Lineare Algebra I
Herbstwintersemester 2019
Mannheim

Claus Hertling

24.07.2019

Inhaltsverzeichnis

hertling@math.uni-mannheim.de

Vorbemerkungen

Lineare Algebra ist normalerweise eine zweisemestrige Vorlesung. Der kanonische Stoff erfordert 1,5 Semester, im hinteren Teil des 2. Semesters ist oft Freiraum für Vertiefungen.
Da es in dieser Vorlesung Lineare Algebra I im HWS 2019 in Mannheim aber eine ganze Reihe Hörer geben wird, die nur diese Vorlesung besuchen und nicht mehr die LA IIa und erst recht nicht die LA IIb, sind die wichtigsten Themen in diesem Manuskript zur LA I aufgenommen.
Allerdings mußten vor allem bei den Kapiteln 8 und 9 einige Abstriche gemacht werden. In Kapitel 8 fehlt eine Diskussion von nicht diagonalisierbaren Endomorphismen, es fehlen Normalformen, insbesondere die Jordan-Normalform, es fehlen die Minimalpolynome und der Satz von Cayley-Hamilton. Im Kapitel 9 werden im wesentlichen nur Euklidische Vektorräume behandelt. Es fehlen Bilinearformen mit Signatur und/oder Radikal, es fehlen Sesquilinearformen über anderen Körpern. All diese Themen werden in den Vorlesungen LA IIa und LA IIb im FSS 2020 behandelt werden.
Weiter sind im Manuskript Quotientenkonstruktionen bei Gruppen, Ringen und Vektorräumen weitgehend ausgeklammert. Nur $\mathbb{Z}/m\mathbb{Z}$ wird in Kapitel 2 vorgestellt.

Das Manuskript kann im Verlauf der Vorlesung noch kleinere Änderungen erfahren. Vor allem in den letzten Kapiteln kann es je nach Schnelligkeit in der Vorlesung zu Abstrichen (oder -weniger wahrscheinlich- Vertiefungen) kommen.
Das Manuskript ist dicht geschrieben. Es hat nicht den Stil eines Lehrbuchs. Es ersetzt nicht den Besuch der Vorlesung. Es soll das Mitschreiben ersparen, aber nicht das Mitarbeiten. Es wird durch die Vorlesung selbst erst lebendig werden. Der Besuch der Vorlesung wird dringend empfohlen.
Fast alle Lehrbücher der Linearen Algebra decken den kanonischen Stoff ab. Hier ist eine Auswahl von 5 schönen Büchern:

S. Bosch: Lineare Algebra.

G. Fischer: Lineare Algebra.

M. Koecher: Lineare Algebra und analytische Geometrie.

W. Klingenberg: Lineare Algebra und Geometrie.

E. Brieskorn: Lineare Algebra und analytische Geometrie I + II.

Sie sind oft in verschiedenen Auflagen in verschiedenen Jahren erschienen, zum Teil auch bei verschiedenen Verlagen.

0 Einige grundlegende Begriffe und Notationen

Mengen: Georg Cantor (Begründer der Mengenlehre):
"Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen."

$\mathbb{N}$ = Menge der natürlichen Zahlen = $\{1, 2, 3, ...\}$
(in Frankreich ist $\mathbb{N} = \{0, 1, 2, 3, ..., \}$);
$\mathbb{Z}$ = Menge der ganzen Zahlen = $\{0, 1, -1, 2, -2, ...\}$;
$\mathbb{Q}$ = Menge der rationalen Zahlen;
$\mathbb{R}$ = Menge der reellen Zahlen;
$\emptyset$ = die leere Menge;
$\{1\},\ \{1, 2, 3\},\ \{\{1\}, 2, 3\}, ...$

"$a \in M$" heißt: a ist Element der Menge M;
$1 \in \{1, 2, 3\},\ 4 \notin \{1, 2, 3\}$.

Seien M_1 und M_2 zwei Mengen;
$M_1 \cup M_2$ ist die Vereinigungsmenge von M_1 und M_2,
$M_1 \cup M_2 = \{a \mid a \in M_1$ oder $a \in M_2\}$,
"die Menge der a, für die gilt: a in M_1 oder a in M_2";
$M_1 \cap M_2$ ist die Schnittmenge von M_1 und M_2,
$M_1 \cap M_2 = \{a \mid a \in M_1$ und $a \in M_2\}$;
$M_1 - M_2 = M_1 \setminus M_2 = \{a \in M_1 \mid a \notin M_2\}$ = die Differenzmenge
("M_1 ohne M_2");
$M_1 \times M_2 = \{(a, b) \mid a \in M_1, b \in M_2\}$
= Produkt der Mengen M_1 und M_2;
(a, b) "geordnetes Paar" aus a und b;
$\mathbb{R} \times \mathbb{R} = \mathbb{R}^2 = \{(x_1, x_2) \mid x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$ = die reelle Ebene;
sei $n \in \mathbb{N}$, $\mathbb{R}^n = \{(x_1, ..., x_n) \mid x_i \in \mathbb{R}$ für alle $i = 1, ..., n\}$;
$(x_1, ..., x_n)$ "geordnetes n-Tupel".

$\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$; $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\} = \mathbb{R}_{>0} \cup \{0\}$.
$\mathbb{R}_{<0} := \{x \in \mathbb{R} \mid x < 0\}$; $\mathbb{R}_{\leq 0} := \mathbb{R}_{<0} \cup \{0\}$;
analog für $\mathbb{Q}$ und $\mathbb{Z}$;
$\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Zwei Mengen M_1 und M_2 heißen disjunkt, falls $M_1 \cap M_2 = \emptyset$.
Eine Menge M_1 ist die *disjunkte Vereinigung* von zwei Mengen M_2 und M_3 falls $M_1 = M_2 \cup M_3$ und $M_2 \cap M_3 = \emptyset$.
$M_1 \subset M_2$ heißt, daß M_1 eine Teilmenge von M_2 ist, d.h. alle Elemente von M_1 sind auch Elemente von M_2.

Ist M eine Menge mit unendlich vielen Elementen, so ist $|M| = \infty$; hat eine Menge M nur endlich viele Elemente, so ist $|M|$ die Anzahl dieser Elemente. In beiden Fällen heißt $|M|$ die *Ordnung* von M.
Die Potenzmenge $\mathcal{P}(M)$ einer Menge M ist die Menge aller Teilmengen von M. Ist M endlich, so auch $\mathcal{P}(M)$, und dann ist $|\mathcal{P}(M)| = 2^{|M|}$.

Abbildungen: Eine Abbildung f von einer Menge X in eine Menge Y ist eine Vorschrift, die jedem Element von X ein Element von Y zuordnet.
Notation: $f : X \to Y, \quad x \mapsto f(x)$;
hier ist $x \in X$, und $f(x) \in Y$ ist das zugeordnete Element.
X ist der *Definitionsbereich*, und Y ist der *Wertebereich* der Abbildung f.

Ist $f : M_1 \to M_2$ eine Abbildung und $M_3 \subset M_1$, so ist $f(M_3) := \{f(x) \mid x \in M_3\}$ *das Bild von* M_3 *unter* f; es ist $f(M_3) \subset M_2$.

Beispiele:

$$
\begin{aligned}
&f_1 : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2;\\
&f_2 : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^3;\\
&f_3 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad x \mapsto \sqrt{x};\\
&f_4 : \{3,4\} \to \{1\}, \quad 3 \mapsto 1,\ 4 \mapsto 1;\\
&f_5 : \{g \mid g : \mathbb{R} \to \mathbb{R} \text{ Abbildung}\} \to \mathbb{R}, \quad g \mapsto g(7);\\
&f_6 : \mathbb{R} \to \{g \mid g : \mathbb{R} \to \mathbb{R} \text{ Abbildung}\},\\
&\quad x \mapsto (\text{die konstante Abbildung mit Wert } x);\\
&f_7 : \mathbb{R} \to \mathbb{R}_{\geq 0}, \quad x \to x^2;\\
&f_8 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad x \to x^2.
\end{aligned}
$$

Definition 0.1 Eine Abbildung $f : X \to Y$ ist
injektiv, falls aus $x_1, x_2 \in X,\ x_1 \neq x_2$ auch $f(x_1) \neq f(x_2)$ folgt, d.h. falls verschiedene Elemente von X unter f verschiedene Bilder in Y haben;
surjektiv, falls zu jedem $y \in Y$ ein $x \in X$ existiert mit $f(x) = y$, d.h. falls das Bild der Menge X unter f die ganze Menge Y ist;
bijektiv, falls f injektiv und surjektiv ist, d.h. falls zu jedem $y \in Y$ *genau ein* $x \in X$ mit $f(x) = y$ existiert.

Ist $f : X \to Y$ bijektiv, so bezeichnet $f^{-1} : Y \to X$ die Abbildung mit

$$f^{-1}(y) := (\text{das eindeutige } x \text{ mit } f(x) = y).$$

f^{-1} ist die *Umkehrabbildung* von f.

Beispiel	injektiv	surjektiv	bijektiv
f_1	nein	nein	nein
f_2	ja	ja	ja
f_3	ja	ja	ja
f_4	nein	ja	nein
f_5	nein	ja	nein
f_6	ja	nein	nein
f_7	nein	ja	nein
f_8	ja	ja	ja

Definition 0.2 Die *Komposition* zweier Abbildungen $f : X \to Y$ und $g : Y \to Z$ ist die Abbildung $g \circ f : X \to Z$, $x \mapsto g(f(x))$.

Lemma 0.3 *(a) Die Komposition zweier Abbildungen ist* assoziativ*, d.h. wenn $f : X \to Y$, $g : Y \to Z$ und $h : Z \to W$ Abbildungen sind, so ist*

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

(b) Ist $f : X \to Y$ bijektiv und $f^{-1} : Y \to X$ die Umkehrabbildung, so ist

$$f^{-1} \circ f = \mathrm{id}_X : X \to X,\ x \mapsto x,$$

die identische Abbildung auf X, und

$$f \circ f^{-1} = \mathrm{id}_Y : Y \to Y,\ y \mapsto y;$$

und natürlich ist

$$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f.$$

Beweis: (a)

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

(b) Die ersten beiden Formelzeilen folgen aus der Definition von f^{-1}, die dritte ist klar. □

"□" bezeichnet das Ende eines Beweises.

Notationen 0.4 (i) Wegen des Lemmas kann man bei einer Komposition $h \circ g \circ f$ von Abbildungen die Klammern weglassen.
(ii) Im Fall einer Abbildung $f : X \to X$ ist die Komposition von f mit sich selbst $f^2 := f \circ f$. Analog sind dann $f^3 := f \circ f \circ f$ und f^n für $n \in \mathbb{N}$ definiert. Weiter ist $f^{-n} := (f^{-1})^n$. Schließlich setzt man $f^0 := \mathrm{id}$.

1 Gruppen

Eine Verknüpfung $*$ auf einer Menge G ist eine Abbildung

$$* : G \times G \to G, \quad (a, b) \mapsto *(a, b).$$

Wir schreiben $a * b$ statt $*(a, b)$.

Definition 1.1 (a) Eine *Gruppe* ist ein Paar $(G, *)$, wobei G eine Menge und $*$ eine Verknüpfung auf G ist, so daß folgende Eigenschaften erfüllt sind:

(G1) *Assoziativität*: für alle $a, b, c \in G$ gilt

$$a * (b * c) = (a * b) * c.$$

(G2) Existenz eines *neutralen Elements*: es gibt ein $e \in G$ mit

$$a * e = e * a = a \text{ für alle } a \in G.$$

(G3) Existenz von *inversen Elementen*: zu jedem $a \in G$ gibt es ein $a' \in G$ mit

$$a * a' = a' * a = e.$$

(b) Eine Gruppe heißt *abelsch* (oder *kommutativ*), falls zusätzlich gilt:

(G4) *Kommutativität*: für alle $a, b \in G$ gilt $a * b = b * a$.

Beispiele 1.2 (i) $(\mathbb{R}, +)$ ist eine abelsche Gruppe mit $e = 0$, $a' = -a$, ebenso $(\mathbb{Q}, +)$ und $(\mathbb{Z}, +)$.
(ii) $(\mathbb{Q} - \{0\}, \cdot)$ mit $\cdot =$ *Multiplikation* ist eine abelsche Gruppe mit $e = 1$, $a' = a^{-1}$, ebenso $(\mathbb{Q}_{>0}, \cdot)$, $(\mathbb{R} - \{0\}, \cdot)$ und $(\mathbb{R}_{>0}, \cdot)$.
(iii) Dagegen ist $(\mathbb{R}_{\geq 0}, +)$ keine Gruppe: zwar ist die Addition eine Verknüpfung auf $\mathbb{R}_{\geq 0}$ (sie schickt $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ auf $\mathbb{R}_{\geq 0}$, denn $a \geq 0$, $b \geq 0 \Rightarrow a + b \geq 0$), und $e = 0$ ist ein neutrales Element, aber zu $x > 0$ gibt es kein inverses Element in $\mathbb{R}_{\geq 0}$.
(iv) $(\mathbb{Q}, \cdot)$ ist keine Gruppe, denn 0 hat kein inverses Element in $\mathbb{Q}$ bezüglich der Multiplikation.
Ebenso ist $(\mathbb{Z} - \{0\}, \cdot)$ keine Gruppe, denn alle Elemente in $\mathbb{Z} - \{-1, 0, 1\}$ haben keine inversen Elemente in $\mathbb{Z} - \{0\}$ bezüglich der Multiplikation.
(v) Auf $G = \mathbb{R}$ definiere

$$*_{am} : G \times G \to G \quad (x, y) \mapsto \frac{x + y}{2},$$

das arithmetische Mittel.

$(G, *_{am})$ ist keine Gruppe: $*_{am}$ ist nicht assoziativ, z.B.

$$\begin{aligned} (1 *_{am} 1) *_{am} 2 &= \frac{\frac{1+1}{2} + 2}{2} = \frac{3}{2} \\ \neq \quad 1 *_{am} (1 *_{am} 2) &= \frac{1 + \frac{1+2}{2}}{2} = \frac{5}{4}, \end{aligned}$$

und überdies existiert kein neutrales Element: aus $x = x *_{am} e = \frac{x+e}{2}$ würde folgen, daß $e = x$ ist; aber man braucht ein gemeinsames e für alle x.

Lemma 1.3 *Sei $(G, *)$ eine Gruppe.*
(a) Es gibt nur ein neutrales Element.
(b) Es gibt zu jedem $a \in G$ nur ein inverses Element.
(c) Kürzungsregel: erfüllen $a, b, c \in G$ die Gleichung $a * b = a * c$ so ist $b = c$.

Beweis: (a) Sind e und $\widetilde{e}$ neutrale Elemente, so ist $e = e * \widetilde{e} = \widetilde{e}$.
(b) Sind a' und $\widetilde{a}'$ inverse Elemente von a, so ist

$$a' = a' * e = a' * (a * \widetilde{a}') = (a' * a) * \widetilde{a}' = e * \widetilde{a}' = \widetilde{a}'.$$

(c)

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

□

Notationen 1.4 (i) Oft wird bei einer Gruppe $(G, *)$ das Verknüpfungssymbol $*$ weggelassen: man schreibt ab oder $a \cdot b$ statt $a * b$; man sagt, man schreibt die Verknüpfung multiplikativ. Das eindeutige neutrale Element heißt e oder 1_G (oder 1), das eindeutige inverse Element zu a heißt a^{-1}.
Man schreibt $a^2 := a \cdot a$, $a^3 := a \cdot a \cdot a$, $a^n := a \cdot ... \cdot a$ (n Faktoren) bei $n \geq 1$, $a^0 := e$, $a^{-n} := a^{-1} \cdot ... \cdot a^{-1}$ (n Faktoren) bei $n \geq 1$. Dann ist $a^{n_1} a^{n_2} = a^{n_1+n_2}$ für $n_1, n_2 \in \mathbb{Z}$.
(ii) Manchmal, aber nur wenn die Gruppe abelsch ist, schreibt man die Verknüpfung als Addition, also $a+b$ statt $a*b$. Dann heißt das neutrale Element 0, und das inverse Element zu a heißt $-a$. Dann schreibt man

$n \cdot a := a + ... + a$	(n Summanden)	bei $n \geq 1$,
$n \cdot a := 0$		bei $n = 0$,
$n \cdot a := (-a) + ... + (-a)$	(n Summanden)	bei $n \leq -1$,

und $a - b = a + (-b)$.
(iii) Oft ergibt sich aus dem Kontext, welche Verknüpfung gemeint ist. Dann spricht man von der Gruppe G.

Satz 1.5 *Sei X eine nichtleere Menge und*

$$Bij(X, X) := \{f : X \to X \mid f \text{ ist bijektiv}\}.$$

$(\mathrm{Bij}(X, X), \circ)$ *ist eine Gruppe mit $e = \mathrm{id}_X$.*

Beweis: Lemma 0.3. □

Bemerkung 1.6 Sei X eine Menge mit mindestens zwei Elementen (d.h. $|X| \geq 2$). Sei $\mathrm{Abb}(X,X)$ die Menge aller Abbildungen von X nach X. Trotz Lemma 0.3 ist $(\mathrm{Abb}(X,X), \circ)$ keine Gruppe. Denn die Menge $\mathrm{Abb}(X,X) - \mathrm{Bij}(X,X)$ ist nicht leer wegen $|X| \geq 2$; und die Elemente in dieser Menge haben keine inversen Elemente: um das zweite einzusehen, muß man zeigen, daß die Gleichungen

$$f' \circ f = \mathrm{id}_X = f \circ f'$$

implizieren, daß f bijektiv ist. Übung, mit Hilfe von Blatt 1, Aufgabe 2.

Definition 1.7 Im Fall $X = \{1, ..., n\}$ für ein $n \in \mathbb{N}$ heißt $(\mathrm{Bij}(X,X), \circ)$ die *symmetrische Gruppe* S_n. Ihre Elemente heißen *Permutationen.*
Eine Notation für ein Element $\sigma \in S_n$: $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$.
Später wird das Verknüpfungssymbol $\circ$ oft weggelassen.

Lemma 1.8 *Für $n \geq 3$ ist die symmetrische Gruppe S_n nicht abelsch. Die Gruppen S_1 und S_2 sind abelsch.*

Beweis: $S_1 = \{\mathrm{id}_{\{1\}}\}$, $S_2 = \{\mathrm{id}_{\{1,2\}}, \varphi\}$ mit $\varphi : 1 \mapsto 2,\ 2 \mapsto 1$.
Sei $n \geq 3$.

$$\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 3 & 1 & 4 & \dots & n \end{pmatrix}, \qquad \tau := \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 1 & 3 & 4 & \dots & n \end{pmatrix},$$
$$\sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 2 & 1 & 4 & \dots & n \end{pmatrix}, \qquad \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 1 & 3 & 2 & 4 & \dots & n \end{pmatrix}.$$

Es ist $\sigma \circ \tau \neq \tau \circ \sigma$. □

Lemma 1.9 *Die symmetrische Gruppe S_n hat*

$$n! := n(n-1) \cdot ... \cdot 2 \cdot 1 \quad (''n \textit{ Fakultät}'')$$

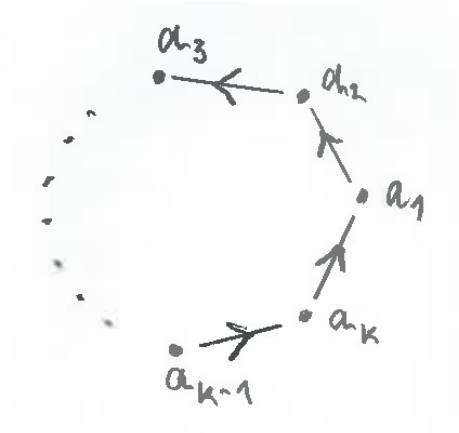
Elemente.

Beweis: Wieviel Freiheit hat man bei der Wahl eines Elementes $\sigma \in S_n$?
Der Wert $\sigma(1)$ kann beliebig in $\{1, 2, ..., n\}$ gewählt werden; also n Möglichkeiten.
Der Wert $\sigma(2)$ kann beliebig in $\{1, 2, ..., n\} - \{\sigma(1)\}$ gewählt werden; also $n-1$ Möglichkeiten.
....
Der Wert $\sigma(n)$ ist das einzige Element von $\{1, 2, ..., n\} - \{\sigma(1), ..., \sigma(n-1)\}$, also 1 Möglichkeit.
Insgesamt hat man $n \cdot (n-1) \cdot ... \cdot 1$ Möglichkeiten. □

Definition 1.10 (a) Eine Permutation $\sigma \in S_n$ heißt *zyklisch*, falls es ein Tupel $(a_1, a_2, ..., a_k)$ gibt mit $2 \leq k \leq n$, $a_1, ..., a_k \in \{1, ..., n\}$, $a_i \neq a_j$ für $i \neq j$, und so, daß

$$\begin{aligned} \sigma(a_i) &= a_{i+1} \text{ für } i = 1, ..., k-1, \\ \sigma(a_k) &= a_1, \\ \sigma(b) &= b \text{ für } b \in \{1, ..., n\} - \{a_1, ..., a_k\} \end{aligned}$$

ist. Notation: Dann schreibt man für σ auch $(a_1 a_2 ... a_k)$.
Bemerkung: Zyklische Permutationen sind die *Bausteine*, aus denen sich alle Permutationen zusammensetzen. Das wird in der großen Übung präzisiert und vertieft.



(b) Eine zyklische Permutation mit $k = 2$ heißt Transposition.

Beispiele 1.11 (i) Im Beweis von Lemma 1.8 waren σ, τ, $\sigma \circ \tau$ und $\tau \circ \sigma \in S_n$ im Fall $n = 3$ zyklisch:

$$\begin{aligned} &\sigma = (123) = (231) = (312), \\ &\tau = (12) = (21) \text{ eine Transposition,} \\ &\text{ebenso } \sigma \circ \tau = (13) = (31),\ \tau \circ \sigma = (23) = (32). \end{aligned}$$

(ii) Wegen $k \geq 2$ in Definition 1.10 ist $e = \mathrm{id} \in S_n$ nicht zyklisch.
(iii) In S_5 ist

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

Vorsicht: Jeden Zykel von links nach rechts durchlaufen, aber die Zykel von rechts nach links abarbeiten (wichtig in Klausuren!).

Definition/Beispiel 1.12 (a) (Definition) Sei $n \geq 2$, $\sigma \in S_n$.

Ein Paar $(i,j) \in \{1,...,n\} \times \{1,...,n\}$ heißt *Fehlstand von* σ, falls $i < j$ und $\sigma(i) > \sigma(j)$ ist.
Das *Signum* von σ ist definiert als

$$\text{sign}(\sigma) := (-1)^{|\{\text{Fehlstände}\}|}.$$

Eine Permutation heißt *gerade*, falls $\text{sign}(\sigma) = +1$ ist, und *ungerade*, falls $\text{sign}(\sigma) = -1$ ist.
(b) (Beispiel)

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} \in S_4$$

hat 5 Fehlstände, also ist $\text{sign}(\sigma) = (-1)^5 = -1$:

(i,j)	$(\sigma(i), \sigma(j))$	Fehlstand
(1,2)	(4,3)	ja
(1,3)	(4,1)	ja
(1,4)	(4,2)	ja
(2,3)	(3,1)	ja
(2,4)	(3,2)	ja
(3,4)	(1,2)	nein

(c) Warnung: Diese Definition ist gut, um Aussagen über das Signum zu beweisen. Zum Ausrechnen in Beispielen sollte man aber nie die Fehlstände bestimmen, sondern immer Eigenschaften in Satz 1.13 benutzen.

Satz 1.13 *Sei* $n \geq 2$. *Es gilt:*

(i) $\text{sign}(\text{id}) = 1$.

(ii) Für $\sigma \in S_n$ *gilt*

$$sign(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{\{a,b\} \subset \{1,..,n\}, a \neq b} \frac{\sigma(a) - \sigma(b)}{a - b}.$$

(iii) $\text{sign}(\widetilde{\sigma} \circ \sigma) = \text{sign}(\widetilde{\sigma}) \cdot \text{sign}(\sigma)$ *für* $\widetilde{\sigma}, \sigma \in S_n$.

(iv) Falls τ *eine Transposition ist, ist* $\text{sign}(\tau) = -1$.

(v) Falls $\tau_1, ..., \tau_k$ *Transpositionen sind, ist* $\text{sign}(\tau_1 \circ ... \circ \tau_k) = (-1)^k$.

(vi) Ein Zykel $(a_1...a_l)$ *erfüllt* $\text{sign}((a_1...a_l)) = (-1)^{l-1}$.

(vii) $k - l$ *ist gerade bei*

$$\tau_1 \circ ... \circ \tau_k = \widetilde{\tau}_1 \circ ... \circ \widetilde{\tau}_l \quad \text{mit } \tau_i, \widetilde{\tau}_j \text{ Transpositionen.}$$

Beweis: (i) gilt, denn id hat keine Fehlstände.
(ii) Wegen

$$\frac{\sigma(b)-\sigma(a)}{b-a} = \frac{\sigma(a)-\sigma(b)}{a-b}$$

kommt es beim zweiten Quotienten nicht auf die Reihenfolge von a und b an, und er ist wohldefiniert. Es reicht, den zweiten Quotienten zu betrachten, denn der erste ist ein Spezialfall des zweiten. Wenn $\{a,b\}$ alle Teilmengen von $\{1,..,n\}$ mit 2 Elementen durchläuft, dann auch $\{\sigma(a),\sigma(b)\}$. Daher ist der Betrag des Produkts aller Zähler gleich dem Betrag des Produkts aller Nenner, und der Quotient ist ± 1. Weil das Vorzeichen des ersten Quotienten gleich $\operatorname{sign}(\sigma)$ ist, sind erster und zweiter Quotient gleich $\operatorname{sign}(\sigma)$.
(iii)

$$\begin{aligned}
\operatorname{sign}(\widetilde{\sigma}\circ\sigma) &= \prod_{\{a,b\}\subset\{1,..,n\},a\neq b} \frac{\widetilde{\sigma}(\sigma(a))-\widetilde{\sigma}(\sigma(b))}{a-b} \\
&= \prod_{\{a,b\}\subset\{1,..,n\},a\neq b} \left(\frac{\widetilde{\sigma}(\sigma(a))-\widetilde{\sigma}(\sigma(b))}{\sigma(a)-\sigma(b)}\cdot\frac{\sigma(a)-\sigma(b)}{a-b}\right) \\
&= \left(\prod_{\{a,b\}\subset\{1,..,n\},a\neq b} \frac{\widetilde{\sigma}(\sigma(a))-\widetilde{\sigma}(\sigma(b))}{\sigma(a)-\sigma(b)}\right)\cdot\left(\prod_{\{a,b\}\subset\{1,..,n\},a\neq b} \frac{\sigma(a)-\sigma(b)}{a-b}\right) \\
&= \operatorname{sign}(\widetilde{\sigma})\cdot\operatorname{sign}(\sigma).
\end{aligned}$$

(iv) Bei der Transposition $\tau = (a\,b)$ mit $1 \le a < b \le n$ ist die Menge der Fehlstände

$$\{(a,j)\,|\,a<j<b\}\cup\{(j,b)\,|\,a<j<b\}\cup\{(a,b)\},$$

also ist ihre Anzahl ungerade.

(v) folgt aus (iii) und (iv).
(vi) folgt aus (v) und $(a_1...a_l) = (a_1a_l)(a_1a_{l-1})...(a_1a_3)(a_1a_2)$.
(vii) folgt aus (v). □

Definition 1.14 Sei $(G,\cdot)$ eine Gruppe und $U \subset G$ eine nichtleere Teilmenge. U heißt Untergruppe von G, falls gilt:

$$\begin{aligned}
a,b\in U &\Rightarrow a\cdot b\in U,\\
a\in U &\Rightarrow a^{-1}\in U.
\end{aligned}$$

Bemerkung 1.15 Dann ist $e \in U$ wegen $a\cdot a^{-1} = e$, und U ist eine Gruppe.

Beispiele 1.16 (i) Für jedes $m \in \mathbb{N}$ sei

$$m\mathbb{Z} := \{mk \mid k\in\mathbb{Z}\} = \{\text{ die durch } m \text{ teilbaren ganzen Zahlen}\}.$$

Folgende Inklusionen geben Untergruppen:

$$(\{0\},+) \subset (m\mathbb{Z},+) \subset (\mathbb{Z},+) \subset (\mathbb{Q},+) \subset (\mathbb{R},+).$$

(ii) Ebenso die Inklusionen

$$\begin{array}{rcl} (\{1\},\cdot) & \subset & (\mathbb{Q}_{>0},\cdot) \subset (\mathbb{Q}-\{0\},\cdot) \subset (\mathbb{R}-\{0\},\cdot) \\ \text{und} & & (\mathbb{Q}_{>0},\cdot) \subset (\mathbb{R}_{>0},\cdot) \subset (\mathbb{R}-\{0\},\cdot). \end{array}$$

(iii) Die Gruppe $S_3 = \{\mathrm{id}, (123), (132), (12), (13), (23)\}$ hat 6 Untergruppen:

$$\begin{array}{l} S_3, \\ A_3 := \{\mathrm{id}, (123), (132)\}, \\ Z_1 := \{\mathrm{id}, (12)\}, \quad Z_2 := \{\mathrm{id}, (13)\}, \quad Z_3 := \{\mathrm{id}, (23)\}, \\ \{\mathrm{id}\}. \end{array}$$

Definition 1.17 Es seien $(G,\cdot)$ und $(H,\cdot)$ Gruppen.
Eine Abbildung $f : G \to H$ heißt *Gruppenhomomorphismus* (oder einfach *Homomorphismus*), falls

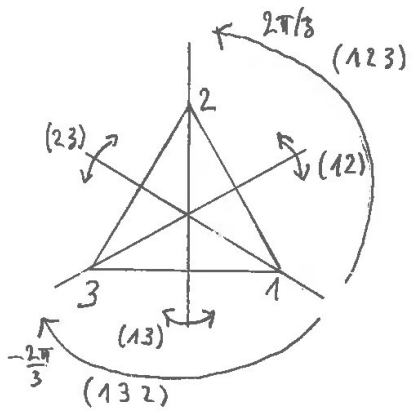
$$f(a \cdot b) = f(a) \cdot f(b) \quad \text{für alle } a, b \in G \text{ gilt.}$$

Falls f darüber hinaus auch bijektiv ist, so heißt f ein *Gruppenisomorphismus*. Dann sind G und H *isomorphe Gruppen*.
Notation: $G \cong H$, "G isomorph H".

Beispiele 1.18 (i) Die Abbildung $\mathrm{sign} : S_n \to \{1,-1\}$, $\sigma \mapsto \mathrm{sign}(\sigma)$, ist ein Gruppenhomomorphismus von S_n in die Gruppe $(\{1,-1\},\cdot)$, wegen Satz 1.13: $\mathrm{sign}(\widetilde{\sigma} \circ \sigma) = \mathrm{sign}(\widetilde{\sigma}) \cdot \mathrm{sign}(\sigma)$.
(ii) Ist U eine Untergruppe einer Gruppe G, so ist die kanonische Inklusion $U \to G$ ein injektiver Gruppenhomomorphismus.
(iii) Die Abbildung

$$\exp : \mathbb{R} \to \mathbb{R}_{>0}, \quad x \mapsto e^x,$$

ist ein Isomorphismus von $(\mathbb{R},+)$ nach $(\mathbb{R}_{>0},\cdot)$, denn $e^{x+y} = e^x \cdot e^y$.
(iv) Die Gruppe S_3 ist isomorph zur Symmetriegruppe eines gleichseitigen Dreiecks: Die Ecken werden mit 1,2,3 bezeichnet, der Mittelpunkt mit 0.



S_3	$\to$	Symmetriegruppe des gleichseitigen Dreiecks
id	$\mapsto$	id
(123)	$\mapsto$	Drehung um $\frac{2\pi}{3}$ mit Fixpunkt 0
(132)	$\mapsto$	Drehung um $\frac{4\pi}{3}$ mit Fixpunkt 0
(12)	$\mapsto$	Spiegelung an der Geraden durch 3 und 0
(13)	$\mapsto$	Spiegelung an der Geraden durch 2 und 0
(23)	$\mapsto$	Spiegelung an der Geraden durch 1 und 0

(v) Zahlreiche Gruppen kann man so interpretieren, als Symmetriegruppen von geometrischen Objekten, oder allgemeiner als Gruppen von Selbstabbildungen (*"Automorphismen"*) von Objekten mit Struktur.

Lemma 1.19 *Sei $f : G \to H$ ein Gruppenhomomorphismus von Gruppen $(G, \cdot)$ und $(H, \cdot)$.*
Dann ist $f(e_G) = e_H$ und $f(a^{-1}) = f(a)^{-1}$ für $a \in G$.
Ist $U \subset G$ eine Untergruppe von G, so ist $f(U) \subset H$ eine Untergruppe von H. Insbesondere ist $f(G)$ eine Untergruppe von H.
Ist $V \subset H$ eine Untergruppe von H, so ist $f^{-1}(V) \subset G$ eine Untergruppe von G. Insbesondere ist die Menge

$$\ker(f) := \{a \in G \mid f(a) = e_H\} = f^{-1}(e_H)$$

eine Untergruppe von G (es gilt mehr: sie ist ein Normalteiler *– Definition nicht hier).*
Ist $f : G \to H$ ein Isomorphismus von Gruppen, so ist auch $f^{-1} : H \to G$ ein Isomorphismus von Gruppen.

Beweis: Die Rechnung $f(e_G) \cdot f(e_G) = f(e_G \cdot e_G) = f(e_G) = f(e_G) \cdot e_H$ und die Kürzungsregel (Lemma 1.3) zeigen $f(e_G) = e_H$.
Aus $f(a^{-1}) \cdot f(a) = f(a^{-1} \cdot a) = f(e_G) = e_H$ und analog $f(a) \cdot f(a^{-1}) = e_H$ folgt $f(a^{-1}) = f(a)^{-1}$.
$f(U)$ Untergruppe von H: zu zeigen ist, daß $f(U)$ abgeschlossen unter dem Produkt und der Inversen-Bildung ist. Das ist es wegen $f(a)^{-1} = f(a^{-1})$ und wegen $f(a) \cdot f(b) = f(a \cdot b)$

$f^{-1}(V)$ Untergruppe von G: zu zeigen ist, daß $f^{-1}(V)$ abgeschlossen unter dem Produkt und der Inversen-Bildung ist. Das ist es wegen

$$\begin{aligned} f(a) \in V, \quad f(b) \in V &\Rightarrow f(ab) = f(a)f(b) \in V, \\ f(a) \in V &\Rightarrow f(a^{-1}) = f(a)^{-1} \in V. \end{aligned}$$

f Isomorphismus $\Rightarrow$ f^{-1} Isomorphismus: sei $f(a) = c, f(b) = d$, also $f^{-1}(c) = a, f^{-1}(d) = b$; es ist

$$f^{-1}(c)f^{-1}(d) = ab = f^{-1}(f(ab)) = f^{-1}(f(a)f(b)) = f^{-1}(cd).$$

□

Beispiel 1.20 Sei $n \geq 2$. Die Teilmenge A_n von S_n,

$$\begin{aligned} A_n &:= \ker(\text{sign} : S_n \to \{1, -1\}) \\ &= \{\sigma \in S_n \mid \text{sign}(\sigma) = 1\} = \{\text{die geraden Permutationen}\}, \end{aligned}$$

ist eine Untergruppe von S_n.
Die Gruppen A_n für $n \geq 2$ heißen *alternierende Gruppen.*

$$S_n = A_n \ \cup \ \{\text{die ungeraden Permutationen}\};$$

die Abbildung $A_n \to \{\text{die ungeraden Permutationen}\},\ a \mapsto (12)a$, ist eine Bijektion; also ist $|A_n| = \frac{n!}{2}$.

