Sebastian Thomas RWTH Aachen, 2018
13. März 2018 *

Lineare Algebra für Informatiker
Manuskript

Die Vorlesung richtet sich in erster Linie an Studierende des Studiengangs Bachelor of Science Informatik.
Es werden die grundlegenden Konzepte der linearen Algebra wie Vektorräume und deren Homomorphismen, der Dimensionsbegriff, der Matrixkalkül für Homomorphismen, Determinanten, etwas Eigenwerttheorie und eine Einführung in die Theorie der Skalarprodukträume behandelt. Als Anwendungen geben wir insbesondere einen Einblick in die lineare Kodierungstheorie.
Benötigte Vorkenntnisse sind einige Inhalte aus der Vorlesung *Diskrete Strukturen*, welche wir in Anhang A zusammenfassen. In Anhang B finden sich als Zusatzmaterial einige Grundlagen zu Algebren und Körpererweiterungen.
Für Hinweise und Korrekturen danke ich meinen Studierenden und Tutoren; insbesondere Daniel Andres, Jens Bürger, Timo Gervens, J. Isabel Klöter, Jan-Frederik Konopka, Mirela D. Mileva, Robin Mross, Erik Müller, Cedric Sodhi, Tomas L. De Stefano, Daria Tripuk und Jan M. Tönshoff für zahlreiche Korrekturen in vorherigen Versionen dieses Manuskripts. Über weitere Hinweise auf Fehler und Unklarheiten freue ich mich.

Aachen, 13. März 2018
Sebastian Thomas

Inhaltsverzeichnis

1 Vektorräume

In diesem Abschnitt werden Vektorräume über einem Körper eingeführt. Neben den zugehörigen Homomorphismen, welche wir etwas später studieren werden, bilden Vektorräume die zentralen Objekte der linearen Algebra.
Aus der Schule kennen wir $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ als Formalisierung der Anschauungsebene, wobei $\mathbb{R}$ die Menge der reellen Zahlen bezeichnet. Punkte dieser Ebene können addiert werden, indem wir sie als sogenannte Pfeilvektoren vom „Ursprung", d.h. dem Punkt $(0, 0)$, zu eben jenen Punkten interpretieren und dann eine

Dieses Vorlesungsmanuskript enthält Inhalte zur Veranstaltung *Lineare Algebra für Informatiker*; vom Autor zuletzt gehalten im Sommersemester 2017 an der RWTH Aachen. Es basiert auf dem Manuskript [14] sowie zu Teilen auf dem Manuskript [13].
*Dies ist eine leicht überarbeitete Version vom 11. Oktober 2018. (Version 2.6.3)

geometrische Operation mit diesen Pfeilen durchführen: Sind zwei Punkte gegeben, so erhalten wir die Summe dieser Punkte, indem wir die zugehörigen Pfeile aneinanderhängen. Ebenso können Punkte mit reellen Zahlen multipliziert werden, indem wir die Richtung des Pfeils zum gegebenen Punkt beibehalten und die Länge um die angegebene reelle Zahl strecken. Bei dieser Multiplikation von reellen Zahlen mit Punkten der Ebene spricht man dann von einer sogenannten *Skalarmultiplikation*.
Formal sind Addition und Skalarmultiplikation *komponentenweise* erklärt, d.h. es ist

$$\begin{aligned}(x, y) + (x', y') &= (x + x', y + y'),\\ a(x, y) &= (ax, ay)\end{aligned}$$

für $a, x, y, x', y' \in \mathbb{R}$. Für diese Operationen lassen sich Eigenschaften nachweisen; zum Beispiel gilt

$$(x, y) + ((x', y') + (x'', y'')) = ((x, y) + (x', y')) + (x'', y'')$$

für $x, y, x', y', x'', y'' \in \mathbb{R}$ sowie

$$a((x, y) + (x', y')) = a(x, y) + a(x', y')$$

für $a, x, y, x', y' \in \mathbb{R}$. Diese Eigenschaften lassen sich direkt aus der Definition von Addition und Skalarmultiplikation herleiten. Manche Eigenschaften, welche man nun zeigen kann, lassen sich (sogar) aus bereits vorher bewiesenen Eigenschaften folgern, und zwar *ohne* die Definition von Addition und Skalarmultiplikation zu benutzen.
Für $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$, die Formalisierung des Anschauungsraums, lassen sich ebenfalls eine Addition und eine Skalarmultiplikation komponentenweise erklären. Diese haben dann analoge geometrische Interpretationen sowie analoge Eigenschaften.
In diesem Abschnitt werden wir die Fälle $\mathbb{R}^2$ und $\mathbb{R}^3$ wie folgt verallgemeinern: Zunächst lassen sich ganz analog eine komponentenweise Addition und eine komponentenweise Skalarmultiplikation auf $\mathbb{R}^n$ für jede natürliche Zahl n definieren, und diese erhalten dann auch analoge Eigenschaften. Zur Herleitung dieser Eigenschaften spielen nur wenige Aspekte der reellen Zahlen eine Rolle: Wenn wir ausschließlich mit rationalen Zahlen, also Brüchen von ganzen Zahlen, arbeiten, so werden wir diesen Zahlbereich nie verlassen. Wir können also eine ähnliche Struktur auch auf $\mathbb{Q}^n$ für jede natürliche Zahl n definieren. Noch allgemeiner gilt dies sogar für K^n, die Menge der n-Tupel mit Einträgen in einem Körper K.
Wir abstrahieren sogar noch weiter: Im Laufe der Zeit hat sich herausgestellt, dass die oben angedeuteten Eigenschaften noch für bedeutend mehr Strukturen gelten, bei denen wir eine Menge zusammen mit einer Addition und einer Skalarmultiplikation gegeben haben, bei denen die Elemente dieser Menge aber nicht wie n-Tupel „aussehen“ müssen. Wir erleben dann ein ähnliches Verhalten wie bei K^n, es gelten analoge Eigenschaften. Ferner lassen sich auch hier einige Eigenschaften aus bereits bewiesenen Eigenschaften folgern. Um alle diese Beispiele, von denen wir einige im späteren Verlauf kennenlernen werden, unter einen Hut zu bringen, wählen wir einen *axiomatischen* Ansatz: Unser Blickpunkt entfernt sich dann von der konkreten „Gestalt“ der Elemente sowie der Definition von Addition und Skalarmultiplikation im einzelnen Beispiel; stattdessen treten einige Eigenschaften der Operationen in den Mittelpunkt.
Konkret bedeutet dies Folgendes: Wir betrachten sogenannte *Vektorräume*, d.h. Strukturen bestehend aus einer Menge, einer Addition und einer Skalarmultiplikation, für welche gewisse Eigenschaften gelten, die uns aus den Beispielen bereits vertraut sind. Diese Eigenschaften, welche für Vektorräume gefordert werden, nennen wir *Vektorraumaxiome*. Alle Eigenschaften, welche sich auf die Vektorraumaxiome zurückführen und bereits in diesem abstrakten Rahmen beweisen lassen, gelten dann für *alle* Vektorräume, unabhängig davon, wie die Elemente jeweils „aussehen“. Der Vorteil dieses axiomatischen Ansatzes liegt auf der Hand: Zum einen entwickeln wir, wie bereits angedeutet, unsere Theorie für alle uns bereits bekannten Beispiele von Vektorräumen simultan. Zum anderen können wir unsere Theorie auf weitere Beispiele von Vektorräumen anwenden, sobald wir uns davon überzeugt haben, dass es sich bei der vorliegenden Struktur um einen Vektorraum handelt. Hierzu müssen wir dann lediglich die Vektorraumaxiome nachweisen, alle anderen Eigenschaften von Vektorräumen folgen bereits aus diesen.
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein Körper K gegeben. Ab Bemerkung (1.7), bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein K-Vektorraum V gegeben.

Begriffsbildung

Wir beginnen mit der Definition eines K-Vektorraums.

(1.1) Definition (Vektorraum)**.** Ein *Vektorraum über* K (oder K*-Vektorraum*) besteht aus einer Menge V zusammen mit einer Abbildung $V \times V \to V$, $(v, w) \mapsto v+w$, genannt *Addition*, und einer Abbildung $K \times V \to V$, $(a, v) \mapsto av = a \cdot v$, genannt *Skalarmultiplikation*, derart, dass folgende Axiome gelten.

- *Assoziativität der Addition.* Für $v, w, x \in V$ ist $v + (w + x) = (v + w) + x$.
- *Existenz des Nullvektors.* Es existiert ein $n \in V$ derart, dass für $v \in V$ stets $n + v = v + n = v$ gilt. Es lässt sich zeigen, dass ein solches n eindeutig bestimmt ist. Wir bezeichnen es mit 0. Dann haben wir also $0 + v = v + 0 = v$ für alle $v \in V$.
- *Existenz der negativen Vektoren.* Für jedes $v \in V$ existiert ein $w \in V$ mit $w + v = v + w = 0$. Es lässt sich zeigen, dass ein solches w durch v eindeutig bestimmt ist. Wir bezeichnen es mit $-v$. Dann haben wir also $(-v) + v = v + (-v) = 0$.
- *Kommutativität der Addition.* Für $v, w \in V$ ist $v + w = w + v$.
- *Assoziativität der Skalarmultiplikation.* Für $a, b \in K$, $v \in V$ ist $a(bv) = (ab)v$.
- *Neutralität der Eins bzgl. der Skalarmultiplikation.* Für $v \in V$ ist $1v = v$.
- *Distributivität.* Für $a, b \in K$, $v \in V$ ist $(a+b)v = (av)+(bv)$. Für $a \in K$, $v, w \in V$ ist $a(v+w) = (av)+(aw)$.

Ein Element von K wird *Skalar* von V genannt. Ein Element von V wird *Vektor* in V genannt. Das Element 0 in V wird *Nullvektor* in V genannt. Für einen Vektor v in V wird $-v$ der *negative Vektor* zu v genannt.

Die ersten vier Axiome besagen gerade, dass jeder K-Vektorraum eine abelsche Gruppe ist.

(1.2) Konvention. Es sei ein K-Vektorraum V gegeben. Meistens lassen wir die Klammern um Produkte aus Skalaren und Vektoren weg, d.h. es gelte *Punkt- vor Strichrechnung.* Da für $a, b \in K$, $v \in V$ stets $(ab)v = a(bv)$ gilt, schreiben wir im außerdem meist kurz $abv := (ab)v = a(bv)$.

(1.3) Notation (Subtraktion)**.** Es sei ein K-Vektorraum V gegeben. Für $v, w \in V$ schreiben wir

$$v - w := v + (-w) = (-w) + v.$$

(1.4) Beispiel.

(a) Es wird K ein K-Vektorraum mit Addition gegeben durch die Addition des Körpers K und Skalarmultiplikation gegeben durch die Multiplikation des Körpers K.

(b) Für $n \in \mathbb{N}_0$ wird K^n ein K-Vektorraum mit Addition gegeben durch

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

für $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in K^n$, und Skalarmultiplikation gegeben durch

$$a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n)$$

für $a \in K$, $(x_1, \ldots, x_n) \in K^n$. Der Nullvektor von K^n ist gegeben durch

$$0 = (0, \ldots, 0).$$

Für $(x_1, \ldots, x_n) \in K^n$ ist

$$-(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n).$$

(c) Für $m, n \in \mathbb{N}_0$ wird $K^{m \times n}$ ein K-Vektorraum mit Addition gegeben durch die Addition von Matrizen und Skalarmultiplikation gegeben durch die Skalarmultiplikation von Matrizen.

(d) Für jede Menge I wird K^I ein K-Vektorraum mit Addition gegeben durch

$$x + y = (x_i + y_i)_{i \in I}$$

für $x, y \in K^I$, und Skalarmultiplikation gegeben durch

$$a \cdot x = (a \cdot x_i)_{i \in I}$$

für $a \in K$, $x \in K^I$. Der Nullvektor von K^I ist gegeben durch

$$0 = (0)_{i \in I}.$$

Für $x \in K^I$ ist

$$-x = (-x_i)_{i \in I}.$$

(e) Für jede Menge X wird $\mathrm{Map}(X, K)$ ein K-Vektorraum mit Addition gegeben durch

$$(f + g)(x) = f(x) + g(x)$$

für $x \in X$, $f, g \in \mathrm{Map}(X, K)$, und Skalarmultiplikation gegeben durch

$$(af)(x) = af(x)$$

für $x \in X$, $a \in K$, $f \in \mathrm{Map}(X, K)$. Die Null von $\mathrm{Map}(X, K)$ ist gegeben durch

$$0(x) = 0$$

für $x \in X$. Für $f \in \mathrm{Map}(X, K)$ ist

$$(-f)(x) = -f(x)$$

für $x \in X$.

(f) Jede einelementige Menge wird ein K-Vektorraum (mit der einzig möglichen Addition und der einzig möglichen Skalarmultiplikation).

(g) Der Polynomring $K[X]$ wird ein K-Vektorraum mit Addition gegeben durch die Addition des Rings $K[X]$ und Skalarmultiplikation gegeben durch Einschränkung der Multiplikation des Rings $K[X]$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(1.5) Konvention.

(a) Sofern nicht anders erwähnt, betrachten wir K als K-Vektorraum wie in Beispiel (1.4)(a).

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Sofern nicht anders erwähnt, betrachten wir K^n als K-Vektorraum wie in Beispiel (1.4)(b).

(c) Es seien $m, n \in \mathbb{N}_0$ gegeben. Sofern nicht anders erwähnt, betrachten wir $K^{m \times n}$ als K-Vektorraum wie in Beispiel (1.4)(c).

(d) Es sei eine Menge I gegeben. Sofern nicht anders erwähnt, betrachten wir K^I als K-Vektorraum wie in Beispiel (1.4)(d).

(e) Es sei eine Menge X gegeben. Sofern nicht anders erwähnt, betrachten wir $\mathrm{Map}(X, K)$ als K-Vektorraum wie in Beispiel (1.4)(e).

(f) Sofern nicht anders erwähnt, betrachten wir $K[X]$ als K-Vektorraum wie in Beispiel (1.4)(g).

Auch für Anwendungen in der Informatik lassen sich Vektorräume nutzen: Bei der JPEG-Komprimierung wird ein Bild zunächst in drei Farbanteile zerlegt. Anschließend werden die einzelnen Farbkomponenten in kleine Blöcke aus 8×8 Pixeln wie folgt unterteilt.

(1.6) Anwendungsbeispiel. Ein digitales Bild mit einer Breite und Höhe aus jeweils acht Pixeln lässt sich als Vektor in $\mathbb{R}^{[0,7]\times[0,7]}$, also als eine durch $[0,7]\times[0,7]$ indizierte Matrix, auffassen.

Wir betrachten einige elementare Eigenschaften von Vektorräumen, welche Verallgemeinerungen der entsprechenden Eigenschaften für Körper sind.
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein K-Vektorraum V gegeben.

(1.7) Bemerkung.

(a) Es seien $v, w, x \in V$ gegeben. Genau dann gilt $v + x = w$, wenn $x = -v + w$ ist.

(b) Es seien $a \in K^\times$ und $v, x \in V$ gegeben. Genau dann gilt $ax = v$, wenn $x = a^{-1}v$ ist.

Beweis.

(a) Dies folgt aus Bemerkung (A.79).

(b) Wenn $ax = v$ gilt, dann auch

$$x = 1x = a^{-1}ax = a^{-1}v.$$

Umgekehrt, wenn $x = a^{-1}v$ ist, dann haben wir nach Proposition (A.78)(c) auch

$$v = (a^{-1})^{-1}x = ax.$$ □

Alternativer Beweis.

(b) Dies gilt nach [12, Bem. (7.5)]. □

(1.8) Korollar.

(a) Es seien $v, x, y \in V$ gegeben. Genau dann gilt $v + x = v + y$, wenn $x = y$ ist.

(b) Es seien $a \in K^\times$, $x, y \in V$ gegeben. Genau dann gilt $ax = ay$, wenn $x = y$ ist.

Beweis.

(a) Dies folgt aus Korollar (A.80).

(b) Wenn $ax = ay$ gilt, dann nach Bemerkung (1.7)(b) auch

$$x = a^{-1}ay = 1y = y.$$ □

Alternativer Beweis.

(b) Dies gilt nach [12, Kor. (7.6)]. □

(1.9) Proposition.

(a) Für $v \in V$ gilt $0v = 0$.

(b) Für $a \in K$ gilt $a \cdot 0 = 0$.

(c) Für $a \in K$, $v \in V$ gilt $(-a)v = a(-v) = -av$.

(d) Für $v \in V$ gilt $(-1)v = -v$.

(e) Für $a \in K$, $v \in V$ gilt $(-a)(-v) = av$.

Beweis.

(a) Für $v \in V$ gilt

$$0v + 0v = (0 + 0)v = 0v,$$

also $0v = -(0v) + 0v = 0$ nach Bemerkung (1.7)(a).

(b) Für $a \in K$ gilt

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

also $a \cdot 0 = -(a \cdot 0) + a \cdot 0 = 0$ nach Bemerkung (1.7)(a).

(c) Es seien $a \in K$ und $v \in V$ gegeben. Dann gilt

$$(-a)v + av = (-a + a)v = 0v = 0$$

nach (a) und damit $-av = (-a)v$. Ferner gilt

$$a(-v) + av = a(-v + v) = a \cdot 0 = 0$$

nach (b) und damit $-av = a(-v)$.

(d) Für $v \in V$ ist

$$(-1)v = -1v = -v$$

nach (c).

(e) Für $a \in K$, $v \in V$ ist

$$(-a)(-v) = -(-a)v = -(-av) = av$$

nach (d). □

(1.10) Lemma. Es seien $a \in K$, $v \in V$ gegeben. Genau dann gilt $av = 0$, wenn $a = 0$ oder $v = 0$ ist.

Beweis. Wenn $av = 0$ und $a \neq 0$ ist, dann folgt $v = a^{-1} \cdot 0 = 0$ nach Bemerkung (1.7) und Proposition (1.9)(b). Umgekehrt, wenn $a = 0$ oder $v = 0$ ist, so folgt $av = 0$ nach Proposition (1.9)(a), (b). □

(1.11) Korollar. Es seien $a, b \in K$, $v \in V \setminus \{0\}$ gegeben. Genau dann gilt $av = bv$, wenn $a = b$ ist.

Beweis. Wenn $a = b$ ist, dann auch $av = bv$. Es gelte umgekehrt $av = bv$, so dass $(a - b)v = av - bv = 0$. Da $v \neq 0$ ist, folgt $a - b = 0$ nach Lemma (1.10), also $a = b$. □

Untervektorräume

Als nächstes betrachten wir sogenannte Untervektorräume, d.h. Vektorräume, die in geeigneter Weise als Teilmengen von gegebenen Vektorräumen auftreten.
Die Idee lässt sich dabei an Hand des folgenden anschaulichen Beispiels erklären: Von der Anschauungsebene können wir zum Anschauungsraum übergehen, indem wir eine weitere Richtung betrachten, welche nicht bereits in der Ebene liegt; wir fügen eine weitere Achse in ein Koordinatensystem hinzu. Sowohl Addition als auch Skalarmultiplikation der Ebene und des Raums lassen sich durch geometrische Operationen interpretieren. Wenn wir nun den Raum wie gerade beschrieben als Erweiterung der Ebene auffassen, so entsprechen sich die geometrischen Operationen der Ebene und des Raums, wenn wir sie nur auf Punkte in der Ebene anwenden. Dies entspricht gerade dem Konzept des *Untervektorraums*: Ein Vektorraum, dessen unterliegende Menge in einem (potentiell) größeren Vektorraum liegt, und zwar gerade so, dass die Anwendung der Operationen (Addition bzw. Skalarmultiplikation) des größeren Vektorraums auf die Vektoren des kleineren Vektorraums gerade den Operationen des kleineren Vektorraums entspricht.
Werden wir nun etwas formaler: Wie bereits zu Anfang dieses Abschnitts erwähnt, lassen sich $\mathbb{R}^2$ als Formalisierung der Anschauungsebene und $\mathbb{R}^3$ als Formalisierung des Anschauungsraums auffassen. Die Erweiterung von $\mathbb{R}^2$ zu $\mathbb{R}^3$ bedeutet, dass wir die injektive Abbildung

$$\iota\colon \mathbb{R}^2 \to \mathbb{R}^3,\ (x, y) \mapsto (x, y, 0)$$

betrachten und $\mathbb{R}^2$ mit deren Bild $\operatorname{Im} \iota = \{(x, y, 0) \mid x, y \in \mathbb{R}\}$ identifizieren. Dabei lassen sich die Operationen von $\mathbb{R}^2$ auf die Elemente von $\operatorname{Im} \iota$ wie folgt übersetzen: Für $x, y, x', y' \in \mathbb{R}$ ist $(x, y) + (x', y') = (x + x', y + y')$ in $\mathbb{R}^2$, durch die Korrespondenz also

$$(x, y, 0) + (x', y', 0) = (x + x', y + y', 0)$$

in $\operatorname{Im} \iota$. Für $a, x, y \in \mathbb{R}$ ist $a(x, y) = (ax, ay)$ in $\mathbb{R}^2$, durch die Korrespondenz also

$$a(x, y, 0) = (ax, ay, 0)$$

in $\operatorname{Im} \iota$. Da sich die Elemente von $\mathbb{R}^2$ und $\operatorname{Im} \iota$ via ι bijektiv entsprechen, wird $\operatorname{Im} \iota$ mit dieser Addition und Skalarmultiplikation ein $\mathbb{R}$-Vektorraum. Andererseits entsprechen diese Operationen gerade den Operationen von $\mathbb{R}^3$, angewandt auf die Elemente von $\operatorname{Im} \iota$, d.h. $\operatorname{Im} \iota$ wird mit diesen Operationen zu einem Untervektorraum von $\mathbb{R}^3$.

(1.12) Definition (Untervektorraum)**.** Ein K-*Untervektorraum* (oder *Untervektorraum* oder *linearer Unterraum* oder *linearer Teilraum*) von V ist ein K-Vektorraum U derart, dass die unterliegende Menge von U eine Teilmenge von V ist, und so, dass für $u, u' \in U$ stets

$$u +^U u' = u +^V u'$$

und für $a \in K$, $u \in U$ stets

$$a \cdot^U u = a \cdot^V u$$

gilt.
Ein Untervektorraum U von V heißt *echt* (oder *strikt*), falls $U \neq V$ gilt.
Es sei ein K-Vektorraum U gegeben. Ist U ein Untervektorraum von V, so schreiben wir $U \leq V$. Ist U kein Untervektorraum von V, so schreiben wir $U \not\leq V$. Ist U ein echter Untervektorraum von V, so schreiben wir $U < V$.

(1.13) Bemerkung. Es ist V ein K-Untervektorraum von V.

Da die Struktur eines Untervektorraums durch die unterliegende Menge festgelegt ist, treffen wir folgende Vereinbarung.

(1.14) Konvention. Es sei eine Teilmenge U von V gegeben. Da die Addition bzw. die Skalarmultiplikation jedes K-Untervektorraums von V vollständig durch die Addition bzw. die Skalarmultiplikation von V bestimmt ist, gibt es höchstens eine Vektorraumstruktur auf U so, dass U mit dieser Vektorraumstruktur ein K-Untervektorraum von V wird. Wir sagen daher auch, dass U ein K-Untervektorraum von V *ist*, falls so eine Vektorraumstruktur auf U existiert.

Unter Verwendung von Konvention (1.14) geben wir nun ein Kriterium zur Erkennung von Untervektorräumen an, welches uns die einfache Betrachtung von Beispielen ermöglichen wird.

