Es ist $A^{-1} \in \mathrm{GL}(n;K)$ und $A^{-1}A = E_n$. Es ist also E_n die zu A gehörige Treppenmatrix, und sie ist vom Rang n. Es folgt $\mathrm{rang}(A) = n$.

(2)⇒(3): Es gelte $\mathrm{rang}(A) = n$. Es sei $T \in \mathrm{M}(n;K)$ die zu A gehörige Treppenmatrix. Für die charakteristischen Spaltenindizes von T muss dann gelten $j(1) = 1$, $j(2) = 2, \ldots, j(n) = n$, und nach Definition einer Treppenmatrix folgt $T = E_n$.

(3)⇒(4): Es sei E_n die Treppenmatrix zu A. Der Gaußsche Algorithmus liefert Elementarmatrizen $P_1, P_2, \ldots, P_s \in \mathrm{M}(n;K)$, so dass $E_n = (P_s \cdot P_{s-1} \cdot \ldots \cdot P_1) \cdot A$ gilt. Es folgt $A = P_1^{-1} \cdot P_2^{-1} \cdot \ldots \cdot P_s^{-1}$, und die P_i^{-1} sind nach Satz 7.8 Elementarmatrizen.

(4)⇒(1): Es sei $A = P_1 \cdot \ldots \cdot P_s$, wobei alle $P_i \in \mathrm{M}(n;K)$ Elementarmatrizen sind. Nach Satz 7.8 sind alle P_i invertierbar, also ist es auch A als deren Produkt. □

Bemerkung 7.25. Es sei $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$.

(1) Will man feststellen, ob A invertierbar ist, so ermittelt man mit Algorithmus A den Rang von A, denn A ist genau dann invertierbar, wenn $\mathrm{rang}(A) = n$ gilt.
(2) Sei A invertierbar. Dann ist E_n die zugehörige Treppenmatrix. Der Gaußsche Algorithmus liefert Elementarmatrizen $P_1, \ldots, P_s \in \mathrm{GL}(n;K)$ mit $E_n = (P_s \cdot \ldots \cdot P_1) \cdot A$. Es folgt

$$A^{-1} = E_n \cdot A^{-1} = P_s \cdot \ldots \cdot P_1 \cdot A \cdot A^{-1} = P_s \cdot \ldots \cdot P_1,$$

und dies ist eine Möglichkeit zur Berechnung der Inversen.

(3) Wir beschreiben eine andere, praktischere Möglichkeit. Setze

$$\widehat{A} \overset{def}{=} (A \mid E_n) \overset{def}{=} \left(\begin{array}{ccc|ccc} \alpha_{11} & \dots & \alpha_{1n} & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \dots & \alpha_{nn} & 0 & \dots & 1 \end{array}\right) \in \mathrm{M}(n, 2n; K).$$

Sei $T \in \mathrm{M}(n; K)$ die zu A gehörige Treppenmatrix. Es gibt ein $P \in \mathrm{GL}(n; K)$ mit $T = PA$. Dann ist

$$\widehat{T} \overset{def}{=} P \cdot \widehat{A} = (PA \mid PE_n) = (T \mid P) \in \mathrm{M}(n, 2n; K).$$

Gilt $T \neq E_n$, so ist A nicht invertierbar. Gilt $T = E_n$, so ist $\widehat{T} = (E_n \mid A^{-1})$ offenbar die zu $\widehat{A}$ gehörige Treppenmatrix.

Dies liefert ein Rezept zum Feststellen, ob A invertierbar ist und zur Berechung der Inversen im invertierbaren Fall. Dazu berechnet man mit dem Gaußschen Algorithmus die Treppenmatrix $\widehat{T} = (T \mid P)$ von $\widehat{A} = (A \mid E_n)$. Ist $T \neq E_n$, so ist A nicht invertierbar. Ist $T = E_n$, so ist A invertierbar und $A^{-1} = P$.

Beispiel 7.26. Sei

$$A = \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix} \in \mathrm{M}(2, 2; \mathbb{R}).$$

Der Gauß-Algorithmus angewendet auf

$$\widehat{A} = \left(\begin{array}{cc|cc} 2 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right)$$

liefert der Reihe nach

$$\left(\begin{array}{cc|cc} 1 & 1 & 1/2 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right), \left(\begin{array}{cc|cc} 1 & 1 & 1/2 & 0 \\ 0 & 2 & -3/2 & 1 \end{array}\right),$$
$$\left(\begin{array}{cc|cc} 1 & 1 & 1/2 & 0 \\ 0 & 1 & -3/4 & 1/2 \end{array}\right), \left(\begin{array}{cc|cc} 1 & 0 & 5/4 & -1/2 \\ 0 & 1 & -3/4 & 1/2 \end{array}\right).$$

Es folgt, dass A invertierbar ist und

$$A^{-1} = \begin{pmatrix} 5/4 & -1/2 \\ -3/4 & 1/2 \end{pmatrix}.$$

Übung 7.27. (a) Man zeige: Für jedes $A \in \mathrm{M}(m, n; K)$ gilt

$$\mathrm{rang}({}^tA) = \mathrm{rang}(A).$$

(b) Man zeige: Für alle $A \in \mathrm{M}(m, n; K)$ und $B \in \mathrm{M}(n, p; K)$ gilt

$$\mathrm{rang}(A \cdot B) \leq \mathrm{rang}(A).$$

(Hinweis: Man zeige dies zunächst für eine Treppenmatrix A und führe dann den allgemeinen Fall darauf zurück.)

(c) Aus (a) und (b) folgere man: Für alle $A \in \mathrm{M}(m, n; K)$ und $B \in \mathrm{M}(n, p; K)$ gilt

$$\mathrm{rang}(A \cdot B) \leq \mathrm{rang}(B).$$

(d) Sei $A \in \mathrm{M}(n; K)$. Man folgere:
- Gibt es ein $B \in \mathrm{M}(n; K)$ mit $A \cdot B = E_n$, so ist A invertierbar, und es gilt $B = A^{-1}$.
- Gibt es ein $C \in \mathrm{M}(n; K)$ mit $C \cdot A = E_n$, so ist A invertierbar, und es gilt $C = A^{-1}$.

KAPITEL 3

Vektorräume und lineare Abbildungen

8. Vektorräume

Im folgenden sei K immer ein Körper.

Definition 8.1. Sei V eine Menge, auf der zwei Abbildungen (Verknüpfungen) gegeben sind:

$$V \times V \to V,\ (x, y) \mapsto x + y,$$
$$K \times V \to V,\ (\alpha, x) \mapsto \alpha \cdot x.$$

V heißt ein K-*Vektorraum*, oder ein *Vektorraum über* K, wenn folgendes gilt:

(VR1) $(V, +)$ ist eine abelsche Gruppe.
(VR2) Für alle α, $\beta \in K$ und alle x, $y \in V$ gilt
- (a) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$.
- (b) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$.
- (c) $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$.
- (d) $1_K \cdot x = x$.

Bemerkung 8.2. Sei V ein K-Vektorraum.

(1) Das neutrale Element in der abelschen Gruppe $(V, +)$ wird mit $0 = 0_V$ bezeichnet und heißt das *Nullelement* von V. Es ist zu unterscheiden vom Nullelement $0 = 0_K$ von K. Für jedes $x \in V$ wird das Inverse mit $-x$ bezeichnet.
(2) Für α, $\beta \in K$ und x, $y \in V$ schreibt man αx statt $\alpha \cdot x$, $-\alpha x$ statt $-(\alpha x)$, $\alpha x + \beta y$ statt $(\alpha \cdot x) + (\beta \cdot y)$, $\alpha\beta x$ statt $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$.
(3) Für $x_1, x_2, \ldots, x_m \in V$ schreibt man

$$\sum_{i=1}^{m} x_i \overset{def}{=} x_1 + x_2 + \cdots + x_m.$$

Ist $m = 0$, so setzt man $\sum_{i=1}^{m} = 0_V$ (leere Summe).
(4) Für $\alpha \in K$ und $x_1, x_2, \ldots, x_m \in V$ gilt

$$\alpha \cdot \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} \alpha \cdot x_i.$$

(Induktion nach m.)

Bemerkung 8.3. Sei V ein K-Vektorraum, und seien $\alpha \in K$ und $x \in V$.

(1) $\alpha \cdot 0_V = 0_V$ und $0_K \cdot x = 0_V$.
(2) Gilt $\alpha \neq 0_K$ und $x \neq 0_V$, so ist $\alpha \cdot x \neq 0_V$.
(3) $(-\alpha) \cdot x = -\alpha \cdot x \ \ (\overset{def}{=} -(\alpha \cdot x))$.
(4) $(-1_K) \cdot x = -x$.

Beweis. (1) Es gilt

$$\begin{aligned} \alpha \cdot 0_V + 0_V &= \alpha \cdot 0_V = \alpha \cdot (0_V + 0_V) \\ &= \alpha \cdot 0_V + \alpha \cdot 0_V, \\ 0_K \cdot x + 0_V &= 0_K \cdot x = (0_K + 0_K) \cdot x \\ &= 0_K \cdot x + 0_K \cdot x, \end{aligned}$$

und Kürzen 2.8 (1) in der Gruppe $(V, +)$ ergibt $\alpha \cdot 0_V = 0_V$ und $0_K \cdot x = 0_V$.

(2) Es gelte $\alpha \neq 0$ und $\alpha x = 0$. Dann folgt

$$x = 1_K \cdot x = (\alpha^{-1}\alpha)x = \alpha^{-1}(\alpha x) = \alpha^{-1} \cdot 0_V \overset{(1)}{=} 0_V.$$

(3) Es ist

$$\alpha x + (-\alpha)x = (\alpha + (-\alpha))x = 0_K \cdot x \overset{(1)}{=} 0_V = \alpha x + (-\alpha x),$$

und 2.8 (1) liefert $(-\alpha)x = -\alpha x$.

(4) Es ist $(-1_K) \cdot x \overset{(3)}{=} -(1_K \cdot x) = -x$. □

Beispiele 8.4. (1) Seien $m, n \in \mathbb{N}$. Es ist $\mathrm{M}(m, n; K)$ mit den Verküpfungen

$$\mathrm{M}(m, n; K) \times \mathrm{M}(m, n; K) \to \mathrm{M}(m, n; K), \ (A, B) \mapsto A + B$$
$$K \times \mathrm{M}(m, n; K) \to \mathrm{M}(m, n; K), \ (\lambda, A) \mapsto \lambda \cdot A$$

aus 6.3 ein K-Vektorraum, vgl. 6.4.

(2) Speziell ist

$$K^n \overset{def}{=} \mathrm{M}(n, 1; K) = \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \mid \alpha_1, \alpha_2, \ldots, \alpha_n \in K \right\}$$

ein K-Vektorraum. Für

$$a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \in K^n$$

und $\lambda \in K$ ist

$$a + b = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} \quad \text{und} \quad \lambda \cdot a = \begin{pmatrix} \lambda\alpha_1 \\ \lambda\alpha_2 \\ \vdots \\ \lambda\alpha_n \end{pmatrix}.$$

(3) Sei $M \neq \emptyset$ eine Menge. Sei

$$V \overset{def}{=} \mathrm{Abb}(M, K) \overset{def}{=} \{f \mid f\colon M \to K \text{ ist eine Abbildung}\}.$$

Für $f, g \in V$ definiert man $f + g \in V$ durch

$$(f + g)(x) \overset{def}{=} f(x) + g(x) \quad \text{für jedes } x \in M.$$

Man sieht, dass $(V, +)$ eine abelsche Gruppe ist. Ihr neutrales Element ist die Abbildung $0_v\colon M \to K$ mit $0_V(x) = 0_K$ für jedes $x \in M$. Zu jedem $f \in V$ ist die Abbildung $-f\colon M \to K$ mit $(-f)(x) = -f(x)$ für jedes $x \in M$ invers.

Für $\alpha \in K$ und $f \in V$ definiert man eine Abbildung $\alpha f\colon M \to K$ durch

$$(\alpha f)(x) \overset{def}{=} \alpha f(x) \quad \text{für jedes } x \in M.$$

Mit diesen beiden Abbildungen wird V zu einem K-Vektorraum. (Vgl. Übungen.)

(4) Es sei L ein Körper und K ein Unterkörper. Mit der in L gegebenen Addition $L \times L \to L$, $(x, y) \mapsto x + y$ und der durch die Multiplikation in L induzierte Abbildung $K \times L \to L$, $(\alpha, x) \mapsto \alpha \cdot x$ ist L ein K-Vektorraum.
(5) Insbesondere ist K ein K-Vektorraum. Er ist (bis auf Schreibweise der Elemente) gleich dem K-Vektorraum $K^1 = \mathrm{M}(1; K)$.
(6) $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum, $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum.

Definition 8.5. Sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ ist ein *Unterraum* von V, wenn gilt

(UR1) $U \neq \emptyset$.
(UR2) Für alle $x, y \in U$ gilt $x + y \in U$.
(UR3) Für alle $\alpha \in K$ und $x \in U$ gilt $\alpha x \in U$.

Bemerkung 8.6. Sei V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann ist U mit den von V vererbten Verknüpfungen $U \times U \to U$, $(x, y) \mapsto x + y$ und $K \times U \to U$, $(\alpha, x) \mapsto \alpha x$ selbst ein K-Vektorraum. Es ist 0_V das Nullelement in U, und für jedes $x \in U$ ist das Inverse $-x$ in V invers zu x in U.

Beweis. Für alle $x, y \in U$ ist $x + y \in U$ und $-x = (-1) \cdot x \in U$. Also ist U eine Untergruppe der abelschen Gruppe $(V, +)$ und daher selbst eine abelsche Gruppe mit neutralem Element 0_V, und zu jedem $x \in U$ ist $-x$ invers in $(U, +)$. Die Aussagen in (VR2) gelten in V, also erst recht in U. □

Beispiele 8.7. (1) Für jeden K-Vektorraum V sind $\{0_V\}$ und V Unterräume.
(2) $U_1 = \{{}^t(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{R}^n \mid \alpha_1 = 0\}$ ist ein Unterraum von $\mathbb{R}^n$.
$U_2 = \{{}^t(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{R}^n \mid \alpha_1 \geq 0\}$ ist *kein* Unterraum von $\mathbb{R}^n$, denn $x = {}^t(1, 0, \dots, 0) \in U_2$, aber $(-1) \cdot x = {}^t(-1, 0, \dots, 0) \notin U_2$.
$U_3 = \{{}^t(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{R}^n \mid \alpha_1 \neq 0\}$ ist *kein* Unterraum von $\mathbb{R}^n$, denn ${}^t(0, 0, \dots, 0) \notin U_3$.
(3) Sei V ein K-Vektorraum, und seien $U_1, U_2, \dots, U_m$ Unterräume von V. Dann ist

$$\sum_{i=1}^{m} U_i = U_1 + U_2 + \cdots + U_m \overset{def}{=} \Big\{\sum_{i=1}^{m} x_i \mid x_1 \in U_1,\, x_2 \in U_2, \dots, x_m \in U_m\Big\}$$

ein Unterraum von V, und zwar der kleinste Unterraum von V, der alle U_i enthält $(i = 1, \dots, m)$. Er heißt die *Summe* der Unterräume $U_1, \dots, U_m$.
(4) Sei V ein K-Vektorraum. Sei I eine Indexmenge, und für jedes $i \in I$ sei U_i ein Unterraum von V. Dann ist

$$\bigcap_{i \in I} U_i = \{x \in V \mid x \in U_i \text{ für jedes } i \in I\}$$

ein Unterraum von V.

Beweis. (3) Es gilt $0 \in \sum_{i=1}^m U_i$. Seien $\lambda \in K$ und $x = \sum_{i=1}^m x_i$, $y = \sum_{i=1}^m y_i \in \sum_{i=1}^m U_i$, also mit $x_i, y_i \in U_i$ für jedes $i \in \{1, \dots, m\}$. Dann gilt $\lambda x_i \in U_i$ und $x_i + y_i \in U_i$ für jedes $i \in \{1, \dots, m\}$, und es folgt $\lambda x = \sum_{i=1}^m \lambda x_i$, $x + y = \sum_{i=1}^m (x_i + y_i) \in \sum_{i=1}^m U_i$. Offensichtlich gilt $U_i \subseteq \sum_{j=1}^m U_j$ für jedes $i \in \{1, \dots, m\}$. Sei $W \subseteq V$ ein Unterraum mit $U_i \subseteq W$ für alle $i \in \{1, \dots, m\}$. Sei $x \in \sum_{i=1}^m U_i$ beliebig, also von der Form $x = \sum_{i=1}^m x_i$ mit $x \in U_i$, also $x_i \in W$ für alle $i \in \{1, \dots, m\}$. Da W ein Unterraum ist, folgt (induktiv) $x \in W$. Es folgt $\sum_{i=1}^m U_i \subseteq W$.

(4) Setze $U = \cap_{i \in I} U_i$. Da jedes U_i den Nullvektor enthält, folgt $U \neq \emptyset$. Sei $\lambda \in K$, und seien $x, y \in U$. Dann gilt $x, y \in U_i$, also auch λx, $x + y \in U_i$ für alle $i \in \{1, \dots, m\}$, und damit λx, $x + y \in U$. □

DEFINITION 8.8. Sei V ein K-Vektorraum. Seien $x_1, \dots, x_m \in V$. Ein Element $x \in V$ der Form

$$x = \sum_{i=1}^{m} \alpha_i x_i$$

mit $\alpha_1, \dots, \alpha_m \in K$ heißt eine (K-) *Linearkombination* von $x_1, \dots, x_m$.

SATZ 8.9. *Sei V ein K-Vektorraum und $X \subseteq V$ eine Menge. Dann ist*

$$\operatorname{Span}(X) \overset{def}{=} \langle X \rangle \overset{def}{=} \Big\{ \sum_{i=1}^{p} \alpha_i x_i \mid p \in \mathbb{N}_0;\ \alpha_1, \dots, \alpha_p \in K;\ x_1, \dots, x_p \in X \Big\}$$

ein Unterraum von V. Es gilt $X \subseteq \operatorname{Span}(X)$, und $\operatorname{Span}(X)$ ist der kleinste Unterraum von V, der X enthält.

BEZEICHNUNG: $\operatorname{Span}(X) = \langle X \rangle$ heißt *der von X erzeugte Unterraum* von V.

BEWEIS. Man sieht, dass $\operatorname{Span}(X)$ ein Unterraum von V ist mit $X \subseteq \operatorname{Span}(X)$. Für jeden Unterraum U von V mit $X \subseteq U$ gilt offenbar $\operatorname{Span}(X) \subseteq U$: Denn ein beliebiges Element aus $\operatorname{Span}(X)$ hat die Form $\sum_{i=1}^{p} \alpha_i x_i$ mit $\alpha_1, \dots, \alpha_p \in K$ und $x_1, \dots, x_p \in X$. Dann gilt $x_1, \dots, x_p \in U$, und da U eine Unterraum ist, folgt $\sum_{i=1}^{p} \alpha_i x_i \in U$. (Induktion nach p.) □

BEISPIELE 8.10. Sei V ein K-Vektorraum.

(1) Es gilt $\operatorname{Span}(V) = V$ und $\operatorname{Span}(\emptyset) = \{0_V\}$ (leere Summe!).

(2) Es sei $x \in V$. Dann ist

$$\operatorname{Span}(x) \overset{def}{=} \operatorname{Span}(\{x\}) = \{\alpha x \mid \alpha \in K\}$$

der kleinste Unterraum von V, der x enthält. Man schreibt hierfür auch $\operatorname{Span}(x) = Kx$.

(3) Seien $x_1, \dots, x_m \in V$. Dann ist

$$\operatorname{Span}(x_1, \dots, x_m) \overset{def}{=} \operatorname{Span}(\{x_1, \dots, x_m\}) = \Big\{ \sum_{i=1}^{m} \alpha_i x_i \mid \alpha_1, \dots, \alpha_m \in K \Big\}$$

der kleinste Unterraum von V, der $x_1, \dots, x_m$ enthält. Es gilt

$$\operatorname{Span}(x_1, \dots, x_m) = Kx_1 + Kx_2 + \dots + Kx_m.$$

9. Erzeugendensysteme und Basen. Lineare Unabhängigkeit

Es sei K stets ein Körper und V ein K-Vektorraum.

DEFINITION 9.1. (1) Eine Teilmenge $E \subseteq V$ heißt *Erzeugendensystem* von V, wenn $\operatorname{Span}(E) = V$ gilt (wenn also zu jedem $x \in V$ Elemente $\alpha_1, \dots, \alpha_p \in K$ und $x_1, \dots, x_p \in V$ existieren mit $x = \sum_{i=1}^{p} \alpha_i x_i$).

(2) V heißt *endlich erzeugt*, wenn es ein endliches Erzeugendensystem E von V gibt, wenn also $x_1, \dots, x_m \in V$ existieren, so dass es zu jedem $x \in V$ Elemente $\alpha_1, \dots, \alpha_m \in K$ gibt mit $x = \sum_{i=1}^{m} \alpha_i x_i$.

BEISPIELE 9.2. (1) Es gilt $\operatorname{Span}(V) = V$.

(2) Sei $n \in \mathbb{N}$. Für jedes $i \in \{1, \dots, n\}$ setze $e_i \overset{def}{=} {}^t(\delta_{i1}, \delta_{i2}, \dots, \delta_{in}) \in K^n$, also

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots,\ e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Für jedes $(\alpha_1, \ldots, \alpha_n) \in K^n$ gilt $\sum_{i=1}^n \alpha_i e_i$. Also gilt $K^n = \operatorname{Span}(e_1, \ldots, e_n)$.

(3) Allgemeiner: Seien $E_{11}, E_{12}, \ldots, E_{mn}$ die Basismatrizen in $\mathrm{M}(m, n; K)$. Dann gilt $\mathrm{M}(m, n; K) = \operatorname{Span}(E_{11}, E_{12}, \ldots, E_{mn})$.

Definition 9.3. (1) $x_1, \ldots, x_m \in V$ heißen *linear unabhängig*, wenn folgendes gilt: Für alle $\alpha_1, \ldots, \alpha_m \in K$ gilt

$$\sum_{i=1}^m \alpha_i x_i = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0.$$

(2) $x_1, \ldots, x_m \in V$ heißen *linear abhängig*, wenn sie nicht linear unabhängig sind (wenn also $\alpha_1, \ldots, \alpha_m \in K$ existieren, die nicht alle $= 0$ sind, und für die $\sum_{i=1}^m \alpha_i x_i = 0$ gilt).

(3) Eine Teilmenge $F \subseteq V$ heißt *frei* (oder *linear unabhängig*, wenn je endlich viele paarweise verschiedene von F linear unabhängig sind.

Bemerkung 9.4. (1) $\emptyset \subseteq V$ ist frei.

(2) Ist $X \subseteq V$ frei und $Y \subseteq X$, so ist auch Y frei.

(3) Sei $x \in V$. Es ist $\{x\}$ frei genau dann, wenn $x \neq 0$ ist. Denn ist $x = 0$, so gilt $1 \cdot x = 0$. Ist $x \neq 0$ und $\alpha \in K$ mit $\alpha x = 0$, so ist $\alpha = 0$ nach 8.3 (2).

(4) Wenn $x_1, \ldots, x_m \in V$ linear unabhängig sind, dann sind $x_1, \ldots, x_m \in V$ paarweise verschieden und $\neq 0$. (Die Umkehrung gilt nicht!)

(5) Sei $X \subseteq V$. Es ist X genau dann frei, wenn jedes $a \in \operatorname{Span}(X)$ *genau eine* Darstellung als Linearkombination von endlich vielen paarweise verschiedenen Elementen aus X ist.

Beweis. (5) Es gelte: X ist frei. Sei $a \in \operatorname{Span}(X)$. Dann ist a eine Linearkombination von endlich vielen Elementen in X, d. h. es gilt $a = \sum_{x \in X} \alpha_x \cdot x$ mit $\alpha_x \in K$ für jedes $x \in X$ und so, dass $\#\{x \in X \mid \alpha_x \neq 0\} < \infty$ ist. Es gelte auch $a = \sum_{x \in X} \beta_x \cdot x$ mit $\beta_x \in K$ für jedes $x \in X$ und so, dass $\#\{x \in X \mid \beta_x \neq 0\} < \infty$. Wir haben also zwei Darstellungen von a als Linearkombination von Elementen in X:

$$\sum_{x \in X} \alpha_x \cdot x = a = \sum_{x \in X} \beta_x \cdot x.$$

Dann gilt $\#\{x \in X \mid \alpha_x - \beta_x \neq 0\} < \infty$ und

$$\sum_{x \in X} (\alpha_x - \beta_x) \cdot x = a - a = 0.$$

Da X frei ist, folgt $\alpha_x - \beta_x = 0$, also $\alpha_x = \beta_x$ für jedes $x \in X$.

Es gelte: Jedes $a \in \operatorname{Span}(X)$ besitzt genau eine Darstellung als Linearkombination von Elementen in X. Seien $x_1, \ldots, x_m \in X$ paarweise verschieden und $\lambda_1, \ldots, \lambda_m \in K$ mit

$$\sum_{i=1}^m \lambda_i \cdot x_i = 0.$$

Es gilt auch $\sum_{i=1}^m 0_K \cdot x_i = 0_V \in \operatorname{Span}(X)$. Aus der Eindeutigkeitsannahme folgt daher $\lambda_i = 0_K$ für alle $i \in \{1, \ldots, m\}$. □

Beispiele 9.5. (1) $\{e_1, \ldots, e_n\}$ ist eine freie Teilmenge im K-Vektorraum K^n.

(2) Allgemeiner: Die Teilmenge der Basismatrizen $E_{11}, E_{12}, \ldots, E_{mn}$ im K-Vektorraum $\mathrm{M}(m, n; K)$ ist frei.

Beweis. Um die Konzepte einzuüben, zeigen wir zunächst den Spezialfall (1): Zu zeigen genügt, dass $e_1, \ldots, e_n$ linear unabhängig sind: Seien dazu $\alpha_1, \ldots, \alpha_n \in K$ mit

$\sum_{i=1}^{n} \alpha_i e_i = 0$. Es ist also

$$(0,0,\dots,0) = \begin{pmatrix}\alpha_1\\0\\0\\\vdots\\0\\0\end{pmatrix} + \begin{pmatrix}0\\\alpha_2\\0\\\vdots\\0\\0\end{pmatrix} + \dots + \begin{pmatrix}0\\0\\0\\\vdots\\0\\\alpha_n\end{pmatrix} = \begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\vdots\\\alpha_{n-1}\\\alpha_n\end{pmatrix},$$

und es folgt $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Allgemeiner nun (2). Die Idee ist dieselbe. Zu zeigen genügt: $E_{11}, E_{12}, \dots, E_{mn}$ sind linear unabhängig: Seien $\alpha_{11}, \alpha_{12}, \dots, \alpha_{mn} \in K$ mit $\sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}E_{ij} = 0$. Da

$$0 = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}E_{ij} = \begin{pmatrix}\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n}\\ \vdots & \vdots & & \vdots\\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn}\end{pmatrix},$$

folgt $\alpha_{ij} = 0$ für alle $i \in \{1,\dots,m\}$ und $j \in \{1,\dots,n\}$. □

Lemma 9.6. (1) *Sei $F \subseteq V$ frei. Ist $a \in V$ mit $a \notin \mathrm{Span}(F)$, so ist $F \cup \{a\}$ frei.*
(2) *Eine Teilmenge $X \subseteq V$ ist genau dann frei, wenn für jedes $x \in X$ gilt:* $\mathrm{Span}(X \setminus \{x\}) \subsetneq \mathrm{Span}(X)$.

Beweis. (1) Seien $\alpha \in K$ und $\alpha_x \in K$ für jedes $x \in F$, von denen nur endlich viele $\neq 0$ sind, mit

$$\sum_{x\in F} \alpha_x \cdot x + \alpha \cdot a = 0.$$

Wäre $\alpha \neq 0$, so würde $a = -\sum_{x\in F} \alpha^{-1}\alpha_x \cdot x \in \mathrm{Span}(F)$ folgen. Da diese aber nach Voraussetzung nicht gilt, folgt $\alpha = 0$. Dann ist $\sum_{x\in F} \alpha_x \cdot x = 0$, es folgt $\alpha_x = 0$ für alle $x \in F$, da F frei ist. Es folgt, dass $F \cup \{a\}$ frei ist.

(2) Sei $X \subseteq V$ frei. Sei $x \in X$. Trivial gilt $\mathrm{Span}(X \setminus \{x\}) \subseteq \mathrm{Span}(X)$. Angenommen, es würde die Gleichheit gelten. Dann ist wegen $x \in \mathrm{Span}(X)$ auch $x \in \mathrm{Span}(X \setminus \{x\})$. Es gibt also $\alpha_y \in K$ für alle $y \in X \setminus \{x\}$, nur endliche viele davon $\neq 0$, mit $x = \sum_{y\in X\setminus\{x\}} \alpha_y \cdot y$. Es ist also $\sum_{y\in X\setminus\{x\}} \alpha_y \cdot y + (-1)x = 0$ eine Darstellung der Null als Linearkombination von Elementen in X, deren Koeffizienten nicht alle $= 0$ sind, was ein Widerspruch ergibt, da X frei ist.

Gelte umgekehrt $\mathrm{Span}(X \setminus \{x\}) \subsetneq \mathrm{Span}(X)$ für alle $x \in X$. Seien $\alpha_x \in K$ für alle $x \in X$, nur endlich viele davon $\neq 0$, und mit $\sum_{x\in X} \alpha_x \cdot x$. Angenommen, es sind dabei nicht alle Koeffizienten $= 0$. Sei etwa $\alpha_x \neq 0$ für ein $x \in X$. Dann folgt

$$x = -\sum_{y\in X,\, y\neq x} \alpha_x^{-1} \cdot \alpha_y \cdot y,$$

also $x \in \mathrm{Span}(X \setminus \{x\})$. Damit enthält $\mathrm{Span}(X \setminus \{x\})$ die Menge X, und es ergibt sich $\mathrm{Span}(X) = \mathrm{Span}(X \setminus \{x\})$, Widerspruch. □

Definition 9.7. Sei V ein K-Vektorraum. Eine Teilmenge $B \subseteq V$ heißt eine *Basis* von V, wenn folgende beiden Eigenschaften gelten

(1) Es ist $\mathrm{Span}(B) = V$ (d. h. B ist ein Erzeugendensystem von V).
(2) B ist frei.

Beispiele 9.8. (1) Im K-Vektorraum $V = \{0_V\}$ ist $\emptyset$ die einzige Basis von V.
(2) $\{e_1,\dots,e_n\}$ ist eine Basis des K-Vektorraums K^n (vgl. 9.2 und 9.5). Sie heißt die *Standardbasis* von K^n.
(3) $\{E_{ij} \mid 1 \le i \le m,\ 1 \le j \le n\}$ ist eine Basis des K-Vektorraums $\mathrm{M}(m,n;K)$ (vgl. 9.2 und 9.5).

Satz 9.9. *Sei V ein K-Vektorraum und $B \subseteq V$. Folgende Eigenschaften sind äquivalent:*

(1) *B ist eine Basis von V.*
(2) *Jedes $a \in V$ besitzt eine und nur eine Darstellung $a = \sum_{b \in B} \alpha_b \cdot b$ mit $\alpha_b \in K$ für jedes $b \in B$ und mit $\#\{b \in B \mid \alpha_b \neq 0\} < \infty$.*
(3) *B ist ein minimales Erzeugendensystem von V (d. h. $\mathrm{Span}(B) = V$, und für jedes $X \subsetneq B$ gilt $\mathrm{Span}(X) \subsetneq V$).*
(4) *B ist eine maximale freie Teilmenge von V (d. h. B ist frei, und jedes $Y \subseteq V$ mit $B \subsetneq Y$ ist nicht frei).*

Beweis. (1)⇒(2): Ist B eine Basis von V, so gilt $\mathrm{Span}(B) = V$, und die Eindeutigkeitsaussage ist Bemerkung 9.4 (5).

(2)⇒(1): Aus (2) folgt $\mathrm{Span}(B) = V$, und die Eindeutigkeitseigenschaft ergibt wegen Bemerkung 9.4 (5), dass B frei ist.

(1)⇒(3): Sei B eine Basis von V. Dann ist B ein Erzeugendensystem. Sei $X \subseteq V$ mit $X \subsetneq B$. Dann gibt es ein $b \in B$ mit $b \notin X$, also $X \subseteq B \setminus \{b\}$. Es folgt $\mathrm{Span}(X) \subseteq \mathrm{Span}(B \setminus \{b\}) \subsetneq \mathrm{Span}(B)$, wobei die Ungleichheit aus Lemma 9.6 (2) folgt. Also ist B ein minimales Erzeugendensystem von V.

(3)⇒(4): Sei B ein minimales Erzeugendensystem von V. Für jedes $b \in B$ ist wegen der Minimalität $\mathrm{Span}(B \setminus \{b\}) \subsetneq V = \mathrm{Span}(B)$, und nach Lemma 9.6 (2) ist B frei. Angenommen, es gibt $X \subseteq V$ frei mit $B \subsetneq X$. Dann gibt es ein $x \in X$ mit $x \notin B$, also mit $B \subseteq X \setminus \{x\}$. Nach Lemma 9.6 (2) gilt $\mathrm{Span}(B) \subseteq \mathrm{Span}(X \setminus \{x\}) \subsetneq \mathrm{Span}(X) = V$, Widerspruch, da B ein Erzeugendensystem von V ist. Also ist B eine maximale freie Teilmenge von V.

(2)⇒(1): Sei B eine maximale freie Teilmenge von V. Sei $v \in V \setminus B$. Dann ist $B \subsetneq B \cup \{v\}$, und wegen der Maximalität ist $B \cup \{v\}$ nicht frei. Aus Lemma 9.6 (1) folgt $v \in \mathrm{Span}(B)$. Ist $v \in B$, so gilt natürlich auch $v \in \mathrm{Span}(B)$. Es ergibt sich also $V = \mathrm{Span}(B)$. Damit ist B ein freies Erzeugendensystem, also eine Basis. □

Beispiel 9.10. Seien $a_1, \ldots, a_p \in K^n$. Man schreibe diese Spalten nebeneinander in eine Matrix $A \overset{def}{=} [a_1, \ldots, a_p] \in \mathrm{M}(n, p; K)$. Sei T die zu A gehörige Treppenmatrix, $r = \mathrm{rang}(T) = \mathrm{rang}(A)$, seien $j(1), \ldots, j(r)$ die charakteristischen Spaltenindizes. Sei $U \overset{def}{=} \mathrm{Span}(a_1, \ldots, a_p) \subseteq K^n$. Dann gilt:

(1) $\{a_{j(1)}, \ldots, a_{j(r)}\}$ ist eine Basis von U.
(2) $a_1, \ldots, a_p$ sind genau dann linear unabhängig, wenn $p = r$ gilt.

Beweis. Wir werden dies im nächsten Abschnitt in etwas allgemeinerer Form beweisen. Wir rechnen hier ein konkretes Beispiel: Seien

$$a_1 = \begin{pmatrix} 2 \\ 4 \\ 5 \\ 2 \end{pmatrix}, \ a_2 = \begin{pmatrix} 2 \\ 6 \\ 6 \\ 3 \end{pmatrix}, \ a_3 = \begin{pmatrix} 0 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \ a_4 = \begin{pmatrix} 2 \\ 7 \\ 7 \\ 4 \end{pmatrix} \in \mathbb{R}^4.$$

Sei $U \overset{def}{=} \mathrm{Span}(a_1, a_2, a_3, a_4) \subseteq \mathbb{R}^4$. Mit obigem Verfahren kann man untersuchen, ob $a_1, a_2, a_3, a_4 \in \mathbb{R}^4$ linear unabhängig sind bzw. welche linearen Abhängigkeiten zwischen diesen Elementen bestehen. Genauer findet man eine maximale freie Teilmenge von $\{a_1, a_2, a_3, a_4\}$.

Setze

$$A = [a_1, a_2, a_3, a_4] = \begin{pmatrix} 2 & 2 & 0 & 2 \\ 4 & 6 & 4 & 7 \\ 5 & 6 & 2 & 7 \\ 2 & 3 & 2 & 4 \end{pmatrix}.$$

Mit dem Gauß-Algorithmus berechnet man die zu A gehörige Treppenmatrix

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

Es folgt $r = \operatorname{rang}(A) = \operatorname{rang}(T) = 3$, $j(1) = 1$, $j(2) = 2$ und $j(3) = 4$, und obige Aussage (1) ergibt, dass eine Basis von U gegeben ist durch $\{a_1, a_2, a_4\}$.

Wir wollen die Aussage (1) aber direkt einsehen: Man sieht an der Treppenmatrix, dass

$$a_3 = (-2) \cdot a_1 + 2 \cdot a_2 \tag{9.1}$$

gilt: Denn es gilt offenbar

$$T_{\bullet 3} = (-2) \cdot T_{\bullet 1} + 2 \cdot T_{\bullet 2}. \tag{9.2}$$

Es gibt ein $P \in \mathrm{GL}(4; \mathbb{R})$ mit $T = PA$, also mit $A = P^{-1}T$. Multipliziert man (9.2) mit P^{-1}, so erhält man die entsprechende Beziehung (9.1) für A.

Es folgt also, dass $\{a_1, a_2, a_4\}$ ein Erzeugendensystem von U ist. Die lineare Unabhängigkeit von a_1, a_2, a_4 zeigt man ganz genauso: Seien α, β, $\gamma \in \mathbb{R}$ mit

$$\alpha \cdot a_1 + \beta \cdot a_2 + \gamma \cdot a_4 = 0. \tag{9.3}$$

Multipliziert man diese Gleichung mit der invertierbaren Matrix P, so erhält man

$$\alpha \cdot T_{\bullet 1} + \beta \cdot T_{\bullet 2} + \gamma \cdot T_{\bullet 4} = 0, \tag{9.4}$$

und es folgt $\alpha = \beta = \gamma = 0$, denn offenbar sind $T_{\bullet 1} = e_1$, $T_{\bullet 2} = e_2, T_{\bullet 4} = e_3$ linear unabhängig in $\mathbb{R}^4$. Damit sehen wir – in diesem Beispiel – die Aussage (1). Der allgemeine Beweis ergibt sich aber ganz genauso. Die Aussage (2) folgt dann aus Aussage (1). □

Satz 9.11. *Sei V ein endlich erzeugter K-Vektorraum. Dann existiert eine Basis von V. Genauer gilt: Jedes endliche Erzeugendensystem von V enthält eine Basis.*

Beweis. Sei E ein endliches Erzeugendensystem von V. (Nach Voraussetzung gibt es ein solches.) Sei

$$M \overset{def}{=} \{X \mid X \subseteq E \text{ mit } \operatorname{Span}(X) = V\} \subseteq 2^E.$$

Es ist $M \neq \emptyset$ (denn $E \in M$), und für jedes $X \in M$ gilt $0 \leq \#(X) \leq \#(E)$. Also gibt es ein $B \in M$ mit minimaler Elementanzahl. Es gilt $B \subseteq E$ und $\operatorname{Span}(B) = V$. Außerdem ist wegen der minimalen Elementanzahl B ein minimales Erzeugendensystem. Nach Satz 9.9 ist B daher eine Basis. □

Satz 9.12. *Jeder K-Vektorraum V besitzt eine Basis.*

Beweis. Ist V endlich erzeugt, so ist dies der vorstehende Satz. Ist V nicht endlich erzeugt, so verwendet man das sogenannte *Zornsche Lemma* (eine Aussage, die äquivalent zum Auswahlaxiom ist). Auf eine detaillierte Beschreibung des Beweises verzichten wir hier. □

Lemma 9.13. *Seien $v_1, \ldots, v_m \in V$ linear unabhängig. Seien $x_1, \ldots, x_r \in \operatorname{Span}(v_1, \ldots, v_m)$ linear unabhängig. Dann gilt $r \leq m$.*

Beweis. Induktion nach m. Für $m = 0$ gilt $\operatorname{Span}(v_1, \ldots, v_m) = \operatorname{Span}(\emptyset) = \{0\}$, und dann folgt auch $r = 0$. Sei nun $m \in \mathbb{N}$, und es sei bereits gezeigt: Sind $v'_1, \ldots, v'_{m-1} \in V$ linear unabhängig und $x'_1, \ldots, x'_s \in \operatorname{Span}(v'_1, \ldots, v'_{m-1})$ linear unabhängig, so gilt $s \leq m - 1$. Seien nun $v_1, \ldots, v_m \in V$ linear unabhängig und $x_1, \ldots, x_r \in \operatorname{Span}(v_1, \ldots, v_m)$ linear unabhängig. Zu jedem $j \in \{1, \ldots, r\}$ existieren $\alpha_{1j}, \ldots, \alpha_{mj} \in K$ mit

$$x_j = \sum_{i=1}^{m} \alpha_{ij} \cdot v_i.$$

Zwei Fälle können auftreten:

1. Fall: Es ist $\alpha_{m1} = \alpha_{m2} = \cdots = \alpha_{mr} = 0$. Dann gilt $x_1, \dots, x_r \in \operatorname{Span}(v_1, \dots, v_{m-1})$, und $v_1, \dots, v_{m-1}$ sind linear unabhängig. Aus der Induktionsvoraussetzung folgt $r \leq m - 1 < m$.

2. Fall: Es gelte $\alpha_{mj} \neq 0$ für ein $j \in \{1, \dots, r\}$. Nach evtl. Umnummerierung kann man $\alpha_{mr} \neq 0$ annehmen. Für jedes $j \in \{1, \dots, r-1\}$ gilt

$$\begin{aligned}
x'_j &\overset{def}{=} x_j - \alpha_{mr}^{-1} \cdot \alpha_{mj} \cdot x_r \\
&= \sum_{i=1}^{m} \big(\alpha_{ij} - \alpha_{mr}^{-1} \cdot \alpha_{mj} \cdot \alpha_{ir}\big) \cdot v_i \\
&= \sum_{i=1}^{m-1} \big(\alpha_{ij} - \alpha_{mr}^{-1} \cdot \alpha_{mj} \cdot \alpha_{ir}\big) \cdot v_i \in \operatorname{Span}(v_1, \dots, v_{m-1}).
\end{aligned}$$

Es sind $x'_1, \dots, x'_{r-1}$ linear unabhängig: Denn seien $\lambda_1, \dots, \lambda_{r-1} \in K$ mit $\sum_{j=1}^{r-1} \lambda_j x'_j = 0$. Dann gilt

$$0 = \sum_{j=1}^{r-1} \lambda_j \big(x_j - \alpha_{mr}^{-1} \cdot \alpha_{mj} \cdot x_r\big) = \sum_{j=1}^{r-1} \lambda_j x_j + \left(-\sum_{j=1}^{r-1} \lambda_j \alpha_{mr}^{-1} \alpha_{mj} \cdot\right) \cdot x_r.$$

Weil $x_1, \dots, x_r$ linear unabhängig sind, folgt $\lambda_1 = \lambda_2 = \cdots = \lambda_{r-1} = 0$. $x'_1, \dots, x'_{r-1}$ sind also linear unabhängig.

Nun folgt aus der Induktionsvoraussetzung $r - 1 \leq m - 1$, also $r \leq m$. □

Satz 9.14 (Satz von der Invarianz der Dimension). *Sei V ein endlich erzeugter K-Vektorraum. Dann gibt es eine Zahl $n \in \mathbb{N}_0$, so dass jede Basis von V die Elementanzahl n hat.*

Bezeichnung: Diese Elementanzahl n heißt die *Dimension* von V. Man schreibt dafür $\dim(V) = \dim_K(V) = n$.

Beweis. Nach Satz 9.11 gibt es eine Basis $B = \{b_1, \dots, b_n\}$ von V (mit $\#(B) = n$). Sei B' eine weitere Basis von V. Sind $b'_1, \dots, b'_r \in B'$ paarweise verschieden, so sind dann $b'_1, \dots, b'_r$ linear unabhängig in $V = \operatorname{Span}(b_1, \dots, b_n)$, und aus dem vorherigen Lemma folgt $r \leq n$. Jede endliche Teilmenge von B' hat also höchstens n Elemente, und es folgt, dass B' endlich ist mit $\#(B') \leq n$. Vertauscht man nun die Rollen von B und B', so ergibt sich auch $n \leq \#(B')$. □

Bemerkung 9.15.
(1) Ein endlich erzeugter K-Vektorraum heißt auch *endlich-dimensionaler* K-Vektorraum.
(2) Ist V ein beliebiger Vektorraum (nicht notwendig endlich erzeugt), so kann man zeigen: Sind B und B' Basen von V, so gibt es eine bijektive Abbildung $f\colon B \to B'$. Insofern hat man auch für nicht endlich erzeugte Vektorräume die Invarianz der Dimension; die Dimension ist um auch unendlichen Fall eine eindeutig definierte Kardinalzahl.

Wir werden später sehen, dass ein K-Vektorraum durch seine Dimension weitgehend bestimmt ist.

Beispiele 9.16.
(1) Sei V ein K-Vektorraum. Dann gilt
$$\dim_K(V) = 0 \quad \Leftrightarrow \quad \emptyset \text{ ist eine Basis von } V \quad \Leftrightarrow \quad V = \{0\}.$$
(2) Sei V ein K-Vektorraum mit $\dim_K(V) = 1$. Dann gibt es ein $b \in V$, so dass $\{b\}$ eine Basis von V ist. Es gilt dann $b \neq 0$, und es ist
$$V = \operatorname{Span}(b) = K \cdot b = \{\lambda \cdot b \mid \lambda \in K\}.$$

Zu jedem $a \in V$ mit $a \neq 0$ existiert ein $\alpha \in K$ mit $\alpha \neq 0$ und mit $a = \alpha \cdot b$. Es folgt $V = \mathrm{Span}(b) = \mathrm{Span}(\alpha^{-1} \cdot b) = \mathrm{Span}(A)$, d. h. $\{a\}$ ist eine Basis von V.

(3) Im K-Vektorraum $\mathrm{M}(m, n; K)$ ist $\{E_{ij} \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$ eine Basis. Es folgt $\dim_K(\mathrm{M}(m, n; K)) = m \cdot n$.
(4) Im K-Vektorraum $K^n = \mathrm{M}(n, 1; K)$ hat man die Standardbasis $\{e_1, \ldots, e_n\}$, also $\dim_K(K^n) = n$.
(5) Speziell für $n = 1$ erhält man den K-Vektorraum $K^1 = K$. Es gilt $\dim_K(K) = 1$; für jedes $a \in K$, $a \neq 0$, ist $\{a\}$ eine Basis.
(6) $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum, und man kann zeigen, dass $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$ gilt.
(7) $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum mit $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, denn $\{1, i\}$ ist eine Basis.

Satz 9.17 (Austauschsatz von Steinitz (E. Steinitz, 1871–1928)). *Sei V ein endlichdimensionaler K-Vektorraum mit $\dim(V) = n$, sei B eine Basis von V. Seien $a_1, \ldots, a_m \in V$ linear unabhängig. Dann gilt $m \leq n$ und es gibt $b_1, \ldots, b_{n-m} \in B$, so dass $\{a_1, \ldots, a_m,\ b_1, \ldots, b_{n-m}\}$ eine Basis von V ist.*

Beweis. Nach Lemma 9.13 gilt $m \leq \#(B) = n$ (die letzte Gleichheit gilt nach der Invarianz der Dimension). Sei $M_0 \overset{def}{=} \{a_1, \ldots, a_m\}$, sei $M \overset{def}{=} M_0 \cup B$. Sei

$$\mathcal{M} \overset{def}{=} \{X \subseteq V \mid X \text{ ist frei mit } M_0 \subseteq X \subseteq M\} \subseteq 2^M.$$

Wegen $M_0 \in \mathcal{M}$ ist $\mathcal{M} \neq \emptyset$, und als endliche Menge besitzt sie ein maximales Element $F \in \mathcal{M}$. Dann ist F frei mit $M_0 \leq F \leq M$.

Es gilt $B \subseteq \mathrm{Span}(F)$: Denn angenommen, es gilt $B \not\subseteq \mathrm{Span}(F)$. Dann gibt es $b \in B$ mit $b \notin \mathrm{Span}(F)$. Nach 9.6 (1) ist dann $F \cup \{b\}$ frei, und es gilt $F \cup \{b\} \in \mathcal{M}$, Widerspruch zur Maximalität von F. – Es folgt $V = \mathrm{Span}(B) \subseteq \mathrm{Span}(F) \subseteq V$, also ist F ein freies Erzeugendensystem, d. h. eine Basis von V. □

Satz 9.18. *Sei V ein endlichdimensionaler K-Vektorraum mit $\dim(V) = n$.*

(1) *Sei E ein endliches Erzeugendensystem von V. Dann gilt $\#(E) \geq n$.*
(2) *Sei E ein Erzeugendensystem von V mit $\#(E) = n$. Dann ist E eine Basis von V.*
(3) *Sei $F \subseteq V$ frei. Dann gibt es eine Basis B von V mit $F \subseteq B$.*
(4) *Sei $F \subseteq V$ frei. Dann gilt $\#(F) \leq n$.*
(5) *Sei $F \subseteq V$ frei mit $\#(F) = n$. Dann ist F eine Basis von V.*

Beweis. Folgt sofort aus den vorherigen Aussagen: (1) und (2) aus Satz 9.11, (3), (4) und (5) aus Satz 9.18; man beachte, dass auch die Invarianz der Dimension 9.14 mehrfach eingeht. □

Satz 9.19 (Basisergänzungssatz). *Sei V ein endlichdimensionaler K-Vektorraum mit $\dim(V) = n$. Sei U ein Unterraum von V. Dann gilt:*

(1) $\dim(U) \leq \dim(V)$.
(2) *Sei B_U eine Basis von U. Dann gibt es eine Basis B_V von V mit $B_U \subseteq B_V$.*

Beweis. Eine Basis B_U von U ist eine freie Teilmenge von V. Sei B'_V eine Basis von V. Nach dem Austauschsatz von Steinitz gibt es (endlich viele) Elemente von B'_V, die zusammen mit B_U eine Basis von B_V bilden. Es folgt Aussage (2) und $\dim(U) = \#(B_U) \leq \#(B_V) = \dim(V)$. □

Folgerung 9.20. *Sei V ein endlichdimensionaler K-Vektorraum. Ist $U \subseteq V$ ein Unterraum mit $\dim(U) = \dim(V)$, so ist $U = V$.*

Beweis. Sei B_U eine Basis von U. Nach dem vorherigen Satz kann man sie zu einer Basis $B_V \supseteq B_U$ von V ergänzen. Wäre $U \subsetneq V$, dann müsste $B_U \subsetneq B_V$ gelten, und damit $\dim(U) < \dim(V)$. □

SATZ 9.21 (Dimensionsformel für Unterräume). *Sei V ein K-Vektorraum. Seien U_1, $U_2 \subseteq V$ eindlichdimensionale Unterräume. Dann sind die Unterräume U_1+U_2 und $U_1 \cap U_2$ endlichdimensional, und es gilt*

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2).$$

BEWEIS. Sei $r = \dim(U_1)$ und $s = \dim(U_2)$. Wegen $U_1 \cap U_2 \subset U_1, U_2$ gilt $d \overset{def}{=} \dim(U_1 \cap U_2) \leq r, s$. Sei $\{a_1, \ldots, a_d\}$ eine Basis von $U_1 \cap U_2$. Nach dem Basisergänzungssatz gibt es $b_1, \ldots, b_{r-d} \in U_1$, so dass $\{a_1, \ldots, a_d, b_1, \ldots, b_{r-d}\}$ eine Basis von U_1 ist, und ebenso gibt es $c_1, \ldots, c_{s-d} \in U_2$, so dass $\{a_1, \ldots, a_d, c_1, \ldots, c_{s-d}\}$ eine Basis von U_2 ist. Setze $B \overset{def}{=} \{a_1, \ldots, a_d, b_1, \ldots, b_{r-d}, c_1, \ldots, c_{s-d}\}$. Dann gilt $B \subseteq U_1 \cup U_2 \subseteq U_1 + U_2$, und es folgt $\operatorname{Span}(B) \subseteq U_1 + U_2$. Ist $u_1 + u_2 \in U_1 + U_2$, so kann man schreiben

$$u_1 = \sum_{i=1}^{d} \alpha_i a_i + \sum_{j=1}^{r-d} \beta_j b_j$$

und

$$u_2 = \sum_{k=1}^{d} \alpha'_k a_k + \sum_{\ell=1}^{s-d} \gamma_\ell c_\ell$$

mit allen α_i, α'_k, β_j, $\gamma_\ell \in K$. Aber dann ist

$$u_1 + u_2 = \sum_{i=1}^{d} (\alpha_i + \alpha'_i) a_i + \sum_{j=1}^{r-d} \beta_j b_j + \sum_{\ell=1}^{s-d} \gamma_\ell c_\ell \in \operatorname{Span}(B).$$

Also ist B ein Erzeugendensystem von $U_1 + U_2$. Zeige, dass

$$a_1, \ldots, a_d, b_1, \ldots, b_{r-d}, c_1, \ldots, c_{s-d}$$

linear unabhängig sind. Seien α_i, β_j, $\gamma_k \in K$ mit

$$\sum_{i=1}^{d} \alpha_i a_i + \sum_{j=1}^{r-d} \beta_j b_j + \sum_{k=1}^{s-d} \gamma_k c_k = 0. \tag{9.5}$$

Dann ist

$$\sum_{k=1}^{s-d} \gamma_k c_k = -\sum_{i=1}^{d} \alpha_i a_i - \sum_{j=1}^{r-d} \beta_j b_j \in U_2 \cap U_1,$$

also von der Form

$$\sum_{k=1}^{s-d} \gamma_k c_k = \sum_{i=1}^{d} \alpha'_i a_i$$

für gewisse $\alpha'_i \in K$. Es folgt

$$\sum_{i=1}^{d} \alpha'_i a_i + \sum_{k=1}^{s-d} (-\gamma_k) c_k = 0.$$

Da $a_1, \ldots, a_d, c_1, \ldots, c_{s-d}$ linear unabhängig sind, folgt, dass alle $\alpha'_i = 0$ und alle $\gamma_k = 0$ sind. Dann wird (9.5) zu

$$\sum_{i=1}^{d} \alpha_i a_i + \sum_{j=1}^{r-d} \beta_j b_j = 0.$$

Da $a_1, \ldots, a_d, b_1, \ldots, b_{r-d}$ linear unabhängig sind, folgt, dass alle $\alpha_i = 0$ und alle $\beta_j = 0$ sind. Es folgt also die behauptete lineare Unabhängigkeit. Es ist also B eine Basis von $U_1 + U_2$, und es folgt

$$\begin{aligned} \dim(U_1 + U_2) &= d + (r-d) + (s-d) = r + s - d \\ &= \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2). \end{aligned}$$

□

Definition 9.22. Sei V ein K-Vektorraum, und seien $U_1, \dots, U_m \subseteq V$ Unterräume. Die Summe $\sum_{i=1}^m U_i = U_1 + \cdots + U_m$ heißt *direkt*, falls sich jedes $v \in \sum_{i=1}^m U_i$ *eindeutig* als $v = u_1 + \cdots + u_m$ schreiben lässt mit $u_i \in U_i$ für alle $i = 1, \dots, m$. Man schreibt dann

$$\sum_{i=1}^m U_i = \bigoplus_{i=1}^m U_i = U_1 \oplus \cdots \oplus U_m.$$

Bemerkung 9.23. Sei V ein K-Vektorraum, und seien $U_1, U_2 \subseteq V$ Unterräume.