Satz 1.21 *(a) Die einzigen Untergruppen von $(\mathbb{Z}, +)$ sind die Untergruppen $m\mathbb{Z}$ für $m \in \mathbb{N}_0$ von Beispiel 1.16 (i).*
(b) Division mit Rest in $\mathbb{Z}$: Zu $a \in \mathbb{Z}$ und $b \in \mathbb{N}$ gibt es eindeutige $q \in \mathbb{Z}$ und $r \in \{0, 1, ..., b-1\}$ mit

$$a = qb + r.$$

Beweis: (b) bekannt oder klar.
(a) Ist $U = \{0\}$, so ist $U = m\mathbb{Z}$ für $m := 0$. Sei nun $U \subset \mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$ mit $U \neq \{0\}$. Sei $m := \min(a \in U \mid a > 0)$. Aus (b) folgt mit $b = m$: zu einem $a \in U$ gibt es eindeutige $q \in \mathbb{Z}$ und $r \in \{0, 1, ..., m-1\}$ mit $a = qm + r$. Es ist $qm = m + ... + m \in U$ und $a \in U$, also auch $r = a - qm \in U$. Aus $0 \leq r < m$ und der Definition von m folgt $r = 0$. Also ist $a = qm \in m\mathbb{Z}$ und $U = m\mathbb{Z}$. □

2 Ringe und Körper

Definition 2.1 (a) Ein *Ring* ist eine Menge R zusammen mit zwei Verknüpfungen, einer Addition $+ : R \times R \to R$ und einer Multiplikation $\cdot : R \times R \to R$ mit folgenden Eigenschaften:

(i) $(R, +)$ ist eine abelsche Gruppe; ihr neutrales Element wird als *Nullelement* oder *Null* bezeichnet und als 0 geschrieben.

(ii) Die Multiplikation ist assoziativ: $(ab)c = a(bc)$ für $a, b, c \in R$.

(iii) Es gelten die *Distributivgesetze:*

$$a(b + c) = ab + ac \quad \text{und} \quad (a + b)c = ac + bc \quad \text{für} \quad a, b, c \in R.$$

(b) Falls ein Element $1_R \in R$ mit $1_R \cdot a = a \cdot 1_R = a$ für alle $a \in R$ existiert, so heißt es *Einselement* oder *Eins*; es wird oft einfach als 1 geschrieben.

(c) Ein Ring heißt *kommutativ*, falls die Multiplikation kommutativ ist, $ab = ba$ für $a, b \in \mathbb{R}$.

(d) Ein *Körper* $(K, +, \cdot)$ ist ein Ring, bei dem $K - \{0\} \neq \emptyset$ ist und $(K - \{0\}, \cdot)$ eine abelsche Gruppe ist.

Bemerkungen 2.2 (i) Ein Ring ist genaugenommen ein Tripel $(R, +, \cdot)$, aber wie bei Gruppen werden wir vom Ring R sprechen, von Elementen und von Teilmengen des Ringes R. Die Menge R wird als primäres Objekt angesehen, die Verknüpfungen darauf als sekundär. Analog bei Körpern.
(ii)

$$\{\text{Körper}\} \subset \{\text{kommutative Ringe mit 1}\} \begin{array}{l} \subset \{\text{komm. Ringe}\} \subset \\ \subset \{\text{Ringe mit 1}\} \subset \end{array} \{\text{Ringe}\}.$$

(iii) Die 1 in einem Ring mit Eins ist eindeutig wegen $1 = 1 \cdot 1' = 1'$.

Beispiele 2.3 (a) $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper, ebenso $(\mathbb{Q}[\sqrt{2}], +, \cdot)$ (Blatt 1, Aufgabe 3: Definition von $\mathbb{Q}[\sqrt{2}]$, und Beweis, dass $(\mathbb{Q}[\sqrt{2}] - \{0\}, \cdot)$ eine abelsche Gruppe ist).

(b) $(\mathbb{C}, +, \cdot)$ ist ein Körper: Satz/Definition 2.12.

(c) $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit 1, aber kein Körper.

(d) Für $m \in \mathbb{N},\ m \geq 2$, ist $(m\mathbb{Z}, +, \cdot)$ ein kommutativer Ring ohne 1.

(e) Satz 2.9: für $m \in \mathbb{N}$ ist $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ ein kommutativer Ring mit 1. Es ist ein Körper genau dann, wenn m eine Primzahl ist.

(f) Zu einem Ring R kann man den Polynomring $R[x]$ definieren. Das wird in der großen Übung ausgeführt. Die Addition und Multiplikation von Polynomen ist bekannt. Beide Verknüpfungen sind assoziativ, es gelten die Distributivgesetze, und die Addition ist kommutativ. Die Multiplikation von Polynomen ist genau dann kommutativ, wenn R ein kommutativer Ring ist.

Das Nullpolynom 0 ist das neutrale Element der Polynomaddition, und zu einem Polynom p ist $-p$ das additive Inverse. Somit ist $R[x]$ bzgl. der Addition eine abelsche Gruppe.

Das Polynom 1 ist ein neutrales Element bzgl. der Polynommultiplikation, und $R[x]$ ist damit ein Ring mit Eins.

$R[x]$ ist nie ein Körper, da das Polynom x kein multiplikatives Inverses haben kann.

(g) Beispiele für nichtkommutative Ringe: später.

(h) $(\{0,\}, +, \cdot)$ ist ein kommutativer Ring mit 1. Er ist der einzige mit $1 = 0$ (Lemma 2.4 (b)).

(i) Für eine Menge M sei $\mathcal{P}(M) := \{ A \mid A \subseteq M \}$ die Potenzmenge von M. Weiter sei für $A, B \in \mathcal{P}(M)$ die symmetrische Differenz definiert durch:

$$A \ominus B := (A \cup B) - (A \cap B).$$

Dann ist $\big(\mathcal{P}(M), \ominus, \cap\big)$ ein kommutativer Ring mit 1.

Lemma 2.4 *Sei R ein Ring.*

(a) Für alle $a \in R$ ist $a \cdot 0 = 0 \cdot a = 0$.

(b) Ist $R \neq \{0\}$ und hat R eine 1, so ist $1 \neq 0$.

(c) $(-a) \cdot b = -(a \cdot b) = a \cdot (-b)$ für $a, b \in R$.

(d) Ist R ein Körper und $a \cdot b = 0$ so ist $a = 0$ oder $b = 0$ (wichtig).

Beweis: (a) $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, also $0 = a \cdot 0$; analog $0 \cdot a = 0$.
(b) Sei $a \in R - \{0\}$. Es ist $a \cdot 0 = 0 \neq a = a \cdot 1$, also $0 \neq 1$.
(c) $a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0$, also $(-a) \cdot b = -(a \cdot b)$; Rest analog.
(d) Ist $a \neq 0$, so ist $b = (a^{-1} \cdot a) \cdot b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0$.

□

Definition 2.5 (a) Eine Teilmenge U eines Ringes R heißt *Unterring*, falls $(U, +)$ eine Untergruppe von $(R, +)$ ist und falls U bezüglich der Multiplikation abgeschlossen ist.

Ein Unterring I von R heißt *Ideal*, falls gilt:

$$x \in R, i \in I \quad \Rightarrow \quad xi \in I \text{ und } ix \in I.$$

(b) Eine Teilmenge U eines Körpers K heißt *Unterkörper*, falls $(U, +)$ eine Untergruppe von $(K, +)$ ist und $(U - \{0\}, \cdot)$ eine Untergruppe von $(K - \{0\}, \cdot)$ ist.
(c) Eine Abbildung $f : R \to S$ von einem Ring R in einen Ring S heißt *Ringhomomorphismus*, falls für $a, b \in R$ sowohl $f(a+b) = f(a)+f(b)$ als auch $f(a \cdot b) = f(a) \cdot f(b)$ gilt. Sie heißt *Ringisomorphismus*, falls sie darüber hinaus auch bijektiv ist.
(d) Ist $f\colon R \to S$ ein Ringhomomorphismus, so ist sein *Kern* definiert durch:

$$\ker(f) := \{\, a \in R \mid f(a) = 0_S \,\}.$$

Bemerkungen 2.6 (i) Einen Ringhomomorphismus $f : K \to L$ zwischen zwei Körpern K und L mit $f(K) \neq \{0\}$ nennt man auch *Körperhomomorphismus*. Analog *Körperisomorphismus*.
(ii) Ein Unterring (Def. 2.5 (a)) ist ein Ring.
(iii) Ein Unterkörper (Def. 2.5 (b)) ist ein Körper.
(iv) Das Bild $f(R) \subset S$ eines Ringhomomorphismus $f : R \to S$ ist ein Unterring von S. Beweis: analog zu Lemma 1.19.
(v) $m\mathbb{Z}$ ist ein Unterring von $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$ und $\mathbb{R}$;
$\mathbb{Z}$ ist ein Unterring von $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$ und $\mathbb{R}$;
$\mathbb{Q}$ ist ein Unterkörper von $\mathbb{Q}[\sqrt{2}]$ und $\mathbb{R}$;
$\mathbb{Q}[\sqrt{2}]$ ist ein Unterkörper von $\mathbb{R}$.

Definition/Lemma 2.7 *(a) (Lemma) Für jeden Körper K hat man einen natürlichen Ringhomomorphismus $\varphi_{\mathbb{Z}} : \mathbb{Z} \to K$, der durch die Abbildung*

$$n \mapsto n \cdot 1_K := 1_K + ... + 1_K \quad (n \text{ Summanden, d.h. } n \text{ mal die } 1_K)$$

gegeben ist.
(b) (Definition) Die Charakteristik $\mathrm{char}(K) \in \mathbb{N}$ *eines Körpers ist*

$$\mathrm{char}(K) := \begin{cases} 0 & \text{falls } n \cdot 1_K \neq 0 \text{ für alle } n \in \mathbb{N} \text{ ist, d.h. falls } \varphi_{\mathbb{Z}} \text{ injektiv ist,} \\ \min(n \in \mathbb{N} \mid n \cdot 1_K = 0) & \text{sonst.} \end{cases}$$

(c) (Lemma) $\mathrm{char}(K)$ *ist entweder 0 oder eine Primzahl.*
(d) (Lemma) $\mathrm{char}(\mathbb{Q}) = \mathrm{char}(\mathbb{Q}[\sqrt{2}]) = \mathrm{char}(\mathbb{R}) = \mathrm{char}(\mathbb{C}) = 0$.

Beweis: (a) Übung. (b) Definition.
(c) Annahme: $\mathrm{char}(K) = a \cdot b$ mit $1 < a, b < \mathrm{char}(K)$.
$0 = \mathrm{char}(K) \cdot 1_K = (ab) \cdot 1_K = (a \cdot 1_K)(b \cdot 1_K)$
$\Rightarrow$ (Lemma 2.4 (d)) $a \cdot 1_K = 0$ oder $b \cdot 1_K = 0$, Widerspruch zur Definition von $\mathrm{char}(K)$.
(d) Klar. $\square$

Bemerkungen 2.8 (i) Wie in den Vorbemerkungen gesagt, werden Quotientenräume erst in der Vorlesung Lineare Algebra IIa behandelt. Nur eine Familie von Quotienten wird hier behandelt, die Quotientenringe $\mathbb{Z}/m\mathbb{Z}$. Denn sie liefern im Fall, wenn m eine Primzahl ist, endliche Körper mit Charakteristik $\neq 0$.

(ii) Die Menge $\mathbb{Z}/m\mathbb{Z}$ wird hier ganz naiv repräsentiert, durch die Menge

$$\mathbb{Z}_m := \{0, 1, ..., m-1\}$$

der ganzen Zahlen von 0 bis $m-1$.

(iii) Für jede ganze Zahl $k \in \mathbb{Z}$ wird der Rest in $\mathbb{Z}_m$ bei Division mit m als $[k]_m$ bezeichnet (vergleiche Satz 1.21 (b)).

Definition/Satz 2.9 *(a) (Definition) Auf $\mathbb{Z}_m$ werden Addition $+_m$ und Multiplikation $\cdot_m$ folgendermaßen definiert.*

$$\begin{aligned} a +_m b &:= [a+b]_m, \\ a \cdot_m b &:= [a \cdot b]_m. \end{aligned}$$

(b) (Satz) Dann ist $(\mathbb{Z}_m, +_m, \cdot_m)$ ein kommutativer Ring mit 1. Das additive Inverse von a ist $[-a]_m$, 0 ist die Null, 1 ist die Eins.

(c) (Satz) Der Ring $\mathbb{Z}_m$ ist genau dann ein Körper, wenn m eine Primzahl ist. Dann ist seine Charakteristik gleich m.

Beweis: (a) Definition.
(b) Man muß zeigen, dass $(\mathbb{Z}_m, +_m)$ eine abelsche Gruppe ist, dass die Multiplikation assoziativ und kommutativ ist und dass die Distributivgesetze erfüllt sind. Das ist alles nicht schwer.
$\mathbb{Z}_m$ *erbt* alle diese Eigenschaften vom kommutativen Ring $\mathbb{Z}$ mit 1.
Die folgenden Rechnungen zeigen die Eigenschaften.

$$\begin{aligned}
(a +_m b) +_m c &= [a+b]_m +_m c = [[a+b]_m + c]_m = [a+b+c]_m \\
&= [a + [b+c]_m]_m = a +_m [b+c]_m = a +_m (b +_m c), \\
a +_m b &= [a+b]_m = [b+a]_m = b +_m a, \\
0 +_m a &= [0+a]_m = [a]_m = a, \\
a +_m [-a]_m &= [a + (-a)]_m = [0]_m = 0, \\
(a \cdot_m b) \cdot_m c &= [a \cdot b]_m \cdot_m c = [[a \cdot b]_m \cdot c]_m = [a \cdot b \cdot c]_m \\
&= [a \cdot [b \cdot c]_m]_m = a \cdot_m [b \cdot c]_m = a \cdot_m (b \cdot_m c), \\
a \cdot_m b &= [a \cdot b]_m = [b \cdot a]_m = b \cdot_m a, \\
1 \cdot_m a &= [1 \cdot a]_m = [a]_m = a, \\
a \cdot_m (b +_m c) &= [a \cdot [b+c]_m]_m = [a(b+c)]_m = [ab + ac]_m \\
&= [ab]_m +_m [ac]_m = a \cdot_m b +_m a \cdot_m c.
\end{aligned}$$

Wegen der Kommutativität der Multiplikation folgt das zweite Distributivgesetz aus dem ersten.
(c) $\Rightarrow$: Wenn $\mathbb{Z}_m$ ein Körper ist, ist offenbar $\operatorname{char} \mathbb{Z}_m = m$. Wegen Lemma 2.7 ist m dann eine Primzahl.
$\Leftarrow$: Diese Richtung ist schwieriger. Zu zeigen ist, dass es zu jeder Zahl $a \in \mathbb{Z}_m - \{0\}$ eine Zahl $b \in \mathbb{Z}_m - \{0\}$ mit $a \cdot_m b = 1$ gibt. Sei $a \in \mathbb{Z}_m - \{0\}$. Man betrachte die Abbildung Multiplikation mit a auf $\mathbb{Z}_m$,

$$\mathbb{Z}_m \to \mathbb{Z}_m,\ b \mapsto a \cdot_m b.$$

Wenn $b \neq 0$ ist, ist wegen folgender Eigenschaft der Primzahl m auch $a \cdot_m b \neq 0$.

Behauptung: Für $c, d \in \mathbb{Z}$ gilt: $m|(c \cdot d) \Rightarrow m|c$ oder $m|d$.

Diese Eigenschaft wird hier als bekannt vorausgesetzt. (Sie ist tatsächlich mit Hilfe der Division mit Rest und des Euklidischen Algorithmus und des Begriffs des ggT nicht so schwer zu beweisen.) Daraus folgt, dass die Multiplikation mit a sich auf eine injektive Abbildung

$$\mathbb{Z}_m - \{0\} \to \mathbb{Z}_m - \{0\},\ b \mapsto a \cdot_m b$$

einschränkt. Denn bei $b_1 \cdot_m a = b_2 \cdot_m a$ ist $(b_1 - b_2) \cdot_m a = 0$. Weil m nicht a teilt, teilt m $b_1 - b_2$. Daraus folgt $b_1 = b_2$.
Eine injektive Abbildung zwischen endlichen Mengen mit gleich vielen Elementen ist bijektiv. Daher ist 1 im Bild. Das Urbild von 1 unter der Abbildung ist das gesuchte Inverse b von a, mit $a \cdot_m b = 1$. □

Beispiele 2.10 (i) In $\mathbb{Z}_5$ hat jedes Element ein multiplikatives Inverses: $1 \cdot_5 1 = 1$, $2 \cdot_5 3 = [6]_5 = 1$, $4 \cdot_5 4 = [16]_5 = 1$.
(ii) In $\mathbb{Z}_6$ ist $2 \cdot_6 3 = [6]_6 = 0$, aber $2 \neq 0$ und $3 \neq 0$; daher ist $\mathbb{Z}_6$ kein Körper. 2 und 3 haben keine multiplikativen Inversen im Ring $\mathbb{Z}_6$.

Satz 2.11 *(nur zur Information; Beweis in einer Algebra-Vorlesung)*
Ist K ein endlicher Körper, so ist seine Ordnung $|K|$ eine Primzahlpotenz, $|K| = p^l$ mit p Primzahl und $l \in \mathbb{N}$.
Zu jeder Primzahlpotenz $q = p^l$ gibt es bis auf Isomorphie genau einen endlichen Körper mit Ordnung q. Er wird $\mathbb{F}_q$ genannt.

Komplexe Zahlen

Vgl. auch Analysis I.

Satz/Definition 2.12 *(a) Die Menge $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ mit den folgendermaßen definierten Verknüpfungen $+$ und $\cdot$ ist ein Körper.*

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2, y_1 + y_2),\\ (x_1, y_1) \cdot (x_2, y_2) &:= (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2).\end{aligned}$$

Seine Elemente heißen komplexe Zahlen. $(0,0) =: 0$ *ist das Nullelement,* $(1,0) =: 1$ *ist das Einselement. Ist* $(x,y) \in \mathbb{C} - \{0\}$*, so ist*

$$(x,y)^{-1} = (\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}).$$

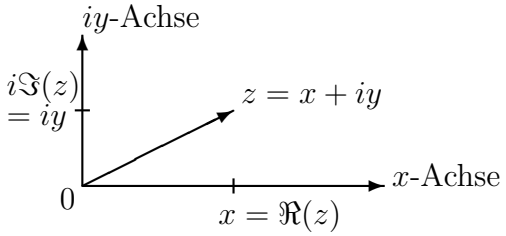
(b) Die Abbildung $\mathbb{R} \to \mathbb{C}$, $x \mapsto (x,0)$*, ist ein Körperhomomorphismus. Mit Hilfe dieser Abbildung wird* $\mathbb{R}$ *mit einem Unterkörper von* $\mathbb{C}$ *identifiziert.*
(c) Das Element $(0,1) =: i$ *erfüllt* $i^2 = (-1,0) = -1$*. Es wird manchmal als* $i = \sqrt{-1}$ *geschrieben.*
(d) Es ist (mit der Identifikation in (c)) für $x, y, x_1, y_1, x_2, y_2 \in \mathbb{R}$

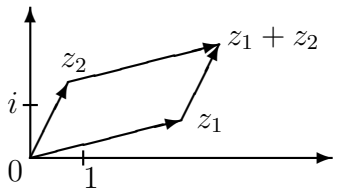
$$\begin{aligned}(x,y) &= (x,0) + (0,y) = (x,0) + (0,1)(y,0) = x + iy \in \mathbb{C},\\ (x_1 + iy_1) + (x_2 + iy_2) &= (x_1 + x_2) + i(y_1 + y_2),\\ (x_1 + iy_1) \cdot (x_2 + iy_2) &= x_1x_2 + x_1 \cdot iy_2 + iy_1 \cdot x_2 + iy_1 \cdot iy_2\\ &= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2).\end{aligned}$$

Man schreibt oft $z = x + iy \in \mathbb{C}$*. Der Realteil von* z *ist* $\Re(z) := x \in \mathbb{R}$*, der Imaginärteil ist* $\Im(z) = y \in \mathbb{R}$*.* z *heißt reell, falls* $\Im(z) = 0$*;* z *heißt rein imaginär, falls* $\Re(z) = 0$*.*

Beweis: (a) $(\mathbb{C}, +)$ abelsche Gruppe: klar.
Die Multiplikation $\cdot$ ist kommutativ und assoziativ, Distributivgesetze: einfache Rechnungen, Übung.
Die Formel für $(x,y)^{-1}$: nachrechnen.
(b) $(0,1) \cdot (0,1) = (-1,0)$.
(c) Klar.
(d) Klar. □

Bemerkung 2.13 Man veranschaulicht sich die komplexen Zahlen in der Gaußschen Zahlenebene. Die Addition ist die komponentenweise Addition im $\mathbb{R}^2$. Multiplikation: siehe Bemerkung 2.15 (iv).





Lemma/Definition 2.14 *(a) Die Abbildung*

$$\bar{\cdot} : \mathbb{C} \to \mathbb{C}, \quad x + iy \mapsto x - iy,$$

ist ein Isomorphismus des Körpers $\mathbb{C}$ *auf sich, also*

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2} \text{ für } z_1, z_2 \in \mathbb{C}.$$

Sie heißt komplexe Konjugation.
(b) Es ist $\overline{z} = z \iff z$ *reell.*
Es ist $\Re(z) = \frac{1}{2}(z + \overline{z})$, $\Im(z) = \frac{1}{2i}(z - \overline{z})$.
Es ist $\overline{\overline{z}} = z$.
Ist $z = x + iy$, *so ist* $z \cdot \overline{z} = x^2 + y^2 \in \mathbb{R}_{\geq 0}$ *und, falls* $z \neq 0$, $z^{-1} = \frac{\overline{z}}{x^2+y^2}$.
(c) Der Absolutbetrag $|z|$ *von* $z = x + iy \in \mathbb{C}$ *ist*

$$|z| := \sqrt{z \cdot \overline{z}} = \sqrt{x^2 + y^2} \in \mathbb{R}_{\geq 0}.$$

Er erfüllt $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$, $\quad |\overline{z}| = |z|$ *und die Dreiecksungleichung:*
$|z_1 + z_2| \leq |z_1| + |z_2|$.

Beweis: (a) Nachrechnen, Übung.
(b) Klar.
(c) Nachrechnen, Übung (vgl. Analysis I). □

Bemerkungen 2.15 (i) Ist $z \in \mathbb{C} - \{0\}$, so ist

$$z = x + iy = |z| \cdot \left(\frac{x}{|z|} + i \frac{y}{|z|} \right) = |z| \cdot (\cos \varphi + i \sin \varphi)$$

mit einem eindeutigen $\varphi \in [0, 2\pi)$, denn $(\frac{x}{|z|})^2 + (\frac{y}{|z|})^2 = 1$. φ heißt das Argument von z. Es ist der Winkel zwischen $x - Achse$ und dem Vektor von 0 nach z in der Gaußschen Zahlenebene.
(ii) In der Analysis wird die Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C} - \{0\}$, $z \mapsto \exp(z) = e^z$ definiert und die *Eulersche Formel*

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad \text{für } \varphi \in \mathbb{R}$$

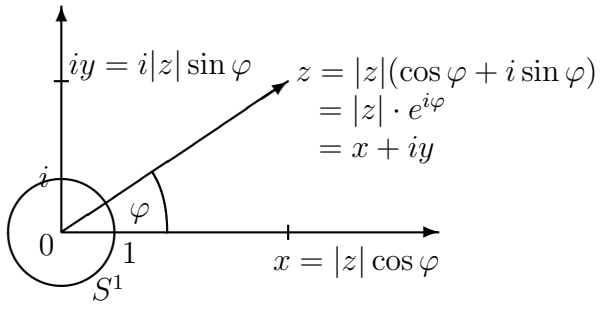
bewiesen. Die Abbildung exp ist ein Gruppenhomomorphismus $\exp : (\mathbb{C}, +) \to (\mathbb{C} - \{0\}, \cdot)$, d.h. sie erfüllt $e^{z_1 + z_2} = e^{z_1} e^{z_2}$.
(iii) Daher ist auch die Abbildung

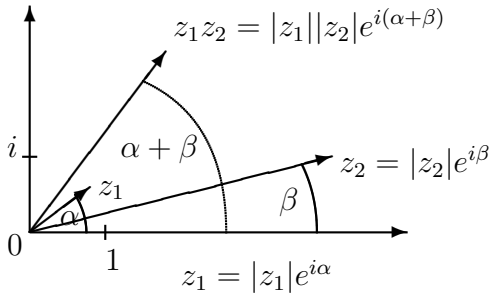
$$\mathbb{R} \to \mathbb{C}, \ \varphi \mapsto e^{i\varphi} = \cos \varphi + i \sin \varphi$$

ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ nach $(\mathbb{C} - \{0\}, \cdot)$. Der Kern ist $2\pi\mathbb{Z} \subset \mathbb{R}$. Das Bild ist die 1-Sphäre $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$. Es ist $(S^1, \cdot)$ eine Untergruppe von $(\mathbb{C} - \{0\}, \cdot)$.
Für $\alpha, \beta \in \mathbb{R}$ erhält man

$$\begin{aligned} & (\cos \alpha + i \sin \alpha) \cdot (\cos \beta + i \sin \beta) \\ = \ & e^{i\alpha} \cdot e^{i\beta} = e^{i(\alpha+\beta)} \\ = \ & \cos(\alpha + \beta) + i \sin(\alpha + \beta). \end{aligned}$$

Die Gleichheit der 1. und 3. Zeile ist äquivalent zu den Additionstheoremen: Gleichheit von Realteil und Imaginärteil sind die Additionstheoreme.
(iv) Sind $z_1 = |z_1|e^{i\alpha}$ und $z_2 = |z_2|e^{i\beta}$, so ist $z_1 z_2 = |z_1||z_2|e^{i(\alpha+\beta)}$, d.h. beim Multiplizieren multipliziert man die Absolutwerte und addiert die Argumente.





3 Vektorräume

Definition 3.1 Sei K ein Körper.
Ein *Vektorraum über K* (oder *K-Vektorraum* oder einfach *Vektorraum*) ist eine abelsche Gruppe $(V, +)$ zusammen mit einer Abbildung

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

mit folgenden Eigenschaften:
(i) ein Distributivgesetz:

$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v \quad \text{für } \lambda, \mu \in K, v \in V;$$

(ii) ein anderes Distributivgesetz:

$$\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w \quad \text{für } \lambda \in K, v, w \in V;$$

(iii) ein Assoziativgesetz:

$$\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v \quad \text{für } \lambda, \mu \in K, v \in V;$$

(iv) das Einselement $1 = 1_K$ von K erfüllt:

$$1 \cdot v = v \quad \text{für } v \in V.$$

Die Abbildung $\cdot : K \times V \to V$ heißt *Multiplikation mit Skalaren* oder *skalare Multiplikation*. Die Elemente des Vektorraums V heißen *Vektoren*.

Bemerkungen 3.2 (i) Wie bei Gruppen, Ringen und Körpern wird die Menge V als das primäre Objekt angesehen, die additive Gruppenstruktur, die skalare Multiplikation und auch der Körper K als sekundär. Daher spricht man vom Vektorraum V und von Elementen des Vektorraums V.
(ii) Die Multiplikation mit Skalaren schreibt man mal mit, mal ohne $\cdot$ (genau wie bei den Multiplikationen in Gruppen, Ringen, Körpern).
(iii) Ersetzt man in Definition 3.1 den Körper K durch einen Ring R mit 1, so heißt V ein *R-Modul*. Die Theorie der R-Moduln ist ein Teil der Algebra.

Beispiele 3.3 (a) Sei K ein Körper und $n \in \mathbb{N}$. Dann ist K^n ein K-Vektorraum mit

$$\begin{aligned} (a_1, ..., a_n) + (b_1, ..., b_n) &:= (a_1 + b_1, ..., a_n + b_n), \\ \lambda \cdot (a_1, ..., a_n) &:= (\lambda a_1, ..., \lambda a_n). \end{aligned}$$

Im Falle $n = 0$ ist $K^0 = \{0\}$. Die Vektorräume K^n, $n \in \mathbb{N}_0$, sind mit Abstand die wichtigsten Vektorräume.

(b) Am allerwichtigsten sind die $\mathbb{R}$-Vektorräume $\mathbb{R}^n$. Bei (fast) allen abstrakten Aussagen über Vektorräume ist es nützlich, an diese Vektorräume zu denken.
(c) Aber es gibt auch andere Vektorräume. Sei $X \neq \emptyset$ eine Menge und K ein Körper. Die Menge $\mathrm{Abb}(X, K)$ ist ein Vektorraum, mit punktweiser Addition und punktweiser skalarer Multiplikation: bei $f, g \in \mathrm{Abb}(X, K)$, $\lambda \in K$ sind $f + g$ und $\lambda \cdot f \in \mathrm{Abb}(X, K)$ definiert durch

$$(f + g)(x) := f(x) + g(x),\ (\lambda \cdot f)(x) := \lambda \cdot f(x) \quad \text{für } x \in X.$$

(d) Die Menge $\mathrm{Abb}([0, 1], \mathbb{R})$ und die Teilmengen

$$\begin{aligned} \mathcal{C}^0([0,1],\mathbb{R}) &:= \{f : [0,1] \to \mathbb{R} \mid f \text{ ist stetig}\}, \\ \mathcal{C}^1([0,1],\mathbb{R}) &:= \{f : [0,1] \to \mathbb{R} \mid f \text{ ist stetig differenzierbar}\} \end{aligned}$$

(Def. von *stetig* und *stetig differenzierbar* in der Analysis) sind $\mathbb{R}$-Vektorräume.
(e) Sei K ein Körper. Der Polynomring $K[t]$ ist auch ein K-Vektorraum.
(f) Ist V ein $\mathbb{R}$-Vektorraum, so ist V mit der Einschränkung der skalaren Multiplikation auf $\mathbb{Q} \times V$ natürlich auch ein $\mathbb{Q}$-Vektorraum.

Lemma 3.4 *Sei V ein K-Vektorraum, $0_K \in K$ die Null in K, $0_V \in V$ die Null in V.*
(a) $0_K \cdot v = 0_V$ *bei* $v \in V$.
(b) $\lambda \cdot 0_V = 0_V$ *bei* $\lambda \in K$.
(c) $\lambda \cdot v = 0_V \Rightarrow \lambda = 0_K$ *oder* $v = 0_V$.
(d) $(-1) \cdot v = -v$ *bei* $v \in V$.

Beweis: (a) $0_k \cdot v = (0_K + 0_K) \cdot v = 0_K \cdot v + 0_K \cdot v$, also $0_V = 0_K \cdot v$.
(b) $\lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V$, also $0_V = \lambda \cdot 0_V$.
(c) Sei $\lambda \cdot v = 0_V$ und $\lambda \neq 0$. Dann ist $v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0_V = 0_V$.
(d) $v + (-1) \cdot v = (1 + (-1)) \cdot v = 0_K \cdot v = 0_V$. □

Von nun an werden die Nullen 0_K und 0_V beide als 0 bezeichnet; die Verwechslungsgefahr ist gering.

Definition/Lemma 3.5 *(a) (Definition) Sei V ein K-Vektorraum. Eine Teilmenge U heißt Untervektorraum, falls $U \neq \emptyset$ ist und falls U abgeschlossen unter der Addition und der skalaren Multiplikation ist, d.h. falls gilt:*

$$\begin{aligned} v \in U, w \in U &\Rightarrow v + w \in U, \\ \lambda \in K, v \in U &\Rightarrow \lambda \cdot v \in U. \end{aligned}$$

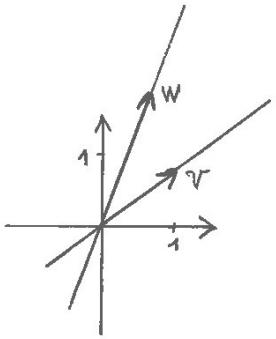
(b) (Lemma) Ein Untervektorraum U eines K-Vektorraums V ist selber ein K-Vektorraum.

Beweis: (a) Definition.
(b) Wegen $U \neq \emptyset$ gibt es ein $v \in U$. Es ist $(-1) \cdot v \in U$ und $0 = v + (-1) \cdot v \in U$. Also ist $(U, +)$ eine abelsche Gruppe. Die Eigenschaften (i) – (iv) von Definition 3.1 gelten in U, weil sie in V gelten. □

Beispiele 3.6 (a) Sei U ein Untervektorraum von $\mathbb{R}^2$ als $\mathbb{R}$-Vektorraum. Ist $v \in U$ und $v \neq 0$, so ist $\{\lambda \cdot v \mid \lambda \in \mathbb{R}\} \subset U$. Ist darüberhinaus $w \in U$ und $w \notin \{\lambda \cdot v \mid \lambda \in \mathbb{R}\}$, so ist

$$U \supset \{\lambda \cdot v + \mu \cdot w \mid \lambda, \mu \in \mathbb{R}\} = \mathbb{R}^2,$$

also $U = \mathbb{R}^2$.



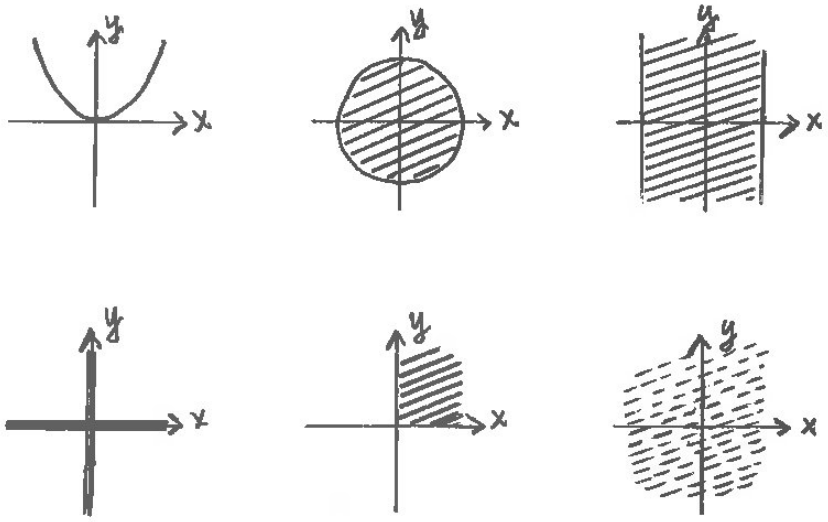
Daher sind die einzigen Untervektorräume von $\mathbb{R}^2$ (als $\mathbb{R}$-Vektorraum) die Mengen

$$\begin{array}{ll} \{0\}, & \\ \{\lambda \cdot v \mid \lambda \in \mathbb{R}\} & \text{mit } v \in \mathbb{R}^2 - \{0\}, \\ \mathbb{R}^2. & \end{array}$$

(b) Daher sind die folgenden Teilmengen alle keine Untervektorräume von $\mathbb{R}^2$ als $\mathbb{R}$-Vektorraum:

$$\begin{array}{l} \{(x,y) \in \mathbb{R}^2 \mid y = x^2\}, \\ \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}, \\ \{(x,y) \in \mathbb{R}^2 \mid |x| \leq 1\}, \\ \{(x,y) \in \mathbb{R}^2 \mid x \cdot y = 0\}, \\ \{(x,y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}. \\ \{(x,y) \in \mathbb{R}^2 \mid (x,y) \in \mathbb{Q}^2\}. \end{array}$$

Man sieht auch direkt, daß sie nicht invariant unter der Addition und/oder der skalaren Multiplikation sind.