(1.15) Lemma (Untervektorraumkriterium)**.** Es sei eine Teilmenge U von V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist U ein K-Untervektorraum von V.

(b) Es gilt:

- *Abgeschlossenheit unter der Addition.* Für $u, u' \in U$ ist

$$u + u' \in U.$$

- *Abgeschlossenheit unter dem Nullvektor.* Es ist

$$0 \in U.$$

- *Abgeschlossenheit unter der Skalarmultiplikation.* Für $a \in K$, $u \in U$ ist

$$au \in U.$$

(c) Es gilt:

- Es ist

$$U \neq \emptyset.$$

- Für $a \in K$, $u, u' \in U$ ist
$$au + u' \in U.$$

Beweis. Wir zeigen zuerst die Äquivalenz von Bedingung (a) und Bedingung (b), danach die Äquivalenz von Bedingung (b) und Bedingung (c).
Zunächst gelte Bedingung (a), d.h. es sei U ein Untervektorraum von V. Für $u, u' \in U$ ist dann $u +^V u' = u +^U u' \in U$ und für $a \in K$, $u \in U$ ist $a \cdot^V u = a \cdot^U u \in U$. Ferner gilt
$$0^U = 0^U +^U 0^U = 0^U +^V 0^U$$
und damit $0^V = 0^U \in U$ nach Korollar (1.8)(a). Folglich gilt Bedingung (b).
Nun gelte umgekehrt Bedingung (b). Da für $u, u' \in U$ stets $u +^V u' \in U$ ist, erhalten wir eine wohldefinierte Abbildung $+^U \colon U \times U \to U$, $(u, u') \mapsto u +^V u'$, und da für $a \in K$, $u \in U$ stets $a \cdot^V u \in U$ ist, erhalten wir eine wohldefinierte Abbildung $\cdot^U \colon K \times U \to U$, $(a, u) \mapsto a \cdot^V u$. Um zu zeigen, dass U ein K-Vektorraum mit Addition $+^U$ und Skalarmultiplikation $\cdot^U$ wird, verifizieren wir die Axiome aus Definition (1.1):

- *Assoziativität der Addition.* Für $u, u', u'' \in U$ ist
$$u +^U (u' +^U u'') = u +^V (u' +^V u'') = (u +^V u') +^V u'' = (u +^U u') +^U u''.$$
- *Kommutativität der Addition.* Für $u, u' \in U$ ist
$$u +^U u' = u +^V u' = u' +^V u = u' +^U u.$$
- *Existenz des Nullvektors.* Es ist $0^V \in U$ und für $u \in U$ gilt
$$0^V +^U u = 0^V +^V u = u.$$
Mit der Kommutativität der Addition folgt, dass $0^U = 0^V$ ist.
- *Existenz der negativen Vektoren.* Für $u \in U$ ist $(-u)^V = (-1)^K \cdot^V u \in U$ und es gilt
$$(-u)^V +^U u = (-u)^V +^V u = 0^V = 0^U.$$
Mit der Kommutativität der Addition folgt, dass für $u \in U$ stets $(-u)^U = (-u)^V$ ist.
- *Assoziativität der Skalarmultiplikation.* Für $a, b \in K$, $u \in U$ ist
$$a \cdot^U (b \cdot^U u) = a \cdot^V (b \cdot^V u) = (ab) \cdot^V u = (ab) \cdot^U u.$$
- *Neutralität der Eins bzgl. der Skalarmultiplikation.* Für $u \in U$ ist
$$1 \cdot^U u = 1 \cdot^V u = u.$$
- *Distributivität.* Für $a, b \in K$, $u \in U$ ist
$$(a + b) \cdot^U u = (a + b) \cdot^V u = a \cdot^V u +^V b \cdot^V u = a \cdot^U u +^U b \cdot^U u.$$
Für $a \in K$, $u, u' \in U$ ist
$$a \cdot^U (u +^U u') = a \cdot^V (u +^V u') = a \cdot^V u +^V a \cdot^V u' = a \cdot^U u +^U a \cdot^U u'.$$

Somit wird U in der Tat ein K-Vektorraum mit Addition $+^U$ und Skalarmultiplikation $\cdot^U$. Nach Definition der Addition und der Skalarmultiplikation von U ist dann U aber sogar ein Untervektorraum von V, d.h. es gilt Bedingung (a).
Als nächstes gelte Bedingung (b). Da U abgeschlossen unter dem Nullvektor ist, gilt $0 \in U$ und damit insbesondere $U \neq \emptyset$. Sind $a \in K$, $u, u' \in U$ gegeben, so ist ferner $au \in U$, da U abgeschlossen unter der Skalarmultiplikation ist, und folglich $au + u' \in U$, da U abgeschlossen unter der Addition ist. Wir haben somit die Gültigkeit von Bedingung (c) gezeigt.
Schließlich gelte Bedingung (c). Für $u, u' \in U$ gilt dann $u + u' = 1u + u' \in U$. Da $U \neq \emptyset$ ist, gibt es ferner ein Element $u_0 \in U$, und es folgt $0 = -u_0 + u_0 = (-1)u_0 + u_0 \in U$ nach Proposition (1.9)(d). Für $a \in K$, $u \in U$ gilt schließlich $au = au + 0 \in U$. Somit haben wir Bedingung (b) gezeigt.
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Um zu zeigen, dass eine gegebene Teilmenge von V ein Untervektorraum ist, müssen wir also insbesondere *keine* Vektorraumaxiome verifizieren; es genügt die Bedingungen aus Lemma (1.15)(b) oder die Bedingungen aus Lemma (1.15)(c) zu zeigen.

(1.16) Beispiel.

(a) Es ist $\{(x, -x) \mid x \in \mathbb{R}\}$ ein $\mathbb{R}$-Untervektorraum von $\mathbb{R}^2$.

(b) Es ist $\{(1 + x, -x) \mid x \in \mathbb{R}\}$ kein $\mathbb{R}$-Untervektorraum von $\mathbb{R}^2$.

Beweis.

(a) Es sei $U := \{(x, -x) \mid x \in \mathbb{R}\}$. Dann ist $0 = (0, 0) = (0, -0) \in U$, also insbesondere $U \neq \emptyset$. Es seien $a \in \mathbb{R}$, $u, u' \in U$ gegeben. Dann gibt es $x, x' \in \mathbb{R}$ mit $u = (x, -x)$, $u' = (x', -x')$, es folgt also

$$au + u' = a(x, -x) + (x', -x') = (ax + x', a(-x) + (-x')) = (ax + x', -(ax + x')) \in U.$$

Nach dem Untervektorraumkriterium (1.15) ist U ein Untervektorraum von $\mathbb{R}^2$.

(b) Es sei $U := \{(1 + x, -x) \mid x \in \mathbb{R}\}$. Für $x \in \mathbb{R}$ ist genau dann $1 + x = 0$, wenn $-x = 1$ ist. Dies bedeutet, dass $0 = (0, 0) \notin U$ ist. Nach dem Untervektorraumkriterium (1.15) ist U kein Untervektorraum von $\mathbb{R}^2$. □

(1.17) Bemerkung.

(a) Für jedes $v \in V$ ist

$$Kv = \{av \mid a \in K\}$$

ein K-Untervektorraum von V.

(b) Es ist $\{0\}$ ein K-Untervektorraum von V.

Beweis.

(a) Zunächst seien $u, u' \in Kv$ gegeben. Dann gibt es $a, a' \in K$ mit $u = av$ und $u' = a'v$. Wir erhalten

$$u + u' = av + a'v = (a + a')v \in Kv.$$

Somit ist Kv abgeschlossen unter der Addition. Ferner ist

$$0 = 0v \in Kv$$

nach Proposition (1.9)(a), d.h. Kv ist abgeschlossen unter dem Nullvektor. Schließlich seien $c \in K$, $u \in Kv$ gegeben. Dann gibt es ein $a \in K$ mit $u = av$ und es folgt

$$cu = cav \in Kv.$$

Folglich ist Kv auch abgeschlossen unter der Skalarmultiplikation. Insgesamt ist Kv ein Untervektorraum von V nach dem Untervektorraumkriterium (1.15).

(b) Nach Proposition (1.9)(b) ist

$$K \cdot 0 = \{a \cdot 0 \mid a \in K\} = \{0 \mid a \in K\} = \{0\},$$

so dass $\{0\}$ nach (a) ein Untervektorraum von V ist. □

Linearkombinationen

Unser nächstes Ziel wird es sein, Untervektorräume effizient zu beschreiben. Hierzu benötigen wir den Begriff der linearen Hülle.

Zunächst wollen wir das Konzept wieder an Hand des Anschauungsraums erläutern. Wie oben fassen wir hierzu die Anschauungsebene als Untervektorraum des Anschauungsraums auf, d.h. wir betrachten die Teilmenge $\{(x, y, 0) \mid x, y \in \mathbb{R}\}$ von $\mathbb{R}^3$. Diese Teilmenge können wir uns als von den Geraden $\{(x, 0, 0) \mid x \in \mathbb{R}\}$ und $\{(0, y, 0) \mid y \in \mathbb{R}\}$ „aufgespannt“ vorstellen. Die Gerade $\{(x, 0, 0) \mid x \in \mathbb{R}\} = \{x(1, 0, 0) \mid x \in \mathbb{R}\} = \mathbb{R}(1, 0, 0)$

wird wiederum durch den Vektor $(1, 0, 0)$ festgelegt, es handelt sich bei den Punkten dieser Gerade um alle Vielfachen von $(1, 0, 0)$ im Sinne der Skalarmultiplikation. Entsprechend für die Gerade $\{(0, y, 0) \mid y \in \mathbb{R}\} = \mathbb{R}(0, 1, 0)$, welche durch den Vektor $(0, 1, 0)$ festgelegt ist. Die Ebene wird also durch die Vektoren $(1, 0, 0)$ und $(0, 1, 0)$ festgelegt. Wir werden sagen, dass die Vektoren $(1, 0, 0)$ und $(0, 1, 0)$ die Ebene *erzeugen* (oder *aufspannen*).
Formal bedeutet dies Folgendes: Haben wir einen Punkt $(x, y, 0)$ für $x, y \in \mathbb{R}$ gegeben, so lässt sich dieser als Summe von Vektoren aus $\mathbb{R}(1, 0, 0)$ und $\mathbb{R}(0, 1, 0)$ beschrieben, es ist nämlich gerade

$$(x, y, 0) = (x, 0, 0) + (0, y, 0).$$

Andererseits ist $(x, 0, 0)$ ein Vielfaches von $(1, 0, 0)$ bzw. $(0, y, 0)$ ein Vielfaches von $(0, 1, 0)$: es gilt

$$(x, 0, 0) = x(1, 0, 0),$$
$$(0, y, 0) = y(0, 1, 0).$$

Insgesamt erhalten wir

$$(x, y, 0) = x(1, 0, 0) + y(0, 1, 0),$$

d.h. $(x, y, 0)$ ist eine *Linearkombination* von $(1, 0, 0)$ und $(0, 1, 0)$ im folgenden Sinn:

(1.18) Definition (Linearkombination)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $s = (s_1, \dots, s_n)$ in V gegeben. Die *K-lineare Hülle* (oder *lineare Hülle* oder *Spann* oder *Erzeugnis*) von s ist definiert als

$$\langle s \rangle = \langle s_1, \dots, s_n \rangle = \langle s_1, \dots, s_n \rangle_K := \sum_{i \in [1,n]} K s_i = K s_1 + \ldots + K s_n = \{ \sum_{i \in [1,n]} a_i s_i \mid a \in K^n \}.$$

Ein Element von $\langle s \rangle$ wird *Linearkombination* über K (oder *K-Linearkombination* oder *Linearkombination*) von s genannt.

(1.19) Beispiel.

(a) In $\mathbb{Q}^3$ ist $(-1, -4, -1)$ eine Linearkombination von $((1, 2, 1), (1, 3, 1))$.

(b) In $\mathbb{R}^4$ ist $(3, 2, -2, 4)$ keine Linearkombination von $((3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1))$.

(c) Jedes $x \in \mathbb{R}^3$ ist eine Linearkombination von $((1, 0, 0), (0, 1, 0), (0, 0, 1))$, d.h. es ist

$$\langle (1, 0, 0), (0, 1, 0), (0, 0, 1) \rangle = \mathbb{R}^3.$$

(d) Jedes $x \in \mathbb{Q}^2$ ist eine Linearkombination von $((1, 1), (1, 2))$, d.h. es ist

$$\langle (1, 1), (1, 2) \rangle = \mathbb{Q}^2.$$

Beweis.

(a) Wir haben

$$(-1, -4, -1) = (1, 2, 1) - 2(1, 3, 1) = 1(1, 2, 1) + (-2)(1, 3, 1).$$

(b) Wäre $(3, 2, -2, 4)$ eine Linearkombination von $((3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1))$, so gäbe es $a, b, c, d \in \mathbb{R}$ mit

$$\begin{aligned}(3, 2, -2, 4) &= a(3, -2, -3, 4) + b(1, 1, -1, 1) + c(-1, -1, 1, 3) + d(-2, 2, 2, 1) \\ &= (3a + b - c - 2d, -2a + b - c + 2d, -3a - b + c + 2d, 4a + b + 3c + d),\end{aligned}$$

es würde also insbesondere

$$1 = 3 - 2 = (3a + b - c - 2d) + (-3a - b + c + 2d) = 0$$

in $\mathbb{R}$ gelten. Im Umkehrschluss ist $(3, 2, -2, 4)$ keine Linearkombination von $((3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1))$.

(c) Für $x \in \mathbb{R}^3$ ist

$$x = (x_1, x_2, x_3) = (x_1, 0, 0) + (0, x_2, 0) + (0, 0, x_3) = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1).$$

(d) Für $x \in \mathbb{Q}^2$ ist

$$\begin{aligned} x = (x_1, x_2) &= ((2x_1 - x_2) + (-x_1 + x_2), (2x_1 - x_2) + 2(-x_1 + x_2)) \\ &= (2x_1 - x_2)(1, 1) + (-x_1 + x_2)(1, 2). \end{aligned}$$

□

Um zu zeigen, dass ein gegebener Vektor v in V Linearkombination eines n-Tupels $(s_1, \ldots, s_n)$ in V für ein $n \in \mathbb{N}_0$ ist, genügt es wie im Beweis von Beispiel (1.19) ein geeignetes $a \in K^n$ mit $v = \sum_{i \in [1,n]} a_i s_i$ anzugeben. Ein solches Koeffiziententupel lässt sich mit Hilfe eines linearen Gleichungssystems berechnen. Die Nichtlösbarkeit eines linearen Gleichungssystems lässt sich hingegen zum systematischen Nachweis, dass ein gegebener Vektor v in V keine Linearkombination eines n-Tupels $(s_1, \ldots, s_n)$ in V für ein $n \in \mathbb{N}_0$ ist, benutzen:

Alternativer Beweis von Beispiel (1.19)(a), (b).

(a) Genau dann ist $(-1, -4, -1)$ eine Linearkombination von $((1, 2, 1), (1, 3, 1))$, wenn es ein $a, b \in \mathbb{Q}$ gibt mit

$$(-1, -4, -1) = a(1, 2, 1) + b(1, 3, 1).$$

Für $a, b \in \mathbb{Q}$ gilt genau dann

$$(-1, -4, -1) = a(1, 2, 1) + b(1, 3, 1) = (a + b, 2a + 3b, a + b),$$

wenn

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \\ -1 \end{pmatrix}$$

gilt. Wir formen die erweiterte Koeffizientenmatrix mittels elementarer Zeilenoperationen um:

$$\left(\begin{array}{cc|c} 1 & 1 & -1 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{array}\right) \xmapsto{\mathrm{add}_{3,1,-1} \circ \mathrm{add}_{2,1,-2}} \left(\begin{array}{cc|c} 1 & 1 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{array}\right) \xmapsto{\mathrm{add}_{1,2,-1}} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{array}\right)$$

Nach Proposition (A.163) gilt also

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

und damit

$$(-1, -4, -1) = 1(1, 2, 1) - 2(1, 3, 1).$$

□

(b) Genau dann ist $(3, 2, -2, 4) \in \langle (3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1) \rangle$, wenn es $a, b, c, d \in \mathbb{R}$ gibt mit

$$(3, 2, -2, 4) = a(3, -2, -3, 4) + b(1, 1, -1, 1) + c(-1, -1, 1, 3) + d(-2, 2, 2, 1).$$

Für $a, b, c, d \in \mathbb{R}$ gilt genau dann

$$\begin{aligned} (3, 2, -2, 4) &= a(3, -2, -3, 4) + b(1, 1, -1, 1) + c(-1, -1, 1, 3) + d(-2, 2, 2, 1) \\ &= (3a + b - c - 2d, -2a + b - c + 2d, -3a - b + c + 2d, 4a + b + 3c + d) \end{aligned}$$

wenn

$$\begin{pmatrix} 3 & 1 & -1 & -2 \\ -2 & 1 & -1 & 2 \\ -3 & -1 & 1 & 2 \\ 4 & 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ -2 \\ 4 \end{pmatrix}$$

gilt. Wir formen die erweiterte Koeffizientenmatrix mittels elementarer Zeilenoperationen um:

$$\left(\begin{array}{cccc|c} 3 & 1 & -1 & -2 & 3 \\ -2 & 1 & -1 & 2 & 2 \\ -3 & -1 & 1 & 2 & -2 \\ 4 & 1 & 3 & 1 & 4 \end{array}\right) \xmapsto{\mathrm{add}_{3,1,1}} \left(\begin{array}{cccc|c} 3 & 1 & -1 & -2 & 3 \\ -2 & 1 & -1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 4 & 1 & 3 & 1 & 4 \end{array}\right)$$

Wegen $0 \neq 1$ gibt es somit nach Proposition (A.163) keine $a, b, c, d \in \mathbb{R}$ mit

$$\begin{pmatrix} 3 & 1 & -1 & -2 \\ -2 & 1 & -1 & 2 \\ -3 & -1 & 1 & 2 \\ 4 & 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ -2 \\ 4 \end{pmatrix},$$

d.h. $(3, 2, -2, 4)$ ist keine Linearkombination von $((3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1))$.

(1.20) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $s = (s_1, \ldots, s_n)$ in V gegeben. Ferner sei

$$\lambda_s \colon K^n \to V,\ a \mapsto \sum_{i \in [1,n]} a_i s_i.$$

Dann ist

$$\langle s \rangle = \operatorname{Im} \lambda_s.$$

Mit ein wenig technischem Aufwand verbunden lassen sich etwas allgemeiner auch Linearkombinationen einer Familie $s = (s_i)_{i \in I}$ über einer beliebigen (möglicherweise unendlichen) Menge I oder Linearkombinationen einer (möglicherweise unendlichen) Menge S definieren, siehe Definition (1.69) und Definition (1.72).

(1.21) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \ldots, s_n)$ in V gegeben. Dann ist $\langle s_1, \ldots, s_n \rangle$ ein K-Untervektorraum von V.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(1.22) Proposition. Es seien $n \in \mathbb{N}_0$, ein n-Tupel $(s_1, \ldots, s_n)$ in V und ein $v \in V$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Der Vektor v ist eine Linearkombination von $(s_1, \ldots, s_n)$.

(b) Es ist $\langle s_1, \ldots, s_n, v \rangle \subseteq \langle s_1, \ldots, s_n \rangle$.

(c) Es ist $\langle s_1, \ldots, s_n, v \rangle = \langle s_1, \ldots, s_n \rangle$.

Beweis. Zunächst gelte Bedingung (a), d.h. es sei v eine Linearkombination von $(s_1, \ldots, s_n)$. Ferner sei ein $w \in \langle s_1, \ldots, s_n, v \rangle$ gegeben. Dann gibt es $a, b \in K^n$, $c \in K$ mit $v = \sum_{i \in [1,n]} a_i s_i$ und $w = \sum_{i \in [1,n]} b_i s_i + cv$. Es folgt

$$w = \sum_{i \in [1,n]} b_i s_i + cv = \sum_{i \in [1,n]} b_i s_i + c \sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,n]} (b_i + c a_i) s_i \in \langle s_1, \ldots, s_n \rangle.$$

Folglich haben wir $\langle s_1, \ldots, s_n, v \rangle \subseteq \langle s_1, \ldots, s_n \rangle$, d.h. es gilt Bedingung (b).
Gilt umgekehrt Bedingung (b), d.h. ist $\langle s_1, \ldots, s_n, v \rangle \subseteq \langle s_1, \ldots, s_n \rangle$, so gilt insbesondere

$$v \in \langle s_1, \ldots, s_n, v \rangle \subseteq \langle s_1, \ldots, s_n \rangle.$$

Dies bedeutet aber, dass v eine Linearkombination von $(s_1, \ldots, s_n)$ ist, d.h. Bedingung (a) gilt.
Wir haben also gezeigt, dass Bedingung (a) und Bedingung (b) äquivalent sind. Da aber ohnehin stets $\langle s_1, \ldots, s_n \rangle \subseteq \langle s_1, \ldots, s_n, v \rangle$ ist, sind auch Bedingung (b) und Bedingung (c) äquivalent.
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Das folgende Korollar gibt eine Antwort auf die Frage, wie wir die Beschreibungen von Erzeugnissen modifizieren können.

(1.23) Korollar. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben.

(a) Für $k, l \in [1, n]$ mit $k < l$ ist

$$\langle s_1, \dots, s_n \rangle = \langle s_1, \dots, s_{k-1}, s_l, s_{k+1}, \dots, s_{l-1}, s_k, s_{l+1}, \dots, s_n \rangle.$$

(b) Für $k, l \in [1, n]$ mit $k \neq l$ und $a \in K$ ist

$$\langle s_1, \dots, s_n \rangle = \langle s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n \rangle.$$

(c) Für $k \in [1, n]$ und $a \in K^\times$ ist

$$\langle s_1, \dots, s_n \rangle = \langle s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n \rangle.$$

Beweis.

(a) Dies folgt aus der Kommutativität der Addition auf V.

(b) Es seien $k, l \in [1, n]$ mit $k \neq l$ und $a \in K$ gegeben. Wegen $s_k + a s_l \in \langle s_1, \dots, s_n \rangle$ gilt

$$\langle s_1, \dots, s_n, s_k + a s_l \rangle = \langle s_1, \dots, s_n \rangle$$

nach Proposition (1.22). Wegen $s_k = (s_k + a s_l) + (-a) s_l \in \langle s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n \rangle$ gilt andererseits aber auch

$$\langle s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n, s_k \rangle = \langle s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n \rangle$$

nach Proposition (1.22). Insgesamt haben wir

$$\begin{aligned}\langle s_1, \dots, s_n \rangle &= \langle s_1, \dots, s_n, s_k + a s_l \rangle = \langle s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n, s_k \rangle \\ &= \langle s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n \rangle.\end{aligned}$$

(c) Es seien $k \in [1, n]$ und $a \in K^\times$ gegeben. Wegen $a s_k \in \langle s_1, \dots, s_n \rangle$ gilt

$$\langle s_1, \dots, s_n, a s_k \rangle = \langle s_1, \dots, s_n \rangle$$

nach Proposition (1.22). Wegen $s_k = a^{-1}(a s_k) \in \langle s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n \rangle$ gilt andererseits aber auch

$$\langle s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n, s_k \rangle = \langle s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n \rangle$$

nach Proposition (1.22). Insgesamt haben wir

$$\begin{aligned}\langle s_1, \dots, s_n \rangle &= \langle s_1, \dots, s_n, a s_k \rangle = \langle s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n, s_k \rangle \\ &= \langle s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n \rangle.\end{aligned}$$ □

Korollar (1.23) liefert uns eine Methode, die Beschreibung eines durch ein Tupel von Vektoren erzeugten Untervektorraums zu vereinfachen. Wir illustrieren dies an einem Beispiel.