(1) Es gilt $U_1 + U_2 = U_1 \oplus U_2$ genau dann, wenn $U_1 \cap U_2 = \{0\}$ ist.
(2) Seien U_1 und U_2 endlichdimensional. Es gilt $U_1 + U_2 = U_1 \oplus U_2$ genau dann, wenn $\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2)$ ist.

(Vgl. Übungen.)

10. Praktische Rechenverfahren

Mit K werde immer ein Körper bezeichnet.

Satz 10.1. *Sei V ein endlichdimensionaler K-Vektorraum mit $\dim(V) = n$ und sei $\{b_1, \dots, b_n\}$ eine Basis von V. Seien $a_1, \dots, a_n \in V$ und sei $A = (\alpha_{ij}) \in \mathrm{M}(n; K)$ die Matrix mit $a_j = \sum_{i=1}^n \alpha_{ij} b_i$ für jedes $j \in \{1, \dots, n\}$. Dann gilt:*

(1) *$\{a_1, \dots, a_n\}$ ist genau dann eine Basis von V, wenn A invertierbar ist.*
(2) *Ist A invertierbar, und ist $A^{-1} = (\widetilde{\alpha}_{ij})$, so ist $b_j = \sum_{i=1}^n \widetilde{\alpha}_{ij} a_i$ für jedes $j \in \{1, \dots, n\}$.*

Beweis. (1) (a) Es gelte, dass $\{a_1, \dots, a_n\}$ eine Basis von V ist. Dann gibt es ein $B = (\beta_{ij}) \in \mathrm{M}(n; K)$ mit $b_j = \sum_{i=1}^n \beta_{ij} a_i$ für jedes $j = 1, \dots, n$. Dann folgt

$$\begin{aligned} b_j &= \sum_{i=1}^n \beta_{ij} a_i = \sum_{i=1}^n \beta_{ij} \sum_{k=1}^n \alpha_{ki} b_k \\ &= \sum_{k=1}^n \sum_{i=1}^n \alpha_{ki} \beta_{ij} b_k = \sum_{k=1}^n AB[k,j] b_k \end{aligned}$$

Da $b_1, \dots, b_n$ linear unabhängig sind, folgt $AB[k,j] = \delta_{kj}$, d. h. $AB = E_n$. Analog folgt $BA = E_n$. Also ist A invertierbar (und $B = A^{-1}$).

(b) Es gelte, dass A invertierbar ist. Seien $\lambda_1, \dots, \lambda_n \in K$ mit $\sum_{j=1}^n \lambda_j a_j = 0$. Setze

$$x \stackrel{def}{=} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Dann gilt also $Ax = 0$. Es folgt $x = A^{-1}Ax = A^{-1} \cdot 0 = 0$, also $\lambda_1 = \cdots = \lambda_n = 0$. Also sind $a_1, \dots, a_n$ linear unabhängig. Da $\dim(V) = n$ ist, folgt, dass $\{a_1, \dots, a_n\}$ eine Basis von V ist.

(2) Dies ergibt sich aus Beweisteil (1)(a). □

Folgerung 10.2. *Sei $A = (\alpha_{ij}) \in \mathrm{M}(n; K)$. Folgende Aussagen sind äquivalent:*

(1) *A ist invertierbar.*
(2) *${}^tA$ ist invertierbar.*
(3) *Die Spalten von A sind linear unabhängig in $\mathrm{M}(n, 1; K)$.*
(4) *Die Zeilen von A sind linear unabhängig in $\mathrm{M}(1, n; K)$.*

Beweis. (1)$\Leftrightarrow$(2): folgt aus 6.13.

(1)$\Leftrightarrow$(3): Für jedes $j \in \{1, \dots, n\}$ gilt $A_{\bullet j} = \sum_{i=1}^n \alpha_{ij} e_i$. Es sind $A_{\bullet 1}, \dots, A_{\bullet n}$ linear unabhängig genau dann, wenn $\{A_{\bullet 1}, \dots, A_{\bullet n}\}$ eine Basis von $\mathrm{M}(n, 1; K)$ ist. Dies ist nach Satz 10.1 äquivalent zur Invertierbarkeit von A.

(2)$\Leftrightarrow$(4): Dies folgt wie oben durch Betrachtung der transponierten Matrix bzw. Zeilen statt Spalten. $\square$

Satz 10.3. *Sei V ein endlichdimensionaler K-Vektorraum mit $\dim(V) = n$ und sei $\{b_1, \dots, b_n\}$ eine Basis von V. Seien $a_1, \dots, a_p \in V$ und sei $A = (\alpha_{ij}) \in \mathrm{M}(n, p; K)$ die Matrix mit $a_j = \sum_{i=1}^n \alpha_{ij} b_i$ für jedes $j \in \{1, \dots, p\}$. Sei T die zu A gehörige Treppenmatrix vom Rang r und mit charakteristischen Spaltenindizes $j(1), \dots, j(r)$. Für den Unterraum $U \overset{def}{=} \mathrm{Span}(a_1, \dots, a_p)$ gilt:*

(1) $\dim(U) = r$.
(2) $\{a_{j(1)}, \dots, a_{j(r)}\}$ *ist eine Basis von U.*
(3) $a_1, \dots, a_p$ *sind genau dann linear unabhängig, wenn $p = r$ gilt.*

Beweis. (a) Da die Spalten $T_{\bullet j(1)}, \dots, T_{\bullet j(r)}$ gerade die r ersten Standardbasisvektoren $e_1, \dots, e_r \in \mathrm{M}(n, 1; K)$ sind, folgt, dass die Menge dieser Spalten eine Basis des von allen Spalten von T erzeugten Unterraums von $\mathrm{M}(n, 1; K)$ ist. Es gibt ein $P \in \mathrm{GL}(n; K)$ mit $T = PA$, und es folgt (vgl. das Argument in Beispiel 9.10), dass $\{A_{\bullet j(1)}, \dots, A_{\bullet j(r)}\}$ eine Basis des von allen Spalten von A erzeugten Unterraums von $\mathrm{M}(n, 1; K)$ ist.

(b) Seien $\lambda_1, \dots, \lambda_r \in K$ mit $\sum_{k=1}^r \lambda_k a_{j(k)} = 0$. Dann gilt

$$0 = \sum_{k=1}^r \lambda_k a_{j(k)} = \sum_{k=1}^r \lambda_k \sum_{i=1}^n \alpha_{i,j(k)} b_i = \sum_{i=1}^n \Big(\sum_{k=1}^r \lambda_k \alpha_{i,j(k)} \Big) b_i,$$

und da $b_1, \dots, b_n$ linear unabhängig sind, folgt $\sum_{k=1}^r \lambda_k \alpha_{i,j(k)} = 0$ für jedes $i = 1, \dots, n$, d. h. in $\mathrm{M}(n, 1; K)$ gilt $\sum_{k=1}^r \lambda_k A_{\bullet j(k)} = 0$. Weil aber $\{A_{\bullet j(1)}, \dots, A_{\bullet j(r)}\}$ nach Teil (a) linear unabhängig sind, folgt $\lambda_1 = \cdots = \lambda_r = 0$. Also sind $a_{j(1)}, \dots, a_{j(r)}$ linear unabhängig.

(c) Sei $T = (\tau_{ij})$. Wegen $A = P^{-1} T$ gilt für jedes $j \in \{1, \dots, p\}$

$$\begin{aligned} A_{\bullet j} &= P^{-1} T_{\bullet j} = P^{-1} \sum_{k=1}^r \tau_{kj} T_{\bullet j(k)} = \sum_{k=1}^r \tau_{kj} P^{-1} T_{\bullet j(k)} \\ &= \sum_{k=1}^r \tau_{kj} A_{\bullet j(k)}. \end{aligned}$$

Es folgt $\alpha_{ij} = \sum_{k=1}^r \tau_{kj} \alpha_{ij(k)}$, und damit

$$\begin{aligned} a_j &= \sum_{i=1}^n \alpha_{ij} b_i = \sum_{i=1}^n \sum_{k=1}^r \tau_{kj} \alpha_{ij(k)} b_i = \sum_{k=1}^r \tau_{kj} \sum_{i=1}^n \alpha_{ij(k)} b_i \\ &= \sum_{k=1}^r \tau_{kj} a_{j(k)}, \end{aligned}$$

also $a_j \in \mathrm{Span}(a_{j(1)}, \dots, a_{j(r)})$ für jedes $j = 1, \dots, p$. Es folgt, dass $\{a_{j(1)}, \dots, a_{j(r)}\}$ ein Erzeugendensystem von U ist. Damit folgt Aussage (2). Die beiden anderen Aussagen (1) und (3) folgen direkt daraus. $\square$

Der folgende Satz beschreibt, wie man eine Basisergänzung praktisch durchführt.

Satz 10.4. *Sei V ein endlichdimensionaler K-Vektorraum mit $\dim(V) = n$. Sei $\{b_1, \dots, b_n\}$ eine Basis von V. Seien $a_1, \dots, a_r \in V$ linear unabhängig. Sei $A = (\alpha_{ij}) \in \mathrm{M}(n, r; K)$ die Matrix mit $a_j = \sum_{i=1}^n \alpha_{ij} b_i$ für jedes $j \in \{1, \dots, r\}$. Setze $\widehat{A} = (A, E_n) \in \mathrm{M}(n, r + n; K)$ und $\widehat{T}$ die zugehörige Treppenmatrix. Dann gilt:*

(1) $\operatorname{rang}(\widehat{A}) = \operatorname{rang}(\widehat{T}) = n$, *und für die charakteristischen Spaltenindizes gilt* $j(1) = 1$, $j(2) = 2, \ldots, j(r) = r$ *und* $r + 1 \leq j(r+1) < \cdots < j(n) \leq r + n$.
(2) *Es ist* $\{a_1, \ldots, a_r, b_{j(r+1)-r}, \ldots, b_{j(n)-r}\}$ *eine Basis von* V.

Beweis. (1) Setze

$$\widehat{a}_i \overset{def}{=} \begin{cases} a_i & i \in \{1, \ldots, r\}, \\ b_{i-r} & i \in \{r+1, \ldots, r+n\}. \end{cases}$$

Sei $\widehat{A} = (A, E_n) = (\widehat{\alpha}_{ij})$. Dann gilt

$$\widehat{a}_j = \sum_{i=1}^{n} \widehat{\alpha}_{ij} \cdot b_i$$

für jedes $j \in \{1, \ldots, r+n\}$. Da schon $\{b_1, \ldots, b_n\}$ ein Erzeugendensystem von V ist, gilt dies erst recht für (die Obermenge) $\{\widehat{a}_1, \ldots, \widehat{a}_{r+n}\}$. Nach Satz 10.3 (1) gilt

$$\operatorname{rang}(\widehat{A}) = \dim(V) = n.$$

(2) Es gibt ein $P \in \mathrm{GL}(n; K)$ mit $\widehat{T} = P\widehat{A} = P \cdot (A, E_n) = (PA, P)$. Dabei ist $PA \in \mathrm{M}(n, r; K)$ offenbar die zu A gehörige Treppenmatrix. Weil $a_1, \ldots, a_r$ linear unabhängig sind, folgt nach Satz 10.3 (3)

$$\operatorname{rang}(PA) = \operatorname{rang}(A) = r.$$

Hieraus folgt notwendig $j(1) = 1, \ldots, j(r) = r$. Klar ist dann $r + 1 \leq j(r+1) < \cdots < j(n) \leq r + n$. Aus Satz 10.3 (2) folgt nun, dass

$$\{\widehat{a}_{j(1)}, \ldots, \widehat{a}_{j(1)}\} = \{a_1, \ldots, a_r, b_{j(r+1)-r}, \ldots, b_{j(n)-r}\}$$

eine Basis von $\operatorname{Span}(\widehat{a}_1, \ldots, \widehat{a}_{r+n}) = V$ ist. $\square$

Beispiel 10.5. Sei V ein $\mathbb{R}$-Vektorraum mit $\dim(V) = 4$. Sei $\{b_1, b_2, b_3, b_4\}$ eine Basis von V. Seien

$$a_1 = b_1 + 2b_2 + 3b_3 + 4b_4, \ a_2 = b_1 + b_3 + 2b_4, \ a_3 = 4b_2 + b_3 - 2b_4.$$

Anwendung des Algorithmus von Gauß auf

$$\widehat{A} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 4 & 0 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 & 0 \\ 4 & 2 & -2 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{M}(4, 7; \mathbb{R})$$

liefert die zugehörige Treppenmatrix

$$\widehat{T} = \begin{pmatrix} 1 & 0 & 0 & -2/3 & 0 & 1/3 & 1/6 \\ 0 & 1 & 0 & 5/3 & 0 & -1/3 & -1/6 \\ 0 & 0 & 1 & 1/3 & 0 & 1/3 & -1/3 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

Man sieht daran: Die zu

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 4 \\ 3 & 1 & 1 \\ 4 & 2 & -2 \end{pmatrix}$$

gehörige Treppenmatrix ist

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Also sind a_1, a_2, a_3 linear unabhängig. Weil $\widehat{T}$ die charakteristischen Spaltenindizes $j(1) = 1$, $j(2) = 2$, $j(3) = 3$, $j(4) = 5$ hat, folgt, dass $\{a_1, a_2, a_3, b_2\}$ eine Basis von V ist.

11. Lineare Abbildungen

Sei K im folgenden immer ein Körper.

Definition 11.1. Seien V und W zwei K-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *linear*, falls

(L1)	$f(x+y) = f(x) + f(y)$	für alle x, $y \in V$,
(L2)	$f(\lambda \cdot x) = \lambda \cdot f(x)$	für alle $\lambda \in K$, $x \in V$.

Man nennt dann f manchmal auch genauer K-*linear*.

Bemerkung 11.2. Sei $f\colon V \to W$ eine lineare Abbildung zwischen den K-Vektorräumen V und W.

(1) Es gilt $f(0_V) = f(0_K \cdot 0_V) = 0_K \cdot f(0_V) = 0_W$. Für jedes $x \in V$ gilt $f(-x) = f((-1_K) \cdot x) = (-1_K) \cdot f(x) = -f(x)$.

(2) Sind $\lambda_1, \dots, \lambda_m \in K$ und $x_1, \dots, x_m \in V$, so gilt
$$f\Big(\sum_{i=1}^{m} \lambda_i x_i\Big) = \sum_{i=1}^{m} \lambda_i f(x_i).$$
(Der einfache Beweis per Induktion nach m.)

(3) Ist f bijektiv, so ist auch die Umkehrabbildung $f^{-1}\colon W \to V$ linear (und bijektiv).
Beweis: Seien x', $y' \in W$ und $\lambda \in K$. Dann gilt
$$f(f^{-1}(x') + f^{-1}(y')) = f(f^{-1}(x')) + f(f^{-1}(y')) = x' + y'$$
und
$$f(\lambda \cdot f^{-1}(x')) = \lambda \cdot f(f^{-1}(x')) = \lambda \cdot x',$$
und es folgt $f^{-1}(x' + y') = f^{-1}(x') + f^{-1}(y')$ und $f^{-1}(\lambda \cdot x') = \lambda \cdot f(x')$.

(4) Sei W' ein weiterer K-Vektorraum und $g\colon W \to W'$ linear. Dann ist $g \circ f\colon V \to W'$ linear.
Beweis: Für alle x, $y \in V$ und $\lambda \in K$ gilt
$$\begin{aligned} g \circ f(x+y) &= g(f(x+y)) = g(f(x) + f(y)) \\ &= g(f(x)) + g(f(y)) = g \circ f(x) + g \circ f(y), \\ g \circ f(\lambda \cdot x) &= g(f(\lambda \cdot x)) = g(\lambda \cdot (f(x))) \\ &= \lambda \cdot g(f(x)) = \lambda \cdot (g \circ f(x)). \end{aligned}$$

Beispiele 11.3. (1) Sei V ein K-Vektorraum. Dann ist die identische Abbildung $1_V\colon V \to V$ linear.

(2) Seien a, $b \in K$. Die Abbildung $f\colon K \to K$, $x \mapsto ax + b$ ist linear genau dann, wenn $b = 0$ ist.

(3) Der Körper $\mathbb{C}$ ist ein $\mathbb{C}$-Vektorraum und auch ein $\mathbb{R}$-Vektorraum. Die Abbildung $\gamma\colon \mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$ (komplexe Konjugation) ist $\mathbb{R}$-linear aber nicht $\mathbb{C}$-linear.

(4) Die Abbildung $\mathrm{sp}\colon \mathrm{M}(n; K) \to K$, $(\alpha_{ij}) \mapsto \sum_{i=1}^{n} \alpha_{ii}$ ist linear (vgl. Aufgabe V.2).

(5) Sei $A \in \mathrm{M}(m, n; K)$. Dann ist
$$f_A\colon K^n \to K^m,\ x \mapsto A \cdot x$$
eine lineare Abbildung. Dies folgt sofort aus Satz 6.6 (2) und (4). Wegen $f_A(e_j) = A \cdot e_j = A_{\bullet j}$ $(j = 1, \dots, n)$ kann man aus f_A die Matrix A zurückgewinnen.

(6) In der Analysis lernt man (später), dass die Menge $C^\infty(\mathbb{R})$ der beliebig oft differenzierbaren Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ ein $\mathbb{R}$-Vektorraum ist (mit den Verknüpfungen aus 8.4 (3)). Die Abbildung

$$D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}),\ f \mapsto f'$$

ist nach den Rechenregeln für Ableitungen linear.

Satz 11.4. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W.*

(1) *Für jeden Unterraum U von V ist $f(U) \stackrel{def}{=} \{f(x) \mid x \in U\}$ ein Unterraum von W.*
(2) *Für jeden Unterraum U' von W' ist das Urbild $f^{-1}(U') \stackrel{def}{=} \{x \in V \mid f(x) \in U'\}$ ein Unterraum von V.*
(3) *Für jede Teilmenge $X \subseteq V$ gilt* $\mathrm{Span}(f(X)) = f(\mathrm{Span}(X))$.

Beweis. (1) Es ist $0_W = f(0_V) \in f(U)$. Seien $x, y \in U$ und $\lambda \in K$. Dann gelten $x + y, \lambda x \in U$, und es folgt $f(x) + f(y) = f(x + y)$, $\lambda \cdot f(x) = f(\lambda \cdot x) \in f(U)$.

(2) Es ist $0_V \in f^{-1}(U')$, da $f(0_V) = 0_W \in U'$. Seien $x, y \in f^{-1}(U')$ und $\lambda \in K$. Dann gilt $f(x), f(y) \in U'$, also auch $f(x + y) = f(x) + f(y)$, $f(\lambda \cdot x) = \lambda \cdot f(x) \in U'$, also $x + y, \lambda \cdot x \in U'$.

(3) Dies folgt sofort aus der Gleichung

$$f\left(\sum_{i=1}^{m} \alpha_i x_i\right) = \sum_{i=1}^{m} \alpha_i f(x_i)$$

für alle $\alpha_1, \dots, \alpha_m \in K$ und $x_1, \dots, x_m \in X$, vgl. 11.2 (2). □

Definition 11.5. Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W.

(1) $\mathrm{Bild}(f) \stackrel{def}{=} f(V) = \{f(x) \mid x \in V\}$ heißt das *Bild* von f.
(2) $\mathrm{Kern}(f) \stackrel{def}{=} f^{-1}(\{0_W\}) = \{x \in V \mid f(x) = 0_W\}$ heißt der *Kern* von f.

Aus dem vorherigen Satz ergibt sich speziell:

Folgerung 11.6. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W.*

(1) $\mathrm{Bild}(f)$ *ist ein Unterraum von W.*
(2) $\mathrm{Kern}(f)$ *ist ein Unterraum von V.*

Offenbar ist eine lineare Abbildung $f\colon V \to W$ genau dann surjektiv, wenn $\mathrm{Bild}(f) = W$ gilt. Für die Injektivität gilt folgendes.

Satz 11.7 (Injektivitätskriterium). *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W. Dann gilt: f ist genau dann injektiv, wenn* $\mathrm{Kern}(f) = \{0\}$ *gilt.*

Beweis. (1) Sei f injektiv. Für alle $x \in V$ mit $x \neq 0_V$ ist dann $f(x) \neq f(0_V) = 0_W$, also $x \notin \mathrm{Kern}(f)$. Also $\mathrm{Kern}(f) = \{0_V\}$.

(2) Es gelte $\mathrm{Kern}(f) = \{0_V\}$. Seien $x, y \in V$ mit $f(x) = f(y)$. Dann gilt $f(x - y) = f(x) - f(y) = 0_W$, d. h. $x - y \in \mathrm{Kern}(f) = \{0_V\}$, also $x - y = 0_V$, d. h. $x = y$. □

Definition 11.8. Seien V und W zwei K-Vektorräume und $f\colon V \to W$ eine lineare Abbildung. Man bezeichnet $\dim(\mathrm{Bild}(f))$ auch als *Rang* von f und schreibt $\mathrm{rang}(f) \stackrel{def}{=} \dim(\mathrm{Bild}(f))$.

Beispiel 11.9. Sei $A \in \mathrm{M}(m,n;K)$ und $f_A\colon K^n \to K^m$, $x \mapsto A \cdot x$ die lineare Abbildung aus 11.3 (5). Dann gilt

$$\begin{aligned} \operatorname{rang}(f_A) &= \dim(\operatorname{Bild}(f_A)) = \dim\bigl(\operatorname{Span}(f_A(e_1),\dots,f_A(e_n))\bigr) \\ &= \dim\bigl(\operatorname{Span}(A\cdot e_1,\dots,A\cdot e_n)\bigr) = \dim\bigl(\operatorname{Span}(A_{\bullet 1},\dots,A_{\bullet n})\bigr) \\ &\overset{10.3}{=} \operatorname{rang}(A). \end{aligned}$$

Satz 11.10. *Seien V und W zwei K-Vektorräume.*

(1) *Ist $\{x_1,\dots,x_m\}$ ein Erzeugendensystem von V, so ist $\{f(x_1),\dots,f(x_m)\}$ ein Erzeugendensystem von* $\operatorname{Bild}(f)$.

(2) *Sei $\{b_1,\dots,b_n\}$ eine Basis von V. Dann gilt*
   (a) *f ist genau dann surjektiv, wenn $\{f(b_1),\dots,f(b_n)\}$ ein Erzeugendensystem von W ist.*
   (b) *f ist genau dann injektiv, wenn $f(b_1),\dots,f(b_n)$ linear unabhängig sind.*

Beweis. (1) Ist $\{x_1,\dots,x_m\}$ ein Erzeugendensystem von V, so gilt $V = \operatorname{Span}(x_1,\dots,x_m)$, und mit 11.4 (3) folgt

$$\operatorname{Bild}(f) = f(V) = f(\operatorname{Span}(x_1,\dots,x_m)) = \operatorname{Span}(f(x_1),\dots,f(x_m)).$$

(2) Es ist $\operatorname{Bild}(f) = f(V) = f(\operatorname{Span}(b_1,\dots,b_n) \overset{11.4}{=} \operatorname{Span}(f(b_1),\dots,f(b_n))$. Es folgt

$$f \text{ surjektiv } \Leftrightarrow \operatorname{Bild}(f) = W \Leftrightarrow \operatorname{Span}(f(b_1),\dots,f(b_n)) = W.$$

(b) $\Rightarrow$: Sei f injektiv. Seien $\alpha_1,\dots,\alpha_n \in K$ mit $\sum_{i=1}^n \alpha_i f(b_i) = 0_W$. Da f linear ist, folgt $f(\sum_{i=1}^n \alpha_i b_i) = 0_W = f(0_V)$. Aus der Injektivität von f folgt $\sum_{i=1}^n \alpha_i b_i = 0_V$. Da $b_1,\dots,b_n$ linear unabhängig sind, folgt $\alpha_1 = \dots = \alpha_n = 0$. Also sind $f(b_1),\dots,f(b_n)$ linear unabhängig.

$\Leftarrow$: Seien $f(b_1),\dots,f(b_n)$ linear unabhängig. Sei $x \in \operatorname{Kern}(f)$. Man kann schreiben $x = \sum_{i=1}^n \alpha_i b_i$ für gewisse $\alpha_1,\dots,\alpha_n \in K$. Es folgt $0_W = f(x) \overset{f \text{ lin.}}{=} \sum_{i=1}^n \alpha_i f(b_i)$, und es folgt $\alpha_1 = \dots = \alpha_n \in K$, und damit $x = 0$. Also gilt $\operatorname{Kern}(f) = \{0_V\}$, d. h. f ist injektiv. □

Der folgende wichtige Satz wird häufig auch *Dimensionsformel für lineare Abbildungen* genannt.

Satz 11.11 (Rangsatz). *Seien V und W zwei K-Vektorräume und $f\colon V \to W$ eine lineare Abbildung. Ist V endlichdimensional, so sind* $\operatorname{Kern}(f)$ *und* $\operatorname{Bild}(f)$ *endlichdimensional, und es gilt*

$$\dim(V) = \dim(\operatorname{Kern}(f)) + \dim(\operatorname{Bild}(f)),$$

also

$$\operatorname{rang}(f) = \dim(V) - \dim(\operatorname{Kern}(f)).$$

Beweis. Sei $n = \dim(V) < \infty$. Als Unterraum von V ist $\operatorname{Kern}(f)$ endlichdimensional (vgl. 9.19). Sei $d = \dim(\operatorname{Kern}(f))$, und sei $\{a_1,\dots,a_d\}$ eine Basis von $\operatorname{Kern}(f)$. Nach dem Basisergänzungssatz 9.19 gibt es $b_1,\dots,b_{n-d} \in V$, so dass $\{a_1,\dots,a_d,\, b_1,\dots,b_{n-d}\}$ eine Basis von V ist.

Behauptung: $\{f(b_1),\dots,f(b_{n-d})\}$ ist eine Basis von $\operatorname{Bild}(f)$. Es gilt

$$\begin{aligned} \operatorname{Bild}(f) &= f(V) = f(\operatorname{Span}(a_1,\dots,a_d,\, b_1,\dots,b_{n-d})) \\ &= \operatorname{Span}\bigl(f(a_1),\dots,f(a_d),\, f(b_1),\dots,f(b_{n-d})\bigr) \\ &\overset{(*)}{=} \operatorname{Span}\bigl(f(b_1),\dots,f(b_{n-d})\bigr), \end{aligned}$$

wobei $(*)$ wegen $f(a_1) = \dots f(a_d) = 0_W$ gilt. Also ist $\{f(b_1),\dots,f(b_{n-d})\}$ ein Erzeugendensystem von $\operatorname{Bild}(f)$.

Zu zeigen ist noch, dass $f(b_1), \dots, f(b_{n-d})$ linear unabhängig sind. Seien $\beta_1, \dots, \beta_{n-d} \in K$ mit $\sum_{j=1}^{n-d} \beta_j f(b_j) = 0_W$. Dann gilt $f\big(\sum_{j=1}^{n-d} \beta_j b_j\big) = 0_W$, d. h. $\sum_{j=1}^{n-d} \beta_j b_j \in \mathrm{Kern}(f)$. Es gibt also $\alpha_1, \dots, \alpha_d \in K$ mit

$$\sum_{j=1}^{n-d} \beta_j b_j = \sum_{i=1}^{d} \alpha_i a_i,$$

oder anders geschrieben

$$\sum_{j=1}^{n-d} \beta_j b_j + \sum_{i=1}^{d} (-\alpha_i) a_i = 0_V.$$

Da aber $a_1, \dots, a_d, b_1, \dots, b_{n-d}$ linear unabhängig sind, folgt $(\alpha_1 = \dots = \alpha_d =)\ \beta_1 = \dots = \beta_{n-d} = 0$.

Insbesondere folgt nun

$$\dim(\mathrm{Bild}(f)) = n - d = \dim(V) - \dim(\mathrm{Kern}(f)).$$

□

Satz 11.12. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W. Es gelte* $\dim(V) = \dim(W) < \infty$. *Dann sind folgende Aussagen äquivalent:*

(1) *f ist bijektiv.*
(2) *f ist injektiv.*
(3) *f ist surjektiv.*

Beweis.

$$\begin{aligned} f \text{ injektiv} &\overset{11.7}{\Leftrightarrow} \mathrm{Kern}(f) = \{0_V\} \overset{9.20}{\Leftrightarrow} \dim(\mathrm{Kern}(f)) = 0 \\ &\overset{11.11}{\Leftrightarrow} \mathrm{rang}(f) = \dim(V) \overset{9.20}{\Leftrightarrow} \mathrm{Bild}(f) = V \\ &\Leftrightarrow f \text{ surjektiv.} \end{aligned}$$

□

Satz 11.13 (Lineare Fortsetzung). *Seien V und W zwei K-Vektorräume. Sei $n =$* $\dim(V) < \infty$. *Sei $\{b_1, \dots, b_n\}$ eine Basis von V. Seien $w_1, \dots, w_n \in W$ beliebige Elemente. Dann gibt es genau eine lineare Abbildung $f\colon V \to W$ mit $f(b_i) = w_i$ für alle $i = 1, \dots, n$.*

Beweis. Jedes $x \in V$ hat genau eine Darstellung der Form $\sum_{i=1}^{n} \alpha_i \cdot b_i$ (mit $\alpha_1, \dots, \alpha_n \in K$). Man setzt

$$f(x) \overset{def}{=} \sum_{i=1}^{n} \alpha_i \cdot w_i.$$

Da f linear sein soll mit $f(b_i) = w_i$, ist klar, dass man f so definieren muss. Andererseits, ist f so definiert, so sieht man sofort, dass f tatsächlich linear ist und $f(b_i) = w_i$ gilt für $i = 1, \dots, n$. □

Definition 11.14.

(1) Eine bijektive, lineare Abbildung $f\colon V \to W$ zwischen K-Vektorräumen V und W heißt ein *Isomorphismus*.
(2) Zwei K-Vektorräume V und W heißen *isomorph*, falls es einen Isomorphismus $f\colon V \to W$ gibt. Man schreibt dann $V \simeq W$. (Offenbar definiert $\simeq$ eine Äquivalenzrelation auf der Klasse der K-Vektorräume.)

Zwei isomorphe K-Vektorräume kann man als *im wesentlichen gleich* ansehen. Ein Isomorphismus stellt nicht nur eine Bijektion zwischen den beiden Vektorräumen her, diese Bijektion "bewahrt" sogar die jeweiligen linearen Strukturen.

Satz 11.15. *Sei V ein endlichdimensionaler K-Vektorraum mit* $\dim(V) = n$. *Sei* $\mathcal{B} = \{b_1, \dots, b_n\}$ *eine Basis von V. Dann definiert*

$$\phi_{\mathcal{B}} \colon K^n \to V, \quad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \sum_{i=1}^{n} \alpha_i b_i$$

einen Isomorphismus von K-Vektorräumen.

Beweis. Nach Satz 11.13 gibt es genau eine lineare Abbildung $f \colon K^n \to V$ mit $f(e_i) = b_i$ für $i = 1, \dots, n$. Dies ist aber gerade die im Satz definierte Abbildung. Da $\{b_1, \dots, b_n\}$ eine Basis von V ist, folgt aus Satz 11.10 (2), dass f injektiv und surjektiv, also bijektiv ist. □

Satz 11.16. *Seien V und W endlichdimensionale K-Vektorräume. Dann sind äquivalent:*

(1) $V \simeq W$.
(2) $\dim(V) = \dim(W)$.

Beweis. (1)⇒(2): Sei $f \colon V \to W$ ein Isomorphismus. Sei $\dim(V) = n$ und $\{b_1, \dots, b_n\}$ eine Basis von V. Nach Satz 11.10 (2) ist $\{f(b_1), \dots, f(b_n)\}$ eine Basis von W mit n (verschiedenen!) Elementen, und daher gilt $\dim(W) = n = \dim(V)$.

(2)⇒(1): Es gelte $\dim(V) = n = \dim(V)$. Nach Satz 11.15 gibt es Isomorphismen $f \colon K^n \to V$ und $g \colon K^n \to W$. Dann ist $g \circ f^{-1} \colon V \to W$ ein Isomorphismus (vgl. 11.2 (3) und (4)). (Man kann natürlich auch Satz 11.13 auf V und W anwenden, ohne den "Umweg" über K^n zu gehen.) □

Bemerkung 11.17. (1) Seien V und W zwei K-Vektorräume. Dann ist $\mathrm{Abb}(V, W) = \{f \mid f \colon V \to W \text{ Abbildung}\}$ ein K-Vektorraum, mit den Verknüpfungen, die definiert sind durch

$$(f + g)(x) \overset{def}{=} f(x) + g(x) \qquad \text{für alle } x \in V,$$

$$(\lambda \cdot f)(x) \overset{def}{=} \lambda \cdot f(x) \qquad \text{für alle } \lambda \in K,\ x \in V$$

(für alle f, $g \in \mathrm{Abb}(V, W)$), vgl. Aufgabe VII.5; $\mathrm{Abb}(V, W) = W^V$.

Offenbar ist

$$\mathrm{Hom}_K(V, W) \overset{def}{=} \{f \in \mathrm{Abb}(V, W) \mid f \text{ ist linear}\}$$

ein Unterraum von $\mathrm{Abb}(V, W)$. (Vgl. Übungen.)

(2) Eine lineare Abbildung $f \colon V \to W$ nennt man auch *Homomorphismus* (von K-Vektorräumen).

12. Lineare Abbildungen und Matrizen

Sei K stets ein Körper. Ist V ein endlichdimensionaler K-Vektorraum, so werden wir im folgenden stillschweigend immer *geordnete Basen* $\mathcal{B} = (b_1, \dots, b_n)$ betrachten, d. h. $\{b_1, \dots, b_n\}$ ist eine Basis von V, und die Reihenfolge der Basiselemente ist festgelegt und darf nachträglich nicht ohne weiteres geändert werden.

12.1. Sei $f \colon V \to W$ eine lineare Abbildung zwischen den Vektorräumen V und W. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von W.

(1) Für jedes $j \in \{1, \dots, n\}$ gibt es eindeutig bestimmte $\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj} \in K$ mit

$$f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i.$$

Dies definiert eine Matrix $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$. Man schreibt

$$M_{\mathcal{B}}^{\mathcal{C}}(f) \overset{def}{=} A$$

und nennt dies die *Darstellungsmatrix* von f bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$. Man beachte, dass die Definition diesr Matrix stark von der Wahl der Basen $\mathcal{B}$ und $\mathcal{C}$ und deren Anordnung abhängt. (Das war auch schon bei der Abbildung $\phi_{\mathcal{B}}$ in 11.15 der Fall.)

Ist $V = W$ und $\mathcal{B} = \mathcal{C}$, so schreibt man abkürzend

$$M_{\mathcal{B}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(f).$$

(2) Sei $x \in V$. Dann gibt es eindeutig bestimmte $\xi_1, \dots, \xi_n \in K$ mit

$$x = \sum_{i=1}^{n} \xi_i v_i.$$

Man nennt

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n$$

den *Koordinatenvektor* von x bzgl. der Basis $\mathcal{B}$. Mit dem Isomorphismus $\phi_{\mathcal{B}} \colon K^n \to V$ aus 11.15 gilt also gerade

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = (\phi_{\mathcal{B}})^{-1}(x).$$

(3) Mit den Bezeichnungen aus (1) und (2) gilt

$$\begin{aligned} f(x) &= \sum_{j=1}^{n} \xi_j f(v_j) \\ &= \sum_{j=1}^{n} \xi_j \sum_{i=1}^{m} \alpha_{ij} w_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} \xi_j \right) w_i, \end{aligned}$$

also ist $A \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^m$ der Koordinatenvektor von $f(x) \in W$ bzgl. $\mathcal{C}$.

(4) Sei nun $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ irgendeine Matrix. Nach Satz 11.13 gibt es genau eine lineare Abbildung $f_A \colon V \to W$ mit

$$f_A(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i$$

für jedes $j \in \{1, \dots, n\}$. (Auch diese lineare Abbildung f_A ist von der Auswahl der Basen $\mathcal{B}$ und $\mathcal{C}$ abhängig; wir haben das in der Notation aber nicht deutlich gemacht.) Offenbar gilt $A = M_{\mathcal{B}}^{\mathcal{C}}(f_A)$.

Satz 12.2. *Seien V und W zwei K-Vektorräume und $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{C} = (w_1, \dots, w_m)$ Basen von V bzw. W. Die Abbildung*

$$\Phi = \Phi_{\mathcal{B}}^{\mathcal{C}} \colon \mathrm{Hom}_K(V, W) \to \mathrm{M}(m, n; K), \ f \mapsto M_{\mathcal{B}}^{\mathcal{C}}(f)$$

ist ein Isomorphismus von K-Vektorräumen.

Man beachte, dass dieser Isomorphismus von der Wahl der Basen $\mathcal{B}$ und $\mathcal{C}$ abhängt.

Beweis. Aus den vorherigen Betrachtungen ergibt sich, dass durch die Zuordnung $A \mapsto f_A$ eine Umkehrabbildung zu Φ definiert wird. Also ist Φ bijektiv. Seien $f, g \in \mathrm{Hom}_K(V, W)$ und $\alpha \in K$. Dann gilt

$$\Phi(f+g) = M_{\mathcal{B}}^{\mathcal{C}}(f+g) \overset{(*)}{=} M_{\mathcal{B}}^{\mathcal{C}}(f) + M_{\mathcal{B}}^{\mathcal{C}}(g) = \Phi(f) + \Phi(g)$$

und

$$\Phi(\alpha f) = M_{\mathcal{B}}^{\mathcal{C}}(\alpha f) \overset{(**)}{=} \alpha M_{\mathcal{B}}^{\mathcal{C}}(f) = \alpha \Phi(f).$$

Zu $(*)$: Ist $M_{\mathcal{B}}^{\mathcal{C}}(f) = A = (\alpha_{ij})$ und $M_{\mathcal{B}}^{\mathcal{C}}(g) = B = (\beta_{ij})$, so gilt $f(v_j) = \sum_{i=1}^m \alpha_{ij} w_i$ und $g(v_j) = \sum_{i=1}^m \beta_{ij} w_i$. Es folgt

$$(f+g)(v_j) = f(v_j) + g(v_j) = \sum_{i=1}^m \alpha_{ij} w_i + \sum_{i=1}^m \beta_{ij} w_i = \sum_{i=1}^m (\alpha_{ij} + \beta_{ij}) w_i,$$

also $M_{\mathcal{B}}^{\mathcal{C}}(f+g) = (\alpha_{ij} + \beta_{ij}) = A + B$. Analog folgt $(**)$. – Also ist Φ linear. □

Folgerung 12.3. *Seien V und W zwei K-Vektorräume mit* $\dim(V) = n$ *und* $\dim(W) = m$. *Dann gilt* $\dim(\mathrm{Hom}_K(V, W)) = m \cdot n$.

Bemerkung 12.4. Seien $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ und $B = (\beta_{ij}) \in \mathrm{M}(n, p; K)$. Seien $f_A \colon K^n \to K^m$, $x \mapsto Ax$ und $f_B \colon K^p \to K^n$, $y \mapsto By$ die zugehörigen linearen Abbildungen. Dann gilt

$$f_A \circ f_B = f_{A \cdot B}.$$

Beweis. Für jedes $x \in K^p$ gilt

$$f_A \circ f_B(x) = f_A(Bx) = A(Bx) = (AB)x = f_{AB}(x).$$

□

Folgerung 12.5. *$A \in \mathrm{M}(n; K)$ ist genau dann invertierbar, wenn die lineare Abbildung $f_A \colon K^n \to K^n$, $x \mapsto Ax$ bijektiv ist.*

Beweis. Gibt es $B \in \mathrm{M}(n; K)$ mit $BA = E_n = AB$, so folgt $f_A \circ f_B = f_{AB} = f_{E_n} = 1_{K^n} = f_{BA} = f_B \circ f_A$, also f_A bijektiv. Ist umgekehrt f_A bijektiv und g die Umkehrabbildung, so gibt es ein $B \in \mathrm{M}(n; K)$ mit $g = f_B$, und es folgt wie oben $f_{AB} = 1_{K^n} = f_{BA}$, und daraus folgt $AB = E_n = BA$, also A invertierbar. □

Ist $A \in \mathrm{M}(m, n; K)$, so schreiben wir für die Abbildung $f_A \colon K^n \to K^m$, $x \mapsto Ax$ im folgenden manchmal auch suggestiv $A\cdot$.

Satz 12.6. *Sei $f \colon V \to W$ eine lineare Abbildung zwischen den Vektorräumen V und W. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \ldots, w_m)$ eine Basis von W. Dann ist das folgende Diagramm linearer Abbildungen* kommutativ

$$\begin{array}{ccc} K^n & \xrightarrow{\phi_{\mathcal{B}}} & V \\ {\scriptstyle M_{\mathcal{B}}^{\mathcal{C}}(f)\cdot}\downarrow & & \downarrow{\scriptstyle f} \\ K^m & \xrightarrow{\phi_{\mathcal{C}}} & W, \end{array}$$

d. h. für jedes $x \in K^n$ gilt

$$f(\phi_{\mathcal{B}}(x)) = \phi_{\mathcal{C}}(M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot x).$$

Durch dieses kommutative Diagramm ist die Matrix $M_{\mathcal{B}}^{\mathcal{C}}(f)$ bestimmt.

Beweis. Es genügt, die Gleichheit für die Standardbasisvektoren $x = e_j \in K^n$ ($j = 1, \ldots, n$) nachzuprüfen. Sei $M_{\mathcal{B}}^{\mathcal{C}}(f) = A = (\alpha_{ij})$. Es gilt

$$\begin{aligned} f(\phi_{\mathcal{B}}(e_j)) &= f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i \\ &= \sum_{i=1}^{m} \alpha_{ij} \phi_{\mathcal{C}}(e_i') = \phi_{\mathcal{C}}\left(\sum_{i=1}^{m} \alpha_{ij} e_i'\right) \\ &= \phi_{\mathcal{C}}(A_{\bullet j}) = \phi_{\mathcal{C}}(A \cdot e_j) \\ &= \phi_{\mathcal{C}}\big(M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot e_j\big). \end{aligned}$$

(Hierbei bezeichne $e_1', \ldots, e_m'$ die Standardbasis von K^m.) Aus dem kommutativen Diagramm ergibt sich $M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot = (\phi_{\mathcal{C}})^{-1} \circ f \circ \phi_{\mathcal{B}}$, und dadurch ist die Matrix $M_{\mathcal{B}}^{\mathcal{C}}(f)$ eindeutig definiert. □

Folgerung 12.7. *Mit den Bezeichnungen aus dem vorherigen Satz gilt*

$$\operatorname{rang}(f) = \operatorname{rang}(M_{\mathcal{B}}^{\mathcal{C}}(f)).$$

Beweis. Der Isomorphismus $\phi_{\mathcal{C}}$ bildet $\operatorname{Bild}(M_{\mathcal{B}}^{\mathcal{C}}(f)\cdot)$ auf $\operatorname{Bild}(f)$ ab. Diese Bilder sind also isomorph, haben insbesondere dieselbe Dimension. Es folgt

$$\begin{aligned} \operatorname{rang}(f) &\overset{def}{=} \dim(\operatorname{Bild}(f)) = \dim(M_{\mathcal{B}}^{\mathcal{C}}(f)\cdot) \\ &\overset{def}{=} \operatorname{rang}(M_{\mathcal{B}}^{\mathcal{C}}(f)\cdot) \overset{11.9}{=} \operatorname{rang}(M_{\mathcal{B}}^{\mathcal{C}}(f)). \end{aligned}$$

□

Satz 12.8. *Seien U, V und W endlichdimensionale K-Vektorräume mit Basen $\mathcal{A}$, $\mathcal{B}$ bzw. $\mathcal{C}$. Seien $f\colon U \to V$ und $g\colon V \to W$ lineare Abbildungen. Dann gilt*

$$M_{\mathcal{A}}^{\mathcal{C}}(g \circ f) = M_{\mathcal{B}}^{\mathcal{C}}(g) \cdot M_{\mathcal{A}}^{\mathcal{B}}(f).$$

Beweis. Gelte $\dim(U) = p$, $\dim(V) = n$ und $\dim(W) = m$.

$$\begin{array}{ccc} K^p & \xrightarrow{\phi_{\mathcal{A}}} & U \\ {\scriptstyle M_{\mathcal{A}}^{\mathcal{B}}(f)\cdot}\downarrow & & \downarrow{\scriptstyle f} \\ K^n & \xrightarrow{\phi_{\mathcal{B}}} & V \\ {\scriptstyle M_{\mathcal{B}}^{\mathcal{C}}(g)\cdot}\downarrow & & \downarrow{\scriptstyle g} \\ K^m & \xrightarrow{\phi_{\mathcal{C}}} & W, \end{array}$$

Da die kleinen Teilquadrate kommutuieren, kommutiert auch das große Rechteck. Weil dies Rechteck auch mit $M_{\mathcal{A}}^{\mathcal{C}}(g \circ f)\cdot$ auf der linken Seite kommutiert, folgt die Behauptung aus der Eindeutigkeitsaussage in Satz 12.6. □

Folgerung 12.9. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen den endlichdimensionalen Vektorräumen V und W. Sei $\mathcal{B}$ eine Basis von V und $\mathcal{C}$ eine Basis von W. Genau dann ist f bijektiv (also ein Isomorphismus), wenn die Matrix $M_{\mathcal{B}}^{\mathcal{C}}(f)$ invertierbar ist. In dem Fall ist*

$$\big(M_{\mathcal{B}}^{\mathcal{C}}(f)\big)^{-1} = M_{\mathcal{C}}^{\mathcal{B}}(f^{-1}).$$

Beweis. Sei f bijektiv. Sei $g\colon W \to V$ die Umkehrabbildung zu f. Auch g ist linear. Es folgt $M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{C}}(f) = M_{\mathcal{B}}(1_V) \overset{klar}{=} E_n$, und ebenso $M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(g) = M_{\mathcal{B}}(1_W) = E_n$, wobei $n \overset{def}{=} \dim(V) = \dim(W)$ gilt. Es folgt, dass $M_{\mathcal{B}}^{\mathcal{C}}(f)$ invertierbar ist.

Sei umgekehrt $A = M_{\mathcal{B}}^{\mathcal{C}}(f)$ invertierbar. Sei $B = A^{-1}$. Nach Satz 12.2 gibt es eine lineare Abbildung $g\colon W \to V$ mit $B = M_{\mathcal{C}}^{\mathcal{B}}(g)$. Es folgt dann leicht $M_{\mathcal{C}}^{\mathcal{C}}(f \circ g) = E_n$ und $M_{\mathcal{B}}^{\mathcal{B}}(g \circ f) = E_n$, und daraus folgt, dass g die Umkehrabbildung zu f ist. □

Definition 12.10. Seien V ein K-Vektorraum. Seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}' = (v'_1, \dots, v'_n)$ Basen von V. Die Matrix

$$T_{\mathcal{B}}^{\mathcal{B}'} \overset{def}{=} M_{\mathcal{B}}^{\mathcal{B}'}(1_V) \in \mathrm{M}(n; K)$$

nennen wir die *Basiswechselmatrix* oder *Transformationsmatrix* bzgl. $\mathcal{B}$ und $\mathcal{B}'$.

Bemerkung 12.11. Es gilt $T_{\mathcal{B}}^{\mathcal{B}'} \in \mathrm{GL}(n; K)$ und $(T_{\mathcal{B}}^{\mathcal{B}'})^{-1} = T_{\mathcal{B}'}^{\mathcal{B}}$.

Beweis. Direkt aus Folgerung 12.9. □

Bemerkung 12.12. Es ist $T_{\mathcal{B}}^{\mathcal{B}'} = (\tau_{ij})$ die Matrix mit $w_j = \sum_{i=1}^{n} \tau_{ij} w'_i$ für jedes $j \in \{1, \dots, n\}$.

Beweis. Unmittelbar aus der Definition von $M_{\mathcal{B}}^{\mathcal{B}'}(1_V)$. □

Die Invertierbarkeit von $T_{\mathcal{B}}^{\mathcal{B}'}$ und die Formel für das Inverse ergibt sich auch hieraus im Verbund mit Satz 10.1.

Satz 12.13 (Transformationsformel). *Sei $f\colon V \to W$ eine lineare Abbildung zwischen den K-Vektorräumen V und W. Seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}' = (v'_1, \dots, v'_n)$ Basen von V, und seien $\mathcal{C} = (w_1, \dots, w_m)$ und $\mathcal{C}' = (w'_1, \dots, w'_m)$ Basen von W. Dann gilt*

$$M_{\mathcal{B}'}^{\mathcal{C}'}(f) = T_{\mathcal{C}}^{\mathcal{C}'} \cdot M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot (T_{\mathcal{B}}^{\mathcal{B}'})^{-1}.$$

Beweis. KOMMUTATIVES DIAGRAMM □

12.14 (Berechnung einer Basis des Bildes einer linearen Abbildung). Sei $f\colon V \to W$ eine lineare Abbildung zwischen den K-Vektorräumen V und W, sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von W. Sei $A = M_{\mathcal{B}}^{\mathcal{C}}(f)$ die Darstellungsmatrix von f bzgl. $\mathcal{B}$ und $\mathcal{C}$. Für jedes $j \in \{1, \dots, n\}$ gilt also $f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i$.

Es gilt $\mathrm{Bild}(f) = \mathrm{Span}(f(v_1), \dots, f(v_n))$. Sei T die zu A gehörige Treppenmatrix vom Rang r, seien $j(1), \dots, j(r)$ die charakteristischen Spaltenindizes. Dann gilt:

(a) $\dim(\mathrm{Bild}(f)) = \mathrm{rang}(f) \overset{12.7}{=} \mathrm{rang}(A) = r$.
(b) Es ist $\{f(v_{j(1)}), \dots, f(v_{j(r)})\}$ eine Basis von $\mathrm{Bild}(f)$.

Für die Aussage (b) schaut man sich nochmal den Beweis von Folgerung 12.7 an: Der Isomorphismus $\phi_{\mathcal{C}}$ bildet eine Basis von $\mathrm{Bild}(M_{\mathcal{B}}^{\mathcal{C}}(f)\cdot)$ (also einer Basis des durch die Spalten von A aufgespannten Unterraums in K^m; vgl. 9.10 zur Berechnung einer solchen) auf eine Basis von $\mathrm{Bild}(f)$ ab. Die Aussage (b) ergibt sich dann aus dem kommutativen Diagramm in Satz 12.6.

12.15 (Berechnung einer Basis des Kerns einer linearen Abbildung). Sei $f\colon V \to W$ eine lineare Abbildung zwischen den K-Vektorräumen V und W, sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von W. Sei $A = M_{\mathcal{B}}^{\mathcal{C}}(f)$ die Darstellungsmatrix von f bzgl. $\mathcal{B}$ und $\mathcal{C}$.

Nach dem Rangsatz gilt $\dim(\mathrm{Kern}(f)) = \dim(V) - \mathrm{rang}(f) = n - r$, wobei r wie in (1) der Rang der Matrix $A = M_{\mathcal{B}}^{\mathcal{C}}(f)$ ist.

Der Isomorphismus $\phi_{\mathcal{B}}$ aus dem kommutativen Diagramm in 12.6 bildet (eine Basis von) $\mathrm{Kern}(M_{\mathcal{B}}^{\mathcal{C}}(f)\cdot)$ auf (eine Basis von) $\mathrm{Kern}(f)$ ab. Es ist also eine Basis des Unterraums

$$R_A \overset{def}{=} \{x \in K^n \mid Ax = 0\}$$

von K^n zu berechnen. (Dies führt also zur Bestimmung der Lösungsmenge eines (homogenen) linearen Gleichungssystems, und dies wird im nächsten Abschnitt diskutiert.)

Sei $T = (\tau_{ij})$ die zu A gehörige Treppenmatrix vom Rang r und seien $j(1), \dots, j(r)$ die charakteristischen Spaltenindizes. Man beachte dass $Ax = 0$ genau dann gilt, wenn $Tx = 0$ ist. (Denn $T = PA$ mit $P \in \mathrm{GL}(m; K)$; $Tx = 0 \;\Leftrightarrow\; PAx = 0 \;\Leftrightarrow\; Ax =$

$P^{-1}PAx = 0$.) Für jedes $i \in \{1, \dots, r\}$ ist $T_{\bullet j(i)} = e_i$ der i-te Standardbasisvektor in K^m. Sei $J \overset{def}{=} \{1, \dots, n\} \setminus \{j(1), \dots, j(r)\}$. Es gilt $\#(J) = n - r$. Für jedes $j \in J$ gilt

$$T_{\bullet j} = \sum_{k=1}^{r} \tau_{kj} T_{\bullet j(k)}$$

und dann auch

$$A_{\bullet j} = \sum_{k=1}^{r} \tau_{kj} A_{\bullet j(k)}.$$

(Vgl. den Beweis von Satz 10.3.) Es gilt also

$$\sum_{s=1}^{n} \eta_s^{(j)} A_{\bullet s} = 0, \tag{12.1}$$

mit

$$\eta_s^{(j)} = \begin{cases} -1 & s = j, \\ 0 & s \in J,\ s \neq j, \\ \tau_{kj} & s = j(k). \end{cases}$$

Dies gilt für jedes der $n - r$ Elemente $j \in J$. Setze

$$y_j \overset{def}{=} \begin{pmatrix} \eta_1^{(j)} \\ \eta_2^{(j)} \\ \vdots \\ \eta_n^{(j)} \end{pmatrix} \in K^n$$

für jedes $j \in J$. Wegen (12.1) gilt also $Ay_j = 0$ für jedes $j \in J$. Man sieht sofort (wegen der -1 und sonst Nullen an den verschiedenen Stellen mit Index in J), dass $\{y_j \mid j \in J\}$ frei ist. Wegen

$$\#(J) = n - r = \dim(\{x \in K^n \mid Ax = 0\})$$

folgt, dass $\{y_j \mid j \in J\}$ eine Basis des Raumes $R_A = \{x \in K^n \mid Ax = 0\}$ ist.