(c) $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum, und $\mathbb{R} \subset \mathbb{C}$ ist ein Untervektorraum von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum.
(d) Die $\mathbb{R}$-Vektorräume $\mathcal{C}^0([0,1],\mathbb{R})$ und $\mathcal{C}^1([0,1],\mathbb{R})$ (vgl. Beispiele 3.3 (d)) sind Untervektorräume von $\mathrm{Abb}([0,1],\mathbb{R})$.
(e) Sei K ein Körper. Es gilt (Beweis und Diskussion in der großen Übung):
Ein Polynom $f(t) \in K[t]$ vom Grad n hat höchstens n verschiedene Nullstellen. Konkreter: Seien $\lambda_1, ..., \lambda_k \in K$ verschiedene Nullstellen von $f(t) = a_n t^n + ... + a_1 t + a_0$. Dann ist $k \leq n$, und es gibt $b_{n-k}, ..., b_1, b_0 \in K$ mit

$$f(t) = (t - \lambda_1) \cdot ... \cdot (t - \lambda_k) \cdot (b_{n-k} t^{n-k} + ... + b_1 t + b_0).$$

Falls $|K| = \infty$ ist, so ist daher die Abbildung

$$\Phi : K[t] \to \mathrm{Abb}(K, K), \quad f(t) \mapsto (a \mapsto f(a)),$$

injektiv. Dann kann $K[t]$ mit seinem Bild $\Phi(K[t]) \subset \mathrm{Abb}(K, K)$ identifiziert werden. Der Polynomring wird dann ein Untervektorraum von $\mathrm{Abb}(K, K)$.
Im Fall $|K| < \infty$ ist Φ nicht injektiv: Sei $K = \{\lambda_1, ..., \lambda_k\}$ und $f(t) = (t - \lambda_1) \cdot ... \cdot (t - \lambda_k)$. Dann ist $f(t) \neq 0$, $\mathrm{grad}\, f(t) = k$, aber $\Phi(f(t)) = 0$.
(f) Im Fall $K = \mathbb{R}$ ist auch die Abbildung $\mathbb{R}[t] \to \mathrm{Abb}([0,1],\mathbb{R})$ injektiv (denn $[0,1]$ hat unendlich viele Elemente). Wieder kann man $\mathbb{R}[t]$ mit seinem Bild identifizieren. Dann hat man folgende Kette von Untervektorräumen von $\mathrm{Abb}([0,1],\mathbb{R})$,

$$\mathbb{R}[t] \subset \mathcal{C}^1([0,1],\mathbb{R}) \subset \mathcal{C}^0([0,1],\mathbb{R}) \subset \mathrm{Abb}([0,1],\mathbb{R}).$$

Bemerkungen 3.7 (i) *Vektorraumhomomorphismen* (=*lineare Abbildungen*) werden erst in Kapitel 5 diskutiert.
(ii) Im folgenden werden die Begriffe *Erzeugendensystem*, *Basis* und *Dimension* etabliert.

Die *Dimension* eines Vektorraums V soll definiert werden als die Anzahl der Elemente einer Basis von V (Def. 3.12).
Dazu muß gezeigt werden, daß alle Basen von V gleich viele Elemente haben (Satz 3.17). Es wird u.a. $\dim_K K^n = n$ herauskommen.

Notationen 3.8 (a) Sei X eine nichtleere Menge und $n \in \mathbb{N}$. Die Menge der n-Tupel, $X^n = \{(x_1, ..., x_n) \mid x_i \in X\}$ wird mit der Menge der Abbildungen $\{1, ..., n\} \to X$ identifiziert: zu einem n-Tupel $(x_1, ..., x_n)$ gehört die Abbildung $i \mapsto x_i$.
(b) Sind I und X nichtleere Mengen, so wird eine Abbildung $I \to X, i \mapsto x_i$, auch *Familie* $(x_i)_{i\in I}$ genannt. Die Menge I wird dann *Indexmenge* genannt. Ein n-Tupel $(y_1, ..., y_n)$ ist also eine Familie $(y_j)_{j\in\{1,...,n\}}$.
(c) Sei $(V, +)$ eine abelsche Gruppe und $(v_1, ..., v_n) \in V^n$, $n \in \mathbb{N}$. Dann ist

$$\sum_{i=1}^{n} v_i := v_1 + ... + v_n.$$

Ist allgemeiner $(w_j)_{j\in J}$ eine Familie mit $w_j \in V$ und J eine *endliche* Indexmenge, so ist $\sum_{j\in J} w_j$ die Summe aller Mitglieder w_j der endlichen Familie $(w_j)_{j\in J}$; genauer: man wählt eine Bijektion $\sigma : \{1, ..., |J|\} \to J$ und definiert

$$\sum_{j\in J} w_j := \sum_{i=1}^{|J|} w_{\sigma(i)} = w_{\sigma(1)} + ... + w_{\sigma(|J|)}.$$

Unendliche Summen werden in der Analysis und Funktionalanalysis behandelt, *nicht* in der linearen Algebra. Die folgende Konvention ist nützlich: für $K := \emptyset \subset J$ ist $\sum_{k\in K} w_k := 0$.
(d) Ist $(G, \cdot)$ eine abelsche Gruppe (mit multiplikativ geschriebener Verknüpfung) und $(w_j)_{j\in J}$ eine Familie mit $w_j \in V$ und mit *endlicher* Indexmenge J, so ist (analog zu (c)) $\prod_{j\in J} w_j$ das Produkt aller Mitglieder der Familie.
(e) Ist I eine beliebige (nicht notwendig endliche) Indexmenge und hat man für jedes $i \in I$ eine Menge M_i, so ist ihre Schnittmenge

$$\bigcap_{i\in I} M_i := \{a \mid a \in M_i \text{ für alle } i \in I\}$$

und ihre Vereinigungsmenge

$$\bigcup_{i\in I} M_i := \{a \mid \text{ es gibt ein } i \in I \text{ mit } a \in M_i\}.$$

Definition 3.9 Sei V ein K-Vektorraum.
(a) Sei $(v_1, ..., v_n) \in V^n$, $n \in \mathbb{N}$. Dann ist

$$\text{span}_K(v_1, ..., v_n) := \langle v_1, ..., v_n\rangle_K := \{\sum_{i=1}^{n} \lambda_i v_i \mid \lambda_i \in K\}.$$

Ist allgemeiner I eine beliebige (nicht notwendig endliche) Indexmenge, so ist

$$\mathrm{span}_K(v_i)_{i\in I} := \langle v_i \mid i \in I\rangle_K := \bigcup_{J\subset I, J \text{ endlich}} \mathrm{span}(v_j)_{j\in J}.$$

Ein Element $\sum_{j\in J}\lambda_j v_j$ mit $J \subset I$ endlich heißt *(endliche) Linearkombination* der v_i, $i \in I$. Die Menge $\mathrm{span}_K(v_i)_{i\in I}$ heißt der von $(v_i)_{i\in I}$ *erzeugte Raum.*
(b) Ist $T \subset V$ eine nichtleere Teilmenge, so ist

$$\mathrm{span}_K T := \mathrm{span}_K(t)_{t\in T}.$$

(Hier dient T selbst als Indexmenge: die Elemente von T sind durch sich selbst indiziert.) Die folgende Konvention ist nützlich: $\mathrm{span}_K \emptyset := \{0\} \subset V$.

Beispiele 3.10 (i) Jedes Element des K-Vektorraums K^n ($n \in \mathbb{N}$) ist eine Linearkombination der Vektoren $e_1 := (1, 0, ..., 0)$, $e_2 := (0, 1, 0, ..., 0)$,...,$e_n := (0, .., 0, 1)$; es ist

$$(x_1, ..., x_n) = \sum_{i=1}^{n} x_i e_i, \quad \mathrm{span}_K(e_1, ..., e_n) = K^n.$$

(ii) Jedes Polynom im K-Vektorraum $K[t]$ ist eine Linearkombination der Monome $1, t, t^2, t^3, ...$ Es ist $\mathrm{span}_K(t^i)_{i\in\mathbb{N}_0} = K[t]$.
(iii) Ist T eine beliebige der 6 Teilmengen von $\mathbb{R}^2$ in Beispiel 3.6 (b), so ist $\mathrm{span}_{\mathbb{R}} T = \mathbb{R}^2$.

Lemma 3.11 *Sei V ein K-Vektorraum und $(v_i)_{i\in I}$ eine Familie von Elementen von V. Der erzeugte Raum $\mathrm{span}_K(v_i)_{i\in I}$ ist ein Untervektorraum von V. Er ist der kleinste Untervektorraum, der alle v_i enthält.*

Beweis: Erinnerung an die Definition von span:

$$\begin{aligned}\mathrm{span}_K(v_i)_{i\in I} &= \bigcup_{J\subset I} \mathrm{span}_K(v_j)_{j\in J} \\ &= \{\sum_{j\in J}\lambda_j v_j \mid J \subset I \text{ endlich}, \lambda_j \in K \text{ für } j \in J\}.\end{aligned}$$

Abgeschlossen unter skalarer Multiplikation:

$$\lambda \cdot (\sum_{j\in J}\lambda_j v_j) = \sum_{j\in J}(\lambda \cdot \lambda_j)v_j.$$

Abgeschlossen unter Addition: sind $a = \sum_{j\in J_1}\lambda_j v_j$ und $b = \sum_{j\in J_2}\mu_j v_j$ mit $J_1, J_2 \subset I$ endlich, so kann man definieren

$$\begin{aligned}\lambda_j &:= 0 \text{ für } j \in J_2 - J_1 \text{ und} \\ \mu_j &:= 0 \text{ für } j \in J_1 - J_2;\end{aligned}$$

dann ist

$$a + b = \sum_{j \in J_1 \cup J_2} (\lambda_j + \mu_j) v_j.$$

Jeder Untervektorraum, der alle v_i, $i \in I$, enthält, enthält auch alle Linearkombinationen, denn er ist abgeschlossen unter Addition und skalarer Multiplikation. Also umfaßt er $\mathrm{span}_K(v_i)_{i \in I}$. □

Definition 3.12 Sei V ein K-Vektorraum und $(v_i)_{i \in I}$ eine Familie von Elementen von V.
(a) Die Familie $(v_i)_{i \in I}$ heißt *Erzeugendensystem* von V, falls $V = \mathrm{span}_K(v_i)_{i \in I}$ ist.
(b) Die Familie $(v_i)_{i \in I}$ heißt *linear unabhängig*, falls für jede endliche Teilmenge $J \subset I$ gilt

$$\sum_{j \in J} \lambda_j v_j = 0 \Rightarrow \lambda_j = 0 \text{ für alle } j \in J.$$

Sonst heißt sie *linear abhängig*.
(c) Die Familie $(v_i)_{i \in I}$ heißt *Basis* von V, falls sie ein Erzeugendensystem von V ist und linear unabhängig ist.

Beispiele 3.13 Sei K ein Körper.
(i) Das n-Tupel $(e_1, ..., e_n)$ (vgl. Bsp. 3.10) ist eine Basis des K-Vektorraums K^n $(n \in \mathbb{N})$.
(ii) Die Familie $(t^i)_{i \in \mathbb{N} \cup \{0\}}$ aller Monome t^i ist eine Basis des K-Vektorraums $K[t]$.
(iii) $(1, i)$ ist eine Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$.
(iv) Für jedes $x \in [0, 1]$ wird eine Abbildung $\chi_x : [0, 1] \to K$ definiert durch $\chi_x(x) := 1$ und $\chi_x(y) := 0$, falls $y \in [0, 1] - \{x\}$. Die Familie $(\chi_x)_{x \in [0,1]}$ ist linear unabhängig, aber sie ist keine Basis von $\mathrm{Abb}([0, 1], K)$. Sie ist eine Basis des Unterraums aller Abbildungen, die nur bei endlich vielen Elementen von $[0, 1]$ Werte $\neq 0$ haben.

Satz 3.14 *Sei V ein K-Vektorraum.*
(a) Eine Familie $(v_i)_{i \in I}$ ist zum Beispiel linear abhängig, falls ein $v_i = 0$ ist oder falls $v_i = v_j$ für zwei Indices $i \neq j$ ist. Sie ist linear abhängig genau dann, wenn ein Mitglied v_i Linearkombination der anderen ist.
(b) Sei $(v_i)_{i \in I}$ eine Familie von Vektoren in V. Die folgenden vier Bedingungen sind äquivalent:

(i) Sie ist eine Basis.

(ii) Jedes Element von V läßt sich mit eindeutigen Koeffizienten als Linearkombination der v_i, $i \in I$, schreiben. Die Eindeutigkeit besagt, dass für jede endliche Teilmenge $J \subset I$ gilt

$$\sum_{j \in J} \lambda_j v_j = \sum_{j \in J} \mu_j v_j \Rightarrow \lambda_j = \mu_j \text{ für alle } j \in J.$$

(iii) Sie ist ein minimales Erzeugendensystem*, d.h. sie ist ein Erzeugendensystem, und wenn man ein v_i wegläßt, so ist die Restfamilie $(v_j)_{j\in I-\{i\}}$ nicht mehr ein Erzeugendensystem.*

(iv) Sie ist maximal linear unabhängig*, d.h. sie ist linear unabhängig, und wenn man ein $v_0 \in V$ mit $0 \notin I$ hinzufügt, so ist die erweiterte Familie $(v_j)_{j\in I\cup\{0\}}$ linear abhängig.*

Beweis: (a) Die Beispielfälle sind klar:

$$\text{Für beliebiges } \lambda \in K \text{ ist } \lambda \cdot v_i = 0 \text{ falls } v_i = 0.$$
$$\text{Für beliebiges } \lambda \in K \text{ ist } \lambda \cdot v_i + (-\lambda) \cdot v_j = 0, \text{ falls } v_i = v_j.$$

Zur "genau dann wenn"-Aussage: "⇒": Ist $\sum_{j\in J} \lambda_j v_j = 0$ und $\lambda_k \neq 0$ für ein $k \in J$, so ist $v_k = \sum_{j\in J-\{k\}} (-\frac{\lambda_j}{\lambda_k}) v_j$.
"⇐": Ist $v_k = \sum_{j\in J-\{k\}} \mu_j v_j$, so ist $0 = (-1)v_k + \sum_{j\in J-\{k\}} \mu_j v_j$.
(b) "(i) ⇒ (ii)": Basis ⇒ Erzeugendensystem ⇒ Jedes Element ist Linearkombination der v_i, $i \in I$.
Eindeutigkeit der Koeffizienten: Man wendet die lineare Unabhängigkeit (d.h. Def. 3.12 (b)) an auf die Differenz von rechter und linker Seite in (ii); man erhält $\lambda_j - \mu_j = 0$.
"(ii) ⇒ (i)": (i) ist der Spezialfall von (ii) mit $\mu_j = 0$ für alle j.
"(i) ⇒ (iii)": Basis ⇒ Erzeugendensystem.
Zu zeigen bleibt, daß es minimal ist. Indirekter Beweis. Annahme: für ein geeignetes $i \in I$ ist $(v_j)_{j\in I-\{i\}}$ immer noch ein Erzeugendensystem von V.
Dann ist v_i Linearkombination der anderen Mitglieder. Nach (a) ist die Familie $(v_j)_{j\in I}$ linear abhängig, also keine Basis.
"(iii) ⇒ (i)": Indirekter Beweis. Annahme: die Familie ist linear abhängig.
Nach (a) gibt es ein v_i, das Linearkombination der anderen Mitglieder ist. Dies v_i kann man weglassen; die Familie $(v_j)_{j\in I-\{i\}}$ ist immer noch ein Erzeugendensystem von V.
"(i) ⇒ (iv)": Basis ⇒ Linear unabhängig.
Zu zeigen bleibt, daß die Familie maximal linear unabhängig ist. Sei $v_0 \in V$ und $0 \notin I$. Basis ⇒ v_0 ist Linearkombination der v_i, $i \in I$. Mit (a) folgt, daß $(v_j)_{j\in I\cup\{0\}}$ linear abhängig ist.
"(iv) ⇒ (i)": $(v_i)_{i\in I}$ ist linear unabhängig laut (iv). Zu zeigen bleibt, daß es ein Erzeugendensystem ist. Sei $v \in V$ beliebig.
Sei $0 \notin I$ und sei $v_0 := v$. Laut (iv) ist $(v_j)_{j\in I\cup\{0\}}$ linear abhängig. Also gibt es eine endliche Menge $J \subset I \cup \{0\}$ und Koeffizienten $\lambda_j \in K$ für $j \in J$, die nicht alle 0 sind und die $0 = \sum_{j\in J} \lambda_j \cdot v_j$ erfüllen. Weil $(v_i)_{i\in I}$ linear unabhängig ist, ist $0 \in J$, *und* es ist $\lambda_0 \neq 0$. Daher ist $v_0 = \sum_{j\in J-\{0\}} \frac{-\lambda_j}{\lambda_0} v_j$. Weil v_0 beliebig war, ist $(v_i)_{i\in I}$ ein Erzeugendensystem. □

Satz 3.15 *(a) Hat ein Vektorraum V ein endliches Erzeugendensystem, so erhält man durch Weglassen geeigneter Mitglieder dieser Familie eine Basis von V. Insbesondere hat ein Vektorraum mit endlichem Erzeugendensystem eine endliche Basis.*
(b) (Verallgemeinerung von (a), ohne Beweis) Jeder Vektorraum hat eine Basis.

Beweis von a): Man läßt so lange Mitglieder des endlichen Erzeugendensystems weg, bis man im Fall (iii) von Satz 3.14 (b) landet. Dann hat man eine Basis. □

Bemerkungen 3.16 (i) Der Beweis von (b) ist viel schwieriger. Er benutzt nichttriviale Aussagen aus der Mengenlehre, das *Auswahlaxiom* oder das *Zornsche Lemma*. In dieser Vorlesung werden Sie gebeten, den Satz einfach zu akzeptieren.
(ii) Beim $\mathbb{Q}$-Vektorraum $\mathbb{Q}[t]$ kann man eine unendliche Basis angeben, die Familie $(t^i)_{i\in\mathbb{N}\cup\{0\}}$. Die Basis ist *abzählbar unendlich*, d.h. es gibt eine Bijektion von $\mathbb{N}$ auf die Indexmenge $\mathbb{N}\cup\{0\}$.
(iii) Nach Satz 3.15 (b) hat auch der $\mathbb{Q}$-Vektorraum $\mathbb{R}$ eine Basis $(v_i)_{i\in I}$. Aber hier ist die Basis bzw. die Indexmenge I *überabzählbar*, d.h. sie ist unendlich und es gibt keine Bijektion $\mathbb{N}\to I$ (Beweis: eventuell Übung). Es ist unmöglich, die Basis "explizit" anzugeben (das kann man präzisieren).
(iv) Tatsächlich haben die Basen von Vektorräumen ohne endliche oder abzählbar unendliche Erzeugendensysteme wenig Bedeutung. Bei ihnen sind "konvergente Reihen" wichtiger als endliche Linearkombinationen (Definition und Diskussion in Analysis und Funktionalanalysis).
(v) (Zu Satz 3.17) Bei einer Basis $(v_i)_{i\in I}$ eines Vektorraums sind alle Mitglieder verschieden wegen Satz 3.14 (a). Daher ist die Abbildung $I\to\{v_i \mid i\in I\}$, $i\mapsto v_i$, eine Bijektion, und die Unterscheidung zwischen der Familie $(v_i)_{i\in I}$ und der Menge $\{v_i \mid i\in I\}$ ist nicht so wichtig. Man nennt die Mitglieder v_i der Familie auch *Elemente der Basis*.

Satz/Definition 3.17 *(a) (Satz) Hat ein Vektorraum eine endliche Basis, so sind alle Basen endlich und haben gleich viele Elemente.*
(b) (Definition) Die Dimension eines K-Vektorraums V ohne endliche Basis ist ∞. Die Dimension eines K-Vektorraums mit einer endlichen Basis ist die Anzahl der Elemente einer Basis. Notation: $\dim_K V\in\{0\}\cup\mathbb{N}\cup\{\infty\}$.
(c) (Satz) Ist U ein Untervektorraum eines K-Vektorraums V, so ist $\dim U\le\dim V$. *Ist* $\dim V<\infty$, *so ist* $\dim U=\dim V \iff U=V$.
(d) $\dim K^n=n,\quad \dim K[t]=\infty$.

Beweis: nach Satz 3.18

Satz 3.18 *(Austauschsatz von Steinitz) Sei V ein K-Vektorraum, $(v_1,...,v_n)$ eine Basis von V und $(w_1,...,w_k)$ eine linear unabhängige Familie in V.*
Dann ist $k\le n$, und es gibt lauter verschiedene Indices $i_1,...,i_k\in\{1,...,n\}$, so daß man nach Austauschen von $v_{i_1},...,v_{i_k}$ gegen $w_1,...,w_k$ wieder eine Basis von V erhält.

Beweis: Fall $V = \{0\}$: $n = 0$, $k = 0$, leere Aussagen. Es bleibt der Fall $V \neq \{0\}$, $n \geq 1$. Nun Induktion nach k. Induktionsanfang: $k = 0$, trivial.
Induktionsschritt, $k-1 \to k$: Nach Induktionsannahme ist $k-1 \leq n$, und wir können annehmen, daß bei geeigneter Numerierung der v_i $(w_1, ..., w_{k-1}, v_k, ..., v_n)$ eine Basis von V ist.
Daher gibt es $\lambda_j \in K$, $j = 1, ..., n$, mit

$$w_k = \lambda_1 w_1 + ... + \lambda_{k-1} w_{k-1} + \lambda_k v_k + ... + \lambda_n v_n.$$

Behauptung: $k - 1 < n$, und es gibt ein $j \geq k$ mit $\lambda_j \neq 0$.
Andernfalls wäre w_k eine Linearkombination der $w_1, ..., w_{k-1}$. Dann wäre nach Satz 3.14 (a) $(w_1, ..., w_k)$ nicht linear unabhängig. Widerspruch. Also stimmt die Behauptung.
Nach Umnumerieren der $v_k, ..., v_n$ können wir annehmen, daß $\lambda_k \neq 0$ ist. Nun wird v_k gegen w_k ausgetauscht.
Behauptung: $(w_1, ..., w_k, v_{k+1}, ..., v_n)$ ist eine Basis von V.
Daß es ein Erzeugendensystem ist, ist klar: man kann

$$v_k = \frac{1}{\lambda_k} w_k - \frac{\lambda_1}{\lambda_k} w_1 - ... - \frac{\lambda_{k-1}}{\lambda_k} w_{k-1} - \frac{\lambda_{k+1}}{\lambda_k} v_{k+1} - ... - \frac{\lambda_n}{\lambda_k} v_n$$

erzeugen und dann mit v_k alle Elemente von V.
Es ist auch eine Basis: Sei

$$0 = \mu_1 w_1 + ... + \mu_k w_k + \mu_{k+1} v_{k+1} + ... + \mu_n v_n.$$

Zu zeigen ist $\mu_1 = ... = \mu_n = 0$. Die rechte Seite ist gleich zu

$$\begin{aligned}&(\mu_1 + \mu_k \lambda_1) w_1 + ... + (\mu_{k-1} + \mu_k \lambda_{k-1}) w_{k-1} + \mu_k \lambda_k v_k \\ &+(\mu_{k+1} + \mu_k \lambda_{k+1}) v_{k+1} + ... + (\mu_n + \mu_k \lambda_n) v_n.\end{aligned}$$

$(w_1, ..., w_{k-1}, v_k, ..., v_n)$ ist eine Basis. Daher sind alle Koeffizienten hier gleich Null. Insbesondere ist $\mu_k \lambda_k = 0$. Wegen $\lambda_k \neq 0$ ist $\mu_k = 0$. Daher sind alle $\mu_1 = ... = \mu_n = 0$. Das beendet den Beweis der Behauptung und den Induktionsbeweis. □

Beweis von Satz 3.17: (a) Sei $(v_1, ..., v_n)$ eine endliche Basis eines Vektorraums V. Hätte V eine unendliche Basis, so könnte man aus dieser eine linear unabhängige Teilfamilie mit $k > n$ Mitgliedern auswählen. Widerspruch zum Austauschsatz. Also ist jede Basis von V endlich.
Ist $(w_1, ..., w_l)$ eine endliche Basis, so ist sie eine linear unabhängige Familie. Also ist nach dem Austauschsatz $l \leq n$. Genauso folgt $n \leq l$. Also ist $l = n$.
(b) Definition.
(c) Im Fall $\dim_K V = \infty$ ist nichts zu zeigen. Sei $\dim_K V = n \in \mathbb{N}_0$. Wie in a) folgt, daß U keine unendliche Basis hat. Ist $(w_1, ..., w_l)$ eine Basis von U, so ist sie eine linear unabhängige Familie in V. Nach dem Austauschsatz ist $l \leq n$.

Wäre $l = n$, aber $U \neq V$, so gäbe es ein $w_{l+1} \in V - U$. Die Familie $(w_1, ..., w_l, w_{l+1})$ wäre linear unabhängig. Nach dem Austauschsatz wäre $l + 1 \leq n$. Widerspruch.
(d) Mit 3.13 (i) und (ii). □

Manchmal ist folgender Satz nützlich.

Satz 3.19 *(Basisergänzungssatz)*
Sei V ein endlichdimensionaler K-Vektorraum mit $\dim_K V = n$. Sei $(w_1, ..., w_k)$ eine linear unabhängige Familie in V mit $k \leq n$.
Dann gibt es $w_{k+1}, ..., w_n$, so daß $(w_1, ..., w_n)$ eine Basis von V ist.

Beweis: Man wählt irgendeine Basis $(v_1, ..., v_n)$, tauscht nach dem Austauschsatz geeignete Elemente der Basis gegen $w_1, ..., w_k$ aus und benennt die anderen um in $w_{k+1}, ..., w_n$. □

4 Matrizen

In diesem Kapitel bezeichnet K stets irgendeinen Körper.

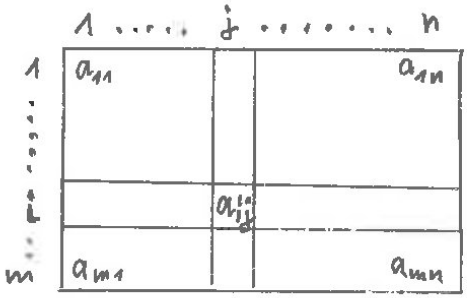
Notation/Definition 4.1 (a) Sei X eine nichtleere Menge und seien $m, n \in \mathbb{N}$. Eine $(m \times n)$-*Matrix A mit Einträgen in X* besteht aus der folgenden Anordnung von $m \cdot n$ Elementen $a_{ij} \in X$, für $i = 1, ..., m$, $j = 1, ..., n$, in einem rechteckigen Schema mit Klammern drumherum:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Eine kürzere Schreibweise ist $A = (a_{ij})_{i=1,...,m;j=1,...,n}$; wenn klar ist, von wo bis wo i und j laufen, schreibt man auch einfach $A = (a_{ij})$. Die *i-te Zeile* von A ist $(a_{i1} \ \cdots \ a_{in})$, die j-te Spalte ist

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Der erste Index des *Koeffizienten* a_{ij} (also i hier) ist der Zeilenindex, der zweite Index (also j hier) ist der Spaltenindex. Der Koeffizient a_{ij} steht in der i-ten Zeile und j-ten Spalte.
Die Einträge einer Matrix A werden auch mit $(A)_{ij}$ bezeichnet; also hier $(A)_{ij} = a_{ij}$.



(b) Die Menge aller $(m \times n)$-Matrizen mit Koeffizienten in X heißt $M(m \times n, X)$. Ist $X = K$ ein Körper, so ist sie (natürlich) ein K-Vektorraum der Dimension $m \cdot n$. Die Elemente von $M(m \times 1, K)$ heißen *Spaltenvektoren*. Die Elemente von $M(1 \times n, K)$ heißen *Zeilenvektoren*.

(c) Der Vektorraum $M(1 \times n, K)$ der Zeilenvektoren $(a_1 \cdots a_n)$ wird mit dem Vektorraum K^n der n-Tupel $(a_1, ..., a_n)$ identifiziert.

Vorsicht: bei Tupeln stehen Kommata zwischen den Einträgen, bei Zeilenvektoren eigentlich nicht. Wir benutzen im Text überwiegend die Notation mit Kommata, in Formeln immer die Notation ohne Kommata.

d) Die Zeilen $v_i := (a_{i1}, ..., a_{in})$ einer $(m \times n)$-Matrix $A = (a_{ij})$ erzeugen einen Untervektorraum $\text{span}(v_i)_{i\in\{1,...,m\}}$ des K-Vektorraums $K^n = M(1 \times n, K)$. Der *Zeilenrang* von A ist

$$\text{Zeilenrang}(A) := \dim_K \text{span}(v_i)_{i\in\{1,...,m\}}.$$

Analog ist der *Spaltenrang* von A

$$\text{Spaltenrang}(A) := \dim_K \text{span}_K(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, ..., \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix})$$

die Dimension des von den Spalten von A erzeugten Untervektorraums von $M(m \times 1, K)$.

(e) Eine $(m \times n)$-Matrix läßt sich verschieden interpretieren:
1.: Eine Liste von Elementen des Vektorraums $K^n = M(1\times n, K)$ von Zeilenvektoren (untereinandergeschrieben).
2.: Eine Liste von Elementen des Vektorraums $M(m \times 1, K)$ von Spaltenvektoren (nebeneinandergeschrieben).
3., in Kapitel 5: Eine *lineare Abbildung* von $M(n \times 1, K)$ nach $M(m \times 1, K)$.
4., in Kapitel 9 und LA II: Eine *Bilinearform* auf $M(m \times 1, K) \times M(n \times 1, K)$.

Satz 4.2 *Sei $A \in M(m \times n, K)$. Es ist*

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A).$$

Beweis: in Kapitel 5.

Bemerkungen 4.3 (i) In der Literatur wird auch öfters $M(m \times 1, K)$ mit K^m identifiziert, aber nicht in dieser Vorlesung.
(ii) Wenn eine Matrix gegeben ist, möchte man den Zeilenrang bestimmen und ein besonders übersichtliches Erzeugendensystem des Vektorraums finden, der von den Zeilen erzeugt wird.
Im folgenden wird dazu ein Algorithmus angegeben, der in mehreren Schritten neue Matrizen konstruiert, deren Zeilen Linearkombinationen der Ausgangsmatrix sind, den gleichen Vektorraum erzeugen und einfacher aussehen.

Definition/Lemma 4.4 *(a)* (Definition) Sei $\lambda \in K - \{0\}$, $i, j \in \{1, ..., m\}$, $i \neq j$. Die folgenden Abbildungen $Z_I(\lambda, i)$, $Z_{II}(\lambda; i, j)$ und $Z_{III}(i, j)$ von $M(m \times n, K)$ auf $M(m \times n, K)$ heißen *elementare Zeilenumformungen.* Die i-te Zeile einer Matrix $A = (a_{ij}) \in M(m \times n, K)$ wird $v_i := (a_{i1}, ..., a_{in})$ genannt.

$Z_I(\lambda; i)$ ersetzt die i-te Zeile v_i durch $\lambda \cdot v_i$.

$Z_{II}(\lambda; i, j)$ ersetzt die j-te Zeile v_j durch $v_j + \lambda \cdot v_i$.

$Z_{III}(i, j)$ vertauscht die i-te und j-te Zeile.

(b) (Lemma) Die Zeilen einer Matrix A erzeugen den gleichen Untervektorraum von $K^n = M(1 \times n, K)$ wie die Zeilen einer Matrix, die man aus A durch eine Folge von elementaren Zeilenumformungen erhält.

Beweis: (a) Definition.
(b) Im Fall einer einzigen Zeilenumformung ist es klar. Der allgemeine Fall folgt mit Induktion. □

Beispiel 4.5

$$\begin{pmatrix} 0 & 1 & 3 & 0 \\ 1 & 2 & 2 & -1 \\ 3 & 4 & 0 & 3 \end{pmatrix} \overset{Z_{III}(1,2)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 2 & -1 \\ 0 & 1 & 3 & 0 \\ 3 & 4 & 0 & 3 \end{pmatrix}$$

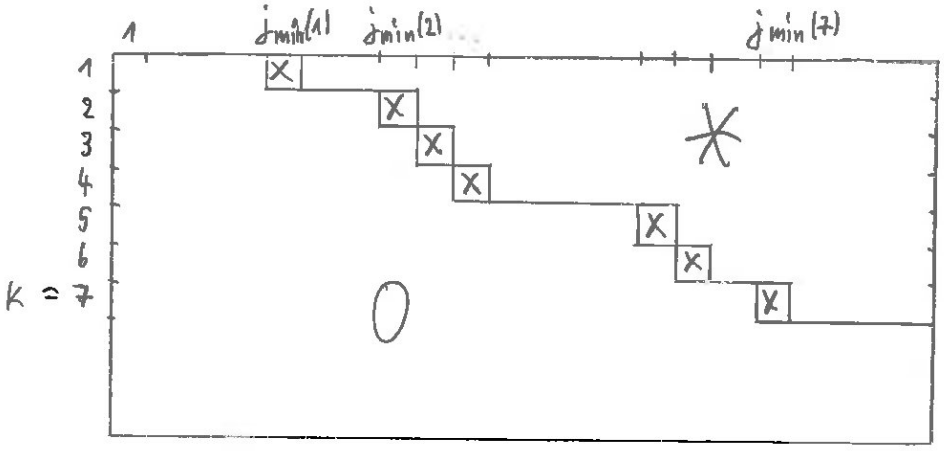
$$\overset{Z_{II}(-3;1,3)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 2 & -1 \\ 0 & 1 & 3 & 0 \\ 0 & -2 & -6 & 6 \end{pmatrix} \overset{Z_{II}(2;2,3)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 2 & -1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

Definition 4.6 Eine Matrix $A = (a_{ij}) \in M(m \times n, K)$ ist in *Zeilenstufenform*, wenn ihre Zeilen $v_i := (a_{i1}, ..., a_{in})$ folgendes erfüllen:

(i) Es gibt ein $k \in \{0, 1, .., m\}$, so daß $v_i = 0$ für $i > k$ (leere Bedingung bei $k = m$) und $v_i \neq 0$ für $i \leq k$ (leere Bedingung bei $k = 0$) ist.