(1.24) Beispiel. In $\mathbb{R}^4$ ist

$$\langle (3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1) \rangle = \langle (1, 0, -1, 0), (0, 1, 0, 0), (0, 0, 0, 1) \rangle.$$

Beweis. Nach Korollar (1.23) und Proposition (1.22) ist

$$\begin{aligned}&\langle (3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1) \rangle \\ &= \langle (3, -2, -3, 4) - 3(1, 1, -1, 1), (1, 1, -1, 1), (-1, -1, 1, 3) + (1, 1, -1, 1), (-2, 2, 2, 1) + 2(1, 1, -1, 1) \rangle \\ &= \langle (0, -5, 0, 1), (1, 1, -1, 1), (0, 0, 0, 4), (0, 4, 0, 3) \rangle = \langle (0, -5, 0, 1), (1, 1, -1, 1), \frac{1}{4}(0, 0, 0, 4), (0, 4, 0, 3) \rangle \\ &= \langle (0, -5, 0, 1), (1, 1, -1, 1), (0, 0, 0, 1), (0, 4, 0, 3) \rangle\end{aligned}$$

$$
\begin{aligned}
&= \langle (0,-5,0,1)-(0,0,0,1),(1,1,-1,1)-(0,0,0,1),(0,0,0,1),(0,4,0,3)-3(0,0,0,1)\rangle \\
&= \langle (0,-5,0,0),(1,1,-1,0),(0,0,0,1),(0,4,0,0)\rangle = \langle -\frac{1}{5}(0,-5,0,0),(1,1,-1,0),(0,0,0,1),(0,4,0,0)\rangle \\
&= \langle (0,1,0,0),(1,1,-1,0),(0,0,0,1),(0,4,0,0)\rangle \\
&= \langle (0,1,0,0),(1,1,-1,0)-(0,1,0,0),(0,0,0,1),(0,4,0,0)-4(0,1,0,0)\rangle \\
&= \langle (0,1,0,0),(1,0,-1,0),(0,0,0,1),(0,0,0,0)\rangle = \langle (1,0,-1,0),(0,1,0,0),(0,0,0,1),(0,0,0,0)\rangle \\
&= \langle (1,0,-1,0),(0,1,0,0),(0,0,0,1)\rangle.
\end{aligned}
$$
□

Die in Korollar (1.23) dargestellten Operationen sind Analoga zu den bei der Gauß-Elimination verwendeten elementaren Zeilenoperationen auf Matrizen, siehe (A.161). Etwas übersichtlicher lässt sich der Beweis von Beispiel (1.24) daher wie folgt formulieren:

Alternativer Beweis. Wir schreiben die Quadrupel als Zeilen in eine Matrix und wenden elementare Zeilenoperationen an:

$$
\begin{pmatrix} 3 & -2 & -3 & 4 \\ 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 3 \\ -2 & 2 & 2 & 1 \end{pmatrix}
\xmapsto{\mathrm{add}_{4,2,2} \circ \mathrm{add}_{3,2,1} \circ \mathrm{add}_{1,2,-3}}
\begin{pmatrix} 0 & -5 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 4 & 0 & 3 \end{pmatrix}
\xmapsto{\mathrm{mul}_{3,\frac{1}{4}}}
\begin{pmatrix} 0 & -5 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 4 & 0 & 3 \end{pmatrix}
$$
$$
\xmapsto{\mathrm{add}_{4,3,-3} \circ \mathrm{add}_{2,3,-1} \circ \mathrm{add}_{1,3,-1}}
\begin{pmatrix} 0 & -5 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 4 & 0 & 0 \end{pmatrix}
\xmapsto{\mathrm{mul}_{1,-\frac{1}{5}}}
\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 4 & 0 & 0 \end{pmatrix}
$$
$$
\xmapsto{\mathrm{add}_{4,1,-4} \circ \mathrm{add}_{2,1,-1}}
\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\xmapsto{\mathrm{sw}_{1,2}}
\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
$$

Nach Korollar (1.23) und Proposition (1.22) ergibt sich

$$
\begin{aligned}
\langle (3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1)\rangle &= \langle (1,0,-1,0),(0,1,0,0),(0,0,0,1),(0,0,0,0)\rangle \\
&= \langle (1,0,-1,0),(0,1,0,0),(0,0,0,1)\rangle.
\end{aligned}
$$
□

Mit Hilfe vereinfachter Beschreibungen lässt sich leichter testen, ob ein gegebener Vektor Linearkombination eines gegebenen Tupels von Vektoren ist:

Alternativer Beweis von Beispiel (1.19)(b). Nach Beispiel (1.24) ist

$$\langle (3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1)\rangle = \langle (1,0,-1,0),(0,1,0,0),(0,0,0,1)\rangle.$$

Wäre $(3,2,-2,4)$ eine Linearkombination von $((1,0,-1,0),(0,1,0,0),(0,0,0,1))$, so gäbe es $a,b,c \in \mathbb{R}$ mit

$$(3,2,-2,4) = a(1,0,-1,0) + b(0,1,0,0) + c(0,0,0,1) = (a,b,-a,c)$$

es würde also insbesondere

$$1 = 3 - 2 = a + (-a) = 0$$

in $\mathbb{R}$ gelten. Im Umkehrschluss ist $(3,2,-2,4)$ keine Linearkombination von $((1,0,-1,0),(0,1,0,0),(0,0,0,1))$ und damit keine Linearkombination von $((3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1))$. □

Erzeugendensysteme

Das Konzept des Erzeugnisses, siehe Definition (1.18), legt folgenden Begriff nahe:

(1.25) Definition (Erzeugendensystem)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1,\ldots,s_n)$ in V gegeben. Wir sagen, dass $(s_1,\ldots,s_n)$ ein *Erzeugendensystem* über K (oder *K-Erzeugendensystem* oder *Erzeugendensystem*) von V ist (oder dass V von $(s_1,\ldots,s_n)$ *erzeugt* wird oder dass V von $(s_1,\ldots,s_n)$ *aufgespannt* wird), wenn $V = \langle s_1,\ldots,s_n\rangle$ gilt.

(1.26) Beispiel.

(a) Das Tripel $((1,0,0),(0,1,0),(0,0,1))$ ist ein Erzeugendensystem von $\mathbb{R}^3$.

(b) Das Paar $((1,1),(1,2))$ ist ein Erzeugendensystem von $\mathbb{Q}^2$.

(c) Es sind

$$\begin{aligned}(s_1,s_2,s_3,s_4) &= ((3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1)),\\ (t_1,t_2,t_3) &= ((1,0,-1,0),(0,1,0,0),(0,0,0,1))\end{aligned}$$

Erzeugendensysteme des $\mathbb{R}$-Untervektorraums $\langle s_1,s_2,s_3,s_4\rangle = \langle t_1,t_2,t_3\rangle$ von $\mathbb{R}^4$.

(d) Das Quadrupel

$$(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix})$$

ist ein Erzeugendensystem von $\mathbb{F}_2^{2\times 2}$.

Beweis.

(c) Dies folgt aus Beispiel (1.24). □

(1.27) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $s = (s_1,\ldots,s_n)$ in V gegeben. Genau dann ist $s = (s_1,\ldots,s_n)$ ein Erzeugendensystem von V, wenn

$$\lambda_s \colon K^n \to V,\ a \mapsto \sum_{i\in[1,n]} a_i s_i$$

surjektiv ist.

(1.28) Bemerkung. Es seien $m,n \in \mathbb{N}_0$, ein n-Tupel $(s_1,\ldots,s_n)$ in V und ein Erzeugendensystem $(t_1,\ldots,t_m)$ von V gegeben. Genau dann ist $(s_1,\ldots,s_n)$ ein Erzeugendensystem von V, wenn t_i für jedes $i \in [1,m]$ eine Linearkombination von $(s_1,\ldots,s_n)$ ist.

Beweis. Wenn $(s_1,\ldots,s_n)$ ein Erzeugendensystem von V ist, so gilt $V = \langle s_1,\ldots,s_n\rangle$ und damit insbesondere $t_i \in V = \langle s_1,\ldots,s_n\rangle$ für $i \in [1,m]$.
Umgekehrt gelte für $i \in [1,m]$ stets $t_i \in \langle s_1,\ldots,s_n\rangle$. Da $(t_1,\ldots,t_m)$ ein Erzeugendensystem von V ist, gilt dies insbesondere auch für $(s_1,\ldots,s_n,t_1,\ldots,t_m)$. Nach Proposition (1.22) ist also

$$\langle s_1,\ldots,s_n\rangle = \langle s_1,\ldots,s_n,t_1,\ldots,t_m\rangle = V,$$

d.h. $(s_1,\ldots,s_n)$ ist ein Erzeugendensystem von V. □

(1.29) Beispiel. Das Tripel $((-3,3,-1),(-1,1,0),(3,-2,1))$ ist ein Erzeugendensystem von $\mathbb{R}^3$.

Beweis. Wir haben

$$\begin{aligned}(1,0,0) &= (-3,3,-1)-(-1,1,0)+(3,-2,1),\\ (0,1,0) &= (-3,3,-1)+(3,-2,1),\\ (0,0,1) &= -(-3,3,-1)+3(-1,1,0),\end{aligned}$$

es gilt also $(1,0,0),(0,1,0),(0,0,1) \in \langle(-3,3,-1),(-1,1,0),(3,-2,1)\rangle$. Da $((1,0,0),(0,1,0),(0,0,1))$ jedoch nach Beispiel (1.26)(a) ein Erzeugendensystem von $\mathbb{R}^3$ ist, ist nach Bemerkung (1.28) auch

$$((-3,3,-1),(-1,1,0),(3,-2,1))$$

ein Erzeugendensystem von $\mathbb{R}^3$. □

(1.30) Definition (endlich erzeugter Vektorraum)**.** Der K-Vektorraum V heißt *endlich erzeugt*, falls ein $n \in \mathbb{N}_0$ und ein Erzeugendensystem $(s_1,\ldots,s_n)$ von V existieren.

(1.31) Beispiel. Der $\mathbb{R}$-Vektorraum $\mathbb{R}^3$ ist endlich erzeugt.

Beweis. Dies folgt aus Beispiel (1.26)(a). □

Definition (1.30) legt nahe, dass es auch Vektorräume gibt, welche nicht endlich erzeugt sind. Beispielsweise ist $K[X]$ nicht endlich erzeugt als K-Vektorraum; stattdessen ist $(X^i)_{i \in \mathbb{N}_0}$ ein (unendliches) Erzeugendensystem von $K[X]$. Um unsere Theorie auch auf solche Vektorräume auszuweiten, müssen wir die Begriffe einer Linearkombination und eines Erzeugendensystems etwas weiter fassen und auch für Familien $(s_i)_{i \in I}$ über beliebigen, möglicherweise unendlichen Mengen I zulassen, siehe Definition (1.69). Auch die im Folgenden aufkommenden Begriffe der linearen (Un)abhängigkeit und der einer Basis machen für allgemeinere Familien Sinn.

Lineare Unabhängigkeit

Wir haben bereits gesehen, dass der $\mathbb{R}$-Untervektorraum $U = \{(x, y, 0) \mid x, y \in \mathbb{R}\} = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$ von $((1, 0, 0), (0, 1, 0))$ aufgespannt wird, d.h. es ist

$$U = \langle (1, 0, 0), (0, 1, 0) \rangle.$$

Andererseits ist $(1, 2, 0) = 1(1, 0, 0) + 2(0, 1, 0)$ eine Linearkombination von $((1, 0, 0), (0, 1, 0))$ (d.h. es gilt $(1, 2, 0) \in \langle (1, 0, 0), (0, 1, 0) \rangle = U$), so dass nach Proposition (1.22) auch

$$U = \langle (1, 0, 0), (0, 1, 0) \rangle = \langle (1, 0, 0), (0, 1, 0), (1, 2, 0) \rangle$$

ist.
Im Folgenden wollen wir den Unterschied zwischen $((1, 0, 0), (0, 1, 0))$ und $((1, 0, 0), (0, 1, 0), (1, 2, 0))$ studieren. Dieser liegt in der *Art* begründet, wie die Linearkombinationen dieser Tupel gebildet werden: Einerseits legen die Linearkombinationen von $((1, 0, 0), (0, 1, 0))$ bereits die Koeffizienten in einer Darstellung als Linearkombination fest; sind nämlich $(x, y, 0) \in U$ und $a, b \in \mathbb{R}$ mit $(x, y, 0) = a(1, 0, 0) + b(0, 1, 0)$ gegeben, so folgt

$$(x, y, 0) = a(1, 0, 0) + b(0, 1, 0) = (a, b, 0)$$

und damit $a = x$ und $b = y$, kurz $(a, b) = (x, y)$. Andererseits sind die Koeffizienten der Linearkombinationen von $((1, 0, 0), (0, 1, 0), (1, 2, 0))$ nicht festgelegt, es gilt etwa

$$\begin{aligned}(3, 2, 0) &= 3(1, 0, 0) + 2(0, 1, 0) + 0(1, 2, 0) \\ &= 2(1, 0, 0) + 0(0, 1, 0) + 1(1, 2, 0) \\ &= 1(1, 0, 0) + (-2)(0, 1, 0) + 2(1, 2, 0),\end{aligned}$$

aber die Koeffiziententripel $(3, 2, 0)$, $(2, 0, 1)$ und $(1, -2, 2)$ sind verschieden.
Für diesen Sachverhalt benutzen wir folgende Terminologie.

(1.32) Definition (lineare (Un)abhängigkeit)**.** Es sei $n \in \mathbb{N}_0$ gegeben. Ein n-Tupel $(s_1, \ldots, s_n)$ in V heißt *linear unabhängig* über K (oder K-*linear unabhängig* oder *linear unabhängig*) in V, wenn für $a, b \in K^n$ aus $\sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,n]} b_i s_i$ stets $a = b$ folgt; ansonsten *linear abhängig* über K (oder K-*linear abhängig* oder *linear abhängig*) in V.

(1.33) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $s = (s_1, \ldots, s_n)$ in V gegeben. Genau dann ist $s = (s_1, \ldots, s_n)$ linear unabhängig in V, wenn

$$\lambda_s \colon K^n \to V,\ a \mapsto \sum_{i \in [1,n]} a_i s_i$$

injektiv ist.

Zum Nachweis der linearen Unabhängigkeit genügt es, ein etwas einfacheres Kriterium zu verifizieren:

(1.34) Lemma (Kriterium für lineare Unabhängigkeit)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \ldots, s_n)$ in V gegeben. Genau dann ist $(s_1, \ldots, s_n)$ linear unabhängig, wenn für $a \in K^n$ aus $\sum_{i \in [1,n]} a_i s_i = 0$ stets $a = 0$ folgt.

Beweis. Zunächst sei $(s_1, \ldots, s_n)$ linear unabhängig, d.h. für $a, b \in K^n$ folge aus $\sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,n]} b_i s_i$ stets $a = b$. Für $a \in K^n$ mit $\sum_{i \in [1,n]} a_i s_i = 0$ gilt dann

$$\sum_{i \in [1,n]} a_i s_i = 0 = \sum_{i \in [1,n]} 0 s_i,$$

es folgt also insbesondere $a = (0)_{i \in [1,n]} = 0$.
Umgekehrt folge für $a \in K^n$ aus $\sum_{i \in [1,n]} a_i s_i = 0$ stets $a = 0$. Für $a, b \in K^n$ mit $\sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,n]} b_i s_i$ folgt dann

$$0 = \sum_{i \in [1,n]} a_i s_i - \sum_{i \in [1,n]} b_i s_i = \sum_{i \in [1,n]} (a_i - b_i) s_i,$$

also $a - b = (a_i - b_i)_{i \in [1,n]} = 0$ und damit $a = b$. Folglich ist $(s_1, \ldots, s_n)$ linear unabhängig. □

(1.35) Beispiel.

(a) In $\mathbb{R}^3$ ist $((1,0,0),(0,1,0),(0,0,1))$ linear unabhängig.

(b) In $\mathbb{Q}^2$ ist $((1,1),(1,2))$ linear unabhängig.

(c) In $\mathbb{R}^4$ ist

$$((3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1))$$

linear abhängig und

$$((1,0,-1,0),(0,1,0,0),(0,0,0,1))$$

linear unabhängig.

(d) In $\mathbb{F}_2^{2 \times 2}$ ist

$$(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$

linear unabhängig.

Beweis.

(a) Für $a, b, c \in \mathbb{R}$ mit

$$a(1,0,0) + b(0,1,0) + c(0,0,1) = 0$$

gilt $(a,b,c) = 0 = (0,0,0)$, also $a = 0$, $b = 0$, $c = 0$. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist somit $((1,0,0),(0,1,0),(0,0,1))$ linear unabhängig in $\mathbb{R}^3$.

(b) Für $a, b \in \mathbb{Q}$ mit

$$a(1,1) + b(1,2) = 0$$

gilt $(a+b, a+2b) = 0 = (0,0)$, also $a + b = 0$ und $a + 2b = 0$ und damit auch

$$b = (a+2b) - (a+b) = 0 - 0 = 0,$$
$$a = (a+b) - b = 0 - 0 = 0.$$

Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist somit $((1,1),(1,2))$ linear unabhängig in $\mathbb{Q}^2$.

(c) Es ist

$$16(3,-2,-3,4)+(-27)(1,1,-1,1)+(-19)(-1,-1,1,3)+20(-2,2,2,1)=(0,0,0,0),$$

aber $(16,-27,-19,20)\neq(0,0,0,0)$. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist somit $((3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1))$ linear abhängig in $\mathbb{R}^4$.

Für $a,b,c\in\mathbb{R}$ mit

$$a(1,0,-1,0)+b(0,1,0,0)+c(0,0,0,1)=0$$

gilt $(a,b,-a,c)=0=(0,0,0,0)$, also $a=0$, $b=0$, $c=0$. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist somit $((1,0,-1,0),(0,1,0,0),(0,0,0,1))$ linear unabhängig in $\mathbb{R}^4$.

(d) Für $a,b,c,d\in\mathbb{F}_2$ mit

$$a\begin{pmatrix}1&0\\0&0\end{pmatrix}+b\begin{pmatrix}0&1\\0&0\end{pmatrix}+c\begin{pmatrix}0&0\\1&0\end{pmatrix}+d\begin{pmatrix}0&0\\0&1\end{pmatrix}=0$$

gilt

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}=0=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

also $a=0$, $b=0$, $c=0$, $d=0$. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist somit

$$(\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix})$$

linear unabhängig in $\mathbb{F}_2^{2\times 2}$. □

Um zu zeigen, dass ein gegebenes Tupel linear unabhängig ist, müssen wir wie im Beweis von Beispiel (1.35)(b) ein lineares Gleichungssystem lösen. Auch die Koeffizienten zum Nachweis der linearen Abhängigkeit im Beweis von Beispiel (1.35)(b) lassen sich mit Hilfe eines linearen Gleichungssystems berechnen:

Alternativer Beweis von Beispiel (1.35)(c). Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist das Quadrupel $((3,-2,-3,4),(1,1,-1,1),(-1,-1,1,3),(-2,2,2,1))$ genau dann linear unabhängig, wenn für $a,b,c,d\in\mathbb{R}$ aus

$$a(3,-2,-3,4)+b(1,1,-1,1)+c(-1,-1,1,3)+d(-2,2,2,1)=0$$

bereits $a=b=c=d=0$ folgt. Für $a,b,c,d\in\mathbb{R}$ gilt genau dann

$$\begin{aligned}0&=a(3,-2,-3,4)+b(1,1,-1,1)+c(-1,-1,1,3)+d(-2,2,2,1)\\&=(3a+b-c-2d,-2a+b-c+2d,-3a-b+c+2d,4a+b+3c+d)\end{aligned}$$

wenn

$$\begin{pmatrix}3&1&-1&-2\\-2&1&-1&2\\-3&-1&1&2\\4&1&3&1\end{pmatrix}\begin{pmatrix}a\\b\\c\\d\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

gilt. Wir formen die Koeffizientenmatrix mittels elementarer Zeilenoperationen um:

$$\begin{aligned}\begin{pmatrix}3&1&-1&-2\\-2&1&-1&2\\-3&-1&1&2\\4&1&3&1\end{pmatrix}&\xmapsto{\operatorname{add}_{4,1,-1}\circ\operatorname{add}_{3,1,1}\circ\operatorname{add}_{2,1,-1}}\begin{pmatrix}3&1&-1&-2\\-5&0&0&4\\0&0&0&0\\1&0&4&3\end{pmatrix}\\&\xmapsto{\operatorname{add}_{2,4,5}\circ\operatorname{add}_{1,4,-3}}\begin{pmatrix}0&1&-13&-11\\0&0&20&19\\0&0&0&0\\1&0&4&3\end{pmatrix}\end{aligned}$$

Nach Proposition (A.163) gilt für $a, b, c, d \in \mathbb{R}$ gegeben durch

$$\begin{aligned}
&c = -19, \\
&d = 20, \\
&a = -4c - 3d = -4(-19) - 3 \cdot 20 = 16, \\
&b = 13c + 11d = 13(-19) + 11 \cdot 20 = -27
\end{aligned}$$

also

$$\begin{pmatrix} 3 & 1 & -1 & -2 \\ -2 & 1 & -1 & 2 \\ -3 & -1 & 1 & 2 \\ 4 & 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

und damit

$$a(3, -2, -3, 4) + b(1, 1, -1, 1) + c(-1, -1, 1, 3) + d(-2, 2, 2, 1) = 0,$$

aber $(a, b, c, d) = (16, -27, -19, 20) \neq (0, 0, 0, 0)$. Somit ist $((3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1))$ linear abhängig in $\mathbb{R}^4$. □

Wir haben bereits gesehen, dass das Paar $((1, 0, 0), (0, 1, 0))$ linear unabhängig in $\mathbb{R}^3$ ist, während das Tripel $((1, 0, 0), (0, 1, 0), (1, 2, 0))$ linear abhängig in $\mathbb{R}^3$ ist. Wie wir nun erkennen werden, liegt dies daran, dass $(1, 2, 0) = (1, 0, 0) + 2(0, 1, 0)$ eine Linearkombination von $((1, 0, 0), (0, 1, 0))$ ist.

(1.36) Bemerkung. Es seien $n \in \mathbb{N}_0$, ein n-Tupel $(s_1, \ldots, s_n)$ in V und $a \in K^n$ mit $\sum_{j \in [1,n]} a_j s_j = 0$ gegeben. Für jedes $i \in [1, n]$ mit $a_i \neq 0$ ist s_i eine Linearkombination von $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$.

Beweis. Für jedes $i \in [1, n]$ mit $a_i \neq 0$ ist $a_i \in K^\times$, wegen $\sum_{j \in [1,n]} a_j s_j = 0$ ist dann also

$$s_i = -a_i^{-1} \sum_{j \in [1,n] \setminus \{i\}} a_j s_j = \sum_{j \in [1,n] \setminus \{i\}} (-a_i^{-1} a_j) s_j$$

eine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{i\}}$. □

(1.37) Proposition. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \ldots, s_n)$ in V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Das n-Tupel $(s_1, \ldots, s_n)$ ist linear unabhängig in V.

(b) Für jedes $i \in [1, n]$ ist $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ linear unabhängig in V und s_i ist keine Linearkombination von $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$.

(c) Es gilt entweder $n = 0$ oder es gibt ein $i \in [1, n]$ derart, dass $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ linear unabhängig in V und s_i keine Linearkombination von $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ ist.

(d) Für jedes $i \in [1, n]$ ist s_i keine Linearkombination von $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$.