Beispiel 12.16. Sei

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 & 7 \\ 2 & 0 & 4 & 0 & -2 & 4 \\ 0 & 1 & -3 & -1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 4 & 6 \end{pmatrix} \in \mathrm{M}(4, 6; \mathbb{R}).$$

Sei $f = f_A \colon \mathbb{R}^6 \to \mathbb{R}^4$ die lineare Abbildung mit $f(x) = Ax$ für alle $x \in \mathbb{R}^6$. Man berechnet mit dem Gauß-Algorithmus die Treppenmatrix von A:

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

Es ist also $r = \mathrm{rang}(A) = 3$, die charakteristischen Spaltenindizes sind $j(1) = 1$, $j(2) = 2$ und $j(3) = 4$. Mit den obigen Bezeichnungen ist also $J = \{3, 5, 6\}$. Es bildet dann

$$y_3 = \begin{pmatrix} 2 \\ -3 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ y_5 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \ y_6 = \begin{pmatrix} 2 \\ 3 \\ 0 \\ 3 \\ 0 \\ -1 \end{pmatrix}$$

eine Basis von $\mathrm{Kern}(f) = \{x \in \mathbb{R}^6 \mid Ax = 0\}$.

Man kann sich den letzten Schritt schematisch auch so vorstellen: In der Treppenmatrix stehen die Elemente $T[k, j(k)] = 1$ $(k = 1, \ldots, r)$ nicht auf der Hauptdiagonalen (abgesehen davon, dass T nicht quadratisch ist). Es werden formal Nullzeilen eingefügt, so dass die genannten Einsen auf der Hauptdiagonalen einer quadratischen $n \times n$-Matrix stehen. Das folgende Verfahren nennt man auch den *Entzerrungsalgorithmus*:

Zunächst streicht man die $m - r$ vielen unteren Nullzeilen in T und erhält

$$T' = \begin{pmatrix} 1 & 0 & 2 & 0 & -1 & 2 \\ 0 & 1 & -3 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{pmatrix}.$$

Danach fügt man $\#(J) = n-r$ viele Nullzeilen ein, und zwar so, dass man eine $n \times n$-Matrix erhält, wobei die Zeilenindizes der neuen Nullzeilen die Elemente aus J durchlaufen. Man erhält also im Beispiel

$$T'' = \begin{pmatrix} 1 & 0 & 2 & 0 & -1 & 2 \\ 0 & 1 & -3 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die 3., die 5. und die 6. Zeile sind also nun die Nullzeilen in T''. Schließlich werden die (eingefügten) Nullen auf der Hauptdiagonalen jeweils durch den Wert -1 ersetzt:

$$T''' = \begin{pmatrix} 1 & 0 & 2 & 0 & -1 & 2 \\ 0 & 1 & -3 & 0 & 2 & 3 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & \underline{0} & 0 & \underline{0} & \underline{-1} \end{pmatrix}.$$

Nun hat man $y_3 = T'''_{\bullet 3}$, $y_5 = T'''_{\bullet 5}$ und $y_6 = T'''_{\bullet 6}$. Also allgemein nimmt man die Spalten, bei denen sich durch das beschriebene Verfahren die Einträge -1 auf der Hauptdiagonalen ergeben haben.

Beispiel 12.17. Seien V und W zwei $\mathbb{R}$-Vektorräume. Sei $\mathcal{B} = (v_1, v_2, v_3, v_4)$ eine Basis von V und $\mathcal{C} = (w_1, w_2, w_3)$ eine Basis von W. Sei $f\colon V \to W$ die lineare Abbildung mit

$$f(v_1) = 2w_1 + 3w_2 + w_3,\ f(v_2) = 2w_1 + 2w_2,\ f(v_3) = w_2 + w_3,\ f(v_4) = 3w_2 + w_3.$$

Es ist dann

$$A = M^{\mathcal{C}}_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 2 & 0 & 1 \\ 3 & 2 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Die Treppenmatrix von A ist

$$T = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Man erhält mit dem Entzerrungsalgorithmus

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

und liest ab, dass

$$y_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \; y_4 = \begin{pmatrix} 1 \\ -1/2 \\ 0 \\ -1 \end{pmatrix}$$

eine Basis des Unterraums $\{x \in \mathbb{R}^4 \mid Ax = 0\}$ bildet. Eine Basis von $\mathrm{Kern}(f)$ erhält man durch Anwendung des Isomorphismus $\phi_{\mathcal{B}} \colon \mathbb{R}^4 \to V$, der den Standardbasisvektor e_i auf v_i schickt ($i = 1, 2, 3, 4$). Also ist

$$\{\phi_{\mathcal{B}}(y_3), \phi_{\mathcal{B}}(y_4)\} = \{v_1 - v_2 - v_3, v_1 - 1/2v_2 - v_4\}$$

eine Basis von $\mathrm{Kern}(f)$.

13. Lineare Gleichungssysteme

Sei K wie immer stets ein Körper.

13.1. (1) Ein *lineares Gleichungssystem* (über K) mit m Gleichungen und n Variablen (Unbestimmten) ist gegeben durch

$$\begin{array}{ccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + \cdots + & \alpha_{1n}x_n & = & b_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + \cdots + & \alpha_{2n}x_n & = & b_2 \\ \vdots & & & \vdots & & \vdots & \\ \alpha_{m1}x_1 & + & \alpha_{12}x_2 & + \cdots + & \alpha_{mn}x_n & = & b_m \end{array} \tag{13.1}$$

Hierbei sind $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{mn} \in K$ und $b_1, \ldots, b_m \in K$ vorgegeben. Gesucht sind $x_1, \ldots, x_n \in K$, so dass (13.1) erfüllt ist.

(2) Die Matrix $A = (\alpha_{ij}) \in \mathrm{M}(m, n; K)$ heißt die zugehörige *Koeffizientenmatrix*. Der Vektor $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$ heißt auch der *Ergebnisvektor*.

(3) Ein Vektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ ist eine *Lösung* des linearen Gleichungssystems (13.1), wenn (13.1) erfüllt ist. Ein $x \in K^n$ ist also genau dann eine Lösung von (13.1), wenn $Ax = b$ gilt. Das lineare Gleichungssystem $Ax = b$ heißt *lösbar*, falls es (mindestens) eine Lösung $x \in K^n$ besitzt.

(4) Die Matrix $(A \mid b) \in \mathrm{M}(m, n+1; K)$ heißt die *erweiterte Koeffizientenmatrix* zu (13.1).

(5) Das lineare Gleichungssystem $Ax = b$ heißt *homogen*, falls $b = 0$ gilt. Andernfalls heißt es *inhomogen*. Zum lineare Gleichnungssystem $Ax = b$ heißt $Ax = 0$ das zugehörige homogene System.

Beispiel 13.2.

$$\begin{array}{rcrcrcrl} 2x_1 & + & 3x_2 & + & 7x_3 & - & 8x_4 = & -7 \\ -2x_1 & + & 5x_2 & - & x_3 & + & 2x_4 = & 3 \\ 3x_1 & + & 3x_2 & + & 12x_3 & + & x_4 = & 2 \end{array}$$

ist ein lineares Gleichungssystem über $\mathbb{R}$ mit Keoffizientenmatrix

$$A = \begin{pmatrix} 2 & 3 & 7 & -8 \\ -2 & 5 & -1 & 2 \\ 3 & 3 & 12 & 1 \end{pmatrix},$$

Ergebnisvektor

$$b = \begin{pmatrix} -7 \\ 3 \\ 2 \end{pmatrix}$$

und erweiterter Koeffizientenmatrix

$$(A \mid b) = \left(\begin{array}{cccc|c} 2 & 3 & 7 & -8 & -7 \\ -2 & 5 & -1 & 2 & 3 \\ 3 & 3 & 12 & 1 & 2 \end{array}\right).$$

Es läßt sich kurz als $Ax = b$ schreiben.

SATZ 13.3. *Sei $A \in \mathrm{M}(m,n;K)$ und $b \in K^m$. Sei $r = \mathrm{rang}(A)$.*

(1) *Das lineare Gleichungssystem $Ax = b$ ist lösbar genau dann, wenn* $\mathrm{rang}(A) = \mathrm{rang}\big((A \mid b)\big)$ *gilt.*
(2) *Die Lösungsmenge $R_A \stackrel{def}{=} \{x \in K^n \mid Ax = 0\}$ des zugehörigen homogenen Systems $Ax = 0$ ist ein Unterraum von K^n der Dimension $n - r$.*
(3) *Sei x^* (irgend-) eine Lösung des linearen Gleichungssystems $Ax = b$. (Man nennt x^* eine* spezielle *Lösung.) Dann ist die Lösungsmenge des Systems $Ax = b$ gegeben durch*

$$\{x \in K^n \mid Ax = b\} = \{x^* + y \mid Ay = 0\} = x^* + R_A.$$

BEWEIS. Sei $f = f_A = A\cdot\colon K^n \to K^m$ die lineare Abbildung, die durch $f(x) = Ax$ gegeben ist.

(1) Das System $Ax = b$ ist lösbar $\Leftrightarrow$ es existiert ein $x \in K^n$ mit $b = Ax = f(x)$ $\Leftrightarrow$ $b \in \mathrm{Bild}(f) = \mathrm{Span}(f(e_1), \ldots, f(e_n)) = \mathrm{Span}(A_{\bullet 1}, \ldots, A_{\bullet n})$ $\Leftrightarrow$ $\mathrm{Span}(A_{\bullet 1}, \ldots, A_{\bullet n}, b) = \mathrm{Span}(A_{\bullet 1}, \ldots, A_{\bullet n})$ $\Leftrightarrow$ $\mathrm{rang}\big((A \mid b)\big) = \mathrm{rang}(A)$ gilt.

(2) Es ist klar, dass $R_A = \mathrm{Kern}(f)$ ein Unterraum von K^n ist. Es folgt

$$\dim(R_A) = \dim(\mathrm{Kern}(f)) = \dim(K^n) - \dim(\mathrm{Bild}(f)) = n - \mathrm{rang}(A) = n - r$$

mit dem Rangsatz.

(3) Es gilt $Ax = b \;\Leftrightarrow\; Ax = Ax^* \;\Leftrightarrow\; A(x - x^*) = 0 \;\Leftrightarrow\; x - x^* \in R_A \;\Leftrightarrow\; x \in x^* + R_A$. □

BEMERKUNG 13.4. (1) Wie man eine Basis des Lösungsraums R_A eines homogenen linearen Gleichungssystems $Ax = 0$ bestimmt, wurde am Ende des vorherigen Abschnitts erläutert.

(2) Zum Bestimmen der Lösungsmenge $\{x \in K^n \mid Ax = b\}$ eines inhomogenen Systems $Ax = b$ geht man wie folgt vor:

(a) Man bestimmt eine spezielle Lösung x^* des Systems $Ax = b$.
(b) Man bestimmt eine Basis $\{y_j \mid j \in J\}$ des zugehörigen homogenen Systems $Ax = 0$.

Die Lösungsmenge besteht dann aus allen Elementen der Form

$$x = x^* + \sum_{j \in J} a_j y_j$$

mit eindeutig bestimmten $a_j \in K$ ($j \in J$). Während das Vorgehen für (b) in (1) schon beschrieben ist, müssen wir noch beschreiben, wie man (systematisch!) eine spezielle Lösung x^* bekommt. Dies geschieht mit einer Variante des Entzerrungsalgorithmus, angewendet auf die erweiterte Koeffizientenmatrix $(A \mid b)$.

13.5 (Bestimmung der Lösungsmenge von $Ax = b$). Sei $A \in \mathrm{M}(m,n;K)$ und $b \in K^m$. Sei $\widetilde{A} = (A \mid b)$. Sei $\widetilde{T}$ die zu $\widetilde{A}$ gehörige Treppenmatrix. Es gibt ein $P \in \mathrm{GL}(m;K)$ mit $\widetilde{T} = P \cdot \widetilde{A} = (P \cdot A \mid P \cdot b) \stackrel{def}{=} (T \mid c)$. Es ist klar, dass T die Treppenmatrix zu A ist. Daher gilt $\mathrm{rang}(A) \leq \mathrm{rang}\big((A \mid b)\big) \leq \mathrm{rang}(A) + 1$.

Gilt $\operatorname{rang}(A) < \operatorname{rang}\big((A \mid b)\big)$, so ist die Lösungsmenge des linearen Gleichungssystems $Ax = b$ leer. Gilt $\operatorname{rang}(A) = \operatorname{rang}\big((A \mid b)\big)$, so ist die Lösungsmenge nichtleer. Wir nehmen dies nun an. Sei r dieser gemeinsame Rang.

Man wendet nun den Entzerrungsalgorithmus auf T wie in 12.16 beschrieben an, und erweitert das Einfügen von Nullzeilen auf die erweiterte Matrix $\widetilde{T}$, also streicht auch in $c = P \cdot b = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$ die untersten $m - r$ Nullen, und fügt in c dann $n - r$ Nullen ein, die nachher an den Stellen j (mit $j \in J$) stehen. Man definiert also $x^* \in K^n$ durch

$$x^*[i] = \begin{cases} c_k & i = j(k), \\ 0 & i \in J \end{cases}.$$

Es gilt damit offenbar $Ax^* = b$. Denn seien $j(1), \ldots, j(r)$ die charakteristischen Spaltenidizes von $\widetilde{A} = (A \mid b)$. Dann gilt nach Voraussetzung $j(1), \ldots, j(r) \in \{1, \ldots, n\}$. Es folgt

$$c = \sum_{k=1}^{r} c_k \cdot \widetilde{T}_{\bullet j(k)} = \sum_{k=1}^{r} c_k \cdot T_{\bullet j(k)},$$

und Multiplikation von links mit P^{-1} liefert

$$b = \sum_{k=1}^{r} c_k \cdot A_{\bullet j(k)} = \sum_{j=1}^{n} x^*[j] \cdot A_{\bullet j} = A \cdot x^*.$$

Also ist das so definierte x^* eine spezielle Lösung von $Ax = b$.

Mit dieser erweiterten Variante des Entzerrungsalgorithmus berechnet man also eine spezielle Lösung x^* von $Ax = b$ und eine Basis von R_A gleichzeitig.

Beispiel 13.6. Es soll untersucht werden, ob das folgende lineare Gleichungssystem über $\mathbb{R}$ lösbar ist. Falls ja, so soll die Lösungsmenge bestimmt werden.

$$\begin{array}{rrrrrcr} 2x_1 & +x_2 & +2x_3 & +x_4 & -2x_5 & = & 5 \\ & x_2 & +2x_3 & +2x_4 & & = & 0 \\ x_1 & & & +x_4 & -x_5 & = & 4 \\ -x_1 & -2x_2 & -4x_3 & +x_4 & +x_5 & = & 2 \end{array}$$

Sei also

$$A = \begin{pmatrix} 2 & 1 & 2 & 1 & -2 \\ 0 & 1 & 2 & 2 & 0 \\ 1 & 0 & 0 & 1 & -1 \\ -1 & -2 & -4 & 1 & 1 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 5 \\ 0 \\ 4 \\ 2 \end{pmatrix}.$$

Die erweiterte Koeffizientenmatrix ist

$$\widetilde{A} = (A \mid b) = \left(\begin{array}{rrrrr|r} 2 & 1 & 2 & 1 & -2 & 5 \\ 0 & 1 & 2 & 2 & 0 & 0 \\ 1 & 0 & 0 & 1 & -1 & 4 \\ -1 & -2 & -4 & 1 & 1 & 2 \end{array}\right).$$

Die hierzu gehörige Treppenmatix ist

$$\widetilde{T} = (T \mid c) = \left(\begin{array}{rrrrr|r} 1 & 0 & 0 & 0 & -1 & 3 \\ 0 & 1 & 2 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Hierbei ist der linke Teil T die Treppenmatrix zu A, und man sieht

$$\operatorname{rang}\big((A \mid b)\big) = 3 = \operatorname{rang}(A).$$

Also ist das lineare Gleichnugssystem $Ax = b$ lösbar. Das oben beschriebene Entzerrungsverfahren macht nun folgendes: Zuerst wird in der Treppenmatrix $\widetilde{T}$ die unterste Nullzeile gelöscht, und dann werden $2 = 5 - 3 = n - r$ Nulzeilen eingefügt, die als Zeilenindex die Spaltenindizes $j \in \{1, \dots, 5\}$ bekommen, die keine charakteristischen sind, hier also $j = 3, 5$; zuletzt werden die eingefügten Nullen auf der Hauptdiagonalen im linken Teil durch Einträge -1 ersetzt. Man erhält:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -1 & 3 \\ 0 & 1 & 2 & 0 & 0 & -2 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right).$$

Es ist dann

$$x^* \stackrel{def}{=} \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

eine spezielle Lösung von $Ax = b$, und die zwei Vektoren

$$\begin{pmatrix} 0 \\ 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

bilden eine Basis des Lösungsraums R_A des zugehörigen homogenen Systems. Genau die Elemente der Form

$$x = \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

mit (eindeutig bestimmten) $\alpha, \beta \in \mathbb{R}$ sind Lösungen des linearen Gleichungssystems $Ax = b$.

KAPITEL 4

Determinanten

14. Der Begriff einer Determinantenfunktion

Sei stets K ein Körper und $n \in \mathbb{N}$ eine natürliche Zahl.

14.1. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}(2; K)$ und $b = \begin{pmatrix} e \\ f \end{pmatrix} \in K^2$. Wir wollen das lineare Gleichungssystem $A \begin{pmatrix} x \\ y \end{pmatrix} = b$ lösen. Dazu berechnen wir die Treppenmatrix von

$$\left(\begin{array}{cc|c} a & b & e \\ c & d & f \end{array} \right).$$

Sei etwa $a \neq 0$. Dann bekommen wir

$$\left(\begin{array}{cc|c} 1 & b/a & e/a \\ 0 & d - bc/a & f - ce/a \end{array} \right),$$

bzw.

$$\left(\begin{array}{cc|c} 1 & b/a & e/a \\ 0 & ad - bc & af - ce \end{array} \right),$$

Die Frage ist nun, ob $ad - bc \neq 0$ ist oder nicht. Wenn nicht, gibt es keine eindeutige Lösung. Wenn doch, dann gibt es eine eindeutige Lösung, die gegeben ist durch

$$y = \frac{af - ec}{ad - bc} \text{ und } x = \frac{ed - bf}{ad - bc}.$$

Definiert man also für eine beliebige Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}(2; K)$ deren *Determinante*

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

so erhält man

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \quad \text{und} \quad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

als eindeutige Lösung, sofern $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$ gilt. Die Voraussetzung $a \neq 0$ galt dabei ohne Einschränkung. (Dies ist der einfachste Fall der sog. Cramerschen Regel.) Man sieht anhand dieser Argumente auch, dass $A \in \mathrm{M}(2; K)$ genau dann invertierbar ist, wenn $\det(A) \neq 0$ gilt.

14.2. Seien $v = \begin{pmatrix} a \\ b \end{pmatrix}$, $w = \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2$. Wir betrachten dass von v und w aufgespannte Parallelogramm:

ZEICHNUNG

Man sieht, dass die Fläche dieses Paralellogramms gegeben ist durch $ad - bc$:

ZEICHNUNG

(Man kann dies auch sauber beweisen, z. B. mit Hilfe von Polarkoordinaten in $\mathbb{C}$.) Es hat also $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ auch eine geometrische Interpretation.

14.3. Mit beiden oben vorgestellten Interpretationen von $\det(A)$ für $A \in \mathrm{M}(n;K)$ (im zweiten Fall $K = \mathbb{R}$) sieht man sofort, dass $\det\colon \mathrm{M}(n;K) \to K$ folgende Eigenschaften besitzt. Sei $A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathrm{M}(2;K)$ mit den Zeilen a_1, $a_2 \in K^2$. Dann gilt:

(D1) (i) $\det(\begin{bmatrix} a_1 + a_1' \\ a_2 \end{bmatrix}) = \det(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \det(\begin{bmatrix} a_1' \\ a_2 \end{bmatrix}$ und

$\det(\begin{bmatrix} a_1 \\ a_2 + a_2' \end{bmatrix}) = \det(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}) + \det(\begin{bmatrix} a_1 \\ a_2' \end{bmatrix}$ für alle $\lambda \in K$; und

(ii) $\det(\begin{bmatrix} \lambda a_1 \\ a_2 \end{bmatrix}) = \lambda \det(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix})$ und

$\det(\begin{bmatrix} a_1 \\ \lambda a_2 \end{bmatrix}) = \lambda \det(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix})$ für alle $\lambda \in K$.

(D2) $\det(\begin{pmatrix} a_1 \\ a_1 \end{pmatrix}) = 0$.

(D3) $\det(E_2) = 1$.

Definition 14.4. Eine Abbildung $d\colon \mathrm{M}(n;K) \to K$ heißt eine *Determinantenfunktion*, falls gilt

(D1) d ist *linear in jeder Zeile*: Für jedes $i \in \{1,\dots,n\}$ gilt:

(i) Ist $a_i = a_i' + a_i''$, so gilt $d(\begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix}) = d(\begin{pmatrix} \vdots \\ a_i' \\ \vdots \end{pmatrix}) + d(\begin{pmatrix} \vdots \\ a_i'' \\ \vdots \end{pmatrix})$.

(ii) Ist $a_i = \lambda a_i'$ ($\lambda \in K$, so gilt $d(\begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix}) = \lambda d(\begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix})$.

(D2) d ist *alternierend*: Hat A zwei gleiche Zeilen, so gilt $d(A) = 0$.

(D3) d ist *normiert*: $d(E_n) = 1$.

(Man kann das Axiomensystem etwas abschwächen, vgl. das Buch von Falko Lorenz.)

Im folgenden wird gezeigt, dass es zu jedem $n \in \mathbb{N}$ genau eine Determinantenfunktion $d\colon \mathrm{M}(n;K) \to K$ gibt (Existenz und Eindeutigkeit). Deswegen werden wir auch immer die Bezeichnung $d(A) = \det(A)$ benutzen, sowie die Schreibweise

$$\det(A) = |A| = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

für $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$.

Für $n = 2$ haben wir die Existenz einer Determinantenfunktion schon gesehen. Für $n = 1$ hat man also einzige Möglichkeit $d((a)) = a =: \det((a))$. Um allgemein die Existenz und Eindeutigkeit einer solchen Funktion zu etablieren, leiten wir erst noch weitere Eigenschaften einer Determinantenfunktion her.

Satz 14.5. *Sei* $\det\colon \mathrm{M}(n;K) \to K$ *eine Determinantenfunktion. Dann gelten (neben (D1), (D2) und (D3) auch) die folgenden Eigenschaften:*

(D4) *Für jedes* $A \in \mathrm{M}(n;K)$ *und* $\lambda \in K$ *gilt* $\det(\lambda A) = \lambda^n \det(A)$.

(D5) *Besitzt* A *eine Nullzeile, so ist* $\det(A) = 0$.

(D6) *Entsteht B aus A durch Vertauschung zweier verschiedener Zeilen, so gilt* $\det(B) = -\det(A)$.

(D7) *Ist $\lambda \in K$ und entsteht B aus A durch Addition des λ-fachen der j-ten Zeile zur i-ten Zeile ($j \neq i$), so gilt* $\det(B) = \det(A)$.

(D8) *Ist A eine* obere Dreieckmatrix, *d. h. gilt*

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ 0 & 0 & \alpha_{33} & \dots & \alpha_{3n} \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & \alpha_{nn} \end{pmatrix},$$

so gilt $\det(A) = \alpha_{11} \cdot \alpha_{22} \cdot \ldots \cdot \alpha_{nn}$. *Analoges gilt für eine* untere Dreiecksmatrix.

(D9) *Ist A von der Form*

$$A = \left(\begin{array}{c|c} A_1 & C \\ \hline 0 & A_2 \end{array} \right),$$

wobei A_1 und A_2 quadratische Matrizen (kleinerer Formate) sind, so gilt $\det(A) = \det(A_1) \cdot \det(A_2)$. *Entsprechendes gilt, wenn die Nullmatrix 0 oben rechts steht.*

(D10) *Ist $A \in \mathrm{M}(n; K)$, so gilt:*

$$\det(A) = 0 \quad \Leftrightarrow \quad \operatorname{rang}(A) < n.$$

(D11) *(Determinanten-Multiplikationssatz) Für alle A, $B \in \mathrm{M}(n; K)$ gilt*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Insbesondere: Ist A invertierbar, so ist

$$\det(A^{-1}) = \det(A)^{-1}.$$

Beweis. (D4) folgt direkt aus Teil (ii) von (D1).

(D5) folgt ebenfalls aus Teil (ii) von (D1), denn man kann aus einer Nullzeile den Faktor 0 herausziehen.

(D6) Entstehe B aus A durch Vertauschung der beiden Zeilen a_i und a_j (mit $i < j$). Dann gilt

$$\begin{aligned}
\det(A) + \det(B) &= \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} \\
&\overset{(D2)}{=} \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \\
&= \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} \\
&\overset{(D1)(i)}{=} \det\begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} \overset{(D2)}{=} 0,
\end{aligned}$$

und es folgt $\det(B) = -\det(A)$.

(D7) Sei etwa $i < j$. Es ist

$$\det(B) = \det\begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{(D1)}{=} \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \lambda \det\begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{(D2)}{=} \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det(A).$$

(D8) Sind alle $\alpha_{ii} \neq 0$ ($i = 1, \ldots, n$), so erhält man durch elementare Zeilenumformungen (vom Typ III Additionsmatrizen) und durch wiederholte Anwendung von (D7)

$$\det(A) \overset{(D7)}{=} \begin{vmatrix} \alpha_{11} & & & \\ & \alpha_{22} & & \\ & & \ddots & \\ & & & \alpha_{nn} \end{vmatrix} \overset{(D1)(ii)}{=} \alpha_{11} \cdot \ldots \cdot \alpha_{nn} \cdot \det(E_n) \overset{(D3)}{=} \alpha_{11} \cdot \ldots \cdot \alpha_{nn}.$$

Ist ein $\lambda_i = 0$, so sei i maximal damit. Durch Anwendung von elementaren Zeilenumformungen (vom Typ III), macht man alle Einträge rechts von α_{ii} zu null, macht also die i-te Zeile zur Nullzeile, und wiederholte Anwendung von (D7) und dann (D5) liefert $\det(A) = 0 = \alpha_{11} \cdot \ldots \cdot \alpha_{nn}$.

(D9) Durch elementare Zeilenumformungen (Typ II und Vertauschungen und Additionen) bringt man zunächst A_1 auf obere Dreiecksgestalt A_1'. Dabei wird C zu C', A_2 bleibt unverändert. Sei k dabei die Anzahl der vorgenommenen Vertauschungen. Nun wird A_2 ebenso durch elementare Zeilenumformungen (Typ II und III) auf Dreiecksgestalt A_2' gebracht. Sei ℓ die Anzahl der dabei vorgenommenen Vertauschungen. Sei

$$A' = \left(\begin{array}{c|c} A_1' & C' \\ \hline 0 & A_2' \end{array}\right).$$

Es geht also A' aus A durch Additionen und $k+\ell$ Vertauschungen hervor, und wegen (D6) und (D7) folgt daher $\det(A) = (-1)^{k+\ell}\det(A')$, und ebenso $\det(A_1') = (-1)^k\det(A_1)$, $\det(A_2') = (-1)^\ell\det(A_2)$. Da A', A_1' und A_2' obere Dreiecksmatrizen sind, folgt aus (D8) unmittelbar $\det(A') = \det(A_1')\cdot\det(A_2')$. Insgesamt:

$$\begin{aligned} \det(A) &= (-1)^{k+\ell}\det(A') = (-1)^k\det(A_1')(-1)^\ell\det(A_2') \\ &= \det(A_1)\cdot\det(A_2). \end{aligned}$$

(D10) Durch elementare Zeilenumformungen vom Typ II und III bringt man A auf obere Dreiecksgestalt

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & * & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

und es gilt dann (nach (D5), (D6), (D7) und (D8)) $\det(A) = \pm\lambda_1\cdot\ldots\cdot\lambda_n$. Es folgt

$$\det(A) = 0 \quad\Leftrightarrow\quad \text{es gibt } i \text{ mit } \lambda_i = 0 \quad\Leftrightarrow\quad \operatorname{rang}(A) < n.$$

(D11) Gilt $\operatorname{rang}(A) < n$, so gilt auch $\operatorname{rang}(A\cdot B) < n$. (Vgl. dazu Aufgabe VIII. 5. Alternativ: $\operatorname{rang}(A\cdot B) = \operatorname{rang}(f_A\circ f_B)$ nach 11.9 und 12.4, und es ist klar, dass $\operatorname{Bild}(f_A\circ f_B)\subseteq\operatorname{Bild}(f_A)$ gilt.) Wegen (D10) folgt in diesem Fall also $\det(AB) = 0 = 0\cdot\det(B) = \det(A)\det(B)$.

Gelte nun $\operatorname{rang}(A) = n$, also $A\in\mathrm{GL}(n;K)$. Nach 7.24 gibt es Elementarmatrizen $P_1,\ldots,P_s\in\mathrm{M}(n;K)$ mit $A = P_1\cdot\ldots\cdot P_s$. Per Induktion genügt es also zu zeigen, dass $\det(PB) = \det(P)\det(B)$ gilt für eine Elementarmatrix P. Es gibt die drei Möglichkeiten

(i) $P = D_k(\lambda)$. Es ist $\det(D_k(\lambda)) \overset{(D8)}{=} \lambda$. Es folgt $\det(D_k(\lambda)\cdot B) \overset{(D1)}{=} \lambda\det(B) = \det(D_k(\lambda))\det(B)$.

(ii) $P = V_{k\ell}$. Es ist $\det(V_{k\ell}) \overset{(D6)}{=} -1$. Es folgt $\det(V_{k\ell}\cdot B) \overset{(D6)}{=} -\det(B) = \det(V_{kl})\cdot\det(B)$.

(iii) $P = W_{k\ell}(\lambda)$. Es ist $\det(W_{k\ell}(\lambda)) \overset{(D8)}{=} 1$. Es folgt $\det(W_{k\ell}(\lambda)\cdot B) \overset{(D7)}{=} \det(B) = \det(W_{kl}(\lambda))\cdot\det(B)$.

□

FOLGERUNG 14.6. *Seien d, $d'\colon \mathrm{M}(n;K)\to K$ zwei Determinantenfunktionen. Dann gilt $d = d'$.*

BEWEIS. Sei $A\in\mathrm{M}(n;K)$. Wie im Beweis des vorherigen Satzes gezeigt, bringt man mit elementaren Zeilenumformungen vom Typ II und III A auf obere Dreiecksgestalt, und die sowohl $d(A)$ wie auch $d'(A)$ ergebens sich aus dem Produkt der Hauptdiagonalelemente dieser Matrix. □

Es bleibt immer noch die Existenz einer Determinantenfunktionen (bzw. Der Determinanten zu zeigen). Dies geschieht erst im nächsten Abschnitt.

Beispiel 14.7. Für

$$A = \begin{pmatrix} 0 & 2 & -1 & 4 \\ 2 & 1 & 3 & -5 \\ -1 & 4 & 2 & 1 \\ 3 & 0 & 2 & 0 \end{pmatrix} \in \mathrm{M}(4;\mathbb{R})$$

hat man

$$\begin{aligned}
|A| &= -\begin{vmatrix} -1 & 4 & 2 & 1 \\ 2 & 1 & 3 & -5 \\ 0 & 2 & -1 & 4 \\ 3 & 0 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & -2 & -1 \\ 2 & 1 & 3 & -5 \\ 0 & 2 & -1 & 4 \\ 3 & 0 & 2 & 0 \end{vmatrix} \\
&= \begin{vmatrix} 1 & -4 & -2 & -1 \\ 0 & 9 & 7 & -3 \\ 0 & 2 & -1 & 4 \\ 0 & 12 & 8 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -4 & -2 & -1 \\ 0 & 9 & 7 & -3 \\ 0 & 2 & -1 & 4 \\ 0 & 12 & 8 & 3 \end{vmatrix} \\
&= -\begin{vmatrix} 1 & -4 & -2 & -1 \\ 0 & 2 & -1 & 4 \\ 0 & 9 & 7 & -3 \\ 0 & 12 & 8 & 3 \end{vmatrix} = -\begin{vmatrix} 1 & -4 & -2 & -1 \\ 0 & 2 & -1 & 4 \\ 0 & 0 & 23/2 & -21 \\ 0 & 0 & 14 & -21 \end{vmatrix} \\
&= -\begin{vmatrix} 1 & -4 & -2 & -1 \\ 0 & 2 & -1 & 4 \\ 0 & 0 & 23/2 & -21 \\ 0 & 0 & 0 & 105/23 \end{vmatrix} = -(1 \cdot 2 \cdot 23/2 \cdot 105/23) \\
&= -105.
\end{aligned}$$

Bemerkung 14.8. Damit hat man ein sehr effizientes Verfahren an der Hand, die Determinante (Existenz momentan unterstellt) eines $A \in \mathrm{M}(n;K)$ zu berechnen: Mit elementaren Zeilenumformungen vom Typ II und III erhält man wie im Beispiel zuvor eine obere Dreiecksmatrix

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & * & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Ist k die Anzahl der dabei durchgeführten Vertauschungen, so gilt

$$|A| = (-1)^k \cdot \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n.$$

Bemerkung 14.9. Obige axiomatische Einführung der Determinante geht auf Karl Weierstraß (1815–1897) zurück, der 1834 sein Abitur am Theodorianum in Paderborn machte. Eine alternative Einführung geht über die Formel von Leibniz, die wir im nächsten Abschnitt behandelt. Mit ihr wird auch die Existenz einer Determinantenfunktion gezeigt.

15. Die Formel von Leibniz

Sei K stets ein Körper und $n \in \mathbb{N}$ eine natürliche Zahl.

15.1. In Abschnitt 2 haben wir schon die Gruppe S_n der Permutationen der Menge $X = \{1,\ldots,n\}$ kennengelernt. Also

$$S_n = \{\sigma \mid \sigma\colon \{1,\ldots,n\} \to \{1,\ldots,n\} \text{ bijektiv}\}.$$

Elemente $\sigma \in S_n$ schreiben wir in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Die Verknüpfung auf S_n ist die Komposition von Abbildungen, und damit gilt

$$\begin{aligned}\tau\sigma \overset{def}{=} \tau\circ\sigma &= \begin{pmatrix}1 & 2 & \dots & n\\ \tau(1) & \tau(2) & \dots & \tau(n)\end{pmatrix}\circ\begin{pmatrix}1 & 2 & \dots & n\\ \sigma(1) & \sigma(2) & \dots & \sigma(n)\end{pmatrix}\\ &= \begin{pmatrix}1 & 2 & \dots & n\\ \tau(\sigma(1)) & \tau(\sigma(2)) & \dots & \tau(\sigma(n))\end{pmatrix}.\end{aligned}$$

Neutrales Element ist die triviale Permutation, gegeben durch die identische Abbildung

$$\varepsilon = 1_{\{1,\dots,n\}} = \begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix}.$$

Eine Permutation $\tau \in S_n$, die Elemente $i,\, j \in \{1,\dots,n\}$ (mit $i \neq j$) vertauscht, und alle anderen Elemente festhält, nennt man eine *Transposition*, und schreibt $\tau = \tau(i,j) = (i\ j)$. Also

$$\tau(k) = \begin{cases} k & k \neq i,\, j,\\ j & k = i,\\ i & k = j.\end{cases}.$$

Die folgende Aussage wurde schon in 2.7 formuliert. Nun folgt auch der Beweis.

Bemerkung 15.2. S_n besteht aus $n! = 1\cdot 2\cdot\ldots\cdot n$ Elementen.

Beweis. Für $n = 1$ (oder auch $n = 0$) ist die Aussage klar; in diesem Fall hat man nur die identsiche Abbildung. Sei nun $n \geq 2$ und die Aussage für $n-1$ bereits gezeigt. Sei $\sigma \in S_n$.

1. Fall: Es gilt $\sigma(n) = n$. Dann gilt $\sigma(i) \in \{1,\dots,n-1\}$, d. h. es ist $\sigma' \overset{def}{=} \sigma|_{\{1,\dots,n-1\}} \in S_{n-1}$. Für σ' gibt es nach Induktionsvoraussetzung $(n-1)!$ Möglichkeiten, also auch für σ selbst.

2. Fall: $\sigma(n) \neq n$. Sei $\sigma(n) = i \neq n$. Dann gilt für $\sigma' \overset{def}{=} (i\ n)\circ\sigma$ nun $\sigma'(n) = n$. Nach dem 1. Fall gibt es für solche σ' nun $(n-1)!$ Möglichkeiten, also auch für σ. (Man beachte $(i\ n)\circ\sigma_1 = (i\ n)\circ\sigma_2 \;\Leftrightarrow\; \sigma_1 = \sigma_2$.) Nun gibt es für i $(\neq n)$ aber $n-1$ Möglichkeiten, und daher gibt es hier im 2. Fall $n-1$ voneinander unabhängige Teilfälle. Insgesamt haben wir (mit dem 1. Fall) nun n unterschiedliche Teilfälle, in jedem gibt es $(n-1)!$ Möglichkeiten, also gibt es insgesamt $n\cdot(n-1)! = n!$ Möglichkeiten. □

Definition 15.3. Sei $\sigma \in S_n$.

(1) Ein Paar $(i,\, j)$ mit $i,\, j \in \{1,\dots,n\}$ und $i < j$ heißt ein *Fehlstand* von σ, falls $\sigma(i) > \sigma(j)$ gilt. (Man beachte die Umkehrung des Ungleichheitszeichens!)

(2) Es heißt

$$\operatorname{sgn}(\sigma) = (-1)^{\#\{(i,j)\mid (i,j) \text{ ist Fehlstand von } \sigma\}} \in \{-1,\, 1\}$$

die *Signatur* von σ. (Es ist also $\operatorname{sgn}(\sigma) = 1$, falls die Anzahl der Fehlstände von σ gerade ist, und $\operatorname{sgn}(\sigma) = -1$, falls die Anzahl der Fehlstände von σ ungerade ist.)

Lemma 15.4.

$$\operatorname{sgn}(\sigma) = \prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i}.$$

Beweis. Es gilt (weil σ eine Permutation von $1,\dots,n$ ist)

$$\begin{aligned}\prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i} &= \operatorname{sgn}(\sigma)\cdot\prod_{i<j,\,\sigma(i)<\sigma(j)}\frac{\sigma(j)-\sigma(i)}{j-i}\cdot\prod_{i<j,\,\sigma(i)>\sigma(j)}\frac{|\sigma(j)-\sigma(i)|}{j-i}\\ &= \operatorname{sgn}(\sigma)\cdot\prod_{i<j}\frac{|\sigma(j)-\sigma(i)|}{j-i} = \operatorname{sgn}(\sigma).\end{aligned}$$

□

Satz 15.5. *Für alle* σ, $\tau \in S_n$ *gilt*

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau).$$

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

Beweis. Es gilt

$$\begin{aligned}\operatorname{sgn}(\sigma\tau) &= \prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{j-i} \\ &= \prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} \cdot \prod_{i<j} \frac{\tau(j) - \tau(i)}{j-i}.\end{aligned}$$

Der rechte Faktor ist $\operatorname{sgn}(\tau)$. Für den linken gilt:

$$\begin{aligned}\prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} &= \prod_{i<j,\,\tau(i)<\tau(j)} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} \cdot \prod_{i<j,\,\tau(i)>\tau(j)} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} \\ &= \prod_{i<j,\,\tau(i)<\tau(j)} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} \cdot \prod_{i>j,\,\tau(i)<\tau(j)} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} \\ &= \prod_{\tau(i)<\tau(j)} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \overset{*}{=} \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j-i} \\ &= \operatorname{sgn}(\sigma).\end{aligned}$$

Dabei gilt $*$, weil τ bijektiv ist. $\square$

Lemma 15.6. *Sei* $n \geq 2$. *Sei* $\sigma \in S_n$. *Dann gibt es Transpositionen* $\tau_1, \ldots, \tau_s \in S_n$ *mit*

$$\sigma = \tau_1 \cdot \ldots \cdot \tau_s.$$

Beweis. Für $\sigma = \varepsilon$ ist die Aussage klar (leeres Produkt bzw. $\varepsilon = (1\ 2)(1\ 2)$. Gelte $\sigma \neq \varepsilon$. Sei $i_1 \in \{1, \ldots, n\}$ minimal mit $\sigma(i_1) \neq i_1$. Es gilt also $\sigma(i_1) > i_1$. Sei $\tau_1 = (i_1\ \sigma(i_1))$. Für $\sigma_1 \overset{def}{=} \tau_1\sigma$ gilt dann $\sigma_1(i) = i$ für $i = 1, \ldots, i_1$. Entweder ist nun $\sigma_1 = \varepsilon$, oder es gibt $i_2 > i_1$ minimal mit $\sigma_1(i_2) \neq i_2$. Dann definiert man τ_2 und σ_2 analog, u.s.w. Schließlich erhält man für ein $s \leq n$ Transpositionen $\tau_1, \ldots, \tau_s$ mit $\varepsilon = \tau_s\tau_{s-1} \cdot \ldots \cdot \tau_1 \cdot \sigma$, und es folgt $\sigma = \tau_1^{-1} \cdot \ldots \cdot \tau_s^{-1} \cdot \varepsilon = \tau_1 \cdot \ldots \cdot \tau_s$. $\square$

Bemerkung 15.7. Sei $n \geq 2$.

(1) Obige Darstellung $\sigma = \tau_1 \cdot \ldots \cdot \tau_s$ ist nicht eindeutig; schon $s \in \mathbb{N}_0$ dabei ist nicht eindeutig.

(2) Sei $\tau_0 = (1\ 2) \in S_n$. Ist $\tau \in S_n$ eine beliebige Transposition, so gibt es ein $\sigma \in S_n$ mit $\tau = \sigma\tau_0\sigma^{-1}$.

Beweis. (2) Sei $\tau = (k\ \ell)$. Sei $\sigma \in S_n$ irgendeine Permutation mit $\sigma(1) = k$ und $\sigma(2) = \ell$. Dann gilt offenbar die behauptete Formel $\tau = \sigma\tau_0\sigma^{-1}$. (Man verifiziert dies an jeder Stelle i, und unterscheidet die Fälle $i = k$, $i = \ell$, $i \neq k, \ell$.) $\square$

Folgerung 15.8. *Sei* $n \geq 2$.

1. *Für jede Transposition* $\tau \in S_n$ *gilt* $\operatorname{sgn}(\tau) = -1$.
2. *Sei* $\sigma \in S_n$. *Ist* $\sigma = \tau_1 \cdot \ldots \cdot \tau_s$ *mit Transpositionen* $\tau_1, \ldots, \tau_s$, *so gilt* $\operatorname{sgn}(\sigma) = (-1)^s$.

Beweis. Offenbar hat die Transposition $\tau_0 = (1\ 2)$ genau einen Fehlstand und daher die Signatur -1. Aus Satz 15.5 folgt mit der vorherigen Bemerkung $\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma\tau_0\sigma^{-1}) = \sigma(\tau_0) = -1$, und (2) folgt dann daraus induktiv. $\square$

15.9. Offenbar ist $A_n \overset{def}{=} \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = +1\}$ wegen Satz 15.5 eine Untergruppe von S_n, die *alternierende Gruppe* (vom Grad n). Ist $\sigma \in A_n$, und ist $\tau \in S_n$ eine beliebige Transposition, so gilt $\operatorname{sgn}(\sigma\tau) = -1$. Es folgt dann leicht $S_n = A_n \cup A_n\tau,\ A_n \cap A_n\tau = \emptyset$; hierbei ist $A_n\tau \overset{def}{=} \{\sigma\tau \mid \sigma \in A_n\}$.

Die folgende explizite Formel der Determinante geht auf den Universalgelehrten Gottfried Wilhelm Leibniz (1646–1716) zurück.

Satz 15.10 (Leibniz-Formel). *Sei $n \in \mathbb{N}$. Für $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$ definiert*

$$\det(A) \overset{def}{=} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \alpha_{1\sigma(1)} \cdot \alpha_{2\sigma(2)} \cdot \ldots \cdot \alpha_{n\sigma(n)} \tag{15.1}$$

die eindeutig bestimmte Determinantenfunktion $\det\colon \mathrm{M}(n;K) \to K$*.*

Beweis. Es genügt zu zeigen, dass durch (15.1) eine Determinantenfunktion $\det\colon \mathrm{M}(n;K) \to K$ definiert wird, dass also (D1)–(D3) gilt (denn die Eindeutigkeit hatten wir schon in 14.6 begründet):

(D1)(i) Es gilt

$$\begin{aligned}
\det\left(\begin{pmatrix} \vdots \\ a_i' + a_i'' \\ \vdots \end{pmatrix}\right) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \alpha_{1\sigma(1)} \cdot \ldots \cdot (\alpha'_{i\sigma(i)} + \alpha''_{i\sigma(i)}) \cdot \ldots \cdot \alpha_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \alpha_{1\sigma(1)} \cdot \ldots \cdot \alpha'_{i\sigma(i)} \cdot \ldots \cdot \alpha_{n\sigma(n)} \\
&\quad + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \alpha_{1\sigma(1)} \cdot \ldots \cdot \alpha''_{i\sigma(i)} \cdot \ldots \cdot \alpha_{n\sigma(n)} \\
&= \det\left(\begin{pmatrix} \vdots \\ a_i' \\ \vdots \end{pmatrix}\right) + \det\left(\begin{pmatrix} \vdots \\ a_i'' \\ \vdots \end{pmatrix}\right).
\end{aligned}$$

Ganz ähnlich zeigt man (D1)(ii).

(D2) Es habe die Matrix A zwei gleiche Zeilen, etwa die k-te und die ℓ-te mit $k < \ell$. Sei $\tau = (k\ \ell)$. Dann gilt $S_n = A_n \cup A_n\tau$ und $A_n \cap A_n\tau = \emptyset$. Es gilt also

$$\det(A) = \sum_{\sigma \in A_n} \alpha_{1\sigma(1)} \cdot \ldots \cdot \alpha_{n\sigma(n)} - \sum_{\sigma \in A_n} \alpha_{1\sigma\tau(1)} \cdot \ldots \cdot \alpha_{n\sigma\tau(n)}. \tag{15.2}$$

Nach Definition von τ, und da k-te und ℓ-te Zeile übereinstimmen, heben sich die beiden Summen in (15.2) gerade weg, denn:

$$\begin{aligned}
&\alpha_{1\sigma\tau(1)} \cdot \ldots \cdot \alpha_{k\sigma\tau(k)} \cdot \ldots \cdot \alpha_{\ell\sigma\tau(\ell)} \cdot \ldots \cdot \alpha_{n\sigma\tau(n)} \\
&= \alpha_{1\sigma\tau(1)} \cdot \ldots \cdot \alpha_{k\sigma(\ell)} \cdot \ldots \cdot \alpha_{\ell\sigma(k)} \cdot \ldots \cdot \alpha_{n\sigma\tau(n)} \\
&= \alpha_{1\sigma\tau(1)} \cdot \ldots \cdot \alpha_{k\sigma(k)} \cdot \ldots \cdot \alpha_{\ell\sigma(\ell)} \cdot \ldots \cdot \alpha_{n\sigma\tau(n)} \\
&= \alpha_{1\sigma(1)} \cdot \ldots \cdot \alpha_{n\sigma(n)}.
\end{aligned}$$

(D3) Für $\sigma \in S_n$ gilt offenbar

$$\delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)} = \begin{cases} 1 & \sigma = \varepsilon; \\ 0 & \sigma \neq \varepsilon. \end{cases}$$

Es folgt

$$\det(E_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)} = 1.$$

□

Es sei nochmal darauf hingewiesen, dass wir jetzt erst (mit der Leibniz-Formel) die Existenz einer Determinatenfunktion nachgewiesen haben. In manchen Lehrbüchern wird statt der Axiomatik die Determinante gleich durch die Leibniz-Formel definiert.

Wir haben zwar die Eindeutigkeit der Determinante schon früher gezeigt. Wir beweisen aber nochmal *direkt*, wie die Formel (15.1) aus den Eigenschaften einer Determinantenfunktion folgt.

Definition 15.11. Sei $\sigma \in S_n$. Definiere

$$P_\sigma \overset{def}{=} {}^t[e_{\sigma(1)}, \ldots, e_{\sigma(n)}] = \begin{bmatrix} {}^te_{\sigma(1)} \\ \vdots \\ {}^te_{\sigma(n)} \end{bmatrix} \in \mathrm{M}(n;K),$$

wobei $e_1, \ldots, e_n$ die Standardbasis von K^n ist. Es ist also

$$P[i,j] = \delta_{j\sigma(i)}.$$

Die Matrizen P_σ heißen *Permutationsmatrizen*.

Lemma 15.12. *Sei $\sigma \in S_n$. Dann gilt*

$$\det(P_\sigma) = \mathrm{sgn}(\sigma).$$

Beweis. Ist $\sigma = \tau_1 \cdot \ldots \cdot \tau_s$ ein Produkt von Transpositionen, so erhält man P_σ offenbar durch s Zeilenvertauschungen aus der Einheitsmatrix E_n. Das Ergebnis ist nun eine Folgerung aus 15.8. □

15.13. Sei

$$A = (\alpha_{ij}) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathrm{M}(n;K)$$

mit den Zeilen $a_1, \ldots, a_n \in \mathrm{M}(n;K)$. Man kann schreiben

$$a_i = \alpha_{i1}{}^te_1 + \cdots + \alpha_{in}{}^te_n.$$

Anwendung von (D1) Zeile für Zeile liefert dann

$$\begin{aligned}
\det(\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}) &= \sum_{i_1}^{n} \alpha_{1i_1} \cdot \det(\begin{pmatrix} {}^te_{i_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix}) = \sum_{i_1}^{n}\sum_{i_2}^{n} \alpha_{1i_1} \cdot \alpha_{2i_2} \cdot \det(\begin{pmatrix} {}^te_{i_1} \\ {}^te_{i_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix}) \\
&= \sum_{i_1}^{n}\sum_{i_2}^{n} \cdots \sum_{i_n}^{n} \alpha_{1i_1} \cdot \alpha_{2i_2} \cdot \ldots \cdot \alpha_{ni_n} \cdot \det(\begin{pmatrix} {}^te_{i_1} \\ {}^te_{i_2} \\ \vdots \\ {}^te_{i_n} \end{pmatrix}).
\end{aligned}$$

Nach dem vorherigen Lemma gilt

$$\det(\begin{pmatrix} {}^te_{i_1} \\ {}^te_{i_2} \\ \vdots \\ {}^te_{i_n} \end{pmatrix}) = \begin{cases} \det(P_\sigma) = \mathrm{sgn}(\sigma) & \text{gibt } \sigma \in S_n \text{ mit } \sigma(j) = i_j,\ j = \{1, \ldots, n\} \\ 0 & \text{sonst.} \end{cases}$$

Es folgt dann (15.1).

BEISPIELE 15.14. (1) Für $n = 1$ erhält man $\det((a)) = a$ für jedes $a \in K$.
(2) Für $n = 2$ erhält man

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

(3) Für $n = 3$ ist berechnet sich die Determinaten einer Matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$$

mit Hilfe der *Sarrus-Regel*: Dazu schreibt man die ersten beiden Spalten nochmal rechts neben die drei Spalten von A:

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{31} & \alpha_{32} \end{pmatrix}$$

und addiert dann die Produkte über die drei Diagonalen von links oben nach rechts unten und zieht ab davon die Summe über die drei Diagonalen von rechts oben nach links unten. (ZEICHNUNG) Die Sarrus-Regel ist nichts anderes als die Leibniz-Formel (für $n = 3$), nur in schematischer Form. Denn die sechs Elemente von S_3 sind gegeben als

$$\varepsilon = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (2\ 3), \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (1\ 2),$$

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

und 15.8 liefert die zugehörigen Signaturen.

BEMERKUNG 15.15. (1) Achtung: Eine der Sarrus-Regel entsprechende Regel gilt für $n \geq 4$ *nicht*! Bei einer solchen Regel hätte man bei $n = 4$ nur 8 Summanden (von Produkten), aber die Leibniz-Formel liefert $4! = 24$ Summanden.

(2) Für "große" n ist die Leibniz-Formel zum praktischen Rechnen der Determinante nicht geeignet, denn $n!$ wächst sehr schnell! Schon für $n = 6$ hat man $6! = 720$ Summanden von Produkten mit jeweils 6 Faktoren.

SATZ 15.16. *Sei* $A \in \mathrm{M}(n;K)$. *Dann gilt* $\det({}^tA) = \det(A)$.