(ii) Für $i \leq k$ sei $j_{min}(i) := \min(j \mid a_{ij} \neq 0)$. Dann ist

$$j_{min}(1) < ... < j_{min}(k).$$

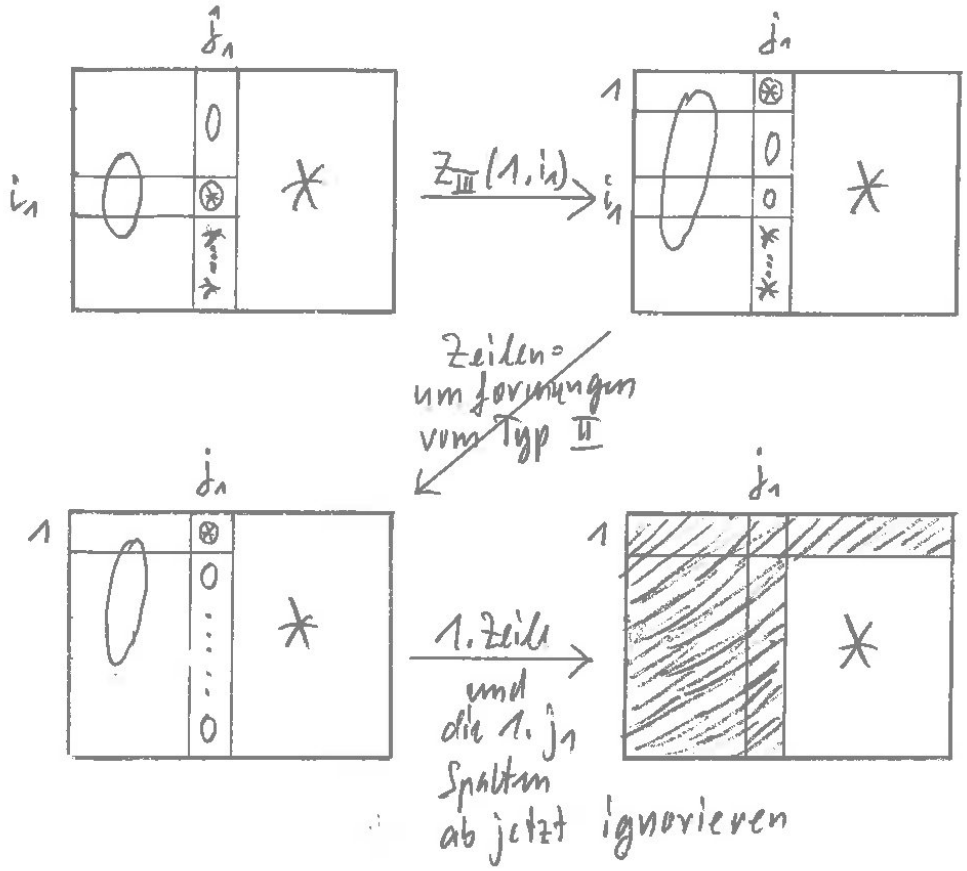


Satz 4.7 *(Gauß-Algorithmus)*
(a) Ist eine Matrix A in Zeilenstufenform, so daß genau die ersten k Zeilen nicht verschwinden, so bilden diese eine Basis des von ihnen erzeugten Untervektorraums, und es ist $\text{Zeilenrang}(A) = k$.

(b) Jede Matrix $A = (a_{ij}) \in M(m \times n, K)$ läßt sich durch eine geeignete Folge von elementaren Zeilenumformungen in Zeilenstufenform bringen. Der von den Zeilen erzeugte Unterraum bleibt dabei gleich. Also bleibt auch der Zeilenrang gleich.
Genauer: Die Aneinanderkettung von Schritten folgenden Typs gibt eine eindeutige Folge von elementaren Zeilenumformungen, die es tut.
1. Schritt:

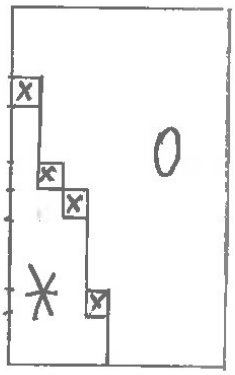
$$\begin{aligned} j_1 &:= \min(j \mid \text{ es gibt ein } i \text{ mit } a_{ij} \neq 0), \\ i_1 &:= \min(i \mid a_{ij_1} \neq 0). \end{aligned}$$

Ist $i_1 \neq 1$, so führt man zuerst $Z_{III}(1, i_1)$ aus. Die neue Matrix (= alte Matrix A bei $i_1 = 1$) nennt man $\widetilde{A} = (\widetilde{a}_{ij})$. Für $i = 2, ..., m$ führt man $Z_{II}(-\widetilde{a}_{i,j_1}/\widetilde{a}_{1,j_1}; 1, i)$ aus.
2. Schritt: *Man streicht die erste Zeile und die ersten j_1 Spalten und führt den 1. Schritt mit der neuen kleineren Matrix aus.*



Beweis: (a) Zu zeigen ist nur, daß die ersten k Zeilen $v_i := (a_{i1}, ..., a_{in})$ linear unabhängig sind. Sei $0 = \sum_{i=1}^{k} \lambda_i v_i$. Es ist $0 = \sum_{i=1}^{k} \lambda_i a_{ij_{min}(1)} = \lambda_1 a_{ij_{min}(1)}$, also $\lambda_1 = 0$. Analog folgt $\lambda_2 = 0$, ..., $\lambda_k = 0$.
(b) Die ersten $j_1 - 1$ Spalten von A und $\widetilde{A}$ sind Null. Offenbar ist $\widetilde{a}_{1j_1} \neq 0$. Nach dem ersten Schritt ist $\widetilde{a}_{1j_1}$ der einzige Eintrag in der j_1-ten Spalte ungleich Null. Der Rest ist klar. (Am Ende wird $j_1 = j_{min}(1)$ sein.) □

Bemerkungen 4.8 (i) Selbstverständlich kann man in konkreten Beispielen von der Schrittfolge oben abweichen, wenn andere Zeilenumformungen günstiger sind.
(ii) Statt Zeilen kann man in 4.4, 4.6 und 4.7 genausogut Spalten betrachten: Man hat den von den Spalten einer Matrix erzeugten Untervektorraum von $M(m \times 1, K)$, *elementare Spaltenumformungen* $S_I(\lambda; i), S_{II}(\lambda, i, j), S_{III}(i, j)$, eine Spaltenstufenform und dafür einen Gauß-Algorithmus.



Definition 4.9 (Matrizen-Multiplikation) Seien $A \in M(l \times m, K)$ und $B \in M(m \times n, K)$ $(l, m, n \in \mathbb{N})$ Matrizen mit

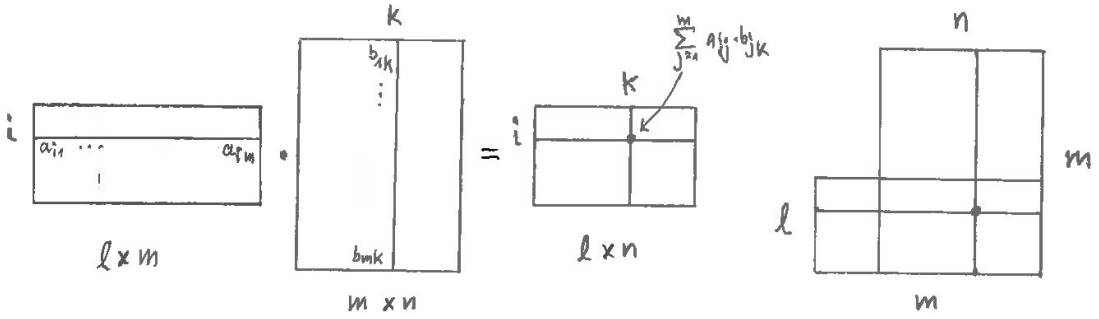
$$(\text{Anzahl der Spalten von } A) = m = (\text{Anzahl der Zeilen von } B).$$

Das *Produkt* $C := A \cdot B$ von A und B ist die $(l \times n)$-Matrix $C = (c_{ik})_{i=1,...,l;k=1,...,n}$ mit

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

Also

$$(l \times m)\text{-Matrix } \cdot (m \times n)\text{-Matrix } = (l \times n)\text{-Matrix}.$$



Eine Anschauung dazu: man dreht die k-te Spalte von B mathematisch positiv um 90 Grad, legt sie auf die i-te Zeile von A, multipliziert aufeinanderliegende Koeffizienten, summiert die Produkte und erhält c_{ik}.

Beispiele 4.10 (i)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 & 3 \\ 0 & 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 & 3 \\ 0 & 1 & 3 & 2 \\ 1 & 2 & 5 & 7 \end{pmatrix}.$$

(ii) Die i-te Zeile von A (in Definition 4.1 (b)) und die k-te Spalte von B sind auch Matrizen; ihr Produkt ist die 1×1-Matrix mit Koeffizient c_{ik}:

$$\begin{pmatrix} a_{i1} & \cdots & a_{im} \end{pmatrix} \cdot \begin{pmatrix} b_{1k} \\ \vdots \\ b_{mk} \end{pmatrix} = (\sum_{j=1}^{m} a_{ij} \cdot b_{jk}).$$

Also: Zeilenvektor $\cdot$ Spaltenvektor $= 1 \times 1$-Matrix.
(iii) Spaltenvektor $\cdot$ Zeilenvektor $= l \times n$-Matrix:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{l1} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & b_{1n} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & \cdots & a_{11}b_{1n} \\ \vdots & \ddots & \vdots \\ a_{l1}b_{11} & \cdots & a_{l1}b_{1n} \end{pmatrix}.$$

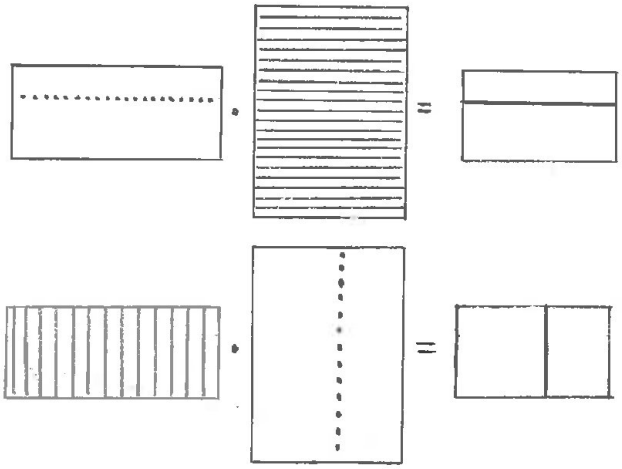
(iv) Das *Kroneckersymbol* δ_{ij} für i und j in einer (gegebenen) Indexmenge ist

$$\delta_{ij} := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j \end{cases}$$

Die $(n \times n)$-*Einheitsmatrix* E_n hat die Einträge $(E_n)_{ij} := \delta_{ij}$, also Einsen in der Diagonalen, Nullen außerhalb. Es ist für $A \in M(m \times n, K)$

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

(v) Die Zeilen einer Produktmatrix $A \cdot B$ sind Linearkombinationen der Zeilen der rechten Matrix B, die Spalten von $A \cdot B$ sind Linearkombinationen der Spalten der linken Matrix A.



(vi) Daher gilt

$$\begin{aligned} \text{Zeilenrang}(A \cdot B) &\leq \text{Zeilenrang}(B), \\ \text{Spaltenrang}(A \cdot B) &\leq \text{Spaltenrang}(A). \end{aligned}$$

(vii) Daher lassen sich die elementaren Zeilenumformungen

$$Z_I(\lambda; i), Z_{II}(\lambda, i, j), Z_{III}(i, j) : M(m \times n, K) \to M(m \times n, K)$$

durch Multiplikation von links mit geeigneten Matrizen $Z_I^{mat}(\lambda, i), Z_{II}^{mat}(\lambda, i, j), Z_{III}^{mat}(i, j) \in M(m \times m, K)$ beschreiben:

$$Z_I(\lambda; i)(A) \;=\; Z_I^{mat}(\lambda; i) \cdot A = \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & & & & & \cdot \\ \cdot & & 1 & & & & \cdot \\ \cdot & & & \lambda & & & \cdot \\ \cdot & & & & 1 & & \cdot \\ \cdot & & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{pmatrix} \cdot A,$$

$$Z_{II}(\lambda,i,j)(A) \;=\; Z_{II}^{mat}(\lambda,i,j)\cdot A = \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & & & & & \cdot \\ \cdot & & 1 & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & \lambda & & 1 & & \cdot \\ \cdot & & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{pmatrix} \cdot A,$$

$$Z_{III}(i,j)(A) \;=\; Z_{III}^{mat}(i,j)\cdot A = \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & & & & & \cdot \\ \cdot & & 0 & & 1 & & \cdot \\ \cdot & & & 1 & & & \cdot \\ \cdot & & 1 & & 0 & & \cdot \\ \cdot & & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{pmatrix} \cdot A;$$

hier ist

$$(Z_I^{mat}(\lambda;i))_{kl} \;=\; \begin{cases} \delta_{kl} & \text{für } (k,l)\neq(i,i) \\ \lambda & \text{für } (k,l)=(i,i), \end{cases}$$

$$(Z_{II}^{mat}(\lambda;i,j))_{kl} \;=\; \begin{cases} \delta_{kl} & \text{für } (k,l)\neq(j,i) \\ \lambda & \text{für } (k,l)=(j,i), \end{cases}$$

$$(Z_{III}^{mat}(i,j))_{kl} \;=\; \begin{cases} \delta_{kl} & \text{für } (k,l)\notin\{(i,i),(i,j),(j,i),(j,j)\}, \\ 1 & \text{für } (k,l)\in\{(i,j),(j,i)\}, \\ 0 & \text{für } (k,l)\in\{(i,i),(j,j)\}. \end{cases}$$

(viii) Analog lassen sich die elementaren Spaltenumformungen (Bemerkung 4.8 (ii)) durch Multiplikation von rechts mit geeigneten Matrizen beschreiben.

Satz 4.11 *(a) Die Multiplikation von Matrizen ist im allgemeinen nicht kommutativ.*
(b) Aber sie ist assoziativ: Sind $A\in M(k\times l,K)$, $B\in M(l\times m,K)$, $C\in M(m\times n,K)$, *so ist*

$$(A\cdot B)\cdot C = A\cdot(B\cdot C).$$

(c) Die Menge $M(n\times n,K)$ *ist ein Ring mit Eins. Die Eins ist* E_n. *Für* $n\geq 2$ *ist der Ring nicht kommutativ.*

Beweis: (a) Wenn $A\in M(p\times q,K)$ und $B\in M(q\times r,K)$ mit $p\neq r$ ist, so existiert $B\cdot A$ nicht einmal. Bei $p=q=r=2$ ist zum Beispiel

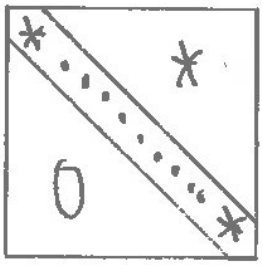
$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

(b) Es ist (mit der Notation $(A)_{ij}$ für die Koeffizienten einer Matrix A)

$$\begin{aligned}((A\cdot B)\cdot C)_{il} &\overset{\text{Def.}}{=} \sum_{k=1}^{m}(A\cdot B)_{ik}\cdot(C)_{kl}\\ &\overset{\text{Def.}}{=} \sum_{k=1}^{m}\left(\sum_{j=1}^{l}(A)_{ij}(B)_{jk}\right)\cdot(C)_{kl}\\ &\overset{!}{=} \sum_{j=1}^{l}(A)_{ij}\cdot\left(\sum_{k=1}^{m}(B)_{jk}(C)_{kl}\right)\\ &\overset{\text{Def.}}{=} \sum_{j=1}^{l}(A)_{ij}\cdot(B\cdot C)_{jl} \overset{\text{Def.}}{=} (A\cdot(B\cdot C))_{il}.\end{aligned}$$

(c) Distributivgesetze: Übung. Der Rest folgt aus (a) und (b). □

Definition/Lemma 4.12 *(a) (Definition) Eine quadratische Matrix $A = (a_{ij}) \in M(n\times n, K)$ ist eine obere Dreiecksmatrix, falls $a_{ij} = 0$ ist für $i > j$.*



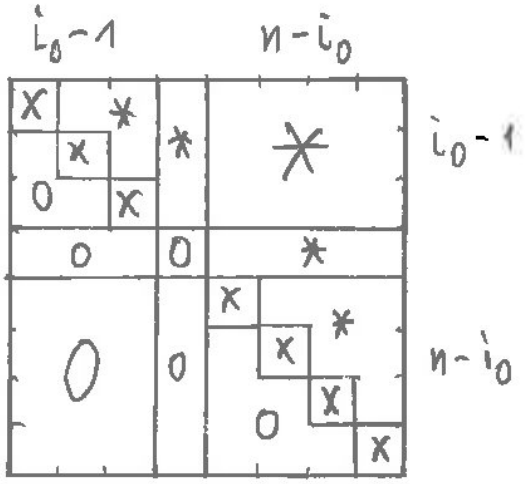
(b) (Lemma) Eine obere Dreiecksmatrix $A = (a_{ij}) \in M(n\times n, K)$ hat genau dann Zeilenrang n, wenn alle Diagonaleinträge a_{ii} ungleich 0 sind.

Beweis: (a) Definition.

(b) "⇐": Dann ist A in Zeilenstufenform und Zeilenrang$(A) = n$ nach Satz 4.7 (a).

"⇒": Indirekter Beweis. Annahme: $a_{i_0 i_0} = 0$ und $a_{ii} \neq 0$ für $1 \leq i < i_0$.

Die hinteren $n - (i_0 - 1)$ Zeilen liegen im $n - i_0$ dimensionalen Unterraum $\{(x_1, ..., x_n) \mid x_1 = ... = x_{i_0} = 0\}$ von K^n. Zusammen mit den ersten $i_0 - 1$ Zeilen erzeugen sie einen höchstens $n - 1$ dimensionalen Unterraum von K^n. Also ist Zeilenrang$(A) \leq n - 1$.



□

Satz/Definition 4.13 *(a) (Satz) Sei $A \in M(n \times n, K)$ eine quadratische Matrix. Die folgenden Bedingungen sind äquivalent.*

(i) Zeilenrang$(A) = n$.

(ii) Es gibt eine Matrix $B \in M(n \times n, K)$ mit $B \cdot A = E_n$.

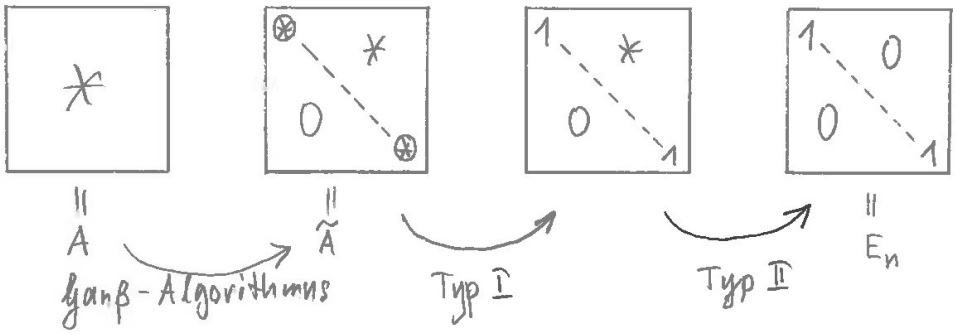
(iii) Es gibt genau eine Matrix $B \in M(n \times n, K)$ mit $B \cdot A = E_n$, und sie erfüllt auch $A \cdot B = E_n$.

(b) (Definition) Eine quadratische Matrix, die die Eigenschaften in (a) erfüllt, heißt invertierbar. Dann heißt die Matrix B in (iii) die inverse Matrix zu A und wird mit A^{-1} bezeichnet.
*(c) (Definition/Satz) Die Menge $GL(n, K)$ aller invertierbaren Matrizen in $M(n \times n, K)$ ist eine Gruppe. Für $n \geq 2$ ist sie nicht abelsch. Es ist $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$. ("GL" steht für "**g**eneral **l**inear (group)".)*
(d) (Satz) Eine obere Dreiecksmatrix ist genau dann invertierbar, wenn alle Diagonaleinträge ungleich 0 sind.
(e) (Satz, wichtige Formel, auswendig lernen) Eine (2×2)-Matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ ist genau dann invertierbar, wenn $ad - bc \neq 0$ ist. Dann ist

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beweis: (a) "(i) ⇒ (ii)": Es reicht zu zeigen, daß A durch Zeilenumformungen in E_n transformiert werden kann. Denn jede Zeilenumformung ist eine Multiplikation mit einer Matrix von links (Bemerkung 4.10 (vi)). Wegen der Assoziativität der Multiplikation ist dann B das Produkt der Matrizen zu den Zeilenumformungen.

Mit dem Gauß-Algorithmus (Satz 4.7) erhält man eine Matrix $\widetilde{A}$ in Zeilenstufenform aus A. Wegen Zeilenrang$(\widetilde{A})$ = Zeilenrang$(A) = n$ und Lemma 4.12 sind alle Diagonaleinträge ungleich 0. Mit Zeilenumformungen vom Typ I normiert man sie zu 1. Mit Zeilenumformungen vom Typ II löscht man alle Einträge oberhalb der Diagonalen.



"(ii) ⇒ (iii)": Aus $B \cdot (A \cdot B) = (B \cdot A) \cdot B = E_n \cdot B = B$ folgt $B \cdot (A \cdot B - E_n) = 0$. Es ist Spaltenrang$(B) \geq$ Spaltenrang$(B \cdot A) = n$, also Spaltenrang$(B) = n$. Also bilden die Spalten von B eine Basis von $M(n \times 1, K)$. Daher folgt aus $B \cdot v = 0$ für $v \in M(n \times 1, K)$ auch $v = 0$. Das gibt $A \cdot B - E_n = 0$, also $A \cdot B = E_n$.
Ist $\widetilde{B} \cdot A = E_n$, so ist $B = (\widetilde{B} \cdot A) \cdot B = \widetilde{B} \cdot (A \cdot B) = \widetilde{B}$.
"(iii) ⇒ (i)": Es ist Zeilenrang$(A) \geq$ Zeilenrang$(B \cdot A) = n$, also Zeilenrang$(A) = n$.
(b) Definition.
(c) Multiplikation assoziativ: 4.11 (b); E_n neutrales Element: 4.10 (iv); inverse Elemente: 4.13 (a). $GL(n, K)$ nicht abelsch: siehe Beweis von Satz 4.11 (a). $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$ gilt wegen

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot A^{-1} = E_n.$$

(d) Lemma 4.12 (b).
(e) Bei $ad - bc = 0$ ist eine Zeile Linearkombination der anderen (falls eine Zeile gleich 0 ist, nur die). Formel: nachrechnen. □

Bemerkung 4.14 Aus dem Beweis von "(i) ⇒ (ii)" erhält man einen Algorithmus zur Berechnung der inversen Matrix: Man schreibt A und E_n nebeneinander und führt an beiden die gleichen Zeilenumformungen durch, so daß man E_n aus A erhält. Dann erhält man A^{-1} aus E_n.
Denn sind $Z_1, ..., Z_k$ die Matrizen zu den Zeilenumformungen (Beispiel 4.10 (vi)), so ist $Z_k \cdot ... \cdot Z_1 \cdot A = E_n$, also $Z_K \cdot ... \cdot Z_1 \cdot E_n = A^{-1}$.
Ein Beispiel: (die Zeilenumformungen sind nach Gefühl gewählt, nicht strikt nach dem Gauß-Algorithmus)

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 9 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_n$$

$$Z_{II}(-1;1,3) \circ Z_{II}(-1;1,2): \quad \begin{pmatrix} 0 & 1 & 2 \\ 3 & 3 & 3 \\ 6 & 6 & 7 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$Z_{II}(-2;2,3): \quad \begin{pmatrix} 0 & 1 & 2 \\ 3 & 3 & 3 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$$

$$Z_{III}(1,2): \quad \begin{pmatrix} 3 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -2 & 1 \end{pmatrix}$$

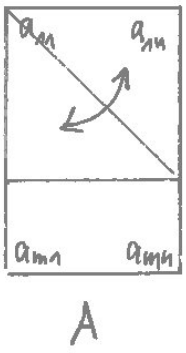
$$Z_I(\frac{1}{3};1): \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} -\frac{1}{3} & \frac{1}{3} & 0 \\ 1 & 0 & 0 \\ 1 & -2 & 1 \end{pmatrix}$$

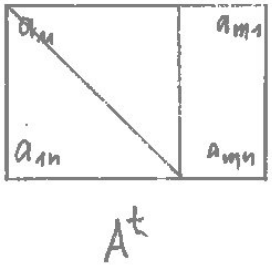
$$Z_{II}(-1;3,1) \circ Z_{II}(-2;3,2): \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} -\frac{4}{3} & \frac{7}{3} & -1 \\ -1 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

$$Z_{II}(-1;2,1): \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} -\frac{1}{3} & -\frac{5}{3} & 1 \\ -1 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} = A^{-1}$$

Definition/Lemma 4.15 *(a) Die transponierte Matrix A^{tr} einer Matrix $A = (a_{ij}) \in M(k \times l, K)$ ist die Matrix $A^{tr} \in M(l \times k, K)$ mit*

$$(A^{tr})_{ij} := a_{ji}.$$

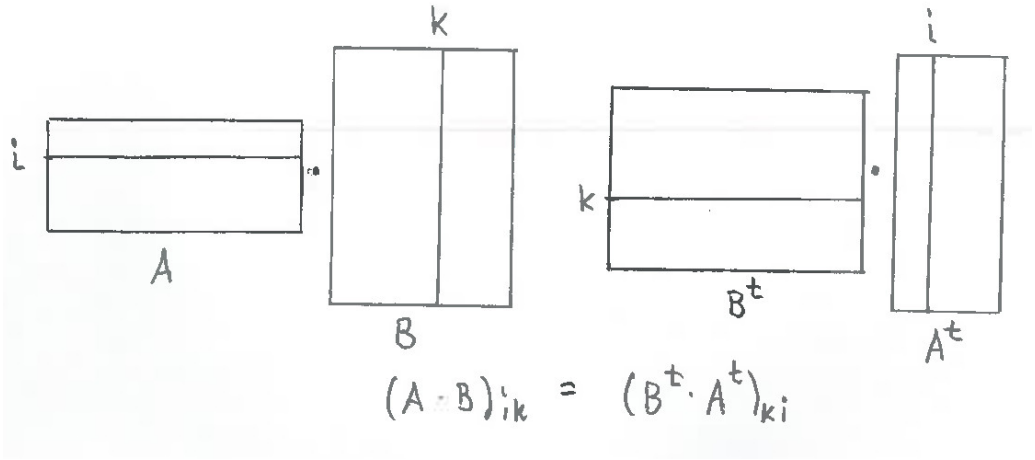




(b) Ist $A \in M(k \times l, K)$ und $B \in M(l \times m, K)$, so ist

$$(A \cdot B)^{tr} = B^{tr} \cdot A^{tr}.$$

Beweis: (a) Definition.
(b) Übung.



□

5 Lineare Abbildungen

In diesem Kapitel bezeichnet K stets irgendeinen Körper.

Definition 5.1 (a) Eine Abbildung $f : V \to W$ von einem K-Vektorraum V in einen K-Vektorraum W heißt *linear* (oder *K-linear*), falls sie folgende Eigenschaften erfüllt:

(i) f ist ein Gruppenhomomorphismus von $(V, +)$ nach $(W, +)$, d.h.

$$f(a + b) = f(a) + f(b) \quad \text{für } a, b \in V;$$

(ii) f ist kompatibel mit den skalaren Multiplikationen von V und W, d.h.

$$f(\lambda \cdot a) = \lambda \cdot f(a) \quad \text{für } \lambda \in K, a \in V.$$

Ein anderer (seltener benutzter) Name für *lineare Abbildung* ist *Vektorraumhomomorphismus*.
(b) Eine lineare Abbildung $f : V \to W$ zwischen zwei Vektorräumen ist
ein *Isomorphismus*, falls sie bijektiv ist,
ein *Endomorphismus*, falls $V = W$ ist,
ein *Automorphismus*, falls sie bijektiv ist und $V = W$ ist.
(c) Zwei Vektorräume V und W heißen *isomorph*, wenn ein Isomorphismus $f : V \to W$ existiert.

Beispiele 5.2 (i) Der einfachste Fall: $f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto a \cdot x$, für ein $a \in \mathbb{R}$.
(ii) (Verallgemeinerung von (i)) $f : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, ..., x_n) \mapsto \sum_{j=1}^n a_j x_j$, für $a_1, ..., a_n \in \mathbb{R}$.
(iii) (Äquivalent zu (ii)) $f : M(n \times 1, \mathbb{R}) \to M(1 \times 1, \mathbb{R})$,

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_1 \cdots a_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\sum_{i=1}^n a_i x_i)$$

für $a_1, ..., a_n \in \mathbb{R}$.
(iv) (Verallgemeinerung von (iii), mit K statt $\mathbb{R}$)
Sei $A = (a_{ij}) \in M(m \times n, K)$. Die Abbildung

$$\begin{array}{rcl} f : M(n \times 1, K) & \to & M(m \times 1, K), \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{1i} x_i \\ \vdots \\ \sum_{i=1}^n a_{mi} x_i \end{pmatrix} \end{array}$$

ist linear. Beweis: Details im Kopf; man muß 5.1 (i) und (ii) zeigen.

Mit $x \in M(n \times 1, K)$ läßt sich die Abbildung ganz kurz schreiben als

$$x \mapsto A \cdot x.$$

In Satz 5.5 (a)+(b) und Satz 5.11 (a)+(b) werden wir sehen, daß jede lineare Abbildung von $M(n \times 1, K)$ nach $M(m \times 1, K)$ von dieser Gestalt ist und daß jede lineare Abbildung zwischen endlich-dimensionalen Vektorräumen "äquivalent" zu einer solchen Abbildung ist.
(v) Sei X eine nichtleere Menge, $x_0 \in X$, K ein Körper. Die Einsetzungsabbildung

$$\Phi_{x_0} : \mathrm{Abb}(X, K) \to K, \quad g \mapsto g(x_0),$$

ist eine lineare Abbildung.
(vi) Die Mengen $\mathcal{C}^0([0,1], \mathbb{R})$, $\mathcal{C}^1([0,1], \mathbb{R})$ und $\mathbb{R}[t]$ sind $\mathbb{R}$-Vektorräume (Beispiel 3.3 (d)). Die Ableitung

$$\frac{\mathrm{d}}{\mathrm{d}x} : \mathcal{C}^1([0,1], \mathbb{R}) \to \mathcal{C}^0([0,1], \mathbb{R}), \quad g \mapsto \frac{\mathrm{d}g}{\mathrm{d}x},$$

ist linear, ebenso ihre Einschränkung $\frac{\mathrm{d}}{\mathrm{d}x} : \mathbb{R}[x] \to \mathbb{R}[x], \quad g \mapsto \frac{\mathrm{d}g}{\mathrm{d}x}$.

Satz/Definition 5.3 *Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen*
(a) (Satz) Dann ist $f(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i f(v_i)$ für $n \in \mathbb{N}$, $\lambda_1, ..., \lambda_n \in K$, $v_1, ..., v_n \in V$.
(b) (Definition) Der Kern von f ist $\ker f := \{x \in V \mid f(x) = 0\}$. Das Bild ist $f(V) \subset W$.
(c) (Satz) $\ker f$ ist ein Untervektorraum von V, $f(V)$ ist ein Untervektorraum von W.
(d) (Satz) f ist genau dann injektiv, wenn $\ker f = \{0\}$ ist (und f ist nach Definition genau dann surjektiv, wenn $f(V) = W$ ist).
(e) (Satz) Ist f ein Isomorphismus, so ist auch die (natürlich ebenfalls bijektive) Umkehrabbildung $f^{-1} : W \to V$ linear und damit ein Isomorphismus von Vektorräumen.
(f) (Definition) Der Rang von f ist $\mathrm{rang}\, f := \dim_K f(V)$.
(g) (Satz)

$$\dim_K V = \dim_K \ker f + \mathrm{rang}\, f$$

(mit $\infty = \infty + n = n + \infty = \infty + \infty$ für $n \in \mathbb{N}_0$).
(h) (Satz) Ist $g : U \to V$ eine zweite lineare Abbildung zwischen K-Vektorräumen, so ist auch die Komposition $f \circ g : U \to W$ eine lineare Abbildung.

Beweis: (a) Klar. (b) Definition. (c) Nach Lemma 1.19 sind $\ker f$ und $f(V)$ Untergruppen von V bzw. W. Es bleibt zu zeigen, daß sie abgeschlossen unter der skalaren Multiplikation sind:

$$\begin{aligned} v_1 \in \ker f, \lambda \in K &\Rightarrow f(\lambda \cdot v_1) = \lambda \cdot f(v_1) = \lambda \cdot 0 = 0 \\ &\Rightarrow \lambda \cdot v_1 \in \ker f; \\ v_2 \in V, \lambda \in K &\Rightarrow \lambda \cdot f(v_2) = f(\lambda \cdot v_2) \in f(V). \end{aligned}$$

(d) Es ist $f(0) = 0$, denn für ein beliebiges $v \in V$ gilt

$$f(0_V) = f(0_K \cdot v) = 0_K \cdot f(v) = 0_V.$$

"$\Rightarrow$": f injektiv und $f(a) = 0 \Rightarrow a = 0$.
"$\Leftarrow$": $f(a) = f(b) \Rightarrow f(a-b) = 0 \Rightarrow$ (wegen $\ker f = 0$) $a - b = 0 \Rightarrow a = b$.
(e) Lemma 1.19 $\Rightarrow f^{-1} : (W, +) \to (V, +)$ ist ein Isomorphismus abelscher Gruppen.
Aus $f(\lambda \cdot a) = \lambda \cdot f(a)$ und $a = f^{-1}(b)$ folgt $\lambda \cdot f^{-1}(b) = f^{-1}(\lambda \cdot b)$.
(f) Definition.
(g) **1. Fall,** $\dim_K \ker f = \infty$:
Nach Satz 3.17 (c) ist $\dim V \geq \dim \ker f = \infty$. Also ist auch $\dim V = \infty$.
2. Fall, $\dim_K f(V) = \infty$:
Wäre $(a_1, ..., a_n)$ ein Erzeugendensystem von V, so wäre $(f(a_1), ..., f(a_n))$ ein Erzeugendensystem von $f(V)$, also $\dim f(V) < \infty$, Widerspruch. Also hat V kein endliches Erzeugendensystem; also ist $\dim V = \infty$.
3. Fall (der interessanteste Fall), $\dim_K \ker f < \infty$ und $\dim_K f(V) < \infty$:
Man wählt eine Basis $(v_1, ..., v_k)$ von $\ker f$, und man wählt $w_1, ..., w_l \in V$, so daß $(f(w_1), ..., f(w_l))$ eine Basis von $f(V)$ ist. Es reicht, folgende Behauptung zu beweisen.
Behauptung: $(v_1, ..., v_k, w_1, ..., w_l)$ ist eine Basis von V.
Beweis: (i) Erzeugendensystem: sei $a \in V$. Es gibt $\mu_1, ..., \mu_l \in K$ mit $f(a) = \sum_{j=1}^{l} \mu_j f(w_j)$. Man sieht sofort

$$a - \sum_{j=1}^{l} \mu_j w_j \in \ker f.$$

Also gibt es $\lambda_1, ..., \lambda_k \in K$ mit

$$a - \sum_{j=1}^{l} \mu_j w_j = \sum_{i=1}^{k} \lambda_i v_i.$$

Also ist a eine Linearkombination der v_i und w_j.
(ii) Linear unabhängig: Sei $\sum_{i=1}^{k} \alpha_i v_i + \sum_{j=1}^{l} \beta_j w_j = 0$. Sein Bild unter f ist $\sum_{j=1}^{l} \beta_j f(w_j) = 0$. Weil $(f(w_1), ..., f(w_l))$ eine Basis von $f(V)$ ist, sind alle $\beta_j = 0$. Weil $(v_1, ..., v_k)$ eine Basis von $\ker f$ ist, sind auch alle $\alpha_i = 0$.
(h) Klar. □

Beispiel 5.4 In Beispiel 5.2 (v) ist $\ker \Phi_{x_0} = \{g \in \mathrm{Abb}(X, K) \mid g(x_0) = 0\}$. In Beispiel 5.2 (vi) ist $\ker \frac{\mathrm{d}}{\mathrm{d}x} = \{$ konstante Abbildungen.$\}$. In beiden Fällen ist die betrachtete Abbildung surjektiv.