Beweis. Wir zeigen zuerst die Äquivalenz von Bedingung (a), Bedingung (b) und Bedingung (c), danach die Äquivalenz von Bedingung (a) und Bedingung (d).
Um zu zeigen, dass Bedingung (b) aus Bedingung (a) folgt, benutzen wir einen Umkehrschluss. Es gelte also Bedingung (b) nicht, d.h. es gebe ein $i \in [1, n]$ derart, dass $(s_j)_{j \in [1,n] \setminus \{i\}}$ linear abhängig in V oder s_i ist eine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{i\}}$ ist.
Wenn $(s_j)_{j \in [1,n] \setminus \{i\}}$ linear abhängig in V ist, so gibt es ein $a \in K^{[1,n] \setminus \{i\}} \setminus \{0\}$ mit $\sum_{j \in [1,n] \setminus \{i\}} a_j s_j = 0$. Da dann aber auch

$$\sum_{j \in [1,n] \setminus \{i\}} a_j s_j + 0 s_i = \sum_{j \in [1,n] \setminus \{i\}} a_j s_j = 0$$

gilt, ist in diesem Fall $(s_j)_{j \in [1,n]}$ linear abhängig in V.

Wenn s_i eine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{i\}}$ ist, so gibt es ein $a \in K^{[1,n] \setminus \{i\}}$ mit $s_i = \sum_{j \in [1,n] \setminus \{i\}} a_j s_j$. Es folgt

$$0 = s_i - \sum_{j \in [1,n] \setminus \{i\}} a_j s_j = s_i + \sum_{j \in [1,n] \setminus \{i\}} (-a_j) s_j,$$

wegen $1 \neq 0$ ist also auch in diesem Fall $(s_j)_{j \in [1,n]}$ linear abhängig in V.
Somit ist in beiden Fällen $(s_j)_{j \in [1,n]}$ linear abhängig in V, d.h. Bedingung (a) gilt nicht.
Im Umkehrschluss folgt: Wenn Bedingung (a) gilt, so gilt auch Bedingung (b).
Wenn Bedingung (b) gilt, dann gilt insbesondere Bedingung (c).
Es gelte Bedingung (c). Wenn $n = 0$ ist, dann ist $(s_1, \dots, s_n)$ aus trivialen Gründen linear unabhängig. Daher sei $n \neq 0$, so dass es ein $i \in [1,n]$ derart gibt, dass $(s_j)_{j \in [1,n] \setminus \{i\}}$ linear unabhängig in V und s_i keine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{i\}}$ ist. Um zu zeigen, dass $(s_j)_{j \in [1,n]}$ linear unabhängig in V ist, sei $a \in K^n$ mit $\sum_{j \in [1,n]} a_j s_j = 0$ gegeben. Da s_i keine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{i\}}$ ist, gilt $a_i = 0$ nach Bemerkung (1.36). Wir erhalten

$$\sum_{j \in [1,n] \setminus \{i\}} a_j s_j = \sum_{j \in [1,n] \setminus \{i\}} a_j s_j + 0 s_i = \sum_{j \in [1,n]} a_j s_j = 0,$$

so dass die lineare Unabhängigkeit von $(s_j)_{j \in [1,n] \setminus \{i\}}$ auch $a_j = 0$ für $j \in [1,n] \setminus \{i\}$ liefert. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist $(s_j)_{j \in [1,n]}$ linear unabhängig in V, d.h. es gilt Bedingung (a).
Wir haben die Äquivalenz von Bedingung (a), Bedingung (b) und Bedingung (c) gezeigt.
Wenn also Bedingung (a) gilt, so gilt auch Bedingung (b) und damit insbesondere Bedingung (d).
Schließlich gelte Bedingung (d), d.h. für jedes $i \in [1,n]$ sei s_i keine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{i\}}$. Für jedes $a \in K^n$ mit $\sum_{j \in [1,n]} a_j s_j = 0$ gilt nach Bemerkung (1.36) dann $a_i = 0$ für $i \in [1,n]$. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist daher $(s_j)_{j \in [1,n]}$ linear unabhängig in V, d.h. es gilt Bedingung (a).
Wir haben gezeigt, dass auch Bedingung (a) und Bedingung (d) äquivalent sind.
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c) und Bedingung (d) äquivalent. □

(1.38) Korollar. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben.

(a) Für $k, l \in [1,n]$ mit $k < l$ gilt: Genau dann ist $(s_1, \dots, s_n)$ linear unabhängig, wenn

$$(s_1, \dots, s_{k-1}, s_l, s_{k+1}, \dots, s_{l-1}, s_k, s_{l+1}, \dots, s_n)$$

linear unabhängig ist.

(b) Für $k, l \in [1,n]$ mit $k \neq l$ und $a \in K$ gilt: Genau dann ist $(s_1, \dots, s_n)$ linear unabhängig, wenn

$$(s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n)$$

linear unabhängig ist.

(c) Für $k \in [1,n]$ und $a \in K^\times$ gilt: Genau dann ist $(s_1, \dots, s_n)$ linear unabhängig, wenn

$$(s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n)$$

linear unabhängig ist.

Beweis.

(a) Dies folgt aus der Kommutativität der Addition auf V.

(b) Es seien $k, l \in [1,n]$ mit $k \neq l$ und $a \in K$ gegeben.

Zunächst sei $(s_1, \dots, s_n)$ linear unabhängig. Nach Proposition (1.37) ist dann auch $(s_j)_{j \in [1,n] \setminus \{k\}}$ linear unabhängig und s_k ist keine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{k\}}$, d.h. es gilt $s_k \notin \langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$. Nach Bemerkung (1.21) ist $\langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$ ein K-Untervektorraum von V. Wegen $k \neq l$ und $s_l \in \langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$ ist also $a s_l \in \langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$ und damit $s_k + a s_l \notin \langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$. Nach Proposition (1.37) impliziert dies bereits die lineare Unabhängigkeit von

$$(s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n).$$

Ist umgekehrt $(s_1, \dots, s_{k-1}, s_k + a s_l, s_{k+1}, \dots, s_n)$ linear unabhängig, so auch

$$(s_1, \dots, s_{k-1}, s_k + a s_l + (-a) s_l, s_{k+1}, \dots, s_n) = (s_1, \dots, s_n).$$

(c) Es seien $k \in [1, n]$ und $a \in K^\times$ gegeben.

Zunächst sei $(s_1, \dots, s_n)$ linear unabhängig. Nach Proposition (1.37) ist dann auch $(s_j)_{j \in [1,n] \setminus \{k\}}$ linear unabhängig und s_k ist keine Linearkombination von $(s_j)_{j \in [1,n] \setminus \{k\}}$, d.h. es gilt $s_k \notin \langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$. Nach Bemerkung (1.21) ist $\langle (s_j)_{j \in [1,n] \setminus \{k\}} \rangle$ ein K-Untervektorraum von V. Wegen $a \in K^\times$ ist also $a s_k \notin \langle s_1, \dots, s_{k-1}, s_{k+1}, \dots, s_n \rangle$. Nach Proposition (1.37) impliziert dies bereits die lineare Unabhängigkeit von

$$(s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n).$$

Ist umgekehrt $(s_1, \dots, s_{k-1}, a s_k, s_{k+1}, \dots, s_n)$ linear unabhängig, so auch

$$(s_1, \dots, s_{k-1}, a^{-1}(a s_k), s_{k+1}, \dots, s_n) = (s_1, \dots, s_n). \qquad \square$$

Basen

Als nächstes studieren wir Tupel, welche sowohl Erzeugendensysteme von V als auch linear unabhängig in V sind, also Tupel derart, dass sich jeder Vektor in V eindeutig als Linearkombination in diesem Tupel schreiben lässt.

(1.39) Definition (Basis)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Wir sagen, dass $(s_1, \dots, s_n)$ eine *Basis* über K (oder K-*Basis* oder *Basis*) von V ist, wenn $(s_1, \dots, s_n)$ ein linear unabhängiges Erzeugendensystem von V ist.

(1.40) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $s = (s_1, \dots, s_n)$ in V gegeben. Genau dann ist $s = (s_1, \dots, s_n)$ eine Basis von V, wenn

$$\lambda_s \colon K^n \to V,\ a \mapsto \sum_{i \in [1,n]} a_i s_i$$

bijektiv ist.

(1.41) Beispiel.

(a) Das Tripel $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ ist eine Basis von $\mathbb{R}^3$.

(b) Das Paar $((1, 1), (1, 2))$ ist eine Basis von $\mathbb{Q}^2$.

(c) Das Tripel $((1, 0, -1, 0), (0, 1, 0, 0), (0, 0, 0, 1))$ ist eine Basis des $\mathbb{R}$-Untervektorraums

$$\langle (3, -2, -3, 4), (1, 1, -1, 1), (-1, -1, 1, 3), (-2, 2, 2, 1) \rangle$$

von $\mathbb{R}^4$.

(d) Das Quadrupel

$$(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$

ist eine Basis von $\mathbb{F}_2^{2 \times 2}$.

Beweis.

(a) Dies folgt aus Beispiel (1.26)(a) und Beispiel (1.35)(a).

(b) Dies folgt aus Beispiel (1.26)(b) und Beispiel (1.35)(b).

(c) Dies folgt aus Beispiel (1.26)(c) und Beispiel (1.35)(c).

(d) Dies folgt aus Beispiel (1.26)(d) und Beispiel (1.35)(d). $\square$

(1.42) Lemma. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $(s_1, \dots, s_n)$ eine Basis von V.

(b) Es ist $(s_1, \dots, s_n)$ ein bzgl. Streichen/Ergänzen minimales Erzeugendensystem von V, d.h. $(s_1, \dots, s_n)$ ist ein Erzeugendensystem von V und für alle $i \in [1, n]$ ist $(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ kein Erzeugendensystem von V.

(c) Es ist $(s_1, \dots, s_n)$ ein bzgl. Streichen/Ergänzen maximal linear unabhängiges Tupel in V, d.h. $(s_1, \dots, s_n)$ ist linear unabhängig in V und für alle $v \in V$ ist $(s_1, \dots, s_n, v)$ linear abhängig in V.

Beweis. Wir zeigen zuerst die Äquivalenz von Bedingung (a) und Bedingung (b), danach die Äquivalenz von Bedingung (a) und Bedingung (c).
Genau dann ist $(s_1, \dots, s_n)$ eine Basis von V, wenn $(s_1, \dots, s_n)$ linear unabhängig in V und ein Erzeugendensystem von V ist. Um die Äquivalenz von Bedingung (a) und Bedingung (b) zu zeigen, nehmen wir an, dass $(s_1, \dots, s_n)$ ein Erzeugendensystem von V ist. Nach Proposition (1.37) ist $(s_1, \dots, s_n)$ genau dann linear unabhängig in V, wenn s_i für $i \in [1, n]$ stets keine Linearkombination von $(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ ist. Dies ist nach Proposition (1.22) aber wiederum dazu äquivalent, dass für $i \in [1, n]$ stets $\langle s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n \rangle \neq \langle s_1, \dots, s_n \rangle = V$ ist, also dazu, dass $(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ kein Erzeugendensystem von V ist.
Um die Äquivalenz von Bedingung (a) und Bedingung (c) zu zeigen, nehmen wir an, dass $(s_1, \dots, s_n)$ linear unabhängig ist. Genau dann ist $(s_1, \dots, s_n)$ ein Erzeugendensystem von V, wenn jedes $v \in V$ eine Linearkombination von $(s_1, \dots, s_n)$ ist. Nach Proposition (1.37) ist dies auf Grund der linearen Unabhängigkeit von $(s_1, \dots, s_n)$ aber äquivalent dazu, dass für $v \in V$ das Tupel $(s_1, \dots, s_n, v)$ stets linear abhängig ist. Insgesamt sind Bedingung (a) und Bedingung (c) äquivalent.
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Standardbasis

Für gewisse Vektorräume gibt es kanonische Basen, welchen wir nun eine eigene Bezeichnung zuweisen wollen.

(1.43) Definition (Standardbasis)**.**

(a) Es sei $n \in \mathbb{N}_0$ gegeben. Die Basis $\mathrm{e} = \mathrm{e}^{K^n} = (\mathrm{e}_1, \dots, \mathrm{e}_n) = (\mathrm{e}_1^{K^n}, \dots, \mathrm{e}_n^{K^n})$ von K^n gegeben durch

$$\mathrm{e}_j = \mathrm{e}_j^{K^n} = (\delta_{i,j})_{i \in [1,n]}$$

für $j \in [1, n]$ heißt *Standardbasis* von K^n.

(b) Es seien $m, n \in \mathbb{N}_0$ gegeben. Die Basis $(\mathrm{e}_{1,1}, \mathrm{e}_{1,2}, \dots, \mathrm{e}_{m,n}) = (\mathrm{e}_{1,1}^{K^{m \times n}}, \mathrm{e}_{1,2}^{K^{m \times n}}, \dots, \mathrm{e}_{m,n}^{K^{m \times n}})$ von $K^{m \times n}$ gegeben durch

$$\mathrm{e}_{k,l} = \mathrm{e}_{k,l}^{K^{m \times n}} = (\delta_{(i,j),(k,l)})_{i \in [1,m], j \in [1,n]}$$

für $k \in [1, m]$, $l \in [1, n]$ heißt (*geordnete*) *Standardbasis* von $K^{m \times n}$.

In Beispiel (1.41)(a), (d) haben wir Standardbasen von $\mathbb{R}^3$ bzw. $\mathbb{F}_2^{2 \times 2}$ betrachtet:

(1.44) Beispiel.

(a) In $\mathbb{R}^3$ ist $\mathrm{e} = (\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3)$ gegeben durch

$$\mathrm{e}_1 = (1, 0, 0),\ \mathrm{e}_2 = (0, 1, 0),\ \mathrm{e}_3 = (0, 0, 1).$$

(b) In $\mathbb{F}_2^{2 \times 2}$ ist $(\mathrm{e}_{1,1}, \mathrm{e}_{1,2}, \mathrm{e}_{2,1}, \mathrm{e}_{2,2})$ gegeben durch

$$\mathrm{e}_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ \mathrm{e}_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ \mathrm{e}_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ \mathrm{e}_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

(1.45) Notation. Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Für $i \in [1, n]$ schreiben wir $\mathrm{e}_i = \mathrm{e}_i^{K^{n \times 1}} := \mathrm{e}_{i,1}^{K^{n \times 1}}$. Ferner schreiben wir $\mathrm{e} = \mathrm{e}^{K^{n \times 1}} = (\mathrm{e}_1, \dots, \mathrm{e}_n)$.

(b) Für $i \in [1, n]$ schreiben wir $\mathrm{e}_i = \mathrm{e}_i^{K^{1 \times n}} := \mathrm{e}_{1,i}^{K^{1 \times n}}$. Ferner schreiben wir $\mathrm{e} = \mathrm{e}^{K^{1 \times n}} = (\mathrm{e}_1, \dots, \mathrm{e}_n)$.

(1.46) Beispiel.

(a) In $\mathbb{Q}^{3\times 1}$ ist $\mathrm{e} = (\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3)$ gegeben durch

$$\mathrm{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathrm{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathrm{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(b) In $\mathbb{F}_3^{1\times 2}$ ist $\mathrm{e} = (\mathrm{e}_1, \mathrm{e}_2)$ gegeben durch

$$\mathrm{e}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \mathrm{e}_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

Existenz von Basen

Im Folgenden wollen wir Eigenschaften von Basen von Vektorräumen studieren. Insbesondere wollen wir zeigen, dass endlich erzeugte Vektorräume stets eine Basis haben. Diese Aussage gilt sogar für beliebige Vektorräume, der Beweis erfordert dann aber einen höheren Aufwand und ist wenig konstruktiv.

(1.47) Satz (Basisauswahlergänzungssatz)**.** Es seien $m, p \in \mathbb{N}_0$, ein linear unabhängiges m-Tupel $(s_1, \ldots, s_m)$ in V und ein Erzeugendensystem $(t_1, \ldots, t_p)$ von V gegeben. Ferner seien $n \in \mathbb{N}_0$ und $i_1, \ldots, i_n \in [1, p]$ so gegeben, dass $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n})$ linear unabhängig in V ist, und so, dass $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n}, t_k)$ für alle $k \in [1, p]$ linear abhängig in V ist. Dann ist $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n})$ eine Basis von V.

Beweis. Für alle $k \in [1, p]$ impliziert die lineare Abhängigkeit von $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n}, t_k)$ auf Grund der linearen Unabhängigkeit von $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n})$ nach Proposition (1.37), dass t_k eine Linearkombination von $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n})$ ist. Nach Bemerkung (1.28) ist daher $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n})$ ein Erzeugendensystem von V. Auf Grund der linearen Unabhängigkeit ist $(s_1, \ldots, s_m, t_{i_1}, \ldots, t_{i_n})$ dann aber sogar eine Basis von V. □

Der Basisauswahlergänzungssatz (1.47) besagt also, dass, sofern V endlich erzeugt ist, jedes linear unabhängige Tupel in V zu einer Basis von V ergänzt werden kann, wobei die ergänzten Vektoren aus einem Erzeugendensystem von V ausgewählt werden können. Algorithmisch lässt sich dieser Satz wie folgt formulieren:

(1.48) Algorithmus.

- Eingabe: linear unabhängiges Tupel $s = (s_1, \ldots, s_m)$ in V, Erzeugendensystem $t = (t_1, \ldots, t_p)$ von V für $m, p \in \mathbb{N}_0$
- Ausgabe: Basis von V
- Verfahren:

```
function basisredext(s, t)
     for k ∈ [1, p] do
          if durch Anhängen von t_k an s entsteht ein linear unabhängiges Tupel then
               hänge t_k an s an;
          end if;
     end for;
     return s;
end function;
```

Der Basisauswahlergänzungssatz (1.47) lässt sich zum Basisauswahlsatz spezialisieren:

(1.49) Korollar (Basisauswahlsatz)**.** Es seien $p \in \mathbb{N}_0$ und ein Erzeugendensystem $(s_1, \ldots, s_p)$ von V gegeben. Ferner seien $n \in \mathbb{N}_0$ und $i_1, \ldots, i_n \in [1, p]$ so gegeben, dass $(s_{i_1}, \ldots, s_{i_n})$ linear unabhängig in V ist, und so, dass $(s_{i_1}, \ldots, s_{i_n}, s_k)$ für alle $k \in [1, p]$ linear abhängig in V ist. Dann ist $(s_{i_1}, \ldots, s_{i_n})$ eine Basis von V.

Beweis. Wegen der linearen Unabhängigkeit von $(\,)$ in V ist $(s_{i_1}, \ldots, s_{i_n})$ eine Basis von V nach Satz (1.47). □

Der Basisauswahlsatz (1.49) besagt also, dass sich aus jedem Erzeugendensystem eines endlich erzeugten Vektorraums eine Basis auswählen lässt. Wir erhalten folgende algorithmische Fassung.

(1.50) Algorithmus.

- Eingabe: Erzeugendensystem $s = (s_1, \dots, s_p)$ von V für $p \in \mathbb{N}_0$
- Ausgabe: Basis von V
- Verfahren:

```
function basisred(s)
     return basisredext((), s);
end function;
```

(1.51) Beispiel. Es sei ein Erzeugendensystem (s_1, s_2, s_3) von $\mathbb{R}^2$ gegeben durch

$$s_1 = (1, -2),\ s_2 = (-2, 4),\ s_3 = (2, 3).$$

Dann sind (s_1, s_3) und (s_2, s_3) Basen von $\mathbb{R}^2$.

(1.52) Korollar. Es sei V endlich erzeugt. Dann gibt es ein $n \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_n)$ von V.

Beweis. Da V endlich erzeugt ist, gibt es ein Erzeugendensystem $(s_1, \dots, s_p)$ von V. Wegen der Endlichkeit von $[1, p]$ gibt es ferner $n \in \mathbb{N}_0$ und $i_1, \dots, i_n \in [1, p]$ so, dass $(s_{i_1}, \dots, s_{i_n})$ linear unabhängig in V ist, und so, dass $(s_{i_1}, \dots, s_{i_n}, s_k)$ für alle $k \in [1, p]$ linear abhängig in V ist. Nach dem Basisauswahlsatz (1.49) ist $(s_{i_1}, \dots, s_{i_n})$ eine Basis von V. □

Der Basisauswahlergänzungssatz (1.47) lässt sich außerdem zum Basisergänzungssatz spezialisieren:

(1.53) Korollar (Basisergänzungssatz)**.** Es sei V endlich erzeugt und es seien $m \in \mathbb{N}_0$ und ein linear unabhängiges m-Tupel $(s_1, \dots, s_m)$ in V gegeben. Dann gibt es ein $n \in \mathbb{N}_0$ und ein n-Tupel $(t_1, \dots, t_n)$ in V so, dass $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist.

Beweis. Da V endlich erzeugt ist, gibt es ein Erzeugendensystem $(t_1, \dots, t_p)$ von V. Wegen der Endlichkeit von $[1, p]$ gibt es ferner $n \in \mathbb{N}_0$ und $i_1, \dots, i_n \in [1, p]$ so, dass $(s_1, \dots, s_m, t_{i_1}, \dots, t_{i_n})$ linear unabhängig in V ist, und so, dass $(s_1, \dots, s_m, t_{i_1}, \dots, t_{i_n}, t_k)$ für alle $k \in [1, p]$ linear abhängig in V ist. Nach Satz (1.47) ist $(s_1, \dots, s_m, t_{i_1}, \dots, t_{i_n})$ eine Basis von V. □

Der Basisergänzungssatz (1.53) besagt also, dass jede linear unabhängige Teilmenge eines endlich erzeugten Vektorraums zu einer Basis dieses Vektorraums ergänzt werden kann.

Alternativer Beweis von Korollar (1.52). Da das 0-Tupel $()$ linear unabhängig in V ist, gibt es nach dem Basisergänzungssatz (1.53) ein $n \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_n)$ von V. □

(1.54) Beispiel. Es sei ein linear unabhängiges Paar (s_1, s_2) in $\mathbb{R}^3$ gegeben durch

$$s_1 = (1, -2, 0),\ s_2 = (2, 3, 0).$$

Dann ist (s_1, s_2, s_3) mit

$$s_3 = (0, 0, 1)$$

eine Basis von $\mathbb{R}^3$.

(1.55) Korollar. Es sei V endlich erzeugt und es seien ein K-Untervektorraum U von V, ein $m \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_m)$ von U gegeben. Dann gibt es ein $n \in \mathbb{N}_0$ und ein n-Tupel $(t_1, \dots, t_n)$ in V so, dass $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist.

Beweis. Als Basis ist $(s_1, \dots, s_m)$ linear unabhängig in U und damit auch in V. Nach dem Basisergänzungssatz (1.53) gibt es ein $n \in \mathbb{N}_0$ und ein n-Tupel $(t_1, \dots, t_n)$ in V so, dass $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist. □

Eindeutigkeit von Basen

In Korollar (1.52) haben wir gesehen, dass endlich erzeugte Vektorräume stets eine Basis haben. Ferner haben wir in Beispiel (1.51) gesehen, dass Basen im Allgemeinen nicht eindeutig sind, ein (endlich erzeugter) Vektorraum hat im Allgemeinen viele verschiedene Basen. Im Folgenden werden wir sehen, dass zumindest die Anzahl der Einträge einer Basis eines endlich erzeugten Vektorraums eindeutig ist. Das Schlüsselresultat zur Beweis dieser Aussage ist der Steinitzsche Austauschsatz (1.58), welcher auf folgendem Lemma beruht.