BEWEIS. In Übung XIII.5 konnte man einen von der Leibniz-Formel unabhängigen Beweis führen (unter der Annahme der Existenz einer Determinantenfunktion; man verwendet insbesondere (D6), (D7), (D8) und (D11), vgl. Zentralübung). Die Aussage folgt aber auch leicht aus der Leibniz-Formel: Sei $A = (\alpha_{ij})$ und ${}^tA = (\alpha'_{ij})$. Es gilt also $\alpha'_{ij} = \alpha_{ji}$. Dann ist

$$\begin{aligned} \det({}^tA) &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \alpha'_{1\sigma(1)} \cdot \alpha'_{2\sigma(2)} \cdot \ldots \cdot \alpha'_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \alpha_{\sigma(1)1} \cdot \alpha_{\sigma(2)2} \cdot \ldots \cdot \alpha_{\sigma(n)n} \\ &\overset{(*)}{=} \sum_{\sigma^{-1} \in S_n} \mathrm{sgn}(\sigma) \cdot \alpha_{1\sigma^{-1}(1)} \cdot \alpha_{2\sigma^{-1}(2)} \cdot \ldots \cdot \alpha_{n\sigma^{-1}(n)} \\ &\overset{(**)}{=} \det(A). \end{aligned}$$

Bei $(*)$ wurde benutzt, dass

$$\alpha_{\sigma(1)1} \cdot \alpha_{\sigma(2)2} \cdot \ldots \cdot \alpha_{\sigma(n)n} = \alpha_{1\sigma^{-1}(1)} \cdot \alpha_{2\sigma^{-1}(2)} \cdot \ldots \cdot \alpha_{n\sigma^{-1}(n)}$$

ist, denn beide Produkte enthalten bis auf die Reihenfolge die gleichen Faktoren, und außerdem wurde $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ verwendet. Bei $(**)$ beachtet man, dass mit σ auch σ^{-1} alle Elemente aus S_n durchläuft. $\square$

Bemerkung 15.17. Aus dem vorstehenden Satz folgt (Beweis zur Übung empfohlen!), dass zu (D1), (D2), (D5), (D6) und (D7) entsprechende Regeln auch für Spalten statt Zeilen gelten. Insbesondere kann man zum Berechnen der Determinante auch elementare Spaltenumformungen vornehmen.

16. Der Entwicklungssatz von Laplace und die Regel von Cramer

Sei K stets ein Körper und n eine natürliche Zahl.

16.1. Sei $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$. Seien $k,\ \ell \in \{1,\dots,n\}$.
(1) Definiere

$$A_{k\ell} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1,\ell-1} & 0 & \alpha_{1,\ell+1} & \dots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{k-1,1} & \dots & \alpha_{k-1,\ell-1} & 0 & \alpha_{k-1,\ell+1} & \dots & \alpha_{k-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \alpha_{k+1,1} & \dots & \alpha_{k+1,\ell-1} & 0 & \alpha_{k+1,\ell+1} & \dots & \alpha_{k+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{n,\ell-1} & 0 & \alpha_{n,\ell+1} & \dots & \alpha_{nn} \end{pmatrix} \in \mathrm{M}(n;K).$$

(2) Definiere $A'_{k\ell} \in \mathrm{M}(n-1;K)$, indem in A (oder auch in $A_{k\ell}$) die k-te Zeile und die ℓ-te Spalte gestrichen wird.

Lemma 16.2. *Seien $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$ mit den Zeilen $a_1,\dots,a_n$ und $k,\ \ell \in \{1,\dots,n\}$. Dann gilt:*

(1)
$$\det(A'_{k\ell}) = (-1)^{k+\ell}\det(A_{k\ell}).$$

(2)
$$\det(A_{k\ell}) = \det(\begin{pmatrix} a_1 \\ \vdots \\ a_{k-1} \\ {}^t e_\ell \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix}).$$

Beweis. (1) Durch $k-1$ Zeilenvertauschungen und $\ell-1$ Spaltenvertauschungen bringt man $A_{k\ell}$ auf die Form

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A'_{k\ell} \end{array}\right),$$

und die Behauptung folgt mit (D6) (auch der Version für Spalten) und (D9).

(2) Durch Addition jeweils geeigneter Vielfacher der k-ten Zeile zu den anderen Zeilen überführt man die Matrix, die in der rechten Determinante steht, in die Matrix $A_{k\ell}$. □

Definition 16.3. Sei $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$. Dann heißt $\mathrm{adj}(A) \stackrel{def}{=} \big((-1)^{k+\ell}\det(A'_{\ell k})\big)_{k\ell} = \big(\det(A_{\ell k})\big)_{k\ell} \in \mathrm{M}(n;K)$ die zu A *adjunkte Matrix* (oder auch: *komplementäre* Matrix). Man beachte die Vertauschung der Indizes $k,\ \ell$.

Satz 16.4. *Sei $A \in \mathrm{M}(n;K)$. Dann gilt*

$$\mathrm{adj}(A)\cdot A = A\cdot \mathrm{adj}(A) = \det(A)\cdot E_n.$$

Insbesondere gilt: Ist A invertierbar, so gilt

$$A^{-1} = \frac{1}{\det(A)}\cdot \mathrm{adj}(A).$$

(Es ergibt sich durch die erste Formel auch nochmal: A invertierbar $\Leftrightarrow$ $\det(A) \neq 0$.)

Beweis. Seien $a_1, \ldots, a_n$ die Zeilen von A. Seien $i, j \in \{1, \ldots, n\}$. Es ist $a_i = \sum_{k=1}^n \alpha_{ik} {}^t e_k$, wobei $e_1, \ldots, e_n$ die Standardbasis von K^n ist. Es folgt

$$\begin{aligned}
\big(\det(A) \cdot E_n\big)[i,j] &= \delta_{ij} \det(A) \overset{(D2)}{=} \det\left(\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ a_i \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix}\right) \\
&= \det\left(\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ \sum_{k=1}^n \alpha_{ik} {}^t e_k \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix}\right) \\
&\overset{(D1)}{=} \sum_{k=1}^n \alpha_{ik} \det\left(\begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ {}^t e_k \\ a_{j+1} \\ \vdots \\ a_n \end{pmatrix}\right) \\
&\overset{16.2}{=} \sum_{k=1}^n \alpha_{ik} \det(A_{jk}) \overset{16.2}{=} \sum_{k=1}^n \alpha_{ik} \cdot \operatorname{adj}(A)[k,j] \\
&= \big(A \cdot \operatorname{adj}(A)\big)[i,j].
\end{aligned}$$

Es folgt $A \cdot \operatorname{adj}(A) = \det(A) \cdot E_n$. Die Gleichheit $\operatorname{adj}(A) \cdot A = \det(A) \cdot E_n$ folgt analog. □

Beispiel 16.5. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}(2; K)$ invertierbar. Dann ist $\det(A) = ad - bc \neq 0$. Es ist $\operatorname{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Es folgt

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(Vgl. Aufgabe IV.4.)

Der folgende Satz ist nach dem französischen Mathematiker und Astronom Pierre-Simon Laplace (1749–1827) benannt. Er führt die Berechnung einer $n \times n$-Determinante zurück auf die Berechnung von $(n-1) \times (n-1)$-Determinanten. Dies ist insbesondere dann nützlich, wenn in einer Zeile oder Spalte viele Nullen stehen.

Satz 16.6 (Entwicklungssatz von Laplace). *Es sei $n \geq 2$ und $A = (\alpha_{ij}) \in \mathrm{M}(n; K)$. Seien $k, \ell \in \{1, \ldots, n\}$. Dann gilt*

(1) *(Entwicklung nach der k-ten Zeile.)*

$$\det(A) = \sum_{j=1}^{n} (-1)^{k+j} \alpha_{kj} \det(A'_{kj}).$$

(2) *(Entwicklung nach der ℓ-ten Spalte.)*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+\ell} \alpha_{i\ell} \det(A'_{i\ell}).$$

Beweis. (1) Nach dem vorherigen Satz ist $\det(A)$ insbesondere der k-te Eintrag auf der Hauptdiagonalen von $A \cdot \operatorname{adj}(A)$, und dies ist

$$A \cdot \operatorname{adj}(A)[k,k] = \sum_{j=1}^{n} A[k,j] \cdot \operatorname{adj}(A)[j,k] = \sum_{j=1}^{n} (-1)^{k+j} \alpha_{kj} \det(A'_{kj}).$$

(2) Folgt aus (1), indem man $\det({}^tA) = \det(A)$ verwendet. Oder analog zum Beweis von (1), die Formel $\operatorname{adj}(A) \cdot A = \det(A) \cdot E_n$ benutzend. □

Beispiel 16.7.

$$\begin{aligned}
\begin{vmatrix} 2 & 3 & 0 & 1 \\ 0 & 1 & 3 & 1 \\ 2 & 3 & 2 & 1 \\ 1 & 1 & 0 & 1 \end{vmatrix} &\overset{(*)}{=} -3 \cdot \begin{vmatrix} 2 & 3 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 3 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \\
&= 2 \cdot \begin{vmatrix} 2 & 3 & 1 \\ -1 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix} \overset{(**)}{=} 2 \cdot 1 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} \\
&= 2(3 \cdot 1 - 1 \cdot 1) = 4.
\end{aligned}$$

Dabei wurde in $(*)$ nach der dritten Spalte und in $(**)$ nach der zweiten Zeile entwickelt.

Die folgende nach dem schweizer Mathematiker Gabriel Cramer (1704–1752) benannte Methode zum Lösen von (eindeutig lösbaren, quadratischen) linearen Gleichungssystemen ist zum praktischen Rechnen ungeeignet, sie ist aber von theoretischem Interesse.

Satz 16.8 (Cramersche Regel für lineare Gleichungssysteme). *Sei $A = [a_1, \dots, a_n] \in \mathrm{M}(n; K)$ invertierbar (mit den Spalten $a_1, \dots, a_n \in K^n$), sei $b \in K^n$. Sei $x = {}^t(x_1, \dots, x_n)$ die eindeutig bestimmte Lösung des linearen Gleichungssystems $Ax = b$. Dann gilt für jedes $j \in \{1, \dots, n\}$*

$$x_j = \frac{\det([a_1, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n])}{\det(A)}.$$

Beweis. Da A invertierbar ist, ist $x = A^{-1}b$ die eindeutig bestimmte Lösung. Nach Satz 16.4 folgt $x = A^{-1}b = \frac{1}{\det(A)} \cdot \operatorname{adj}(A) \cdot b$. Es folgt

$$\begin{aligned}
x_j &= \frac{1}{\det(A)} \sum_{k=1}^{n} \operatorname{adj}(A)[j,k] \cdot b[k] = \sum_{k=1}^{n} b[k] \cdot \frac{\det(A_{kj})}{\det(A)} \\
&\overset{16.2,\, 15.16}{=} \sum_{k=1}^{n} b[k] \cdot \frac{\det([a_1, \dots, a_{j-1}, e_j, a_{j+1}, \dots, a_n])}{\det(A)} \\
&\overset{(D1),\, 15.16}{=} \frac{\det([a_1, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n])}{\det(A)}.
\end{aligned}$$

□

Beispiel 16.9. Seien

$$A = [a_1,\, a_2,\, a_3] = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 0 & 2 & 1 \end{pmatrix},\ b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

über dem Körper $\mathbb{R}$ der reellen Zahlen. Es gilt

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 0 & 2 & 1 \end{vmatrix} = -2.$$

Mit

$$x'_1 = |b,\, a_2,\, a_3| = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 1 & 2 & 1 \end{vmatrix} = -4,$$

$$x'_2 = |a_1,\, b,\, a_2| = \begin{vmatrix} 1 & 1 & 3 \\ 1 & 0 & 4 \\ 0 & 1 & 1 \end{vmatrix} = -2$$

und

$$x'_3 = |a_1,\, a_2,\, b| = \begin{vmatrix} 1 & 2 & 1 \\ 1 & 2 & 0 \\ 0 & 2 & 1 \end{vmatrix} = 2$$

ergibt

$$x = \frac{1}{|A|} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} -4 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

die eindeutig bestimmte Lösung von $Ax = b$.

17. Die Determinante eines Endomorphismus

Sei K ein Körper und n eine natürliche Zahl.

Definition 17.1. Sei V ein K-Vektorraum. Eine lineare Abbildung $f\colon V \to V$ (also von V in sich) nennt man auch einen *Endomorphismus* von V. Sei

$$\mathrm{End}_K(V) \overset{def}{=} \{f \mid f \text{ ist Endomorphismus von } V\}.$$

Definition 17.2. Sei V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus von V. Sei $\mathcal{B}$ eine Basis von V. Man definiert die *Determinante* von f durch

$$\det(f) \overset{def}{=} \det\bigl(\mathrm{M}_{\mathcal{B}}(f)\bigr).$$

Bemerkung 17.3. Diese Definition ist unabhängig von der Wahl der Basis $\mathcal{B}$. (Vgl. Aufgabe XIII.4. Man verwendet die Transformationsformel und den Determinantenmultiplikationssatz.)

Die folgende Aussage ergibt sich direkt aus schon bekannten Resultaten für (quadratische) Matrizen.

Satz 17.4. *Sei V ein n-dimensionaler K-Vektorraum und seien $f, g \in \mathrm{End}_K(V)$ Endomorphismen von V.*

(1) $\det(1_V) = 1$.
(2) *Es gilt* $\det(f \circ g) = \det(f) \cdot \det(g)$.
(3) *Folgende Aussagen sind äquivalent:*
   (a) *f ist ein Isomorphismus.*
   (b) $\mathrm{rang}(f) = n$.
   (c) $\det(f) \neq 0$.

Ist dies erfüllt, so gilt

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

Ausblick auf das 2. Semester

Im zweiten Teil werden u. a. folgende Themen behandelt:

(1) Eigenwerttheorie eines Endomorphismus.
(Als Hilfsmittel dafür: Polynome)

(2) Normalformen von Endomorphismen. Insbesondere: Die Jordansche Normalform.

(3) Vektorräume mit Skalarprodukt. ($K = \mathbb{R}$, $\mathbb{C}$.)

(4) Normalformen orthogonaler Endomorphismen (die das Skalarprodukt bewahren), symmetrischer Endomorphismen.

(5) Funktionsweise eines Computeralgebrasystems (Bachelor Mathe/Technomathe)

KAPITEL 5

Ringe und Polynome

18. Ringe

18.1. Im ersten Teil haben wir mit den Körpern algebraische Strukuren betrachtet, auf denen eine Addition und eine Multiplikation definiert sind. Allerdings weisen Körper Besonderheiten auf:

- Die Multiplikation ist kommutativ.
- Jedes Element $\neq 0$ hat ein multiplikativen Inverses.

Wir kennen aber schon Beispiele von algebraischen Strukuren, die eine Addition und eine Multiplikation besitzen, bei denen diese Bedingungen nicht erfüllt sind. Bei der Menge $\mathbb{Z}$ der ganzen Zahlen zum Beispiel, haben nur die Elemente ± 1 ein multiplikatives Inverses. Aber auch die Kommutativität ist nicht erfüllt, zum Beispiel nicht für $\mathrm{M}(n; K)$ für $n \neq 2$, oder (aus denselben Gründen) $\mathrm{End}_K(V)$ (der Menge der Endomorphismen von V), sofern $\dim_K V \geq 2$ gilt. In all diesen Fällen bilden die Menge bzgl. der Addition weiterhin eine abelsche Gruppe.

Definition 18.2. (1) Es sei R eine Menge, auf der zwei Verknüpfungen gegeben sind, eine "Addition" $+\colon K \times K \to K$ und eine "Multiplikation" $\cdot\colon K \times K \to K$, für die die folgenden Eigenschaften gelten:
- (R1) $(R, +)$ ist eine abelsche Gruppe; das neutrale Elemente wird mit 0 ("null") bezeichnet.
- (R2) $(R, \cdot)$ ist assoziativ.
- (R3) In R gelten die Distributivgesetz: Für alle a, b, $c \in R$ gilt $a \cdot (b+c) = a \cdot b + a \cdot c$ und $(a + b) \cdot c = a \cdot c + b \cdot c$.

Dann heißt R ein *Ring*.

(2) Enhält R außerdem ein Einselement, d. h. ein neutrales Element der Multiplikation, so heißt R ein Ring mit Eins.

(3) Ist $(R, \cdot)$ kommutativ, so heißt R ein *kommutativer Ring*.

Bemerkung 18.3. (1) Wir werden hier nur Ringe mit Eins betrachten. Daher bedeutet "Ring" im folgenden immer "Ring mit Eins".

(2) In Ringen gelten dieselben Konventionen bzgl. der Schreibweisen wie in 3.2 (1), (2) und (6).

(3) Sei R ein Ring. Wie in 2.3 zeigt man, dass Nullelemente und Einselemente eindeutig bestimmt sind. Wie bei Körpern schreibt man $0 = 0_R$ bzw. $1 = 1_R$.

(4) Sei R ein Ring. Wie in 3.2 (3) zeigt man, dass $r \cdot 0 = 0 = 0 \cdot r$ für jedes $r \in R$ gilt.

(5) Wie in 3.7 gilt auch in Ringen ein allgemeines Distributuivgesetz. In Ringen, die nicht kommutativ sind, muss man allerdings streng auf die Multiplikationsreihenfolge achten.

Beispiel 18.4. (1) Jeder Körper ist ein kommutativer Ring.

(2) Die Menge $\mathbb{Z}$ bildet mit der üblichen Addition $+$ und Multiplikation $\cdot$ einen kommutativen Ring.

(3) Sei K ein Körper und $n \geq 2$. Dann bildet die Menge $\mathrm{M}(n;K)$ mit der Matrizenaddition und Matrizenmultiplikation wegen Satz 6.4 und Satz 6.6 ein Ring. Das Einselement ist die Einheitsmatrix E_n.
(4) Sei V ein K-Vektorraum. Dann bildet $\mathrm{End}_K(V)$ mit der Addition von Abbildungen und der Komposition von Abbildungen ein Ring. Das Einselement ist die identsiche Abbildung 1_V.

Definition 18.5. Sei R ein Ring.

(1) Ein Element $r \in R$ heißt eine *Einheit* (oder *invertierbar*), wenn es ein $s \in R$ gibt mit $rs = 1$ und mit $sr = 1$.
(2) Die Menge der Einheiten von R wird mit $E(R)$ bezeichnet.
(3) Ist $r \in R$ eine Einheit, so ist ein $s \in R$ wie in (1) eindeutig bestimmt (vgl. 6.13 (1)); es wird wie üblich mit r^{-1} bezeichnet.

Satz 18.6. *Sei R ein Ring. Dann bildet $E(R)$ eine Gruppe, die sog. Einheitengruppe von R.*

Beweis. Offenbar gilt $1 \in E(R)$. Sind r, $s \in R$ Einheiten, so zeigt man genau wie in 6.13 (3), dass rs invertierbar ist (und zwar mit $(rs)^{-1} = s^{-1}r{-}1)$), und dass r^{-1} invertierbar ist (und zwar mit $(r^{-1})^{-1} = r$). □

Beispiel 18.7. (1) Ist K ein Körper, so ist $E(K) = K \setminus \{0\}$.
(2) $E(\mathbb{Z}) = \{\pm 1\}$.
(3) $E(\mathrm{M}(n;K)) = \mathrm{GL}(n;K)$.

Definition 18.8. Sei K ein Körper. Ein K-Vektorraum R heißt eine K-Algebra, falls R ein Ring ist, so dass außerdem gilt

$$a(rs) = (ar)s = r(as)$$

für alle $a \in K$ und alle r, $s \in R$.

Beispiel 18.9. Sei K ein Körper.

(1) Es ist $\mathrm{M}(n;K)$ eine K-Algebra.
(2) Sei V ein K-Vektorraum. Dann ist $\mathrm{End}_K(V)$ eine K-Algebra.

Definition 18.10. (1) Seien R und S Ringe. Eine Abbildung $\varphi\colon R \to S$ heißt *Ringhomomorphismus*, falls
  (a) $\varphi(r + r') = \varphi(r) + \varphi(r')$ für alle r, $r' \in R$;
  (b) $\varphi(r \cdot r') = \varphi(r) \cdot \varphi(r')$ für alle r, $r' \in R$;
  (c) $\varphi(1_R) = 1_S$.
  Ist φ zusätzlich bijektiv, so heißt φ ein *Isomorphismus* von Ringen. Die Ringe R und S heißen *isomorph*, wenn es einen Isomorphismus $\varphi\colon R \to S$ gibt.
(2) Sind R und S sogar K-Algebren über dem Körper K, so heißt ein Ringhomomorphismus $\varphi\colon R \to S$ ein *K-Algebrenhomomorphismus*, falls f zusätzlich K-linear ist. Ist φ zusätzlich bijektiv, so heißt φ ein *Isomorphismus* von K-Algebren. Entsprechend heißen die K-Algebren bei der Existenz eines solchen *isomorph*.

Satz 18.11. *Sei K ein Körper und V ein K-Vektorraum. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Die Abbildung*

$$\Phi = \Phi_{\mathcal{B}}\colon \mathrm{End}_K(V) \to \mathrm{M}(n;K),\ f \mapsto M_{\mathcal{B}}(f)$$

ist ein Isomorphismus von K-Algebren.

Beweis. Nach Satz 12.2 ist Φ ein Isomorphismus von K-Vektorräumen. Es ist also noch zu zeigen, dass Φ ein Ringhomomorphismus ist. Dies ist aber in neuer Terminologie gerade die Aussage von Satz 12.8: Für alle f, $g \in \mathrm{End}_K(V)$ gilt

$$\Phi(g \circ f) = M_{\mathcal{B}}(g \circ f) \overset{12.8}{=} M_{\mathcal{B}}(g) \cdot M_{\mathcal{B}}(f) = \Phi(g) \cdot \Phi(f).$$

□

DEFINITION 18.12. Es sei S ein Körper. Eine Teilmenge $R \subseteq S$ heißt *Unterring* (oder *Teilring*) von S, wenn folgendes gilt:

(1) R ist Untergruppe von $(S, +)$.
(2) Es ist $1_S \in R$, und für alle r, $r' \in R$ gilt $r \cdot r' \in R$ (das Produkt gebildet in S).

BEISPIEL 18.13. $\mathbb{Z}$ ist ein Teilring von $\mathbb{Q}$.

DEFINITION 18.14. Ein Ring R heißt *Integritätsbereich*, falls gilt

(1) R ist kommutativ;
(2) $1 \neq 0$ (d. h. $R \neq \{0\}$);
(3) Für alle r, $r' \in R$ gilt: Aus $rr' = 0$ folgt $r = 0$ oder $r' = 0$. (Man sagt, R sei nullteilerfrei.)

BEISPIEL 18.15. (1) Jeder Körper ist ein Integritätsbereich.
(2) $\mathbb{Z}$ ist ein Integritätsbereich.
(3) Für $n \geq 2$ ist $\mathrm{M}(n; K)$ ist *kein* Integritätsbereich.

18.16 (Matrizen und Determinanten über kommutativen Ringen.). Sei R ein kommutativer Ring (mit Eins).

(1) Wie über Körpern kann man $m \times n$-Matrizen $A = (\alpha_{ij})$ mit Einträgen $\alpha_{ij} \in R$ betrachten. Dies führt zu den Mengen $\mathrm{M}(m, n; R)$, und im Fall $m = n$ zu $\mathrm{M}(n; R)$. Auf diesen Mengen kann man dann wie zuvor auch Addition und Multiplikation definieren. Beides ist assoziativ, und $(\mathrm{M}(m, n; R), +)$ ist eine abelsche Gruppe. Ebenso wird Invertierbarkeit von Matrizen definiert.
(2) Für $A = (\alpha_{ij}) \in \mathrm{M}(n; R)$ definiert man die *Determinante* durch die Leibniz-Formel
$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \alpha_{1\sigma(1)} \cdot \alpha_{2\sigma(2)} \cdot \ldots \cdot \alpha_{n\sigma(n)}.$$
Dies definiert eine Determinatenfunktion $\det\colon \mathrm{M}(n; R) \to R$, es gilt also (D1), (D2) und (D3). Es gelten auch die weiteren Eigenschaften, wie für eine Determinantenfunktion über einem Körper, sofern bei dem Beweis die Division keine Rolle gespielt hat. Insbesondere kann man aus (D1), (D2) und (D3) auch die Leibniz-Formel herleiten, man vergleiche den Beweis von Satz 15.10 (der allerdings von der Kommutativität Gebrauch macht.) Die Berechnung von $\det(A)$ führt man wie über Körpern mit Hilfe dieser Rechenregeln durch.

ÜBUNG 18.17. Man überzeuge sich, welche der Regeln (D4)–(D9) aus Satz 14.5 und weiteren Sätze für Determinanten über Körpern auch über kommutativen Ringen gelten.

Der folgende Satz gilt allgemein über kommutativen Ringen. Aus beweistechnischen Gründen formulieren wir ihn nur für Integritätsbereiche. Für unsere Anwendungen wird dies ausreichen.

SATZ 18.18. *Sei R ein Integritätsbereich, seien A, $B \in \mathrm{M}(n; R)$. Dann gilt:*

(1) $\det(A \cdot B) = \det(A) \cdot \det(B)$. *(Determinanten-Multiplikationssatz.)*
(2) $\operatorname{adj}(A) \cdot A = A \cdot \operatorname{adj}(A) = \det(A) \cdot E_n$.
(3) *A ist invertierbar in $\mathrm{M}(n; R)$ genau dann, wenn $\det(A) \in E(R)$.*

BEWEIS. (2) folgt wie über einem Körper, und (3) folgt leicht aus (1) und (2). Wir zeigen nun (1). Wir zeigen, dass R Teilring eines Körpers K ist. Dann folgt (1) aus dem Determinanten-Multiplikationssatz (D11) über Körpern. Wir skizzieren die Konstruktion des Körpers K. Dies ist analog der Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$. Man beachte, dass für a, b, c, $d \in \mathbb{Z}$ mit b, $d \neq 0$ gilt:
$$\frac{a}{b} = \frac{c}{d} \;\Leftrightarrow\; ad = bc.$$

Dies führt zu der folgenden Definition:

Sei X die Menge aller Paare (a, b) mit $a, b \in R$, $b \neq 0$. Wir erklären eine Äquivalenzrelation $\sim$ auf X durch

$$(a, b) \sim (c, d) \quad \Leftrightarrow \quad ad = bc.$$

Sei $K = X/\sim$ die Menge der Äquivalenzklassen. Wir schreiben $\left[\frac{a}{b}\right]$ für die Klasse von (a, b).

Auf K wird nun eine Addition und eine Multiplikation wie folgt erklärt:

$$\left[\frac{a}{b}\right] + \left[\frac{c}{d}\right] \overset{def}{=} \left[\frac{ad + bc}{bd}\right]$$

und

$$\left[\frac{a}{b}\right] \cdot \left[\frac{c}{d}\right] \overset{def}{=} \left[\frac{ac}{bd}\right].$$

Man prüft nach, dass dies *wohldefinierte* Verknüpfungen liefert.

Man sieht, dass diese "Bruchrechenregeln" K zu einem kommutativen Ring machen mit Nullelement $\left[\frac{0}{1}\right]$ und Einselement $\left[\frac{1}{1}\right]$. Es gilt $\left[\frac{a}{b}\right] = 0$, genau dann, wenn $a = 0$ ist. Daher ist jedes $\left[\frac{a}{b}\right] \neq 0$ invertierbar mit Inversem $\left[\frac{b}{a}\right]$.

Es ist offenbar $\iota : R \to K$, $a \mapsto \left[\frac{a}{1}\right]$ ein injektiver Ringhomomorphismus. Wir können daher R mit seinem Bild $\iota(R) \subseteq K$ identifizieren. □

19. Polynomringe

Im folgenden sei K immer ein Körper.

Definition 19.1. Ein *Polynom* über K (oder: mit Koeffizienten in K) ist eine Abbildung $f : \mathbb{N}_0 \to K$ mit endlichem Träger, d. h. f ist eine Folge $(a_n)_{n\geq 0}$ mit $a_n \in K$, und es gilt $a_n = 0$ bis auf endlich viele $n \geq 0$. Wir schreiben $0 = (0, 0, 0, \dots)$ (Nullpolynom) und $1 = (1, 0, 0, \dots)$.

Satz 19.2. *Die Menge aller Polynome über K wird zu einem kommutativen Ring mit Nullelement 0 und Einselement 1 durch folgende Addition und Multiplikation*

$$(f_n) + (g_n) = (f_n + g_n)$$

und

$$(f_n) \cdot (g_n) = \left(\sum_{i=0}^{n} f_i \cdot g_{n-i}\right)_{n\geq 0}.$$

Beweis. Einfaches Nachrechnen. Wir verzichten auf die Ausführung. (Vgl. Übungen.) □

Bemerkung 19.3. Sei R der Ring der Polynome über K.

(1) Es ist $a \mapsto (a, 0, 0, \dots)$ ein injektiver Ringhomomorphismus $K \to R$. Wir identifizieren K damit als Teilring (-körper) von R vermöge dieses Homomorphismus. Die Elemente aus K heißen auch konstante Polynome.

(2) Setze $T := (0, 1, 0, 0, \dots) \in R$. Dann gilt

$$T^0 = 1 \in R,\ T^2 = (0, 0, 1, 0, 0, \dots),\ T^3 = (0, 0, 0, 1, 0, \dots), \dots$$

Satz 19.4. *Jedes vom Nullpolynom verschiedene Polynom f über K hat eine Darstellung $f = a_0 + a_1T + a_2T^2 + \cdots + a_nT^n$ mit $n \geq 0$ und eindeutigen Koeffizienten $a_0, a_1, \dots, a_n \in K$, wobei $a_n \neq 0$.*

Beweis. Nach Teil (2) der vorherigen Bemerkung ist dies klar. □

Bezeichnung 19.5. (1) Wir bezeichnen den oben definierten Polynomring mit $K[T]$; es ist der Polynomring in einer *Unbestimmten* (oder *Variablen* T mit Koeffizienten in K.

(2) Ist $f \in K[T]$ nicht das Nullpolynom, so bezeichnet der Index n aus dem vorherigen Satz den *Grad* von f, also $n = \operatorname{grad} f$. Es heißt a_n der *Leitkoeffizient* von f.

Multipliziert man zwei Polynome in der Unbestimmten T nach den “üblichen” Rechenregeln aus und sortiert das Ergebnis nach den Potenzen von T, so sieht man, dass die Koeffizienten dabei gerade nach obiger Multiplikationsregel ausgrechnet werden. Die zunächst etwas künstlich anmutende Definition der Multiplikation ist also gerade das gewohnte Ausmultiplizieren. [Konkretes Beispiel behandelt.]

Satz 19.6. *Seien $f,\ g \in K[T]$ vom Nullpolynom verschieden. Dann gilt*

(1) $\operatorname{grad}(f \cdot g) = \operatorname{grad}(f) + \operatorname{grad}(g)$. *(Insbesondere $fg \neq 0$.)*
(2) *$K[T]$ ist ein Integritätsbereich.*

Beweis. (1) Seien $f = a_0 + a_1T + \cdots + a_nT^n$ und $g = b_0 + b_1T + \cdots + b_mT^m$ mit $a_n,\ b_m \neq 0$, also mit $n = \operatorname{grad} f$ und $m = \operatorname{grad} g$. Aus der Definition der Multiplikation erhält man sofort

$$fg = a_0b_0 + \cdots + a_nb_mT^{n+m},$$

und in K gilt $a_nb_m \neq 0$. Es folgt $fg \neq 0$ und $\operatorname{grad}(fg) = n + m = \operatorname{grad}(f) + \operatorname{grad}(g)$.
(2) Dies folgt unmittelbar aus (1). □

Satz 19.7 (Polynomdivision mit Rest). *Sei K ein Körper. Seien $f,\ g \in K[T]$ mit $g \neq 0$. Dann gibt es eindeutig bestimmte $q,\ r \in K[T]$ mit*

$$f = qg + r,$$

wobei entweder $r = 0$ gilt oder $r \neq 0$ und $\operatorname{grad}(r) < \operatorname{grad}(g)$.

Beweis. Existenz: Schreibe $f = a_0 + a_1T + \cdots + a_nT^n$ und $g = b_0 + b_1T + \cdots + b_mT^m$ mit $b_m \neq 0$, also $\operatorname{grad} g = m$.
Ist $f = 0$, so setze $q = 0$ und $r = f$ $(= 0)$.
Sei nun $f \neq 0$, also ohne Einschränkung $a_n \neq 0$, $\operatorname{grad} f = n$. Wir beweisen die Aussage nun per Induktion nach n.
Sei $n = 0$.

(1) Ist $\operatorname{grad} g > \operatorname{grad} f = 0$, so setze $q = 0$ und $r = f$.
(2) Ist $\operatorname{grad} g = \operatorname{grad} f = 0$, so setze $q = a_nb_m^{-1}$ und $r = 0$.

Sei nun $n > 0$.

(1) Ist $\operatorname{grad} g > \operatorname{grad} f$, so setze $q = 0$ und $r = f$.
(2) Sei nun $\operatorname{grad} g \leq \operatorname{grad} f$. Dann ist
$$f = a_nb_m^{-1}T^{n-m} \cdot g + \widetilde{f}$$
mit $\widetilde{f} = 0$ oder $\operatorname{grad} \widetilde{f} < n$. Die Induktionsvoraussetzung auf $\widetilde{f}$ angewandt ergibt
$$\widetilde{f} = \widetilde{q} \cdot g + r$$
mit $r = 0$ oder $\operatorname{grad} r < \operatorname{grad} g$. Dies ergibt
$$f = \left(a_nb_m^{-1}T^{n-m} + \widetilde{q}\right) \cdot g + r.$$
Setze $q = a_nb_m^{-1}T^{n-m} + \widetilde{q}$.

Eindeutigkeit: Es gelte $f = q_1g + r_1 = q_2g + r_2$, mit $r_1 = 0$ oder $\operatorname{grad} r_1 < \operatorname{grad} g$ und $r_2 = 0$ oder $\operatorname{grad} r_2 < \operatorname{grad} g$. Es folgt

$$(q_1 - q_2) \cdot g = r_2 - r_1.$$

Aus Gradgründen folgt dann $q_1 - q_2 = 0$, und dann auch $r_2 - r_1 = 0$, also $q_1 = q_2$ und $r_1 = r_2$. □

Offenbar ist $K[T]$ nicht nur ein Ring, sondern sogar eine K-Algebra.

Satz 19.8 (Universelle Eigenschaft des Polynomrings). *Der Polynomring $K[T]$ hat folgende universelle Eigenschaft: Sei S eine K-Algebra und $s \in S$. Dann gibt es genau einen Homomorphismus von K-Algebren $\varphi\colon K[T] \to S$ mit $\varphi(T) = s$.*

Man schreibt: $f(s) \stackrel{def}{=} \varphi(f)$.

Beweis. Definiere $\varphi\big(\sum_{i=1}^n a_i T^i\big) = \sum_{i=1}^n a_i s^i$. □

19.9 (Einsetzen). Seien die Voraussetzungen wie im vorigen Satz. Man sagt, dass man in die Unbestimmte T das Element $s \in S$ *einsetzt*. Der KAlgebrenhomomorphismus $\varphi\colon f \mapsto f(s)$, $K[T] \longrightarrow S$ heißt *Einsetzungshomomorphismus*. Ein $s \in S$ heißt *Nullstelle* von f (in S), falls $f(s) = 0$ gilt.

Beispiel 19.10. Sei $f = T^2 - 2T - 3 \in \mathbb{R}[T]$. Sei

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in \mathrm{M}(2; \mathbb{R}).$$

Man kann A in f einsetzen und erhält

$$f(A) = A^2 - 2A - 3E_2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} - \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Es ist also A eine Nullstelle von f in $\mathrm{M}(2; \mathbb{R})$.

Satz 19.11. *Sei $f \in K[T]$ ($f \neq 0$) ein Polynom vom Grad n. Sei $\alpha \in K$ eine Nullstelle von f in K. Dann gilt*

$$f = q \cdot (T - \alpha),$$

mit $q \in K[T]$ vom Grad $n - 1$. Insbesondere hat f in K höchstens n Nullstellen.

Beweis. Division mit Rest ergibt

$$f = q \cdot (T - \alpha) + r$$

mit $r = 0$ oder $\mathrm{grad}(r) < \mathrm{grad}(T - \alpha) = 1$. Also ist r ein konstantes Polynom. Setzt man in diese Polynomgleichung das Element α ein, so ergibt sich

$$0 = f(\alpha) = q(\alpha) \cdot (\alpha - \alpha) + r(\alpha) = r(\alpha).$$

Es folgt $r = 0$. Da nun $\mathrm{grad}(q) = n - 1$ gilt, folgt der Zusatz induktiv. □

Folgerung 19.12. *Sei K ein unendlicher Körper. Dann ist $K[T]$ isomorph zum Ring* $\mathrm{Pol}(K, K)$ *aller Polynomfunktionen $f\colon K \to K$, $t \mapsto a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ mit Koeffizienten in K.*

Beweis. Jedes $f \in K[T]$ liefert eine eindeutige Polynomfunktion $t \mapsto f(t)$. Diese Zuordnung ist offenbar surjektiv und ein Homorphismus von Ringen. Weil K unendlich ist, ist diese Zuordnung auch injektiv nach dem vorigen Satz. □

Beispiel 19.13. Sei $\mathbb{F}_2$ der Körper mit zwei Elementen 0 und 1, vgl. Beispiel 3.3 (3). Sei $f = T + 1$, $g = T^2 + 1$. Dann gilt $f(0) = 1 = g(0)$ und $f(1) = 0 = g(1)$. Es gilt also $f(t) = g(t)$ für alle $t \in \mathbb{F}_2$, d. h. f und g sind verschiedene Polynome, liefern aber dieselben Polynom*funktionen* $\mathbb{F}_2 \to \mathbb{F}_2$, $t \mapsto f(t) = g(t)$.

Auch aus diesem Grund ist es wichtig, zwischen Polynomen und Polynomfunktionen zu unterscheiden.

Bemerkung 19.14. Sei $f \in K[T]$ ($f \neq 0$) ein Polynom vom Grad n. Sei $\alpha \in K$. Ist α eine Nullstelle von f, so gibt es, wie oben gezeigt, ein $q \in K[T]$ mit $f = q \cdot (T - \alpha)$. Ist α auch Nullstelle von q, so kann man dies wiederholen. Es gibt also ein eindeutig bestimmtes $m \in \mathbb{N}$ mit $0 \leq m \leq n$ und $g \in K[T]$ mit

$$f = g \cdot (T - \alpha)^m \quad \text{und } g(\alpha) \neq 0.$$

Es heißt m die *Vielfachheit der Nullstelle* α. Wir schreiben $m = v_\alpha(f)$. (Natürlich ist α nur für $m \geq 1$ tatsächlich eine Nullstelle.)

Bemerkung 19.15. Sei $f \in K[T]$ ($f \neq 0$) ein Polynom. Die vorige Bemerkung zeigt: Sind $\alpha_1, \dots, \alpha_r$ die verschiedenen Nullstellen von f in K, so kann man schreiben

$$f = a \cdot (T - \alpha_1)^{m_1} \cdot \ldots \cdot (T - \alpha_r)^{m_r} \cdot g$$

mit $a \in K \setminus \{0\}$, wobei $g \in K[T]$ normiert ist und keine Nullstellen in K hat und mit $m_i = v_{\alpha_i}(f)$. Offenbar sind $r, \alpha_1, \dots, \alpha_r$ (bis auf Reihenfolge), a und g eindeutig bestimmt. (Vgl. Übungen.)

Anhang: Zerlegung in Primfaktoren.

Definition 19.16. Sei $f \in K[T]$ mit $f \neq 0$.

(1) f heißt *irreduzibel*, wenn $\mathrm{grad}(f) \geq 1$ gilt, und wenn aus $f = gh$ (mit $g, h \in K[T]$) folgt, dass $\mathrm{grad}(f) = 0$ oder $\mathrm{grad}(g) = 0$ gilt.
(2) f heißt *normiert*, wenn 1 der Leitkoeffizient von f ist.

Lemma 19.17. *Sei $f \in K[T]$, $f \neq 0$. Dann gibt es irreduzible, normierte Polynome $p_1, \dots, p_r \in K[T]$ und ein $a \in K \setminus \{0\}$ mit $f = a \cdot p_1 \cdot \ldots \cdot p_r$.*

Beweis. Gilt $\mathrm{grad}(f) = 0$, so sei $a = f$ und $r = 0$ (leeres Produkt = 1). Sei nun $\mathrm{grad}(f) \geq 1$. Ist a der Leitkoeffizient von f, so kann man $f = a\widetilde{f}$ schreiben, wobei $\widetilde{f} \in K[T]$ normiert ist. Ist $\widetilde{f}$ selbst schon irreduzibel, so ist man fertig. Andernfalls kann man schreiben $\widetilde{f} = gh$, wobei $g, h \in K[T]$ gilt, ohne Einschränkung normiert, mit $\mathrm{grad}(g) \geq 1$ und $\mathrm{grad}(h) \geq 1$. Sind jetzt g und h irreduzibel, so ist man fertig. Andernfalls zerlegt man g und/oder h weiter in Polynome kleineren Grades. Da die Grade immer kleiner werden, muss dieses Verfahren nach endlichen Schritten abbrechen, und dann sind alle auftretenden Faktoren irreduzibel. □

Wir kümmern uns nun um die Eindeutigkeit solcher Zerlegungen in irreduzible Polynome.

Definition 19.18. Seien $f, g \in K[T]$. Wir schreiben $f \mid g$ (und sagen: f *teilt* g), wenn es ein $h \in K[T]$ gibt mit $hf = g$. Es heißt dann f ein *Teiler* von g, und umgekehrt g ein *Vielfaches* von f.

Bemerkung 19.19. Seien $f, g \in K[T]$ nicht beide das Nullpolynom. Dann gibt es ein $d \in K[T]$, das man ohne Einschränkung als normiert annehmen kann, mit

$$I \overset{def}{=} \{rf + sg \mid r, s \in K[T]\} = \{hd \mid h \in K[T]\}.$$

Denn es gibt ein $d \in I$, $d \neq 0$ (normiert) von minimalem Grad. Dann ist die Teilmengenbeziehung "$\supseteq$" klar. Sei umgekehrt $k \in I$ beliebig. Division mit Rest ergibt $k = qd + r$ mit $r = 0$ oder $\mathrm{grad}(r) < \mathrm{grad}(d)$. Da aber $r = k - qd \in I$ gilt, kann nur (nach Wahl von d) $r = 0$ gelten. Es gilt also $k = qd$. –

Daraus folgt nun:

(1) $d \mid f$ und $d \mid g$. Es ist d also ein gemeinsamer Teiler von f und g.
(2) Jeder andere gemeinsame Teiler von f und g teilt schon d. Es ist also d ein *größter gemeinsamer Teiler* von f und g.
(3) Es gibt $a, b \in K[T]$ mit $d = af + bg$.

(Beachte: (2) folgt aus (3).)

Lemma 19.20 ("Euklids Lemma"). *Sei $f \in K[T]$, $f \neq 0$ mit $\mathrm{grad}(f) \geq 1$. Äquivalent sind:*

(1) *f ist irreduzibel.*
(2) *Für alle $g, h \in K[T]$ gilt: $f \mid gh \Rightarrow f \mid g$ oder $f \mid h$.*

Beweis. “(2)⇒(1)” Es gelte (2). Sei $f = gh$. Dann gilt insbesondere $f \mid gh$, und daher wegen (2) $f \mid g$ oder $f \mid h$. Es folgt $\mathrm{grad}(g) \geq \mathrm{grad}(f)$, und dann $\mathrm{grad}(h) = 0$, oder $\mathrm{grad}(h) \geq \mathrm{grad}(f)$, und dann $\mathrm{grad}(g) = 0$. Also ist f irreduzibel.

“(1)⇒(2)” Seien $g, h \in K[T]$ mit $f \mid gh$. Zu f und g sei $d \in K[T]$ wie in der vorigen Bemerkung. Es gilt also $d \mid f$ und $d \mid g$. Sei etwa $ed = f$. Da f irreduzibel ist, folgt $e \in K$ oder $d \in K$. Im ersten Fall folgt $f = ed \mid g$. Im zweiten Fall folgt aber, da d normiert ist, $d = 1$ ist, also gibt es $a, b \in K[T]$ mit $1 = af + bg$. Es folgt $h = h \cdot 1 = ahf + bgh$. Da f nach Annahme gh teilt, folgt daraus $f \mid h$. □

Satz 19.21. *Sei $f \in K[T]$, $f \neq 0$. Dann gibt es (bis auf Reihenfolge) eindeutig bestimmte irreduzible, normierte Polynome $p_1, \dots, p_r \in K[T]$ und ein eindeutig bestimmtes $a \in K \setminus \{0\}$ mit $f = a \cdot p_1 \cdot \ldots \cdot p_r$.*

Beweis. Wir zeigen: Gilt $f = a \cdot p_1 \cdot \ldots \cdot p_r = a' \cdot p'_1 \cdot \ldots \cdot p'_{r'}$, so gilt $a = a'$, $r = r'$ und bis auf Umnummerierung $p_i = p'_i$. Offenbar stimmt sowohl a wie auch a' mit dem Leitkoeffizienten von f überein. Es folgt $a = a'$, also auch $p_1 \cdot \ldots \cdot p_r = p'_1 \cdot \ldots \cdot p'_{r'}$. Es wird nun Induktion nach r gemacht. Ist $r = 0$ oder $r = 1$, so folgt auch $r' = 0$ bzw. $r' = 1$, da p_1 irreduzibel ist. Sei nun $r \geq 2$. Es ist p_r ein Teiler von $p'_1 \cdot \ldots \cdot p'_{r'}$, daher teilt p_r nach Lemma 19.20 eines der Elemente p'_i. Nach evtl. Umnummerierung können wir $p_r \mid p'_{r'}$ annehmen. Da aber $p'_{r'}$ irreduzibel ist, folgt $p_r = p'_{r'}$. Also kann man (in dem *Integritätsbereich* $K[T]$) kürzen, und erhält $p_1 \cdot \ldots \cdot p_{r-1} = p'_1 \cdot \ldots \cdot p'_{r'-1}$. Per Induktion kann man $r - 1 = r' - 1$, also $r = r'$ folgern, und (nach evtl. Umnummerierung) $p_i = p'_i$ $(i = 1, \dots, r-1)$. □

Bemerkung 19.22. Irreduzible Polynome heißen auch *Primpolynome*. Die Zerlegung im vorigen Satz wir auch (eindeutige) *Primfaktorzerlegung* genannt. Man vgl. die Analogie zu den ganzen Zahlen $\mathbb{Z}$.

Bemerkung 19.23. (1) Jedes Polynom vom Grad 1 ist irreduzibel.

(2) Sei $K = \mathbb{C}$ der Körper der komplexen Zahlen. Da man stets Nullstellen als Linearfaktor abspalten kann, folgt aus dem Fundamentalsatz der Algebra (vgl. Satz 4.7), dass hier *nur* die Polynome vom Grad 1 irreduzibel sind. Man sagt auch, dass (über $\mathbb{C}$) jedes Polynom in Linearfaktoren zerfällt.

(3) Ein Körper, über dem jedes Polynom in Linearfaktoren zerfällt, heißt *algebraisch abgeschlossen*. Der Fundamentalsatz der Algebra kann deshalb auch so formuliert werden: *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

(4) Der Körper $\mathbb{R}$ der reellen Zahlen ist *nicht* algebraisch abgeschlossen. Denn etwa das Polynom $f = T^2 + 1$ zerfällt in $\mathbb{R}[T]$ nicht in Linearfaktoren. (Beweis als Übung.) Daher ist dieses Polynom auch irreduzibel in $\mathbb{R}[T]$. In $\mathbb{C}[T]$ zerfällt dieses Polynom jedoch in Linearfaktoren. Irreduzibilität hängt daher vom Grundkörper ab, den man betrachtet!

(5) Man zeigt leicht (vgl. Übungen), dass die normierten irreduziblen Polynome über $\mathbb{R}$ gerade die folgenden sind:

$$T - \alpha \quad (\alpha \in \mathbb{R})$$
$$(T - z) \cdot (T - \overline{z}) = T^2 - (z + \overline{z})T + z\overline{z} \quad (z \in \mathbb{C} \setminus \mathbb{R}).$$

(6) Man zeigt in der *Algebra*, dass man jeden Körper in einen algebraisch abgeschlossenen Körper einbetten kann.

KAPITEL 6

Eigenwerttheorie

20. Eigenwerte und Eigenvektoren

Im folgenden sei K stets ein Körper und V stets K-Vektorraum.

DEFINITION 20.1. Sei $f \in \mathrm{End}_K(V)$ ein Endomorphismus.

(1) Ein Element $\lambda \in K$ heißt ein *Eigenwert* von f (in K), falls es ein $x \in V$ gibt mit $x \neq 0$ und $f(x) = \lambda x$.
(2) Ist $\lambda \in K$ ein Eigenwert von f, so heißt ein $x \in V$ mit $x \neq 0$ und mit $f(x) = \lambda x$ ein *Eigenvektor* von f.
(3) Die Menge $E(f, \lambda) \overset{def}{=} \{x \in V \mid f(x) = \lambda x\} \subseteq V$ heißt der *Eigenraum* zu f (zum Eigenwert λ).

BEMERKUNG 20.2. Sei $f \in \mathrm{End}_K(V)$ ein Endomorphismus und $\lambda \in K$.

(1) Es gilt $E(f, \lambda) = \mathrm{Kern}(\lambda 1_V - f)$. Insbesondere ist $E(f, \lambda)$ ein Unterraum von V.
(2) Es gilt $E(f, \lambda) \neq \{0_V\} \;\Leftrightarrow\; \lambda$ ist Eigenwert von f.

BEWEIS. Klar. □

BEISPIEL 20.3. Sei $V \neq \{0\}$.

(1) Der einzige Eigenwert von 1_V ist 1_K, und es gilt $E(1_V, 1_K) = V$.
(2) Der einzige Eigenwert von der Nullabbildung $0\colon V \to V$ ist 0_K, und es gilt $E(0, 0_K) = V$.

SATZ 20.4. *Sei $f \in \mathrm{End}_K(V)$, seien $x_1, \ldots, x_m$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_m \in K$. Dann sind $x_1, \ldots, x_m$ linear unabhängig. Insbesondere gibt es höchstens* $\dim(V)$ *verschiedene Eigenwerte.*

BEWEIS. Induktion nach m. Für $m = 1$ ist dies wegen $x_1 \neq 0$ klar. Sei nun $m \geq 2$. Seien $\alpha_1, \ldots, \alpha_m \in K$ mit

$$\alpha_1 x_1 + \ldots + \alpha_m x_m = 0.$$

Wende einmal f an, das andere mal mutlipliziere mit λ_1:

$$\begin{aligned} \alpha_1\lambda_1 x_1 + \alpha_2\lambda_2 x_2 + \ldots + \alpha_m\lambda_m x_m &= 0 \\ \alpha_1\lambda_1 x_1 + \alpha_2\lambda_1 x_2 + \ldots + \alpha_m\lambda_1 x_m &= 0. \end{aligned}$$

Subtrahiert man die Gleichungen, so erhält man

$$\alpha_2(\lambda_2 - \lambda_1)x_2 + \ldots + \alpha_m(\lambda_m - \lambda_1)x_m = 0.$$

Nach Induktionsvoraussetzung sind $x_2, \ldots, x_m$ linear unabhängig, und es folgt

$$\alpha_2(\lambda_2 - \lambda_1) = \cdots = \alpha_m(\lambda_m - \lambda_1) = 0.$$

Da die Eigenwerte verschieden sind, folgt

$$\alpha_2 = \cdots = \alpha_m = 0.$$

Es folgt dann auch $\alpha_1 x_1 = 0$, also $\alpha_1 = 0$ wegen $x_1 \neq 0$. □

DEFINITION 20.5. Ein Endomorphismus $f \in \mathrm{End}_K(V)$ heißt *diagonalisierbar*, wenn es eine Basis von V gibt, die nur aus Eigenvektoren von f besteht.

Satz 20.6. *Es gelte* $\dim(V) = n$. *Sei* $f \in \mathrm{End}_K(V)$. *Folgende Aussagen sind äquivalent:*
(1) *f ist diagonalisierbar.*
(2) *Es gibt eine Basis $\mathcal{B}$ von V, so dass*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist.
(3) $V = \sum_{\lambda \in K} E(f, \lambda)$.
(4) $V = \bigoplus_{\lambda \in K} E(f, \lambda)$.
(5) $\sum_{\lambda \in K} \dim_K(E(f, \lambda)) = n$.

Gilt dies, so sind $\lambda_1, \ldots, \lambda_n$ aus (2) die einzigen (nicht notwendig verschiedenen) Eigenwerte von f in K.

(Man beachte, dass in den Summen nur endlich viele Summanden $\neq \{0\}$ bzw. $\neq 0$ sind.)

Beweis. "(1)⇒(2)" Bei $\mathcal{B} = (x_1, \ldots, x_n)$ eine Basis, die aus Eigenvektoren x_i besteht, d. . es gibt zugehörige Eigenwerte λ_i, also $f(x_i) = \lambda_i x_i$ für $i = 1, \ldots, n$. Es folgt, dass $M_{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ist.

"(2)⇒(1)" Sei $\mathcal{B} = (x_1, \ldots, x_n)$ eine Basis, so dass $M_{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ gilt. Dann folgt aber $f(x_i) = \lambda_i x_i$, d. h. x_i ist Eigenvektor $(i = 1, \ldots, n)$.

"(1)⇔(4)" ist klar (die Basisvektoren, die zum selben Eigenwert λ gehören, bilden eine Basis von $E(f, \lambda)$). Die Implikationen "(4)⇒(3)", "(4)⇒(5)" und "(5)⇒(3)" sind ebenfalls klar.

"(3)⇒(4)" folgt aus dem vorigen Satz. □

Folgerung 20.7. *Es gelte* $\dim(V) = n$, *sei* $f \in \mathrm{End}_K(V)$. *Wenn f n verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n \in K$, so ist f diagonalisierbar.*

Beweis. Es gilt

$$n \leq \sum_{i=1}^{n} \dim_K(E(f, \lambda_i)) \leq n,$$

und nach (5) aus vorigem Satz ist f diagonalisierbar. □

Da sich bei endlichdimensionalen Vektorräumen mit gewählter Basis Endomorphismen und quadratische Matrizen entsprechen, ist es nicht überraschend, dass die zuvor besprochenen Begriffe analog für quadratische Matrizen existieren.

Definition 20.8. Sei $A \in \mathrm{M}_n(K)$.
(1) Ein Element $\lambda \in K$ heißt ein *Eigenwert* von A (in K), falls es ein $x \in K^n$ gibt mit $x \neq 0$ und $Ax = \lambda x$.
(2) Ist $\lambda \in K$ ein Eigenwert von A, so heißt ein $x \in K^n$ mit $x \neq 0$ und mit $Ax = \lambda x$ ein *Eigenvektor* von A.
(3) Die Menge $E(A, \lambda) \overset{def}{=} \{x \in V \mid Ax = \lambda x\} \subseteq V$ heißt der *Eigenraum* zu A (zum Eigenwert λ).