Satz/Definition 5.5 *(Matrizen und lineare Abbildungen, 1. Teil)*

(a) (Satz) Zu jeder linearen Abbildung $f : M(n\times 1, K) \to M(m\times 1, K)$ *gibt es genau eine Matrix* $A \in M(m \times n, K)$ *mit*

$$f(x) = A \cdot x.$$

(Definition) Diese Matrix wird $\mathrm{Mat}(f)$ *genannt. Also ist* $f(x) = \mathrm{Mat}(f) \cdot x$.
(b) (Definition/Satz) Die Menge

$$\begin{aligned} &\mathrm{Hom}_K(M(n \times 1, K), M(m \times 1, K)) \\ := \ &\{f : M(n \times 1, K) \to M(m \times 1, K) \mid f \text{ ist linear}\} \end{aligned}$$

ist ein K*-Vektorraum, und die Abbildung*

$$Mat : \mathrm{Hom}_K(M(n \times 1, K), M(m \times 1, K)) \to M(m \times n, K), \quad f \mapsto \mathrm{Mat}(f),$$

ist ein Isomorphismus von K*-Vektorräumen.*
(c) Sind

$$f : M(n \times 1, K) \to M(m \times 1, K)$$

und

$$g : M(m \times 1, K) \to M(l \times 1, K)$$

lineare Abbildungen, so ist auch

$$g \circ f : M(n \times 1, K) \to M(l \times 1, K)$$

linear (Satz 5.3 (h)), und es ist

$$\mathrm{Mat}(g \circ f) = \mathrm{Mat}(g) \cdot \mathrm{Mat}(f).$$

(d) Eine lineare Abbildung $f : M(n \times 1, K) \to M(n \times 1, K)$ *ist genau dann ein Isomorphismus, wenn* $\mathrm{Mat}(f)$ *invertierbar ist. Dann ist* $\mathrm{Mat}(f^{-1}) = (\mathrm{Mat}(f))^{-1}$.

Beweis: (a) Es sei $e_1^{(n)} = (1, 0, ..., 0)^{tr}$,...,$e_n^{(n)} = (0, ..., 0, 1)^{tr}$ die Standardbasis von $M(n\times 1, K)$ und $e_1^{(m)}$,...,$e_m^{(m)}$ die Standardbasis von $M(m\times 1, K)$. Für jedes $e_j^{(n)}$ (mit $j = 1, ..., n$) gibt es *eindeutige* $a_{ij} \in K$ (mit $i = 1, ..., m$) mit

$$f(e_j^{(n)}) = \sum_{i=1}^{m} a_{ij} e_i^{(m)},$$

denn $e_1^{(m)}$,...,$e_m^{(m)}$ ist eine Basis von $M(m \times 1, K)$. Also ist

$$f(e_j^{(n)}) = \sum_{i=1}^{m} a_{ij} e_i^{(m)} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = (a_{ij}) \cdot e_j^{(n)}.$$

Daraus folgt schon die Eindeutigkeit der Matrix $A = (a_{ij})$.
Es ist

$$\begin{aligned} f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) &= f\left(\sum_{j=1}^{n} x_j e_j^{(n)}\right) = \sum_{j=1}^{n} x_j \cdot f(e_j^{(n)}) \\ &= \sum_{j=1}^{n} x_j \cdot (a_{ij}) \cdot e_j^{(n)} = (a_{ij}) \cdot \left(\sum_{j=1}^{n} x_j e_j^{(n)}\right) = (a_{ij}) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

Das sagt, daß $f(x) = (a_{ij}) \cdot x$ ist.
(b) $\mathrm{Mat}(f)$ bestimmt f. Daher ist die Abbildung $f \mapsto \mathrm{Mat}(f)$ injektiv.
Sie ist surjektiv, denn die Multiplikation von links mit einer Matrix ist eine lineare Abbildung (Bemerkung 5.2 (iv)).
Die Abbildung $f \mapsto \mathrm{Mat}(f)$ ist linear:

$$\begin{aligned} (f+g)(x) &\overset{\text{Def.}}{=} f(x) + g(x) = \mathrm{Mat}(f) \cdot x + \mathrm{Mat}(g) \cdot x \\ &= (\mathrm{Mat}(f) + \mathrm{Mat}(g)) \cdot x, \end{aligned}$$

und weil $\mathrm{Mat}(f+g)$ eindeutig ist, ist $\mathrm{Mat}(f+g) = \mathrm{Mat}(f) + \mathrm{Mat}(g)$.
Genauso folgt aus

$$(\lambda \cdot f)(x) \overset{\text{Def.}}{=} \lambda \cdot f(x) = \lambda \cdot (\mathrm{Mat}(f) \cdot x) = (\lambda \cdot \mathrm{Mat}(f)) \cdot x,$$

daß $\mathrm{Mat}(\lambda \cdot f) = \lambda \cdot \mathrm{Mat}(f)$ ist.
(c) Es ist

$$\begin{aligned} (g \circ f)(x) &= g(f(x)) = g\left(\mathrm{Mat}(f) \cdot x\right) = \mathrm{Mat}(g) \cdot (\mathrm{Mat}(f) \cdot x) \\ &= (\mathrm{Mat}(g) \cdot \mathrm{Mat}(f)) \cdot x. \end{aligned}$$

Weil $\mathrm{Mat}(g \circ f)$ eindeutig ist, ist

$$\mathrm{Mat}(g \circ f) = \mathrm{Mat}(g) \cdot \mathrm{Mat}(f).$$

(d) "$\Rightarrow$": Sei f ein Isomorphismus. Nach Satz 5.3 ist $f^{-1} : M(n\times 1, K) \to M(n\times 1, K)$ linear. Es ist $f^{-1} \circ f = \mathrm{id} = f \circ f^{-1}$. Mit (c) folgt

$$\mathrm{Mat}(f^{-1}) \cdot \mathrm{Mat}(f) = \mathrm{Mat}(f^{-1} \circ f) = \mathrm{Mat}(\mathrm{id}) = E_n.$$

Also ist $\mathrm{Mat}(f)$ invertierbar und $\mathrm{Mat}(f^{-1})$ die inverse Matrix.

"$\Leftarrow$": Sei $\mathrm{Mat}(f)$ invertierbar und A die inverse Matrix, also

$$\mathrm{Mat}(f) \cdot A = E_n = A \cdot \mathrm{Mat}(f).$$

Wegen (b) gibt es eine eindeutige Abbildung $g : M(n \times 1, K) \to M(n \times 1, K)$ mit $A = \mathrm{Mat}(g)$. (g ist die Multiplikation mit A von links.) Also ist

$$Mat(g \circ f) = \mathrm{Mat}(g) \cdot \mathrm{Mat}(f) = E_n = \mathrm{Mat}(f) \cdot \mathrm{Mat}(g) = \mathrm{Mat}(f \circ g),$$

also

$$g \circ f = \mathrm{id} = f \circ g.$$

Also ist f invertierbar und g das Inverse, und f ist ein Isomorphismus. □

Bemerkungen 5.6 (i) Ein Beispiel ist die Abbildung $f : M(2\times1, K) \to M(3\times1, K)$ mit

$$f(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_2 \\ x_1 - x_2 \\ 3x_1 + 4x_2 \end{pmatrix}.$$

(ii) Im Beweis von (c) wurde die Assoziativität der Matrizenmultiplikation benutzt. Umgekehrt folgt aus (c) und der Assoziativität der Komposition von Abbildungen, $(h \circ g) \circ f = h \circ (g \circ f)$, die Assoziativität der Matrizenmultiplikation sofort.
(iii) Oft hat man lineare Abbildungen zwischen abstrakten endlich-dimensionalen Vektorräumen. Um dann Matrizen zu erhalten, muß man Basen der Vektorräume wählen.
Eine Korrespondenz ist in Definition 5.8 (a) und Satz 5.11 (b) formuliert. Ihre Eigenschaften sind in Satz 5.11 diskutiert. Er verallgemeinert Satz 5.5. Mit Lemma 5.10 kann man ihn weitgehend auf Satz 5.5 zurückspielen.

Notation 5.7 (Eine Verallgemeinerung der Matrizenmultiplikation)
Sei V ein K-Vektorraum und $m, n \in \mathbb{N}$. Die Menge V^m wird mit der Menge $M(1 \times m, V)$ der Zeilenvektoren mit Einträgen in V identifiziert (vgl. Notation 4.1 (c)). Die Abbildung

$$V^m \times M(m \times n, K) \to V^n,$$
$$((b_1, ..., b_m), (a_{ij})) \mapsto (\sum_{i=1}^{m} a_{i1} b_i, ..., \sum_{i=1}^{m} a_{in} b_i)$$

wird als eine Verallgemeinerung der Matrizenmultiplikation aufgefaßt:

$$(\sum_{i=1}^{m} a_{i1} b_i \cdots \sum_{i=1}^{m} a_{in} b_i) = (b_1 \cdots b_m) \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

(Die Verallgemeinerung besteht darin, daß hier $(b_1, ..., b_m)$ Einträge in V und nicht in K hat und daß die Körpermultiplikation durch die skalare Multiplikation ersetzt ist und deshalb die a_{ij} links von den b_i stehen.)

Mit $\mathcal{B} := (b_1, ..., b_m) \in V^m = M(1 \times m, V)$ und $A = (a_{ij})$ läßt sich die Abbildung $V^m \times M(m \times n, K) \to V^n$ sehr kurz schreiben als

$$(\mathcal{B}, A) \mapsto \mathcal{B} \cdot A.$$

Definition 5.8 (Matrizen und lineare Abbildungen, 2. Teil)
(a) Es sei $f : U \to V$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorräumen U und V. Es sei $\mathcal{A} = (a_1, ..., a_n)$ eine Basis von U und $\mathcal{B} = (b_1, ..., b_m)$ eine Basis von V.
Weil $\mathcal{B}$ eine Basis von V ist, gibt es eindeutige Koeffizienten $\lambda_{ij} \in K$ mit

$$f(a_j) = \sum_{i=1}^{m} \lambda_{ij} b_i.$$

Mit der Verallgemeinerung der Matrixmultiplikation oben lassen sich diese Gleichungen für $j = 1, ..., n$ schön kompakt zusammenfassen zu

$$(f(a_1), ..., f(a_n)) = (b_1, ..., b_m) \cdot (\lambda_{ij}).$$

Die Matrix (λ_{ij}) wird $M(\mathcal{B}, f, \mathcal{A})$ genannt. Mit ihrer Hilfe wird das Bild von $\mathcal{A}$ unter f in $\mathcal{B}$ ausgedrückt. Wenn wir (etwas unsauber, aber elegant) $f(\mathcal{A}) := (f(a_1), ..., f(a_n))$ schreiben, wird die Gleichung oben noch kompakter,

$$f(\mathcal{A}) = \mathcal{B} \cdot M(\mathcal{B}, f, \mathcal{A}).$$

(b) Im Spezialfall $U = V$ und $f = \mathrm{id}$ heißt die Matrix $M(\mathcal{B}, \mathrm{id}, \mathcal{A})$ *Basiswechselmatrix* und wird auch mit $M(\mathcal{B}, \mathcal{A})$ bezeichnet.

Beispiele 5.9 (i) Die Menge $\mathbb{R}[t]_{\le n} := \{f \in \mathbb{R}[t] \mid \deg f \le n\}$ ist ein Vektorraum der Dimension $n + 1$ mit Basis $\mathcal{B}_n := (1, t, t^2, ..., t^n)$. Die Ableitung $\frac{\mathrm{d}}{\mathrm{d}t}$ kann man auffassen als eine lineare Abbildung $\mathbb{R}[t]_{\le n} \to \mathbb{R}[t]_{\le n-1}$. Für $n = 3$ ist

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} 1 & t & t^2 & t^3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2t & 3t^2 \end{pmatrix} = \begin{pmatrix} 1 & t & t^2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

$$\text{also} \qquad M(\mathcal{B}_2, \frac{\mathrm{d}}{\mathrm{d}t}, \mathcal{B}_3) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(ii) $\mathbb{C}$ als $\mathbb{R}$-Vektorraum hat die Standardbasis $(1, i)$. Die Multiplikation $m_z : \mathbb{C} \to \mathbb{C}$ mit einer komplexen Zahl $z = |z|e^{i\alpha}$ ist ein Endomorphismus auf $\mathbb{C}$ als $\mathbb{C}$-Vektorraum und erst recht als $\mathbb{R}$-Vektorraum. Wegen $z = |z| \cos\alpha + i|z| \sin\alpha$ und $z \cdot i = -|z| \sin\alpha + i|z| \cos\alpha$ ist

$$\begin{pmatrix} z \cdot 1 & z \cdot i \end{pmatrix} = \begin{pmatrix} 1 & i \end{pmatrix} \cdot \begin{pmatrix} |z| \cos\alpha & -|z| \sin\alpha \\ |z| \sin\alpha & |z| \cos\alpha \end{pmatrix},$$

$$\text{also} \qquad M((1, i), m_z, (1, i)) = \begin{pmatrix} |z| \cos\alpha & -|z| \sin\alpha \\ |z| \sin\alpha & |z| \cos\alpha \end{pmatrix}.$$

(iii) Eine Basis des K-Vektorraum $K[t]_{\leq 3}$ ist $\mathcal{B} := (1, t, t^2, t^3)$, eine andere ist $\mathcal{A} := (1 + 2t, 3 - t + 2t^2 + t^3, 5t^3, t)$. Die Basiswechselmatrizen sind

$$\begin{aligned} M(\mathcal{B}, \mathcal{A}) &= \begin{pmatrix} 1 & 3 & 0 & 0 \\ 2 & -1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 5 & 0 \end{pmatrix} \quad \text{und} \\ M(\mathcal{A}, \mathcal{B}) &= M(\mathcal{B}, \mathcal{A})^{-1} \quad \text{(vgl. Satz 5.11 (e))} \end{aligned}$$

(iv) Der K-Vektorraum $M(n \times 1, K)$ hat die Standardbasis

$$\mathcal{B}^{(n)} = (e_1^{(n)}, ..., e_n^{(n)}) = (\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}).$$

Sei $f : M(n \times 1, K) \to M(m \times 1, K)$ eine lineare Abbildung. Die Konstruktion von $\mathrm{Mat}(f)$ im Beweis von Satz 5.5 (a) zeigt

$$f(\mathcal{B}^{(n)}) = \mathcal{B}^{(m)} \cdot \mathrm{Mat}(f),$$

also

$$\mathrm{Mat}(f) = M(\mathcal{B}^{(m)}, f, \mathcal{B}^{(n)}).$$

Das gibt eine direkte Beziehung zwischen den Notationen $\mathrm{Mat}(f)$ und $M(\mathcal{B}, f, \mathcal{A})$. Satz 5.11 (a) verallgemeinert diese Formel.

Lemma 5.10 *(a) Sei V ein n-dimensionaler K-Vektorraum und $\mathcal{B} = (b_1, ..., b_n)$ eine Basis von V. Die Abbildung*

$$l_{\mathcal{B}} : M(n \times 1, K) \to V, \quad x \mapsto \mathcal{B} \cdot x,$$

ist ein Isomorphismus von K-Vektorräumen.
(b) Da $M(n \times 1, K)$ natürlich isomorph zu $K^n = M(1 \times n, K)$ ist, ist jeder n-dimensionale K-Vektorraum isomorph zu K^n.

Beweis: (a) $l_{\mathcal{B}}$ ist bijektiv, denn $\mathcal{B}$ ist eine Basis von V.
$l_{\mathcal{B}}$ ist linear: klar.
(b) Klar. □

Satz 5.11 *(Matrizen und lineare Abbildungen, 3. Teil)*
Es seien U und V endlich-dimensionale K-Vektorräume. Es sei $\mathcal{A} = (a_1, ..., a_n)$ eine Basis von U und $\mathcal{B} = (b_1, ..., b_m)$ eine Basis von V.

(a) (Bemerkung) Es sei $f : U \to V$ eine lineare Abbildung. Mit den Isomorphismen $l_{\mathcal{A}} : M(n \times 1, K) \to U$ und $l_{\mathcal{B}} : M(m \times 1, K) \to V$ induziert $f : U \to V$ eine lineare Abbildung

$$l_{\mathcal{B}}^{-1} \circ f \circ l_{\mathcal{A}} : M(n \times 1, K) \to M(m \times 1, K).$$

Sie ist gerade so definiert, daß das Diagramm

$$\begin{array}{ccccc} M(n\times 1,K) & \xrightarrow{l_{\mathcal{A}}} & U & , x\mapsto & \mathcal{A}\cdot x \\ \downarrow l_{\mathcal{B}}^{-1}\circ f\circ l_{\mathcal{A}} & & \downarrow f & & \\ M(m\times 1,K) & \xrightarrow{l_{\mathcal{B}}} & V & , y\mapsto & \mathcal{B}\cdot y, \end{array}$$

kommutiert, *d.h. beide Wege von links oben nach rechts unten geben die gleiche Abbildung.*
(Satz) Es ist

$$\mathrm{Mat}(l_{\mathcal{B}}^{-1}\circ f\circ l_{\mathcal{A}}) = M(\mathcal{B},f,\mathcal{A}).$$

Das heißt, daß die Abbildung f mit Hilfe von $l_{\mathcal{A}}$ und $l_{\mathcal{B}}$ gerade in die Matrixmultiplikation mit $M(\mathcal{B},f,\mathcal{A})$ übergeht. Die folgende Gleichung sagt dasselbe etwas anders,

$$f(\mathcal{A}\cdot x) = \mathcal{B}\cdot M(\mathcal{B},f,\mathcal{A})\cdot x.$$

(b) (Verallgemeinerung von Satz 5.5 (b)) Die Menge $\mathrm{Hom}_K(U,V) := \{f : U \to V \mid f \text{ linear}\}$ *ist ein K-Vektorraum, und die Abbildung*

$$\mathrm{Hom}_K(U,V)\to M(m\times n,K),\quad f\mapsto M(\mathcal{B},f,\mathcal{A}),$$

ist ein Isomorphismus von K-Vektorräumen

(c) (Verallgemeinerung von Satz 5.5 (c)) Ist W ein K-Vektorraum mit einer Basis $\mathcal{C}=(c_1,...,c_l)$ und sind $f: U\to V$ und $g: V\to W$ linear, so ist auch $g\circ f: U\to W$ linear (Satz 5.3 (h)), und es ist

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

(d) (Verallgemeinerung von Satz 5.5 (d)) Eine Abbildung $f: U\to V$ ist genau dann ein Isomorphismus, wenn die Matrix $M(\mathcal{B},f,\mathcal{A})$ invertierbar ist. Dann ist $M(\mathcal{B},f,\mathcal{A})^{-1} = M(\mathcal{A},f^{-1},\mathcal{B})$.

(e) Im Fall $U=V$ und $f=\mathrm{id}$ heißt $M(\mathcal{B},\mathrm{id},\mathcal{A}) =: M(\mathcal{B},\mathcal{A})$ ja Basiswechselmatrix (Definition 5.8 (b)). Sie ist invertierbar; die inverse Matrix ist $M(\mathcal{A},\mathcal{B})$.

(f) Die Menge $\mathrm{End}_K(V) := \mathrm{Hom}_K(V,V)$ *der Endomorphismen von V ist ein Ring und ein K-Vektorraum. Die Abbildung*

$$\mathrm{End}_K(V)\to M(m\times m,K),\quad f\mapsto M(\mathcal{B},f,\mathcal{B})$$

ist ein Isomorphismus von Ringen und K-Vektorräumen. Für $m\geq 2$ sind die Ringe nicht kommutativ.

(g) Die Menge $\mathrm{Aut}_K(V)$ *der Automorphismen von V ist eine Gruppe. Die Abbildung*

$$\mathrm{Aut}_K(V)\to GL(m,K),\quad f\mapsto M(\mathcal{B},f,\mathcal{B})$$

ist ein Isomorphismus von Gruppen (Def. von $GL(m,K)$ in Satz 4.13 (c)). Für $m\geq 2$ sind die Gruppen nicht kommutativ.

Beweis: (a) Mit der Notation $f(\mathcal{A}) = (f(a_1), ..., f(a_n))$ von Definition 5.8 ist

$$f(\mathcal{A} \cdot x) \overset{f \text{ linear}}{=} f(\mathcal{A}) \cdot x \overset{\text{Def. 5.8}}{=} \mathcal{B} \cdot M(\mathcal{B}, f, \mathcal{A}) \cdot x.$$

Das ist äquivalent zur Behauptung

$$l_{\mathcal{B}}^{-1} \circ f \circ l_{\mathcal{A}} = \text{ Matrixmultiplikation von links mit } M(\mathcal{B}, f, \mathcal{A}).$$

(b) Wegen (a) ist die Abbildung

$$\mathrm{Hom}_K(U, V) \to M(m \times n, K), \quad f \mapsto M(\mathcal{B}, f, \mathcal{A}),$$

die Komposition der beiden Abbildungen

$$\mathrm{Hom}_K(U, V) \to \mathrm{Hom}_K(M(n \times 1, K), M(m \times 1, K)), \quad f \mapsto l_{\mathcal{B}}^{-1} \circ f \circ l_{\mathcal{A}}$$

und

$$\mathrm{Hom}_K(M(n \times 1, K), M(m \times 1, K)) \to M(m \times n, K), \quad g \mapsto \mathrm{Mat}(g).$$

Die zweite ist ein Vektorraumisomorphismus nach Satz 5.5 (b). Die erste ist ein Vektorraumisomorphismus, weil $l_{\mathcal{A}}$ und $l_{\mathcal{B}}$ Vektorraumisomorphismen sind: sie ist bijektiv, denn ein Element $g \in \mathrm{Hom}_K(M(n \times 1, K), M(m \times 1, K))$ hat genau ein Urbild, nämlich $l_{\mathcal{B}} \circ g \circ l_{\mathcal{A}}^{-1}$. Die Linearität ist auch klar.
(c) Es ist

$$\begin{aligned} M(\mathcal{C}, g \circ f, \mathcal{A}) &= \mathrm{Mat}(l_{\mathcal{C}}^{-1} \circ g \circ f \circ l_{\mathcal{A}}) \quad \text{(mit (a))} \\ &= \mathrm{Mat}(l_{\mathcal{C}}^{-1} \circ g \circ l_{\mathcal{B}}) \cdot \mathrm{Mat}(l_{\mathcal{B}}^{-1} \circ f \circ l_{\mathcal{A}}) \quad \text{(Satz 5.5 (c))} \\ &= M(\mathcal{C}, g, \mathcal{B}) \cdot M(\mathcal{B}, f, \mathcal{A}) \quad \text{(mit (a))}. \end{aligned}$$

(d) Wie Satz 5.5 (d) (bzw. mit Satz 5.5 (d) und Satz 5.11 (a)).
(e) folgt aus (d).
(f) folgt aus (b) und (c).
(g) folgt aus (c) und (d). □

Bemerkungen 5.12 (i) Es seien $U, V, f : U \to V, \mathcal{A}$ und $\mathcal{B}$ wie in Satz 5.11. Es sei $\widetilde{\mathcal{A}}$ eine andere Basis von U und $\widetilde{\mathcal{B}}$ eine andere Basis von V. Dann ist wegen Satz 5.11 (c)

$$M(\widetilde{\mathcal{B}}, f, \widetilde{\mathcal{A}}) = M(\widetilde{\mathcal{B}}, \mathcal{B}) \cdot M(\mathcal{B}, f, \mathcal{A}) \cdot M(\mathcal{A}, \widetilde{\mathcal{A}}).$$

Diese Formel zeigt, wie $M(\mathcal{B}, f, \mathcal{A})$ von $\mathcal{B}$ und $\mathcal{A}$ abhängt, bzw. wie sich die Matrix transformiert, wenn man $\mathcal{B}$ und $\mathcal{A}$ ändert.
Die Abhängigkeit von $M(\mathcal{B}, f, \mathcal{A})$ von f ist linear (Satz 5.11 (b)).

(ii) Satz 5.13 (a) zeigt, daß man $M(\mathcal{B}, f, \mathcal{A})$ durch Wahl geeigneter Basen $\mathcal{B}$ und $\mathcal{A}$ auf eine sehr einfache Gestalt bringen kann. Aus $M(\mathcal{B}, f, \mathcal{A})$ allein kann man nur $\mathrm{rang}(f)$ ablesen.

(iii) Viel reicher und interessanter wird die Situation, wenn f eine Endomorphismus ist und wenn man nur die Matrizen $M(\mathcal{B}, f, \mathcal{B})$ ansieht. Das kommt in Kapitel 8 und in LA IIa.

Satz 5.13 *(a) Zu jeder linearen Abbildung $f : U \to V$ von endlich-dimensionalen K-Vektorräumen gibt es eine Basis $\mathcal{A} = (a_1, ..., a_n)$ von U und eine Basis $\mathcal{B} = (b_1, ..., b_m)$ von V, so daß für $k := \operatorname{rang} f$ gilt:*

$$M(\mathcal{B}, f, \mathcal{A}) = \begin{pmatrix} (\delta_{ij})_{i,j=1,...,k} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdot & 0 \\ 0 & \ddots & & \\ \cdot & & 1 & 0 \\ 0 & \cdot & 0 & 0 \end{pmatrix};$$

d.h. die Einheitsmatrix in den ersten k Zeilen und Spalten und 0 außerhalb.
(b) (Matrix-Version von a)) Zu jeder Matrix $C \in M(m \times n, K)$ gibt es Matrizen $A \in GL(n, K)$ und $B \in GL(m, K)$, so daß mit $k :=$ Spaltenrang(C) gilt:

$$B \cdot C \cdot A = \begin{pmatrix} (\delta_{ij})_{i,j=1,...,k} & 0 \\ 0 & 0 \end{pmatrix}.$$

(c) Durch Links- oder Rechtsmultiplikation mit invertierbaren Matrizen ändern sich Zeilen- und Spaltenrang einer gegebenen Matrix nicht.
(d) (=Satz 4.2) Für jede Matrix $C \in M(m \times n, K)$ ist

$$\text{Zeilenrang}(C) = \text{ Spaltenrang}(C).$$

Beweis: (a) Man wählt $a_1, ..., a_k \in U$ so, daß $(b_1, ..., b_k) := (f(a_1), ..., f(a_k))$ eine Basis von $f(U) \subset V$ ist.
Man ergänzt diese Basis zu einer Basis $\mathcal{B} = (b_1, ..., b_m)$ von V (Satz 3.19).
Wegen Satz 5.3 (g) ist $\dim \ker(f) = \dim U - \operatorname{rang}(f) = n - k$. Man wählt eine Basis $(a_{k+1}, ..., a_n)$ von $\ker(f)$.
Aus dem Beweis von Satz 5.3 (g) folgt, daß $(a_1, ..., a_n)$ eine Basis von U ist. Dann sieht $M(\mathcal{B}, f, \mathcal{A})$ aus wie gewünscht.
(b) [Es gibt einen Beweis, der nur Zeilen- und Spaltenumformungen benutzt; der hier gegebene Beweis benutzt (a).]
Sei $U := M(n \times 1, K)$, $\widetilde{\mathcal{A}}$ seine Standardbasis, $V := M(m \times 1, K)$, $\widetilde{\mathcal{B}}$ seine Standardbasis, $l_C : U \to V$ die Linksmultiplikation mit C. Dann ist

$$M(\widetilde{\mathcal{B}}, l_C, \widetilde{\mathcal{A}}) = C$$

und

$$\text{Spaltenrang}(C) = \operatorname{rang}(l_C).$$

Laut (a) gibt es Basen $\mathcal{A}$ von U und $\mathcal{B}$ von V mit

$$M(\mathcal{B}, l_C, \mathcal{A}) = \begin{pmatrix} (\delta_{ij})_{i,j=1,...,k} & 0 \\ 0 & 0 \end{pmatrix}.$$

Mit den Bezeichnungen $A := M(\widetilde{\mathcal{A}}, \mathcal{A})$ und $B := M(\mathcal{B}, \widetilde{\mathcal{B}})$ für die Basiswechselmatrizen ist

$$M(\mathcal{B}, l_C, \mathcal{A}) = M(\mathcal{B}, \widetilde{\mathcal{B}}) \cdot M(\widetilde{\mathcal{B}}, l_C, \widetilde{\mathcal{A}}) \cdot M(\widetilde{\mathcal{A}}, \mathcal{A}) = B \cdot C \cdot A.$$

(c) **1. Schritt**: Sei $B \in GL(m,K)$ und $C \in M(m \times n, K)$. Die Zeilen von $B \cdot C$ sind Linearkombinationen der Zeilen von C, und die Zeilen von $C = B^{-1} \cdot (B \cdot C)$ sind Linearkombinationen der Zeilen von $B \cdot C$. Daher ist

$$\text{Zeilenrang}(B \cdot C) = \text{Zeilenrang}(C).$$

2. Schritt: Ist $f : W \to W$ ein Automorphismus eines Vektorraums W und ist $W_1 \subset W$ ein endlich-dimensionaler Untervektorraum, so ist $\dim W_1 = \dim f(W_1)$. Denn das Bild einer Basis von W_1 ist eine Basis von $f(W_1)$.
3. Schritt: Sei $C \in M(m \times n, K)$ mit den Spalten $w_1, ..., w_n \in M(m \times 1, K)$. Ist nun $B \in GL(m,K)$, so ist die Linksmultiplikation l_B ein Automorphismus von $M(m \times 1, K)$. Also ist

$$\begin{aligned} \text{Spaltenrang}(B \cdot C) &= \dim_K(\text{span}_K(B \cdot w_1, ..., B \cdot w_n)) \\ &= \dim_K(\text{span}_K(l_B(w_1), ..., l_B(w_n))) \\ &= \dim_K l_B(\text{span}_K(w_1, ..., w_n)) \\ &= \dim_K(\text{span}_K(w_1, ..., w_n)) = \text{ Spaltenrang}(C). \end{aligned}$$

4. Schritt: Sei $A \in GL(n,K)$ und C wie oben. Analog zum 1. Schritt zeigt man $\text{Spaltenrang}(C \cdot A) = \text{Spaltenrang}(C)$; analog zum 3. Schritt zeigt man $\text{Zeilenrang}(C \cdot A) = \text{Zeilenrang}(C)$.
(d) folgt aus (c) und (b), denn bei der Matrix $B \cdot C \cdot A$ in (b) sind Zeilenrang und Spaltenrang gleich. □

Satz 5.14 *(a) Ist I eine nichtleere Menge und K ein Körper, so ist die Menge*

$$\text{Abb}_{endlich}(I,K) := \{g : I \to K \mid \text{die Menge } \{j \in I \mid g(j) \neq 0\} \text{ ist endlich}\}$$

ein Untervektorraum von $\text{Abb}(I,K)$.
(b) (Verallgemeinerung von Lemma 5.10 (b) auf Vektorräume mit beliebig großen Basen) Sei V ein K-Vektorraum mit Basis $(v_i)_{i \in I}$. Die Abbildung

$$\begin{aligned} V &\to \text{Abb}_{endlich}(I,K) \\ \sum_{j \in J} \lambda_j v_j &\mapsto g, \quad \text{mit } g(j) := \begin{cases} \lambda_j & \text{für } j \in J, \\ 0 & \text{für } j \in I - J, \end{cases} \end{aligned}$$

ist ein Isomorphismus von Vektorräumen.

Beweis: Übung. □

6 Lineare Gleichungssysteme

In diesem Kapitel bezeichnet K irgendeinen Körper.

Definition 6.1 Ein *lineares Gleichungssystem* ist ein Gleichungssystem der Gestalt

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m. \end{aligned}$$

Hier sind $A = (a_{ij}) \in M(m \times n, K)$ und $b = (b_1, ..., b_m)^{tr} \in M(m \times 1, K)$ gegeben, und $x = (x_1, ..., x_n)^{tr}$ ist ein Spaltenvektor von "Unbestimmten". Man kann es kürzer schreiben, in der Form

$$A \cdot x = b.$$

Es heißt *inhomogenes lineares Gleichungssystem*, falls $b \neq 0$ ist, sonst *homogenes lineares Gleichungssystem*.
Einem inhomogenen linearen Gleichungssystem $A \cdot x = b$ ist das homogene lineare Gleichungssystem $A \cdot x = 0$ zugeordnet.
Man möchte die *Lösungsmengen*

$$\begin{aligned} & \text{Lös}(A, b) &:= \{x \in M(n \times 1, K) \mid A \cdot x = b\} \\ \text{und} \quad & \text{Lös}(A, 0) &:= \{x \in M(n \times 1, K) \mid A \cdot x = 0\} \end{aligned}$$

bestimmen.

Beispiel 6.2

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & 7 & 4 & 0 \\ 2 & 4 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

In der Matrix $(A\ b)$ trennen wir b von A durch einen senkrechten Strich, um anzuzeigen, daß b die rechte Seite des linearen Gleichungssystems ist. Bei Zeilenumformungen ändert sich $\text{Lös}(A, b)$ nicht.

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 3 & 7 & 4 & 0 & 1 \\ 2 & 4 & 1 & 0 & 1 \end{array}\right) \overset{\text{Z.umf.}}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 0 & 1 & 4 & -3 & -2 \\ 0 & 0 & 1 & -2 & -1 \end{array}\right)$$

$$\overset{\text{Z.umf.}}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -9 & -3 \\ 0 & 1 & 0 & 5 & 2 \\ 0 & 0 & 1 & -2 & -1 \end{array}\right)$$

Bei den Lösungen ist $x_4 =: t$ beliebig, und x_1, x_2, x_3 sind dann eindeutig. $(x_1, x_2, x_3, x_4)^{tr}$ ist genau dann eine Lösung, wenn

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 9t-3 \\ -5t+2 \\ 2t-1 \\ t \end{pmatrix} = \begin{pmatrix} 9 \\ -5 \\ 2 \\ 1 \end{pmatrix} \cdot t + \begin{pmatrix} -3 \\ 2 \\ -1 \\ 0 \end{pmatrix}$$

ist mit $t \in K$ beliebig.