(1.56) Lemma (Austauschlemma von Steinitz)**.** Es seien $n \in \mathbb{N}_0$, ein Erzeugendensystem $(s_1, \dots, s_n)$ von V, ein $v \in V$ und ein $a \in K^n$ mit $v = \sum_{j \in [1,n]} a_j s_j$ gegeben. Für jedes $i \in [1, n]$ mit $a_i \neq 0$ ist

$$(s_1, \dots, s_{i-1}, v, s_{i+1}, \dots, s_n)$$

ein Erzeugendensystem von V.

Beweis. Es sei $i \in [1, n]$ mit $a_i \neq 0$ gegeben, so dass $a_i \in K^\times$ gilt. Nach Korollar (1.23)(c), (b) folgt

$$\begin{aligned} V &= \langle s_1, \dots, s_n \rangle = \langle s_1, \dots, s_{i-1}, a_i s_i, s_{i+1}, \dots, s_n \rangle = \langle s_1, \dots, s_{i-1}, a_i s_i + \sum_{j \in [1,n] \setminus \{i\}} a_j s_j, s_{i+1}, \dots, s_n \rangle \\ &= \langle s_1, \dots, s_{i-1}, \sum_{j \in [1,n]} a_j s_j, s_{i+1}, \dots, s_n \rangle = \langle s_1, \dots, s_{i-1}, v, s_{i+1}, \dots, s_n \rangle, \end{aligned}$$

d.h. $(s_1, \dots, s_{i-1}, v, s_{i+1}, \dots, s_n)$ ist ein Erzeugendensystem von V. □

(1.57) Beispiel. Es sei ein Erzeugendensystem (s_1, s_2, s_3) von $\mathbb{R}^3$ gegeben durch

$$s_1 = (1, 1, 0),\ s_2 = (1, -1, 0),\ s_3 = (0, 0, 1).$$

Ferner sei $x \in \mathbb{R}^3$ gegeben durch

$$x = (-1, 5, 0).$$

Dann sind (x, s_2, s_3) und (s_1, x, s_3) Erzeugendensysteme von $\mathbb{R}^3$.

Beweis. Es ist

$$x = (-1, 5, 0) = 2(1, 1, 0) + (-3)(1, -1, 0) = 2s_1 + (-3)s_2 + 0s_3.$$

Nach dem Austauschlemma von Steinitz (1.56) ist (x, s_2, s_3) ein Erzeugendensystem von $\mathbb{R}^3$ wegen $2 \neq 0$ in $\mathbb{R}$, und es ist (s_1, x, s_3) ein Erzeugendensystem von $\mathbb{R}^3$ wegen $-3 \neq 0$ in $\mathbb{R}$. □

(1.58) Satz (Steinitzscher Austauschsatz)**.** Es seien $m, n \in \mathbb{N}_0$, ein linear unabhängiges m-Tupel $(s_1, \dots, s_m)$ in V und ein Erzeugendensystem $(t_1, \dots, t_n)$ von V gegeben. Dann gibt es $i_1, \dots, i_m \in [1, n]$ mit $i_1 < \ldots < i_m$ so, dass

$$(t_1, \dots, t_{i_1 - 1}, s_1, t_{i_1 + 1}, \dots, t_{i_m - 1}, s_m, t_{i_m + 1}, \dots, t_n)$$

ein Erzeugendensystem von V ist.

Beweis. Wir führen Induktion nach m.
Ist $m = 0$, so ist $(t_j)_{j \in [1,n]}$ ein Erzeugendensystem von V.
Es sei also $m \geq 1$ und es sei angenommen, dass es $i_1, \dots, i_{m-1} \in [1, n]$ so gibt, dass

$$(s_j)_{j \in [1,m-1]} \cup (t_j)_{j \in [1,n] \setminus \{i_1, \dots, i_{m-1}\}}$$

ein Erzeugendensystem von V ist. Dann gibt es ein $a \in K^{[1,m-1]}$ und ein $b \in K^{[1,n] \setminus \{i_1, \dots, i_{m-1}\}}$ so, dass

$$s_m = \sum_{j \in [1,m-1]} a_j s_j + \sum_{j \in [1,n] \setminus \{i_1, \dots, i_{m-1}\}} b_j t_j.$$

Wegen der linearen Unabhängigkeit von $(s_j)_{j \in [1,m]}$ ist s_m nach Proposition (1.37) keine Linearkombination von $(s_j)_{j \in [1,m-1]}$. Folglich gibt es ein $i_m \in [1, n] \setminus \{i_1, \dots, i_{m-1}\}$ mit $b_{i_m} \neq 0$. Nach dem Austauschlemma (1.56) ist dann aber auch $(s_j)_{j \in [1,m]} \cup (t_j)_{j \in [1,n] \setminus \{i_1, \dots, i_m\}}$ ein Erzeugendensystem von V. □

(1.59) Korollar. Es seien $m, n \in \mathbb{N}_0$, ein linear unabhängiges m-Tupel $(s_1, \dots, s_m)$ in V und ein Erzeugendensystem $(t_1, \dots, t_n)$ von V gegeben. Dann ist

$$m \leq n.$$

(1.60) Korollar. Es seien $m, n \in \mathbb{N}_0$ und Basen $(s_1, \dots, s_m)$ und $(t_1, \dots, t_n)$ von V gegeben. Dann ist

$$m = n.$$

(1.61) Korollar. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $(s_1, \dots, s_n)$ eine Basis von V.

(b) Es ist $(s_1, \dots, s_n)$ ein bzgl. Länge minimales Erzeugendensystem von V, d.h. es ist $(s_1, \dots, s_n)$ ein Erzeugendensystem von V und für alle $p \in \mathbb{N}_0$ und jedes Erzeugendensystem $(t_1, \dots, t_p)$ von V gilt $n \leq p$.

(c) Es ist $(s_1, \dots, s_n)$ ein bzgl. Länge maximales linear unabhängiges Tupel in V, d.h. es ist $(s_1, \dots, s_n)$ linear unabhängig in V und für alle $m \in \mathbb{N}_0$ und jedes linear unabhängige Tupel $(t_1, \dots, t_m)$ in V gilt $m \leq n$.

Beweis. Wenn $(s_1, \dots, s_n)$ eine Basis von V ist, so ist $(s_1, \dots, s_n)$ ein linear unabhängiges Erzeugendensystem und damit ein bzgl. Länge minimales Erzeugendensystem nach Korollar (1.59). Ist umgekehrt $(s_1, \dots, s_n)$ ein bzgl. Länge minimales Erzeugendensystem von V, so ist $(s_1, \dots, s_n)$ insbesondere minimal bzgl. Streichen/Ergänzen und damit eine Basis nach Lemma (1.42). Dies zeigt die Äquivalenz von Bedingung (a) und Bedingung (b).
Wenn $(s_1, \dots, s_n)$ eine Basis von V ist, so ist $(s_1, \dots, s_n)$ ein linear unabhängiges Erzeugendensystem und damit ein bzgl. Länge maximales linear unabhängiges Tupel nach Korollar (1.59). Ist umgekehrt $(s_1, \dots, s_n)$ ein bzgl. Länge maximales linear unabhängiges Tupel in V, so ist $(s_1, \dots, s_n)$ insbesondere maximal bzgl. Streichen/Ergänzen und damit eine Basis nach Lemma (1.42). Dies zeigt die Äquivalenz von Bedingung (a) und Bedingung (c).
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Dimension

Nach Korollar (1.52) hat jeder endlich erzeugte Vektorraum eine (endliche) Basis, und nach Korollar (1.60) sind die Längen von verschiedenen Basen gleich. Man sagt, dass diese Kardinalität eine *Invariante* des Vektorraums ist – sie ist unabhängig von der betrachteten Basis und hängt nur vom Vektorraum ab. Wir geben dieser Invarianten im Folgenden einen Namen.

(1.62) Definition ((un)endlichdimensional)**.** Wir sagen, dass V *endlichdimensional* ist, wenn es ein $n \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_n)$ von V gibt, und ansonsten, dass V *unendlichdimensional* ist. ([1])

(1.63) Bemerkung. Genau dann ist V endlichdimensional, wenn V endlich erzeugt ist.

Beweis. Zunächst sei V endlichdimensional, d.h. es gebe ein $n \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_n)$ von V. Dann ist $(s_1, \dots, s_n)$ insbesondere ein Erzeugendensystem von V und damit V endlich erzeugt.
Umgekehrt, wenn V endlich erzeugt ist, so gibt es nach Korollar (1.52) ein $n \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_n)$ von V, d.h. V ist endlichdimensional. □

(1.64) Definition (Dimension)**.** Es sei V endlichdimensional. Ferner sei $n \in \mathbb{N}_0$ derart gegeben, dass es eine Basis $(s_1, \dots, s_n)$ von V gibt. Wir nennen

$$\dim V = \dim_K V := n$$

die *Dimension* von V über K (oder die *K-Dimension* von V oder die *Dimension* von V). Wir sagen auch, dass V ein *n-dimensionaler* K-Vektorraum ist.

[1] Wenn V nicht endlichdimensional ist, so lässt sich zeigen, dass es eine unendliche Menge I und eine Basis $s = (s_i)_{i \in I}$ von V im Sinne von Definition (1.69)(d) gibt.

(1.65) Beispiel.

(a) Für $n \in \mathbb{N}_0$ ist

$$\dim_K K^n = n.$$

(b) Für $m, n \in \mathbb{N}_0$ ist

$$\dim_K K^{m \times n} = mn.$$

Beweis.

(a) Es ist $(\mathrm{e}_1, \dots, \mathrm{e}_n)$ eine Basis von K^n.

(b) Es ist $(\mathrm{e}_{1,1}, \mathrm{e}_{1,2}, \dots, \mathrm{e}_{m,n})$ eine Basis von $K^{m \times n}$. □

(1.66) Bemerkung. Es sei V endlichdimensional.

(a) Für jedes $n \in \mathbb{N}_0$ und jedes Erzeugendensystem $(s_1, \dots, s_n)$ von V gilt $n \geq \dim_K V$.

(b) Für jedes $n \in \mathbb{N}_0$ und jedes linear unabhängige n-Tupel $(s_1, \dots, s_n)$ in V gilt $n \leq \dim_K V$.

Beweis. Dies folgt aus Korollar (1.59). □

(1.67) Korollar. Es sei V endlichdimensional und es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $(s_1, \dots, s_n)$ eine Basis von V.

(b) Es ist $(s_1, \dots, s_n)$ ein Erzeugendensystem von V und $n = \dim_K V$.

(c) Es ist $(s_1, \dots, s_n)$ linear unabhängig in V und $n = \dim_K V$.

Beweis. Dies folgt aus Bemerkung (1.66) und Korollar (1.61). □

(1.68) Proposition. Es sei V endlichdimensional und es sei ein K-Untervektorraum U von V gegeben. Dann ist U endlichdimensional und es gilt

$$\dim_K U \leq \dim_K V.$$

Ferner gilt genau dann $U = V$, wenn

$$\dim_K U = \dim_K V$$

ist.

Beweis. Nach Bemerkung (1.66)(b) sind die Längen linear unabhängiger Tupel in V durch $\dim V$ beschränkt. Es sei m die größte nichtnegative ganze Zahl derart, dass es ein linear unabhängiges Tupel $(s_1, \dots, s_m)$ in V gibt mit der Eigenschaft, dass alle Einträge in U liegen. Dann ist $(s_1, \dots, s_m)$ nach Korollar (1.61) eine Basis von U, d.h. U ist endlichdimensional und es gilt $\dim U = m \leq \dim V$ nach Bemerkung (1.66)(b).
Wenn $U = V$ ist, so gilt insbesondere auch $\dim U = \dim V$. Es gelte also umgekehrt $\dim U = \dim V$. Dann ist $m = \dim U = \dim V$ und damit $(s_1, \dots, s_m)$ eine Basis von V nach Korollar (1.67), also insbesondere ein Erzeugendensystem von V und damit $U = \langle s_1, \dots, s_m \rangle = V$. □

Linearkombinationen beliebiger Familien und von Mengen

Zum Abschluss dieses Abschnitts werden wir unsere Begriffe zur Beschreibung von Untervektorräumen wie der einer Basis noch verallgemeinern. Zunächst betrachten wir allgemeinere Familien und führen eine alternative Notation ein. Hierdurch erhalten wir insbesondere auch das Konzept einer unendlichen Basis im Sinne einer Familie über einer unendlichen Menge. Es lässt sich zeigen, dass ein Vektorraum, welcher eine solche Basis besitzt, nicht gleichzeitig eine Basis im Sinne von Definition (1.39) besitzt und damit nicht endlichdimensional ist.

Es seien eine Menge I und Familie $s = (s_i)_{i \in I}$ in V gegeben. Im Fall $I = [1, n]$ für ein $n \in \mathbb{N}_0$ haben wir einen Vektor der Form $\sum_{i \in I} a_i s_i = \sum_{i \in [1,n]} a_i s_i$ für ein $a \in K^I = K^n$ eine Linearkombination von s genannt. Wenn nun $I \neq [1, n]$ für $n \in \mathbb{N}_0$ ist, so möchten wir gerne, dass eine Linearkombination weiterhin ein Vektor der Form $\sum_{i \in I} a_i s_i$ für ein $a \in K^I$ ist. Hierbei stellt sich nun folgendes Problem: Wenn I unendlich ist, so ist nicht klar, wie eine „unendliche Summe" definiert sein soll. Hierfür gibt es verschiedene Ansätze. Während man in der Analysis (*absolut*) *konvergente Reihen* einführt, werden wir uns eines kleinen Tricks bedienen und nur solche Koeffizientenfamilien betrachten, bei denen de facto eine endliche Summe übrig bleibt, vgl. Notation (A.113). Auf Grund der Assoziativität und der Kommutativität der Addition ist dann die Verwendung des Summenzeichens auch in diesem allgemeineren Fall gerechtfertigt.

(1.69) Definition (lineare Hülle, Erzeugendensystem, linear (un)abhängig, Basis)**.** Es seien eine Menge I und eine Familie $s = (s_i)_{i \in I}$ in V gegeben und es sei

$$\lambda_s \colon K^{(I)} \to V,\ a \mapsto \sum_{i \in I} a_i s_i.$$

(a) Die *lineare Hülle* über K (oder K*-lineare Hülle* oder *lineare Hülle* oder *Spann* oder *Erzeugnis*) von s ist definiert als

$$\langle s \rangle = \langle s_i \mid i \in I \rangle = \langle s_i \mid i \in I \rangle_K := \operatorname{Im} \lambda_s.$$

Ein Element von $\langle s \rangle$ wird *Linearkombination* über K (oder K*-Linearkombination* oder *Linearkombination*) von s genannt.

(b) Wir sagen, dass s ein *Erzeugendensystem* über K (oder K*-Erzeugendensystem* oder *Erzeugendensystem*) von V ist (oder dass V von s *erzeugt* wird oder dass V von s *aufgespannt* wird), wenn λ_s surjektiv ist.

(c) Die Familie s heißt *linear unabhängig* über K (oder K*-linear unabhängig* oder *linear unabhängig*) in V, wenn λ_s injektiv ist; ansonsten *linear abhängig* über K (oder K*-linear abhängig* oder *linear abhängig*) in V.

(d) Wir sagen, dass s eine *Basis* über K (oder K*-Basis* oder *Basis*) von V ist, wenn λ_s bijektiv ist.

(1.70) Beispiel.

(a) Die (durch $\mathbb{N}_0$ indizierte) Folge $(X^i)_{i \in \mathbb{N}_0}$ ist eine Basis von $K[X]$.

(b) Für $m, n \in \mathbb{N}_0$ ist $(\mathrm{e}_{i,j})_{i \in [1,m], j \in [1,n]}$ eine Basis von $K^{m \times n}$.

Für den $\mathbb{R}$-Vektorraum $\mathbb{R}^{[0,7] \times [0,7]}$ aus Anwendungsbeispiel (1.6) gibt es bspw. folgende Basen:

(1.71) Beispiel.

(a) Es ist $(\mathrm{e}_{k,l})_{k,l \in [0,7]}$ gegeben durch

$$\mathrm{e}_{k,l} = (\delta_{(i,j),(k,l)})_{i,j \in [0,7]}$$

für $k, l \in [0, 7]$ eine Basis von $\mathbb{R}^{[0,7] \times [0,7]}$.

(b) Es ist $(s_{k,l})_{k,l \in [0,7]}$ gegeben durch

$$s_{k,l} = (\cos(\frac{2i+1}{16} k\pi) \cos(\frac{2j+1}{16} l\pi))_{i,j \in [0,7]}$$

für $k, l \in [0, 7]$ eine Basis von $\mathbb{R}^{[0,7] \times [0,7]}$.

Ohne Beweis. □

Nun werden wir für die betrachteten Konzepte noch alternative Notationen mit Hilfe von Teilmengen an Stelle von Familien kennenlernen. Kurz gesagt sind die in Definition (1.72) eingeführten Begriffe für eine Teilmenge S von V definiert als die entsprechenden Begriffe für die Familie $(s)_{s \in S}$ in V über S.

Der Vorteil der Mengennotation liegt in erster Linie auf sprachlicher Ebene: Auf Grund der vorhandenen Konzepte für Mengen wie die einer Teilmenge oder der Vereinigung von Mengen lassen sich manche Sätze wie etwa der Basisergänzungssatz mit Hilfe von Mengen einfacher formulieren als mit Tupeln und allgemeineren Familien. Nichtsdestotrotz sind Familien aus technischen Gründen an vielen Stellen einfacher handhabbar, weswegen wir ab dem nächsten Abschnitt wieder ausschließlich mit Familien (und in der Regel sogar lediglich mit n-Tupeln für ein $n \in \mathbb{N}_0$) arbeiten werden.

(1.72) Definition (lineare Hülle, Erzeugendensystem, linear (un)abhängig, Basis)**.** Es sei eine Teilmenge S von V gegeben.

(a) Die *lineare Hülle* über K (oder K*-lineare Hülle* oder *lineare Hülle* oder *Spann* oder *Erzeugnis*) von S ist definiert als

$$\langle S \rangle := \langle s \mid s \in S \rangle.$$

Die Elemente von $\langle S \rangle$ werden *Linearkombinationen* über K (oder K*-Linearkombinationen* oder *Linearkombinationen*) von S genannt.

(b) Wir sagen, dass S ein *Erzeugendensystem* über K (oder K*-Erzeugendensystem* oder *Erzeugendensystem*) von V ist (oder dass V von S *erzeugt* wird oder dass V von S *aufgespannt* wird), wenn $(s)_{s \in S}$ ein Erzeugendensystem von V ist.

(c) Die Teilmenge S heißt *linear unabhängig* über K (oder K*-linear unabhängig* oder *linear unabhängig*) in V, wenn $(s)_{s \in S}$ linear unabhängig ist; ansonsten *linear abhängig* über K (oder K*-linear abhängig* oder *linear abhängig*) in V.

(d) Wir sagen, dass S eine *Basis* über K (oder K*-Basis* oder *Basis*) von V ist, wenn $(s)_{s \in S}$ eine Basis von V ist.

Unabhängigkeit von Untervektorräumen und innere direkte Summe

Schließlich führen wir noch einige Begriffe für Untervektorräume ein, welche die Begriffe des Erzeugendensystems, der linearen Unabhängigkeit und der Basis in einem präzisen Sinn auf Untervektorräume portieren, vgl. Bemerkung (1.76).

(1.73) Definition (innere Summe, (un)abhängig, innere direkte Summe)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $U = (U_1, \dots, U_n)$ von K-Untervektorräumen von V gegeben und es sei

$$\sigma_U \colon U_1 \times \dots \times U_n \to V,\ u \mapsto \sum_{i \in [1,n]} u_i.$$

(a) Die *innere Summe* (oder *Summe*) von $(U_1, \dots, U_n)$ ist definiert als

$$U_1 + \dots + U_n = \sum_{i \in [1,n]} U_i := \operatorname{Im} \sigma_U.$$

(b) Das n-Tupel $(U_1, \dots, U_n)$ heißt *unabhängig* in V, wenn σ_U injektiv ist; ansonsten *abhängig* in V.

(c) Wenn $(U_1, \dots, U_n)$ unabhängig ist, so sagen wir, dass $\sum_{i \in [1,n]} U_i$ eine *innere direkte Summe* (oder *direkte Summe*) von $(U_1, \dots, U_n)$ ist, und schreiben

$$U_1 \dotplus \dots \dotplus U_n = \dot{\sum_{i \in [1,n]}} U_i := \sum_{i \in [1,n]} U_i.$$

Zum Nachweis der Unabhängigkeit von Vektorräumen genügt es, analog zum Kriterium für lineare Unabhängigkeit (1.34), ein etwas einfacheres Kriterium zu verifizieren:

(1.74) Lemma (Kriterium für Unabhängigkeit von Untervektorräumen)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(U_1, \dots, U_n)$ von K-Untervektorräumen von V gegeben. Genau dann ist $(U_1, \dots, U_n)$ unabhängig in V, wenn für $u \in U_1 \times \dots \times U_n$ aus $\sum_{i \in [1,n]} u_i = 0$ stets $u = 0$ folgt.

Beweis. Zunächst sei $(U_1, \dots, U_n)$ unabhängig. Für $u \in U_1 \times \dots \times U_n$ mit $\sum_{i \in [1,n]} u_i = 0$ gilt dann

$$\sum_{i \in [1,n]} u_i = 0 = \sum_{i \in [1,n]} 0,$$

es folgt also $u = (0)_{i \in [1,n]} = 0$. Umgekehrt folge für $u \in U_1 \times \ldots \times U_n$ aus $\sum_{i \in [1,n]} u_i = 0$ stets $u = 0$. Für $u, u' \in U_1 \times \ldots \times U_n$ mit $\sum_{i \in [1,n]} u_i = \sum_{i \in [1,n]} u'_i$ gilt dann

$$0 = \sum_{i \in [1,n]} u_i - \sum_{i \in [1,n]} u'_i = \sum_{i \in [1,n]} (u_i - u'_i),$$

also $u - u' = (u_i - u'_i)_{i \in [1,n]} = 0$ und damit $u = u'$. Folglich ist $(U_1, \ldots, U_n)$ unabhängig. □

(1.75) Beispiel.

(a) Es ist

$$\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} \dot{+} \mathbb{R}(0,0,1).$$

(b) Es ist

$$\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} \dot{+} \mathbb{R}(1,1,1).$$

(c) Es ist $(\{(x,y,0) \mid x,y \in \mathbb{R}\}, \{(0,y,z) \mid y,z \in \mathbb{R}\})$ abhängig in $\mathbb{R}^3$ und

$$\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} + \{(0,y,z) \mid y,z \in \mathbb{R}\}.$$

Beweis.