Bemerkung 20.9. Sei V ein endlichdimensionaler K-Vektorraum, sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V und $\phi_{\mathcal{B}} \colon K^n \to V$ der dazu definierte Isomorphismus. Sei f ein Endomorphismus und $A \in \mathrm{M}(n; K)$ die zugehörige Darstellungsmatrix. Sei $\lambda \in K$. Dann gilt

$$\lambda \text{ ist Eigenwert von } f \iff \lambda \text{ ist Eigenwert von } A;$$

die zugehörigen Eigenvektoren bzw. Eigenräume korrespondieren unter $\phi_{\mathcal{B}}$.

Beweis. Klar. □

Bemerkung 20.10. Sei $A \in \mathrm{M}(n;K)$. Sei $f_A\colon K^n \to K^n$ die zugehörige lineare Abbbildung. Genau dann ist f_A diagonalisierbar, wenn es ein $P \in \mathrm{GL}(n;K)$ gibt, so dass $P^{-1}AP$ eine Diagonalmatrix ist.

Beweis. Sei $\mathcal{B} = (e_1, \dots, e_n)$ die Standardbasis von K^n, sei $\mathcal{C}$ irgendeine andere Basis von K^n. Dann besagt die Transformationsformel 12.13, dass

$$M_{\mathcal{C}}(f_A) = T^{\mathcal{C}}_{\mathcal{B}} \cdot M_{\mathcal{B}}(f_A) \cdot (T^{\mathcal{C}}_{\mathcal{B}})^{-1} = P^{-1}AP,$$

mit $P \overset{def}{=} (T^{\mathcal{C}}_{\mathcal{B}})^{-1}$. Nun besteht $\mathcal{C}$ genau dann aus Eigenvektoren von f_A, wenn $M_{\mathcal{C}}(f_A)$ eine Diagonalmatrix ist. □

Definition 20.11. Zwei Matrizen A, $B \in \mathrm{M}(n;K)$ heißen *ähnlich*, und man schreibt $A \approx B$, wenn es ein $P \in \mathrm{GL}(n;K)$ gibt mit $B = P^{-1}AP$. Dies definiert offenbar eine Äquivalenzrelation auf $\mathrm{M}(n;K)$.

Definition 20.12. $A \in \mathrm{M}(n;K)$ heißt *diagonalisierbar*, wenn A ähnlich zu einer Diagonalmatrix $D \in \mathrm{M}(n;K)$ ist.

Bemerkung 20.13.
(1) Im folgenden hat fast jede Aussage über Endomorphismen (eines endlichdimensionalen Raums) eine entsprechende Aussage für Matrizen, und umgekehrt. Wir werden nicht jedesmal beide Fassungen formulieren.
(2) Eines der Ziele der Linearen Algebra ist es, einen vorgegebenen Endomorphismus f bzgl. einer geeignten Basis durch eine Matrix mit "möglichst einfacher Gestalt" darzustellen. Eine solche Gestalt, die man f auf eindeutige Weise zuordnet, nennt man auch eine *Normalform* von f. Für Matrizen $A \in \mathrm{M}(n;K)$ bedeutet das, dass man innerhalb der Äquivalenzklasse von A (bzgl. $\approx$) einen möglichst einfachen Repräsentanten finden möchte.

    Ein Analogon ist uns schon in Satz 7.23 bzgl. der Äquivalenzrelation $\sim$ in $\mathrm{M}(m,n;K)$ begegnet. Solche einfache Normalformen sind für Endomorphismen aber nur ganz selten möglich. Das liegt grob gesprochen, dass man hier nur eine Basis abändern darf, während man in 7.23 im Definitionsbereich und im Bildbereich Basen separat abändern durfte.
(3) Diagonalisierbare Endomorphismen z. B. haben eines besonders einfache Normalform, nämlich eine Diagonalmatrix. Aber die wenigsten Endomorphismen sind diagonalisierbar. Daher ist erstens ein Kriterium wichig, mit dem man leicht entscheiden kann, ob ein Endomorphismus diagonalisierbar ist oder nicht, und zweitens, sollen auch Normalformen von nicht-diagonalisierbaren Endomorphismen entwickelt werden.

21. Minimalpolynom und charakteristisches Polynom

Es sei K stets ein Körper und V ein K-Vektorraum. Wir nehmen stets $n = \dim_K(V) < \infty$ an.

Definition 21.1. Es sei $f \in \mathrm{End}_K(V)$. Sei $\mathcal{B}$ eine Basis von V, und sei $A = (\alpha_{ij}) = M_{\mathcal{B}}(f) \in \mathrm{M}(n;K)$. Betrachte die quadratische Matrix

$$T \cdot E_n - A = \begin{pmatrix} T-\alpha_{11} & -\alpha_{12} & \dots & -\alpha_{1n} \\ -\alpha_{11} & T-\alpha_{12} & \dots & -\alpha_{1n} \\ \vdots & & & \vdots \\ -a_{11} & -\alpha_{12} & \dots & T-\alpha_{nn} \end{pmatrix} \in \mathrm{M}(n;K[T]).$$

Dann heißt

$$\chi_f \overset{def}{=} \det(T \cdot E_n - A) \in K[T]$$

das *charakteristische Polynom* von f. Analog: für eine beliebige Matrix $A \in \mathrm{M}(n; K)$ heißt $\chi_A = \det(T \cdot E_n - A)$ das charakteristische Polynom von A.

LEMMA 21.2. *Das charakteristische Polynom χ_f eines Endomorphismus $f \in \mathrm{End}_K(V)$ ist unabhängig von der Wahl der Basis $\mathcal{B}$.*

BEWEIS. Sei auch $\mathcal{B}'$ eine Basis von V und $A' = M_{\mathcal{B}'} \in \mathrm{M}(n; K)$. Dann gibt es ein $p \in \mathrm{GL}(n; K)$ mit $A' = P^{-1}AP$. Es ist P insbesondere eine invertierbare Matrix in $\mathrm{M}(n; K[T])$. Es folgt aus dem Determinanten-Multiplikationssatz 18.18

$$\begin{aligned}\det(T \cdot E_n - A') &= \det(T \cdot E_n - P^{-1}AP) = \det(P^{-1}(T \cdot E_n - A)P)\\ &= \det(P^{-1})\det(T \cdot E_n - A)\det(P)\\ &= \det(T \cdot E_n - A).\end{aligned}$$

□

BEMERKUNG 21.3. Matrizentheoretisch bedeutet die vorige Aussage: Ähnliche quadratische Matrizen haben dasselbe charaketristische Polynom: sind A, $B \in \mathrm{M}(n; K)$, so gilt

$$A \approx B \quad \Rightarrow \quad \chi_A = \chi_B.$$

BEISPIEL 21.4.
(1) Sei $n = 1$. Sei $A = (\alpha)$. Dann gilt $\chi_A = |T - \alpha| = T - \alpha$.
(2) Sei $n = 2$ und $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann gilt

$$\chi_A = \begin{vmatrix} T-a & -b \\ -c & T-d \end{vmatrix} = (T-a)(T-d) - (-b)(-c) = T^2 - (a+d)T + (ad - bc).$$

(3) Sei

$$A = \begin{pmatrix} 2 & -2 & -4 & -2 \\ -8 & 5 & 1 & -8 \\ 0 & 3 & -1 & 0 \\ -2 & -1 & 5 & 2 \end{pmatrix} \in \mathrm{M}(4; \mathbb{C}).$$

Dann gilt

$$\begin{aligned}\chi_A &= \begin{vmatrix} T-2 & 2 & 4 & 2 \\ 8 & T-5 & -1 & 8 \\ 0 & -3 & T+1 & 0 \\ 2 & 1 & -5 & T-2 \end{vmatrix}\\ &= \begin{vmatrix} T & 0 & 0 & 0 \\ 8 & T-5 & -1 & 8 \\ 0 & -3 & T+1 & 0 \\ 2 & 1 & -5 & T-2 \end{vmatrix} + \begin{vmatrix} -2 & 2 & 4 & 2 \\ 8 & T-5 & -1 & 8 \\ 0 & -3 & T+1 & 0 \\ 2 & 1 & -5 & T-2 \end{vmatrix}\\ &= T \cdot \begin{vmatrix} T-5 & -1 & 8 \\ -3 & T+1 & 0 \\ 1 & -5 & T-2 \end{vmatrix} + \begin{vmatrix} -2 & 2 & 4 & 2 \\ 0 & T+3 & 15 & 16 \\ 0 & -3 & T+1 & 0 \\ 0 & 3 & -1 & T \end{vmatrix}\\ &= \ldots\\ &= T^4 - 8T^3 - 16T^2 + 128T.\end{aligned}$$

BEMERKUNG 21.5.
(1) Man beachte, dass zur Definition von χ_f vom Polynombegriff über einem Integritätsbereich Gebrauch gemacht wurde.
(2) Die zu χ_f gehörige Polynomfunktion $K \to K$ ist durch $\lambda \mapsto \chi_f = \det(\lambda \cdot 1_V - f)$ gegeben. Dies folgt aus der Leibniz-Formel und der Tatsache, dass der Einsetzungshomomorphismus ein Ringhomomorphismus ist.

SATZ 21.6. *Das charakteristische Polynom χ_A einer Matrix $A \in \mathrm{M}(n;K)$ ist normiert und vom Grad n. Genauer gilt: Ist $\chi_A = a_nT^n + a_{n-1}T^{n-1} + \ldots + a_1T + a_0 \in K[T]$, so gilt $a_n = 1$, $a_{n-1} = -\operatorname{sp}(A)$ und $a_0 = (-1)^n \det(A)$.*

BEWEIS. Es gilt $a_0 = \chi_A(0) = \det(-A) = (-1)^n \det(A)$. Sei $A = (\alpha_{ij})$. Die Leibniz-Formel ergibt

$$\chi_A = \prod_{i=1}^{n}(T - \alpha_{ii}) + \sum_{\sigma\in S_n,\, \sigma\neq\varepsilon} \operatorname{sgn}(\sigma) \prod_{i=1}^{n}(\delta_{i,\sigma(i)}T - \alpha_{i,\sigma(i)}).$$

Der zweite Term hat einen Grad $\leq n-2$, denn für $\sigma \neq \varepsilon$ gibt es mindestens zwei verschiedene i und j mit $\sigma(i) \neq i$ und $\sigma(j) \neq j$. Die Koeffizienten a_n und a_{n-1} sind also dieselben, wie die im ersten Summanden, und damit folgt sofort $a_n = 1$ und $a_{n-1} = -\sum_{i=1}^{n} \alpha_{ii}$. □

LEMMA 21.7. *Jedes normierte Polynom $f \in K[T]$ ist das charakteristische Polynom einer quadratischen Matrix. Genau: Ist*

$$f = T^n + a_{n-1}T^{n-1} + \ldots + a_1T + a_0 \in K[T],$$

so ist

$$A = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & \ldots & 0 & -a_1 \\ 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & 0 & & \ddots & \vdots & \\ \vdots & \vdots & & 1 & 0 & -a_{n-2} \\ 0 & 0 & \ldots & 0 & 1 & -a_{n-1} \end{pmatrix} \in \mathrm{M}(n;K)$$

mit

$$f = \chi_A.$$

BEWEIS. Entwicklung nach der letzten Spalte. (Vgl. Übung.) □

DEFINITION 21.8. Die zum normierten Polynom $f \in K[T]$ wie im vorigen Lemma definerte quadratische Matrix A heißt die *Begleitmatrix* zu f.

SATZ 21.9. *Sei $f \in \operatorname{End}_K(V)$. Sei $\lambda \in K$. Äquiavlent sind:*

(1) *λ ist ein Eigenwert von f.*
(2) *λ ist eine Nullstelle des charakteristischen Polynoms χ_f.*

BEWEIS. Sei A die Darstellungsmatrix von f bezüglich einer Basis von V. Es gilt

$$\begin{aligned} \lambda \text{ Eigenwert von } f \quad &\Leftrightarrow \quad \operatorname{Kern}(\lambda \cdot 1_V - f) \neq \{0\} \\ &\Leftrightarrow \quad \lambda \cdot 1_V - f \text{ ist nicht bijektiv} \\ &\Leftrightarrow \quad \det(\lambda \cdot 1_V - f) = 0 \\ &\Leftrightarrow \quad \det(\lambda \cdot E_n - A) = 0 \\ &\Leftrightarrow \quad \chi_f(\lambda) = 0. \end{aligned}$$

□

DEFINITION 21.10. Sei f ein Endomorphismus von V. Ein Unterraum $U \subseteq V$ heißt *f-invariant*, falls $f(U) \subseteq U$ gilt.

BEMERKUNG 21.11. Sei f ein Endomorphismus von V.

(1) Sei $\lambda \in K$ ein Eigenwert von f. Dann ist der Eigenraum $E(f,\lambda)$ ein f-invarianter Unterraum von V.
(2) Sei V endlichdimensional. Die Einschränkung $f\,|_{E(f,\lambda)}$ hat bzgl. einer beliebigen Basis $(b_1, \ldots, b_k)$ die Darstellungsmatrix $\lambda \cdot E_k$.

Beweis. (1) Sei $U = E(f,\lambda)$. Für alle $u \in U$ gilt $f(f(u)) = f(\lambda u) = \lambda f(u)$, also gilt $f(u) \in E(f,\lambda) = U$.

(2) Dies ist klar, da für jedes $u \in U$ gilt $f(u) = \lambda u$. □

21.12. Sei $n = \dim(V)$, sei f ein Endomorphismus von V. Sei $x \in V$, $x \neq 0$ beliebig. Man betrachte $f(x)$, $(f(x)) = f^2(x)$, $f^3(x), \ldots$ Wegen $\dim(V) = n$ müssen die $n+1$ Elemente

$$x,\ f(x),\ f^2(x), \ldots,\ f^n(x)$$

linear abhängig sein. Es gibt also eine natürliche Zahl r mit $1 \leq r \leq n$, so dass

$$x,\ f(x),\ f^2(x), \ldots,\ f^{r-1}(x)$$

linear unabhängig, aber

$$x,\ f(x),\ f^2(x), \ldots,\ f^r(x)$$

linear abhängig sind. Es gibt dann $c_0, \ldots, c_{r-1}$ in K mit

$$f^r(x) = c_0 x + c_1 f(x) + \ldots + c_{r-1} f^{r-1}(x). \tag{21.1}$$

Der Unterraum

$$U \overset{def}{=} \operatorname{Span}\big(x,\ f(x), \ldots,\ f^{r-1}(x)\big)$$

hat die Basis $b_1 = x$, $b_2 = f(x), \ldots, b_r = f^{r-1}(x)$. Wendet man f auf die Basiselemente an, so erhält man

$$f(b_i) = b_{i+1} \quad i = 1, \ldots, r-1$$

und

$$f(b_r) = c_0 b_1 + c_1 b_2 + \ldots + c_{r-1} b_r.$$

Es ist also U ein f-invarianter Unterraum, und die Einschränkung $f\,|_U : U \to U$ hat bzgl. der Basis $(b_1, \ldots, b_r)$ die Darstellungsmatrix

$$A = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & c_0 \\ 1 & 0 & \ldots & \ldots & 0 & c_1 \\ 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & 0 & & \ddots & \vdots & \\ \vdots & \vdots & & 1 & 0 & c_{r-2} \\ 0 & 0 & \ldots & 0 & 1 & c_{r-1} \end{pmatrix}.$$

Aus Lemma 21.7 folgt

$$\chi_{f|_U} = T^r - c_{r-1} T^{r-1} - \ldots - c_1 T - c_0.$$

Aus (21.1) folgt dann

$$\big(\chi_{f|_U}(f)\big)(x) = 0. \tag{21.2}$$

Lemma 21.13. *Sei*

$$A = \left(\begin{array}{c|c} A_1 & * \\ \hline 0 & A_2 \end{array}\right).$$

Dann gilt

$$\chi_A = \chi_{A_1} \cdot \chi_{A_2}.$$

Beweis. Wir zeigen, dass generell die Formel

$$\left|\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right| = |A| \cdot |B|$$

für Matrizen A, $B = (\beta_{ij})$ über einem kommutativen Ring R gilt. Daraus folgt offenbar die Behauptung. Es gelte $A \in \mathrm{M}(p;R)$ und $B \in \mathrm{M}(q;R)$ (mit $p+q=n$). Wir machen Induktion nach q. Für $q = 0$ ist die Aussage klar. Auch für $q = 1$ folgt die Behauptung sofort durch Entwicklung nach der letzten Zeile. Sei nun $q \geq 2$. Entwicklung nach der letzten Zeile ergibt

$$\begin{aligned} \left|\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right| &= \sum_{j=p+1}^{n} (-1)^{j+n}\beta_{nj} \cdot \left|\begin{array}{c|c} A & * \\ \hline 0 & B'_{nj} \end{array}\right| \\ &\overset{IV}{=} \sum_{j=p+1}^{n} (-1)^{j+n}\beta_{nj}|B'_{nj}| \cdot |A| \\ &= |B| \cdot |A|. \end{aligned}$$

□

Satz 21.14 (Cayley-Hamilton). *Sei f ein Endomorphismus des endlichdimensionalen K-Vektorraum V. Dann gilt*

$$\chi_f(f) = 0.$$

Beweis. Zu zeigen ist, dass für jedes $x \in V$ die Gleichung $\chi_f(f)(x) = 0$ gilt. Man kann $x \neq 0$ annehmen. Sei U dazu der wie eben definierte f-invariante Unterraum mit der wie eben definierten Basis $b_1, \ldots, b_r$. Wir ergänzen dies mit $b_{r+1}, \ldots, b_n$ zu einer Basis von V. Bzgl. dieser Basis hat die Darstellungsmatrix die Form

$$\left(\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right),$$

wobei A auch wie oben definiert ist. Es folgt $\chi_f = \chi_A \cdot \chi_B = \chi_{f|_U} \cdot \chi_B$. Setzt man f ein und wendet dies auf x an, so erhält man aus (21.2)

$$\chi_f(f)(x) = 0.$$

□

Satz und Definition 21.15. *Sei f ein Endomorphismus des endlichdimensionalen K-Vektorraums V. Dann gibt es ein eindeutig bestimmtes normiertes Polynom $\mu_f \in K[T]$ kleinsten Grades mit $\mu_f(f) = 0$. Dies heißt das* Minimalpolynom *von f. Ferner gilt:*

(1) *Ist $P \in K[T]$ ein (normiertes) Polynom mit $P(f) = 0$, so gilt $\mu_f \mid P$.*
(2) *Insbesondere gilt $\mu_f \mid \chi_f$.*

Beweis. Nach dem Satz von Cayley-Hamilton gibt es ein normiertes Polynom μ_f mit $\mu_f(f) = 0$, und dann gibt es auch ein solches ($\neq 0$) kleinsten Grades. Die Eindeutigkeit wird aus (1) folgen, und (2) folgt trivialerweise aus (1). Ist nun $P \in K[T]$ mit $P(f) = 0$, so ergibt Polynomdivision mit Rest

$$P = q \cdot \mu_f + r$$

mit $r = 0$ oder $\operatorname{grad}(r) < \operatorname{grad}(\mu_f)$. Da aber

$$0 = P(f) = q(f) \cdot \mu_f(f) + r(f) = r(f)$$

gilt, kann, aufgrund der Minimalität des Grades von μ_f, nur $r = 0$ gelten. □

Satz 21.16. *Sei f ein Endomorphismus des n-dimensionalen K-Vektorraums V. Dann gilt für das Minimalpolynom μ_f und das charakteristische Polynom χ_f in $K[T]$:*

(1) $\mu_f \mid \chi_f$.
(2) $\chi_f \mid (\mu_f)^n$.
(3) *μ_f und χ_f haben dieselben Primfaktoren.*
(4) *μ_f und χ_f haben dieselben Nullstellen in K.*
(5) *χ_f zerfällt genau dann in Linearfaktoren, wenn dies für μ_f gilt.*

Beweis. (1) wurde schon gezeigt, und (3), (4) und (5) folgen sofort aus (1) und (2). Also bleibt (2) zu zeigen. Wir können K in einen algebraisch abgeschlossenen Körper L einbetten (vgl. Bemerkung 19.23). Über L zerfällt das χ_f in Linearfaktoren:

$$\chi_f = (T-\lambda_1)^{m_1} \cdot \ldots \cdot (T-\lambda_s)^{m_s},$$

wobei $\lambda_1, \ldots, \lambda_s$ in L liegen. Nach (1) gilt

$$\mu_f = (T-\lambda_1)^{k_1} \cdot \ldots \cdot (T-\lambda_s)^{k_s}$$

mit $k_i \leq m_i$ für $i = 1, \ldots, s$. Wir zeigen nun $k_i \geq 1$ für $i = 1, \ldots, s$. Denn ist etwa $k_i = 0$, und ist x_i ein Eigenvektor zum Eigenwert λ_i, so ergibt eine einfache Rechnung $\mu_f(f)x_i \neq 0$, aber andererseits ist $\mu_f(f) = 0$, Widerspruch.

Setze nun

$$q = (T-\lambda_1)^{n \cdot k_1 - m_1} \cdot \ldots \cdot (T-\lambda_s)^{n \cdot k_s - m_s}.$$

Damit gilt $q \cdot \chi_A = (\mu_A)^n$ in $L[T]$. Da aber χ_A, $(\mu_A)^n \in K[T]$ gilt, folgt auch leicht $q \in K[T]$. □

Lemma 21.17. (1) *Sei f ein Endomorphismus von V und $U \subseteq V$ ein f-invarianter Unterraum. Dann gilt $\chi_{f|_U} \mid \chi_f$.*

(2) *Sei $\pi \colon V \to W$ ein surjektiver Homomorphismus. Seien f und g Endomorphismen von V bzw. W mit $g \circ \pi = \pi \circ f$. Dann gilt $\chi_g \mid \chi_f$. (KOMMUTATIVES DIAGRAMM)*

Beweis. (1) Die Beweisidee tauchte schon im Beweis des Satzes von Cayley-Hamilton auf. Ergänzt man eine Basis von U zu einer Basis von V, so hat die diesbzgl. Darstellungsmatrix von f die Form

$$\left(\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right),$$

und es folgt $\chi_f = \chi_A \cdot \chi_B = \chi_{f|_U} \cdot \chi_B$.

(2) Man verfährt wie im Beweis des Rangsatzes. Eine Basis des Kerns von π wird ergänzt zu einer Basis von V. Die Bilder dieser ergänzten Vektoren unter π bilden eine Basis von W, und π induziert einen Isomorphismus zwischen dem Unterraum U von V, der durch die ergänzten Vektoren erzeugt wird, und W. Der komplementäre Summand $\mathrm{Kern}(\pi)$ ist wegen $\pi \circ f = g \circ \pi$ offenbar ein f-invarianter Unterraum. Es ist dann $\chi_f = \chi_{f|_{\mathrm{Kern}\,\pi}} \cdot \chi_B$ für eine quadratische Matrix B nach (1), und für diese Matrix gilt wegen $\pi \circ f = g \circ \pi$ $\chi_B = \chi_g$. □

Definition 21.18. Sei f ein Endomorphismus von V und $\lambda \in K$ ein Eigenwert.

(1) Die Vielfachheit $v_\lambda(\chi_f)$ der Nullstelle λ (vgl. 19.14) von χ_f, heißt die *algebraische* Vielfachheit von λ. Wir schreiben dafür $e(f,\lambda)$.

(2) Die Dimension des Eigenraums $E(f,\lambda)$ heißt die *geometrische Vielfachheit* von λ. Wir schreiben dafür $d(f,\lambda)$.

Lemma 21.19. *Sei $\lambda \in K$ ein Eigenwert des Endomorphismus f von V. Dann gilt*

$$1 \leq d(f,\lambda) \leq e(f,\lambda).$$

Beweis. Setze zur Abkürzung $d = d(f,\lambda)$ und $e = e(f,\lambda)$. $d \geq 1$ ist klar. Sei $x_1, \ldots, x_d$ eine Basis von $U = E(f,\lambda)$. Ergänze mit $x_{d+1}, \ldots, x_n$ zu einer Basis von V. Es ist U ein f-invarianter Unterraum von V, und $f\,|_U$ hat bzgl. $x_1, \ldots, x_d$ die Darstellungsmatrix $\lambda \cdot E_d$. Es folgt $\chi_{f|_U} = (T-\lambda)^d$, und nach Lemma 21.17 ist dies ein Teiler von χ_f, und es folgt $d \leq e$. □

Satz 21.20 (Diagonalisierbarkeitskriterium). *Sei $f \in \mathrm{End}_K(V)$. Äquivalent sind:*

(1) *f ist diagonalisierbar (über K).*

(2) (a) *Das charakteristische Polynom $\chi_f \in K[T]$ zerfällt in Linearfaktoren, und*
(b) *für jeden Eigenwert $\lambda \in K$ von f gilt $d(f,\lambda) = e(f,\lambda)$.*

(3) (a) *Das Minimalpolynom $\mu_f \in K[T]$ zerfällt in Linearfaktoren, und*
(b) *es hat nur einfache Nullstellen (Vielfachheit =1).*

Beweis. (1)⇒(2): Es gilt $V = E(f, \lambda_1) \oplus \ldots \oplus E(f, \lambda_s)$, wobei $\lambda_1, \ldots, \lambda_s \in K$ die paarweise verschiedenen Eigenvektoren sind. Setzt man Basen der einzelnen Eigenräume zu einer Basis von V zusammen, so ergibt sich als Darstellungsmatrix

$$A = \begin{pmatrix} \lambda_1 E_{n_1} & & & 0 \\ & \lambda_2 E_{n_2} & & \\ & & \ddots & \\ 0 & & & \lambda_s e_{n_s} \end{pmatrix}, \tag{21.3}$$

wobei $n_i = \dim(E(f, \lambda_i))$ gilt $(i = 1, \ldots, n)$. Es folgt dann

$$\chi_f = (T - \lambda_1)^{n_1} \cdot \ldots \cdot (T - \lambda_s)^{n_s},$$

und die beiden Aussagen aus (2) ergeben sich damit.

(2)⇒(1): Seien $\lambda_1, \ldots, \lambda_s \in K$ die verschiedenen Eigenwerte von f. Aus (2) folgt sofort

$$n = \operatorname{grad}(\chi_f) = e(f, \lambda_1) + \ldots + e(f, \lambda_s) = d(f, \lambda_1) + \ldots + d(f, \lambda_s),$$

und daher ist f nach Satz 20.6 diagonalisierbar.

(1)⇒(3): Wie oben können wir eine Darstellungsmatrix von f von der Form (21.3) annehmen. Es folgt sofort $(f - \lambda_1 1_V) \cdot \ldots \cdot (f - \lambda_s 1_V) = 0$. Also ist μ_f ein Teiler von $(T - \lambda_1) \cdot \ldots \cdot (T - \lambda_s) = 0$. Andererseits kommen nach Satz 21.16 alle Nullstellen von χ_f auch tatsächlich vor, also ist μ_f das Produkt der (einfachen) Linearfaktoren.

(3)⇒(1): Zunächst eine einfache Vorbemerkung: Sei $V = U_1 \oplus U_2$ mit f-invarianten Unterräumen U_1 und U_2. Einzelne Basen von U_1 und U_2 setzt man zu einer Basis von V zusammen. Die zugehörige Darstellungsmatrix hat die Blockdiagonalform

$$A = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right).$$

Für jedes $p \in K[T]$ gilt

$$p(A) = \left(\begin{array}{c|c} p(A_1) & 0 \\ \hline 0 & p(A_2) \end{array}\right).$$

Ist $p(A) = 0$, so gilt auch $p(A_1) = 0$ und $p(A_2) = 0$, und daher gilt $\mu_{f|_{U_1}} \mid \mu_f$ und $\mu_{f|_{U_2}} \mid_f$. Insbesondere: Zerfällt μ_f in einfache Linearfaktoren, so auch $\mu_{f|_{U_1}}$ und $\mu_{f|_{U_2}}$.

Sei nun

$$\mu_f = (T - \lambda_1) \cdot \ldots \cdot (T - \lambda_s)$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_s \in K$. Wir zeigen nun, dass f diagonalisierbar ist per Induktion nach $n = \dim(V)$. Für $n = 1$ ist nichts zu zeigen. Sei nun $n \geq 2$. Wir zeigen

$$V = \operatorname{Kern}(f - \lambda_1 1_V) \oplus \operatorname{Bild}(f - \lambda_1 1_V).$$

Dies ist eine direkte Zerlegung in f-invariante Unterräume: Für $\operatorname{Kern}(f - \lambda_1 1_V) = E(f, \lambda_1)$ ist dies bekannt. Sei $y \in \operatorname{Bild}(f - \lambda_1 1_V)$. Es gibt ein $x \in V$ mit $y = f(x) - \lambda_1 x$. Es gilt $f(y) = f^2(x) - \lambda_1 f(x) = (f - \lambda_1 1_V)(f(x)) \in \operatorname{Bild}(f - \lambda_1 1_V)$. – Hat man diese direkte Zerlegung, so folgt die Aussage mit der Vorbemerkung per Induktion. Es genügt also, diese direkte Zerlegung zu zeigen.

Da λ_1 verschieden ist von $\lambda_2, \ldots, \lambda_s$, haben die Polynome $T - \lambda_1$ und

$$(T - \lambda_2) \cdot \ldots \cdot (T - \lambda_s)$$

sicherlich keine gemeinsamen Teiler. Der größte (normierte) gemeinsame Teiler ist daher 1. Aus 19.19 folgt daher, dass es p, $q \in K[T]$ gibt mit

$$1 = p \cdot (T - \lambda_1) + q \cdot (T - \lambda_2) \cdot \ldots \cdot (T - \lambda_s).$$

Setzt man f ein, so erhhält man

$$1_V = p(f) \circ (f - \lambda_1 1_V) + q(f) \circ (f - \lambda_2 1_V) \circ \ldots \circ (f - \lambda_s 1_V).$$

Für jedes $x \in V$ folgt

$$x = p(f) \circ (f - \lambda_1 1_V)(x) + q(f) \circ (f - \lambda_2 1_V) \circ \ldots \circ (f - \lambda_s 1_V)(x).$$

Der erste Summand liegt in $\mathrm{Bild}(f - \lambda_1 1_V)$, der zweite in $\mathrm{Kern}(f - \lambda_1 1_V)$. Letzteres folgt aus $\mu_f(f) = 0$. Damit bekommt man $V = \mathrm{Kern}(f - \lambda_1 1_V) + \mathrm{Bild}(f - \lambda_1 1_V)$. Dass die Summe direkt ist, folgt aus Dimensiongründen aus dem Rangsatz. □

Definition 21.21.
(1) $f \in \mathrm{End}_K(V)$ heißt *trigonalisierbar*, wenn es eine Basis von V gibt, bzgl. welcher f durch eine obere Dreicksmatrix dargestellt wird.
(2) Entsprechend heißt $A \in \mathrm{M}(n; K)$ *trigonalisierbar*, wenn es ein $P \in \mathrm{GL}(n; K)$ gibt, so dass $P^{-1}AP$ eine obere Dreiecksmatrix ist.

Man könnte in der Definition auch untere statt obere Dreiecksmatrizen verwenden. Wir kommen darauf später noch zurück.

Satz 21.22 (Trigonalisierbarkeitskriterium). *Für $f \in \mathrm{End}_K(V)$ sind folgende Aussagen äquivalent:*

(1) *f ist trigonalisierbar (über K).*
(2) *Das charakteristische Polynom $\chi_f \in K[T]$ zerfällt in Linearfaktoren.*
(3) *Das Minimalpolynom $\mu_f \in K[T]$ zerfällt in Linearfaktoren.*

Beweis. Die Aussagen (2) und (3) sind nach 21.16 äquivalent.

(1)⇒(2): Dies folgt aus der folgenden offensichtlichen Aussage für Matrizen: Ist

$$A = \begin{pmatrix} \lambda_1 & & * & \\ & \lambda_2 & & \\ & & \ddots & \\ & 0 & & \lambda_n \end{pmatrix} \in \mathrm{M}(n; K),$$

so ist $\chi_A = (T - \lambda_1) \cdot \ldots \cdot (T - \lambda_n)$.

(2)⇒(1): Induktion nach $n = \dim(V)$. Für $n = 1$ ist nichts zu zeigen. Gelte nun $n \geq 2$. Da χ_f in Linearfaktoren zerfällt, hat f mindestens einen Eigenwert λ_1. Sei x_1 ein zugehöriger Eigenvektor. Ergänze diesen zu einer Basis $x_1, ß, x_2, \ldots, x_n$ von V. Dann hat die zugehörige Darstellungsmatrix die Form

$$A = \left(\begin{array}{c|ccc} \lambda_1 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$

mit einem $B \in \mathrm{M}(n-1; K)$, und es folgt

$$\chi_f = \chi_A = (T - \lambda_1) \cdot \chi_B.$$

Es folgt, dass auch χ_B in Linearfaktoren zerfällt. Sei U der von $x_2, \ldots, x_n$ erzeugte Unterraum. Ist $u \in U$, so gibt es eine eindeutige Darstellung $f(u) = \alpha x_1 + g(u)$ mit $g(u) \in U$. Dies definiert offenbar eine lineare Abbildung $g\colon U \to U$. Bzgl. der Basis $x_2, \ldots, x_n$ von U hat g die Darstellungsmatrix B. Per Induktionsvoraussetzung ist g trigonalisierbar. Es gibt also eine Basis $x'_2, \ldots, x'_n$, bzgl. der die Darstellungsmatrix von g eine obere Dreiecksmatrix ist. Es folgt dann, dass bzgl. der Basis $x_1, x'_2, \ldots, x'_n$ von V die Darstellungsmatrix von f eine obere Dreiecksmatrix ist. □

Folgerung 21.23. *Ist K ein algebraisch abgeschlossener Körper (z. B. $K = \mathbb{C}$), so ist jeder Endomorphismus von V trigonalisierbar.*

Der Induktionsbeweis des vorigen Satzes liefert ein Konstruktionsverfahren für eine Basis, bzgl. derer ein vorgegebener trigonalisierbarer Endomorphismus durch eine obere Dreiecksmatrix dargestellt wird. Die wird durch folgendes Beispiel illustriert.

Beispiel 21.24. Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ die lineare Abbildung $x \mapsto Ax$, wobei

$$A = \begin{pmatrix} 17 & 9 & -6 \\ -13 & -6 & 5 \\ 17 & 10 & -5 \end{pmatrix}.$$

Man rechnet

$$\begin{aligned} \chi_A &= \begin{vmatrix} T-17 & -9 & 6 \\ 13 & T+6 & -5 \\ -17 & -10 & T+55 \end{vmatrix} = \dots \\ &= T^3 - 6T^2 + 12T - 8 = (T-2)^3. \end{aligned}$$

Einziger Eigenwert ist also $\lambda = 3$, und f ist trigonalisierbar. Zur Bestimmung einer Basis des Eigenraums $E(A, 2)$ berechnet man die Treppenmatrix zu

$$\begin{pmatrix} 2-17 & -9 & 6 \\ 13 & 2+6 & -5 \\ -17 & -10 & 2+55 \end{pmatrix},$$

und die ist durch

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

gegeben. Damit bildet $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ ein Basis von $E(A, 2)$. Wir ergänzen diese nun zu einer Basis von $\mathbb{R}^3$, etwa:

$$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(Allgemein wird man das Verfahren aus Satz 10.4 benutzen.) Sei

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Dann gilt

$$P_1^{-1} A P_1 = \left(\begin{array}{c|cc} 2 & 9 & -6 \\ \hline 0 & 3 & -1 \\ 0 & 1 & 1 \end{array} \right).$$

Jetzt betrachtet man die 2×2-Untermatrix $B = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$ und wendet das Verfahren erneut auf diese kleinere Matrix an. Es ist weiterhin $\lambda = 2$ der einzige Eigenwert von B. Es ergibt sich, dass $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ eine Basis von $E(B, 2)$ bildet. Dies wird zu einer Basis von $\mathbb{R}^2$ ergänzt, z. B. durch $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Setzt man

$$P_2 = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 1 & 0 \end{array} \right)$$

und

$$P = P_1 P_2 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

so erhält man

$$\begin{aligned} P^{-1}AP &= P_2^{-1}P_1^{-1}AP_1P_2 \\ &= \begin{pmatrix} 2 & 3 & 9 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}. \end{aligned}$$

Bzgl. der durch

$$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

gegebenen Basis von $\mathbb{R}^3$ wird f durch diese obere Dreiecksmatrix beschrieben.

Ferner haben wir gesehen:

- Die algebraische Vielfachheit von $\lambda = 2$ ist $e(f, 2) = 3$, die geometrische Vielfachheit ist $d(f, 2) = \dim(E(f, 2)) = 1$. Also ist f *nicht* diagonalisierbar.
- Es ist μ_f ein Teiler von $\chi_f = (T-2)^3$, muss also eines von $T-2$, $(T-2)^2$ und $(T-2)^3$ sein. Berechnet man die Potenzen von $A - 2 \cdot E_3$, so sieht man, dass $\chi_f = (T-2)^3 = \chi_f$ gilt.

Bemerkung 21.25. (1) Allgemein stimmen Minimalpolynom und charakteristisches Polynom nicht überein. Z. B. ist für die Nullmatrix $A = 0 \in \mathrm{M}(n; K)$ offenbar $\chi_A = T^n$ und $\mu_A = T$. Hier ist der Unterschied also sehr drastisch. (Vgl. auch Übungen IV.1.) Später werdn wir sehen, wie allgemein das Minimalpolynom einer trigonalisierbaren Matrix bestimmt werden kann.

(2) Ein Rechenbeispiel zur Diagonalisierung lassen wir hier beiseite. (Vgl. Übungen.) Ist nämlich ein Endomorphismus f von V diagonalisierbar, so haben wir Basen der Eigenräume zu berechnen (voraussetzend, dass man alle Eigenwerte ermittelt hat). Dies ist nichts anderes als die Bestimmung der Basis eines Kerns bzw. des Lösungsraumes eines homogenen linearen Gleichungssystems. Die einzelnen Basen fasst man dann zu einer Gesamtbasis von V zusammen.

22. Endomorphismen als Moduln

Es bezeichne K stets einen Körper. Zu weiteren Strukturuntersuchungen von Endomorphismen führen wir nun eine neue Definition ein, die ganz praktisch ist.

Definition 22.1. Sei ein K-Vektorraum V zusammen mit einem Endomorphismus f, also ein geordnetes Paar (V, f), heißt ein *Modul*.

Diese Terminologie ist etwas provisorisch, aber in diesem Zusammenhang recht nützlich. Allgemein nennt man die den Vektorräumen über Körpern entsprechenden Gebilde über Ringen Moduln. Man kann zeigen, dass ein Modul (V, f) wie zuvor definiert nichts anderes als ein Modul über dem Polynomring $K[T]$ ist. Aber das ist im folgenden unerheblich.

Im folgenden nehmen wir stets an, dass V endlichdimensional ist, auch wenn das nicht in jeder Aussage benötigt wird.

Definition 22.2. Sei (V, f) ein Modul. Ein f-invarianter Unterraum U, genauer das Paar bzw. der Modul $(U, f|_U)$, heißt ein *Untermodul* von (V, f).

Untermoduln von (V, f) sind also gerade die f-invarianten Unterräume von V, für die wir uns ja interessieren. Wichtig ist nun, dass man für Moduln strukturverträgliche Abbildungen definieren kann.

DEFINITION 22.3. Seien (V, f) und (W, g) Moduln über dem Grundkörper K. Eine K-lineare Abbildung $h\colon V \to W$ heißt *Modulhomomorphismus*, falls $g \circ h = h \circ f$ gilt. (KOMMUTATIVES DIAGRAMM)

Wir schreiben dann auch $h\colon (V, f) \to (W, g)$.

Ist h zusätzlich bijektiv, so heißt h ein Modulisomorphismus.

Ist h ein Modulisomorphismus, so ist auch h^{-1} ein Modulisomorphismus. (Gilt $g \circ h = h \circ f$, so folgt $h^{-1} \circ g = f \circ h^{-1}$.)

Die folgenden Aussagen sind Routine.

BEMERKUNG 22.4. Sei $h\colon (V, f) \to (W, g)$ ein Modulhomomorphismus.
(1) Ist U ein Untermodul von (V, f), so ist $h(U)$ ein Untermodul von (W, g).
(2) Ist X ein Untermodul von (W, g), so ist das Urbild $h^{-1}(X)$ ein Untermodul von (V, f).
(3) $\mathrm{Bild}(h)$ ist ein Untermodul von (W, g).
(4) $\mathrm{Kern}(h)$ ist ein Untermodul von (V, f).

Die Aussagen (3) und (4) sind offenbar Spezialfälle der Aussagen (1) bzw. (2), sind uns als Aussagen über f-invariante Unterräume in anderer Terminologie in Spezialfällen schon begegnet.

Offensichtlich gilt außerdem:

BEMERKUNG 22.5. (1) $1_V\colon (V, f) \to (V, f)$ ist ein Modulhomomorphismus.
(2) Komposition von Modulhomomorphismen ist wieder ein Modulhomomorphismus.

BEMERKUNG 22.6. (1) Wir fassen jede Matrix $A \in \mathrm{M}(n; K)$ als Modul (K^n, f) auf, wobei f der Endomorphismus von K^n ist, der durch $x \mapsto Ax$ gegeben ist. Wir schreiben hierfür auch einfach (K^n, A).
(2) Aus der Definition der Isomorphie von Moduln folgt: Zwei Moduln (K^n, A) und (K^n, B) sind isomorph genau dann, wenn die Matrizen A und B ähnlich sind.
(3) Zwei Moduln (V, f) und (V, g) sind genau dann isomorph, wenn es Basen $\mathcal{B}$ und $\mathcal{C}$ von V gibt mit $M_{\mathcal{B}}(f) = M_{\mathcal{C}}(g)$. (Dies folgt aus der Transformationsformel.)

Das bisherige Problem – Klassifikation von quadratsichen Matrizen bis auf Ähnlichkeit – hat sich nun in das Problem übersetzt, Moduln bis auf Isomorphie zu klassifizieren. Wir zerlegen ein solches Problem in kleinere Teilprobleme:

DEFINITION 22.7. Sei (V, f) ein Modul.
(1) Ist $V = U_1 \oplus \ldots \oplus U_s$ direkte Summe von f-invarianten Unterräumen, so schreiben wir
$$(V, f) = (U_1, f\,|_{U_1}) \oplus \ldots \oplus (U_s, f\,|_{U_s})$$
und nennen dies eine direkte Zerlegung von (V, f) in die Untermoduln $(U_1, f\,|_{U_1}), \ldots, (U_s, f\,|_{U_s})$.
(2) (V, f) heißt *unzerlegbar*, wenn $V \neq 0$ gilt, und wenn aus einer direkten Zerlegung $(V, f) = (U, f\,|_U) \oplus (W, f\,|_W)$ folgt, dass $U = 0$ oder $W = 0$ gilt.

BEISPIEL 22.8. Der Modul (K^n, f), wobei f bzgl. der Standardbasis $e_1, \ldots, e_n$ durch die Matrix
$$J_n(0) \overset{def}{=} \begin{pmatrix} 0 & 0 & \ldots & & 0 \\ 1 & 0 & & & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & & 1 & 0 \end{pmatrix}$$
gegeben ist, ist unzerlegbar.

Beweis. Es gilt $f(e_1) = e_2$, $f(e_2) = e_3, \ldots, f(e_{n-1}) = e_n$, $f(e_n) = 0$. Nehme an $V = U_1 \oplus U_2$ mit f-invarianten Unterräumen U_1, $U_2 \neq \{0\}$. Sei $x = \alpha_k e_k + \ldots + \alpha_n e_n \in U_1$, mit $\alpha_k \neq 0$. Es folgt $\alpha_k e_n = f^{n-k}(x) \in U_1$, also $e_n \in U_1$. Genauso folgt $e_n \in U_2$, Widerspruch zu $U_1 \cap U_2 = 0$. $\square$

Satz 22.9. *Sei*

$$(V, f) = (U_1, f\,|_{U_1}) \oplus \ldots \oplus (U_s, f\,|_{U_s})$$

ein direkte Zerlegung des Moduls (V, f). *Sei* $\mathcal{B}_i$ *eine Basis von* U_i *und* A_i *die Darstellungsmatrix von* $f\,|_{U_i}$ *bzgl.* $\mathcal{B}_i$ $(i = 1, \ldots, s)$.

(1) $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_s$ *eine Basis von* V.

(2) *Es ist*

$$M_{\mathcal{B}}(f) = \left(\begin{array}{c|c|c|c} A_1 & & & \\ \hline & A_2 & & \\ \hline & & \ddots & \\ & & \quad\ddots & \\ \hline & & & A_s \end{array}\right)$$

von Blockdiagonalgestalt.

Beweis. (1) Da $V = U_1 \oplus \ldots \oplus U_s$ eine direkte Summe ist, folgt leicht, dass $\mathcal{B}_1 \cup \ldots \cup \mathcal{B}_s$ eine Basis von V ist.

(2) Dies folgt dann auch sofort wegen der f-Invarianz der Unterräume U_i. $\square$

Satz 22.10. *Sei* (V, f) *ein Modul. (*V *endlichdimensional). Dann gilt*

$$(V, f) = (U_1, f\,|_{U_1}) \oplus \ldots \oplus (U_s, f\,|_{U_s})$$

mit unzerlegbaren *Untermoduln* $(U_1, f\,|_{U_1}), \ldots, (U_s, f\,|_{U_s})$.

Beweis. Induktion nach $n = \dim(V)$. Für $n = 0$ gilt die Aussage (mit $s = 0$). Ist $n = 1$, so ist (V, f) offenbar selbst unzerlegbar aus Dimensionsgründen (die Aussage gilt also mit $s = 1$). Sei nun $n \geq 2$. Entweder ist (V, f) selbst unzerlegbar (und dann gilt die Aussage mit $s = 1$), oder man kann schreiben $(V, f) = (U, f\,|_U) \oplus (W, f\,|_W)$, wobei $\dim(U)$, $\dim(W) \geq 1$ gilt, also $\dim(U)$, $\dim(W) \leq n - 1$. Nach Induktionsvoraussetzung zerlegen sich sowohl der Untermodul $(U, f\,|_U)$ wie auch der Untermodul $(W, f\,|_W)$ in eine direkte Summe von unzerlegbaren Untermoduln, und dann gilt dies auch für (V, f). $\square$

Bei dem Beweis wurde im letzten Schritt ein Argument über direkte Summen verwendet, das man sich leicht klar macht:

$$V = U \oplus W,\ U = U_1 \oplus U_2 \quad \Rightarrow \quad V = U_1 \oplus U_2 \oplus W.$$

Durch die beiden vorigen Sätze, ist das Problem der Beschreibung der Struktur eines Endomorphismus reduziert auf den unzerlegbaren Fall.

Definition 22.11. Ein Endomorphismus f von V heißt *nilpotent*, wenn es ein $q \in \mathbb{N}$ gibt mit $f^q = 0$. Entsprechend werden nilpotente (quadratische) Matrizen definiert.

Satz 22.12 (Fittings Lemma). *Sei* (V, f) *unzerlegbar. Dann ist jeder Endomorphismus* g *von* (V, f) *bijektiv oder nilpotent.*

Beweis. Für jedes $i \in \mathbb{N}$ sei $B_i = \text{Bild}(g^i)$ und $K_i = \text{Kern}(g^i)$. Man bekommt damit eine fallende bzw. eine aufsteigende Kette von f-invarianten Unterräumen von V:

$$B_1 \supseteq B_2 \supseteq \ldots \supseteq B_i \supseteq B_{i+1} \supseteq \ldots$$
$$K_1 \subseteq K_2 \subseteq \ldots \subseteq K_i \subseteq K_{i+1} \subseteq \ldots$$

Wegen $\dim(V) < \infty$ muss es ein $n \in \mathbb{N}$ geben mit $B_n = B_{n+1} = \ldots$ und mit $K_n = K_{n+1} = \ldots$.

Behauptung: Es gilt $V = B_n \oplus K_n$. Beweis: Sei $x \in V$. Wegen $B_n = B_{2n}$ gibt es ein $y \in V$ mit $g^n(x) = g^{2n}(y)$. Es folgt $x = g^n(y) + (x - g^n(y) \in B_n + K_n$. Es folgt also $V = B_n + K_n$. Sei $x \in B_n \cap K_n$. Dann gibt es ein $y \in V$ mit $x = g^n(y)$, und es gilt $g^n(x) = 0$, und somit $g^{2n}(y) = g^n(x) = 0$. Es folgt $y \in K_{2n} = K_n$, also folgt schon $x = g^n(y) = 0$. Dies zeigt $B_n \cap K_n = \{0\}$.

Insgesamt folgt also $V = B_n \oplus K_n$. Dies ist eine direkte Zerlegung in f-invariante Unterräume. Da (V, f) unzerlegbar is, folgt dann entweder $B_n = \{0\}$ oder $K_n = \{0\}$. Im ersten Fall folgt $g^n = 0$, d. h. g ist nilpotent, im zweiten Fall folgt, dass g^n injektiv ist; dann ist aber g selbst injektiv, und dann als Endomorphismus des endlichdimensionalen Vektorraums V auch surjektiv. □

23. Die Jordansche Normalform

Fittings Lemma legt uns nahe, mit nilpotenten Endomorphismen anzufangen.

Lemma 23.1. *Sei f ein Endomorphismus von V, und es gelte $f^n = 0$, $f^{n-1} \neq 0$. Sei S ein direktes Komplemente von* $\text{Kern}(f^{n-1})$*, es gelte also* $V = S \oplus \text{Kern}(f^{n-1})$*. Dann gilt*

(1) *Die Einschränkungen*

$$S \xrightarrow{f} f(S) \xrightarrow{f} \dots \xrightarrow{f} f^{n-1}(S)$$

sind allesamt Isomorphismen.

(2) *Die Summe* $\langle S \rangle \overset{def}{=} S \oplus f(S) \oplus \dots \oplus f^{n-1}(S)$ *ist direkt und bildet einen f-invarianten Unterraum von V.*

(3) $\langle S \rangle$ *besitzt ein f-invariantes Komplement in V, d. h. es gibt einen f-invarianten Unterraum H von V mit $V = \langle S \rangle \oplus H$.*

Beweis. (1) Es genügt zu zeigen, dass der Kern jeder dieser Einschränkungen trivial ist, dass also $f^i(S) \cap \text{Kern}(f) = 0$ gilt $(i = 0, \dots, n-2)$. Sei $s \in S$ und $x = f^i(s)$ mit $f(x) = 0$. Es folgt $0 = f^{n-i-1}(x) = f^{n-1}(s)$ (denn es ist der Exponent $n-i-1 \geq 1$), also $s \in S \cap \text{Kern}(f^{n-1}) = \{0\}$, also $s = 0$ und damit auch $x = 0$.

(2) Die f-Invarianz von $\langle S \rangle$ folgt wegen $f^n = 0$ unmittelbar aus der Definition von $\langle S \rangle$. Zur Direktheit: Angenommen, es gibt eine Darstellung des Nullelements als $f^i(s_i) + \dots f^{n-1}(s_{n-1}) = 0$ mit $s_i \neq 0$. Anwendung von f^{n-i-1} liefert $0 = f^{n-i-1}(f^i(s_i)) = f^{n-1}(s_i)$, also $s_i \in S \cap \text{Kern}(f^{n-1}) = \{0\}$, Widerspruch.

(3) Induktion nach n. $n = 1$ bedeutet $f = 0$, und dann ist jeder Unterraum von V f-invariant. Gelte nun $f^{n+1} = 0$, $f^n \neq 0$ und $V = S \oplus \text{Kern}(f^n)$. Schreibe $\text{Kern}(f^n) = T \oplus \text{Kern}(f^{n-1})$. Wegen $f(S) \cap \text{Kern}(f^{n-1}) = \{0\}$ (folgt aus (1)) kann man

$$f(S) \subseteq T \tag{23.1}$$

annehmen. Die Induktionsvoraussetzung, angewandt auf $\text{Kern}(f^n)$, liefert einer Zerlegung

$$\text{Kern}(f^n) = [T \oplus f(T) \oplus \dots \oplus f^{n-1}(T)] \oplus H$$

für einen f-invarianten Unterraum H von H von $\text{Kern}(f^n)$. Wegen (23.1) gibt es eine Zerlegung (von Vektorräumen)

$$T = f(S) \oplus U.$$

Es folgt

$$\begin{aligned} V &= S \oplus [f(S) \oplus f^2(S) \oplus \dots \oplus f^n(S)] \oplus [U \oplus f(U) \oplus \dots \oplus f^{n-1}(U)] \oplus H \\ &= \langle S \rangle \oplus \langle U \rangle \oplus H. \end{aligned}$$

Also ist $\langle U \rangle \oplus H$ ein f-invariantes Komplement von $\langle S \rangle$ in V. □

Satz 23.2 (Normalform für nilpotente Endomorphismen (unzerlegbarer Fall)). *Sei (V, f) unzerlegbar mit $f^n = 0$, $f^{n-1} \neq 0$. Dann gibt es eine n-elementige Basis von V, bzgl. welcher f durch die Matrix*

$$J_n(0) = \begin{pmatrix} 0 & 0 & \dots & & 0 \\ 1 & 0 & & & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & & 1 & 0 \end{pmatrix}$$

beschrieben wird.

Beweis. Sei S ein direktes Komplement von $\mathrm{Kern}(f^{n-1})$, also $V = S \oplus \mathrm{Kern}(f^{n-1})$. Aus dem vorigen Lemma folgt zusammen mit der Unzerlegbarkeit von (V, f), dass

$$V = \langle S \rangle = S \oplus f(S) \oplus \ldots \oplus f^{n-1}(S)$$

gilt. Jede nichttriviale Zerlegung $S = S' \oplus S''$ (in Unterräume) liefert damit eine Zerlegung $V = \langle S \rangle = \langle S' \rangle \oplus \langle S'' \rangle$ in f-invariante Unterräume. Da aber (V, f) unzerlegbar ist, muss also S eindimensional sein, $S = Ke$ für ein $e \in S$. Es ist dann

$$e,\ f(e), \ldots, f^{n-1}(e)$$

eine Basis von V, und bzgl. dieser hat f offenbar die Darstellungsmatrix $J_n(0)$. □

Satz 23.3 (Jordansche Normalform (unzerlegbarer Fall)). *Sei (V, f) unzerlegbar. Wir nehmen an, dass f einen Eigenwert $\lambda \in K$ besitzt. Dann gibt es eine Basis von V (bestehend aus $n = \dim(V)$ Elementen) bzgl. welcher f durch die Matrix*

$$J_n(\lambda) \overset{def}{=} \begin{pmatrix} \lambda & 0 & \dots & & 0 \\ 1 & \lambda & & & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & & 1 & \lambda \end{pmatrix}$$

beschrieben wird.