Satz 6.3 *Sei $A \cdot x = b$ ein lineares Gleichungssystem wie in Definition 6.1.*
(a) $\text{Lös}(A, 0)$ *ist ein Untervektorraum von $M(n \times 1, K)$ der Dimension*

$$\dim_K \text{Lös}(A, 0) = n - \operatorname{rang} A.$$

(b) Sei $U := \operatorname{span}_K(\text{Spalten von } A) \subset M(m \times 1, K)$.
1. Fall, *$b \notin U$***:** *dann ist* $\text{Lös}(A, b) = \emptyset$.
2. Fall, *$b \in U$***:** *dann ist* $\text{Lös}(A, b)$ *nicht leer und*

$$\text{Lös}(A, b) = \text{Lös}(A, 0) + v,$$

wobei $v \in \text{Lös}(A, b)$ eine beliebige Lösung ist.
(c) Ist $B \in GL(m, K)$, so ist $\text{Lös}(B \cdot A, B \cdot b) = \text{Lös}(A, b)$. *Mit anderen Worten: Bei Zeilenumformungen der Matrix $(A\ b) \in M(m \times (n+1), K)$ ändert sich die Lösungsmenge des Gleichungssytems nicht.*

Beweis: (a) Die Multiplikation von links mit A ist eine lineare Abbildung

$$l_A : M(m \times 1, K) \to M(n \times 1, K), \quad x \mapsto A \cdot x.$$

Es ist $\text{Lös}(A, 0) = \ker(l_A)$, also ein Untervektorraum (Satz 5.3 (c)). Nach Satz 5.3 (g) ist

$$\dim \ker(l_A) = n - \operatorname{rang} l_A = n - \text{Spaltenrang}(A) = n - \operatorname{rang} A.$$

(b) $A \cdot v = b$ sagt gerade, daß b eine Linearkombination der Spalten von A ist, mit Koeffizienten $v_1, ..., v_n$. Daher ist $b \in U$ natürlich äquivalent zur Lösbarkeit von $A \cdot x = b$.
Ist $v \in \text{Lös}(A, b)$, so ist

$$\text{Lös}(A, b) = \{x \mid l_A(x) = b\} = v + \ker l_a,$$

denn l_A ist linear.
(c) Klar. □

Beispiel 6.4 Hier sollen für jedes $a \in K$ die Lösungen des Gleichungssystems $A \cdot x = b$ bestimmt werden, wo

$$(A|b) = \left(\begin{array}{cc|c} 1 & -a & -1 \\ a+1 & 0 & -1 \end{array} \right).$$

Eine Zeilenumformung gibt

$$\left(\begin{array}{cc|c} 1 & -a & -1 \\ 0 & a(a+1) & a \end{array} \right).$$

1. Fall, $a = -1$: $\text{Lös}(A, b) = \emptyset$.

2. Fall, $a = 0$: $\text{Lös}(A, b) = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot t + \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mid t \in K \right\}$.

3. Fall, $a \notin \{0, -1\}$: $\text{Lös}(A, b) = \left\{ \begin{pmatrix} \frac{-1}{a+1} \\ \frac{1}{a+1} \end{pmatrix} \right\}$.

(Bei $K = \mathbb{F}_2$ ist $K = \{0, -1\}$, und der 3. Fall tritt nicht ein.)

Bemerkungen 6.5 (Methoden zur Lösung eines linearen Gleichungssystems)
(i) Wenn $m = n$ ist und A invertierbar ist, ist l_A ein Isomorphismus, und zu jedem b gibt es genau eine Lösung von $A \cdot x = b$. Sie ist einfach $A^{-1} \cdot b$. Ein Weg, A^{-1} zu berechnen, ist in Bemerkung 4.14 beschrieben. Ein anderer kommt in Kapitel 7.
Man muß aber nicht wirklich A^{-1} ausrechnen, um $A^{-1} \cdot b$ auszurechnen. Es reicht, die $(n \times (n+1))$-Matrix $(A\ b)$ durch Zeilenumformungen in die Gestalt $(E_n\ \widetilde{b})$ zu bringen. Dann ist $\widetilde{b} = A^{-1} \cdot b$.

(ii) Im allgemeinen Fall bestimmt man $\text{Lös}(A, b)$ so: Mit dem Gauß-Algorithmus bringt man die Matrix $(A\ b) \in M(m \times (n+1), K)$ in Zeilenstufenform $(A^{(1)}\ b^{(1)})$. Es sei $k \in \mathbb{N}_0$ so, daß genau die ersten k Zeilen von $A^{(1)}$ nicht verschwinden.

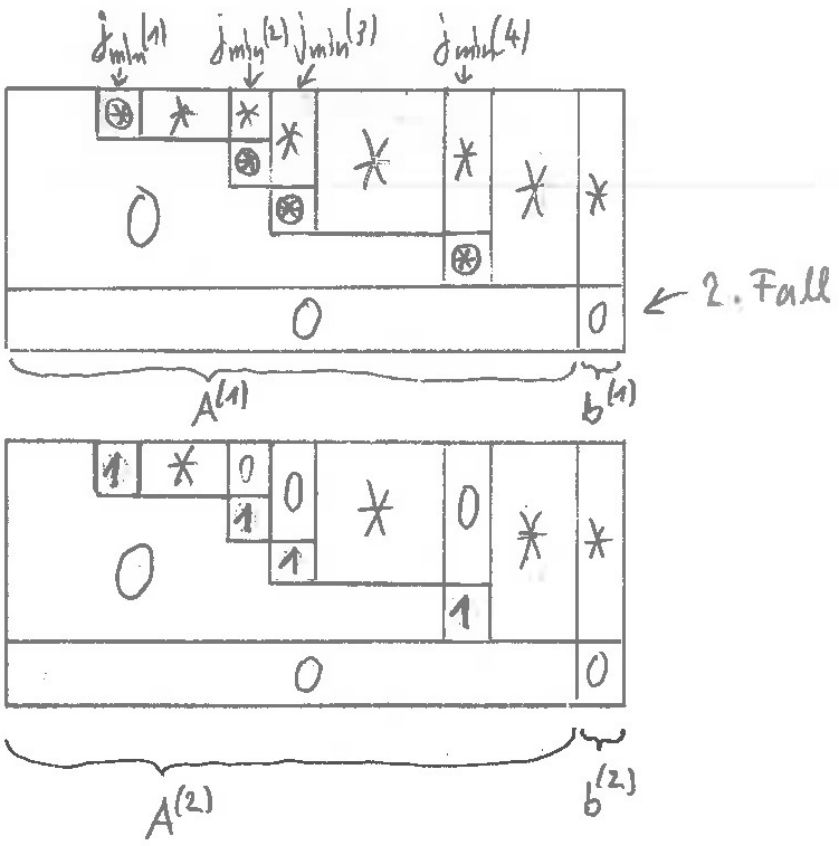
1. Fall: Sei $(b^{(1)}_{k+1}, ..., b^{(1)}_n) \neq 0$. Dann ist $\text{Lös}(A, b) = \emptyset$.

2. Fall: Sei $(b^{(1)}_{k+1}, ..., b^{(1)}_n) = 0$. Für $i \leq k$ sei

$$j_{min}(i) := \min(j \mid a^{(1)}_{ij} \neq 0),$$

Es ist $1 \leq j_{min}(1) < ... < j_{min}(k) \leq n$. Es sei

$$J := \{1, ..., n\} - \{j_{min}(1), ..., j_{min}(k)\}.$$



Mit Zeilenumformungen vom Typ II löscht man alle Einträge in den Spalten mit Spaltenindices $j_{min}(i)$, $i = 1, ..., k$ außer dem Eintrag $a^{(1)}_{ij_{min}(i)}$.

Mit Zeilenumformungen vom Typ I normiert man alle Einträge $a^{(1)}_{ij_{min}(i)}$ zu 1.

Man erhält eine Matrix $(A^{(2)}\ b^{(2)})$. Die i-te Zeile des neuen Gleichungssystems ist

$$x_{j_{min}(i)} + \sum_{j \in J, j > j_{min}(i)} a^{(2)}_{ij} x_j = b^{(2)}_i.$$

Der Lösungsraum ist nun leicht beschreibbar:

$$\begin{aligned} \text{Lös}(A,b) = \{x \in M(n \times 1, K) \quad | \quad & x_j \text{ für } j \in J \text{ ist beliebig,} \\ & x_{j_{min}(i)} \text{ für } i = 1, ..., k \text{ ist durch} \\ & \text{die Gleichung oben bestimmt}\}. \end{aligned}$$

Eine besonders schöne Lösung $v_{inhom} \in \text{Lös}(A^{(2)}, b^{(2)})$ des inhomogenen Gleichungssystems erhält man nun, wenn man $x_j := 0$ für alle $j \in J$ setzt, sie ist

$$v_{inhom} = \sum_{i=1}^{k} b_i^{(2)} \cdot e_{j_{min}(i)} = \begin{pmatrix} 0 \\ \vdots \\ b_1^{(2)} \\ \vdots \\ b_k^{(2)} \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow j_{min}(1) \\ \\ \leftarrow j_{min}(k) \\ \\ \end{matrix}$$

Und eine besonders schöne Basis $v_j, j \in J$, von Lösungen in $\text{Lös}(A^{(2)}, 0)$ (d.h. des homogenen Gleichungssystems) erhält man, indem man für v_j $x_j := -1$ setzt und $x_k := 0$ für $k \in J - \{j\}$, man erhält für $j \in J$

$$v_j = -e_j + \sum_{i=1}^{k} a_{ij}^{(2)} \cdot e_{j_{min}(i)} = \begin{pmatrix} 0 \\ \vdots \\ -1 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \leftarrow j + \begin{pmatrix} 0 \\ \vdots \\ a_{1j}^{(2)} \\ \vdots \\ a_{kj}^{(2)} \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow j_{min}(1) \\ \\ \leftarrow j_{min}(k) \\ \\ \end{matrix}.$$

In Prosa und etwas vage: Man erhält v_{inhom} aus dem Spaltenvektor b, indem man die ersten k Einträge von b auf die Stellen $j_{min}(1), ..., j_{min}(k)$ eines Spaltenvektors der Länge n verteilt und die anderen Stellen als 0 ansetzt.
Man erhält $v_j, j \in J$, als Summe des Spaltenvektors $-e_j$ und eines Spaltenvektors, den man analog zu v_{inhom} durch Verteilen der ersten k Einträge der j-ten Spalte von $A^{(2)}$ auf die Stellen $j_{min}(1), ..., j_{min}(k)$ erhält.
(Man muss mit all den Indices aufpassen, aber man braucht nicht mehr zu rechnen für die Bestimmung von v_{inhom} und $v_j, j \in J$).
Vorsicht: Man *muß* nicht ganz genau so rechnen. In einfachen Fällen reicht es oft, die Matrix A irgendwie auf Zeilenstufenform zu bringen.

Beispiel 6.6

$$(A|b) = (A^{(2)}|b^{(2)}) = \left(\begin{array}{ccccccc|c} 1 & 2 & 0 & 3 & 5 & 0 & 7 & 10 \\ 0 & 0 & 1 & 4 & 6 & 0 & 8 & 11 \\ 0 & 0 & 0 & 0 & 0 & 1 & 9 & 12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \in M(m \times (n+1), \mathbb{Q}),$$

mit $m = 5, n = 7$ und $k := \operatorname{rang} A = 3$, also $\dim \text{Lös}(A, 0) = n - k = 4$. Es sind $(j_{min}(1), j_{min}(2), j_{min}(3)) = (1, 3, 6)$ und $J := \{1, ..., n\} - \{j_{min}(1), j_{min}(2), j_{min}(3)\} =$

$\{2,4,5,7\}$. Hier sind die Lösung $v_{inhom} \in \text{Lös}(A,b)$ und die Basis $v_j \in \text{Lös}(A,0)$, $j \in J$, von $\text{Lös}(A,0)$

$$v_{inhom} = \begin{pmatrix} 10 \\ 0 \\ 11 \\ 0 \\ 0 \\ 12 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 3 \\ 0 \\ 4 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 5 \\ 0 \\ 6 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad v_7 = \begin{pmatrix} 7 \\ 0 \\ 8 \\ 0 \\ 0 \\ 9 \\ -1 \end{pmatrix}.$$

Es ist

$$\begin{aligned} \text{Lös}(A,b) &= \{v_{inhom} + t_1 v_2 + t_2 v_4 + t_3 v_5 + t_4 v_7 \mid t_1,t_2,t_3,t_4 \in \mathbb{Q}\} \\ &= v_{inhom} + \text{span}(v_2, v_4, v_5, v_7) = v_{inhom} + \text{Lös}(A,0). \end{aligned}$$

Beispiel 6.7 Oft ist es schwerer, aus einer Textaufgabe ein lineares Gleichungssystem herauszudestillieren, als es zu lösen. Ein Beispiel:

"A famous puzzle of Sam Loyd: The combined ages of Mary and Ann are 44 years, and Mary is twice as old as Ann was when Mary was half as old as Ann will be when Ann ist three times as old as Mary was when Mary was three times as old as Ann. How old is Ann?"

1. Schritt: Geeignete Variablen einführen.
The combined ages of Mary (Alter jetzt: m) and Ann (Alter jetzt: a) are 44 years, and Mary is twice as old as Ann was (Alter zu dem Zeitpunkt: $a - t_1$) when Mary was (A.z.d.Z.: $m - t_1$) half as old as Ann will be (A.z.d.Z.: $a + t_2$) when Ann ist three times as old as Mary was (A.z.d.Z.: $m - t_3$) when Mary was three times as old as Ann (A.z.d.Z.: $a - t_3$).

2. Schritt: Die Bedingungen als lineare Gleichungen schreiben.

$$\begin{aligned} &m + a = 44, \quad m = 2(a - t_1), \quad m - t_1 = \frac{1}{2}(a + t_2), \\ &a + t_2 = 3(m - t_3), \quad m - t_3 = 3(a - t_3). \end{aligned}$$

3. Schritt: Das lineare Gleichungssystem lösen. Angesichts seiner Gestalt wird es hier durch Zeilenumformungen von unten nach oben vereinfacht.

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 44 \\ 1 & -2 & 2 & 0 & 0 & 0 \\ 1 & -\frac{1}{2} & -1 & -\frac{1}{2} & 0 & 0 \\ -3 & 1 & 0 & 1 & 3 & 0 \\ 1 & -3 & 0 & 0 & 2 & 0 \end{array}\right) \to \left(\begin{array}{ccccc|c} 0 & \frac{8}{3} & 0 & 0 & 0 & 44 \\ -\frac{3}{2} & \frac{5}{2} & 0 & 0 & 0 & 0 \\ -\frac{5}{4} & \frac{9}{4} & -1 & 0 & 0 & 0 \\ -\frac{9}{2} & \frac{11}{2} & 0 & 1 & 0 & 0 \\ 1 & -3 & 0 & 0 & 2 & 0 \end{array}\right),$$

die eindeutige Lösung ist

$$(m, a, t_1, t_2, t_3)^{tr} = (\frac{55}{2}, \frac{33}{2}, \frac{11}{4}, 33, 11)^{tr}.$$

7 Determinanten

In diesem Kapitel bezeichnet K irgendeinen Körper und R irgendeinen kommutativen Ring mit Eins.
Eine Motivation für Determinanten besteht in der Beziehung zum Volumen eines Parallelotops, siehe Bemerkung 7.20.

Definition 7.1 (Leibniz-Formel) Die *Determinante* $\det A$ einer quadratischen Matrix $A = (a_{ij}) \in M(n \times n, R)$ (mit $n \geq 1$) ist

$$\det A := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}.$$

Notation: manchmal schreibt man $|A| = \det A$ (aber nur bei $n \geq 2$).

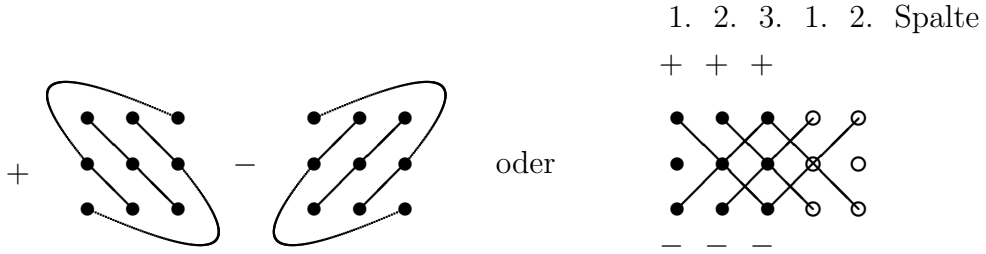
Beispiele 7.2 (a) $n = 2$: $S_2 = \{\mathrm{id}, (1\ 2)\}$,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(b) $n = 3$: $S_3 = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2), (1\ 2), (1\ 3), (2\ 3)\}$,

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}. \end{aligned}$$

Sarrussche Regel:



(c) $A = (a_{ij}) \in M(n \times n, R)$ obere Dreiecksmatrix, d.h. $a_{ij} = 0$ für $i > j$:

$$\det A = a_{11} \cdot \ldots \cdot a_{nn},$$

denn für $\sigma \in S_n - \{id\}$ gibt es ein $i \in \{1, \ldots, n\}$ mit $i > \sigma(i)$.

(d) (Verallgemeinerung von (c)) Sei $A = (a_{ij}) \in M((n+m) \times (n+m), R)$ eine Matrix der Gestalt

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

mit $B \in M(n \times n, R)$, $C \in M(n \times m, R)$, $D \in M(m \times m, R)$ und 0 die Nullmatrix in $M(m \times n, R)$. Dann ist

$$\det A = \det B \cdot \det D.$$

Beweis: Die Null links unten in A besagt $a_{ij} = 0$, falls $i > n, j \leq n$. Daher verschwinden in der Leibniz-Formel alle Summanden für $\sigma \in S_n$ mit $\sigma(i) \leq n$ für irgendein $i > n$.
Es bleiben nur die, für die gilt: σ bildet $\{n+1, ..., n+m\}$ auf sich ab (also bijektiv), und bildet daher auch $\{1, ..., n\}$ auf sich ab. Solche σ lassen sich eindeutig schreiben als $\sigma = \sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$ mit

$$\sigma_1|_{\{n+1,...,n+m\}} = \mathrm{id} \quad \text{und} \quad \sigma_2|_{\{1,...,n\}} = \mathrm{id}\,.$$

Die Summe in der Leibniz-Formel läßt sich aufspalten in eine Doppelsumme, über die möglichen σ_1 und über die möglichen σ_2. Es ist auch

$$\begin{aligned} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{(n+m)\sigma(n+m)} &= \mathrm{sign}(\sigma_1) \cdot a_{1\sigma_1(1)} \cdot ... \cdot a_{n\sigma_1(n)} \\ &\cdot \quad \mathrm{sign}(\sigma_2) \cdot a_{(n+1)\sigma_2(n+1)} \cdot ... \cdot a_{(n+m)\sigma_2(n+m)} \end{aligned}$$

Das zeigt $\det A = \det B \cdot \det D$. □

Definition 7.3 Eine Abbildung $\delta : M(n \times n, R) \to R$ heißt

(1) *multilinear* bezüglich der Zeilen, falls gilt:

(a) Entsteht A' aus A durch Multiplikation einer Zeile mit $\lambda \in R$, so ist

$$\delta(A') = \lambda \cdot \delta(A).$$

(b) Unterscheiden sich A, A', A'' nur in der i-ten Zeile und ist die i-te Zeile von A die Summe der i-ten Zeilen von A' und A'', so ist

$$\delta(A) = \delta(A') + \delta(A'').$$

(2) *alternierend* bezüglich der Zeilen, falls gilt:

Ist A eine Matrix mit zwei gleichen Zeilen, so ist $\delta(A) = 0$.

(3) *schiefsymmetrisch* bezüglich der Zeilen, falls gilt:
Entsteht A' aus A durch Vertauschen von zwei Zeilen, so ist

$$\delta(A') = -\delta(A).$$

(4) *normiert*, falls gilt:

$$\delta(E_n) = 1.$$

Lemma 7.4 *Sei $\delta : M(n \times n, R) \to R$ eine Abbildung.*
(a) Erfüllt sie (1) und (2), so auch (3).
(b) Ist $2 = 1_R + 1_R \in R$ in R invertierbar (im Fall $R = K$ bedeutet das, daß $\mathrm{char}(K) \neq 2$ sein muß) und erfüllt δ (1) und (3), so auch (2).

Beweis: Es sei $1 \leq k < l \leq n$ und $A(x, y)$ die Matrix mit i-ter Zeile $(a_{i1}, ..., a_{in})$ für $i \notin \{k, l\}$, k-ter Zeile $x = (x_1, ..., x_n)$ und l-ter Zeile $y = (y_1, ..., y_n)$.
(a) Die folgende Gleichung zeigt (a),

$$\begin{aligned} 0 &= \delta(A(x+y, x+y)) \\ &= \delta(A(x,x)) + \delta(A(x,y)) + \delta(A(y,x)) + \delta(A(y,y)) \\ &= \delta(A(x,y)) + \delta(A(y,x)). \end{aligned}$$

(b) Die folgende Gleichung zeigt (b),

$$\delta(A(x,x)) = -\delta(A(x,x)), \text{ also } \quad 0 = 2 \cdot \delta(A(x,x)).$$

□

Satz 7.5 *Die Abbildung* $\det : M(n \times n, R) \to R$ *ist multilinear bezüglich der Zeilen, alternierend bezüglich der Zeilen und normiert.*

Beweis: (1) (a)

$$\begin{aligned} \det A' &= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot (\lambda a_{i\sigma(i)}) \cdot ... \cdot a_{n\sigma(n)} \\ &= \lambda \cdot \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{i\sigma(i)} \cdot ... \cdot a_{n\sigma(n)} \\ &= \lambda \cdot \det A. \end{aligned}$$

(1) (b)

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \cdot ... \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} ... + \sum_{\sigma \in S_n} ... = \det A' + \det A''. \end{aligned}$$

(2) Die k-te und die l-te Zeile von $A = (a_{ij})$ seien gleich (mit $k \neq l$ natürlich), d.h. $a_{ki} = a_{li}$ für alle i. Es sei $\tau := (k\ l) \in S_n$. Es ist $S_n = A_n \cup A_n\tau$.

$$\begin{aligned} \det A &= \sum_{\sigma \in A_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{k\sigma(k)} \cdot ... \cdot a_{l\sigma(l)} \cdot ... \cdot a_{n\sigma(n)} \\ &+ \sum_{\sigma \in A_n} \mathrm{sign}(\sigma \circ \tau) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{k\sigma\circ\tau(k)} \cdot ... \cdot a_{l\sigma\circ\tau(l)} \cdot ... \cdot a_{n\sigma(n)} \\ &= 0, \end{aligned}$$

denn

$$\operatorname{sign}(\sigma) = +1, \qquad \operatorname{sign}(\sigma \circ \tau) = (-1) \cdot \operatorname{sign}(\sigma) = -1,$$

und

$$a_{k\sigma\circ\tau(k))} = a_{k\sigma(l)} = a_{l\sigma(l)} \quad \text{und} \quad a_{l\sigma\circ\tau(l)} = a_{l\sigma(k)} = a_{k\sigma(k)}.$$

(4) Klar nach Beispiel 7.2 (c). □

Bemerkungen 7.6 (i) Aus Satz 7.5 und Lemma 7.4 folgt, daß det in folgender Weise mit den elementaren Zeilenumformungen $Z_I(\lambda; i), Z_{II}(\lambda; i, j)$ und $Z_{III}(i, j) : M(n \times n, R) \to M(n \times n, R)$ von Definition 4.4 verträglich ist:

$$\begin{aligned} \det(Z_I(\lambda; i)(A)) &= \lambda \cdot \det A, \\ \det(Z_{II}(\lambda; i, j)(A)) &= \det A, \\ \det(Z_{III}(i, j)(A)) &= -\det A. \end{aligned}$$

(ii) Im Fall $R = K$ (und bei günstigen Koeffizienten auch in anderen Fällen) kann man A mit elementaren Zeilenumformungen der Typen Z_{II} und Z_{III} auf obere Dreiecksgestalt bringen (Gauß-Algorithmus, Satz 7.4). Z_{II} ändert dabei nichts an der Determinante, Z_{III} ändert das Vorzeichen. Danach wird die Berechnung von $\det A$ wegen Beispiel 7.2 (c) trivial.

(iii) **Methoden zur Berechnung von** det**:**

- **Gauß-Algorithmus:** Die Methode in (ii) ist nur bei $R = K$ anwendbar (aber das ist meistens der Fall), oder wenn man Glück mit den Koeffizienten hat. Aber dann ist sie fast immer die schnellste.

- **Leibniz-Formel:** Nur bei $n = 2$ oder $n = 3$ (Beispiel 7.2) schnell. Für großes n ist $n!$ horrend groß.

- **Laplacescher Entwicklungssatz:** Satz 7.15; bei $n = 3$ okay; sonst nur gut, wenn in einer Zeile oder Spalte viele Einträge Null sind.

(iv) Die Leibniz-Formel und der Laplacesche Entwicklungssatz sind aber für theoretische Aussagen über det wichtig (die Leibniz-Formel zum Beispiel zur Definition von det).
(v) Im folgenden Satz muß man Koeffizienten in einem Körper K betrachten, denn nur dann ist $\operatorname{rang}(A)$ wohldefiniert.

Satz 7.7 *Sei $A \in M(n \times n, K)$ eine Matrix mit Koeffizienten in einem Körper K. Dann gilt:*

$$\det(A) \neq 0 \iff \operatorname{rang}(A) = n \iff A \text{ ist invertierbar.}$$

Beweis: Spezialfall: Für obere Dreiecksmatrizen gelten die Äquivalenzen oben wegen Beispiel 7.2 (c), Lemma 4.12 und Satz 4.13.
Allgemeiner Fall: Durch elementare Zeilenumformungen der Typen Z_{II} und Z_{III} läßt sich eine Matrix auf obere Dreiecksgestalt bringen. Ihre Determinante ändert dabei höchstens das Vorzeichen. □

Satz 7.8 *(a) Für $A \in M(n \times n, R)$ gilt*

$$\det A = \det A^{tr}.$$

(b) Daher ist det *multilinear und alternierend auch bezüglich der Spalten, und Formeln analog zu denen in Bemerkung 7.6 (i) gelten für die elementaren Spaltenumformungen der Typen S_I, S_{II} und S_{III} (Bemerkung 4.8 (ii)).*

Beweis: (a) Die folgende Rechnung benutzt drei Aussagen:

(i) Weil σ eine Bijektion auf $\{1, ..., n\}$ ist, ist

$$\{(\sigma(1), 1), ..., (\sigma(n), n)\} = \{(1, \sigma^{-1}(1)), ..., (n, \sigma^{-1}(n))\}.$$

(ii) Wenn σ die Menge S_n durchläuft, durchläuft auch σ^{-1} sie, denn die Abbildung $S_n \to S_n, \quad \sigma \mapsto \sigma^{-1}$, ist bijektiv.

(iii) $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$.

$$\begin{aligned} \det A^{tr} &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{\sigma(1),1} \cdot ... \cdot a_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdot ... \cdot a_{n\sigma^{-1}(n)} = \det A. \end{aligned}$$

(b) Klar. □

Beispiel 7.9 Daher kann man Matrizen mit elementaren Zeilen- **und** Spaltenumformungen vereinfachen, wenn man ihre Determinanten ausrechnen will. Im folgenden Beispiel wurden erst untereinanderstehende Zeilen subtrahiert (von oben nach unten); dann wurde die erste Spalte zu allen anderen addiert.

$$\begin{vmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 5 & 4 & 3 & 2 \\ 3 & 4 & 5 & 4 & 3 \\ 2 & 3 & 4 & 5 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 2 & 2 & 0 & 0 \\ 1 & 2 & 2 & 2 & 0 \\ 1 & 3 & 4 & 5 & 6 \end{vmatrix} = 48.$$

Satz 7.10 (Weierstraß-Axiome)
Ist $\delta : M(n \times n, R) \to R$ multilinear und alternierend bezüglich der Zeilen und normiert, so ist $\delta =$ det*. Das heißt,* det *ist durch diese Eigenschaften (=Weierstraß-Axiome) eindeutig charakterisiert.*

Beweis (eventuell nicht in der Vorlesung): Sei $\delta : M(n \times n, R) \to R$ multilinear und alternierend. Mit $e_j = (0, ..., 0, 1, 0, ..., 0) = (\delta_{jk})_{k=1,...,n}$ (für $j = 1, ..., n$) wird der Zeilenvektor mit einer Eins an der j-ten Stelle und Nullen sonst bezeichnet.

1. Teil: Für $\sigma : \{1, ..., n\} \to \{1, ..., n\}$ sei $P_\sigma \in M(n \times n, R)$ die Matrix mit

$$(P_\sigma)_{ij} := \delta_{\sigma(i),j}.$$

Ihre i-te Zeile ist $e_{\sigma(i)}$. Sind $1 \leq k < l \leq n$ und ist $\tau = (k\ l)$ die Transposition, die k und l vertauscht, so ist

$$P_{\sigma\circ\tau} = Z_{III}(k, l)(P_\sigma).$$

Wegen Lemma 7.4 ist δ schiefsymmetrisch. Daher ist

$$\delta(P_{\sigma\circ\tau}) = -\delta(P_\sigma),$$

Im Fall $\sigma \in S_n$ folgt

$$\delta(P_\sigma) = \text{sign}(\sigma) \cdot \delta(E_n).$$

Im Fall $\sigma \notin S_n$ (d.h. σ ist nicht bijektiv), gibt es zwei Zeilen von P_σ, die gleich sind; weil σ alternierend ist, ist dann

$$\delta(P_\sigma) = 0.$$

2. Teil: Sei $A = (a_{ij}) \in M(n \times n, R)$. Mit $a_i = (a_{ij})_{j=1,...,n} \in M(1 \times n, R)$ wird die i-te Zeile von A bezeichnet. Aus der Multilinearität von δ folgt

$$\begin{aligned}
\delta(A) &= \sum_{j_1=1}^{n} a_{1j_1} \cdot \delta\left(\begin{pmatrix} e_{j_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix}\right) = \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} a_{1j_1}a_{2j_2} \cdot \delta\left(\begin{pmatrix} e_{j_1} \\ e_{j_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix}\right) \\
&= \ ... = \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}...\sum_{j_n=1}^{n} a_{1j_1}a_{2j_2}...a_{nj_n} \cdot \delta\left(\begin{pmatrix} e_{j_1} \\ e_{j_2} \\ \vdots \\ e_{j_n} \end{pmatrix}\right) \\
&= \sum_{\sigma:\{1,...,n\}\to\{1,...,n\}} a_{1\sigma(1)}a_{2\sigma(2)}...a_{n\sigma(n)} \cdot \delta(P_\sigma)
\end{aligned}$$

Mit dem 1. Teil und der Leibniz-Formel folgt

$$\delta(A) = \det(A) \cdot \delta(E_n).$$

□

Satz 7.11 *(a)* **(Cauchy)** *Seien A und B in $M(n \times n, R)$. Es ist*

$$\det(A \cdot B) = \det A \cdot \det B.$$

(b) (Achtung, hier ist der Koeffizientenbereich ein Körper K.) Die Einschränkung von det *auf $GL(n, K)$ ist ein Gruppenhomomorphismus*

$$\det : (GL(n, K), \circ) \to (K - \{0\}, \cdot).$$

Insbesondere ist

$$\det(A^{-1}) = (\det A)^{-1}.$$

Beweis (eventuell nicht in der Vorlesung): (a) Sei $B \in M(n \times n, R)$ fest und $X \in M(n \times n, R)$ variabel. Die Abbildung

$$\delta : M(n \times n, R) \to M(n \times n, R), \quad X \mapsto \det(X \cdot B)$$

ist multilinear und alternierend bezüglich der Zeilen; denn elementare Zeilenumformungen vertauschen mit der Multiplikation von rechts mit B. Nach dem Beweis von Satz 7.10 ist daher

$$\det(A \cdot B) \overset{\text{Def. von } \delta}{=} \delta(A) = \det(A) \cdot \delta(E_n) = \det(A) \cdot \det(B).$$

(b) Das folgt aus (a). □

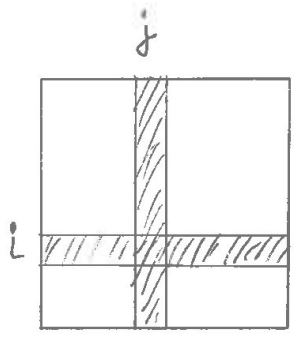
Definition/Lemma 7.12 *Die Menge*

$$SL(n, K) := \{A \in M(n \times n, K) \mid \det A = 1\}$$

ist eine Untergruppe (sogar ein Normalteiler*) von $(GL(n, K), \circ)$. Sie heißt* spezielle lineare Gruppe.

Beweis: Satz 7.11 (b) und Lemma 1.19. □

Definition 7.13 Sei $A \in M(n \times n, R)$ und $i, j \in \{1, ..., n\}$.
Mit $A[i, j]$ wird die $((n-1) \times (n-1))$-Matrix bezeichnet, die man aus A erhält, indem man die i-te Zeile und die j-te Spalte streicht.



Die Matrix $A^\sharp \in M(n \times n, R)$ hat die Einträge

$$(A^\sharp)_{ij} := (-1)^{i+j} \det A[j,i]$$

(Vorsicht; $[j,i]$, nicht $[i,j]$) und wird *Komplementärmatrix zur Matrix A* genannt (Satz 7.15 (c) zeigt, warum).

Die $n \times n$-Matrix mit Eintrag $(-1)^{i+j}$ an der Stelle (i,j) sieht so aus (die Einsen sind weggelassen):

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Beispiele 7.14 (i) Im Fall einer (2×2)-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, R)$ ist

$$A^\sharp = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

also

$$A \cdot A^\sharp = A^\sharp \cdot A = (ad - bc) \cdot E_2 = \det A \cdot E_2.$$

(ii) (Vgl. Bemerkung4.14)

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 9 \end{pmatrix}, \\ \det A &= \begin{vmatrix} 0 & 1 & 2 \\ 3 & 3 & 3 \\ 0 & 0 & 1 \end{vmatrix} = (-1) \cdot \begin{vmatrix} 3 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = -3. \end{aligned}$$

$$A^\sharp = \begin{pmatrix} \begin{vmatrix} 4 & 5 \\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \\ -\begin{vmatrix} 3 & 5 \\ 6 & 9 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 6 & 9 \end{vmatrix} & -\begin{vmatrix} 0 & 2 \\ 3 & 5 \end{vmatrix} \\ \begin{vmatrix} 3 & 4 \\ 6 & 7 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 6 & 7 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ 3 & 4 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 5 & -3 \\ 3 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}$$

Beispiel 4.14 zeigt

$$A^\sharp = (-3) \cdot A^{-1}, \quad \text{also } A \cdot A^\sharp = A^\sharp \cdot A = \det A \cdot E_3.$$

Satz 7.15 *Sei $A \in M(n \times n, R)$.*
*(a) (***Laplacescher Entwicklungssatz** *für die i-te Zeile) Sei $i \in \{1, ..., n\}$.*

$$\det A = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} \cdot \det A[i,j] = \sum_{j=1}^{n} a_{ij} \cdot (A^\sharp)_{ji}.$$

*(b) (***Laplacescher Entwicklungssatz** *für die j-te Spalte) Sei $j \in \{1, ..., n\}$.*

$$\det A = \sum_{i=1}^{n} a_{ij} \cdot (-1)^{i+j} \cdot \det A[i,j] = \sum_{i=1}^{n} a_{ij} \cdot (A^\sharp)_{ji}.$$

(c)

$$A \cdot A^\sharp = A^\sharp \cdot A = \det A \cdot E_n.$$

*(d) (***Cramersche Regel***, allgemeine Version für R und für $\det A$ beliebig)*
Sei $b \in M(n \times 1, R)$. Eine Lösung $y \in M(n \times 1, R)$ des linearen Gleichungssystems

$$A \cdot x = \det A \cdot b$$

ist offenbar

$$y = A^\sharp \cdot b.$$

Die Koeffizienten y_j dieser Lösung $y = (y_1, ..., y_n)^{tr}$ sind gegeben durch

$$y_j = \det B_j \quad \text{für } j = 1, ..., n,$$

wo B_j aus A entsteht, indem man die j-te Spalte von A durch b ersetzt.