(a) Für $x, y, z \in \mathbb{R}$ gilt

$$(x,y,z) = (x,y,0) + (0,0,z) = (x,y,0) + z(0,0,1),$$

und aus $(x,y,0) + z(0,0,1) = 0$ folgt $(x,y,z) = (0,0,0)$, also $x = 0$, $y = 0$, $z = 0$ und damit $(x,y,0) = 0$ und $(0,0,z) = 0$. Folglich ist $\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} + \mathbb{R}(0,0,1)$ und $(\{(x,y,0) \mid x,y \in \mathbb{R}\}, \mathbb{R}(0,0,1))$ ist unabhängig in $\mathbb{R}^3$, es gilt also

$$\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} \dot{+} \mathbb{R}(0,0,1).$$

(b) Für $x, y, z \in \mathbb{R}$ gilt

$$(x,y,z) = (x-z, y-z, 0) + (z,z,z) = (x-z, y-z, 0) + z(1,1,1).$$

und aus $(x,y,0) + z(1,1,1) = 0$ folgt $(x+z, y+z, z) = (0,0,0)$, also $x + z = 0$, $y + z = 0$, $z = 0$ und damit $(0,0,z) = 0$ und

$$(x,y,0) = (x+z, y+z, 0) - (0,0,z) = 0 - 0 = 0.$$

Folglich ist $\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} + \mathbb{R}(1,1,1)$ und $(\{(x,y,0) \mid x,y \in \mathbb{R}\}, \mathbb{R}(1,1,1))$ ist unabhängig in $\mathbb{R}^3$, es gilt also

$$\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} \dot{+} \mathbb{R}(1,1,1).$$

(c) Für $x, y, z \in \mathbb{R}$ gilt

$$(x,y,z) = (x,0,0) + (0,y,z).$$

Folglich ist

$$\mathbb{R}^3 = \{(x,y,0) \mid x,y \in \mathbb{R}\} + \{(0,y,z) \mid y,z \in \mathbb{R}\}.$$

Ferner gilt $(0,1,0) \in \{(x,y,0) \mid x,y \in \mathbb{R}\} \setminus \{0\}$ und $(0,-1,0) \in \{(0,y,z) \mid y,z \in \mathbb{R}\} \setminus \{0\}$ sowie

$$(0,1,0) + (0,-1,0) = 0,$$

es ist also $(\{(x,y,0) \mid x,y \in \mathbb{R}\}, \{(0,y,z) \mid y,z \in \mathbb{R}\})$ abhängig in $\mathbb{R}^3$. □

(1.76) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \ldots, s_n)$ in V gegeben.

(a) Es ist $\langle s_1, \ldots, s_n \rangle = \sum_{i \in [1,n]} K s_i$.

(b) Genau dann ist $(s_1, \ldots, s_n)$ ein Erzeugendensystem von V, wenn $V = \sum_{i \in [1,n]} K s_i$ ist.

(c) Genau dann ist $(s_1, \ldots, s_n)$ linear unabhängig in V, wenn $s_i \neq 0$ für $i \in [1,n]$ und $(K s_1, \ldots, K s_n)$ unabhängig in V ist.

(d) Genau dann ist $(s_1, \ldots, s_n)$ eine Basis von V, wenn $s_i \neq 0$ für $i \in [1,n]$ und $V = \dot{\sum}_{i \in [1,n]} K s_i$ ist.

Beweis.

(c) Zunächst sei $(s_1, \ldots, s_n)$ linear unabhängig in V. Dann ist insbesondere $s_i \neq 0$ für $i \in [1,n]$. Um zu zeigen, dass $(K s_1, \ldots, K s_n)$ unabhängig in V ist, seien $u, w \in K s_1 \times \ldots \times K s_n$ mit $\sum_{i \in [1,n]} u_i = \sum_{i \in [1,n]} w_i$ gegeben. Für jedes $i \in [1,n]$ gibt es dann $a_i, b_i \in K$ mit $u_i = a_i s_i$ und $w_i = b_i s_i$; es gilt also $\sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,n]} b_i s_i$. Auf Grund der linearen Unabhängigkeit von $(s_1, \ldots, s_n)$ impliziert dies aber bereits $a_i = b_i$ und damit insbesondere $u_i = a_i s_i = b_i s_i = w_i$ für $i \in [1,n]$, also $u = w$. Folglich ist $(K s_1, \ldots, K s_n)$ unabhängig in V.

Nun gelte umgekehrt $s_i \neq 0$ für $i \in [1,n]$ und es sei $(K s_1, \ldots, K s_n)$ unabhängig in V. Um zu zeigen, dass $(s_1, \ldots, s_n)$ linear unabhängig in V ist, seien $a, b \in K^n$ mit $\sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,n]} b_i s_i$ gegeben. Wegen $a_i s_i, b_i s_i \in K s_i$ für $i \in [1,n]$ impliziert die Unabhängigkeit von $(K s_1, \ldots, K s_n)$ nun $a_i s_i = b_i s_i$ für $i \in [1,n]$. Da für $i \in [1,n]$ aber $s_i \neq 0$ ist, folgt hieraus bereits $a_i = b_i$ für $i \in [1,n]$ nach Korollar (1.11), also $a = b$. Folglich ist $(s_1, \ldots, s_n)$ linear unabhängig in V. □

(1.77) Beispiel.

(a) Für $n \in \mathbb{N}_0$ ist $K^n = \dot{\sum}_{i \in [1,n]} K \mathrm{e}_i$.

(b) Für $m, n \in \mathbb{N}_0$ ist $K^{m \times n} = \dot{\sum}_{(i,j) \in [1,m] \times [1,n]} K \mathrm{e}_{i,j}$.

Die nachfolgende Proposition lässt sich als Verallgemeinerung von Proposition (1.37) auffassen.

(1.78) Proposition. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(U_1, \ldots, U_n)$ von K-Untervektorräumen von V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Das n-Tupel $(U_1, \ldots, U_n)$ ist unabhängig in V.

(b) Für $i \in [1,n]$ ist $(U_1, \ldots, U_{i-1}, U_{i+1}, \ldots, U_n)$ unabhängig in V und $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$.

(c) Es gilt entweder $n = 0$ oder es gibt ein $i \in [1,n]$ derart, dass $(U_1, \ldots, U_{i-1}, U_{i+1}, \ldots, U_n)$ unabhängig in V und $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$ ist.

(d) Für $i \in [1,n]$ ist $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$.

Beweis. Wir zeigen zuerst die Äquivalenz von Bedingung (a), Bedingung (b) und Bedingung (c), danach die Äquivalenz von Bedingung (a) und Bedingung (d).
Zunächst gelte Bedingung (a), d.h. $(U_j)_{j \in [1,n]}$ sei unabhängig in V, und es sei $i \in [1,n]$ gegeben. Für jede Familie $(u_j)_{j \in [1,n] \setminus \{i\}} \in \bigtimes_{j \in [1,n] \setminus \{i\}} U_j$ mit $\sum_{j \in [1,n] \setminus \{i\}} u_j = 0$ gilt dann

$$\sum_{j \in [1,i-1]} u_j + 0 + \sum_{j \in [i+1,n]} u_j = \sum_{j \in [1,n] \setminus \{i\}} u_j = 0,$$

und da U_i ein Untervektorraum von V und damit $0 \in U_i$ ist, folgt $u_j = 0$ für $j \in [1,n] \setminus \{i\}$ wegen der Unabhängigkeit von $(U_1, \ldots, U_n)$. Folglich ist $(U_j)_{j \in [1,n] \setminus \{i\}}$ unabhängig in V. Um zu zeigen, dass $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$ ist, sei $v \in U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j$ gegeben. Wegen $v \in \sum_{j \in [1,n] \setminus \{i\}} U_j$ gibt es eine Familie $(u_j)_{j \in [1,n] \setminus \{i\}} \in \bigtimes_{j \in [1,n] \setminus \{i\}} U_j$ mit $v = \sum_{j \in [1,n] \setminus \{i\}} u_j$. Es folgt

$$0 = v - \sum_{j \in [1,n] \setminus \{i\}} u_j = v + \sum_{j \in [1,n] \setminus \{i\}} (-u_j)$$

und damit $v = 0$ und $-u_j = 0$ für $j \in [1, n] \setminus \{i\}$ wegen der Unabhängigkeit von $(U_1, \dots, U_n)$. Da $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j$ ein Untervektorraum von V ist, gilt also $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$. Folglich gilt Bedingung (b).
Wenn Bedingung (b) gilt, dann gilt insbesondere Bedingung (c).
Es gelte Bedingung (c) und es sei $n \neq 0$, d.h. es gebe ein $i \in [1, n]$ derart, dass $(U_j)_{j \in [1,n] \setminus \{i\}}$ unabhängig in V und $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$ ist. Um zu zeigen, dass $(U_j)_{j \in [1,n]}$ unabhängig in V ist, sei $u \in \bigtimes_{j \in [1,n]} U_j$ mit $\sum_{j \in [1,n]} u_j = 0$ gegeben. Dann folgt

$$u_i = - \sum_{j \in [1,n] \setminus \{i\}} u_j = \sum_{j \in [1,n] \setminus \{i\}} (-u_j) \in U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$$

und damit $u_i = 0$. Dann ist aber auch $\sum_{j \in [1,n] \setminus \{i\}} u_j = 0$, so dass die Unabhängigkeit von $(U_j)_{j \in [1,n] \setminus \{i\}}$ auch $u_j = 0$ für $j \in [1, n] \setminus \{i\}$ liefert. Folglich ist $(U_j)_{j \in [1,n]}$ unabhängig in V, d.h. es gilt Bedingung (a).
Wir haben die Äquivalenz von Bedingung (a), Bedingung (b) und Bedingung (c) gezeigt.
Wenn also Bedingung (a) gilt, so gilt auch Bedingung (b) und damit insbesondere Bedingung (d).
Schließlich gelte Bedingung (d), d.h. für jedes $i \in [1, n]$ sei $U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$. Um zu zeigen, dass $(U_j)_{j \in [1,n]}$ unabhängig in V ist, sei $u \in \bigtimes_{j \in [1,n]} U_j$ mit $\sum_{j \in [1,n]} u_j = 0$ gegeben. Für jedes $i \in [1, n]$ ist dann

$$u_i = - \sum_{j \in [1,n] \setminus \{i\}} u_j = \sum_{j \in [1,n] \setminus \{i\}} (-u_j) \in U_i \cap \sum_{j \in [1,n] \setminus \{i\}} U_j = \{0\}$$

und damit $u_i = 0$. Folglich ist $(U_j)_{j \in [1,n]}$ unabhängig in V, d.h. es gilt Bedingung (a).
Wir haben gezeigt, dass auch Bedingung (a) und Bedingung (d) äquivalent sind.
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c) und Bedingung (d) äquivalent. □

(1.79) Proposition. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(U_1, \dots, U_n)$ von K-Untervektorräumen von V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $(U_1, \dots, U_n)$ unabhängig in V.

(b) Für jedes $i \in [1, n]$, jedes $k_i \in \mathbb{N}_0$ und jedes linear unabhängige k_i-Tupel $s_i = (s_{i,1}, \dots, s_{i,k_i})$ in U_i ist

$$(s_{1,1}, \dots, s_{1,k_1}, \dots, s_{n,1}, \dots, s_{n,k_n})$$

linear unabhängig in V.

(c) Für alle $u \in U_1 \times \ldots \times U_n$ ist

$$(u_i)_{\substack{i \in [1,n] \\ u_i \neq 0}}$$

linear unabhängig in V.

Beweis. Zunächst gelte Bedingung (a), d.h. $(U_i)_{i \in [1,n]}$ sei unabhängig. Ferner seien für jedes $i \in [1, n]$ ein $k_i \in \mathbb{N}_0$ und ein linear unabhängiges k_i-Tupel $s_i = (s_{i,1}, \dots, s_{i,k_i})$ in U_i gegeben. Um zu zeigen, dass

$$(s_{1,1}, \dots, s_{1,k_1}, \dots, s_{n,1}, \dots, s_{n,k_n})$$

linear unabhängig ist, seien $a_i \in K^{k_i}$ für $i \in [1, n]$ mit

$$\sum_{i \in [1,n]} \sum_{j \in [1,k_i]} a_{i,j} s_{i,j} = 0$$

gegeben. Da $(U_1, \dots, U_n)$ unabhängig ist, folgt $\sum_{j \in [1,k_i]} a_{i,j} s_{i,j} = 0$ für $i \in [1, n]$, und da $s_i = (s_{i,1}, \dots, s_{i,k_i})$ für jedes $i \in [1, n]$ linear unabhängig in U_i ist, folgt $a_i = 0$ für $i \in [1, n]$. Folglich ist

$$(s_{1,1}, \dots, s_{1,k_1}, \dots, s_{n,1}, \dots, s_{n,k_n})$$

in der Tat linear unabhängig, d.h. es gilt Bedingung (b).

Als nächstes gelte Bedingung (b), d.h. für jedes $i \in [1, n]$, jedes $k_i \in \mathbb{N}_0$ und jedes linear unabhängige k_i-Tupel $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ in U_i sei $(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$ linear unabhängig in V, und es sei $u \in \bigtimes_{i \in [1,n]} U_i$ gegeben. Für $i \in [1, n]$ ist $(\,)$ stets linear unabhängig in U_i und (u_i) ist genau dann linear unabhängig in U_i, wenn $u_i \neq 0$ ist. Nach unserer Annahme ist daher

$$(u_i)_{\substack{i \in [1,n] \\ u_i \neq 0}}$$

linear unabhängig in V. Folglich gilt Bedingung (c).
Schließlich gelte Bedingung (c), d.h. für $u \in U_1 \times \ldots \times U_n$ sei $(u_i)_{\substack{i \in [1,n] \\ u_i \neq 0}}$ stets linear unabhängig in V. Für $u \in U_1 \times \ldots \times U_n$ mit $\sum_{i \in [1,n]} u_i = 0$ gilt

$$\sum_{\substack{i \in [1,n] \\ u_i \neq 0}} u_i = 0,$$

so dass aus der linearen Unabhängigkeit von $(u_i)_{\substack{i \in [1,n] \\ u_i \neq 0}}$ bereits $\{i \in [1, n] \mid u_i \neq 0\} = \emptyset$ und damit $u = 0$ folgt.
Nach dem Kriterium für Unabhängigkeit von Untervektorräumen (1.74) ist somit $(U_1, \ldots, U_n)$ unabhängig in V, d.h. es gilt Bedingung (a).
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

(1.80) Korollar. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(U_1, \ldots, U_n)$ von endlichdimensionalen K-Untervektorräumen von V gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist

$$V = \dot{\sum_{i \in [1,n]}} U_i.$$

(b) Für $i \in [1, n]$, jedes $k_i \in \mathbb{N}_0$ und jede Basis $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ von U_i ist

$$(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$$

eine Basis von V.

(c) Für jedes $i \in [1, n]$ gibt es ein $k_i \in \mathbb{N}_0$ und eine Basis $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ von U_i so, dass

$$(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$$

eine Basis von V ist.

Beweis. Zunächst gelte Bedingung (a), d.h. es gelte $V = \dot{\sum}_{i \in [1,n]} U_i$. Ferner seien für $i \in [1, n]$ ein $k_i \in \mathbb{N}_0$ und eine Basis $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ von U_i gegeben. Dann ist $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ für $i \in [1, n]$ insbesondere ein Erzeugendensystem von U_i, d.h. es gilt $U_i = \sum_{j \in [1,k_i]} K s_{i,j}$. Dies impliziert

$$V = \sum_{i \in [1,n]} U_i = \sum_{i \in [1,n]} \sum_{j \in [1,k_i]} K s_{i,j},$$

es ist also $(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$ ein Erzeugendensystem von V. Da $(U_1, \ldots, U_n)$ unabhängig und s_i für $i \in [1, n]$ insbesondere linear unabhängig in U_i ist, ist $(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$ nach Proposition (1.79) linear unabhängig in V. Insgesamt ist $(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$ eine Basis von V. Folglich gilt Bedingung (b).
Wenn Bedingung (b) gilt, so gilt wegen der Endlichdimensionalität von U_i für $i \in [1, n]$ auch Bedingung (c).
Schließlich gelte Bedingung (c), d.h. für jedes $i \in [1, n]$ gebe es ein $k_i \in \mathbb{N}_0$ und eine Basis $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ von U_i so, dass $(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$ eine Basis von $\sum_{i \in [1,n]} U_i$ ist. Da $s_i = (s_{i,1}, \ldots, s_{i,k_i})$ für jedes $i \in [1, n]$ ein Erzeugendensystem von U_i ist und da $(s_{1,1}, \ldots, s_{1,k_1}, \ldots, s_{n,1}, \ldots, s_{n,k_n})$ ein Erzeugendensystem von V ist, gilt

$$V = \sum_{i \in [1,n]} \sum_{j \in [1,k_i]} K s_{i,j} = \sum_{i \in [1,n]} U_i.$$

Um zu zeigen, dass $(U_1, \dots, U_n)$ unabhängig ist, sei $u \in U_1 \times \ldots \times U_n$ mit $\sum_{i \in [1,n]} u_i = 0$ gegeben. Da $s_i = (s_{i,1}, \dots, s_{i,k_i})$ für $i \in [1,n]$ ein Erzeugendensystem von U_i ist, gibt es für jedes $i \in [1,n]$ ein $a_i \in K^{k_i}$ mit $u_i = \sum_{j \in [1,k_i]} a_{i,j} s_{i,j}$. Es folgt

$$0 = \sum_{i \in [1,n]} u_i = \sum_{i \in [1,n]} \sum_{j \in [1,k_i]} a_{i,j} s_{i,j}.$$

Die lineare Unabhängigkeit von $(s_{1,1}, \dots, s_{1,k_1}, \dots, s_{n,1}, \dots, s_{n,k_n})$ impliziert $a_i = 0$ und somit

$$u_i = \sum_{j \in [1,k_i]} a_{i,j} s_{i,j} = 0$$

für alle $i \in [1,n]$. Nach dem Kriterium für Unabhängigkeit von Untervektorräumen (1.74) ist $(U_1, \dots, U_n)$ unabhängig. Somit ist sogar

$$V = \dot{\sum_{i \in [1,n]}} U_i,$$

d.h. es gilt Bedingung (a).
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

(1.81) Korollar. Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(U_1, \dots, U_n)$ von endlichdimensionalen K-Untervektorräumen von V mit $V = \dot{\sum}_{i \in [1,n]} U_i$ gegeben. Dann ist

$$\dim_K V = \sum_{i \in [1,n]} (\dim_K U_i).$$

Beweis. Für jedes $i \in [1,n]$ ist U_i ein endlichdimensionaler K-Untervektorraum von V, es gibt also ein $k_i \in \mathbb{N}_0$ und eine Basis $s_i = (s_{i,1}, \dots, s_{i,k_i})$ von U_i. Wegen $V = \dot{\sum}_{i \in [1,n]} U_i$ ist dann

$$(s_{1,1}, \dots, s_{1,k_1}, \dots, s_{n,1}, \dots, s_{n,k_n})$$

nach Korollar (1.80) eine Basis von V. Es folgt

$$\dim V = \sum_{i \in [1,n]} k_i = \sum_{i \in [1,n]} (\dim U_i).$$ □

(1.82) Proposition. Es seien K-Untervektorräume U und U' von V, $m, n, n' \in \mathbb{N}_0$, ein m-Tupel $(s_1, \dots, s_m)$ in $U \cap U'$, ein n-Tupel $(t_1, \dots, t_n)$ in U und ein n'-Tupel $(t'_1, \dots, t'_{n'})$ in U' gegeben. Wenn drei der folgenden Bedingungen erfüllt sind, dann auch die vierte.

(a) Es ist $(s_1, \dots, s_m)$ eine Basis von $U \cap U'$.

(b) Es ist $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von U.

(c) Es ist $(s_1, \dots, s_m, t'_1, \dots, t'_{n'})$ eine Basis von U'.

(d) Es ist $(s_1, \dots, s_m, t_1, \dots, t_n, t'_1, \dots, t'_{n'})$ eine Basis von $U + U'$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(1.83) Korollar (Inklusions-Exklusionsprinzip für Untervektorräume)**.** Es seien K-Untervektorräume U und U' von V gegeben. Genau dann ist $U + U'$ endlichdimensional, wenn U und U' endlichdimensional sind. In diesem Fall gibt es $m, n, n' \in \mathbb{N}_0$, ein m-Tupel $(s_1, \dots, s_m)$, ein n-Tupel $(t_1, \dots, t_n)$ und ein n'-Tupel $(t'_1, \dots, t'_{n'})$ in V derart, dass $(s_1, \dots, s_m)$ eine Basis von $U \cap U'$ und $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von U und $(s_1, \dots, s_m, t'_1, \dots, t'_{n'})$ eine Basis von U' und $(s_1, \dots, s_m, t_1, \dots, t_n, t'_1, \dots, t'_{n'})$ eine Basis von $U + U'$ ist. Insbesondere gilt

$$\dim_K (U + U') = \dim_K U + \dim_K U' - \dim_K (U \cap U').$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Jede innere direkte Summe von Vektorräumen gibt Anlass zu kanonischen Abbildungen:

(1.84) Definition (Projektionen)**.** Es seien $n \in \mathbb{N}_0$ und ein n-Tupel $(U_1, \dots, U_n)$ von K-Untervektorräumen von V mit

$$V = \dot{\sum_{i \in [1,n]}} U_i$$

gegeben. Für $i \in [1, n]$ sei

$$\mathrm{pr}_i = \mathrm{pr}_i^V : V \to U_i$$

definiert durch $\mathrm{pr}_i(\sum_{j \in [1,n]} u_j) := u_i$ für $(u_1, \dots, u_n) \in U_1 \times \dots \times U_n$. Wir nennen pr_i für $i \in [1, n]$ die *Projektion* von V auf U_i (bzgl. $(U_1, \dots, U_n)$).

Die Projektionen in Definition (1.84) sind wohldefiniert, da es auf Grund der Unabhängigkeit von $(U_1, \dots, U_n)$ für jedes $v \in V$ genau ein $(u_1, \dots, u_n) \in U_1 \times \dots \times U_n$ mit $v = \sum_{j \in [1,n]} u_j$ gibt.

(1.85) Beispiel.

(a) Es seien $U_1 := \{(x, y, 0) \mid x, y \in \mathbb{R}\}$ und $U_2 := \mathbb{R}(0, 0, 1)$. Dann ist

$$\mathrm{pr}_1 : \mathbb{R}^3 \to \{(x, y, 0) \mid x, y \in \mathbb{R}\},\ (x, y, z) \mapsto (x, y, 0),$$
$$\mathrm{pr}_2 : \mathbb{R}^3 \to \mathbb{R}(0, 0, 1),\ (x, y, z) \mapsto (0, 0, z).$$

(b) Es seien $U_1 := \{(x, y, 0) \mid x, y \in \mathbb{R}\}$ und $U_2 := \mathbb{R}(1, 1, 1)$. Dann ist

$$\mathrm{pr}_1 : \mathbb{R}^3 \to \{(x, y, 0) \mid x, y \in \mathbb{R}\},\ (x, y, z) \mapsto (x - z, y - z, 0),$$
$$\mathrm{pr}_2 : \mathbb{R}^3 \to \mathbb{R}(1, 1, 1),\ (x, y, z) \mapsto (z, z, z).$$

(1.86) Definition (Untervektorraumkomplement)**.** Es sei ein K-Untervektorraum U von V gegeben. Ein *Untervektorraumkomplement* (oder *Komplement*) von U in V ist ein K-Untervektorraum U' von V mit

$$V = U \dot{+} U'.$$

(1.87) Beispiel.

(a) Der $\mathbb{R}$-Untervektorraum $\mathbb{R}(0, 0, 1)$ von $\mathbb{R}^3$ ist ein Untervektorraumkomplement von $\{(x, y, 0) \mid x, y \in \mathbb{R}\}$ in $\mathbb{R}^3$.

(b) Der $\mathbb{R}$-Untervektorraum $\mathbb{R}(1, 1, 1)$ von $\mathbb{R}^3$ ist ein Untervektorraumkomplement von $\{(x, y, 0) \mid x, y \in \mathbb{R}\}$ in $\mathbb{R}^3$.

Beweis.

(a) Dies folgt aus Beispiel (1.75)(a).

(b) Dies folgt aus Beispiel (1.75)(b). □

(1.88) Bemerkung. Es seien ein K-Untervektorraum U von V, $m, n \in \mathbb{N}_0$, eine Basis $(s_1, \dots, s_m)$ von U und ein n-Tupel $(t_1, \dots, t_n)$ in V so gegeben, dass $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist. Dann ist $\langle t_1, \dots, t_n \rangle$ ein Untervektorraumkomplement von U in V.