Wir nennen $J_n(\lambda)$ ein *Jordankästchen* (vom Format n) zum Eigenwert λ.

Beweis. Sei $g \overset{def}{=} f - \lambda \cdot 1_V$. Es gilt $g \circ f = f \circ g$, also ist g ein Endomorphismus von (V, f). Wegen $\mathrm{Kern}(g) = \mathrm{Kern}(f - \lambda \cdot 1_V) = E(f, \lambda) \neq \{0\}$ ist g nicht bijektiv, nach Fittings Lemma also nilpotent. Nach dem vorigen Satz gibt es eine Basis $\mathcal{B} = (x_1, \ldots, x_n)$ mit $M_{\mathcal{B}}(g) = J_n(0)$. Es folgt dann

$$M_{\mathcal{B}}(f) = M_{\mathcal{B}}(g + \lambda \cdot 1_V) = M_{\mathcal{B}}(g) + M_{\mathcal{B}}(\lambda \cdot 1_V) = J_n(0) + \lambda \cdot E_n = J_n(\lambda).$$

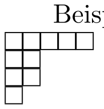
□

Definition 23.4. Sei $n \in \mathbb{N}$.

(1) Eine Folge $\mathbf{n} = (n_1, \ldots, n_s)$ mit natürlichen Zahlen $n_1 \geq n_2 \geq \ldots \geq n_s \geq 1$ heißt *Partition* von n, falls $n_1 + n_2 + \ldots + n_s = n$ gilt.

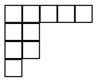
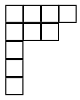
(2) Das *Young-Diagramm* einer Partition $\mathbf{n} = (n_1, \ldots, n_s)$ von n ist ein Schema, dass aus s Zeilen besteht. Die i-Zeile besteht aus n_i Kästchen, an den Positionen $(i, 1)$, $(i, 2), \ldots, (i, n_i)$.

Beispiel: Ist $\mathbf{n} = (5, 2, 2, 1)$, so ist das Young-Diagramm von $\mathbf{n}$ gegeben als



(3) Ist $\mathbf{n} = (n_1, \ldots, n_s)$ eine Partition von n, so ist die duale Partition $\mathbf{n}^*$ diejenige Partition von n, deren Young-Diagramm man aus dem Young-Diagramm von $\mathbf{n}$ durch "Transponieren" erhält.

Beispiel: Zu $(5, 2, 2, 1)$ ist $(4, 3, 1, 1, 1)$ dual:

 wird durch Transponieren zu 

Satz 23.5 (Normalform für nilpotente Endomorphismen (allgemeiner Fall)). *Sei V ein n-dimensionaler K-Vektorraum und f ein nilpotenter Endomorphismus von V. Dann gibt es eine Basis $\mathcal{B}$ von V und eine Partition $\mathbf{n} = (n_1, dots, n_s)$ von n, so dass*

$$M_{\mathcal{B}}(f) = \left(\begin{array}{c|c|c|c} J_{n_1}(0) & & & \\ \hline & J_{n_2}(0) & & \\ \hline & & \ddots & \\ & & & \ddots \\ \hline & & & J_{n_s}(0) \end{array}\right) \stackrel{def}{=} N(n_1, \ldots, n_s)$$

gilt. Ferner gilt:

(1) $\chi_f = T^n$, $\mu_f = T^{n_1}$.

(2) $\dim E(f, 0) = s$. *(Dies ist die geometrische Vielfachheit von $\lambda = 0$.)*

(3) $\mathbf{n} = (n_1, \ldots, n_s)$ *ist die zur Partition $\mathbf{m} = (m_1, \ldots, m_q)$ duale Partition von n, wobei*

$$m_i \stackrel{def}{=} \dim \mathrm{Kern}(f^i) - \dim \mathrm{Kern}(f^{i-1})$$

und $f^q = 0$, $f^{q-1} \neq 0$ gilt. (Es gilt $q = n_1$.) Die Kästchengrößen sind also eindeutig durch f bestimmt (sofern absteigend angeordnet; sonst nur bis auf Reihenfolge).

Beweis. Der Modul (V, f) zerlegt sich in eine direkte Summe unzerlegbarer Moduln, $(V, f) = (V_1, f_1) \oplus \ldots \oplus (V_s, f_s)$. Als Einschränkungen von f sind alle $f_i \colon V_i \to V_i$ nilpotent. Ist $\dim(V_i) = n_i$, so gibt es nach Satz 23.2 eine Basis $\mathcal{B}_i$ von V_i mit $M_{\mathcal{B}_i}(f_i) = J_{n_i}(0)$. Für die Basis $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_s$ gilt dann

$$M_{\mathcal{B}} = \mathrm{diag}\big(J_{n_1}(0),\, J_{n_2}(0), \ldots, J_{n_s}(0)\big),$$

wobei man ohne Einschränkung $n_1 \geq \ldots \geq n_s$ annehmen kann.

Zur Aussage (1): $\chi_f = T^n$ ist klar. Ist $A = \mathrm{diag}\big(J_{n_1}(0),\, J_{n_2}(0), \ldots, J_{n_s}(0)\big)$, so gilt offenbar $A^k = \mathrm{diag}\big(J_{n_1}(0)^k,\, J_{n_2}(0)^k, \ldots, J_{n_s}(0)^k\big)$ für jede natürliche Zahl k. Da n_1 das maximale Format ist, folgt $J_{n_i}(0)^{n_1} = 0$ für $i = 1, \ldots, s$, aber $J_{n_1}(0)^{n_1 - 1} \neq 0$. Daher ist T^{n_1} das Minimalpolynom von f.

Zur Aussage (2): Es gilt offenbar $\mathrm{rang}(J_{n_i}(0)) = n_i - 1$ für jedes $i = 1, \ldots, s$, und man sieht, dass sich für $\mathrm{diag}\big(J_{n_1}(0),\, J_{n_2}(0), \ldots, J_{n_s}(0)\big)$ die Ränge aufaddieren, d. h.

$$\mathrm{rang}(f) = \sum_{i=1}^{s} (n_i - 1) = n - s.$$

Mit dem Rangsatz folgt

$$\dim E(f, 0) = \dim \mathrm{Kern}(f) = \dim(V) - \mathrm{rang}(f) = n - (n - s) = s.$$

Zur Aussage (3): Gelte $f^q = 0$ und $f^{q-1} \neq 0$. Dann gilt

$$\{0\} = \mathrm{Kern}(f^0) \subsetneq \mathrm{Kern}(f^1) \subsetneq \ldots \subsetneq \mathrm{Kern}(f^{q-1}) \subsetneq \mathrm{Kern}(f^q) = V.$$

Wir werden weiter unten zeigen, dass für jedes $i \in \{1, \ldots, q\}$ gilt:

(23.2) $$\mathrm{Kern}(f^i) = \mathrm{Kern}(f^{i-1}) \oplus W_i \text{ mit } f\,|_{W_i} \colon W_i \to W_{i-1} \text{ ist injektiv.}$$

Setzt man $m_i = \dim(W_i) = \dim \operatorname{Kern}(f^i) - \dim \operatorname{Kern}(f^{i-1})$, so folgt daraus

$$\begin{aligned} \sum_{i=1}^{q} m_i &= \sum_{i=1}^{q} \Big(\dim \operatorname{Kern}(f^i) - \dim \operatorname{Kern}(f^{i-1}) \Big) \\ &= -\dim \operatorname{Kern}(f^0) + \dim \operatorname{Kern}(f^q) \\ &= \dim(V) = n, \end{aligned}$$

und außerdem $m_1 \geq m_2 \geq \ldots \geq m_q \geq 1$. Es ist also $\mathbf{m} = (m_1, \ldots, m_q)$ eine Partition von n.

Ferner gilt nach dem Rangsatz

$$\begin{aligned} m_i &= \operatorname{rang}(f^{i-1}) - \operatorname{rang}(f^i) \\ &\overset{(*)}{=} \sum_{j=1}^{s} \Big(\operatorname{rang}(J_{n_j}(0)^{i-1} - \operatorname{rang}(J_{n_j}(0)^i \Big) \\ &= \#\{ j \mid 1 \leq j \leq s,\, n_j \geq i \}. \end{aligned}$$

(Gleichung $(*)$ ist klar; vgl. auch nachstehendes Lemma.) Denn mit jeder höheren Potenz fällt der Rang eines Jordankästchens um 1, sofern nicht vorher schon die Nullmatrix erreicht war. Diese Gleichheit sagt aber nichts anderes, als dass $\mathbf{m} = (m_1, \ldots, m_q)$ die zu $\mathbf{n} = (n_1, \ldots, n_s)$ duale Partition ist, d. h. $\mathbf{m} = \mathbf{n}^*$, was aber auch $\mathbf{n} = \mathbf{m}^*$ bedeutet.

Es ist also noch (23.2) zu zeigen: Es werden die Unterräume $W_q, W_{q-1}, \ldots, W_1$ induktiv definiert. Zunächst sei W_q irgendein Unterraum von V mit $V = \operatorname{Kern}(f^{q-1}) \oplus W_q$. Gelte $2 \leq i \leq q$ und sei W_i mit $\operatorname{Kern}(f^i) = \operatorname{Kern}(f^{i-1}) \oplus W_i$ schon geeignet definiert. Offenbar gilt

(a) $f(\operatorname{Kern}(f^i)) \subseteq \operatorname{Kern}(f^{i-1})$.
(b) $f\,|_{W_i} \colon W_i \to \operatorname{Kern}(f^{i-1})$ ist injektiv. (Denn $\operatorname{Kern}(f) \subseteq \operatorname{Kern}(f^{i-1})$, und $\operatorname{Kern}(f^{i-1}) \cap W_i = \{0\}$.)
(c) $f(W_i) \cap \operatorname{Kern}(f^{i-2}) = \{0\}$. (Folgt aus (b).)

Also gibt es eine Zerlegung $\operatorname{Kern}(f^{i-1}) = \operatorname{Kern}(f^{i-2}) \oplus W_{i-1}$ mit $f(W_i) \subseteq W_{i-1}$, und hiermit gelten die gewünschten Eigenschaften wie in (23.2). □

Lemma 23.6. *Sei $(V, f) = (V_1, f_1) \oplus \ldots \oplus (V_r, f_r)$ eine direkte Zerlegung des Moduls (V, f) in Untermoduln. Dann gilt*

(1) $\operatorname{Kern}(f) = \operatorname{Kern}(f_1) \oplus \ldots \oplus \operatorname{Kern}(f_r)$.
(2) $\operatorname{Bild}(f) = \operatorname{Bild}(f_1) \oplus \ldots \oplus \operatorname{Bild}(f_r)$.

Beweis. Jedes $x \in V$ lässt sich eindeutig schreiben als $x = x_1 + \ldots + x_r$ mit $x_i \in V_i$ $(i = 1, \ldots, r)$. Ferner gilt

$$f(x) = f(x_1) + \ldots + f(x_r) = f_1(x_1) + \ldots + f_r(x_r),$$

mit $f_i(x_i) \in V_i$. Daraus folgt einerseits die Aussage für das Bild und andererseits

$$f(x) = 0 \quad \Leftrightarrow \quad f_i(x_i) = 0 \text{ für alle } i = 1, \ldots, r,$$

und damit auch die Aussage für den Kern. □

Das Hauptergebnis dieses Kapitels ist nun der folgende Satz:

Satz 23.7 (Jordansche Normalform (allgemeiner Fall)). *Sei f ein Endomorphimus eines n-dimensionalen K-Vektorraums V. Es gelte, dass f trigonalisierbar ist, etwa*

$$\chi_f = (T - \lambda_1)^{e_1} \cdot \ldots \cdot (T - \lambda_r)^{e_r}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r \in K$, und $e_1, \ldots, e_r \geq 1$. Dann gibt es eine Basis $\mathcal{B}$ von V, so dass

$$M_{\mathcal{B}}(f) = \left(\begin{array}{c|c|ccc|c} J_1 & & & & & \\ \hline & J_2 & & & & \\ \hline & & \ddots & & & \\ & & & & \ddots & \\ \hline & & & & & J_r \end{array}\right) \tag{23.3}$$

gilt, wobei für jedes $i = 1, \ldots, r$

(1) $J_i = \operatorname{diag}\big(J_{n_1^{(i)}}(\lambda_i), \ldots, J_{n_{s_i}^{(i)}}(\lambda_i)\big) \in \mathrm{M}(e_i; K)$ *für eine Partition* $\mathbf{n}^{(i)} = (n_1^{(i)}, \ldots, n_{s_i}^{(i)})$ *von e_i. gilt.*

Hieraus ergeben sich automatisch die folgenden Eigenschaften:

(2) *Die geometrische Vielfachheit von λ_i ist* $\dim(E(f, \lambda_i)) = s_i$.

(3) $\mathbf{n}^{(i)} = (n_1^{(i)}, \ldots, n_{s_i}^{(i)})$ *ist die zu* $\mathbf{m}^{(i)} = (m_1^{(i)}, \ldots, m_{q_i}^{(i)})$ *duale Partition von e_i, wobei*

$$m_j^{(i)} = \dim \operatorname{Kern}((f - \lambda_i \cdot 1_V)^j) - \dim \operatorname{Kern}((f - \lambda_i \cdot 1_V)^{j-1})$$

und $q_i = n_1^{(i)}$ gilt.

(4) *Für das Minimalpolynom gilt*

$$\mu_f = (T - \lambda_1)^{n_1^{(1)}} \cdot \ldots \cdot (T - \lambda_r)^{n_1^{(r)}}.$$

Die (bis auf Reihenfolge der Bestandteile $J_1, \ldots J_r$ zu den Eigenwerte $\lambda_1, \ldots, \lambda_r$ eindeutig bestimmte) Matrix in (23.3) heißt die *Jordansche Normalform* von f. Generell nennen wir eine Matrix der Form (23.3), die der Bedingung (1) genügt, eine *Jordan-Matrix.*

Beweis. Man zerlegt (V, f) in unzerlegbare Bestandteile,

$$(V, f) = (V_1, f_1) \oplus \ldots \oplus (V_t, f_t).$$

Da f trigonalisierbar ist, gilt dies auch für jedes $f_k \colon V_k \to V_k$ (denn $\chi_{f_k} \mid \chi_f$), jedes f_k besitzt also einen Eigenwert λ_i und V_k hat nach dem Satz über die Jordansche Normalform im unzerlegbaren Fall eine Basis, so dass f_k bzgl. dieser Basis durch ein Jordankästchen zum Eigenwert λ_i dargestellt wird. Es wird nun die direkte Summe aller V_k zum *selben* Eigenwert λ_i zusammengefasst, und dies ergibt einen f-invarianten Unterraum H_i von V. Setzt man die Teilbasen der jeweiligen V_k zu einer Basis von H_i zusammen, so liefern Satz 22.9 und Satz 23.3 die Darstellungsmatrix $J_i = \operatorname{diag}\big(J_{n_1^{(i)}}(\lambda_i), \ldots, J_{n_{s_i}^{(i)}}(\lambda_i)\big) \in \mathrm{M}(e_i; K)$, wobei man (nach evtl. Umnummerierung der Basiselemente) $n_1^{(i)} \geq \ldots \geq n_{s_i}^{(i)}$ annehmen kann. Die Aussage (1) ist somit klar. Wegen

$$V = H_1 \oplus \ldots \oplus H_r$$

bekommt man wiederum mit Satz 22.9 eine Basis $\mathcal{B}$ von V mit $M_{\mathcal{B}}(f) = \operatorname{diag}\big(J_1, \ldots, J_r\big)$.

Zu den Aussagen (2) und (3): Sei $i \in \{1, \ldots, r\}$. Setze $g_i = f - \lambda_i \cdot 1_V$. Die Unterräume $H_1, \ldots, H_r$ sind allesamt g_i-invariant. Offenbar ist $g_i \mid_{H_i} \colon H_i \to H_i$ nilpotent. Nach dem Normalformensatz für nilpotente Endomorphismen ergibt sich die Kästchengrößen innerhalb der Matrix J_i durch Betrachtung der Differenzen

$$\dim \operatorname{Kern}((g_i \mid_{H_i})^j) - \dim \operatorname{Kern}((g_i \mid_{H_i})^{j-1}).$$

Diese Zahlen stimmen aber nach dem vorigen Lemma mit $\dim \operatorname{Kern}((g_i)^j) - \dim \operatorname{Kern}((g_i)^{j-1})$ überein, da $g_i \mid_{H_j} \colon H_j \to H_j$ für $j \neq i$ (also $\lambda_j \neq \lambda_i$) ein Isomorphismus ist ($\lambda_j - \lambda_i \neq 0$

stehen auf der Hauptdiagonalen) und die entsprechenden Kerne trivial sind. Es folgt also (3). Aussage (2) folgt ganz analog, denn

$$\dim E(f,\lambda_i) = \dim \mathrm{Kern}(g_i) \overset{(*)}{=} \dim \mathrm{Kern}(g_i\,|_{H_i}) \overset{(**)}{=} s_i,$$

wobei bei $(*)$ wie zuvor das vorige Lemma benutzt wurde und $(**)$ aus dem Normalformensatz für nilpotente Endomorphismen folgt.

Zur Aussage (4): Definiere $p = (T-\lambda_1)^{n_1^{(1)}} \cdot \ldots \cdot (T-\lambda_r)^{n_1^{(r)}}$. Für jedes $i \in \{1,\ldots,r\}$ ist

$$(J_i - \lambda_i \cdot 1_V)^{n_1^{(i)}} = 0,$$

und daher gilt $p(J_i) = 0$. Es folgt

$$p\big(\mathrm{diag}(J_1,\ldots,J_r)\big) = \mathrm{diag}\big(p(J_1),\ldots,p(J_r)\big) = 0.$$

Damit folgt $\mu_f \mid p$. Für jedes $i \in \{1,\ldots,r\}$ ist aber andererseits $(J_i - \lambda_i \cdot 1_V)^{n_1^{(i)}-1} \neq 0$, und daraus folgt, dass in μ_f keine kleineren Exponenten als in p auftreten können, also $\mu_f = p$. □

Um explizit Basen zu bestimmen, ist die folgende Feststellung nützlich.

Bemerkung 23.8 (Haupträume). Die im vorigen Beweis zu den verschiedenen Eigenwerten $\lambda_1,\ldots,\lambda_r$ konstruierten Räume $H_1,\ldots,H_r$ nennt man die *Haupträume* von f, schreibt auch $H_i = H(f,\lambda_i)$, und nennt dementsprechend $V = H_1 \oplus \ldots \oplus H_r$ bzw.

$$(V,f) = (H_1, f\,|_{H_1}) \oplus \ldots \oplus (H_r, f\,|_{H_r})$$

die *Hauptraumzerlegung*. Für jedes $i \in \{1,\ldots,r\}$ gilt (mit den Bezeichnungen des vorigen Satzes)

$$H_i = \mathrm{Kern}\big((f-\lambda_i \cdot 1_V)^{e_i}\big) = \mathrm{Kern}\big((f-\lambda_i \cdot 1_V)^{n_1^{(i)}}\big).$$

Beweis. Setze $g_i = f - \lambda_i \cdot 1_V$. Wegen $n_1^{(i)} \le e_i$ und $(g_i\,|_{H_i})^{n_1^{(i)}} = 0$ gilt

$$H_i = \mathrm{Kern}\big((g_i\,|_{H_i})^{n_1^{(i)}}\big) = \mathrm{Kern}\big((g_i\,|_{H_i})^{e_i}\big).$$

Da $g_i\,|_{H_j} \colon H_j \to H_j$ für $j \neq i$ wegen $\lambda_j \neq \lambda_i$ ein Isomorphismus ist, folgt die Aussage aus dem vorigen Lemma. □

Folgerung 23.9. *Sei f ein trigonalisierbarer Endomorphismus eines n-dimensionalen K-Vektorraums. Folgende Aussagen sind äquivalent:*

(1) *(V,f) ist unzerlegbar.*
(2) *f hat nur einen Eigenwert $\lambda \in K$, und es gilt $\chi_f = \mu_f$.*

Beweis. (V,f) ist unzerlegbar genau dann, wenn in der Jordanschen Normalform nur ein Jordankästchen $J_k(\lambda)$ auftaucht. □

Folgerung 23.10. *Mit den Bezeichnungen aus dem Satz (insbesondere sei f trigonalisierbar) gilt die Äquivalenz der folgenden Aussagen:*

(1) *f ist diagonalisierbar.*
(2) *Die Jordansche Normalform von f ist eine Diagonalmatrix.*
(3) *$s_i = e_i$ für alle $i = 1,\ldots,r$.*
(4) *$n_1^{(i)} = 1$ für alle $i = 1,\ldots,r$.*

Bemerkung 23.11. (1) Die Jordansche Normalform eines trigonalisierbaren Endomorphismus f von V bzw. einer quadratischen Matrix $A \in \mathrm{M}(n;K)$ bestimmt man wie folgt:

(a) Zunächst wird das charakteristische Polynom bestimmt, und daraus die Eigenwerte $\lambda_1, \ldots, \lambda_r \in K$. (Letzteres kann eine nicht-triviale Aufgabe sein und u. U. numerische Methoden erforden.) Zerfällt das charakteristische Polynom, so ist f trigonalisierbar, sonst nicht. Falls f trigonalisierbar ist, fährt man fort:

(b) Zu jedem Eigenwert λ_i berechnet man die Zahlen

$$\begin{aligned} m_j^{(i)} &= \dim \operatorname{Kern}((f - \lambda_i \cdot 1_V)^j) - \dim \operatorname{Kern}((f - \lambda_i \cdot 1_V)^{j-1}) \\ &= \operatorname{rang}((f - \lambda_i \cdot 1_V)^{j-1}) - \operatorname{rang}((f - \lambda_i \cdot 1_V)^j) \end{aligned}$$

für $j = 1, 2, \ldots,$ bis man das erste mal Null erhält. Zu der daraus definierten Partition $\mathbf{m}^{(i)}$ bildet man die duale Partition und die daraus resultierende Jordan-Matrix J_i zum Eigenwert λ_i.

(c) Schließlich bildet man $\operatorname{diag}(J_1, \ldots, J_r)$.

Vgl. Besipiel 23.12.

(2) Das Minimalpolynom μ_f eines trigonalisierbaren Endomorphismus wird man üblicherweise anhand der Jordanschen Normalform wie in Satz 23.7 bestimmen.

Beispiel 23.12. Sei $f\colon \mathbb{R}^8 \to \mathbb{R}^8$ die lineare Abbildung, deren Darstellungsmatrix bzgl. der Standardbasis von $\mathbb{R}^8$ gegeben ist durch

$$A = \begin{pmatrix} 1 & 0 & 2 & -1 & 1 & -1 & -1 & -3 \\ 2 & -1 & 3 & -1 & 1 & -2 & -2 & -5 \\ 3 & -1 & 5 & -2 & 1 & -3 & -3 & -8 \\ 1 & 0 & 2 & -1 & 0 & -1 & -1 & -3 \\ -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ 3 & -1 & 6 & -2 & 2 & -3 & -4 & -10 \\ 1 & 0 & 1 & -1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{M}(8; \mathbb{R}).$$

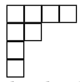
Man findet: $\chi_f = T^8$, d. h. f ist nilpotent. Wir berechnen die Potenzen von A:

$$A^2 = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 1 & 2 \\ -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 2 \\ 2 & -1 & 3 & -1 & 1 & -2 & -2 & -5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A^3 = \begin{pmatrix} -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ -2 & 1 & -4 & 1 & -1 & 2 & 3 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

und schließlich $A^4 = 0$. Man rechnet: $\operatorname{rang}(A) = 4$, $\operatorname{rang}(A^2) = 2$, $\operatorname{rang}(A^3) = 1$. Es ergibt sich

$$\begin{aligned} m_1 &= \operatorname{rang}(A^0) - \operatorname{rang}(A) = 8 - 4 = 4, \\ m_2 &= \operatorname{rang}(A^1) - \operatorname{rang}(A^2) = 4 - 2 = 2, \\ m_3 &= \operatorname{rang}(A^2) - \operatorname{rang}(A^3) = 2 - 1 = 1, \\ m_4 &= \operatorname{rang}(A^3) - \operatorname{rang}(A^4) = 1 - 0 = 1, \end{aligned}$$

also $\mathbf{m} = (4,2,1,1)$, und das Young-Diagramm dazu ist 

Offenbar ist in diesem Beispiel (!) $\mathbf{m}$ zu sich selbst dual, d. h. die Kästchengrößen sind gegeben durch $\mathbf{n} = \mathbf{m}^* = \mathbf{m} = (4,2,1,1)$. Es folgt, dass die Jordansche Normalform von f gegeben ist durch die Matrix

$$N(4,2,1,1) = \left(\begin{array}{cccc|cc|c|c} 0 & & & & & & & \\ 1 & 0 & & & & & & \\ & 1 & 0 & & & & & \\ & & 1 & 0 & & & & \\ \hline & & & & 0 & & & \\ & & & & 1 & 0 & & \\ \hline & & & & & & 0 & \\ \hline & & & & & & & 0 \end{array}\right).$$

Man liest hieraus auch noch ab (hat dies aber vorher schon gesehen), dass $\mu_A = T^4$ gilt.

Bemerkung 23.13. Oftmals ist man nicht nur an der Jordanschen Normalform an sich interessiert, sondern auch an einer Basis $\mathcal{B}$, bzgl. welcher $\mathrm{M}_{\mathcal{B}}(f)$ eine Jordan-Matrix ist, bzw. an einem $P \in \mathrm{GL}(n;K)$, so dass $P^{-1}AP$ eine Jordanmatrix ist. Der Beweis von Satz 23.2 gibt eine Idee, wie man eine solche Basis bestimmen kann. Für *jeden* Eigenwert $\lambda = \lambda_i$ betrachtet man $g \overset{def}{=} f - \lambda \cdot 1_V$, und mit den Unterräumen $U_j \overset{def}{=} \mathrm{Kern}(g^j)$ die aufsteigende Kette

$$\{0\} = U_0 \subsetneq U_1 \subsetneq \ldots \subsetneq U_{q-1} \subsetneq U_q = H(f,\lambda), \tag{23.4}$$

zum Hauptraum von f zum Eigenwert λ. (Wir unterdrücken dabei durchweg den Index i, um nicht zu viele Indizes zu haben.) Man bestimmt Basen $\mathcal{B}_j$ all dieser Unterräume U_j $(j < q)$. Man ergänzt (mit dem Verfahren 10.4) dann $\mathcal{B}_{q-1}$ mit $v_1^{(1)}, \ldots, v_{s_1}^{(1)}$ zu einer Basis von $U_q = H(f,\lambda)$. U_{q-1}. In 23.2 genügte hier ein einziger Vektor, da es sich um den unzerlegbaren Fall handelte. Allgemein können hier aber mehr Vektoren nötig sein. Auf die ergänzenden Vektoren wendet man nun f an: die Bildvektoren $g(v_1^{(1)}), \ldots, g(v_{s_1}^{(1)})$ bilden (wegen (23.2)) zusammen mit der Basis $\mathcal{B}_{q-2}$ eine freie Teilmenge von U_{q-1}, die man ggfs. durch $v_1^{(2)}, \ldots, v_{s_2}^{(2)}$ zu einer Basis von U_{q-1} ergänzt (s_2 kann $= 0$ sein). Auf die schon gebildeten Bildvektoren wendet man erneut g an und auch auf die neu ergänzten Vektoren, und diese Bildvektoren bilden dann (wegen (23.2)) mit $\mathcal{B}_{q-3}$ eine freie Teilmenge in U_{q-2}, u.s.w. Auf diese Weise arbeitet man sich durch sukzessive Anwendung von g von rechts nach links durch die Kette (23.4) bis man links angelangt ist. Schließlich sammelt man alle Vektoren in der folgenden Form Zeile für Zeile auf:

$$\begin{array}{lcl} v_1^{(1)},\ g(v_1^{(1)}), & \ldots, & g^{q-1}(v_1^{(1)}), \\ \vdots & & \vdots \\ v_{s_1}^{(1)},\ g(v_1^{(1)}), & \ldots, & g^{q-1}(v_{s_1}^{(1)}), \\ v_1^{(2)},\ g(v_1^{(2)}), & \ldots, & g^{q-2}(v_1^{(2)}), \\ \vdots & & \vdots \\ v_{s_2}^{(2)},\ g(v_{s_2}^{(2)}), & \ldots, & g^{q-2}(v_{s_2}^{(2)}), \\ \vdots & & \vdots \end{array}$$

wobei aber einige der Zahlen $s_2, s_3, \ldots$ auch $= 0$ sein können.

Dies wird für jeden Eigenwert λ von f durchgeführt, und anschließend werden die so konstruierten Basen der einzelnen Haupträume zu einer Basis von V zusammengesetzt. – Dies wird am folgenden kleinen Beispiel illustriert.

Beispiel 23.14. Sei

$$A = \begin{pmatrix} 3 & 5 & -1 & 0 & 0 & -7 \\ 0 & -2 & 0 & 0 & 0 & 10 \\ 0 & 0 & 3 & 0 & 0 & 1 \\ 0 & 7 & 32 & 3 & 15 & -77 \\ 0 & 1 & -10 & 0 & -2 & 16 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \in \mathrm{M}_6(\mathbb{R}).$$

Sei $f\colon \mathbb{R}^6 \to \mathbb{R}^6$, $x \mapsto Ax$ die zugehörige lineare Abbildung. Wir berechnen eine Basis $\mathcal{B}$ von $V = \mathbb{R}^6$ (bzw. ein $P \in \mathrm{GL}_6(\mathbb{R})$), so dass $M_{\mathcal{B}}(f)$ (bzw. $P^{-1}AP$) eine Jordansche Normalform ist.

Man berechnet das charakteristische Polynom:

$$\chi_f(\lambda) = \lambda^6 - 8\lambda^5 + 10\lambda^4 + 60\lambda^3 - 135\lambda^2 - 108\lambda + 324 = (\lambda+2)^2(\lambda-3)^4.$$

Damit zerfällt χ_f, also ist f trigonalisierbar. $\lambda_1 = -2$ und $\lambda_2 = 3$ sind die Eigenwerte von f.

Da hier 2 die algebraische Vielfachheit von λ_1 und 4 die algebraische Vielfachheit von λ_2 ist, berechnet man die Unterräume

$$H(f,-2) = \mathrm{Kern}((f - \lambda_1 \cdot 1)^2)$$

und

$$H(f,3) = \mathrm{Kern}((f - \lambda_2 \cdot 1)^4).$$

Es gilt $\mathbb{R}^6 = H(f,-2) \oplus H(f,3)$, und $H(f,-2)$ und $H(f,3)$ sind f-invariante Unterräume von $\mathbb{R}^6$ (Hauptraumzerlegung).

Seien $f_1\colon H(f,-2) \to H(f,-2)$, $x \mapsto f(x)$ bzw. $f_2\colon H(f,3) \to H(f,3)$, $x \mapsto f(x)$ die Einschränkungen von f auf die Haupträume.

Es ist

$$(A + 2E_6)^2 = \begin{pmatrix} 25 & 25 & -10 & 0 & 0 & -21 \\ 0 & 0 & 0 & 0 & 0 & 50 \\ 0 & 0 & 25 & 0 & 0 & 10 \\ 0 & 50 & 170 & 25 & 75 & -428 \\ 0 & 0 & -50 & 0 & 0 & 80 \\ 0 & 0 & 0 & 0 & 0 & 25 \end{pmatrix}.$$

Die dazu gehörige Treppenmatrix ist

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

Es ist somit

$$\begin{aligned}
H(f,-2) &= \mathrm{Kern}((f+2\cdot 1)^2) \\
&= \{x\in\mathbb{R}^6 \mid (f+2\cdot 1)^2(x)=0\} \\
&= \{x\in\mathbb{R}^6 \mid (A+2E_n)^2x=0\} \\
&= \{x\in\mathbb{R}^6 \mid x = r\begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} -\frac{3}{2} \\ \frac{3}{2} \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \text{ mit } r,s\in\mathbb{R}\}
\end{aligned}$$

Wir multiplizieren die beiden Vektoren jeweils mit -2 und erhalten

$$v_1 := \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} \text{ und } v_2 := \begin{pmatrix} 3 \\ -3 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}.$$

Somit ist $\mathcal{B}_{12} := (v_1, v_2)$ eine geordnete Basis von $H(f,-2)$. Wegen

$$\begin{aligned}
f_1(v_1) = f(v_1) &= Av_1 &= \begin{pmatrix} -2 \\ 2 \\ 0 \\ -1 \\ -1 \\ 0 \end{pmatrix} = -\frac{1}{2}v_1 - \frac{1}{2}v_2 \\
f_1(v_2) = f(v_2) &= Av_2 &= \begin{pmatrix} -6 \\ 6 \\ 0 \\ 9 \\ -7 \\ 0 \end{pmatrix} = \frac{9}{2}v_1 - \frac{7}{2}v_2
\end{aligned}$$

ist

$$A_1 = M_{\mathcal{B}_{12}}(f_1) = \begin{pmatrix} -\frac{1}{2} & \frac{9}{2} \\ -\frac{1}{2} & -\frac{7}{2} \end{pmatrix}.$$

Sei $g_1 = f_1 - \lambda_1 \cdot 1\colon H(f,-2) \to H(f,-2)$. Offenbar ist 0 der einzige Eigenwert von g_1, und daher ist g_1 nilpotent. Damit

$$C_1 := A_1 + 2E_2 = M_{\mathcal{B}_{12}}(g_1) = \begin{pmatrix} \frac{3}{2} & \frac{9}{2} \\ -\frac{1}{2} & -\frac{3}{2} \end{pmatrix}.$$

Es ist

$$C_1^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Man berechnet $U_{1i} = \mathrm{Kern}(g_1^i)$ und Basen davon:

$$\{0\} = U_{10} \subset U_{11} \subset U_{12} = H(f,-2).$$

Die Treppenmatrix zu C_1 ist

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

Wir erhalten hieraus den Koordinatenvektor $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$. Hieraus erhält man den Vektor $v := 3v_1 - 1v_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 6 \\ 2 \\ 0 \end{pmatrix}$ als Basis $\mathcal{B}_{11}$ von U_{11}.

Nun ergänzt man $\mathcal{B}_{11}$ z. B. mittels v_1 zu einer Basis von U_{12}. Desweiteren ist

$$g_1(v_1) = (f_1 + 2 \cdot 1)(v_1) = (f + 2 \cdot 1)(v_1) = (A + 2E_n)v_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \\ -1 \\ 0 \end{pmatrix}.$$

$\{g_1(v_1)\}$ ist in U_{11} linear unabhängig und wegen $\dim U_{11} = 1$ eine Basis von U_{11}. Es ist also keine Ergänzung mehr notwendig. Insgesamt folgt also, dass $\mathcal{B}_1 = (v_1, g_1(v_1))$ wieder eine Basis von $H(f, -2)$ ist. Bzgl. dieser wird g_1 (bzw. f_1) dargestellt durch

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ (bzw. } \begin{pmatrix} -2 & 0 \\ 1 & -2 \end{pmatrix} \text{).}$$

Wir haben noch $H(f, 3)$ zu behandeln: Es ist

$$(A - 3E_6)^4 = \begin{pmatrix} 0 & -625 & 0 & 0 & 0 & 1250 \\ 0 & 625 & 0 & 0 & 0 & -1250 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 250 & -3750 & 0 & -1875 & 7000 \\ 0 & -500 & 1250 & 0 & 625 & -1500 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die zugehörige Treppenmatrix ist

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

Seien $w_1 := \begin{pmatrix} 1\\0\\0\\0\\0\\0 \end{pmatrix}, w_2 := \begin{pmatrix} 0\\0\\0\\1\\0\\0 \end{pmatrix}, w_3 := \begin{pmatrix} 0\\0\\-\frac{1}{2}\\0\\1\\0 \end{pmatrix}$ und $w_4 := \begin{pmatrix} 0\\2\\2\\0\\0\\1 \end{pmatrix}$. Es ist $\mathcal{B}_{23} :=$ (w_1, w_2, w_3, w_4) eine geordnete Basis von $H(f,3)$. Wegen

$$\begin{aligned}
f_2(w_1) = f(w_1) &= Aw_1 = \begin{pmatrix} 3\\0\\0\\0\\0\\0 \end{pmatrix} &= 3w_1 + 0w_2 + 0w_3 + 0w_4 \\
f_2(w_2) = f(w_2) &= Aw_2 = \begin{pmatrix} 0\\0\\0\\3\\0\\0 \end{pmatrix} &= 0w_1 + 3w_2 + 0w_3 + 0w_4 \\
f_2(w_3) = f(w_3) &= Aw_3 = \begin{pmatrix} \frac{1}{2}\\0\\-\frac{3}{2}\\-1\\3\\0 \end{pmatrix} &= \frac{1}{2}w_1 - 1w_2 + 3w_3 + 0w_4 \\
f_2(w_4) = f(w_4) &= Aw_4 = \begin{pmatrix} 1\\6\\7\\1\\-2\\3 \end{pmatrix} &= 1w_1 + 1w_2 - 2w_3 + 3w_4
\end{aligned}$$

ist

$$A_2 = M_{\mathcal{B}_{23}}(f_2) = \begin{pmatrix} 3 & 0 & \frac{1}{2} & 1 \\ 0 & 3 & -1 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Sei $g_2 = f_2 - \lambda_2 \cdot 1 \colon H(f,3) \to H(f,3)$. Offenbar ist 0 der einzige Eigenwert von g_2, und daher ist g_2 nilpotent. Damit

$$C_2 := A_2 - 3E_4 = M_{\mathcal{B}_{23}}(g_2) = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es ist

$$C_2^2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

und

$$C_2^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Man berechnet $U_{2i} = \text{Kern}(g_2^i)$ und Basen davon:

$$\{0\} = U_{20} \subset U_{21} \subset U_{22} \subset U_{23} = H(f, 3).$$

Die Treppenmatrizen zu C_2 und C_2^2 sind

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ bzw. } \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Daher erhält man als Basen $\mathcal{B}_{21} := \{w_1, w_2\}$ von U_{21} sowie $\mathcal{B}_{22} := \{w_1, w_2, w_3\}$ von U_{22}.

Man ergänzt nun $\mathcal{B}_{22}$ z. B. mittels w_4 zu einer Basis von U_{23}.

Dann ist

$$g_2(w_4) = (f_2 - 3 \cdot 1_V)(w_4) = (f - 3 \cdot 1_V)(w_4) = (A - 3E_4)w_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ -2 \\ 0 \end{pmatrix}.$$

$\mathcal{B}_{21} \cup \{g_2(w_4)\}$ ist dann ein linear unabhängiges System in U_{22} und wegen $\dim U_{22} = 3$ eine Basis von U_{22}.

Desweiteren ist

$$g_2^2(w_4) = (f_2 - 3 \cdot 1)(w_4) = (f - 3 \cdot 1)(w_4) = (A - 3E_4)^2 w_4 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

und $\mathcal{B}_{20} \cup \{g_2^2(w_4)\} = \{g_2^2(w_4)\}$ ist linear unabhängig in U_{21}. Da $\dim U_{21} = 2$ ist, müssen wir $\{g_2^2(w_4)\}$ mittels eines hierzu linear unabhängigen Vektors aus U_{21} ergänzen. Z. B. ist dann $\{g_2^2(w_4), w_1\}$ eine Basis von U_{21}.
Insgesamt erhalten wir hieraus die geordnete Basis

$$\mathcal{B}_2 := (w_4, g_2(w_4), g_2^2(w_4), w_1),$$

bezüglich welcher g_2 (bzw. f_2) dargestellt wird durch die Matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \left(\text{bzw. } \begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}\right).$$

Setzt man nun die gefundenen Basen $\mathcal{B}_1$ von $H(f, -2)$ und $\mathcal{B}_2$ von $H(f, 3)$ zu einer Basis $\mathcal{B}$ zusammen, so erhält man: Bzgl. der Basis

$$\mathcal{B} := (v_1, g_1(v_1), w_4, g_2(w_4), g_2^2(w_4), w_1)$$

wird f dargestellt durch die Matrix

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

Setzt man die Basisvektoren als Spalten zu einer (invertierbaren) Matrix P zusammen, so erhält man

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

damit

$$P^{-1} = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -2 \\ 0 & 1 & \frac{5}{2} & \frac{1}{2} & \frac{3}{2} & -7 \\ 1 & 2 & \frac{3}{2} & \frac{1}{2} & \frac{3}{2} & -7 \end{pmatrix}$$

und

$$P^{-1}AP = \begin{pmatrix} -2 & & & & & \\ 1 & -2 & & & & \\ & & 3 & & & \\ & & 1 & 3 & & \\ & & & 1 & 3 & \\ & & & & & 3 \end{pmatrix}.$$

Bemerkung 23.15. (1) Die Voraussetzung in Satz 23.7, dass f trigonalisierbar ist, ist überflüssig bzw. gilt automatisch, wenn der Körper K algebraisch abgeschlossen ist, also z. B. für $K = \mathbb{C}$.

(2) Satz 23.7 (und die anderen) haben wir nur in der Version für Endomorphismen formuliert. Es gilt natürlich auch die Matrixversion.

23.16. (1) Ausführliche Wiederholung und Zusammenfassung der Ergebnisse. [...]

- Von Interesse: f-invariante Unterräume. Führt zum Modulbegriff.
- Idee: Basis mit Bestandteilen der Form x, $f(x)$, $f^2(x), \ldots$ finden.
- Reduktion auf unzerlegbaren Fall.
- Im unzerlegbaren Fall: Fittings-Lemma.
- Dies führt zur Untersuchung des nilpotenten Falls.
- Der allgemeine Fall wird auf den nilpotenten zurückgeführt.
- Statt die (theoretische Zerlegung) in Unzerlegbare konkret zu berechnen, verwendet man die Haupträume.

(2) Ein (abstraktes) Fallbeispiel. Sei $f\colon V \to V$ ein Endomorphismus. Sei $\dim(V) = 10$. Es sei f trigonalisierbar, und wir nehmen an, dass f nur einen Eigenwert $\lambda \in K$ hat. Wir nehmen weiter an, dass für $U_j \overset{def}{=} \mathrm{Kern}((f - \lambda \cdot 1_V)^j)$ folgendes gilt:

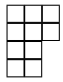
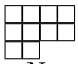
$$\{0\} = U_0 \subseteq U_1 \subsetneq U_2 \subsetneq U_3 \subsetneq U_4 = V$$

mit

$$\dim(U_1) = 3,\ \dim(U_2) = 6,\ \dim(U_3) = 8$$

gilt. Daraus folgt

$$\begin{aligned} m_1 &= 3-0=3, \\ m_2 &= 6-3=3, \\ m_3 &= 8-6=2, \\ m_4 &= 10-8=2. \end{aligned}$$

$\mathbf{m} = (3,3,2,2)$ ist eine Partition von 10, dessen Young-Diagramm  wird durch Transponieren zu . Die duale Partition ist also $\mathbf{n} = \mathbf{m}^* = (4,4,2)$. Es folgt damit, dass die Jordansche Normalform von f die folgende Form hat:

$$J = \left(\begin{array}{cccc|cccc|cc} \lambda &&&&&&&&& \\ 1 & \lambda &&&&&&&& \\ & 1 & \lambda &&&&&&& \\ && 1 & \lambda &&&&&& \\ \hline &&&& \lambda &&&&& \\ &&&& 1 & \lambda &&&& \\ &&&&& 1 & \lambda &&& \\ &&&&&& 1 & \lambda && \\ \hline &&&&&&&& \lambda & \\ &&&&&&&& 1 & \lambda \end{array}\right)$$

Wir demonstrieren nochmal (abstrakt) das Verfahren, wie man eine Basis $\mathcal{B}$ von V bestimmt mit $M_{\mathcal{B}}(f) = J$:

Setze $G \overset{def}{=} f - \lambda \cdot 1_V$.

Man berechnet jeweils Basen von U_1, U_2, U_3 und U_4 (Berechnung der Basis eines Kerns). U_3 ist 8-dimensional im 10-dimensionalen Raum U_4. Man ergänzt die berechnete Basis von U_3 (mit dem Verfahren aus 10.4) durch zwei Elemente x, y zu einer Basis von V.

Nun wendet man f an: Es liegen dann $g(x)$, $g(y)$ in U_3, und bilden zusammen mit den 6 Elementen der berechneten Basis von U_2 eine freie Teilmenge von U_3. Wegen $6+2=8$ ist dies dann schon eine Basis von U_3, es ist also keine Basisergänzung nötig.

Man wendet wieder g an. Wir bilden also $g^2(x)$, $g^2(y)$. Diese liegen in U_2 und bilden zusammen mit den 3 Elementen der berechneten Basis von U_1 eine freie, 5-elementige Teilmenge von U_2. Dies folgt aus dem Beweis der Aussage (23.2) im Beweis von Satz 23.5. Wir ergänzen diese mit einem Element z zu einer Basis von U_2.

Nun wenden wir erneut g an. Nicht nur auf die Elemente $g^2(x)$ und $g^2(y)$, sondern auch auf den weiteren ergänzenden Vektor z: Wir bilden also $g^3(x)$, $g^3(y)$, $g(z)$. Diese Elemente liegen in U_1 und bilden dort (zusammen mit der Basis $\emptyset$ von $U_0 = \{0\}$) ein freie Teilmenge, und wegen $\dim(U_1) = 3$ schon eine Basis von U_1. Es ist keine Ergänzung nötig.

Da wir bereits "links" angelangt sind, endet das Verfahren. Wir sammeln nun die ergänzenden Vektoren und deren Bilder auf, und ordnen sie wie folgt an:

$$x,\ g(x),\ g^2(x),\ g^3(x), \quad y,\ g(y),\ g^2(y),\ g^3(y), \quad z,\ g(z),$$

und bzgl. dieser Basis ist (offenbar) J die Darstellungsmatrix.

Man sieht hierbei auch, wie die duale Partition zustande kommt.

Ordnet man die Teile x, $g(x)$, $g^2(x), \ldots, g^q(x)$ einer solchen Basis umgekehrt an (was durch eine Permutation zustande kommt), so sieht man, dass ein Jordankästchen $J_q(\lambda)$ zu ${}^t(J_q(\lambda))$ ähnlich ist, und daraus folgt, dass jede Jordanmatrix J zu ${}^tJ$ ähnlich ist. Weil Ähnlichkeit eine Äquivalenzrelation ist, folgt daraus sofort:

SATZ 23.17. *Ist $A \in \mathrm{M}(n;K)$ trigonalisierbar (etwa: $K = \mathbb{C}$), so ist A ähnlich zu ${}^tA$.*

Beweis. Nach dem Satz über die Jordansche Normalform gibt es ein $P \in \mathrm{GL}(n;K)$, so dass $P^{-1}AP = J$ eine Jordanmatrix ist. Es folgt $Q^{-1\,t}AQ = {}^tJ$, mit $Q = {}^t(P^{-1}) \in \mathrm{GL}(n;K)$. Also ist A ähnlich zu J, ${}^tA$ ähnlich zu ${}^tJ$, und nach der Vorbemerkung folgt die Behauptung. □

Wir zeigen noch eine Anwendung der Hauptraumzerlegung:

23.18. Sei f ein trigonalisierbarer Endomorphismus von V. Es gelte

$$\chi_f = (T-\lambda_1)^{e_1} \cdot \ldots \cdot (T-\lambda_r)^{e_r}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r \in K$, und $e_1, \ldots, e_r \geq 1$. Sei

$$(V,f) = (H_1, f\,|_{H_1}) \oplus \ldots \oplus (H_r, f\,|_{H_r})$$

die Hauptraumzerlegung. Für jedes $i \in \{1, \ldots, r\}$ sei $P_i \colon V \to V$ der Endomorphismus mit $P(x) = x_i$ für jedes $x = x_1 + \ldots + x_r$, wobei $x_j \in H_j$ gilt. Es gilt dann offenbar

$$f = \sum_{i=1}^{r} f \circ P_i = \sum_{i=1}^{r} P_i \circ f.$$

Dies nennt man auch die *Spektralzerlegung* von f.

Es ist

$$g \overset{def}{=} \sum_{i=1}^{r} \lambda_i P_i \colon V \to V$$

diagonalisierbar, und es ist

$$h \overset{def}{=} f - g = \sum_{i=1}^{r} (f - \lambda_i \cdot 1_V) \circ P_i$$

nilpotent. Letzteres muss man nur auf den Haupträumen H_i nachprüfen, aber hierauf ist $\sum_{j=1}^{r}(f-\lambda_j \cdot 1_V)\circ P_j$ gleich $f - \lambda_i \cdot 1$. (Man muss dann das Maximum der Nilpotenzindizes über alle i nehmen.)

Wegen $f \circ P_i = P_i \circ f$ für alle $i = 1, \ldots, r$, gilt $f \circ g = g \circ f$, daher auch $h \circ g = (f-g) \circ g = g \circ (f-g) = g \circ h$.

Man erhält:

Satz 23.19 (Jordan-Zerlegung). *Sei f ein trigonalisierbarer Endomorphismus von V. Dann gibt es eindeutig bestimmte g, $h \in \mathrm{End}_K(V)$ mit:*

(1) $f = g + h$.
(2) *g ist diagonalisierbar und h nilpotent.*
(3) $g \circ h = h \circ g$.

Man nennt g den diagonalisierbaren und h den nilpotenten Bestandteil von f.

Beweis. Die Existenz von g und h wurde gerade zuvor gezeigt. Es ist noch die Eindeutigkeit zu zeigen. Seien auch g', $h' \in \mathrm{End}_K(V)$, die obige Bedingungen erfüllen. Aus (3) folgt, dass g' und h' mit f vertauschen und sich daher auf die Haupträume H_i einschränken lassen, und diese Einschränkungen erfüllen ebenfalls (2). Auf H_i gilt $g' - \lambda_i \cdot 1 = f - h' - \lambda_i \cdot 1 = h - h'$. Nun ist h' mit $g' - \lambda_i \cdot 1$, also auch mit h vertauschbar. Aus dem binomischen Lehrsatz (der sich auf h und h' (eingeschränkt auf H_i anwenden lässt), folgt leicht, dass auch $h - h'$ auf H_i nilpotent ist. Es ist also $g' - \lambda_i \cdot 1$ auf H_i nilpotent und diagonalisierbar, also Null auf H_i, also $g' = \lambda_i \cdot 1$ auf H_i, und es folgt $g' = \sum_{i=1}^{r} \lambda_i P_i = g$. Dann folgt aber auch $h' = (f - g') = (f - g) = h$. □

KAPITEL 7

Euklidische und unitäre Vektorräume

24. Skalarprodukte

Wir nehmen nun an, dass $K = \mathbb{R}$ oder $K = \mathbb{C}$ gilt.

Definition 24.1. Sei V ein K-Vektorraum. Eine Abbildung $V \times V \to K$, $(x, y) \mapsto \langle x \mid y \rangle$ heißt eine *hermitesche Form*, wenn für alle x, x', y und $\alpha \in K$ gilt:

(1) $\langle x + x' \mid y \rangle = \langle x \mid y \rangle + \langle x' \mid y \rangle$
(2) $\langle \alpha x \mid y \rangle = \alpha \langle x \mid y \rangle$
(3) $\langle y \mid x \rangle = \overline{\langle x \mid y \rangle}$.

Im Fall $K = \mathbb{R}$ vereinfacht sich die dritte Bedingung zu $\langle y \mid x \rangle = \langle x \mid y \rangle$, und man sagt dann, dass $\langle - \mid - \rangle$ *symmetrisch* ist.

Bemerkung 24.2. Sei $\langle - \mid - \rangle$ eine hermitsche Form auf dem K-Vektorraum V. Dann gilt für alle x, y, y' und $\beta \in K$:

(1') $\langle x \mid y + y' \rangle = \langle x \mid y \rangle + \langle x \mid y' \rangle$
(2') $\langle x \mid \beta y \rangle = \overline{\beta} \langle x \mid y \rangle$.
(4) $\langle x \mid x \rangle \in \mathbb{R}$
(5) $\langle 0 \mid y \rangle = 0 = \langle x \mid 0 \rangle$.

Beweis. (1') Es ist

$$\begin{aligned} \langle x \mid y + y' \rangle &= \overline{\langle y + y' \mid x \rangle} = \overline{\langle y \mid x \rangle + \langle y' \mid x \rangle} \\ &= \overline{\langle y \mid x \rangle} + \overline{\langle y' \mid x \rangle} = \langle x \mid y \rangle + \langle x \mid y' \rangle. \end{aligned}$$

(2') Es ist

$$\begin{aligned} \langle x \mid \beta y \rangle &= \overline{\langle \beta y \mid x \rangle} = \overline{\beta \langle y \mid x \rangle} \\ &= \overline{\beta}\,\overline{\langle y \mid x \rangle} = \overline{\beta} \langle x \mid y \rangle. \end{aligned}$$

(4) Wegen $\overline{\langle x \mid x \rangle} = \langle x \mid x \rangle$ ist $\langle x \mid x \rangle \in \mathbb{R}$.

(5) Es ist $\langle 0 \mid y \rangle = \langle 0_K \cdot 0_V \mid y \rangle = 0_K \langle 0_V \mid y \rangle = 0_K$, und ebenso $\langle x \mid 0 \rangle = \langle x \mid 0_K \cdot 0_V \rangle = \overline{0_K} \langle x \mid 0_V \rangle = 0_K$. □

Definition 24.3. (1) Sei V ein K-Vektorraum. Eine Abbildung

$$V \times V \to K, \ (x, y) \mapsto \langle x \mid y \rangle$$

heißt *Skalarprodukt* (oder *inneres Produkt*) auf V, falls gilt:

(SP 1) $\langle - \mid - \rangle$ ist eine hermitesche Form.
(SP 2) $\langle - \mid - \rangle$ ist *positiv definit*, d. h. für jedes $x \in V$ mit $x \neq 0$ gilt $\langle x \mid x \rangle > 0$.