Beweis nach Korollar 7.17.

Beispiel 7.16 (Vgl. Beispiel 7.14 (ii)) Laplace-Entwicklung nach der zweiten Zeile,

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 9 \end{pmatrix} &= -3 \cdot \begin{vmatrix} 1 & 2 \\ 7 & 9 \end{vmatrix} + 4 \cdot \begin{vmatrix} 0 & 2 \\ 6 & 9 \end{vmatrix} - 5 \cdot \begin{vmatrix} 0 & 1 \\ 6 & 7 \end{vmatrix} \\ &= -3 \cdot (-5) + 4 \cdot (-12) - 5 \cdot (-6) = -3. \end{aligned}$$

Korollar 7.17 ist eine unmittelbare Folgerung von Satz 7.15.

Korollar 7.17 *Sei $A \in M(n \times n, K)$ mit $\det A \neq 0$.*
(a)

$$A^{-1} = (\det A)^{-1} \cdot A^\sharp.$$

*(b) (***Cramersche Regel***, klassische Version für K und für $\det A \neq 0$)*
Sei $b \in M(n \times 1, K)$. Die eindeutige Lösung $y = (y_1, ..., y_n)^{tr} = A^{-1} \cdot b \in M(n \times 1, K)$ des linearen Gleichungssystems

$$A \cdot x = b$$

ist gegeben durch

$$y_j = \frac{\det B_j}{\det A} \quad \text{für } j = 1, ..., n,$$

wo B_j aus A entsteht, indem man die j-te Spalte von A durch b ersetzt.

Beweis von Satz 7.15: (a) Das zweite Gleichheitszeichen folgt aus der Definition von $A^\sharp$.
Zum ersten Gleichheitszeichen: Für einen Moment wird mit $\widetilde{A}[i,j] \in M(n \times n, R)$ die Matrix bezeichnet, die aus A entsteht, indem man die i-te Zeile von A durch die Zeile $e_j = (\delta_{jk})_{k=1,\dots,n}$ ersetzt. Aus der Multilinearität von det folgt

$$\det A = \sum_{j=1}^{n} a_{ij} \cdot \det \widetilde{A}[i,j].$$

Durch $i-1$ Zeilenvertauschungen und $j-1$ Spaltenvertauschungen erhält man aus $\widetilde{A}[i,j]$ eine Matrix, die so aussieht,

$$\begin{pmatrix} 1 & 0 & \cdots \\ * & & \\ \vdots & & A[i,j] \end{pmatrix}.$$

Nun zeigen Beispiel 7.2 (d) und die Eigenschaft, dass sich bei einer Zeilen- oder Spaltenvertauschung nur das Vorzeichen einer Determinante ändert,

$$\det \widetilde{A}[i,j] = (-1)^{i+j} \cdot \det A[i,j].$$

(b) Analog zu (a).

(c) Wegen (a) ist der Diagonaleintrag $(A \cdot A^\sharp)_{ii}$ von $A \cdot A^\sharp$

$$(A \cdot A^\sharp)_{ii} = \sum_{j=1}^{n} a_{ij} \cdot A^\sharp_{ji} = \det A = (\det A \cdot E_n)_{ii}.$$

Ersetzt man in (a) die Matrix A durch die Matrix, die man aus A erhält, indem man die i-te Zeile durch die k-te Zeile ersetzt (mit $k \neq i$), so hat sie zwei gleiche Zeilen. Also ist ihre Determinante 0, und (a) gibt $\overset{\text{a)}}{=}$ in

$$(A \cdot A^\sharp)_{ki} = \sum_{j=1}^{n} a_{kj} \cdot A^\sharp_{ji} \overset{\text{a)}}{=} 0 = (\det A \cdot E_n)_{ki}.$$

Also ist $A \cdot A^\sharp = \det A \cdot E_n$. Analog zeigt man $A^\sharp \cdot A = \det A \cdot E_n$.

(d) Laplace-Entwicklung von B_j nach der j-ten Spalte gibt

$$\det B_j = \sum_{k=1}^{n} b_k \cdot (-1)^{k+j} \cdot \det A[k,j] = \sum_{k=1}^{n} b_k \cdot (A^\sharp)_{jk},$$

also

$$(\det B_1, ..., \det B_n)^{tr} = A^\sharp \cdot b = y.$$

Wegen $A \cdot A^\sharp = \det A \cdot E_n$ ist das eine Lösung des Gleichungssystems $A \cdot x = \det A \cdot b$.
□

Bemerkungen 7.18 (i) $A^\sharp$ ist mehr aus theoretischen als aus praktischen Gründen wichtig. Ein Pluspunkt ist, daß $A^\sharp$ für beliebige kommutative Ringe R mit Eins definiert ist, ohne Annahmen über $\det A$ und ohne daß man dividieren muß.
Aber meistens ist $R = K$ ein Körper, und dann berechnet man A^{-1} am besten wie in Bemerkung 4.14.

(ii) Das gleiche gilt für die Cramersche Regel, beide Versionen. *Praktisch* löst man lineare Gleichungssysteme besser mit dem Gauß-Algorithmus als mit der Cramerschen Regel.
Aber die geschlossene Formel in der Cramerschen Regel zeigt zum Beispiel, daß die Lösung y stetig von den Koeffizienten in A und b abhängt.

Satz/Definition 7.19 *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums.*
(a) (Satz) Sind $\mathcal{A}$ und $\mathcal{B}$ zwei Basen von V, so ist

$$\begin{aligned} \det M(\mathcal{A}, f, \mathcal{A}) &= \det\left(M(\mathcal{A}, \mathcal{B}) \cdot M(\mathcal{B}, f, \mathcal{B}) \cdot M(\mathcal{B}, \mathcal{A})\right) \\ &= \det M(\mathcal{B}, \mathcal{A})^{-1} \cdot \det M(\mathcal{B}, f, \mathcal{B}) \cdot \det M(\mathcal{B}, \mathcal{A}) \\ &= \det M(\mathcal{B}, f, \mathcal{B}). \end{aligned}$$

Daher ist $\det M(\mathcal{B}, f, \mathcal{B})$ *unabhängig von der Wahl der Basis $\mathcal{B}$.*
(b) (Definition) Daher ist die Determinante von f wohldefiniert durch

$$\det f := \det M(\mathcal{B}, f, \mathcal{B}).$$

Beweis: Klar. □

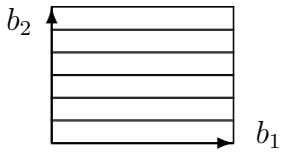
Bemerkungen 7.20 (Volumen und Orientierung im $\mathbb{R}^n$)
(i) Es seien $b_1, ..., b_n$ (Zeilen-)Vektoren im $\mathbb{R}^n$. Das von ihnen erzeugte *Parallelotop* ist

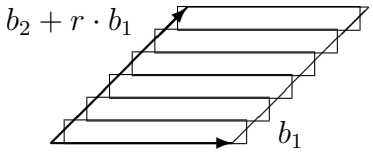
$$P(b_1, ..., b_n) = \{\sum_{j=1}^{n} x_j b_j \mid 0 \le x_j \le 1 \text{ für alle } i = 1, ..., n\}.$$

Sein *Volumen* ist

$$\text{Volumen } (P(b_1, ..., b_n)) = |\det(\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix})|.$$

Das wird hier nicht bewiesen, denn zuerst bräuchte man eine Definition von "Volumen", die hier auch nicht gegeben wird.
Immerhin ist klar, daß das Volumen unter "Scherungen" invariant sein soll, und das paßt zur Multilinearität von det.
Ein Parallelotop heißt im Fall $n = 2$ *Parallelogramm*, im Fall $n = 3$ *Spat*.





(ii) Beim Volumen oben vergißt man das Vorzeichen von $\det(b_1, ..., b_n)^{tr}$. Was ist seine Rolle?
Jeder Basis $(b_1, ..., b_n)$ des $\mathbb{R}^n$ ist eine *Orientierung* zugeordnet, die Zahl

$$\frac{\det(b_1, ..., b_n)^{tr}}{|\det(b_1, ..., b_n)^{tr}|}.$$

Es gibt also nur zwei Orientierungen, $+1$ oder -1.
Bei einer Permutation $(e_{\sigma(1)}, ..., e_{\sigma(n)})$ $(\sigma \in S_n)$ der Standardbasis ist die Orientierung gerade $\text{sign}(\sigma)$.
Man kann zeigen, daß die Gruppe $GL(n, \mathbb{R})$ zwei "Zusammenhangskomponenten hat", die Teilmengen

$$\{A \in GL(n, \mathbb{R}) \mid \det A > 0\} \text{ und } \{A \in GL(n, \mathbb{R}) \mid \det A < 0\}.$$

Daher lassen sich je zwei Basen mit gleicher Orientierung durch eine "stetige Familie" von Basen verbinden.
Beim $\mathbb{C}^n$ hat man das Phänomen der Orientierung nicht: im Gegensatz zu $\mathbb{R} - \{0\}$ ist $\mathbb{C} - \{0\}$ "zusammenhängend".

8 Eigenvektoren und Eigenwerte

In diesem Kapitel bezeichnet K irgendeinen Körper. Hier geht es um Normalformen für Endomorphismen.

Definition/Lemma 8.1 *(a) (Definition) Sei $A \in M(n \times n, K)$. Das* charakteristische Polynom *von A ist das Polynom*

$$\begin{aligned} P_A(t) := \det(t \cdot E_n - A) &= (-1)^n \det(A - t \cdot E_n) \\ &= (-1)^n \cdot \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} - t \end{vmatrix}. \end{aligned}$$

(Lemma) Es ist

$$\begin{aligned} P_A(t) &= t^n - (a_{11} + a_{22} + ... + a_{nn})t^{n-1} \\ &\quad +(...)t^{n-2} + ... + (...)t + (-1)^n \det A \in K[t]. \end{aligned}$$

Es ist also ein unitäres (d.h. Leitkoeffizient 1) Polynom vom Grad n. Der Koeffizient

$$\mathrm{Spur}(A) := a_{11} + a_{22} + ... + a_{nn}$$

von $-t^{n-1}$ heißt Spur *von A.*

(b) (Lemma) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Sei $\mathcal{B}$ eine Basis von V. Das Polynom

$$P_f(t) := (-1)^n \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) = P_{M(\mathcal{B},f,\mathcal{B})}(t)$$

ist unabhängig von der Wahl der Basis $\mathcal{B}$.
(Definition) Es heißt charakteristisches Polynom *von f.*
Auch die Zahl

$$\mathrm{Spur}(f) := \mathrm{Spur}(M(\mathcal{B}, f, \mathcal{B}))$$

ist unabhängig von der Wahl von $\mathcal{B}$. Sie heißt Spur *von f.*

Beweis: (a) Rechnet man $P_A(t)$ mit der Leibniz-Formel aus, so sieht man, daß nur der Summand $(-1)^n(a_{11} - t)(a_{22} - t)...(a_{nn} - t)$ zu $\mathrm{id} \in S_n$ Beiträge zu t^n und t^{n-1} liefert. Daher sind deren Koeffizienten 1 und $-(a_{11} + ... + a_{nn})$. Der konstante Koeffizient ist natürlich $(-1)^n \det A$.
(b) Ist $\widetilde{\mathcal{B}}$ eine zweite Basis, so sei $B := M(\mathcal{B}, \widetilde{\mathcal{B}})$ die Basiswechselmatrix. Es ist

$$\begin{aligned} \det(M(\widetilde{\mathcal{B}}, f, \widetilde{\mathcal{B}}) - t \cdot E_n) &= \det\left[B^{-1} \cdot M(\mathcal{B}, f, \mathcal{B}) \cdot B - t \cdot E_n\right] \\ &= \det\left[B^{-1} \cdot (M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) \cdot B\right] \\ &= \det(B^{-1}) \cdot \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) \cdot \det B \\ &= \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n). \end{aligned}$$

Daher ist $P_f(t)$ unabhängig von der Wahl einer Basis $\mathcal{B}$. Das gleiche gilt für Spur(f), da es der Koeffizient von $-t^{n-1}$ in $P_f(t)$ ist. □

Beispiele 8.2 (i) (Obere Dreiecksmatrix) Ist $A = (a_{ij}) \in M(n \times n, K)$ eine obere Dreiecksmatrix (also $a_{ij} = 0$ für $i > j$), so ist auch $t \cdot E_n - A$ eine obere Dreiecksmatrix, mit Diagonaleinträgen $t - a_{11}, ..., t - a_{nn}$. Daher ist

$$P_A(t) = \det(t \cdot E_n - A) = \prod_{i=1}^{n} (t - a_{ii}).$$

Ein wichtiger Spezialfall: A eine *Diagonalmatrix*, d.h. $a_{ij} = 0$ für $i \neq j$,

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

(ii) (Jordanblock) Eine Matrix $A \in M(n \times n, K)$ ist ein *Jordanblock*, falls sie die Gestalt hat

$$A = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Das ist auch ein Spezialfall einer oberen Dreiecksmatrix, mit

$$P_A(t) = (t - \lambda)^n.$$

(iii) (Obere Blockdreiecksmatrix) Ist $A \in M(n \times n, K)$ eine obere Blockdreiecksmatrix mit Blöcken $A_i \in M(r_i \times r_i, K)$ in der Diagonalen, so ist auch $t \cdot E_n - A$ eine obere Blockdreiecksmatrix:

$$A = \begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix}, \; t \cdot E_n - A = \begin{pmatrix} tE_{r_1} - A_1 & & & * \\ & tE_{r_2} - A_2 & & \\ & & \ddots & \\ 0 & & & tE_{r_k} - A_k \end{pmatrix}$$

Daher ist dann

$$P_A(t) = \det(t \cdot E_n - A) = \prod_{i=1}^{k} \det(t \cdot E_{r_i} - A_i) = \prod_{i=1}^{k} P_{A_i}(t).$$

(iv)(Drehmatrix) Sei $\alpha \in \mathbb{R}$. Die Matrix

$$A := \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

hat das charakteristische Polynom

$$P_A(t) = (\cos\alpha - t)(\cos\alpha - t) + \sin^2\alpha = t^2 - 2\cos\alpha \cdot t + 1 = (t - e^{i\alpha})(t - e^{-i\alpha}).$$

(v) (Typische Klausurmatrix) $A := \begin{pmatrix} 1 & 3 & 1 \\ -1 & 5 & 2 \\ 0 & 0 & 3 \end{pmatrix} \in M(3 \times 3, \mathbb{Q})$.

$$\begin{aligned} P_A(t) &= (-1)^3 \cdot \det \begin{pmatrix} 1-t & 3 & 1 \\ -1 & 5-t & 2 \\ 0 & 0 & 3-t \end{pmatrix} = (t-3) \cdot \det \begin{pmatrix} 1-t & 3 \\ -1 & 5-t \end{pmatrix} \\ &= (t-3) \cdot ((1-t)(5-t)+3) = (t-3)(t^2 - 6t + 8) \\ &= (t-3)(t-2)(t-4). \end{aligned}$$

(vi) (Telefonmatrix) Die Matrix

$$\text{Tel} := \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}$$

hat das charakteristische Polynom

$$P_{\text{Tel}}(t) = ... = t^3 - 12t^2 - 18t = t(t - 6 + 3\sqrt{6})(t - 6 - 3\sqrt{6}).$$

Definition 8.3 (a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Element $v \in V - \{0\}$ heißt *Eigenvektor* von f, falls es ein $\lambda \in K$ gibt mit

$$f(v) = \lambda \cdot v.$$

Ein solches λ heißt *Eigenwert* von f.

(b) (Lemma) Für jedes $\lambda \in K$ ist die Menge

$$\text{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda \cdot v\} = \ker(f - \lambda \cdot \text{id})$$

offenbar ein Untervektorraum von V. Und offenbar ist λ genau dann ein Eigenwert von f, wenn $\text{Eig}(f, \lambda) \neq \{0\}$ ist.
(Definition) Der Vektorraum $\text{Eig}(f, \lambda)$ heißt *Eigenraum* von f bezüglich λ.

(c) Eigenwerte, Eigenvektoren und Eigenräume einer Matrix $A \in M(n \times n, K)$ sind die Eigenwerte, Eigenvektoren und Eigenräume des Endomorphismus

$$l_A : M(n \times 1, K) \to M(n \times 1, K), \quad b \mapsto A \cdot b.$$
($l_A =$ Linksmultiplikation mit A)

Äquivalent und konkreter: $v \in M(n \times 1, K) - \{0\}$ heißt *Eigenvektor* von A, falls es ein $\lambda \in K$ gibt mit

$$A \cdot v = \lambda \cdot v.$$

Ein solches λ heißt *Eigenwert* von A.
Für *jedes* $\lambda \in K$ ist

$$\begin{aligned} \mathrm{Eig}(A, \lambda) &:= \{v \in M(n \times 1, K) \mid A \cdot v = \lambda \cdot v\} \\ &= \ker(l_A - \lambda \cdot \mathrm{id}) = \mathrm{Lös}(A - \lambda \cdot E_n, 0) \end{aligned}$$

der Eigenraum von A zum Wert λ. Es ist offenbar ein Untervektorraum von $M(n \times 1, K)$. Und offenbar ist $\lambda \in K$ genau dann ein Eigenwert von A, wenn $\mathrm{Eig}(A, \lambda) \neq \{0\}$ ist.

Satz 8.4 (Bestimmung von Eigenwerten und Eigenvektoren)
(a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Dann gilt folgende Äquivalenz:

$$\lambda \in K \text{ ist ein Eigenwert von } f \iff P_f(\lambda) = 0.$$

(b) Sei $A \in M(n \times n, K)$. Dann gilt folgende Äquivalenz:

$$\lambda \in K \text{ ist ein Eigenwert von } A \iff P_A(\lambda) = 0.$$

Beweis: Zuerst b):

$$\begin{aligned} & \lambda \text{ ist ein Eigenwert von } A \\ \iff\ & \mathrm{Eig}(A, \lambda) \neq \{0\} \\ \iff\ & \mathrm{Lös}(A - \lambda \cdot E_n, 0) \neq \{0\} \\ \overset{6.3}{\iff}\ & A - \lambda \cdot E_n \text{ ist nicht invertierbar} \\ \overset{7.7}{\iff}\ & 0 = \det(\lambda \cdot E_n - A) = P_A(\lambda). \end{aligned}$$

Nun a):

$$\begin{aligned} & \lambda \text{ ist ein Eigenwert von } f \\ \iff\ & \mathrm{Eig}(f, \lambda) \neq \{0\} \\ \iff\ & \ker(f - \lambda \cdot \mathrm{id}) \neq \{0\} \\ \iff\ & f - \lambda \cdot \mathrm{id} \text{ ist nicht invertierbar} \\ \overset{*}{\iff}\ & 0 = \det(\lambda \cdot \mathrm{id} - f) \overset{!}{=} P_f(\lambda). \end{aligned}$$

$\overset{*}{\iff}$ benutzt: ein Endomorphismus (hier $\lambda \cdot \mathrm{id} - f$) ist invertierbar $\iff$ seine Determinante ist Null. Das folgt mit 7.19 und 5.11 (d) aus der analagen Aussage 7.7 für Matrizen. □

Beispiele 8.5 Es werden dieselben Beispiele wie in 8.2 betrachtet.
(i) (Obere Dreiecksmatrix) Wegen $P_A(t) = \prod_{i=1}^{n}(t - a_{ii})$ sind die Eigenwerte $a_{11}, ..., a_{nn}$.
Aber die Eigenvektoren sind schwerer zu bestimmen. Sofort sieht man nur den Eigenvektor $e_1 = (1, 0..., 0)^{tr}$ zum Eigenwert a_{11}. Wenn mehrere a_{ii} übereinstimmen, gibt es zu ihnen eventuell nur einen Eigenvektor, siehe ii).
Nur im Fall einer Diagonalmatrix sieht man sofort viele Eigenvektoren: e_i ist Eigenvektor zum Eigenwert a_{ii}. Mehr dazu kommt in 8.6.

(ii) (Jordanblock) Hier ist $P_A(t) = (t - \lambda)^n$, also hat man nur einen Eigenwert. Tatsächlich ist hier

$$A - \lambda \cdot E_n = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix},$$

also $\operatorname{rang}(A - \lambda \cdot E_n) = n - 1$, also $\dim \operatorname{Eig}(A, \lambda) = 1$.
Genauer:

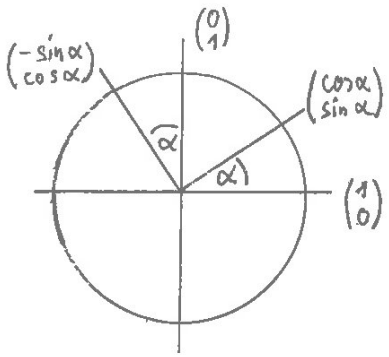
$$\operatorname{Eig}(A, \lambda) = \operatorname{Lös}(A - \lambda \cdot E_n, 0) = K \cdot e_1.$$

(iii) (Blockdiagonalmatrix) Wenn A_i einen Eigenwert λ und Eigenvektor $v \in M(r_i \times 1, K)$ hat, so ist $(0, ..., 0, v^{tr}, 0, ...0)^{tr} \in M(n \times 1, K)$ (mit den richtigen Anzahlen von Nullen) ein Eigenvektor von A mit Eigenwert λ.

(iv) (Vgl. Beispiel 5.9 (ii)) Sei $\alpha \in \mathbb{R}$. Die Linksmultiplikation l_A mit der Matrix

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

beschreibt eine Drehung des $\mathbb{R}$-Vektorraums $M(2 \times 1, \mathbb{R}) \cong \mathbb{R}^2$ um 0 um den Winkel α.



Für $\alpha \notin \pi \cdot \mathbb{Z}$ hat sie offenbar keinen Eigenvektor und keinen Eigenwert. Das paßt dazu, daß

$$P_A(t) = t^2 - 2\cos\alpha \cdot t + 1 = (t - e^{i\alpha})(t - e^{-i\alpha})$$

keine reelle Nullstelle hat.
Aber die Linksmultiplikation mit derselben Matrix auf dem $\mathbb{C}$-Vektorraum $M(2 \times 1, \mathbb{C}) \cong \mathbb{C}^2$ hat die Eigenwerte $e^{i\alpha}$ und $e^{-i\alpha}$ und die Eigenvektoren $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ und $\begin{pmatrix} 1 \\ i \end{pmatrix}$:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \mp i \end{pmatrix} = \begin{pmatrix} \cos\alpha \pm i\sin\alpha \\ \sin\alpha \mp i\cos\alpha \end{pmatrix} = e^{\pm i\alpha} \cdot \begin{pmatrix} 1 \\ \mp i \end{pmatrix}.$$

(v) (Typische Klausurmatrix) Wegen $P_A(t) = (t-3)(t-2)(t-4)$ sind die Eigenwerte $2, 3$ und 4. Die Eigenräume sind:

$$Eig(A,2) = \text{Lös}(A - 2 \cdot E_3, 0) = \text{Lös}(\begin{pmatrix} -1 & 3 & 1 \\ -1 & 3 & 2 \\ 0 & 0 & 1 \end{pmatrix}, 0) = \text{span}_{\mathbb{Q}}(\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}),$$

$$Eig(A,3) = \text{Lös}(A - 3 \cdot E_3, 0) = \text{Lös}(\begin{pmatrix} -2 & 3 & 1 \\ -1 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}, 0) = \text{span}_{\mathbb{Q}}(\begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix}),$$

$$Eig(A,4) = \text{Lös}(A - 4 \cdot E_3, 0) = \text{Lös}(\begin{pmatrix} -3 & 3 & 1 \\ -1 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}, 0) = \text{span}_{\mathbb{Q}}(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}).$$

(vi) (Telefonmatrix) Wegen

$$P_{\text{Tel}}(t) = ... = t^3 - 12t^2 - 18t = t(t - 6 + 3\sqrt{6})(t - 6 - 3\sqrt{6})$$

hat Tel als Matrix in $M(3 \times 3, \mathbb{R})$ die drei Eigenwerte 0, $6 - 3\sqrt{6}$ und $6 + 3\sqrt{6}$.
Mit einiger Mühe rechnet man aus: Die Eigenräume sind hier alle eindimensional und werden erzeugt durch die drei Eigenvektoren

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 5 \\ 8 - 3\sqrt{6} \\ 11 - 3\sqrt{6} \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 5 \\ 8 + 3\sqrt{6} \\ 11 + 3\sqrt{6} \end{pmatrix}.$$

Als Matrix in $M(3 \times 1, \mathbb{Q})$ hat Tel dagegen nur den Eigenwert 0 und den zugehörigen eindimensionalen Eigenraum $\text{Eig}(Tel, 0)$, der vom Eigenvektor $(1, -2, 1)^{tr}$ erzeugt wird.

Definition/Lemma 8.6 *(a) (Definition) Ein Endomorphismus $f : V \to V$ eines K-Vektorraums V mit $\dim V = n \in \mathbb{N}$ heißt* diagonalisierbar*, falls es eine Basis von V aus Eigenvektoren von f gibt.*
(b) (Lemma) Ist $\mathcal{B} = (b_1, ..., b_n)$ eine solche Basis mit $f(b_i) = \lambda_i$, so ist

$$M(\mathcal{B}, f, \mathcal{B}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n, \end{pmatrix}.$$

(c) (Definition) Eine Matrix $A \in M(n \times n, K)$ heißt diagonalisierbar*, falls es es eine Basis von $M(n \times 1, K)$ aus Eigenvektoren von A gibt.*
(d) (Lemma) Eine Matrix A ist diagonalierbar genau dann, wenn es eine Matrix $B \in GL(n, K)$ gibt mit

$$B^{-1} \cdot A \cdot B = \text{ Diagonalmatrix.}$$

Beweis: (a)+(c): Definitionen.
(b): Klar.
(d) Die Spalten von B bilden eine Basis von Eigenvektoren. □

Bemerkungen 8.7 (i) Wenn ein Endomorphismus diagonalisierbar ist, ist das etwas ganz besonderes. Viele Endomorphismen lassen sich nicht diagonalisieren.
Aber auch dann gibt es *Normalformen*. Das sind Matrizen von besonders guter Gestalt, die nach Wahl von besonders guten Basen die Endomorphismen repräsentieren. Wenn das charakteristische Polynom in Linearfaktoren zerfällt, hat man die *Jordannormalform*. Aber auch, wenn es nicht zerfällt, hat man interessante und nützliche Normalformen. Das wird ein Thema der Linearen Algebra IIa sein.

(ii) Der letzte Satz des Kapitels, Satz 8.11, soll einen Eindruck davon geben, was jenseits der diagonalisierbaren Endomorphismen los ist. Er ist ein wichtiger Schritt zur Jordannormalform.

(iii) Er erfordert den Begriff der *direkten Summe* von Untervektorräumen eines Vektorraums, der in der folgenden Kombination 8.8 aus Satz und Definition entwickelt wird.
Dieser Begriff ist aber auch für sich allein wichtig. Er wird zum Beispiel auch im Satz 9.22 (a) benutzt.

Satz/Definition 8.8 *(a) (Notation) Seien $V_i \subset V$ $(i = 1, ..., k)$ Untervektorräume eines Vektorraums V. Der von ihnen erzeugte Untervektorraum von V ist*

$$\sum_{i=1}^{k} V_i := \{v_1 + ... + v_k \mid v_i \in V_i\}.$$

(b) (Satz) Folgende drei Bedingungen sind äquivalent:

- *(α) Jedes Element von $\sum_{i=1}^{k} V_i$ lässt sich auf eindeutige Weise als Linearkombination von Elementen der V_i schreiben, d.h.*

$$v_1 + ... + v_k = \widetilde{v}_1 + ... + \widetilde{v}_k \text{ mit } v_i, \widetilde{v}_i \in V_i \Rightarrow \text{ für alle } i \quad v_i = \widetilde{v}_i.$$

- *(β) Beliebige Basen der V_i bilden zusammen eine Basis von $\sum_{i=1}^{k} V_i$.*
- *(γ) Für alle $i = 1, ..., k$ ist $V_i \cap (\sum_{j \neq i} V_j) = \{0\}$.*

(c) (Definition) Wenn die drei äquivalenten Bedingungen in (b) erfüllt sind, ist der von den V_i erzeugte Untervektorraum die direkte Summe *der V_i; Notation: $\bigoplus_{i=1}^{k} V_i$ oder $V_1 \oplus ... \oplus V_k$.*

Beweis von (b): $(\alpha) \iff (\beta)$: Leicht, Übung.
$(\alpha) \Rightarrow (\gamma)$: Wäre $v \in V_i \cap (\sum_{j \neq i} V_j)$ und $v \neq 0$, so ließe sich v auf zwei Weisen als Linearkombination der Elemente der V_k schreiben, einmal als $v \in V_i$, einmal als $v \in \sum_{j \neq i} V_j$.
$(\gamma) \Rightarrow (\alpha)$: Indirekt. Sei $\alpha)$ nicht erfüllt; sei $\sum_j v_j = \sum_j \widetilde{v}_j$ mit $v_j, \widetilde{v}_j \in V_j$ und $v_i \neq \widetilde{v}_i$ für irgendein (mindestens ein) i.
Dann ist

$$0 \neq v_i - \widetilde{v}_i = \sum_{j \neq i} (\widetilde{v}_j - v_j) \in V_i \cap (\sum_{j \neq i} V_j).$$

Also ist auch $\gamma)$ nicht erfüllt. □

Beispiel 8.9 Sei $V = K^4$ mit Standardbasis (e_1, e_2, e_3, e_4). Sei

$$V_1 := Ke_1 + Ke_2,\ V_2 := Ke_3,\ V_3 := Ke_1 + Ke_4,\ V_4 := Ke_4.$$

Dann hat man direkte Summen $V_1 \oplus V_2,\ V_1 \oplus V_4,\ V_2 \oplus V_3,\ V_2 \oplus V_4, V_1 \oplus V_2 \oplus V_4$.
Aber die Summen $V_1 + V_3,\ V_3 + V_4$ und alle Summen, die diese enthalten $(V_1 + V_2 + V_3,\ V_1 + V_3 + V_4,\ V_2 + V_3 + V_4, V_1 + V_2 + V_3 + V_4)$, sind nicht direkt.

Definition 8.10 (a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Element $v \in V - \{0\}$ heißt *verallgemeinerter Eigenvektor* oder *Hauptvektor*, falls es ein $\lambda \in K$ und ein $m \in \mathbb{N}$ gibt mit

$$(f - \lambda \cdot \mathrm{id})^m(v) = 0.$$

(b) Für jedes $\lambda \in K$ ist die Menge

$$\mathrm{Hau}(f, \lambda) := \{v \in V \mid \text{ es gibt ein } m \in \mathbb{N} \text{ mit } (f - \lambda \cdot \mathrm{id})^m(v) = 0\}$$

offenbar ein Untervektorraum von V. Offenbar ist $\mathrm{Hau}(f, \lambda) \supset \mathrm{Eig}(f, \lambda)$. Also gilt

$$\lambda \text{ Eigenwert} \iff \mathrm{Eig}(f, \lambda) \neq \{0\} \iff \mathrm{Hau}(f, \lambda) \neq \{0\}.$$

$\mathrm{Hau}(f, \lambda)$ heißt *Hauptraum* von f bezüglich λ.

Satz 8.11 *(In dieser Vorlesung ohne Beweis) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums mit $\dim_K V = n \in \mathbb{N}$. Sein charakteristisches Polynom zerfalle in Linearfaktoren,*

$$P_f(t) = \prod_{\lambda \text{ Eigenwert}} (t - \lambda)^{d(\lambda)} \qquad (\text{mit } d(\lambda) \in \mathbb{N}).$$

Dann gilt:

$$V = \bigoplus_{\lambda \text{ Eigenwert}} \mathrm{Hau}(f, \lambda) \quad \text{und} \quad \dim \mathrm{Hau}(f, \lambda) = d(\lambda).$$

9 Euklidische Vektorräume

Am Anfang dieses Kapitels wird mit einem beliebigen Körper K gearbeitet, später wird nur noch der Körper $\mathbb{R}$ betrachtet.

Definition 9.1 (a) Seien V und W K-Vektorräume. Eine Abbildung $\phi : V \times W \to K$ ist eine *Bilinearform*, falls gilt:

(i) für jedes $y \in W$ ist folgende Abbildung linear:

$$V \to K, \quad x \mapsto \phi(x, y).$$

(ii) für jedes $x \in V$ ist folgende Abbildung linear:

$$W \to K, \quad y \mapsto \phi(x, y).$$

(b) Eine *Bilinearform auf einem Vektorraum* V ist eine Bilinearform

$$\phi : V \times V \to K.$$

[Die meisten interessanten Bilinearformen sind von diesem Typ, d.h. $V = W$.]
(c) Eine *symmetrische Bilinearform* ist eine Bilinearform $\phi : V \times V \to K$ mit

$$\phi(x, y) = \phi(y, x).$$

(d) Eine (*alternierende* oder) *schiefsymmetrische Bilinearform* ist eine Bilinearform $\phi : V \times V \to K$ mit

$$\phi(x, y) = -\phi(y, x).$$

Beispiele 9.2 (i) Das *Standard-Skalarprodukt* auf $M(n \times 1, K)$:

$$\begin{aligned} \langle\ ,\rangle : M(n \times 1, K) \times M(n \times 1, K) &\to K, \\ (x, y) &\mapsto \langle x, y\rangle := \sum_{i=1}^{n} x_i \cdot y_i = x^{tr} \cdot y. \end{aligned}$$

$\langle\ ,\rangle$ ist eine symmetrische Bilinearform. Sie wird in Kapitel 10 studiert.

(ii) Auf dem $\mathbb{R}$-Vektorraum $\mathcal{C}^0([0, 1], \mathbb{R})$ der stetigen Funktionen auf dem Intervall $[0, 1]$ (vgl. Beispiel 3.3 (d)) hat man die symmetrische Bilinearform

$$\begin{aligned} \phi_{int} : \mathcal{C}^0([0, 1], \mathbb{R}) \times \mathcal{C}^0([0, 1], \mathbb{R}) &\to \mathbb{R}, \\ (f, g) &\mapsto \int_0^1 f(x)g(x)\mathrm{d}x. \end{aligned}$$

(iii) Eine *Linearform* auf einem K-Vektorraum V ist ein Vektorraumhomomorphismus $f : V \to K$.