Beweis. Da $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist, ist $(t_1, \dots, t_n)$ linear unabhängig in V und damit in $\langle t_1, \dots, t_n \rangle$. Ferner ist $(t_1, \dots, t_n)$ ein Erzeugendensystem von $\langle t_1, \dots, t_n \rangle$, also insgesamt eine Basis von $\langle t_1, \dots, t_n \rangle$. Da aber $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist, gilt $V = U \dot{+} \langle t_1, \dots, t_n \rangle$ nach Korollar (1.80), d.h. $\langle t_1, \dots, t_n \rangle$ ist ein Komplement von U. □

(1.89) Korollar. Es sei V endlichdimensional. Dann hat jeder K-Untervektorraum U von V ein Untervektorraumkomplement in V.

Beweis. Es sei ein Untervektorraum U von V gegeben. Nach Proposition (1.68) ist dann U ebenfalls ein endlichdimensionaler Vektorraum, d.h. es gibt ein $m \in \mathbb{N}_0$ und eine Basis $(s_1, \dots, s_m)$ von U. Ferner gibt es nach Korollar (1.55) ein $n \in \mathbb{N}_0$ und ein n-Tupel $(t_1, \dots, t_n)$ in V so, dass $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Basis von V ist. Nach Bemerkung (1.88) ist $\langle t_1, \dots, t_n \rangle$ ein Komplement von U in V. □

Anwendung: Verlustbehaftete Datenkompression

Zum Schluss dieses Abschnitts erläutern wir, wie die Konzepte dieses Abschnitts für eine verlustbehaftete Datenkompression wie etwa das JPEG-Komprimierungsverfahren verwendet werden können. Die Grundidee dabei ist folgende: Die zu komprimierenden Daten werden wie in Anwendungsbeispiel (1.6) als ein Vektor v eines geeigneten K-Vektorraums V modelliert. Ferner betrachtet man, ggf. in Abhängigkeit von v, ein Paar (U_1, U_2) von Untervektorräumen von V derart, dass

$$V = U_1 \dot{+} U_2$$

ist und so, dass im betrachteten Kontext $\mathrm{pr}_1(v) \in U_1$ einem „relevanten Anteil“ und $\mathrm{pr}_2(v) \in U_2$ einem „irrelevanten Anteil“ der modellierten Daten entspricht:

$$v = \mathrm{pr}_1(v) + \mathrm{pr}_2(v)$$

Die dem Vektor $\mathrm{pr}_1(v)$ entsprechenden Daten bilden dann das Ergebnis der Kompression.
Es seien $m, n \in \mathbb{N}_0$ und eine geeignete Basis $(s_1, \ldots, s_n)$ von V so gegeben, dass $(s_1, \ldots, s_m)$ eine Basis von U_1 und $(s_{m+1}, \ldots, s_n)$ eine Basis von U_2 ist. Dann gibt es genau ein $a \in K^n$ mit

$$v = \sum_{i \in [1,n]} a_i s_i = \sum_{i \in [1,m]} a_i s_i + \sum_{i \in [m+1,n]} a_i s_i.$$

Die Speicherung des Vektors v kann durch die Speicherung von $a = (a_1, \ldots, a_n) = (a_1, \ldots, a_m, a_{m+1}, \ldots, a_n)$ erfolgen. Wegen

$$\mathrm{pr}_1(v) = \sum_{i \in [1,m]} a_i s_i$$

ist $\mathrm{pr}_1(v)$ bereits durch das kleinere Tupel $(a_1, \ldots, a_m)$ bestimmt. Für $\mathrm{pr}_1(v)$ genügt es daher, weniger Informationen zu speichern als für v. Das Verhältnis

$$\frac{\dim_K U}{\dim_K V} = \frac{m}{n}$$

wird *Kompressionsrate* genannt.
Kommen wir zu Anwendungsbeispiel (1.6) zurück: Ein digitales Bild mit einer Breite und Höhe aus jeweils acht Pixeln sei als Vektor in $\mathbb{R}^{[0,7] \times [0,7]}$ modelliert. Offenbar ist die Standardbasis $\mathrm{e} = (\mathrm{e}_{k,l})_{k,l \in [0,7]}$ von $\mathbb{R}^{[0,7] \times [0,7]}$, gegeben durch

$$\mathrm{e}_{k,l} = (\delta_{(i,j),(k,l)})_{i,j \in [0,7]}$$

für $k, l \in [0,7]$, für eine Datenkompression völlig ungeeignet, unabhängig vom betrachteten digitalen Bild: Möchte man hier eine Kompressionsrate von $75\% = \frac{48}{64}$ erzielen, so „verliert“ man 16 Pixel des Bildes. Geeigneter zur Datenkompression ist die Basis $(s_{k,l})_{k,l \in [0,7]}$ von $\mathbb{R}^{[0,7] \times [0,7]}$ gegeben durch

$$s_{k,l} = (\cos(\frac{2i+1}{16} k\pi) \cos(\frac{2j+1}{16} l\pi))_{i,j \in [0,7]}$$

für $k, l \in [0,7]$. Für kleine Indizes entsprechen die Koeffizienten hier groben Strukturen des Bildes, für große Indizes entsprechen sie feinen Details. Der Übergang von den Koeffizienten der Standardbasis e zu den Koeffizienten der Basis s ist als (*zweidimensionale*) *diskrete Kosinustransformation* bekannt.

2 Vektorraumhomomorphismen

Im letzten Abschnitt wurde das Konzept eines Vektorraums über einem Körper eingeführt. Neben Strukturen wie der eines Vektorraums sind oftmals auch die zugehörigen strukturerhaltenden Abbildungen, die sogenannten Homomorphismen, von Interesse. Zum einen helfen Sie uns, die Strukturen besser zu verstehen, zum anderen haben sie aber auch auf Grund ihres natürlichen Auftretens eine Daseinsberechtigung innerhalb unseres Studiums der Mathematik. Daher werden wir uns im Folgenden mit strukturerhaltenden Abbildungen von Vektorräumen beschäftigen.

Zur Motivation betrachten wir noch einmal das einführende Beispiel für Untervektorräume. In diesem haben wir den $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ in den $\mathbb{R}$-Vektorraum $\mathbb{R}^3$ eingebettet, indem wir die injektive Abbildung

$$\iota\colon \mathbb{R}^2 \to \mathbb{R}^3,\ (x, y) \mapsto (x, y, 0)$$

betrachtet haben und die Vektorraumstruktur von $\mathbb{R}^2$ auf $\operatorname{Im}\iota$ übersetzt haben: Für $x, y, x', y' \in \mathbb{R}$ gilt

$$(x, y, 0) + (x', y', 0) = (x + x', y + y', 0)$$

in $\operatorname{Im}\iota$, und für $a, x, y \in \mathbb{R}$ gilt

$$a(x, y, 0) = (ax, ay, 0)$$

in $\operatorname{Im}\iota$. Diese beiden Operationen entsprechen gerade den jeweiligen Operationen von $\mathbb{R}^3$, angewandt auf die Elemente von $\operatorname{Im}\iota$, d.h. $\operatorname{Im}\iota$ wird hierdurch zu einem $\mathbb{R}$-Untervektorraum von $\mathbb{R}^3$.
Um dieses Phänomen zu erklären, schreiben wir die beiden Gleichungen mit Hilfe der Abbildung ι um: Für $x, y, x', y' \in \mathbb{R}$ gilt

$$\iota((x, y)) + \iota((x', y')) = (x, y, 0) + (x', y', 0) = (x + x', y + y', 0) = \iota((x + x', y + y')) = \iota((x, y) + (x', y')),$$

und für $a, x, y \in \mathbb{R}$ gilt

$$a\,\iota((x, y)) = a(x, y, 0) = (ax, ay, 0) = \iota((ax, ay)) = \iota(a(x, y)).$$

Wie wir sehen, ist $\iota\colon \mathbb{R}^2 \to \mathbb{R}^3$ verträglich mit den Additionen und den Skalarmultiplikationen auf der Startmenge $\mathbb{R}^2$ und der Zielmenge $\mathbb{R}^3$: Es ist egal, ob wir zunächst Vektoren (x, y) und (x', y') aus $\mathbb{R}^2$ einzeln unter ι nach $\mathbb{R}^3$ abbilden und dort addieren, oder ob wir diese zuerst einzeln in $\mathbb{R}^2$ addieren und dann das Bild der Summe unter ι betrachten; in beiden Fällen erhalten wir dasselbe Ergebnis in $\mathbb{R}^3$. Genauso ist es für einen Skalar a aus $\mathbb{R}$ und einen Vektor (x, y) aus $\mathbb{R}^2$ unerheblich, ob wir das Bild $\iota((x, y))$ mit dem Skalar a multiplizieren, oder ob wir zuerst das skalare Vielfache $a(x, y)$ in $\mathbb{R}^2$ betrachten und dieses unter ι abbilden; auch in diesen beiden Fällen erhalten wir wieder dasselbe Ergebnis in $\mathbb{R}^3$. Diese Strukturverträglichkeit entspricht gerade der Idee eines Homomorphismus.
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein Körper K gegeben.

Begriffsbildung

Wir beginnen mit der Definition eines Homomorphismus. Dabei fordern wir zunächst noch etwas mehr Strukturverträglichkeit als in der Einleitung erläutert, werden aber in Bemerkung (2.2) sehen, dass diese zusätzlichen Forderungen redundant sind.

(2.1) Definition (Vektorraumhomomorphismus)**.** Es seien K-Vektorräume V und W gegeben. Ein *Vektorraumhomomorphismus* über K (oder *K-Vektorraumhomomorphismus* oder *Homomorphismus* von *K-Vektorräumen* oder *K-Homomorphismus* oder *Homomorphismus* oder *K-lineare Abbildung* oder *lineare Abbildung*) von V nach W ist eine Abbildung $\varphi\colon V \to W$ derart, dass folgende Axiome gelten.

- *Verträglichkeit mit den Additionen.* Für $v, v' \in V$ ist $\varphi(v + v') = \varphi(v) + \varphi(v')$.
- *Verträglichkeit mit den Nullvektoren.* Es ist $\varphi(0) = 0$.
- *Verträglichkeit mit den negativen Vektoren.* Für $v \in V$ ist $\varphi(-v) = -\varphi(v)$.
- *Verträglichkeit mit den Skalarmultiplikationen.* Für $a \in K$, $v \in V$ ist $\varphi(av) = a\,\varphi(v)$.

Die *Menge der K-Vektorraumhomomorphismen* von V nach W bezeichnen wir mit

$$\operatorname{Hom}(V, W) = \operatorname{Hom}_K(V, W) := \{\varphi \in \operatorname{Map}(V, W) \mid \varphi \text{ ist ein } K\text{-Vektorraumhomomorphismus}\}.$$

(2.2) Bemerkung (Kriterium für Vektorraumhomomorphismen)**.** Es seien K-Vektorräume V und W und eine Abbildung $\varphi\colon V \to W$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist φ ein K-Vektorraumhomomorphismus.

(b) Es gilt:

- *Verträglichkeit mit den Additionen (Additivität).* Für $v, v' \in V$ ist $\varphi(v + v') = \varphi(v) + \varphi(v')$.
- *Verträglichkeit mit den Skalarmultiplikationen (Homogenität).* Für $a \in K$, $v \in V$ ist $\varphi(av) = a\,\varphi(v)$.

(c) Für $a \in K$, $v, v' \in V$ ist

$$\varphi(av + v') = a\,\varphi(v) + \varphi(v').$$

Beweis. Wenn Bedingung (a) gilt, so insbesondere auch Bedingung (b).
Es gelte Bedingung (b), d.h. es gelte $\varphi(v + v') = \varphi(v) + \varphi(v')$ für $v, v' \in V$ sowie $\varphi(av) = a\,\varphi(v)$ für $a \in K$, $v \in V$. Wir erhalten

$$\varphi(av + v') = \varphi(av) + \varphi(v') = a\,\varphi(v) + \varphi(v')$$

für $a \in K$, $v, v' \in V$, d.h. Bedingung (c) ist erfüllt.
Schließlich gelte Bedingung (c), d.h. es gelte $\varphi(av + v') = a\,\varphi(v) + \varphi(v')$ für $a \in K$, $v, v' \in V$. Dann haben wir

$$\varphi(v + v') = \varphi(1v + v') = 1\,\varphi(v) + \varphi(v') = \varphi(v) + \varphi(v')$$

für $v, v' \in V$. Insbesondere gilt $\varphi(0) + \varphi(0) = \varphi(0 + 0) = \varphi(0)$ und damit $\varphi(0) = 0$. Für $v \in V$ folgt weiter

$$\varphi(-v) + \varphi(v) = \varphi(v) + \varphi(-v) = \varphi(v + (-v)) = \varphi(0) = 0$$

und damit $\varphi(-v) = -\varphi(v)$. Schließlich haben wir

$$\varphi(av) = \varphi(av + 0) = a\,\varphi(v) + \varphi(0) = a\,\varphi(v)$$

für $a \in K$, $v \in V$. Folglich ist φ ein K-Vektorraumhomomorphismus, d.h. Bedingung (a) gilt.
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Wir betrachten einige Beispiele und Gegenbeispiele zu Vektorraumhomomorphismen:

(2.3) Beispiel.

(a) Die Abbildung

$$\varphi\colon K^3 \to K^2,\ (x, y, z) \mapsto (z, x)$$

ist ein K-Vektorraumhomomorphismus.

(b) Die Abbildung

$$\varphi\colon K \to K^2,\ x \mapsto (0, x)$$

ist ein K-Vektorraumhomomorphismus.

(c) Die Abbildung

$$\varphi\colon \mathbb{Q}^{2\times 2} \to \mathbb{Q}^{1\times 3},\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 2a - 2b + c + 5d & -a + b + 3c + d & a - b - c + d \end{pmatrix}$$

ist ein $\mathbb{Q}$-Vektorraumhomomorphismus.

(d) Die Abbildung

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2,\ (x, y) \mapsto (x, y + 1)$$

ist kein $\mathbb{R}$-Vektorraumhomomorphismus.

(e) Die Abbildung

$$\varphi\colon \mathbb{Q} \to \mathbb{Q}^2,\ x \mapsto (x^2, x^3)$$

ist kein $\mathbb{Q}$-Vektorraumhomomorphismus.

Beweis.

(a) Für $(x,y,z),(x',y',z') \in K^3$ gilt

$$\begin{aligned}\varphi((x,y,z)+(x',y',z')) &= \varphi(x+x',y+y',z+z') = (z+z',x+x') = (z,x)+(z',x') \\ &= \varphi(x,y,z)+\varphi(x',y',z').\end{aligned}$$

Für $a \in K$, $(x,y,z) \in K^3$ gilt

$$\varphi(a(x,y,z)) = \varphi(ax,ay,az) = (az,ax) = a(z,x) = a\,\varphi(x,y,z).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein K-Vektorraumhomomorphismus.

(b) Für $a,x,y \in K$ gilt

$$\varphi(ax+y) = (0,ax+y) = a(0,x)+(0,y) = a\,\varphi(x)+\varphi(y).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein K-Vektorraumhomomorphismus.

(c) Für $a,b,c,d,a',b',c',d' \in \mathbb{Q}$ gilt

$$\begin{aligned}&\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}) = \varphi(\begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix}) \\ &= \big(2(a+a')-2(b+b')+(c+c')+5(d+d') \quad -(a+a')+(b+b')+3(c+c')+(d+d') \\ &\qquad (a+a')-(b+b')-(c+c')+(d+d')\big) \\ &= \big(2a-2b+c+5d+2a'-2b'+c'+5d' \quad -a+b+3c+d-a'+b'+3c'+d' \\ &\qquad a-b-c+d+a'-b'-c'+d'\big) \\ &= \big(2a-2b+c+5d \quad -a+b+3c+d \quad a-b-c+d\big) \\ &\quad + \big(2a'-2b'+c'+5d' \quad -a'+b'+3c'+d' \quad a'-b'-c'+d'\big) \\ &= \varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) + \varphi(\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}).\end{aligned}$$

Für $e,a,b,c,d \in \mathbb{Q}$ gilt

$$\begin{aligned}\varphi(e\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \varphi(\begin{pmatrix} ea & eb \\ ec & ed \end{pmatrix}) &= \big(2ea-2eb+ec+5ed \quad -ea+eb+3ec+ed \quad ea-eb-ec+ed\big) \\ &= \big(e(2a-2b+c+5d) \quad e(-a+b+3c+d) \quad e(a-b-c+d)\big) \\ &= e\,\big(2a-2b+c+5d \quad -a+b+3c+d \quad a-b-c+d\big) = e\,\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}).\end{aligned}$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein $\mathbb{Q}$-Vektorraumhomomorphismus.

(d) Es ist

$$\varphi(0^{\mathbb{R}^2}) = \varphi(0,0) = (0,0+1) = (0,1) \neq (0,0) = 0^{\mathbb{R}^2}.$$

Somit ist φ kein $\mathbb{R}$-Vektorraumhomomorphismus.

(e) Es ist $\varphi(1) = (1^2,1^3) = (1,1)$ und $\varphi(2) = (2^2,2^3) = (4,8)$, also

$$\varphi(1+1) = \varphi(2) = (4,8) \neq (1,1)+(1,1) = \varphi(1)+\varphi(1).$$

Somit ist φ kein $\mathbb{Q}$-Vektorraumhomomorphismus. □

Wir geben den Homomorphismen, bei denen Start und Ziel übereinstimmen, eine eigene Terminologie:

(2.4) Definition (Vektorraumendomorphismus)**.** Es sei ein K-Vektorraum V gegeben. Ein *Vektorraumendomorphismus über K* (oder *K-Vektorraumendomorphismus* oder *Endomorphismus von K-Vektorräumen* oder *K-Endomorphismus* oder *Endomorphismus*) von V ist ein K-Vektorraumhomomorphismus von V nach V.
Die *Menge der K-Vektorraumendomorphismen* von V bezeichnen wir mit

$$\mathrm{End}(V) = \mathrm{End}_K(V) := \mathrm{Hom}_K(V,V).$$

(2.5) Beispiel. Die Abbildung

$$\varphi\colon K^2 \to K^2,\ (x, y) \mapsto (y, 0)$$

ist ein K-Vektorraumendomorphismus.

Beweis. Für $(x, y), (x', y') \in K^2$ gilt

$$\varphi((x, y) + (x', y')) = \varphi(x + x', y + y') = (y + y', 0) = (y, 0) + (y', 0) = \varphi(x, y) + \varphi(x', y').$$

Für $a \in K$, $(x, y) \in K^2$ gilt

$$\varphi(a(x, y)) = \varphi(ax, ay) = (ay, 0) = a(y, 0) = a\,\varphi(x, y).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein K-Vektorraumhomomorphismus und wegen $\operatorname{Source}\varphi = K^2 = \operatorname{Target}\varphi$ somit ein K-Vektorraumendomorphismus. □

Komposition von Vektorraumhomomorphismen

Gewisse Paare von Abbildungen können komponiert werden, gewisse Abbildungen können invertiert werden. Wir wollen nun Komposition und Inversion von Vektorraumhomomorphismen studieren.

(2.6) Bemerkung.

(a) Für alle K-Vektorraumhomomorphismen $\varphi\colon V \to W$ und $\psi\colon W \to X$ ist $\psi \circ \varphi\colon V \to X$ ein K-Vektorraumhomomorphismus.

(b) Für jeden K-Vektorraum V ist $\mathrm{id}_V\colon V \to V$ ein K-Vektorraumhomomorphismus.

Beweis.

(a) Es seien K-Vektorraumhomomorphismen $\varphi\colon V \to W$ und $\psi\colon W \to X$ gegeben. Für $v, v' \in V$ gilt dann

$$(\psi \circ \varphi)(v + v') = \psi(\varphi(v + v')) = \psi(\varphi(v) + \varphi(v')) = \psi(\varphi(v)) + \psi(\varphi(v')) = (\psi \circ \varphi)(v) + (\psi \circ \varphi)(v')$$

und für $a \in K$, $v \in V$ gilt

$$(\psi \circ \varphi)(av) = \psi(\varphi(av)) = \psi(a\,\varphi(v)) = a\,\psi(\varphi(v)) = a\,(\psi \circ \varphi)(v).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist somit $\psi \circ \varphi\colon V \to X$ ein K-Vektorraumhomomorphismus.

(b) Es sei ein K-Vektorraum V gegeben. Für $v, v' \in V$ gilt dann

$$\mathrm{id}_V(v + v') = v + v' = \mathrm{id}_V(v) + \mathrm{id}_V(v')$$

und für $a \in K$, $v \in V$ gilt

$$\mathrm{id}_V(av) = av = a\,\mathrm{id}_V(v).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist somit $\mathrm{id}_V\colon V \to V$ ein K-Vektorraumhomomorphismus. □

Gewisse Abbildungen können bzgl. der Komposition invertiert werden. Dabei können Start- und Zielmenge einer solchen invertierbaren Abbildung über diese miteinander identifiziert werden: Die Abbildung liefert eine Korrespondenz zwischen den einzelnen Elementen. Mit anderen Worten: Unter rein quantitativen Gesichtspunkten sind diese Mengen nicht unterscheidbar.

Im Folgenden sind wir interessiert an Vektorraumhomomorphismen, die nicht nur als Abbildungen invertiert werden können, sondern sogar als Vektorraumhomomorphismen. Hierdurch wird sichergestellt, dass nicht nur die Elemente von Start- und Zielvektorraum eines solchen Homomorphismus einander entsprechen, sondern dass sich auch die Vektorraumstrukturen entsprechen. Bei strukturerhaltenden Abbildungen, welche als solche invertierbar sind, spricht man von sogenannten *Isomorphismen*; im Fall von Vektorräumen also von Vektorraumisomorphismen.

(2.7) Definition (Vektorraumisomorphismus)**.** Es seien K-Vektorräume V und W gegeben.

(a) Ein *Vektorraumisomorphismus über K* (oder *K-Vektorraumisomorphismus* oder *Isomorphismus von K-Vektorräumen* oder *K-Isomorphismus* oder *Isomorphismus*) von V nach W ist ein K-Vektorraumhomomorphismus $\varphi \colon V \to W$ so, dass φ eine invertierbare Abbildung und $\varphi^{-1} \colon W \to V$ ein K-Vektorraumhomomorphismus ist.

Die *Menge der K-Vektorraumisomorphismen* von V nach W bezeichnen wir mit

$$\operatorname{Iso}(V, W) = \operatorname{Iso}_K(V, W) := \{\varphi \in \operatorname{Hom}(V, W) \mid \varphi \text{ ist ein } K\text{-Vektorraumisomorphismus}\}.$$

(b) Wir sagen, dass V *isomorph* zu W (*als K-Vektorräume*) ist, geschrieben $V \cong W$, falls ein K-Vektorraumisomorphismus von V nach W existiert.

(2.8) Beispiel. Die Abbildung

$$K^2 \to K^{2 \times 1},\ (x, y) \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

ist ein K-Vektorraumisomorphismus.

Die folgende Bemerkung besagt, dass die Forderung der Linearität von φ^{-1} in Definition (2.7)(a) redundant ist; alle Vektorraumhomomorphismen, welche als Abbildungen invertierbar sind, sind de facto bereits Vektorraumisomorphismen.

(2.9) Bemerkung. Es sei ein K-Vektorraumhomomorphismus $\varphi \colon V \to W$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist φ ein K-Vektorraumisomorphismus.