(2) Ein $\mathbb{R}$-Vektorraum V mit Skalarprodukt heißt *euklidischer Vektorraum.*

(3) Ein $\mathbb{C}$-Vektorraum V mit Skalarprodukt heißt *unitärer Vektorraum.*

Satz 24.4 (Ungleichung von Cauchy und Schwarz). *Sei $(V, \langle - \mid - \rangle)$ ein K-Vektorraum mit Skalarprodukt. Für alle x, $y \in V$ gilt*

$$|\langle x \mid y \rangle|^2 \leq \langle x \mid x \rangle \cdot \langle y \mid y \rangle,$$

und es gilt Gleichheit genau dann, wenn x und y linear abhängig sind.

Beweis. Für $y = 0$ haben wir offenbar Gleichheit. Also nehmen wir nun $y \neq 0$ an. Für alle $\alpha, \beta \in K$ gilt

$$0 \leq \langle \alpha x + \beta y \mid \alpha x + \beta y \rangle = \alpha\overline{\alpha}\langle x \mid x \rangle + \beta\overline{\beta}\langle y \mid y \rangle + \alpha\overline{\beta}\langle x \mid y \rangle + \beta\overline{\alpha}\langle y \mid x \rangle.$$

Ist speziell $\alpha \overset{def}{=} \langle y \mid y \rangle$, so dividiert man durch $\alpha > 0$ und erhält

$$0 \leq \langle x \mid x \rangle\langle y \mid y \rangle + \beta\overline{\beta} + \overline{\beta}\langle x \mid y \rangle + \beta\langle y \mid x \rangle.$$

Setzt man weiter $\beta \overset{def}{=} -\langle x \mid y \rangle$, so folgt

$$0 \leq \langle x \mid x \rangle\langle y \mid y \rangle - \langle x \mid y \rangle\overline{\langle x \mid y \rangle} = \langle x \mid x \rangle\langle y \mid y \rangle - |\langle x \mid y \rangle|^2,$$

und dies zeigt die behauptete Ungleichung.

Für die Frage, wann Gleichheit gilt, können wir wieder $y \neq 0$ annehmen. Gilt die Gleichheit, so folgt mit den oben definierten $\alpha, \beta \in K$ die Gleichung

$$0 = \langle \alpha x + \beta y \mid \alpha x + \beta y \rangle$$

und damit $\alpha x + \beta y = 0$.

Sind umgekehrt x und y linear abhängig, so gibt es $\lambda \in K$ mit $x = \lambda y$, und dann

$$\begin{aligned} |\langle x \mid y \rangle|^2 &= |\lambda|^2 \cdot |\langle y \mid y \rangle|^2 = \lambda \cdot \overline{\lambda} \cdot \langle y \mid y \rangle \cdot \langle y \mid y \rangle \\ &= \langle \lambda y \mid \lambda y \rangle\langle y \mid y \rangle = \langle x \mid x \rangle\langle y \mid y \rangle. \end{aligned}$$

□

Satz 24.5. *Sei $(V, \langle - \mid - \rangle)$ ein K-Vektorraum mt Skalarprodukt. Für die Abbildung*

$$V \to \mathbb{R},\ x \mapsto \|x\| \overset{def}{=} {}_{+}\sqrt{\langle x \mid x \rangle}$$

gilt:

(1) *Für jedes $x \in V$ gilt $\|x\| \geq 0$, und es gilt $\|x\| = 0$ genau für $x = 0$.*
(2) *Für alle $x \in V$ und alle $\alpha \in K$ gilt $\|\alpha x\| = |\alpha| \cdot \|x\|$.*
(3) *Für alle $x, y \in V$ gilt $|\langle x \mid y \rangle| \leq \|x\| \cdot \|y\|$.*
(4) *Für alle $x, y \in V$ gilt $\|x + y\| \leq \|x\| + \|y\|$. (Dreiecksungleichung)*

Man nennt $\| - \|$ die zu $\langle - \mid - \rangle$ gehörige *Norm.*

Beweis. (1) und (2) sind klar, und (3) ist gerade die Cauchy-Schwarzsche Ungleichung.

Zu (4). Es ist

$$\begin{aligned} \|x + y\|^2 &= \langle x + y \mid x + y \rangle = \langle x \mid x \rangle + \langle x \mid y \rangle + \langle y \mid x \rangle + \langle y \mid y \rangle \\ &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\big(\langle x \mid y \rangle\big) \overset{klar}{\leq} \|x\|^2 + \|y\|^2 + 2|\langle x \mid y \rangle| \\ &\overset{(3)}{\leq} \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = \big(\|x\| + \|y\|\big)^2. \end{aligned}$$

□

Folgerung 24.6. *Sei $(V, \langle - \mid - \rangle)$ ein euklidischer Vektorraum. Wir können für $x, y \in V$ mit $x, y \neq 0$ durch*

$$\cos\alpha = \frac{\langle x \mid y \rangle}{\|x\| \cdot \|y\|}$$

im Intervallbereich $0 \leq \alpha \leq \pi$ liegenden Winkel $\alpha = \sphericalangle(x, y)$ zwischen x und y erklären. Ferner gilt:

$$\sphericalangle(x, y) = \pi/2 \quad \Leftrightarrow \quad \langle x \mid y \rangle = 0.$$

Auf dieses Konzept von *Orthogonalität* werden wir im nächsten Abschnitt eingehen.

Beweis. Nach der Cauchy-Schwarzschen Ungleichung 24.5 (3) gilt

$$-1 \leq \frac{\langle x \mid y \rangle}{\|x\| \cdot \|y\|} \leq 1.$$

Der Rest ist dann klar. □

Beispiel 24.7 (Der $\mathbb{R}^n$ mit dem Standardskalarprodukt). Es sei $K = \mathbb{R}$ und $V = \mathbb{R}^n$. Seien

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \; y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n.$$

Definiere

$$(x \mid y) = \sum_{i=1}^{n} x_i \cdot y_i = {}^t x \cdot y \in \mathbb{R}.$$

Man prüft leicht nach, dass dadurch ein Skalarprodukt auf $\mathbb{R}^n$ definiert wird, und dies heißt das *Standardskalarprodukt*. Für die zugehörige Norm gilt

$$\|x\| = +\sqrt{\sum_{i=1}^{n} x_i^2}.$$

Man nennt dies auch die *euklidische Norm* auf $\mathbb{R}^n$.

Die Cauchy-Schwarzsche Ungleichung wird zu

$$\left| \sum_{i=1}^{n} x_i y_i \right|^2 \leq \left(\sum_{i=1}^{n} x_i^2 \right) \cdot \left(\sum_{i=1}^{n} y_i^2 \right).$$

Beispiel 24.8 (Der $\mathbb{C}^n$ mit dem Standardskalarprodukt). Es sei $K = \mathbb{C}$ und $V = \mathbb{C}^n$. Seien

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \; y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n.$$

Definiere

$$(x \mid y) = \sum_{i=1}^{n} x_i \cdot \overline{y_i} = {}^t x \cdot \overline{y} \in \mathbb{C}.$$

Man prüft leicht nach, dass dadurch ein Skalarprodukt auf $\mathbb{C}^n$ definiert wird, und dies heißt das *Standardskalarprodukt*. Für die zugehörige Norm gilt

$$\|x\| = +\sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Man nennt dies auch die *euklidische Norm* auf $\mathbb{C}^n$.

Die Cauchy-Schwarzsche Ungleichung wird zu

$$\left| \sum_{i=1}^{n} x_i \overline{y_i} \right|^2 \leq \left(\sum_{i=1}^{n} |x_i|^2 \right) \cdot \left(\sum_{i=1}^{n} |y_i|^2 \right).$$

Beispiel 24.9. Sei $V = \mathbb{R}^3$. Sei

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathrm{M}(3; \mathbb{R}).$$

Für $x, y \in V$ definiere

$$\langle x \mid y \rangle = {}^t x A y.$$

Dies definiert ein Skalarprodukt auf $\mathbb{R}^3$. (Beweis als Übung.)

Beispiel 24.10. Sei $V = C([0,1])$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$. Für alle $f, g \in V$ setzt man

$$\langle f \mid g \rangle \overset{def}{=} \int_0^1 f(x)g(x)\,dx.$$

Dies definiert ein Skalarprodukt auf V: (SP 1) folgt leicht aus der Linearität der Integralbildung. Zu (SP 2): Es ist

$$\langle f \mid f \rangle = \int_0^1 f(x)^2\,dx.$$

Sei $f \neq 0$, d. h. nicht die Nullfunktion. Es gibt also (wegen der Stetigkeit) ein $x_0 \in]0,1[$, ein $a > 0$ und ein $b > 0$, so dass $[x_0 - a, x_0 + a]$ in $[0,1]$ liegt in und $f(x)^2 \geq b$ gilt für alle $x \in [x_0 - a, x_0 + a]$. Es folgt $\int_0^1 f(x)^2\,dx \geq 2ab > 0$.

Hier lautet die Cauchy-Schwarzsche Ungleichung

$$\left| \int_0^1 f(x)g(x)\,dx \right|^2 \leq \left(\int_0^1 f(x)^2\,dx \right) \cdot \left(\int_0^1 g(x)^2\,dx \right).$$

25. Orthogonalität

Sei $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Definition 25.1. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Es heißen $x, y \in V$ zueindander *orthogonal*, wenn $\langle x \mid y \rangle = 0$ gilt.

Definition 25.2. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Ein System von Vektoren $x_1, \ldots, x_s \in V$ heißt

(1) *Orthogonalsystem*, falls $\langle x_i \mid x_j \rangle = 0$ für alle $i \neq j$ und $x_i \neq 0$ für alle $i = 1, \ldots, s$ gilt;
(2) *Orthonormalsystem*, falls $\langle x_i \mid x_j \rangle = 0$ für alle $i \neq j$ gilt und $\|x_i\| = 1$ gilt für alle $i = 1, \ldots, s$. Kurz: $\langle x_i \mid x_j \rangle = \delta_{ij}$.

Satz 25.3. *Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt und $x_1, \ldots, x_s$ ein Orthogonalsystem. Dann sind $x_1, \ldots, x_s$ linear unabhängig.*

Beweis. Seien $\alpha_1, \ldots, \alpha_s \in K$ mit $\sum_{i=1}^s \alpha_i x_i = 0$. Für jedes $j = 1, \ldots, s$ gilt

$$0 = \langle 0 \mid x_j \rangle = \langle \sum_{i=1}^s \alpha_i x_i \mid x_j \rangle = \sum_{i=1}^s \alpha_i \langle x_i \mid x_j \rangle = \alpha_j \langle x_j \mid x_j \rangle,$$

und wegen $\langle x_j \mid x_j \rangle \neq 0$ folgt $\alpha_j = 0$. □

Definition 25.4. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt, und es gelte $n \overset{def}{=} \dim(V) < \infty$.

(1) Ein Orthogonalsystem $x_1, \ldots, x_s$ mit $s = n$ heißt eine *Orthogonalbasis* von V.
(2) Ein Orthonormalsystem $x_1, \ldots, x_s$ mit $s = n$ heißt eine *Orthonormalbasis* (ONB) von V.

Beispiel 25.5. Im $\mathbb{R}^n$ bildet die Standardbasis $e_1, \ldots, e_n$ bzgl. des Standardskalarprodukts $(- \mid -)$ eine ONB:

$$(e_i \mid e_j) = \sum_{k=1}^n (e_i)_k \cdot (e_j)_k = \sum_{k=1}^n \delta_{ik} \cdot \delta_{jk} = \delta_{ij}.$$

Entsprechendes gilt für $\mathbb{C}^n$.

Definition 25.6. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Sei $X \subseteq V$ eine Teilmenge. Definiere

$$X^\perp \overset{def}{=} \{v \in V \mid \langle v \mid x \rangle = 0 \text{ für alle } x \in X\}.$$

Ist $U \subseteq V$ ein Unterraum, so heißt $U^\perp$ das *orthogonale Komplement* von U in V.

Bemerkung 25.7. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Sei $X \subseteq V$ eine Teilmenge. Dann ist $X^\perp$ ein Unterraum von V mit $X \cap X^\perp \subseteq \{0\}$.

Beweis. Es gilt $\langle 0 \mid x \rangle = 0$ für alle $x \in X$, also ist $0 \in X^\perp$. Sind $v, v' \in X^\perp$, so gilt $\langle v + v' \mid x \rangle = \langle v \mid x \rangle + \langle v' \mid x \rangle = 0 + 0 = 0$ für alle $x \in X$, also $v + v' \in X^\perp$. Ist $v \in X^\perp$ und $\alpha \in K$, so gilt $\langle \alpha v \mid x \rangle = \alpha \langle v \mid x \rangle = \alpha \cdot 0 = 0$ für alle $x \in X$, also $\alpha v \in X^\perp$. Sei $x \in X \cap X^\perp$. Dann gilt $\langle x \mid x \rangle = 0$, und aus (SP 2) folgt $x = 0$. □

Alternativer Beweis der Unterraumeigenschaft: Es ist offenbar

$$X^\perp = \bigcap_{x \in X} \mathrm{Kern}(\varphi_x)$$

als Durchschnitt von Unterräumen selbst ein Unterraum, wobei für jedes $x \in X$ die lineare Abbildung φ_x gegeben ist durch $\langle - \mid x \rangle$, also durch $\varphi_x(v) = \langle v \mid x \rangle$ für jedes $v \in V$.

Satz 25.8 (Projektionslemma). *Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Sei $U \subseteq V$ ein Unterraum mit einer Orthonormalbasis $u_1, \dots, u_r$. Dann lässt sich jedes $v \in V$ eindeutig schreiben als*

$$v = u + w \quad \textit{mit} \quad u \in U,\ w \in U^\perp.$$

Es gilt dabei $u = \sum_{i=1}^r \langle v \mid u_i \rangle u_i$.

Beweis. *Existenz.* Setze $u \overset{def}{=} \sum_{i=1}^r \langle v \mid u_i \rangle u_i$ und $w \overset{def}{=} v - u$. Dann gilt $v = u + w$, $u \in U$, und es gilt

$$\langle w \mid u_i \rangle = \langle v - u \mid u_i \rangle = \langle v - \langle v \mid u_i \rangle u_i \mid u_i \rangle = \langle v \mid u_i \rangle - \langle v \mid u_i \rangle \langle u_i \mid u_i \rangle = 0$$

für alle $i = 1, \dots, r$, und es folgt $w \in U^\perp$.

Eindeutigkeit. Gelte

$$v = u + w = u' + w' \quad \text{mit} \quad u,\, u' \in U,\ w,\, w' \in U^\perp.$$

Es folgt

$$u - u' = w' - w \in U \cap U^\perp = \{0\},$$

also $u = u'$ und $w = w'$. □

Es wird nun die Frage beantwortet, inwieweit Orthonormalbasen existieren. Das Projektionslemma spielt dabei die Rolle des Induktionsschritts.

Satz 25.9. *Jeder endlichdimensionale Vektorraum mit Skalarprodukt besitzt eine Orthonormalbasis.*

Beweis. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Sei $x_1, \dots, x_n$ eine Basis von V. Induktion nach $n = \dim(V)$.

Sei $n = 1$. Dann bildet $e_1 \overset{def}{=} x_1 / \|x_1\|$ eine Orthonormalbasis von V.

$n-1 \to n$. Nach Induktionsvoraussetzung bildet besitzt der von $x_1, \dots, x_{n-1}$ erzeugte Unterraum U von V eine Orthonormalbasis $e_1, \dots, e_{n-1}$. Mit dem Projektionslemma kann man x_n schreiben als $x_n = u + w$ mit $u \in U$ und $w \in U^\perp$, wobei $u = \sum_{i=1}^{n-1} \langle x_n \mid e_i \rangle e_i$ gilt. Es ist $x_n \notin U$, also $w \neq 0$. Normieren von w liefert einen weiteren Einheitsvektor

$$e_n \overset{def}{=} \frac{x_n - \sum_{i=1}^{n-1} \langle x_n \mid e_i \rangle e_i}{\|x_n - \sum_{i=1}^{n-1} \langle x_n \mid e_i \rangle e_i\|}.$$

Für jedes $j = 1, \dots, n-1$ gilt $\langle e_n \mid e_j \rangle = 0$. Also bildet $e_1, \dots, e_{n-1}, e_n$ ein Orthonormalsystem von V aus $n = \dim(V)$ Vektoren. □

Bemerkung 25.10. Explizite Durchformung des Induktionsarguments liefert das sogenannte *Schmidtsche Orthonormalisierungsverfahren*, welches konstruktiv aus einer vorgegebenen Basis $x_1, \ldots, x_n$ von V eine Orthonormalbasis $e_1, \ldots, e_n$ von V macht.

Satz 25.11. *Sei $(V, \langle - \mid - \rangle)$ ein endlichdimensionaler Vektorraum mit Skalarprodukt. Sei $U \subseteq V$ ein Unterraum. Dann gilt:*

(1) $V = U + U^\perp$, $U \cap U^\perp = \{0\}$, $\langle U \mid U^\perp \rangle = 0$.
(2) $(U^\perp)^\perp = U$.

Man sagt in der Situation von (1), dass U die *orthogonale Summe* von U und $U^\perp$ ist und schreibt $V = U \perp U^\perp$.

Beweis. (1) Mit V ist auch U endlichdimensional, besitzt also nach Satz 25.9 eine Orthonormalbasis. Also kann man auf U das Projektionslemma anwenden, und es folgt damit $V = U + U^\perp$. Der Rest ist klar.

(2) Nach Definition gilt $U \subseteq (U^\perp)^\perp$. Sei $v \in (U^\perp)^\perp$. Wegen $V = U + U^\perp$ kann man schreiben $v = u + w$ mit $u \in U$ und $w \in U^\perp$. Es folgt

$$0 = \langle v \mid w \rangle = \langle u + w \mid w \rangle = \langle u \mid w \rangle + \langle w \mid w \rangle = \langle w \mid w \rangle,$$

also $w = 0$, und damit $v = u \in U$. □

Bemerkung 25.12. Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt. Seien $U, W \subseteq V$ Unterräume mit

(1) $V = U + W$, und
(2) $\langle U \mid W \rangle = 0$, d. h. $\langle u \mid w \rangle = 0$ für alle $u \in U$ und alle $w \in W$.

Dann gilt $W = U^\perp$ und $V = U \perp W$.

Beweis. BEWEIS ALS ÜBUNG.

Sei $w \in W$. Dann gilt $\langle u \mid w \rangle = 0$ für alle $u \in U$, also $w \in U^\perp$. Sei umgekehrt $v \in U^\perp$. Schreibe $v = u + w$ mit $u \in U$ und $w \in W$. Dann gilt $0 = \langle v \mid u \rangle = \langle u + w \mid u \rangle = \langle u \mid u \rangle + \langle w \mid u \rangle = \langle u \mid u \rangle$, und es folgt $u = 0$. Damit $v = w \in W$. Insgesamt $W = U^\perp$. □

Satz 25.13. *Sei $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt und $e_1, \ldots, e_n$ eine ONB von V. Dann gilt:*

(1) *Für jedes $x \in V$ gilt $x = \sum_{i=1}^n \langle x \mid e_i \rangle e_i$.*
(2) $\left\| \sum_{i=1}^n \alpha_i e_i \right\| = +\sqrt{\sum_{i=1}^n |\alpha_i|^2}$.
(3) $\langle \sum_{i=1}^n \alpha_i e_i \mid \sum_{i=1}^n \beta_i e_i \rangle = \sum_{i=1}^n \alpha_i \overline{\beta_i}$.

Beweis. Direktes Ausrechnen. Übung. □

26. Der adjungierte Endomorphismus

Sei $K = \mathbb{R}$ oder $K = \mathbb{C}$. Sei im folgenden $(V, \langle - \mid - \rangle)$ ein Vektorraum mit Skalarprodukt der endlichen Dimension n. Im Fall $K = \mathbb{R}$ ist also V ein euklidischer Vektorraum, im Fall $K = \mathbb{C}$ ein unitärer Vektorraum. Sei $K \to K$, $x \mapsto \overline{x}$ die komplexe Konjugation im Fall $K = \mathbb{C}$ und die identische Abbildung im Fall $K = \mathbb{R}$.

Lemma 26.1. *Sei $f\colon V \to K$ eine lineare Abbildung. Dann gibt es ein eindeutig bestimmtes $a \in V$ mit $f(x) = \langle x \mid a \rangle$ für jedes $x \in V$.*

Beweis. *Existenz.* Sei $e_1, \ldots, e_n$ eine ONB von V. Setze

$$a \overset{def}{=} \sum_{i=1}^n \overline{f(e_i)} e_i.$$

Für jedes $x \in V$ ist $x = \sum_{i=1}^n \langle x \mid e_i \rangle e_i$ nach Satz 25.13, und es folgt

$$f(x) = \sum_{i=1}^n \langle x \mid e_i \rangle f(e_i) = \langle x \mid \sum_{i=1}^n \overline{f(e_i)} e_i \rangle = \langle x \mid a \rangle.$$

Eindeutigkeit. Ist auch $a' \in V$ mit $f(x) = \langle x \mid a' \rangle$ für alle $x \in V$, so folgt $0 = \langle x \mid a - a' \rangle$, insbesondere $0 = \langle a - a' \mid a - a' \rangle$, und daher $a - a' = 0$, d.'h. $a = a'$. □

SATZ 26.2. *Zu jedem $f \in \mathrm{End}_K(V)$ gibt es ein eindeutig bestimmtes $f^* \in \mathrm{End}_K(V)$ mit*

$$\langle f(x) \mid y \rangle = \langle x \mid f^*(y) \rangle$$

für alle $x, y \in V$.

f^* heißt der zu f *adjungierte* Endomorphismus von V.

BEWEIS. Sei $f \in \mathrm{End}_K(V)$. Sei $y \in V$. Dann ist die Abbildung $V \to K,\ x \mapsto \langle f(x) \mid y \rangle$ linear, und daher gibt es nach dem vorigen Lemma ein eindeutig bestimmtes $y^* \in V$ mit $\langle f(x) \mid y \rangle = \langle x \mid y^* \rangle$ für alle $x \in V$. Definiere $f^* \colon V \to V$ durch $f^*(y) \stackrel{def}{=} y^*$ für jedes $y \in V$. Nach Konstruktion erfüllt f^* die behauptete Formel für das Skalarprodukt. Es ist f^* linear: Seien $y, z \in V$, sei $\alpha \in K$. Dann gilt

$$\begin{aligned}\langle x \mid f^*(y+z) \rangle &= \langle f(x) \mid y + z \rangle = \langle f(x) \mid y \rangle + \langle f(x) \mid z \rangle \\ &= \langle x \mid f^*(y) \rangle + \langle x \mid f^*(z) \rangle = \langle x \mid f^*(y) + f^*(z) \rangle\end{aligned}$$

und

$$\begin{aligned}\langle x \mid f^*(\alpha y) \rangle &= \langle f(x) \mid \alpha y \rangle = \overline{\alpha} \langle f(x) \mid y \rangle \\ &= \overline{\alpha} \langle x \mid f^*(y) \rangle = \langle x \mid \alpha f^*(y) \rangle,\end{aligned}$$

und es folgt $f^*(y + z) = f^*(y) + f^*(z)$ sowie $f^*(\alpha y) = \alpha f^*(y)$.

Zur Eindeutigkeit: Sei auch $\widetilde{f} \in \mathrm{End}_K(V)$ mit $\langle f(x) \mid y \rangle = \langle x \mid \widetilde{f}(y) \rangle$ für alle $x, y \in V$. Dann gilt $\langle x \mid (f^* - \widetilde{f})(y) \rangle = 0$ für alle $x, y \in V$, und es folgt $(f^* - \widetilde{f})(y) = 0$ für alle $y \in V$, also $f^* = \widetilde{f}$. □

BEMERKUNG 26.3. Für den adjungierten Endomorphismus gelten folgende Eigenschaften:

(1) $(1_V)^* = 1_V$.
(2) Für $(f \circ g)^* = g^* \circ f^*$.
(3) Für alle $f \in \mathrm{End}_K(V)$ gilt $(f^*)^* = f$.

BEWEIS. (1) ist klar.
(2) Für alle $x, y \in V$ gilt

$$\begin{aligned}\langle f \circ g(x) \mid y \rangle &= \langle f(g(x)) \mid y \rangle = \langle g(x) \mid f^*(y) \rangle \\ &= \langle x \mid g^*(f^*(y)) \rangle = \langle x \mid g^* \circ f^*(y) \rangle.\end{aligned}$$

(3) Für alle $x, y \in V$ gilt

$$\langle x \mid (f^*)^*(y)) \rangle = \langle f^*(x) \mid y \rangle = \overline{\langle y \mid f^*(x) \rangle} = \overline{\langle f(y) \mid x \rangle} = \langle x \mid f(y) \rangle.$$

□

SATZ 26.4. *Sei $f \in \mathrm{End}_K(V)$.*

(1) *Es ist* $\mathrm{Kern}(f^*) = \mathrm{Bild}(f)^{\perp}$ *und* $\mathrm{Bild}(f^*)) = \mathrm{Kern}(f)^{\perp}$.
(2) *Ist U ein f-invarianter Unterraum von V, so ist $U^{\perp}$ ein f^*-invarianter Unterraum von V.*

BEWEIS. (1) Sei $y \in V$. Dann gilt

$$\begin{aligned}y \in \mathrm{Bild}(f)^{\perp} &\Leftrightarrow \langle f(x) \mid y \rangle = 0 \ \ \forall\, x \in V \\ &\Leftrightarrow \langle x \mid f^*(y) \rangle = 0 \ \ \forall\, x \in V \ \Leftrightarrow\ f^*(y) = 0 \\ &\Leftrightarrow y \in \mathrm{Kern}(f^*).\end{aligned}$$

Wendet man dies nun auf f^* statt f an, so folgt $\mathrm{Bild}(f^*)^\perp = \mathrm{Kern}((f^*)^*) = \mathrm{Kern}(f)$, und damit

$$\mathrm{Bild}(f^*) = (\mathrm{Bild}(f^*)^\perp)^\perp = \mathrm{Kern}(f)^\perp.$$

(2) Sei $x \in U^\perp$. Dann gilt $\langle u \mid f^*(x)\rangle = \langle f(u) \mid x\rangle = 0$ für alle $u \in U$, also $f^*(x) \in U^\perp$. □

Satz 26.5. *Sei $\mathcal{B} = (e_1, \ldots, e_n)$ eine ONB von V. Sei $f \in \mathrm{End}_K(V)$. Ist $A = M_\mathcal{B}(f) \in \mathrm{M}(n; K)$, so gilt*

$$M_\mathcal{B}(f^*) = A^* \overset{def}{=} \overline{{}^tA}.$$

Beweis. Für jedes $x \in V$ gilt $x = \sum_{i=1}^n \langle x \mid e_i\rangle e_i$, insbesondere $f(x) = \sum_{i=1}^n \langle f(x) \mid e_i\rangle e_i$. Wendet man das auf $x = e_j$ an, so erhält man $f(e_j) = \sum_{i=1}^n \langle f(e_j) \mid e_i\rangle e_i$, und damit $A = \big(\langle f(e_j) \mid e_i\rangle\big)_{i,\,j}$. Es folgt

$$M_\mathcal{B}(f^*) = \big(\langle f^*(e_j) \mid e_i\rangle\big)_{i,\,j} = \big(\langle e_j \mid f(e_i)\rangle\big)_{i,\,j} = \big(\overline{\langle f(e_i) \mid e_j\rangle}\big)_{i,\,j} = \overline{{}^tA} = A^*.$$

□

Folgerung 26.6. *Seien f, $g \in \mathrm{End}_K(V)$, sei $\alpha \in K$. Dann gilt*

(1) $(f + g)^* = f^* + g^*$.
(2) $(\alpha f)^* = \overline{\alpha} f^*$.

Beweis. Dies folgt einerseits genau wie in 26.2, andererseits aber auch durch Betrachtung der darstellenden Matrizen bzgl. einer ONB. □

Folgerung 26.7. *Sei $f \in \mathrm{End}_K(V)$. Dann gilt:*

(1) $\mathrm{rang}(f^*) = \mathrm{rang}(f)$.
(2) *Ist f bijektiv, so ist f^* bijektiv, und es gilt $(f^*)^{-1} = (f^{-1})^*$.*

Beweis. (1) folgt sofort aus Satz 26.5. Der erste Teil von (2) folgt sofort daraus, der zweite Teil folgt mit 26.2. □

27. Isometrien

Es gelten die gleichen Voraussetzungen wie im vorigen Abschnitt.

Definition 27.1. (1) $f \in \mathrm{End}_K(V)$ heißt *Isometrie*, falls

$$\langle f(x) \mid f(y)\rangle = \langle x \mid y\rangle$$

für alle x, $y \in V$ gilt. (f ist “winkeltreu”)

(2) Gilt $K = \mathbb{C}$, so nennt man eine Isometrie f auch *unitäre* Abbildung.
(3) Gilt $K = \mathbb{R}$, so nennt man eine Isometrie f auch *orthogonale* Abbildung.

Bemerkung 27.2. (1) 1_V ist eine Isometrie.
(2) Sind f, $g \in \mathrm{End}_K(V)$ Isometrien, so auch $f \circ g$.

Beweis. (1) ist klar. Zu (2). Für alle x, $y \in V$ gilt

$$\langle f \circ g(x) \mid f \circ g(y)\rangle = \langle f(g(x)) \mid f(g(y))\rangle = \langle f(x) \mid f(y)\rangle = \langle x \mid y\rangle.$$

□

Lemma 27.3. *Sei $f \in \mathrm{End}_K(V)$. Äquivalent sind:*

(1) *f ist Isometrie.*
(2) *Es gilt $\|f(x)\| = \|x\|$ für alle $x \in V$. (f ist “längentreu”)*

Beweis. (1)⇒(2): Für alle $x \in V$ ist $\|f(x)\|^2 = \langle f(x) \mid f(x)\rangle = \langle x \mid x\rangle = \|x\|^2$.
(2)⇒(1): Es gelte $K = \mathbb{R}$: Für alle $x, y \in V$ gilt

$$\langle x \mid y\rangle = \frac{1}{2}\Big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\Big).$$

Dann ist klar, wie (1) aus (2) folgt.

Es gelte $K = \mathbb{C}$. Es gilt

$$\operatorname{Re}(\langle x \mid y\rangle) = \frac{1}{2}\Big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\Big)$$

und $\operatorname{Im}(\langle x \mid y\rangle) = \operatorname{Re}(-i\langle x \mid y\rangle)$, und auch in diesem Fall folgt dann (1) aus (2). □

Satz 27.4. *Sei $f \in \operatorname{End}_K(V)$ eine Isometrie. Dann gilt:*
(1) *f ist bijektiv.*
(2) *f^{-1} ist eine Isometrie.*
(3) *Es gilt $f^* = f^{-1}$.*

Beweis. f ist injektiv: Denn gilt $f(x) = 0$, so ist $0 = \|f(x)\| = \|x\|$, also auch $x = 0$. Wegen $\dim(V) < \infty$ folgt, dass $f\colon V \to V$ auch surjektiv ist. Weiter folgt für alle $x, y \in V$

$$\langle f^{-1}(x) \mid f^{-1}(y)\rangle = \langle f(f^{-1}(x)) \mid f(f^{-1}(y))\rangle = \langle x \mid y\rangle,$$

also ist auch f^{-1} Isometrie, und ferner

$$\langle x \mid y\rangle = \langle f(x) \mid f(y)\rangle = \langle x \mid f^*(f(y))\rangle,$$

was $f^* \circ f = 1_V$ impliziert. Damit folgt weiter

$$f^* = f^* \circ 1_V = f^* \circ (f \circ f^{-1}) = (f^* \circ f) \circ f^{-1} = 1_V \circ f^{-1} = f^{-1}.$$

□

Satz 27.5. *Sei $f \in \operatorname{End}_K(V)$. Sei $\mathcal{B} = (e_1, \ldots, e_n)$ eine ONB von V. Sei Sei $A = M_{\mathcal{B}}(f) \in \mathrm{M}(n;K)$. Äquivalent sind:*
(1) *f ist eine Isometrie.*
(2) *f ist bijektiv und es gilt $f^{-1} = f^*$.*
(3) *Es gilt $f^* \circ f = 1_V$.*
(4) *Es gilt $f \circ f^* = 1_V$.*
(5) *$(f(e_1), \ldots, f(e_n))$ ist eine ONB von V.*
(6) *Es ist $A \in \mathrm{GL}(n;K)$ mit $A^{-1} = A^*$.*
(7) *Es gilt $A^* \cdot A = E_n$.*
(8) *Es gilt $A \cdot A^* = E_n$.*
(9) *Die Spalten von A bilden eine ONB von $(K^n, (- \mid -))$.*
(10) *Die Zeilen von A bilden eine ONB von $(\mathrm{M}(1,n;K), (- \mid -))$ (analog definiert).*

Definition 27.6. (1) $A \in \mathrm{M}(n;\mathbb{C})$ heißt *unitär*, falls $A^* \cdot A = E_n$ gilt.
(2) $A \in \mathrm{M}(n;\mathbb{R})$ heißt *orthogonal*, falls ${}^tAA = E_n$ gilt.

Folgerung 27.7. (1) *Für eine unitäre Matrix $A \in \mathrm{M}(n;\mathbb{C})$ gilt $|\det(A)| = 1$.*
(2) *Für eine orthogonale Matrix $A \in \mathrm{M}(n;\mathbb{R})$ gilt $\det(A) = \pm 1$.*

Beweis. Klar. □

Beweis von Satz 27.5. Die Äquivalenz von (1), (2), (3), (4), (6), (7), und (8) ist zusammen mit dem zuvor gesagten klar. Auch ist klar, dass (5) aus (1) folgt.

Gelte (5). Seien $x = \sum_{i=1}^n \alpha_i e_i$, $y = \sum_{i=1}^n \beta_i e_i$. Nach 25.13 gilt $\langle x \mid y\rangle = \sum_{i=1}^n \alpha_i \overline{\beta_i}$. Außerdem gilt $f(x) = \sum_{i=1}^n \alpha_i f(e_i)$, $f(y) = \sum_{i=1}^n \beta_i f(e_i)$, und wiederum aus 25.13, nun angewandt auf die ONB $f(e_1), \ldots f(e_n)$, folgt $\langle f(x) \mid f(y)\rangle = \sum_{i=1}^n \alpha_i \overline{\beta_i} = \langle x \mid y\rangle$. Also ist f Isometrie.

(9) folgt aus (5), angewandt auf die Standardbasis $e_1, \ldots, e_n$. Analog folgt (10) aus (5). Ebenso folgt (5) aus (9) bzw. aus (10). □

Satz und Definition 27.8. (1) *Die Menge der unitären Matrizen bildet eine Untergruppe von* $\mathrm{GL}(n;\mathbb{C})$*, die mit* $\mathrm{U}(n)$ *bezeichnet wird. "Unitäre Gruppe vom Grad* n*".*

(2) $\mathrm{SU}(n) \stackrel{def}{=} \{A \in \mathrm{U}(n) \mid \det(A) = 1\}$ *ist eine Untergruppe von* $\mathrm{O}(n)$*. Deren Elemente heißen spezielle unitäre Matrizen. "Spezielle unitäre Gruppe vom Grad* n*".*

(3) *Die Menge der orthogonalen Matrizen bildet eine Untergruppe von* $\mathrm{GL}(n;\mathbb{R})$*, die mit* $\mathrm{O}(n)$ *bezeichnet wird. "Orthogonale Gruppe vom Grad* n*".*

(4) $\mathrm{SO}(n) \stackrel{def}{=} \{A \in \mathrm{O}(n) \mid \det(A) = 1\}$ *ist eine Untergruppe von* $\mathrm{O}(n)$*. Deren Elemente heißen speziell orthogonale Matrizen. "Spezielle orthogonale Gruppe vom Grad* n*".*

Beweis. Das ist nun klar. □

Bemerkung 27.9. Analog bildet die Menge der Isometrien f von V eine Gruppe, die im Fall $K = \mathbb{C}$ mit $\mathrm{U}(V)$ (bzw. $SU(V)$, falls $\det(f) = 1$), und im Fall $K = \mathbb{R}$ mit $\mathrm{O}(V)$ (bzw. $\mathrm{SO}(V)$, falls $det(f) = 1$) bezeichnet wird.

Satz 27.10. *Sei* $f \in \mathrm{End}_K(V)$ *eine Isometrie. Sei* $\lambda \in K$ *ein Eigenwert von* f*. Dann gilt* $|\lambda| = 1$*.*

Beweis. Sei x ein zu λ gehöriger Eigenvektor. Dann gilt

$$\langle x \mid x \rangle = \langle f(x) \mid f(x) \rangle = \langle \lambda x \mid \lambda x \rangle = \lambda\overline{\lambda}\langle x \mid x \rangle,$$

und wegen $\langle x \mid x \rangle \neq 0$ folgt $|\lambda|^2 = \lambda\overline{\lambda} = 1$. □

Unitäre Endomorphismen. Wir behandeln zunächst den Fall $K = \mathbb{C}$, was wir nun voraussetzen.

Satz 27.11. *Sei* $f \in \mathrm{U}(V)$*. Dann gibt es eine ONB bestehend aus Eigenvektoren von* f*.*

Beweis. Induktion nach $n = \dim(V)$.

Sei $n = 1$. Es existiert ein Eigenwert $\lambda \in \mathbb{C}$, und normiert man einen zugehörigen Eigenvektor, so liefert dieser eine ONB von V bestehend aus Eigenvektoren.

Sei nun $n \geq 2$. In $\mathbb{C}$ existiert ein Eigenwert von f (Fundamentalsatz der Algebra), und (nach Normierung) ein zugehöriger normierter Eigenvektor $e_1 \in V$. Sei $U \stackrel{def}{=} \mathrm{Span}(e_1)$. Dann gilt $V = U \perp U^\perp$, und es ist $\dim(U^\perp) = n - 1$. Es ist U offenbar f-invariant, also ist (wegen $f^* = f^{-1}$) offenbar auch $U^\perp$ ein f-invarianter Unterraum. Nach Induktionsvoraussetzung hat $U^\perp$ eine ONB $e_2, \ldots, e_n$ bestehend aus Eigenvektoren von $f\,|_{U^\perp} : U^\perp \to U^\perp$. Dies sind dann auch Eigenvektoren von f, und wegen $V = U \perp U^\perp$ ist $e_1, e_2, \ldots, e_n$ eine ONB von V, bestehend aus Eigenvektoren von f. □

Folgerung 27.12. *Sei* $A \in \mathrm{U}(n)$*. Dann gibt es ein* $S \in \mathrm{U}(n)$*, so dass* $S^{-1}AS = S^*AS$ *eine Diagonalmatrix ist.*

Man sagt auch, A sei *unitär-diagonalisierbar*.

Orthogonale Endomorphismen. Wir behandeln jetzt den etwas komplizierteren Fall $K = \mathbb{R}$, was wir nun voraussetzen.

Zunächst schauen wir uns für kleine $n = \dim(V)$ die orthogonalen Endomorphismen auf V an.

27.13 ($n = 1$). Für jedes $x \in V$ ist nur $f(x) = \pm x$ möglich, und es folgt $f(x) = x$ für alle $x \in V$ oder $f(x) = -x$ für alle $x \in V$.

27.14 ($n = 2$). Es genügt, die orthogonalen Matrizen $A \in \mathrm{M}(2;\mathbb{R})$ zu bestimmen. Sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}(2;\mathbb{R})$$

orthogonal, d. h. es gilt $A \cdot {}^t A = E_2$. Dies bedeutet

$$a^2 + b^2 = 1,\ c^2 + d^2 = 1,\ ac + bd = 0.$$

Die ersten beiden Bedingungen bedeuten, dass es α, $\beta \in [0, 2\pi[$ gibt mit

$$a = \cos\alpha,\ b = \sin\alpha \text{ und } c = \sin\beta,\ d = \cos\beta.$$

ZEICHNUNG (Einheitskreis)

Aus der dritten Bedingung folgt $\sin(\alpha+\beta) = \cos\alpha\sin\beta + \sin\alpha\cos\beta = 0$, und daher ist $\alpha + \beta$ ein ganzahliges Vielfaches von π, d. h. entweder ein geradzahliges oder ein ungeradzahliges Vielfaches von π. Es folgt:

(1) entweder $c = \sin\beta = -\sin\alpha$ und $d = \cos\beta = \cos\alpha$,
(2) oder $c = \sin\beta = \sin\alpha$ und $d = \cos\beta = -\cos\alpha$.

Es gibt also ein eindeutiges $\alpha \in [0, 2\pi[$ mit

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

oder

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

Ferner gilt: Ist $\det(A) = 1$ (oberer Fall), so hat A nur dann reelle Eigenwerte, wenn $\alpha = 0$ oder $\alpha = \pi$ gilt. Denn es ist $\chi_A = T^2 - 2\cos(\alpha)T + 1$, und die (komplexen) Nullstellen sind $\cos\alpha \pm \sqrt{\cos^2\alpha - 1}$, also reell genau dann, wenn $\cos^2\alpha = 1$, d. h. $\cos\alpha = \pm 1$ gilt.

Ist hingegen $\det(A) = -1$ (unterer Fall), so hat A die Eigenwerte $+1$ und -1, und die zugehörigen Eigenvektoren stehen senkrecht aufeinander. Denn es ist $\chi_A = T^2 - 1 = (T-1)(T+1)$; ist x Eigenvektor zum EW $+1$ und y Eigenvektor -1, so gilt $(x \mid y) = (Ax \mid Ay) = (x \mid -y) = -(x \mid y)$, also $(x \mid y) = 0$.

Geometrische Interpretation: Es ist

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix},$$

und man sieht, dass dadurch eine Drehung in $\mathbb{R}^2$ um den Nullpunkt um den Winkel α durchgeführt wird. Hingegen beschreibt

$$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$

eine Spiegelung an der x-Achse, und damit beschreibt $\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$ eine "Drehspiegelung", d. h. eine Spiegelung gefolgt von einer Drehung (oder auch umgekehrt).

27.15 ($n = 3$). Sei $(V, \langle - \mid - \rangle)$ ein dreidimensionaler euklidischer Vektorraum. Sei $f\colon V \to V$ orthogonal. Das charakteristische Polynom $\chi_f \in \mathbb{R}[T]$ hat nach dem Zwischenwertsatz aus der Analysis mindestens eine reelle Nullstelle λ_1. Nach Satz 27.10 gilt $\lambda_1 = \pm 1$. Sei x_1 ein Eigenvektor zu λ_1. Ohne Einschränkung gelte $\|x_1\| = 1$. Man kann x_1 zu einer ONB $\mathcal{B} = \{x_1, x_2, x_3\}$ von V ergänzen. Es ist dann

$$M_{\mathcal{B}}(f) = A = \left(\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & & \\ 0 & & A' \end{array}\right).$$

Mit A ist offenbar auch $A' \in \mathrm{M}(2;\mathbb{R})$ eine orthogonale Matrix, d. h. $A' \in \mathrm{O}(2)$. Zwei Fälle:

(1) $\det(f) = \det(A) = +1$. Ist $\lambda_1 = -1$, so muss $\det(A') = -1$ sein. Wie in der Behandlung des Falls $n = 2$ ersichtlich, kann man x_2, x_3 so wählen, dass $\lambda_2 = +1$ und $\lambda_3 = -1$ gilt, und dann

$$A = \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix}.$$

Dies ist eine Drehung mit Achse $\mathbb{R}e_2$ und Winkel π. (Hierbei ist e_1, e_2, e_3 die Standardbasis von $\mathbb{R}^3$.)

Gilt aber $\lambda_1 = +1$, so muss $\det(A') = +1$ sein, also gibt es ein (eindeutiges) $\alpha \in [0, 2\pi[$ mit

$$A = \begin{pmatrix} 1 & & \\ & \cos\alpha & -\sin\alpha \\ & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Dies ist eine Drehung mit Achse $\mathbb{R}e_1$ und Winkel α.

(2) $\det(f) = \det(A) = -1$. Analog zum vorigen Fall gibt es nach geeigneter Wahl von x_2, x_3 die Möglichkeiten

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

(Spiegelung an der e_1-e_2-Ebene) und

$$A = \begin{pmatrix} -1 & & \\ & \cos\alpha & -\sin\alpha \\ & \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} -1 & & \\ & \cos\alpha & -\sin\alpha \\ & \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix},$$

eine Drehspiegelung, hier eine Spiegelung an der e_2-e_3-Ebene gefolgt von einer Drehung mit Achse $\mathbb{R}e_1$ und Winkel α. (Als Spezialfall $\alpha = \pi$ erhält man $A = -E_3$, eine Punktspiegelung am Nullpunkt.

SATZ 27.16 (Normalform orthogonaler Endomorphismen). *Sei $(V, \langle - \mid - \rangle)$ ein n-dimensionaler euklidischer Vektorrau. Sei $f \in \mathrm{O}(V)$. Dann gibt es eine ONB $\mathcal{B}$, so dass*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} +1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & +1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{A_1} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \boxed{A_t} \end{pmatrix}$$

mit $r \geq 0$ vielen $+1$, mit $s \geq 0$ vielen -1, mit $t \geq= 0$ mit $r + s + 2t = n$, und zu jedem $i \in \{1, \ldots, t\}$ gibt es $\alpha_i \in [0, 2\pi[$, $\alpha_i \neq 0, \pi$ und

$$A_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix} \in \mathrm{SO}(2).$$

Zum Beweis verwenden wir folgendes

LEMMA 27.17. *Sei $f \in \mathrm{End}_{\mathbb{R}}(V)$ ein orthogonaler Endomorphismus, wobei* $\dim(V) \geq 1$ *gelte. Dann gibt es einen f-invarianten Unterraum U von V mit* $1 \leq \dim(U) \leq 2$.

Beweis von Satz 27.16. Induktion nach $n = \dim(V)$. Für $n = 1$ ist die Sache klar. Gelte nun $n \geq 2$. Nach dem Lemma gibt es einen f-invarianten Unterraum U von V, mit $\dim(U) = 1$ oder $\dim(U) = 2$. Es gilt $V = U \oplus U^\perp$. Wegen $f^* = f^{-1}$ ist $U^\perp$ auch ein f-invarianter Unterraum von V, und $f\,|_{U^\perp} \colon U^\perp \to U^\perp$ ist orthogonal, so dass man auf $f\,|_{U^\perp}$ die Induktionsvoraussetzung anwenden kann: Es gibt eine ONB $\mathcal{B}'$ von $U^\perp$, so dass die diesbzgl. Darstellungsmatrix die behauptete Form hat. Wir müssen uns noch um U selbst kümmern.

Ist $\dim(U) = 1$, so enthält U einen normierten Eigenvektor zu einem Eigenwert $\lambda_1 = +1$ oder $\lambda_2 = -1$. Mit diesem Eigenvektor ergänzt man $\mathcal{B}'$ zu einer ONB $\mathcal{B}$ von V.

Ist $\dim(U) = 2$, so hat U eine ONB e_1, e_2, so dass $f\,|_U \colon U \to U$ bzgl. dieser Basis dargestellt wird durch eine Matrix der Form

$$\begin{pmatrix} \pm 1 & \\ & \pm 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \text{mit } \alpha \neq 0,\, \pi.$$

Mit e_1, e_2 ergänzt man $\mathcal{B}'$ zu einer ONB von V mit den gewünschten Eigenschaften. □

Beweis des Lemmas. Sei $A \in \mathrm{M}(n;\mathbb{R})$ die Darstellungsmatrix von f bzgl. einer gewählten ONB. Es genügt zeigen, dass es einen Unterraum U von $\mathbb{R}^n$ gibt mit $A \cdot U \subseteq U$ und mit $\dim(U) = 1$ oder $\dim(U) = 2$.

Wir "komplexifizieren": Definiere $g \colon \mathbb{C}^n \to \mathbb{C}^n$ durch $g(z) = Az$ (für $z \in \mathbb{C}^n$). Bzgl. der Standardbasis wird g dargestellt durch A, ist also unitär. Es gibt einen Eigenwert $\lambda \in \mathbb{C}$, also dazu ein $z \in \mathbb{C}^n$ mit $z \neq 0$ und $g(z) = \lambda z$. Außerdem gilt, da A reell ist,

$$A\overline{z} = \overline{Az} = \overline{\lambda z} = \overline{\lambda}\,\overline{z}. \tag{27.1}$$

Also ist $\overline{z}$ ein Eigenvektor zu $\overline{\lambda}$. Daraus werden nun zwei reelle Vektoren definiert, durch

$$x \overset{def}{=} \frac{1}{2}(z + \overline{z}),\ \ y \overset{def}{=} \frac{1}{2i}(z - \overline{z}) \in \mathbb{R}^n.$$

Es gilt $z = x + iy$. Sei $U \overset{def}{=} \mathrm{Span}(x,\, y) \subseteq \mathbb{R}^n$. Dann gilt $1 \leq \dim_{\mathbb{R}}(U) \leq 2$. Schreibe $\lambda = a + bi$, mit a, $b \in \mathbb{R}$. Dann folgt

$$\begin{aligned} Ax &= \frac{1}{2}(Az + A\overline{z}) = \frac{1}{2}(\lambda z + \overline{\lambda}\overline{z}) = \mathrm{Re}(\lambda z) = ax - by, \\ Ay &= \frac{1}{2i}(Az - A\overline{z}) = \frac{1}{2i}(\lambda z - \overline{\lambda}\overline{z}) = \mathrm{Im}(\lambda z) = bx + ay. \end{aligned}$$

Es folgt $A \cdot U \subseteq U$. □

Folgerung 27.18. *Sei $A \in \mathrm{O}(n)$. Dann gibt es $S \in \mathrm{O}(n)$, so dass $S^{-1}AS = {}^tSAS$ eine Matrix wie in Satz 27.16 ist.*

Rechenbeispiele.

Lemma 27.19. *Sei $f \in \mathrm{U}(V)$ oder $f \in \mathrm{O}(V)$. Sind x, $y \in V$ Eigenvektoren zu* ver- schiedenen *Eigenwerten λ, $\mu \in K$, so gilt $\langle x \mid y \rangle = 0$.*

Beweis. Es gilt

$$\langle x \mid y \rangle = \langle f(x) \mid f(y) \rangle = \langle \lambda x \mid \mu y \rangle = \lambda\overline{\mu}\langle x \mid y \rangle.$$

Nimmt man $\langle x \mid y \rangle \neq 0$ an, so folgt $\lambda\overline{\mu} = 1$. Wegen $|\mu| = 1$ folgt aber $\overline{\mu} = \mu^{-1}$, und damit $\lambda = \mu$. □

27.20. Sei A unitär. Dann ist A diagonalisierbar. Man berechnet Basen aller Eigenräume und normalisiert diese (einzelnen) Basen mit dem Schimdtschen Orthonormalisierungsverfahren. Nach dem vorigen Lemma erhält man damit eine ONB $e_1, \ldots, e_n$ von $\mathbb{C}^n$. Sei $S = (e_1, \ldots, e_n) \in \mathrm{M}(n;\mathbb{C})$ die aus den Spaltenvektoren $e_1, \ldots, e_n$ zusammengesetzte Matrix. Diese ist unitär, und es ist S^*AS eine Diagonalmatrix, bei der die Eigenwerte von A (entsprechend ihrer Vielfachheit) auf der Hauptdiagonalen stehen.

Beispiel 27.21. Sei

$$A = \begin{pmatrix} 2/\sqrt{5} & i/\sqrt{5} & 0 \\ i/\sqrt{5} & 2/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{M}(3; \mathbb{C}).$$

Offenbar bilden die Spalten eine ONB von $(\mathbb{C}^3, (- \mid -))$, daher ist A unitär. Es gilt

$$\chi_A = \big((T - 2/\sqrt{5})(T - 2\sqrt{5}) + 1/5)\big)(T - 1) = (T^2 - 4/\sqrt{5}T + 1)(T - 1).$$

Eigenwerte sind also 1, $2/\sqrt{5} \pm i/\sqrt{5}$. Man rechnet:

$$\begin{aligned} E(A, 1) &= \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle \\ E(A, 2/\sqrt{5} + i/\sqrt{5}) &= \langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \rangle \\ E(A, 2/\sqrt{5} - i/\sqrt{5}) &= \langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix} \rangle. \end{aligned}$$

Nach dem Lemma bilden die normierten Vektoren zusammen eine ONB von $\mathbb{C}^n$. Es ist also

$$S = \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 1 & 0 & 0 \end{pmatrix} \in \mathrm{U}(n),$$

und es gilt

$$S^* A S = \begin{pmatrix} 1 & & \\ & 2/\sqrt{5} + i/\sqrt{5} & \\ & & 2/\sqrt{5} - i/\sqrt{5} \end{pmatrix}.$$

27.22. Sei $A \in \mathrm{O}(n)$. Wir fassen A als komplexe Matrix auf. Dann gilt $A \in \mathrm{U}(n)$. Es gibt also eine ONB von $\mathbb{C}^n$ bestehend aus Eigenvektoren $e_1, \ldots, e_n$ von A. Seien ohne Einschränkung $e_1, \ldots, e_r$ die Eigenvektoren zu reellen Eigenwerten $\lambda_1, \ldots, \lambda_r \in \{\pm 1\}$, die restlichen Basiselemente seien Eigenvektoren zu den komplexen, nicht-reellen Eigenwerten, und wegen (27.1) sowie dem vorigen Lemma kann man ohne Einschränkung annehmen, dass diese Basiselemente von der Form e_{r+1}, $\overline{e_{r+1}}, \ldots, e_{r+k}$, $\overline{e_{r+k}}$ sind ($r + 2k = n$). Für $j = r + 1, \ldots, n$ setzt man dann

$$x_j \overset{def}{=} \frac{1}{2}(e_j + \overline{e_j}),\ y_j \overset{def}{=} \frac{1}{2i}(e_j - \overline{e_j}) \in \mathbb{R}^n.$$

Offenbar gilt $\|x_j\| = 1/\sqrt{2} = \|y_j\|$. Man sieht:

(1) $e_1, \ldots, e_r$, $\sqrt{2}x_{r+1}$, $\sqrt{2}y_{r+1}, \ldots, \sqrt{2}x_n$, $\sqrt{2}y_n$ ist eine ONB von $\mathbb{R}^n$.
(2) Die aus deren Spaltenvektoren zusammengesetzte Matrix $S \in \mathrm{M}(n; K)$ ist folglich orthogonal, und es ist ${}^tSAS$ eine Matrix wie in Satz 27.16.