[Das paßt zum Begriff *Bilinearform* in Definition 9.1.]
Der *Dualraum* zu V ist der K-Vektorraum

$$V^* := \mathrm{Hom}(V, K).$$

Die folgende Bilinearform auf $V^* \times V$ ist zugleich natürlich, etwas abstrakt und ziemlich trivial:

$$\phi : V^* \times V \to K, \quad (f, x) \mapsto \phi(f, x) := f(x).$$

(iv) Die folgende Definition von Bil_A und das Lemma 9.3 geben eine erste Verbindung zwischen Bilinearformen und Matrizen,für den Fall von Spaltenvektorräumen.
Sie ist analog zum 1. Teil der Verbindung zwischen linearen Abbildungen und Matrizen im Kapitel 5. Die Analoga des 2. und 3. Teils werden aber erst in der Linearen Algebra IIa behandelt.
Eine $(m \times n)$-Matrix $A = (a_{ij}) \in M(m \times n, K)$ definiert (offenbar) eine Bilinearform auf $M(m \times 1, K) \times M(n \times 1, K)$ durch

$$\begin{aligned} \mathrm{Bil}_A : M(m \times 1, K) \times M(n \times 1, K) &\to K \\ (x, y) &\mapsto x^{tr} \cdot A \cdot y = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \cdot a_{ij} \cdot y_j. \end{aligned}$$

Lemma 9.3 *(a) Zu jeder Bilinearform ϕ auf $M(m \times 1, K) \times M(n \times 1, K)$ gibt es eine eindeutige Matrix $A \in M(m \times n, K)$ mit $\phi = \mathrm{Bil}_A$.*
(b) Die Menge der Bilinearformen auf $M(m \times 1, K) \times M(n \times 1, K)$ ist ein K-Vektorraum, und die Abbildung $A \mapsto \mathrm{Bil}_A$ ist ein Vektorraumisomorphismus von $M(m \times n, K)$ in diese Menge.
(c) Im Fall von $m = n$ ist Bil_A genau dann symmetrisch [bzw. schiefsymmetrisch], wenn A symmetrisch [bzw. schiefsymmetrisch] ist.

Beweis: Übung. □

Ab hier wird in diesem Kapitel nur noch der Körper $\mathbb{R}$ betrachtet.

Definition 9.4 (a) Eine symmetrische Bilinearform $\phi : V \times V \to \mathbb{R}$ auf einem $\mathbb{R}$-Vektorraum V heißt *positiv definit*, falls gilt:

$$\phi(v, v) > 0 \text{ für alle } v \in V - \{0\}.$$

Sie heißt *negativ definit*, falls gilt:

$$\phi(v, v) < 0 \text{ für alle } v \in V - \{0\}.$$

Sie heißt *positiv semidefinit* [bzw. *negativ semidefinit*], falls nur gilt:

$$\phi(v, v) \geq 0 \text{ [bzw } \phi(v, v) \leq 0] \text{ für alle } v \in V.$$

(b) Eine symmetrische Matrix A heißt positiv oder negativ (semi)definit, wenn die Bilinearform Bil_A auf $M(n \times 1, \mathbb{R})$ es ist.

Definition 9.5 (a) Ein $\mathbb{R}$-Vektorraum V zusammen mit einer positiv definiten symmetrischen Bilinearform $\phi : V \times V \to \mathbb{R}$ heißt *Euklidischer Vektorraum*. Die Bilinearform ϕ ist sein *Skalarprodukt*.
(b) Dann ist die *Norm* (oder *Länge*) eines Vektors $v \in V$

$$\|v\| := \sqrt{\phi(v,v)} \geq 0.$$

Bemerkungen 9.6 (i) Die Formulierung "zusammen mit" ist nicht präzise. Genaugenommen ist ein Euklidischer Vektorraum das Paar (V, ϕ). Aber der Vektorraum V wird als das primäre Objekt angesehen.
(ii) Bei "natürlich" gegebenen Skalarprodukten gibt es in der Literatur neben $\phi(x,y)$ eine ganze Reihe von gebräuchlichen Notationen, zum Beispiel $x \cdot y$, $\langle x, y\rangle$, (x,y).
(iii) Eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ ist nach Definition positiv definit, falls

$$(x_1 \cdots x_n) \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} > 0 \quad \text{für alle } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M(n \times 1, \mathbb{R}) - \{0\} \text{ ist.}$$

Am Ende des Kapitels werden zwei andere Charakterisierungen positiv definiter Matrizen gegeben.

Beispiele 9.7 (i) Der wichtigste Euklidische Vektorraum ist der $\mathbb{R}^n$ mit dem Standard-Skalarprodukt ϕ_{st},

$$\phi_{st} : ((x_1, ..., x_n), (y_1, ..., y_n)) \overset{\phi_{st}}{\mapsto} (x_1 \cdots x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i.$$

(ii) Die Vektorräume $\mathbb{R}^n$ und $M(1 \times n, \mathbb{R})$ waren in Definition 4.1 (c) identifiziert worden, aber nicht $\mathbb{R}^n$ und $M(n \times 1, \mathbb{R})$. Das Standard-Skalarprodukt auf $M(n \times 1, \mathbb{R})$ ist Bil_{E_n},

$$(x,y) \mapsto \mathrm{Bil}_{E_n}(x,y) = x^{tr} \cdot E_n \cdot y = x^{tr} \cdot y = \sum_{i=1}^{n} x_i y_i.$$

(iii) Der $\mathbb{R}$-Vektorraum $\mathcal{C}^0([0,1], \mathbb{R})$ mit der Bilinearform ϕ_{int},

$$\phi_{int} : (f,g) \mapsto \int_0^1 f(x)g(x)\mathrm{d}x,$$

ist ein Euklidischer Vektorraum wegen (Beweis: Analysis)

$$\int_0^1 f^2(x)\mathrm{d}x > 0, \quad \text{falls } f \neq 0.$$

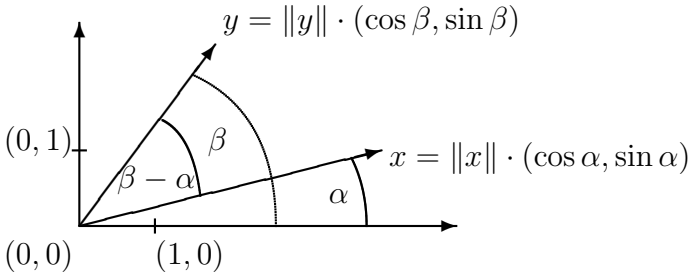
Bemerkungen 9.8 (i) Längen und Winkel im $\mathbb{R}^2$:
Nach dem Satz von Pythagoras ist die Länge des Vektors $(x_1, x_2) \in \mathbb{R}^2$ tatsächlich

$$\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}.$$

Sind x und $y \in \mathbb{R}^2 - \{0\}$, so haben $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ Länge 1, und es gibt eindeutige α und $\beta \in [0, 2\pi)$ mit

$$x = \|x\| \cdot (\cos\alpha, \sin\alpha) \text{ und } y = \|y\| \cdot (\cos\beta, \sin\beta).$$

Skizze:



Es ist

$$\begin{aligned} \phi_{st}(x,y) &= \|x\| \cdot \|y\| \cdot (\cos\alpha\cos\beta + \sin\alpha\sin\beta) \\ &= \|x\| \cdot \|y\| \cdot \cos(\beta - \alpha) \text{ (nach einem Additionstheorem).} \end{aligned}$$

Es gibt ein eindeutiges $\gamma \in [0, \pi]$ mit $\cos(\beta - \alpha) = \cos\gamma$. Es ist $\gamma \equiv \pm(\beta - \alpha)$ mod 2π. Man kann γ als *den Winkel* zwischen x und y auffassen. Er wird eindeutig bestimmt durch $\gamma \in [0, \pi]$ und

$$\phi_{st}(x,y) = \|x\| \cdot \|y\| \cdot \cos(\gamma).$$

(ii) Aufgrund der Cauchy-Schwarzschen Ungleichung (Satz 9.9) wird es möglich sein, diesen Begriff eines Winkels zwischen zwei Vektoren auf beliebige Euklidische Vektorräume zu verallgemeinern (Definition 9.10). Danach wird auch der Begriff der Länge im allgemeinen diskutiert werden (Definition 9.11 und Korollar 9.12).

Satz 9.9 *Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Dann gilt die Cauchy-Schwarzsche Ungleichung:*

$$|\phi(x,y)| \leq \|x\| \cdot \|y\| \text{ für } x, y \in V,$$

und Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Beweis: Falls $x = 0$ oder $y = 0$ ist, sind beide Seiten gleich Null, und x und y sind linear abhängig.
Seien $x \neq 0$ und $y \neq 0$. Dann haben die Vektoren $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ Länge 1. Es ist (mit $\pm$ = plus oder minus)

$$\begin{aligned} 0 &\leq \phi(\frac{x}{\|x\|} \pm \frac{y}{\|y\|}, \frac{x}{\|x\|} \pm \frac{y}{\|y\|}) \\ &= \phi(\frac{x}{\|x\|}, \frac{x}{\|x\|}) \pm 2\phi(\frac{x}{\|x\|}, \frac{y}{\|y\|}) + \phi(\frac{y}{\|y\|}, \frac{y}{\|y\|}) \\ &= 1 \pm 2\frac{\phi(x,y)}{\|x\| \cdot \|y\|} + 1, \end{aligned}$$

also

$$-(\pm 1)\phi(x,y) \leq \|x\| \cdot \|y\|,$$

also

$$|\phi(x,y)| \leq \|x\| \cdot \|y\|$$

Bei Gleichheit ist $-(\pm 1)\phi(x,y) = \|x\| \cdot \|y\|$, also

$$0 = \phi(\frac{x}{\|x\|} \pm \frac{y}{\|y\|}, \frac{x}{\|x\|} \pm \frac{y}{\|y\|}),$$

also

$$\frac{x}{\|x\|} \pm \frac{y}{\|y\|} = 0.$$

Bei $x \neq 0$ und $y \neq 0$ ist das äquivalent dazu, daß x und y linear abhängig sind. □

Definition 9.10 Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ.
(a) Wegen der Cauchy-Schwarzschen Ungleichung gibt es zu je zwei Vektoren x und $y \in V - \{0\}$ ein eindeutiges $\gamma \in [0, \pi]$ mit

$$\phi(x,y) = \|x\| \cdot \|y\| \cdot \cos\gamma.$$

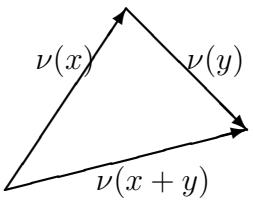
Dieses γ heißt der *Winkel* zwischen x und y.
(b) Zwei Vektoren x und y in V heißen *orthogonal* ($x \perp y$) genau dann, wenn der Winkel zwischen ihnen $\pi/2$ ist, d.h. wenn $\varphi(x,y) = 0$ ist. (Also ist der Vektor 0 zu jedem Vektor orthogonal.)
(c) Bemerkung: Zwei Vektoren x und y in $V - \{0\}$ sind genau dann linear abhängig, wenn der Winkel zwischen ihnen 0 oder π ist.

Definition 9.11 Sei V ein $\mathbb{R}$-Vektorraum. Eine *Norm* auf V ist eine Abbildung $\nu : V \to \mathbb{R}_{\geq 0}$ mit den Eigenschaften:

(N1) $\nu(x) = 0 \iff x = 0$.

(N2) $\nu(\lambda \cdot x) = |\lambda| \cdot \nu(x)$ für $\lambda \in \mathbb{R}$, $x \in V$.

(N3) $\nu(x+y) \leq \nu(x) + \nu(y)$ (Dreiecksungleichung).



Korollar 9.12 *(zur Cauchy-Schwarzschen Ungleichung, Satz 9.9) Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Die Längen-Abbildung*

$$\|.\| : V \to \mathbb{R}_{\geq 0},\ x \mapsto \sqrt{\phi(x,x)}$$

ist eine Norm.

Beweis: (N1) und (N2) sind klar. (N3) folgt aus

$$\begin{aligned} \|x+y\|^2 &= \phi(x+y,x+y) = \phi(x,x) + 2\phi(x,y) + \phi(y,y) \\ &= \|x\|^2 + 2\phi(x,y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Definition/Lemma 9.13 *(a) (Definition) Den Begriff des Abstandes erfaßt man abstrakt mit dem Begriff einer Metrik:*
Sei X irgendeine nichtleere Menge. Eine Abbildung $d : X \times X \to \mathbb{R}_{\geq 0}$ heißt Metrik, falls sie erfüllt:

(M1) $d(x,y) = 0 \iff x = y$.

(M2) $d(y,x) = d(x,y)$.

(M3) $d(x,z) \leq d(x,y) + d(y,z)$ *(Dreiecksungleichung)*

(b) Sei V ein $\mathbb{R}$-Vektorraum mit einer Norm $\|.\| : V \to \mathbb{R}_{\geq 0}$. Dann ist die Abbildung $d : V \times V \to \mathbb{R}_{\geq 0}$ mit $d(x,y) := \|x - y\|$ eine Metrik auf V.

Beweis: (a) Definition. (b) Klar. □

Bemerkung 9.14 Also hat man auf einem Euklidischen Vektorraum den Begriff eines Winkels nach 9.10 und den Begriff des Abstandes nach 9.13 und 9.12.

Definition 9.15 Sei V ein endlich-dimensionaler Euklidischer Vektorraum mit Skalarprodukt $\phi : V \times V \to \mathbb{R}$.
Eine Basis $(b_1, ..., b_n)$ heißt *Orthogonalbasis*, falls $b_i \perp b_j$ für $i \neq j$ ist.
Eine Basis $(b_1, ..., b_n)$ heißt *Orthonormalbasis* oder kürzer *ON-Basis*, falls $b_i \perp b_j$ für $i \neq j$ und $\|b_i\| = 1$ für alle i ist. Die gleiche Bedingung kürzer geschrieben:

$$\phi(b_i, b_j) = \delta_{ij}.$$

Satz 9.16 *Sei V ein endlich-dimensionaler Euklidischer Vektorraum mit Skalarprodukt $\phi : V \times V \to \mathbb{R}$.*
(a) Er besitzt eine Orthonormalbasis.
(b) (Gram-Schmidtsches Orthogonalisierungsverfahren)
Aus einer gegebenen Basis $(a_1, ..., a_n)$ erhält man in folgender Weise eine Orthogonalbasis $(b_1, ..., b_n)$.

$$\begin{aligned} b_1 &:= a_1, \\ b_2 &:= a_2 - \frac{\phi(a_2, b_1)}{\phi(b_1, b_1)} \cdot b_1, \\ b_3 &:= a_3 - \sum_{i=1}^{2} \frac{\phi(a_3, b_i)}{\phi(b_i, b_i)} \cdot b_i, \\ &\vdots \\ b_n &:= a_n - \sum_{i=1}^{n-1} \frac{\phi(a_n, b_i)}{\phi(b_i, b_i)} \cdot b_i. \end{aligned}$$

(c) Aus einer Orthogonalbasis $(b_1, ..., b_n)$ erhält man eine Orthonormalbasis $(c_1, ..., c_n)$ durch Normieren:

$$c_i := \frac{b_i}{\|b_i\|}.$$

(d) Ist $(b_1, ..., b_n)$ eine Orthogonalbasis von V und $x \in V$ beliebig, so ist

$$x = \sum_{i=1}^{n} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot b_i.$$

Im Spezialfall einer Orthonormalbasis $(c_1, ..., c_n)$ ist dann natürlich

$$x = \sum_{i=1}^{n} \phi(x, c_i) \cdot c_i.$$

Beweis: (a) folgt aus (b) und (c).
(b) Induktiv.
Induktionsanfang: $b_1 = a_1$.
Induktionsannahme: für ein $j \in \{2, ..., n\}$ ist $(b_1, ..., b_{j-1})$ eine Orthogonalbasis des von $(a_1, ..., a_{j-1})$ erzeugten Untervektorraums.

Induktionsschritt: Die Definition von b_j für $j = 2, ..., n$ zeigt, daß b_j orthogonal zu allen b_k mit $k < j$ ist:

$$\begin{aligned} b_j &:= a_j - \sum_{i=1}^{j-1} \frac{\phi(a_j, b_i)}{\phi(b_i, b_i)} \cdot b_i, \\ \phi(b_j, b_k) &= \phi(a_j, b_k) - \sum_{i=1}^{j-1} \frac{\phi(a_j, b_i)}{\phi(b_i, b_i)} \cdot \phi(b_i, b_k) \\ &= \phi(a_j, b_k) - \frac{\phi(a_j, b_k)}{\phi(b_k, b_k)} \cdot \phi(b_k, b_k) = 0. \end{aligned}$$

Daher ist $(b_1, ..., b_j)$ eine Orthogonalbasis des von $(a_1, ..., a_j)$ erzeugten Untervektorraums.
(c) Klar.
(d) Weil $(b_1, ..., b_n)$ eine Basis von V ist, gibt es $x_i \in \mathbb{R}$ mit $x = \sum_{i=1}^n x_i \cdot b_i$. Es ist

$$\phi(x, b_k) = \sum_{i=1}^{n} x_i \cdot \phi(b_i, b_k) = x_k \cdot \phi(b_k, b_k).$$

□

Beispiele 9.17 (i) $n = 3$, $a_1 = (1, 1, 1)$, $a_2 = (1, 1, 0)$, $a_3 = (1, 0, 0)$.

$$\begin{aligned} b_1 &= a_1, \quad \|b_1\|^2 = 3. \\ b_2 &= a_2 - \frac{\phi(a_2, b_1)}{\|b_1\|^2} b_1 = (1, 1, 0) - \frac{2}{3}(1, 1, 1) = (\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}), \\ &\quad \|b_2\|^2 = \frac{1}{9} + \frac{1}{9} + \frac{4}{9} = \frac{6}{9} = \frac{2}{3}. \\ b_3 &= a_3 - \frac{\phi(a_3, b_1)}{\|b_1\|^2} b_1 - \frac{\phi(a_3, b_2)}{\|b_2\|^2} b_2 = (1, 0, 0) - \frac{1}{3}(1, 1, 1) - \frac{1/3}{2/3}(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}) \\ &= (1, 0, 0) - (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) - (\frac{1}{6}, \frac{1}{6}, -\frac{1}{3}) = (\frac{1}{2}, -\frac{1}{2}, 0), \ \|b_3\|^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \\ c_1 &= \frac{1}{\sqrt{3}}(1, 1, 1), \ c_2 = \frac{1}{\sqrt{6}}(1, 1, -2), \ c_3 = \frac{1}{\sqrt{2}}(1, -1, 0). \end{aligned}$$

(ii) $n = 3$, $a_1 = (1, 0, 0)$, $a_2 = (1, 1, 0)$, $a_3 = (1, 1, 1)$
(gegenüber dem Beispiel in (i) sind nur a_1 und a_3 vertauscht).

$$\begin{aligned} b_1 &= a_1, \quad \|b_1\|^2 = 1. \\ b_2 &= a_2 - \frac{\phi(a_2, b_1)}{\|b_1\|^2} b_1 = (1, 1, 0) - \frac{1}{1}(1, 0, 0) = (0, 1, 0), \ \|b_2\|^2 = 1. \\ b_3 &= a_3 - \frac{\phi(a_3, b_1)}{\|b_1\|^2} b_1 - \frac{\phi(a_3, b_2)}{\|b_2\|^2} b_2 = (1, 1, 1) - \frac{1}{1}(1, 0, 0) - \frac{1}{1}(0, 1, 0) = (0, 0, 1), \\ &\quad \|b_3\|^2 = 1. \\ c_1 &= b_1, \ c_2 = b_2, \ c_3 = b_3. \end{aligned}$$

Korollar 9.18 *(QR-Zerlegung, Korollar zum Gram-Schmidtschen Orthogonalisierungsverfahren)*
Sei $k \leq n$ und sei $A \in M(n \times k, \mathbb{R})$ eine Matrix mit linear unabhängigen Spalten. Dann existieren Matrizen $Q \in M(n \times k, \mathbb{R})$ und $R \in M(k \times k, \mathbb{R})$ mit den Eigenschaften:

(i)
$$A = Q \cdot R;$$

(ii) die Spalten von Q bilden eine ON-Basis des von ihnen erzeugten Untervektorraums von $M(n \times 1, \mathbb{R})$ bezüglich des Standard-Skalarproduktes;

(iii) R ist eine invertierbare obere Dreiecksmatrix.

Beweis: Die Spalten von A werden mit $a_1, ..., a_k$ bezeichnet. Der von ihnen erzeugte Untervektorraum $V := \text{span}\langle a_1, ..., a_k\rangle$ von $M(n \times 1, \mathbb{R})$ ist zusammen mit der Einschränkung des Standard-Skalarproduktes $\langle\ ,\ \rangle$ ein Euklidischer Vektorraum. Das Gram-Schmidtsche Orthogonalisierungsverfahren liefert eine ON-Basis $c_1, ...c_k$.
Die Basiswechselmatrix $M((a_1, ..., a_k), (c_1, ..., c_k))$ und die inverse Matrix $M((c_1, ..., c_k), (a_1, ..., a_k)) =: R$ sind invertierbare obere Dreiecksmatrizen. Die Matrix mit den Spalten $c_1, ..., c_k$ wird Q genannt. Dann gelten $A = Q \cdot R$ und die anderen Eigenschaften. □

Beispiel 9.19 A ist die 4×3-Matrix
$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$
mit den Spalten a_1, a_2, a_3.
Durch Hingucken findet man die Matrix B, deren Spalten die Orthogonalbasis von $\text{span}\langle a_1, a_2, a_3\rangle$ bilden, die man mit Gram-Schmidt bekommt:
$$B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$
Wieder im Kopf findet man die Matrix Q, deren Spalten c_1, c_2, c_3 man aus denen von B durch Normieren erhält,
$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Ebenfalls im Kopf findet man die obere Dreiecksmatrix R mit $(a_1, a_2, a_3) = (c_1, c_2, c_3) \cdot R$, d.h. mit $A = Q \cdot R$,

$$R = \begin{pmatrix} 2 & 1 & 1 \\ 0 & \sqrt{2} & \sqrt{2} \\ 0 & 0 & \sqrt{3} \end{pmatrix}.$$

Bemerkungen 9.20 (i) $x = (x_1 \cdots x_n)^{tr} \in M(n \times 1, \mathbb{R})$ hat Norm 1 bezüglich des Standard-Skalarproduktes genau dann, wenn $\sum_{i=1}^{n} x_i^2 = 1$ ist. Dann gilt $|x_i| \leq 1$ für alle Einträge von x.
(ii) Die QR-Zerlegung ist interessant für numerische Verfahren, weil die Matrix Q robust gegen kleine Störungen ist (unter anderem wegen (i)) und weil man Störungen bei oberen Dreiecksmatrizen gut kontrollieren kann.

Definition 9.21 Sei V ein Euklidischer Vektorraum mit Skalarprodukt $\phi : V \times V \to \mathbb{R}$. Sei $U \subset V$ ein Untervektorraum.
Der *orthogonale* Untervektorraumraum $U^\perp \subset V$ ist

$$U^\perp := \{x \in V \,|\, \text{für alle } y \in U \text{ ist } x \perp y\}.$$

Es ist klar, dass die so definierte Menge ein Untervektorraum ist.

Definition/Satz 9.22 *Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Sei $U \subset V$ ein endlich-dimensionaler Untervektorraum.*
(a)

$$V = U \oplus U^\perp.$$

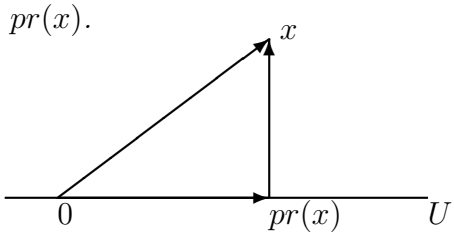
Deswegen heißt $U^\perp$ hier auch orthogonales Komplement *von U.*
(b) (Definition) Für jedes $x \in V$ gibt es eindeutige $y \in U$ und $z \in U^\perp$ mit $x = y + z$. Die Abbildung

$$pr_U : V \to U, \qquad x \mapsto y$$

heißt Projektion *von V auf U. Sie wird pr genannt, wenn klar ist, welcher Unterraum U gemeint ist.*
(c) Man kann $pr(x)$ folgendermaßen ausrechnen. Man wählt eine Orthogonalbasis $(b_1, ..., b_k)$ von U. Dann ist für jedes $x \in V$

$$pr(x) = \sum_{i=1}^{k} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot b_i.$$

(d) Für jedes $x \in V$ gibt es genau ein $\widetilde{y} \in U$, so daß $\|x - \widetilde{y}\|$ minimal ist. Es ist $\widetilde{y} = pr(x)$.



Beweis: (a) Aus ϕ positiv definit folgt $U \cap U^\perp = \{0\}$, also $U + U^\perp = U \oplus U^\perp$. Daher reicht es, $U + U^\perp = V$ zu zeigen.
Sei $(b_1, ..., b_k)$ eine Orthogonalbasis von U. Man definiert eine Abbildung $\widetilde{pr} : V \to U$ durch

$$\widetilde{pr}(x) := \sum_{i=1}^{k} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot b_i.$$

Dann ist

$$\begin{aligned} \phi(x - \widetilde{pr}(x), b_j) &= \phi(x, b_j) - \sum_{i=1}^{k} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot \phi(b_i, b_j) \\ &= \phi(x, b_j) - \phi(x, b_j) = 0, \end{aligned}$$

also $x - \widetilde{pr}(x) \perp b_j$, also $x - \widetilde{pr}(x) \perp U$, also $x - \widetilde{pr}(x) \in U^\perp$. Mit

$$x = \widetilde{pr}(x) + (x - \widetilde{pr}(x))$$

und $\widetilde{pr}(x) \in U$ folgt $V = U + U^\perp$.
(b) Definition.
(c) Der Beweis von (a) zeigt $pr = \widetilde{pr}$.
(d) Sei $x \in V$ und $y \in U$. Es ist

$$\begin{aligned} \|x - y\|^2 &= \|(x - pr(x)) + (pr(x) - y)\|^2 \\ &= \|x - pr(x)\|^2 + 2\phi(x - pr(x), pr(x) - y) + \|pr(x) - y\|^2 \\ &= \|x - pr(x)\|^2 + \|pr(x) - y\|^2 \end{aligned}$$

denn $x - pr(x) \in U^\perp$, $pr(x) - y \in U$. Der Abstand $\|x - y\|$ ist also genau dann minimal, wenn $pr(x) = y$ ist. □

Bemerkungen 9.23 (i) Satz 9.22 (c)+(d) hat viele Anwendungen. Normalerweise sucht man innerhalb eines endlich-dimensionalen Unterraums $U \subset V$ das Element $\widetilde{y}$, das ein Element $x \in V$ am besten approximiert.
Das ist offensichtlich relevant in praktischen Problemen aus der Wirtschaft.
Innerhalb der Mathematik ist V häufig ein Vektorraum von gewissen Funktionen und U ist ein Unterraum von spezielleren Funktionen, z.B.

$$V = \mathcal{C}^0([-1, 1], \mathbb{R}) \quad \text{und} \quad U = \mathbb{R}[t]_{\leq n} \text{ für ein } n \in \mathbb{N},$$

oder

$$\begin{aligned} V &= \{f : \mathbb{R} \to \mathbb{R} \text{ stetig } \mid f(x + 2\pi) = f(x)\} \quad \text{und} \\ U &= \{f = a_0 + \sum_{p=1}^{n} (a_p \cos(px) + b_p \sin(px)) \mid a_p, b_p \in \mathbb{R}\}. \end{aligned}$$

Das erste Beispiel führt zu den Legendre-Polynomen, das zweite zu trigonometrischen Polynomen und Fourier-Reihen.

(ii) Im Satz 9.22 ist es wichtig, dass U ein *endlich-dimensionaler* Unterraum ist. Im folgenden Beispiel 9.24 ist das nicht erfüllt. Da ist zwar $U + U^\perp = U \oplus U^\perp$, aber $U \oplus U^\perp$ ist kleiner als V.

(iii) Aber falls V ein *Hilbertraum* und $U \subset V$ ein *abgeschlossener Unterraum* ist, gilt $U \oplus U^\perp = V$ (Definitionen und Beweis in der Funktionalanalysis).

Beispiel 9.24 Sei $V := C^0([0,1], \mathbb{R})$, $U := \mathbb{R}[t]$. Dann ist $U \subsetneq V$.
Behauptung: $U^\perp = \{0\}$. Also ist $U + U^\perp = U \oplus U^\perp = U \subsetneq V$.
(Zitat aus der Analysis:) *Approximationssatz von Weierstrass:*

$$\forall\ f \in V \quad \forall\ \varepsilon > 0 \quad \exists\ g \in U \quad \forall\ t \in [0,1] \quad |f(t) - g(t)| < \varepsilon.$$

Indirekter Beweis der Behauptung: Sei $U^\perp \supsetneq \{0\}$ und $f \in U^\perp - \{0\}$. Sei $a := \phi_{int}(f,f) > 0$. Sei $b := \max(|f(t)| \,|\, t \in [0,1])$. Sei $\varepsilon := \frac{a}{2b}$. Wähle ein $g \in U$ mit $\forall\ t \in [0,1] \quad |f(t) - g(t)| < \varepsilon$.
Dann ist

$$\begin{aligned}
\phi_{int}(f,g) &= \int_0^1 f(t)g(t)\mathrm{d}t = \int_0^1 f(t)\left(f(t) - (f(t) - g(t))\right)\mathrm{d}t \\
&= \int_0^1 f^2(t)\mathrm{d}t - \int_0^1 f(t)(f(t) - g(t))\mathrm{d}t \\
&\geq \phi_{int}(f,f) - \int_0^1 |f(t)| \cdot |f(t) - g(t)|\mathrm{d}t \\
&\geq a - b \cdot \varepsilon = a - \frac{a}{2} = \frac{a}{2},
\end{aligned}$$

also $f \notin U^\perp$, Widerspruch. □

Die letzten drei Sätze 9.25, 9.26 und 9.27 behandeln reelle symmetrische Matrizen, die nicht notwendig positiv definit sind. Man kann ihre Aussagen auch übersetzen in Aussagen über reelle symmetrische Bilinearformen. Die Beweise dieser drei Sätze werden in dieser Vorlesung LA I nicht gegeben, aber in der Vorlesung LA IIa oder der Vorlesung LA IIb.

Satz 9.25 *(Spektralsatz für reelle symmetrische Matrizen, hier ohne Beweis)*
Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Es gilt:

(i) $P_A(t) = \prod_{i=1}^n (t - \lambda_i)$ mit $\lambda_i \in \mathbb{R}$. Also sind alle Eigenwerte von A (als komplexer Matrix) reell.

(ii) Die Matrix A ist diagonalisierbar. Mit Satz 8.11 folgt: $M(n \times 1, \mathbb{R})$ ist direkte Summe aller Eigenräume.

(iii) Die Eigenräume zu verschiedenen Eigenwerten sind orthogonal bezüglich des Standardskalarproduktes auf $M(n \times 1, \mathbb{R})$.

(iv) Daher gibt es eine ON-Basis von $M(n \times 1, \mathbb{R})$ *(bezüglich des Standardskalarproduktes) aus Eigenvektoren von A. Ist T eine Matrix, deren Spalten eine solche ON-Basis bilden, so ist* $T^{-1} = T^{tr}$ *(dann heißt T orthogonal), und es ist*

$$T^{tr} \cdot A \cdot T = T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Satz/Definition 9.26 *(Satz von Sylvester, hier ohne Beweis)*

(a) (Matrix-Version) Sei $A \in M(n \times n, \mathbb{R})$ *symmetrisch. Seien* $\lambda_1, ..., \lambda_n \in \mathbb{R}$ *die Eigenwerte von A. Jedes* $T \in GL(n, \mathbb{R})$ *mit*

$$T^{tr} \cdot A \cdot T = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

für gewisse $\alpha_1, ..., \alpha_n \in \mathbb{R}$ *erfüllt*

$$\begin{aligned} (\text{die Anzahl der positiven } \alpha_i) &= (\text{die Anzahl der positiven } \lambda_i) =: r_+, \\ (\text{die Anzahl der } \alpha_i = 0) &= (\text{die Anzahl der } \lambda_i = 0) =: r_0, \\ (\text{die Anzahl der negativen } \alpha_i) &= (\text{die Anzahl der negativen } \lambda_i) =: r_-. \end{aligned}$$

Das Tripel (r_+, r_0, r_-) *heißt* Signatur *von A. Die Zahlen erfüllen auch*

$$r_+ + r_0 + r_- = n \quad \text{und} \quad r_0 = n - \operatorname{rang} A.$$

(b) (Abstrakte Version) Sei V ein endlich-dimensionaler $\mathbb{R}$*-Vektorraum und* $\phi : V \times V \to \mathbb{R}$ *eine symmetrische Bilinearform. Sei*

$$\begin{aligned} r_+ &:= \max(\dim U \mid U \text{ ist ein Untervektorraum von } V, \\ &\qquad\qquad \text{auf dem } \phi \text{ positiv definit ist}, \\ r_0 &:= \dim \operatorname{Rad}(\phi), \\ r_- &:= \max(\dim U \mid U \text{ ist ein Untervektorraum von } V, \\ &\qquad\qquad \text{auf dem } \phi \text{ negativ definit ist}. \end{aligned}$$

Es gibt Unterräume $U_1, U_2 \subset V$ *mit den Eigenschaften*

$$\begin{aligned} &V = U_1 \oplus \operatorname{Rad}(\phi) \oplus U_2, \\ &U_1 \perp U_2 \text{ (d.h. } u_1 \perp u_2 \text{ für alle } u_1 \in U_1, u_2 \in U_2), \\ &\phi \text{ ist positiv definit auf } U_1 \text{ und negativ definit auf } U_2. \end{aligned}$$

Es gilt immer:

$$\dim U_1 = r_+, \quad \dim U_2 = r_-.$$

Das Tripel (r_+, r_0, r_-) *heißt* Signatur *von* ϕ.

Satz 9.27 *(Hauptminorenkriterium für Definitheit, hier ohne Beweis)*
Sei $A = (a_{ij}) \in M(n \times n, \mathbb{R})$ *eine symmetrische Matrix.*
(a) (Definition) Ihre Hauptminoren sind die Determinanten $\det A_k$ *der Untermatrizen*

$$A_k := (a_{ij})_{i,j=1,...,k} \in M(k \times k, \mathbb{R}) \text{ für } k = 1, 2, ..., n.$$

(b) (Satz) A ist genau dann positiv definit, wenn alle Hauptminoren positiv sind.

Beispiel 9.28 Sei $r > 1$. Die Hauptminoren der Matrix

$$\begin{pmatrix} \frac{1}{r-1} & \frac{1}{r} \\ \frac{1}{r} & \frac{1}{r+1} \end{pmatrix}$$

sind $\frac{1}{r-1} > 0$ und

$$\frac{1}{r-1} \cdot \frac{1}{r+1} - \frac{1}{r} \cdot \frac{1}{r} = \frac{1}{r^2-1} - \frac{1}{r^2} > 0.$$

Daher ist die Matrix positiv definit.