Beispiel 27.23. Sei

$$A = \begin{pmatrix} 1/2 & -1/\sqrt{(2)} & 1/2 \\ 1/\sqrt{(2)} & 0 & -1/\sqrt{(2)} \\ 1/2 & 1/\sqrt{(2)} & 1/2 \end{pmatrix} \in \mathrm{M}(3; \mathbb{R}).$$

Es ist

$$\chi_A = T^3 - T^2 + T - 1 = (T - 1)(T^2 + 1) = (T - 1)(T - i)(T + i).$$

Die komplexen Eigenwerte sind also 1, i, $-i = \bar{i}$. Man rechnet: Es wird $E(A,1)$ durch den Vektor $v_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}$ erzeugt, und daher auch durch den normierten Vektor $e_1 = \begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{pmatrix}$.

Rechnet man in $\mathbb{C}^n$, so ergibt sich, dass $E(A,i) = \langle \begin{pmatrix}-1\\i\sqrt{2}\\1\end{pmatrix} \rangle = \langle e_2 \rangle$ mit $e_2 = \begin{pmatrix}-1/2\\i/\sqrt{2}\\1/2\end{pmatrix}$.

Es folgt $E(A,-i) = \langle \overline{e_2} \rangle = \langle \begin{pmatrix}-1/2\\-i/\sqrt{2}\\1/2\end{pmatrix} \rangle$. Man bildet

$$x_2 = \frac{1}{2}(e_2 + \overline{e_2}) = \begin{pmatrix}-1/2\\0\\1/2\end{pmatrix}, \; y_2 = \frac{1}{2i}(e_2 - \overline{e_2}) = \begin{pmatrix}0\\1/\sqrt{2}\\0\end{pmatrix}$$

und multipliziert diese noch mit $\sqrt{2}$. Man erhält die orthogonale Matrix

$$S = [e_1, \sqrt{2}x_2, \sqrt{2}y_2] = \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2} & 0\\ 0 & 0 & 1\\ 1/\sqrt{2} & 1/\sqrt{2} & 0\end{pmatrix},$$

und damit gilt

$${}^tSAS = \left(\begin{array}{c|cc} 1 & & \\ \hline & & 1 \\ & -1 & \end{array}\right) = \begin{pmatrix}1 & & \\ & \cos 3\pi/2 & -\sin 3\pi/2\\ & \sin 3\pi/2 & \cos 3\pi/2\end{pmatrix}.$$

28. Selbstadjungierte Endomorphismen

Es gelten die gleichen Voraussetzungen wie im vorigen Abschnitt.

Definition 28.1. (1) $f \in \operatorname{End}_K(V)$ heißt *selbstadjungiert*, falls $f^* = f$ gilt. Gilt $K = \mathbb{R}$, so nennt man f auch *symmetrisch*.

(2) $A \in \mathrm{M}(n;K)$ heißt *hermitesch*, wenn $A^* = A$ gilt. Im Fall $K = \mathbb{R}$ bedeutet dies gerade ${}^tA = A$, und A heißt *symmetrisch*.

Bemerkung 28.2. Sei $f \in \operatorname{End}_K(V)$. Sei $\mathcal{B} = (e_1, \dots, e_n)$ eine ONB von V. Sei $A = M_{\mathcal{B}}(f) \in \mathrm{M}(n;K)$. Äquivalent sind:

(1) f ist selbstadjungiert (bzw. symmetrisch).
(2) Es gilt $\langle f(x) \mid y \rangle = \langle x \mid f(y) \rangle$ für alle $x, y \in V$.
(3) A ist hermitesch (bzw. symmetrisch).

Satz 28.3. *Sei $f \in \operatorname{End}_K(V)$ selbstadjungiert. Dann hat das charakteristische Polynom χ_f reelle Koeffizienten und zerfällt in $\mathbb{R}[T]$ vollständig in Linearfaktoren.*

Beweis. Sei zunächst $K = \mathbb{C}$. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von f, und $x \in V$ ein zugehöriger Eigenvektor. Dann gilt

$$\lambda\langle x \mid x \rangle = \langle \lambda x \mid x \rangle = \langle f(x) \mid x \rangle = \langle x \mid f(x) \rangle = \langle x \mid \lambda x \rangle = \overline{\lambda}\langle x \mid x \rangle,$$

und wegen $\langle x \mid x \rangle \neq 0$ folgt $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. Da $\mathbb{C}$ algebraisch abgeschlossen ist (Fundamentalsatz der Algebra), zerfällt χ_f in $\mathbb{C}[T]$ in Linearfaktoren. Da die Nullstellen von χ_f wie gezeigt alle reell sind, folgt $\chi_f \in \mathbb{R}[T]$.

Sei nun $K = \mathbb{R}$. Sei A die Darstellungsmatrix von f bzgl. einer ONB von V. Dann gilt $A^* = {}^tA = A$. Definiere nun $g\colon \mathbb{C}^n \to \mathbb{C}^n$, $x \mapsto Ax$. Ist $\mathbb{C}^n$ mit dem Standardskalarprodukt ausgestattet, so ist g offenbar ein selbstadjungierter Endomorphismus, denn bzgl. der Standardbasis wird g dargestellt durch die Matrix A. Da aber $\chi_f = \chi_A = \chi_g$ gilt, folgt die Aussage nun aus dem ersten Teil des Beweises. $\square$

FOLGERUNG 28.4. (1) *Eine hermitesche Matrix* $A \in \mathrm{M}(n;\mathbb{C})$ *hat nur reelle Eigenwerte.*
(2) *Eine reelle symmetrische Matrix* $A \in \mathrm{M}(n;\mathbb{R})$ *hat auch als Matrix aufgefasst in* $\mathrm{M}(n;\mathbb{C})$ *nur reelle Eigenwerte.*

SATZ 28.5 (Spektralsatz für selbstadjungierte Endomorphismen). *Sei* $f \in \mathrm{End}_K(V)$ *ein selbstadjungierter Endomorphismus. Dann gibt es eine ONB von* V*, die nur aus Eigenvektoren von* f *besteht.*

BEWEIS. Induktion nach $n = \dim(V)$.
$n = 1$. Da f selbstadjungiert ist, besitzt f einen (reellen) Eigenwert λ. Normiert man einen zugehörigen Eigenvektor x, so bildet dieser eine ONB von V, die nur aus Eigenvektoren von f besteht.
Sei nun $n \geq 2$. Da f selbstadjungiert ist, existiert ein (reeller) Eigenwert von f, und (nach Normierung) ein zugehöriger normierter Eigenvektor $e_1 \in V$. Sei $U \stackrel{def}{=} \mathrm{Span}(e_1)$. Dann gilt $V = U \perp U^\perp$, und es ist $\dim(U^\perp) = n-1$. Es ist U offenbar f-invariant, also ist (wegen $f^* = f$) auch $U^\perp$ ein f-invarianter Unterraum. Nach Induktionsvoraussetzung hat $U^\perp$ eine ONB $e_2, \ldots, e_n$ bestehend aus Eigenvektoren von $f\,|_{U^\perp} \colon U^\perp \to U^\perp$. Dies sind dann auch Eigenvektoren von f, und wegen $V = U \perp U^\perp$ ist $e_1, e_2, \ldots, e_n$ eine ONB von V, bestehend aus Eigenvektoren von f. □

Das folgende ist eine etwas ausführlichere Formulierung des Spektralsatzes.

SATZ 28.6. *Sei* $f \in \mathrm{End}_K(V)$*. Folgende Aussagen sind äquivalent:*
(1) f *ist selbstadjungiert.*
(2) f *hat nur reelle Eigenwerte und* V *hat eine ONB bestehend aus Eigenvektoren von* f*.*
(3) *Es gibt eine ONB* $\mathcal{B}$*, so dass* $M_{\mathcal{B}}(f)$ *eine Diagonalmatrix mit reellen Einträgen ist.*

BEWEIS. (1)⇒(2) folgt aus den vorigen Sätzen, (2)⇒(3) ist klar. Gelte nun (3). Sei $\mathcal{B}$ eine ONB, so dass $A = M_{\mathcal{B}}(f)$ eine Diagonalmatrix mit reellen Einträgen ist. Dann gilt $A^* = A$, und daher ist $f^* = f$ selbstadjungiert nach obiger Bemerkung. □

FOLGERUNG 28.7. *Sei* $A \in \mathrm{M}(n;K)$ *eine Matrix mit* $A^* = A$ *(also hermitesch oder symmetrisch). Dann gibt es eine unitäre bzw. orthogonale Matrix* $S \in \mathrm{GL}(n;K)$*, so dass* $S^{-1}AS = S^*AS$ *eine Diagonalmatrix mit reellen Einträgen ist.*

Konkrete Berechnungen werden genau wie im unitären Fall durchgeführt. Man berechnet Basen der einzelnen Eigenräume und orthonormalisiert diese. Auch hier gilt:

LEMMA 28.8. *Sei* f *ein selbstadjungierter Endomorphismus von* V*. Sind* x, $y \in V$ *Eigenvektoren zu verschiedenen Eigenwerten* λ *bzw.* μ*, so gilt* $\langle x \mid y\rangle = 0$.

BEWEIS. ÜBUNG.
$\lambda\langle x \mid y\rangle = \langle \lambda x \mid y\rangle = \langle f(x) \mid y\rangle = \langle x \mid f(y)\rangle = \langle x \mid \mu y\rangle = \mu\langle x \mid y\rangle$, und aus $\lambda \neq \mu$ folgt $\langle x \mid y\rangle = 0$. □

29. Symmetrische Bilinearformen

Es gelten die gleichen Voraussetzungen wie im vorigen Abschnitt. (Zu Beginn kann der Körper ganz beliebig sein; später werden wir nur $K = \mathbb{R}$ betrachten.)

DEFINITION 29.1. Sei V ein K-Vektorraum. Eine Abbildung $s \colon V \times V \to K$, $(x, y) \mapsto s(x, y)$ heißt eine *Bilinearform* oder *bilinear*, falls sie linear in der linken und in der rechten Komponenten ist, d. h. falls also für alle x, x', y, $y' \in V$ und α, $\beta \in K$ gilt:
(1) $s(x + x', y) = s(x, y) + s(x', y)$, $s(\alpha x, y) = \alpha s(x, y)$,
(2) $s(x, y + y') = s(x, y) + s(x, y')$, $s(x, \beta y) = \beta s(x, y)$,

Eine Bilinearform s heißt *symmetrisch*, falls für alle x, $y \in V$ gilt $s(x,y) = s(y,x)$.

Beispiel 29.2.
(1) Ein Skalarprodukt auf einem euklidischen Vektorraum ist eine symmetrische Bilinearform. Dies gilt nicht für unitäre Vektorräume.
(2) Sei $A \in \mathrm{M}(n;K)$. Durch $s(x,y) \stackrel{def}{=} {}^t x A y$ wird eine Bilinearform auf K^n definiert. Genau dann ist s symmetrisch, wenn A es ist.
(3) Betrachte konkret $n = 2$ und $A = -E_2$. Dann definiert $s(x,y) = {}^t x A y = -{}^t x y$ *kein* Skalarprodukt, denn wegen $s(x,x) = -{}^t x x \leq 0$ liegt hier keine positive Definitheit vor.

Definition 29.3. Sei $s\colon V \times V \to K$ eine Bilinearform. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Dann heißt die Matrix

$$M_{\mathcal{B}}(s) = M_{\mathcal{B}}(\langle - \mid - \rangle) = \Big(s(v_i, v_j)\Big)_{i,j} \in \mathrm{M}(n;K)$$

die *Darstellungsmatrix* zu s.

Satz 29.4. *Ist die Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V fixiert, so definiert die Zuordnung $s \mapsto A = M_{\mathcal{B}}(s)$ eine bijektive Abbildung von der Menge der Bilinearformen auf V und $\mathrm{M}(n;K)$. Die Umkehrabbildung wird durch die Zuordnung $A \mapsto s$ gegeben, wobei $s(v,w) = {}^t x A y$ für alle v, $w \in V$ gilt, wobei x, $y \in K^n$ die Koordinatenvektoren von v bzw. w bzgl. $\mathcal{B}$ sind.*

Ferner gilt:

$$s \text{ ist symmetrisch} \quad \Leftrightarrow \quad A \text{ ist symmetrisch.}$$

Beweis. Sei s eine Bilinearform und $A = M_{\mathcal{B}}(s)$. Seien $v = \sum_{i=1}^n \alpha_i v_i$ und $w = \sum_{i=1}^n \beta_i v_i$. Dann gilt

$$s(v,w) = s\Big(\sum_{i=1}^n \alpha_i v_i, \sum_{i=1}^n \beta_i v_i\Big) = \sum_{i=1}^n \alpha_i \sum_{j=1}^n \beta_j s(v_i, v_j) = (\alpha_1, \ldots, \alpha_n) \cdot A \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Also ist "$s \mapsto A \mapsto s$" die identische Abbildung. Ist umgekehrt $A = (\alpha_{ij}) \in \mathrm{M}(n;K)$ und s definiert durch die Formel ${}^t x A y$ wie im Satz beschrieben, so gilt $s v_i v_j = {}^t e_i A e_j = \alpha_{ij}$, wobei $e_1, \ldots, e_n$ die Standardbasis von K^n ist. Es ist also auch "$A \mapsto s \mapsto M_{\mathcal{B}}(s)$" die Identität.

Der Zusatz ist unmittelbar klar. □

Definition 29.5. Eine symmetrische Matrix $A \in \mathrm{M}(n;\mathbb{R})$ heißt *positiv definit*, wenn für alle $x \in \mathbb{R}^n$ mit $x \neq 0$ gilt ${}^t x A x = (x \mid Ax) > 0$.

Bemerkung 29.6. Ist $K = \mathbb{R}$, so gilt unter obiger Bijektion offenbar, dass eine Bilinearform s genau dann ein Skalarprodukt ist, wenn deren Darstellungsmatrix bzgl. einer (aller) Basis (Basen) positiv definit ist. Wir werden dafür noch ein anderes Kriterium angeben.

Satz 29.7 (Transformationsformel). *Seien $\mathcal{B}$ und $\mathcal{B}'$ zwei (endliche) Basen von V. Sei $T = T_{\mathcal{B}}^{\mathcal{B}'}$ die zugehörige Übergangsmatrix. Für jede Bilinearform s auf V gilt dann*

$$M_{\mathcal{B}'}(s) = {}^t T \cdot M_{\mathcal{B}}(s) \cdot T.$$

Beweis. Seien v, $w \in V$ und seien x, $y \in K^n$ bzw. x', $y' \in K^n$ die Koordinatenvektoren von v bzw. w bzgl. der Basis $\mathcal{B}$ bzw. $\mathcal{B}'$. Es gilt $x = Tx'$ und $y = Ty'$. Setze noch $A = M_{\mathcal{B}}(s)$ und $A' = M_{\mathcal{B}'}(s)$. Dann gilt

$${}^t x' A' y' = s(v,w) = {}^t x A y = {}^t (Tx') A (Ty) = {}^t x' ({}^t T A T) y',$$

und weil das für alle x', $y' \in K^n$ gilt (nehme insbesondere die Standardbasisvektoren), folgt $A' = {}^t T A T$. □

Die Gestalt der Transformationsformel für Bilinearformen ermöglicht dem Spektralsatz Zugang im symmetrischen Fall:

SATZ 29.8 (Hauptachsentransformation). *Sei $(V, \langle - \mid - \rangle)$ ein (endlichdimensionaler) euklidischer Vektorraum. Sei s eine symmetrische Bilinearform auf V.*

(1) *Dann gibt es eine ONB $\mathcal{B}$ von V, so dass*

$$M_{\mathcal{B}}(s) = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

mit reellen Zahlen $\lambda_1, \dots, \lambda_n$ gilt.

(2) *Es gibt eine Basis $\mathcal{B}'$ von V, so dass*

$$M_{\mathcal{B}'}(s) = \begin{pmatrix} E_k & & \\ & -E_\ell & \\ & & 0 \end{pmatrix}$$

mit Einheitsmatrizen E_k, E_ℓ gilt.

BEWEIS. (1) Sei A die Darstellungsmatrix von s bzgl. irgendeiner ONB von V. Es ist A symmetrisch, und daher gibt es nach Satz 28.5 ein $S \in \mathrm{O}(n)$, so dass ${}^tSAS = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ ist mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von A, die allesamt reell sind. Ist dann aber $\mathcal{B} = (b_1, \dots, b_n)$ die Basis, so dass $S = T^{\mathcal{B}}_{\mathcal{B}'}$ ist, so ist (da S orthogonal ist) auch $\mathcal{B}$ eine ONB und es folgt die Formel in (1) aus der Transformationsformel.

(2) Mit den Bezeichnungen aus (1) seien die $\lambda_1, \dots, \lambda_n$ ohne Einschränkung so angeordnet, dass $\lambda_1, \dots, \lambda_k > 0$ sind, $\lambda_{k+1}, \dots, \lambda_{k+\ell} < 0$, und $\lambda_{k+\ell+1} = \ldots = \lambda_n = 0$ sind. Man setzt

$$b'_i \overset{def}{=} \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} \cdot b_i & i = 1, \dots, k+\ell \\ b_i & i = k+\ell+1, \dots, n. \end{cases}$$

Dann ist $\mathcal{B}' = (b'_1, \dots, b'_n)$ eine Basis von V, es gilt offenbar

$$s(b'_i, b'_i) = \begin{cases} \frac{s(b_i,b_i)}{|\lambda_i|} = \frac{\lambda_i}{|\lambda_i|} = +1 & i = 1, \dots, k \\ \frac{s(b_i,b_i)}{|\lambda_i|} = \frac{\lambda_i}{|\lambda_i|} = -1 & i = k+1, \dots, k+\ell \\ 0 & \text{sonst} \end{cases}$$

und daher folgt, dass $M_{\mathcal{B}'}(s)$ von der behaupteten Gestalt ist. □

BEMERKUNG 29.9. Es folgt aus dem *Trägheitssatz von Sylvester*, dass die Anzahlen k und ℓ in (2) durch s (und nicht nur durch A) eindeutig definiert sind.

SATZ 29.10 (Positiv-Definitheits-Kriterium). *Sei $A = (\alpha_{ij}) \in \mathrm{M}(n; \mathbb{R})$ symmetrisch. Äquivalent sind:*

(1) *A ist positiv definit;*
(2) *Alle Eigenwert von A sind positiv (> 0).*
(3) *Für alle $k = 1, \dots, n$ gilt für $A_k \overset{def}{=} (\alpha_{ij})_{1 \le i,j \le k} \in \mathrm{M}(k; \mathbb{R})$, dass $\det(A_k) > 0$ ist.*

BEWEIS. (1)⇔(2): Seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A. Es gibt $S \in \mathrm{O}(n)$ mit ${}^tSAS = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Für $x = {}^t(x_1, \dots, x_n) \in \mathbb{R}^n$ gilt daher

$${}^txAx = \sum_{i=1}^n \lambda_i y_i^2$$

mit $y \overset{def}{=} {}^tSx$, und dieser Ausdruck ist für alle $x \neq 0$ genau dann positiv, wenn alle λ_i positiv sind.

(1)⇒(3): Ist A positiv definit, so gilt $\det(A) > 0$. Denn es gibt $S \in \mathrm{O}(n)$ mit ${}^tSAS = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, und wie in "(1)⇔(2)" gezeigt wurde, gilt $\lambda_1, \dots, \lambda_n > 0$. Es folgt $\det(A) = \det(A) \cdot \det(S)^2 = \lambda_1 \cdot \ldots \cdot \lambda_n > 0$.

Sei $k \in \{1, \dots, n\}$. Sei $x \in \mathbb{R}^k$, $x \neq 0$. Sei $x' \in \mathbb{R}^n$ der Vektor, dessen erste k-Komponenten aus x bestehen und die restlichen $= 0$ sind. Dann gilt ${}^txA_kx = {}^tx'Ax' > 0$. Es ist also A_k positiv definit, und daher gilt $\det(A_k) > 0$.

(3)⇒(2): Sei $A \in \mathrm{M}(n; \mathbb{R})$ symmetrisch, so dass $\det(A_k) > 0$ gilt für $k = 1, \dots, n$. Wir zeigen per Induktion nach n, dass dann alle Eigenwerte von A positiv sind per Induktion nach n. Der Fall $n = 1$ ist klar. Sei nun $n \geq 2$. Nach Induktionsvoraussetzung folgt, dass $A_{n-1} \in \mathrm{M}(n-1; \mathbb{R})$ nur positive Eigenwerte hat. Es gibt also ein $T \in \mathrm{O}(n-1)$ mit ${}^tTA_{n-1}T = \mathrm{diag}(\alpha_1, \dots, \alpha_{n-1})$, wobei $\alpha_1, \dots, \alpha_{n-1} > 0$ gilt. Setze

$$S \overset{def}{=} \left(\begin{array}{ccc|c} & & & 0 \\ & T & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & 1 \end{array}\right) \in \mathrm{O}(n).$$

Dann gilt

$${}^tSAS = \left(\begin{array}{ccc|c} \alpha_1 & & 0 & \beta_1 \\ & \ddots & & \vdots \\ 0 & & \alpha_{n-1} & \beta_{n-1} \\ \hline \beta_1 & \dots & \beta_{n-1} & \beta_n \end{array}\right) =: B.$$

Setzt man noch

$$P \overset{def}{=} \left(\begin{array}{ccc|c} & & & \gamma_1 \\ & E_{n-1} & & \vdots \\ & & & \gamma_{n-1} \\ \hline 0 & \dots & 0 & 1 \end{array}\right) \in \mathrm{GL}(n; \mathbb{R}) \quad \text{mit} \quad \gamma_i \overset{def}{=} -\frac{\beta_i}{\alpha_i}, \tag{29.1}$$

so folgt

$${}^tPBP = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} =: D \quad (= {}^t(SP)A(SP)\,).$$

Nach Voraussetzung gilt $\det(A) = \det(A_n) > 0$, damit auch $\det(B) = \det(A) > 0$, und aus $\frac{\alpha_1 \cdot \ldots \cdot \alpha_n}{\det(P)^2}$ folgt, dass auch $\alpha_n > 0$ gilt. Also sind alle $\alpha_1, \dots, \alpha_n$ positiv. Es folgt, dass mit D auch $A = {}^tQDQ$ (mit $Q = (SP)^{-1}$) positiv definit ist, und aus "(1)⇒(2)" folgt, dass alle Eigenwerte von A positiv sind. (Das letzte Argument folgt auch einfacher aus dem Sylvesterschen Trägheitssatz). □

Definition 29.11. Sei $s \colon V \times V \to K$ eine symmetrische Bilinearform. Die Abbildung $q \colon V \to K$, die durch $q(x) = s(x, x)$ definiert ist, nennt man eine *quadratische Form* bzw. die zu s gehörige quadratische Form.

Beispiel 29.12. Sei $V = \mathbb{R}^2$. Sei A die symmetrische Matrix $A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \in \mathrm{M}(2; \mathbb{R})$. Es ist $\chi_A = (T-\alpha)(T-\alpha) - \beta^2 = T^2 - 2\alpha T + (\alpha^2 - \beta^2)$, und es ergeben sich die Eigenwerte $\lambda_1 = \alpha + \beta$ und $\lambda_2 = \alpha - \beta$. Normierte Eigenvektoren dazu sind

$$e_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \text{bzw.} \quad e_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}.$$

Sei s die zu A gehörige symmetrische Bilinearform. Die zu s gehörige quadratische Form ist gegeben durch

$$q(x) = q(x_1, x_2) = \alpha x_1^2 + 2\beta x_1 x_2 + \alpha x_2^2.$$

Bzgl. der neuen Koordinaten e_1, e_2 hat sie die Form

$$q(z) = q(z_1, z_2) = (\alpha + \beta)z_1^2 + (\alpha - \beta)z_2^2.$$

Es sind also die gemischten Terme $x_1 x_2$ verschwunden.

Betrachte die Kurve $C = \{x \in \mathbb{R}^2 \mid q(x) = 1\}$. Es gelte $\lambda_1 > 0$ und $\lambda_2 \neq 0$. Setzt man $a = 1/\sqrt{\lambda_1}$ und $b = 1/\sqrt{|\lambda_2|}$, so lautet die Gleichung in den z-Koordinaten:

$$\frac{z_1^2}{a^2} \pm \frac{z_2^2}{b^2} = 1.$$

Im Fall $+$ beschreibt dies eine Ellipse, im Fall $-$ eine Hyperbel. Es sind a und b die sog. *Hauptachsen.*

ZEICHNUNG

Bemerkung 29.13. (1) Eine Abbildung $f\colon K^n \to K^m$ heißt *affin*, wenn es ein $b \in K^m$ und ein $A \in \mathrm{M}(m, n; K)$ gibt mit

$$f(x) = Ax + b \quad \text{für alle } x \in K^n.$$

Es ist dabei f genau dann bijektiv, wenn $(m = n)$ und $A \in \mathrm{GL}(n; K)$ gilt. In dem Fall gilt $f^{-1}(y) = A^{-1}y - A^{-1}b$ für alle $y \in K^n$. Man nennt dann f ein *Affinität.*

29.14 (Quadriken). Wir behandeln hier beispielhaft nur den Fall $n = 2$. Sei $V = \mathbb{R}^2$, sei

$$P(x, y) = \alpha_{11}x^2 + \alpha_{12}xy + \alpha_{22}y^2 + \alpha_{01}x + \alpha_{02}y + \alpha_{00}$$

mit reellen Koeffizienten. Die Nullstellenmenge

$$C = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid P(x, y) = 0 \right\}$$

nennt man eine *(ebene) Quadrik* (eben wegen $n = 2$). Wir können dabei das *Polynom* ohne Einschränkung in der Form

$$P(x, y) = \alpha_{11}x^2 + 2\alpha_{12}xy + \alpha_{22}y^2 + 2\alpha_{01}x + 2\alpha_{02}y + \alpha_{00}$$

schreiben. Setzt man noch

$$\alpha_{10} \overset{def}{=} \alpha_{01},\ \alpha_{21} \overset{def}{=} \alpha_{12} \text{ und } \alpha_{20} \overset{def}{=} \alpha_{02},$$

so ist

$$A \overset{def}{=} (\alpha_{ij}) \in \mathrm{M}(3; \mathbb{R})$$

eine symmetrische Matrix. Sei $v = \begin{pmatrix} x \\ y \end{pmatrix}$. Man erweitert v zu

$$\widetilde{v} \overset{def}{=} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} \in \mathbb{R}^3.$$

Damit gilt offenbar

$${}^t\widetilde{v} A \widetilde{v} = (1, x, y) \cdot A \cdot \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = P(x, y).$$

Behauptung: *Ist $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ eine Affinität, so ist $f(C)$ auch eine Quadrik.*

Denn wird f durch $w = f(v) = Pv + b$ beschrieben ($P \in \mathrm{GL}(2; \mathbb{R})$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$), und sind $\widetilde{v}$, $\widetilde{w} \in \mathbb{R}^3$ wie oben beschrieben die (um 1) erweiterten Vektoren und

$$\widetilde{P} \overset{def}{=} \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline b_1 & & \\ b_2 & & P \end{array}\right) \in \mathrm{GL}(3; \mathbb{R}),$$

so lässt sich f beschreiben durch $\widetilde{w} = \widetilde{P}\widetilde{v}$ und für $Q = P^{-1}$ ist

$$\widetilde{Q} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline -Qb_1 & & \\ -Qb_2 & & Q \end{array}\right) \in \mathrm{GL}(3;\mathbb{R});$$

es folgt $\widetilde{v} = \widetilde{Q}\widetilde{w}$ und damit

$$w = f(v) \in f(C) \;\Leftrightarrow\; v \in C \;\Leftrightarrow\; 0 = {}^t\widetilde{v}A\widetilde{v} = {}^t\widetilde{w}({}^t\widetilde{Q}A\widetilde{Q})\widetilde{w}.$$

Es wird also die Quadrik $f(C)$ beschrieben (in den w-Koordinaten) durch die quadratische Gleichung ${}^t\widetilde{w}({}^t\widetilde{Q}A\widetilde{Q})\widetilde{w} = 0$. —

Wir klassifizieren die Quadriken nun bis auf eine Affinität, durch die wir die Quadrik durch eine einfache Gleichung beschreiben wollen. Es ist

$$A = \left(\begin{array}{c|cc} \alpha_{00} & \alpha_{01} & \alpha_{02} \\ \hline \alpha_{01} & & \\ \alpha_{02} & & A' \end{array}\right),$$

wobei $A' \in \mathrm{M}(2;\mathbb{R})$ symmetrisch ist. Es gibt, wie im Beispiel zuvor, ein $S' \in \mathrm{O}(2)$ mit ${}^tS'A'S' = \mathrm{diag}(\lambda_1, \lambda_2)$. Setzt man

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

dann ist $S \in \mathrm{O}(3)$ und es gilt

$$B \overset{def}{=} {}^tSAS = \left(\begin{array}{c|cc} \beta_{00} & \beta_{01} & \beta_{02} \\ \hline \beta_{01} & \lambda_1 & 0 \\ \beta_{02} & 0 & \lambda_2 \end{array}\right),$$

was bedeutet, dass nun der "gemischte" Term xy verschwunden ist. Wir wollen nun noch die linearen Terme vereinfachen. Sicherlich ist der Fall $\lambda_1 = \lambda_2$ nicht von Interesse, denn in dem Fall tauchen keine quadratischen Terme auf.

<u>1. Fall</u>: $\lambda_1, \lambda_2 \neq 0$. Man kann dann ohne Einschränkung $\lambda_1 > 0$ annehmen. Ferner könnte man statt $S \in \mathrm{O}(3)$ wie oben $S \in \mathrm{GL}(3;\mathbb{R})$ wählen, so dass sogar

$$B \overset{def}{=} {}^tSAS = \left(\begin{array}{c|cc} \beta_{00} & \beta_{01} & \beta_{02} \\ \hline \beta_{01} & 1 & 0 \\ \beta_{02} & 0 & \pm 1 \end{array}\right)$$

gilt. Man bildet nun $P \in \mathrm{GL}(3;\mathbb{R})$ analog zu in (29.1):

$$P \overset{def}{=} \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline -\beta_{01} & 1 & 0 \\ \mp\beta_{02} & 0 & 1 \end{array}\right).$$

Es folgt dann

$${}^tPBP = \left(\begin{array}{c|cc} \gamma_{00} & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & \pm 1 \end{array}\right).$$

Ist dabei $\gamma_{00} = 0$ so erhalten wir die Quadriken

$$x^2 + y^2 = 0 \quad \text{ein Punkt}$$

bzw.

$$x^2 - y^2 = 0 \quad \text{zwei sich schneidende Geraden.}$$

Ist $\gamma_{00} \neq 0$, so kann man (nach evtl. Multiplikation der gesamten Gleichung mit -1) annehmen, dass $\gamma_{00} < 0$ gilt. Eine weitere Transformation mit der invertierbaren Matrix

$$Q \overset{def}{=} \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \sqrt{-\gamma_{00}} & 0 \\ 0 & 0 & \sqrt{-\gamma_{00}} \end{array}\right)$$

und anschließender Division der gesamten Gleichung durch γ_{00} führt dann zu den Quadriken

$$x^2 + y^2 = 1 \qquad \text{Ellipse},$$

$$x^2 - y^2 = 1 \qquad \text{Hyperbel},$$

man muss aber auch den Fall

$$x^2 + y^2 + 1 = 0 \qquad \text{leere Quadrik}$$

zulassen (da obige Bedingung, dass eines der $\lambda_i > 0$ ist, verletzt worden sein kann).

2. Fall: $\lambda_1 \neq 0$, $\lambda_2 = 0$. Wie oben erhält man $S \in \mathrm{GL}(3;\mathbb{R})$ mit

$$B \overset{def}{=} {}^tSAS = \left(\begin{array}{c|cc} \beta_{00} & \beta_{01} & \beta_{02} \\ \hline \beta_{01} & \pm 1 & 0 \\ \beta_{02} & 0 & 0 \end{array}\right)$$

gilt. Man bildet nun

$$P \overset{def}{=} \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline \mp\beta_{01} & 1 & 0 \\ 0 & 0 & 1 \end{array}\right).$$

Es folgt dann

$${}^tPBP = \left(\begin{array}{c|cc} \gamma_{00} & 0 & \gamma_{01} \\ \hline 0 & \pm 1 & 0 \\ \gamma_{01} & 0 & 0 \end{array}\right),$$

also eine Gleichung

$$\pm x^2 + 2\gamma_{01} y + \gamma_{00} = 0.$$

Eine weitere Analyse liefert dann nach evtl. weiterer Transformation die möglichen Fälle

$$\begin{array}{rl} x^2 = 2ay \ (a > 0) & \text{Parabel} \\ x^2 = a^2 (a \neq 0) & \text{Paar paralleler Geraden} \\ x^2 = -a^2 (a \neq 0) & \text{leere Quadrik} \\ x^2 = 0 & \text{Doppelgerade} \end{array}$$

Fazit: Jede ebene Quadrik ist ein *Kegelschnitt*.

ZEICHNUNG, Begründung...

30. Normale Endomorphismen

Es gelten die Voraussetzungen wie in den Abschnitten über Isometrien und selbstadjungierten Endomorphismen.

Lemma 30.1. *Sei $f \in \mathrm{End}_K(V)$. Dann gilt*

$$\chi_{f^*} = \overline{\chi_f},$$

d. h. die Koeffizienten von χ_f werden konjugiert. Insbesondere: Ist $\lambda \in K$ Eigenwert von f, so ist $\overline{\lambda}$ Eigenwert von f^*.

BEWEIS. Sei A Darstellungsmatrix von f bzgl. einer ONB. Dann ist ${}^t\overline{A}$ Darstellungsmatrix von f^* bzgl. derselben Basis. Es folgt

$$\begin{aligned}\chi_{f^*} &= \det(TE_n - {}^t\overline{A}) = \det(TE_n - \overline{A}) \\ &= \overline{\chi_f}.\end{aligned}$$

□

DEFINITION 30.2. $f \in \mathrm{End}_K(V)$ heißt *normal*, falls $f \circ f^* = f^* \circ f$ gilt.

BEISPIEL 30.3. (1) Jede Isometrie ist normal.
(2) Jeder selbstadjungierte Endomorphismus ist normal.

LEMMA 30.4. *Sei $f \in \mathrm{End}_K(V)$. Folgende Aussagen sind äquivalent:*
(1) *f ist normal.*
(2) *$\langle f(x) \mid f(y)\rangle = \langle f^*(x) \mid f^*(y)\rangle$ für alle x, $y \in V$.*
(3) *$\|f(x)\| = \|f^*(x)\|$ für alle $x \in V$.*

BEWEIS. Als Übung. □

FOLGERUNG 30.5. *Sei $f \in \mathrm{End}_K(V)$ normal. Dann gilt* $\mathrm{Kern}(f^*) = \mathrm{Kern}(f)$ *und* $\mathrm{Bild}(f^*) = \mathrm{Bild}(f)$.

BEWEIS. Es gilt

$$\begin{aligned} x \in \mathrm{Kern}(f^*) &\Leftrightarrow f^*(x) = 0 \Leftrightarrow \|f^*(x)\| = 0 \\ &\Leftrightarrow \|f(x)\| = 0 \Leftrightarrow f(x) = 0 \\ &\Leftrightarrow x \in \mathrm{Kern}(f).\end{aligned}$$

Weiter folgt

$$\mathrm{Bild}(f^*) = \mathrm{Kern}(f)^\perp = \mathrm{Kern}(f^*)^\perp = (\mathrm{Bild}(f)^\perp)^\perp = \mathrm{Bild}(f).$$

□

FOLGERUNG 30.6. *Sei $f \in \mathrm{End}_K(V)$ normal. Dann gilt für jedes $\lambda \in K$ für die Eigenräume*

$$E(f, \lambda) = E(f^*, \overline{\lambda}).$$

BEWEIS. Setze $g \overset{def}{=} f - \lambda \cdot 1_V$. Es ist $g^* = f^* - \overline{\lambda} \cdot 1_V$. Man rechnet leicht nach, dass $g \circ g^* = g^* \circ g$, also auch g normal ist. Es folgt $E(f, \lambda) = \mathrm{Kern}(g) = \mathrm{Kern}(g^*) = E(f^*, \overline{\lambda})$. □

FOLGERUNG 30.7. *Sei $f \in \mathrm{End}_K(V)$ normal. Sind x, $y \in V$ Eigenvektoren zu* verschiedenen *Eigenwerten λ, $\mu \in K$, so gilt $\langle x \mid y\rangle = 0$.*

BEWEIS. Mit der vorigen Folgerung und dem Lemma gilt

$$\begin{aligned}\lambda\overline{\mu}\langle x \mid y\rangle &= \langle \lambda x \mid \mu y\rangle = \langle f(x) \mid f(y)\rangle \\ &= \begin{cases}\langle x \mid f^*f(y)\rangle = \mu\overline{\mu}\langle x \mid y\rangle & \text{und} \\ \langle f^*f(x) \mid y\rangle = \lambda\overline{\lambda}\langle x \mid y\rangle & ,\end{cases}\end{aligned}$$

und aus $\langle x \mid y\rangle \neq 0$ würde folgen $\lambda\overline{\mu} = \mu\overline{\mu} = \lambda\overline{\lambda}$, und damit $\lambda = \mu$. □

SATZ 30.8. *Sei $(V, \langle - \mid - \rangle)$ ein (endlichdimensionaler) unitärer Vektorraum. Sei $f \in \mathrm{End}_{\mathbb{C}}(V)$. Dann sind äquivalent:*
(1) *f ist normal.*
(2) *Es gibt eine ONB von V bestehend aus Eigenvektoren von f.*

Beweis. (2)⇒(1): Sei $e_1, \dots, e_n$ eine ONB bestehend aus Eigenvektoren von f. Seien $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die zugehörigen Eigenwerte. Bzgl. dieser ONB wird f dargestellt durch die Diagonalmatrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ und f^* durch $D^* = \operatorname{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$. Offenbar gilt $D \cdot D^* = D^* \cdot D$, und daher auch $f \circ f^* = f^* \circ f$, d. h. f ist normal.

(1)⇒(2): Man führt einen Induktionsbeweis nach $n = \dim(V)$ wie für unitäre oder selbstadjungierte Endomorphismen: Ist e_1 ein normierte Eigenvektor zum Eigenwert $\lambda_1 \in \mathbb{C}$ (Existenz nach dem Fundamentalsatz der Algebra), so gilt $f^*(e_1) = \overline{\lambda_1} e_1$, also ist $\langle e_1 \rangle$ ein f^*-invarianter Unterraum, und daher ist $U \overset{def}{=} \langle e_1 \rangle^\perp$ ein f-invarianter Unterraum. Da außerdem U ein f^*-invarianter Unterraum und daher $f\,|_U : U \to U$ normal ist, ergibt sich die Behauptung nun per Induktion. □

Satz 30.9. *Sei $(V, \langle - \mid - \rangle)$ ein euklidischer Vektorraum der Dimension n. Sei $f \in \operatorname{End}_{\mathbb{R}}(V)$. Dann sind äquivalent:*

(1) *f ist normal.*
(2) *Es gibt eine ONB $\mathcal{B}$ von V, so dass*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_s & & & \\ & & & \boxed{M_1} & & \\ & & & & \ddots & \\ & & & & & \boxed{M_t} \end{pmatrix}$$

gilt, wobei $n = s + 2t$, $\lambda_1, \dots, \lambda_s \in \mathbb{R}$ und

$$M_i = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix} \in \mathrm{M}(2; \mathbb{R})$$

mit $b_i \neq 0$ gilt $(i = 1, \dots, t)$.

Beweis. Beweis als Übungsaufgabe. (Vgl. die Beweise von Satz 27.16, Lemma 27.17 und Übungsaufgabe XI.4.) □

KAPITEL 8

Dualität

31. Dualräume

Es sei jetzt K wieder ein beliebiger Körper.

DEFINITION 31.1. Sei V ein K-Vektorraum. Dann heißt der Vektorraum

$$V^* \overset{def}{=} \operatorname{Hom}_K(V, K) = \{\varphi \mid \varphi\colon V \to K \text{ linear}\}$$

der *Dualraum* von V. Die Elemente von V^* heißen *Lineare Funktionale* oder *Linearformen* auf V.

BEMERKUNG 31.2. Sie V endlichdimensional. Dann gilt $\dim(V^*) = \dim(V)$. (Vgl. 12.3.)

BEISPIEL 31.3.

(1) Sei $s\colon V \times V \to K$ eine Bilinearform auf dem K-Vektorraum V. Dann ist für jedes $a \in V$ die Abbildung $x \mapsto s(x, a)$ ein lineares Funktional. (Ebenso $x \mapsto s(a, x)$.)

(2) Sei $K = \mathbb{R}$ oder $K = \mathbb{C}$ und V ein endlichdimensionaler Raum mit Skalarprodukt $\langle - \mid - \rangle$. Dann ist jedes lineare Funktional $\varphi \in V^*$ von der Form $\varphi = \langle - \mid a \rangle$ für ein eindeutig bestimmtes $a \in V$. (Dies ist gerade Lemma 26.1.) Es gilt, im Fall $K = \mathbb{R}$, sogar, dass die Abbildung $V \to V^*$, $a \mapsto \langle - \mid a \rangle$ ein Isomorphismus von K-Vektorräumen ist.

(3) Sei $V = C([0, 1]) = \{f\colon [0, 1] \to \mathbb{R} \mid f \text{ stetig}\}$. Dann ist die Abbildung

$$C([0, 1]) \to \mathbb{R},\ f \mapsto \int_0^1 f(x)\, dx$$

ein lineares Funktional.

BEMERKUNG 31.4. Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V. Zu jedem $i \in \{1, \ldots, n\}$ gibt es nach Satz 11.13 genau eine lineare Abbildung $b_i^*\colon V \to K$ mit $b_i^*(b_j) = \delta_{ij}$.

SATZ UND DEFINITION 31.5. *Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V. Dann ist $b_1^*, \ldots, b_n^*$ eine Basis von V^*. Sie heißt die* Dualbasis *zu $\mathcal{B}$.*

BEWEIS. Seien $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum_{i=1}^n \alpha_i b_i^* = 0$. Das bedeutet, es gilt $\sum_{i=1}^n \alpha_i b_i^*(x) = 0$ für alle $x \in V$. Für $x = b_j$ erhält man insbesondere

$$0 = \sum_{i=1}^n \alpha_i b_i^*(b_j) = \alpha_j$$

für $j = 1, \ldots, n$. Also gilt $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Damit sind $b_1^*, \ldots, b_n^*$ linear unabhängig. Wegen $\dim(V^*) = \dim(V) = n$ folgt, dass dies eine Basis ist. Wir zeigen aber die Erzeugungseigenschaft nochmal direkt:

Sei $\varphi \in V^*$. Für jedes $i \in \{1, \ldots, n\}$ sei $\beta_i \stackrel{def}{=} \varphi(b_i) \in K$. Sei $x \in V$, schreibe $x = \sum_{i=1}^n \alpha_i b_i$. Dann gilt

$$\begin{aligned} \varphi(x) &= \varphi(\sum_{i=1}^n \alpha_i b_i) = \sum_{i=1}^n \alpha_i \varphi(b_i) \\ &= \sum_{i=1}^n \alpha_i \beta_i = \sum_{i=1}^n \beta_i \alpha_i \\ &= \sum_{i=1}^n \beta_i b_i^* \Big(\sum_{j=1}^n \alpha_j b_j \Big) \\ &= \sum_{i=1}^n \beta_i b_i^*(x), \end{aligned}$$

und es folgt $\varphi = \sum_{i=1}^n \beta_i b_i^*$. □

Folgerung 31.6. *Zu jeder Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von V gibt es einen Isomorphismus $\psi_{\mathcal{B}} \colon V \to V^*$ mit $\psi_{\mathcal{B}}(b_i) = b_i^*$ $(i = 1, \ldots, n)$.*

Folgerung 31.7. *Sei V endlichdimensional. Zu jedem $v \in V$ mit $v \neq 0$ gibt es ein $\varphi \in V^*$ mit $\varphi(v) \neq 0$.*

Beweis. Sei $0 \neq x = \sum_{i=1}^n \alpha_i b_i$. Sei etwa $\alpha_{i_0} \neq 0$. Dann gilt $b_{i_0}^*(x) = b_{i_0}^*(\sum_{i=1}^n \alpha_i b_i) = \alpha_{i_0} \neq 0$. □

Definition 31.8. Seien $f \colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Definiere $f^* \colon W^* \to V^*$ durch $f^*(\varphi) \stackrel{def}{=} \varphi \circ f$ (für jedes $\varphi \in W^*$). Es heißt f^* die zu f *duale Abbildung*.

Bemerkung 31.9.
(1) Es ist $f^* \colon W^* \to V^*$ wieder linear: Sind φ, $\psi \in W^*$ und $\alpha \in K$, so gilt $f^*(\varphi + \psi) = (\varphi + \psi) \circ f = \varphi \circ f + \psi \circ f = f^*(\varphi) + f^*(\psi)$ und $f^*(\alpha\varphi) = (\alpha\varphi) \circ f = \alpha(\varphi \circ f) = \alpha f^*(\varphi)$.
(2) Es gilt $(1_V)^* = 1_{V^*}$.
(3) Sind $f \colon V \to W$ und $g \colon W \to U$ linear, so gilt
$$(g \circ f)^* = f^* \circ g^*.$$
Denn für jedes $\varphi \in U^*$ gilt $(g \circ f)^*(\varphi) = \varphi \circ (g \circ f) = (\varphi \circ g) \circ f = f^*(\varphi \circ g) = f^*(g^*(\varphi))$.
(4) Sind f, $g \colon V \to W$ linear und $\alpha \in K$, so gilt
$$(f + g)^* = f^* + g^* \quad \text{und} \quad (\alpha f)^* = \alpha f^*.$$
Mit anderen Worten: Die Abbildung
$$\mathrm{Hom}(V, W) \to \mathrm{Hom}(W^*, V^*), \;\; f \mapsto f^*$$
ist linear.
(5) Trotz selber Bezeichnung ist die duale Abbildung von der adjungierten Abbildung eines Endomorphimsus aus dem vorigen Kapitel zu unterscheiden.

Satz 31.10. *Sei $f \colon V \to W$ eine lineare Abbildung. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \ldots, w_m)$ eine Basis von W. Seien $\mathcal{B}^* = (b_1^*, \ldots, b_n^*)$ und $\mathcal{C}^* = (w_1^*, \ldots, w_m^*)$ die zugehörigen Dualbasen von V^* bzw. W^*. Ist $A = M_{\mathcal{B}}^{\mathcal{C}}(f) \in \mathrm{M}(m, n; K)$, so gilt*
$$M_{\mathcal{C}^*}^{\mathcal{B}^*}(f^*) = {}^t A \in \mathrm{M}(n, m; K).$$

Beweis. Sei $A = (\alpha_{ij})$. Es gilt also $f(v_j) = \sum_{i=1}^m \alpha_{ij} w_i$. Es folgt

$$f^*(w_i^*)(v_j) = w_i^* \circ f(v_j) = w_i^*\big(\sum_{k=1}^m \alpha_{kj} w_k\big) = \alpha_{ij}.$$

Ist andererseits $B = (\beta_{ij}) = M_{\mathcal{C}^*}^{\mathcal{B}^*}(f^*)$, so gilt

$$f^*(w_i^*) = \sum_{k=1}^n \beta_{ki} v_k^*,$$

und es folgt

$$f^*(w_i^*)(v_j) = \beta_{ji},$$

insgesamt $\alpha_{ij} = \beta_{ji}$, damit $B = {}^tA$. □

Definition 31.11. Eine Folge von zwei linearen Abbildungen $U \xrightarrow{f} V \xrightarrow{g} W$ heißt *exakt*, wenn $\operatorname{Bild}(f) = \operatorname{Kern}(g)$ gilt.

Bemerkung 31.12. (1) $\operatorname{Bild}(f) \subseteq \operatorname{Kern}(g)$ ist gleichbedeutend mit $g \circ f = 0$.

(2) Sei $f \colon V \to W$ linear. Es ist f injektiv genau dann, wenn die Folge $0 \to V \xrightarrow{f} W$ exakt ist.

(3) Sei $f \colon V \to W$ linear. Es ist f surjektiv genau dann, wenn die Folge $V \xrightarrow{f} W \to 0$ exakt ist.

Satz 31.13. *Sei $U \xrightarrow{f} V \xrightarrow{g} W$ eine exakte Folge linearer Abbildungen. Dann ist die Folge $W^* \xrightarrow{g^*} V^* \xrightarrow{f^*} U^*$ exakt.*

Beweis. Aus $g \circ f = 0$ folgt $0 = 0^* = (g \circ f)^* = f^* \circ g^*$, und daher $\operatorname{Bild}(g^*) \subseteq \operatorname{Kern}(f^*)$.

Sei umgekehrt $\varphi \in \operatorname{Kern}(f^*)$, d. h. mit $\varphi \circ f = 0$, was mit $\operatorname{Bild}(f) \subseteq \operatorname{Kern}(\varphi)$ gleichbedeutend ist, und wegen $\operatorname{Bild}(f) = \operatorname{Kern}(g)$ mit $\operatorname{Kern}(g) \subseteq \operatorname{Kern}(\varphi)$. Zu zeigen ist, dass es ein $\psi \in \operatorname{Hom}(V, K)$ gibt mit $\varphi = g^*(\psi) = \psi \circ g$. Sei $g_1 \colon V \to \operatorname{Bild}(g)$ die Einschränkung von g auf sein Bild. Definiere $\psi_1 \colon \operatorname{Bild}(g_1) \to K$ durch $\psi_1(g(x)) \overset{def}{=} \varphi(x)$. Dies ist wohldefiniert, denn ist $g_1(x) = g_1(x')$, so gilt $x - x' \in \operatorname{Kern}(g_1) = \operatorname{Kern}(g) \subseteq \operatorname{Kern}(\varphi)$, und es folgt $\varphi(x) = \varphi(x')$. — Nach Konstruktion folgt $\psi_1 \circ g_1 = \varphi$. Indem man ψ_1 zu $\psi \in \operatorname{Hom}(W, K)$ auf ganz W beliebig fortsetzt, erhält man $\psi \circ g = \varphi$. □

Folgerung 31.14. *Sei $f \colon V \to W$ linear. Es gilt:*

(1) *f injektiv $\Rightarrow$ f^* surjektiv.*
(2) *f surjektiv $\Rightarrow$ f^* injektiv.*

Beweis. (1) Es gilt

$$\begin{aligned} f \text{ injektiv} \quad &\Leftrightarrow \quad 0 \to V \xrightarrow{f} W \text{ exakt} \\ &\Rightarrow \quad W^* \xrightarrow{f^*} V^* \to 0 \text{ exakt} \\ &\Leftrightarrow \quad f^* \text{ surjektiv.} \end{aligned}$$

(2) geht analog. □

Satz 31.15. *Seien V und W endlichdimensionale K-Vektorräume. Dann ist*

$$\operatorname{Hom}(V, W) \to \operatorname{Hom}(W^*, V^*), \ f \mapsto f^*$$

ein Isomorphismus.

Beweis. Gelte $f^* = 0$. Dann gilt $\varphi \circ f = 0$ für alle $\varphi \in W^*$. Ist $f \neq 0$, so gibt es ein $x \in V$ mit $f(x) \neq 0$. Aber dann gibt es nach Folgerung 31.7 ein $\varphi \in W^*$ mit $\varphi(f(x)) \neq 0$, Widerspruch. Also ist die Abbildung injektiv. Da beide Hom-Räume gleiche Dimension haben, folgt auch die Bijektivität. □

Definition 31.16. Für jeden K-Vektorraum V bezeichne

$$V^{**} \stackrel{def}{=} (V^*)^* = \mathrm{Hom}_K(V^*, K)$$

den Dualraum des Dualraum von V, den *Bidualraum* von V.

Satz 31.17. *Sei V ein K-Vektorraum. Die* Auswertungsabbildung

$$e\colon V \to V^{**},\ x \mapsto \big(\ \varphi \mapsto \varphi(x)\ \big)$$

ist linear und injektiv. Ist V endlichdimensional, so liefert e einen natürlichen Isomorphismus $V \xrightarrow{\sim} V^{**}$.

Beweis. Seien $x, y \in V$ und $\alpha \in K$. Dann gilt für alle $\varphi \in V^*$:

$$\begin{aligned} e(x+y)(\varphi) &= \varphi(x+y) = \varphi(x) + \varphi(y) \\ &= e(x)(\varphi) + e(y)(\varphi) \end{aligned}$$

und

$$\begin{aligned} e(\alpha x)(\varphi) &= \varphi(\alpha x) = \alpha\varphi(x) \\ &= (\alpha e(x))(\varphi), \end{aligned}$$

und es folgt $e(x+y) = e(x) + e(y)$ sowie $e(\alpha x) = \alpha e(x)$.

Sei $x \in \mathrm{Kern}(e)$. Es gilt also $e(x) = 0_{V^*}$, d. h. $\varphi(x) = e(x)(\varphi) = 0$ für alle $\varphi \in V^*$. Aus Folgerung 31.7 ergibt sich $x = 0$. Also ist e injektiv.

Ist V endlichdimensional, so auch V^* (mit $\dim(V^*) = \dim(V)$), und es folgt $\dim(V^{**}) = \dim(V^*) = \dim(V)$. Aus den Folgerungen des Rangsatzes folgt, dass e bijektiv, also ein Isomorphismus ist. $\square$

Literaturverzeichnis

[1] Siegfried Bosch: *Lineare Algebra*, Springer 2008.
[2] Gerd Fischer: *Lineare Algebra*, Vieweg 2000.
[3] Klaus Jänich: *Lineare Algebra* (Mit 110 Testfragen), Springer 2008.
[4] Serge Lang: *Linear Algebra*, Addison-Wesley 1972.
[5] Falko Lorenz: *Lineare Algebra*, Spektrum Akademischer Verlag 2003.