(b) Es ist φ eine invertierbare Abbildung.

(c) Es ist φ bijektiv.

Beweis. Nach Satz (A.49)(c) ist Bedingung (b) äquivalent zu Bedingung (c). Ferner ist jeder Isomorphismus insbesondere eine invertierbare Abbildung, d.h. Bedingung (a) impliziert Bedingung (b). Um zu zeigen, dass die drei Bedingungen äquivalent sind, genügt es zu zeigen, dass Bedingung (b) Bedingung (a) impliziert.
Es gelte also Bedingung (b), d.h. es sei φ eine invertierbare Abbildung. Da φ ein Homomorphismus ist, haben wir

$$\varphi(a\,\varphi^{-1}(w) + \varphi^{-1}(w')) = a\,\varphi(\varphi^{-1}(w)) + \varphi(\varphi^{-1}(w')) = aw + w'$$

und damit $\varphi^{-1}(aw + w') = a\,\varphi^{-1}(w) + \varphi^{-1}(w')$ für $a \in K$, $w, w' \in W$. Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ^{-1} ein Homomorphismus und damit φ ein Isomorphismus, d.h. es gilt Bedingung (a). Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Isomorphismen, bei denen Start und Ziel übereinstimmen, welche also zugleich Endomorphismen sind, erhalten eine eigene Bezeichnung:

(2.10) Definition (Vektorraumautomorphismus)**.** Es sei ein K-Vektorraum V gegeben. Ein *Vektorraumautomorphismus über K* (oder *K-Vektorraumautomorphismus* oder *Automorphismus von K-Vektorräumen* oder *K-Automorphismus* oder *Automorphismus*) von V ist ein K-Vektorraumisomorphismus von V nach V.
Die *Menge der K-Vektorraumautomorphismen* von V bezeichnen wir mit

$$\operatorname{GL}(V) = \operatorname{Aut}(V) = \operatorname{Aut}_K(V) := \operatorname{Iso}_K(V, V).$$

(2.11) Beispiel. Die Abbildung

$$\varphi \colon K^2 \to K^2,\ (x, y) \mapsto (y, x)$$

ist ein K-Vektorraumautomorphismus.

Beweis. Für $(x, y), (x', y') \in K^2$ gilt

$$\varphi((x, y) + (x', y')) = \varphi(x + x', y + y') = (y + y', x + x') = (y, x) + (y', x') = \varphi(x, y) + \varphi(x', y').$$

Für $a \in K$, $(x, y) \in K^2$ gilt

$$\varphi(a(x, y)) = \varphi(ax, ay) = (ay, ax) = a(y, x) = a\,\varphi(x, y).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein K-Vektorraumhomomorphismus. Ferner gilt

$$\varphi(\varphi(x, y)) = \varphi(y, x) = (x, y)$$

für $(x, y) \in K^2$, also $\varphi \circ \varphi = \mathrm{id}_{K^2}$ nach dem Gleichheitskriterium für Abbildungen (A.35). Somit ist φ invertierbar mit $\varphi^{-1} = \varphi$. Insgesamt ist φ ein K-Vektorraumautomorphismus. □

Bild und Kern

Als nächstes werden wir sehen, dass jedem Vektorraumhomomorphismus zwei Untervektorräume zugeordnet werden können.

(2.12) Bemerkung. Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben.

(a) Es ist $\operatorname{Im}\varphi$ ein K-Untervektorraum von W.

(b) Es ist

$$\operatorname{Ker}\varphi = \{v \in V \mid \varphi(v) = 0\}$$

ein K-Untervektorraum von V.

Beweis.

(a) Für $v, v' \in V$ ist $\varphi(v) + \varphi(v') = \varphi(v + v') \in \operatorname{Im}\varphi$. Ferner ist $\varphi(0) = 0 \in \operatorname{Im}\varphi$. Für $a \in K$, $v \in V$ ist schließlich $a\,\varphi(v) = \varphi(av) \in \operatorname{Im}\varphi$. Nach dem Untervektorraumkriterium (1.15) ist also $\operatorname{Im}\varphi$ ein Untervektorraum von W.

(b) Für $v, v' \in \operatorname{Ker}\varphi$ gilt $\varphi(v) = 0$ und $\varphi(v') = 0$, also auch

$$\varphi(v + v') = \varphi(v) + \varphi(v') = 0 + 0 = 0$$

und damit $v + v' \in \operatorname{Ker}\varphi$. Wegen $\varphi(0) = 0$ ist $0 \in \operatorname{Ker}\varphi$. Schließlich gilt für $v \in \operatorname{Ker}\varphi$ stets $\varphi(v) = 0$, also für $a \in K$ auch

$$\varphi(av) = a\,\varphi(v) = a\,0 = 0$$

und damit $av \in \operatorname{Ker}\varphi$. Nach dem Untervektorraumkriterium (1.15) ist also $\operatorname{Ker}\varphi$ ein Untervektorraum von V. □

(2.13) Definition (Kern)**.** Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Der K-Untervektorraum

$$\operatorname{Ker}\varphi = \{v \in V \mid \varphi(v) = 0\}$$

von V heißt *Kern* von φ.

Wir ermitteln als Beispiel den Kern des Vektorraumhomomorphismus aus Beispiel (2.3)(c):

(2.14) Beispiel. Es sei der $\mathbb{Q}$-Vektorraumhomomorphismus

$$\varphi\colon \mathbb{Q}^{2\times 2} \to \mathbb{Q}^{1\times 3},\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 2a - 2b + c + 5d & -a + b + 3c + d & a - b - c + d \end{pmatrix}$$

gegeben.

(a) Es ist

$$\operatorname{Im}\varphi = \mathbb{Q}\begin{pmatrix}7 & 0 & 2\end{pmatrix} + \mathbb{Q}\begin{pmatrix}3 & 2 & 0\end{pmatrix}.$$

(b) Es ist

$$\operatorname{Ker}\varphi = \mathbb{Q}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix} + \mathbb{Q}\begin{pmatrix}-2 & 0\\ -1 & 1\end{pmatrix}.$$

Beweis.

(a) Es ist

$$\begin{aligned}\operatorname{Im}\varphi &= \{\varphi(\begin{pmatrix}a & b\\ c & d\end{pmatrix}) \mid \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2}\}\\ &= \{\begin{pmatrix}2a-2b+c+5d & -a+b+3c+d & a-b-c+d\end{pmatrix} \mid \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2}\}\\ &= \{a\begin{pmatrix}2 & -1 & 1\end{pmatrix} + b\begin{pmatrix}-2 & 1 & -1\end{pmatrix} + c\begin{pmatrix}1 & 3 & -1\end{pmatrix} + d\begin{pmatrix}5 & 1 & 1\end{pmatrix} \mid a,b,c,d \in \mathbb{Q}\}\\ &= \langle \begin{pmatrix}2 & -1 & 1\end{pmatrix}, \begin{pmatrix}-2 & 1 & -1\end{pmatrix}, \begin{pmatrix}1 & 3 & -1\end{pmatrix}, \begin{pmatrix}5 & 1 & 1\end{pmatrix}\rangle.\end{aligned}$$

Wir schreiben die vier angegebenen Erzeuger als Zeilen in eine Matrix und formen diese mittels elementarer Zeilenoperationen um:

$$\begin{pmatrix}2 & -1 & 1\\ -2 & 1 & -1\\ 1 & 3 & -1\\ 5 & 1 & 1\end{pmatrix} \xmapsto{\mathrm{add}_{4,1,-1}\circ\mathrm{add}_{3,1,1}\circ\mathrm{add}_{2,1,1}} \begin{pmatrix}2 & -1 & 1\\ 0 & 0 & 0\\ 3 & 2 & 0\\ 3 & 2 & 0\end{pmatrix} \xmapsto{\mathrm{mul}_{1,2}} \begin{pmatrix}4 & -2 & 2\\ 0 & 0 & 0\\ 3 & 2 & 0\\ 3 & 2 & 0\end{pmatrix}$$

$$\xmapsto{\mathrm{add}_{4,3,-1}\circ\mathrm{add}_{1,3,1}} \begin{pmatrix}7 & 0 & 2\\ 0 & 0 & 0\\ 3 & 2 & 0\\ 0 & 0 & 0\end{pmatrix}$$

Nach Korollar (1.23) und Proposition (1.22) ergibt sich

$$\operatorname{Im}\varphi = \langle \begin{pmatrix}7 & 0 & 2\end{pmatrix}, \begin{pmatrix}3 & 2 & 0\end{pmatrix}\rangle = \mathbb{Q}\begin{pmatrix}7 & 0 & 2\end{pmatrix} + \mathbb{Q}\begin{pmatrix}3 & 2 & 0\end{pmatrix}.$$

(b) Es ist

$$\begin{aligned}\operatorname{Ker}\varphi &= \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2} \mid \varphi(\begin{pmatrix}a & b\\ c & d\end{pmatrix}) = 0\}\\ &= \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2} \mid \begin{pmatrix}2a-2b+c+5d & -a+b+3c+d & a-b-c+d\end{pmatrix} = 0\}\\ &= \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2} \mid \begin{pmatrix}2a-2b+c+5d\\ -a+b+3c+d\\ a-b-c+d\end{pmatrix} = 0\}\\ &= \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2} \mid \begin{pmatrix}2 & -2 & 1 & 5\\ -1 & 1 & 3 & 1\\ 1 & -1 & -1 & 1\end{pmatrix}\begin{pmatrix}a\\ b\\ c\\ d\end{pmatrix} = 0\}\\ &= \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{Q}^{2\times 2} \mid \begin{pmatrix}a\\ b\\ c\\ d\end{pmatrix} \in \operatorname{Sol}(\begin{pmatrix}2 & -2 & 1 & 5\\ -1 & 1 & 3 & 1\\ 1 & -1 & -1 & 1\end{pmatrix}, 0)\}.\end{aligned}$$

Wir wenden elementare Zeilenoperationen an:

$$\begin{pmatrix}2 & -2 & 1 & 5\\ -1 & 1 & 3 & 1\\ 1 & -1 & -1 & 1\end{pmatrix} \xmapsto{\mathrm{add}_{2,3,1}\circ\mathrm{add}_{1,3,-2}} \begin{pmatrix}0 & 0 & 3 & 3\\ 0 & 0 & 2 & 2\\ 1 & -1 & -1 & 1\end{pmatrix} \xmapsto{\mathrm{mul}_{2,\frac{1}{2}}} \begin{pmatrix}0 & 0 & 3 & 3\\ 0 & 0 & 1 & 1\\ 1 & -1 & -1 & 1\end{pmatrix}$$

$$\xmapsto{\mathrm{add}_{3,2,1} \circ \mathrm{add}_{1,2,-3}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 2 \end{pmatrix} \xmapsto{\mathrm{sw}_{1,3}} \begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Nach Proposition (A.163) ist

$$\mathrm{Sol}(\begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}, 0) = \mathrm{Sol}(\begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, 0) = \mathbb{Q} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

und damit

$$\begin{aligned} \operatorname{Ker} \varphi &= \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Q}^{2 \times 2} \mid \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \in \mathbb{Q} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}\} \\ &= \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Q}^{2 \times 2} \mid \text{es gibt } e, f \in \mathbb{Q} \text{ mit } \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = e \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + f \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}\} \\ &= \{\begin{pmatrix} e - 2f & e \\ -f & f \end{pmatrix} \mid e, f \in \mathbb{Q}\} = \mathbb{Q} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}. \end{aligned}$$

□

Mit Hilfe des Kerns, also der Faser über dem Nullvektor, können wir alle nicht-leeren Fasern eines Vektorraumhomomorphismus beschreiben:

(2.15) Proposition. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ und $v \in V$ gegeben. Dann ist

$$\operatorname{Ker} \varphi \to \varphi^{-1}(\{\varphi(v)\}),\ u \mapsto v + u$$

eine wohldefinierte Bijektion. Insbesondere ist

$$\varphi^{-1}(\{\varphi(v)\}) = v + \operatorname{Ker} \varphi.$$

Beweis. Auf Grund der Negierbarkeit von v ist $f\colon V \to V$, $u \mapsto v + u$ eine Bijektion. Ferner ist

$$\begin{aligned} f(\operatorname{Ker} \varphi) &= \{v' \in V \mid \text{es gibt ein } u \in \operatorname{Ker} \varphi \text{ mit } v' = f(u)\} \\ &= \{v' \in V \mid \text{es gibt ein } u \in \operatorname{Ker} \varphi \text{ mit } v' = v + u\} \\ &= \{v' \in V \mid \text{es gibt ein } u \in \operatorname{Ker} \varphi \text{ mit } v' - v = u\} = \{v' \in V \mid v' - v \in \operatorname{Ker} \varphi\} \\ &= \{v' \in V \mid \varphi(v' - v) = 0\} = \{v' \in V \mid \varphi(v') - \varphi(v) = 0\} = \{v' \in V \mid \varphi(v') = \varphi(v)\} \\ &= \varphi^{-1}(\{\varphi(v)\}), \end{aligned}$$

so dass f zu einer surjektiven Abbildung $f|_{\operatorname{Ker} \varphi}^{\varphi^{-1}(\{\varphi(v)\})}\colon \operatorname{Ker} \varphi \to \varphi^{-1}(\{\varphi(v)\})$ einschränkt. Die Injektivität von f impliziert ferner die Injektivität von $f|_{\operatorname{Ker} \varphi}^{\varphi^{-1}(\{\varphi(v)\})}$. Insgesamt ist $f|_{\operatorname{Ker} \varphi}^{\varphi^{-1}(\{\varphi(v)\})}$ eine Bijektion. Insbesondere ist

$$\varphi^{-1}(\{\varphi(v)\}) = \operatorname{Im} f|_{\operatorname{Ker} \varphi}^{\varphi^{-1}(\{\varphi(v)\})} = f(\operatorname{Ker} \varphi) = v + \operatorname{Ker} \varphi.$$

□

(2.16) Korollar. Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Der K-Vektorraumhomomorphismus $\varphi\colon V \to W$ ist injektiv.

(b) Es ist $\operatorname{Ker} \varphi = \{0\}$.

(c) Es gibt ein $w \in W$ derart, dass $\varphi^{-1}(\{w\})$ einelementig ist.

Beweis. Wir zeigen zuerst die Äquivalenz von Bedingung (a) und Bedingung (b) und danach die Äquivalenz von Bedingung (b) und Bedingung (c).
Zunächst gelte Bedingung (a), d.h. φ sei injektiv. Für $v \in \operatorname{Ker} \varphi$ ist $\varphi(v) = 0 = \varphi(0)$, wegen der Injektivität also $v = 0$. Somit ist $\operatorname{Ker} \varphi \subseteq \{0\}$. Da aber 0 in jedem Untervektorraum enthalten ist, impliziert dies bereits $\operatorname{Ker} \varphi = \{0\}$, d.h. es gilt Bedingung (b).
Nun gelte umgekehrt Bedingung (b), d.h. es sei $\operatorname{Ker} \varphi = \{0\}$. Ferner sei $w \in W$ gegeben. Wenn $w \notin \operatorname{Im} \varphi$ ist, so ist $\varphi^{-1}(\{w\}) = \emptyset$. Daher sei im Folgenden $w \in \operatorname{Im} \varphi$, so dass es ein $v \in V$ mit $w = \varphi(v)$ gibt. Nach Proposition (2.15) ist dann aber

$$\varphi^{-1}(\{w\}) = \varphi^{-1}(\{\varphi(v)\}) = v + \operatorname{Ker} \varphi = v + \{0\} = \{v\}.$$

Somit besitzt jede Faser von φ höchstens ein Element. Nach Satz (A.49)(a) ist φ injektiv, d.h. es gilt Bedingung (a).
Folglich sind Bedingung (a) und Bedingung (b) äquivalent.
Wenn Bedingung (b) gilt, so ist $\varphi^{-1}(\{0\}) = \operatorname{Ker} \varphi = \{0\}$ einelementig, es gilt also insbesondere auch Bedingung (c).
Schließlich gelte Bedingung (c), d.h. es gebe ein $w \in W$ derart, dass $\varphi^{-1}(\{w\})$ einelementig ist. Ferner sei $v \in V$ mit $\varphi^{-1}(\{w\}) = \{v\}$ gegeben. Dann ist $w = \varphi(v)$, nach Proposition (2.15) gilt also

$$v + \operatorname{Ker} \varphi = \varphi^{-1}(\{\varphi(v)\}) = \varphi^{-1}(\{w\}) = \{v\}$$

und damit $\operatorname{Ker} \varphi = \{0\}$, d.h. es gilt Bedingung (b).
Folglich sind auch Bedingung (b) und Bedingung (c) äquivalent.
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

Nach Korollar (2.16) ist der Kern eines Vektorraumhomomorphismus also ein „Maß“ dafür, wie weit der Homomorphismus davon „entfernt“ ist, injektiv zu sein – ganz analog zur Tatsache, dass das Bild einer Abbildung ein „Maß“ für deren Surjektivität ist.

(2.17) Korollar. Ein K-Vektorraumhomomorphismus $\varphi \colon V \to W$ ist genau dann ein K-Vektorraumisomorphismus, wenn $\operatorname{Ker} \varphi = \{0\}$ und $\operatorname{Im} \varphi = W$ ist.

Beweis. Nach Bemerkung (2.9) ist ein Homomorphismus $\varphi \colon V \to W$ genau dann ein Isomorphismus, wenn er bijektiv, also injektiv und surjektiv, ist. Nach Korollar (2.16) ist φ genau dann injektiv, wenn $\operatorname{Ker} \varphi = \{0\}$ ist. Ferner ist φ genau dann surjektiv, wenn $\operatorname{Im} \varphi = W$ ist. Insgesamt ist φ genau dann ein Isomorphismus, wenn $\operatorname{Ker} \varphi = \{0\}$ und $\operatorname{Im} \varphi = W$ ist. □

Fortsetzen von Vektorraumhomomorphismen

Zur effizienten Beschreibung von (Unter-)Vektorräumen wurden in Abschnitt 1 einige Begriffe, wie etwa der der linearen Hülle, siehe Definition (1.18), oder der einer Basis, siehe Definition (1.39), eingeführt. Im Folgenden werden wir sehen, dass diese Begriffe auch zur Beschreibung von Vektorraumhomomorphismen dienen.

(2.18) Bemerkung. Es seien ein K-Vektorraum V, ein $n \in \mathbb{N}_0$ und ein n-Tupel $s = (s_1, \ldots, s_n)$ in V gegeben. Dann ist

$$\lambda_s \colon K^n \to V,\, a \mapsto \sum_{i \in [1,n]} a_i s_i$$

ein K-Vektorraumhomomorphismus mit $\lambda_s(\mathrm{e}_j) = s_j$ für $j \in [1, n]$.

Beweis. Für $a, a' \in K^n$ gilt

$$\lambda_s(a + a') = \sum_{i \in [1,n]} (a + a')_i s_i = \sum_{i \in [1,n]} (a_i + a'_i) s_i = \sum_{i \in [1,n]} a_i s_i + \sum_{i \in [1,n]} a'_i s_i = \lambda_s(a) + \lambda_s(a').$$

Für $c \in K$, $a \in K^n$ gilt

$$\lambda_s(ca) = \sum_{i \in [1,n]} (ca)_i s_i = \sum_{i \in [1,n]} c a_i s_i = c \sum_{i \in [1,n]} a_i s_i = c \, \lambda_s(a).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist also λ_s ein Homomorphismus. Für $j \in [1,n]$ gilt ferner

$$\lambda_s(\mathrm{e}_j) = \sum_{i\in[1,n]} (\mathrm{e}_j)_i s_i = \sum_{i\in[1,n]} \delta_{i,j} s_i = s_j. \qquad \square$$

(2.19) Proposition. Es seien ein endlichdimensionaler K-Vektorraum V, $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1,\dots,s_n)$ in V gegeben.

(a) Wenn $(s_1,\dots,s_n)$ ein Erzeugendensystem von V ist, so gibt es für jeden K-Vektorraum W und jedes n-Tupel $(t_1,\dots,t_n)$ in W höchstens einen K-Vektorraumhomomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für alle $i \in [1,n]$.

(b) Wenn $(s_1,\dots,s_n)$ linear unabhängig in V ist, so gibt es für jeden K-Vektorraum W und jedes n-Tupel $(t_1,\dots,t_n)$ in W mindestens einen K-Vektorraumhomomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für alle $i \in [1,n]$.

(c) Wenn $(s_1,\dots,s_n)$ eine Basis von V ist, so gibt es für jeden K-Vektorraum W und jedes n-Tupel $(t_1,\dots,t_n)$ in W genau einen K-Vektorraumhomomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für alle $i \in [1,n]$.

Beweis.

(a) Es sei $(s_1,\dots,s_n)$ ein Erzeugendensystem von V und es seien ein K-Vektorraum W und ein n-Tupel $(t_1,\dots,t_n)$ in W und ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für $i \in [1,n]$ gegeben. Ferner sei $v \in V$ gegeben. Da $(s_1,\dots,s_n)$ ein Erzeugendensystem von V ist, gibt es ein $a \in K^n$ mit $v = \sum_{i\in[1,n]} a_i s_i$. Wir erhalten

$$\varphi(v) = \varphi(\sum_{i\in[1,n]} a_i s_i) = \sum_{i\in[1,n]} a_i\,\varphi(s_i) = \sum_{i\in[1,n]} a_i t_i.$$

Da $v \in V$ beliebig war, ist φ durch $(t_1,\dots,t_n)$ bereits eindeutig bestimmt.

(c) Es sei $s = (s_1,\dots,s_n)$ eine Basis von V und es seien ein K-Vektorraum W und ein n-Tupel $t = (t_1,\dots,t_n)$ in W gegeben. Nach (a) gibt es höchstens einen Homomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für $i \in [1,n]$.

Wir müssen also noch die Existenz eines solchen Homomorphismus zeigen. Nach Bemerkung (2.18) sind $\lambda_s\colon K^n \to V,\ a \mapsto \sum_{i\in[1,n]} a_i s_i$ und $\lambda_t\colon K^n \to W,\ a \mapsto \sum_{i\in[1,n]} a_i t_i$ Homomorphismen. Da s eine Basis von V ist, ist ferner λ_s bijektiv und damit ein Isomorphismus nach Bemerkung (2.9). Folglich ist $\varphi := \lambda_t \circ \lambda_s^{-1}$ ein Homomorphismus nach Bemerkung (2.6)(a).

$$\begin{tikzcd} K^n \arrow[r, "\lambda_t"] \arrow[d, "\lambda_s"', "\cong"] & W \\ V \arrow[ur, "\varphi"'] & \end{tikzcd} \qquad \begin{tikzcd} \mathrm{e}_j \arrow[r, mapsto, "\lambda_t"] \arrow[d, mapsto, "\lambda_s"'] & t_j \\ s_j \arrow[ur, mapsto, "\varphi"'] & \end{tikzcd}$$

Für $i \in [1,n]$ gilt nach Bemerkung (2.18) schließlich

$$\varphi(s_i) = \lambda_t(\lambda_s^{-1}(s_i)) = \lambda_t(\mathrm{e}_i) = t_i.$$

(b) Es sei $(s_1,\dots,s_n)$ linear unabhängig in V und es seien ein K-Vektorraum W und n-Tupel $(t_1,\dots,t_n)$ in W gegeben. Nach dem Basisergänzungssatz (1.53) gibt es ein $p \in \mathbb{N}_0$ und ein p-Tupel $(s'_1,\dots,s'_p)$ in V so, dass $(s_1,\dots,s_n,s'_1,\dots,s'_p)$ eine Basis von V ist. Nun gibt es aber nach (c) (genau) einen Homomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für $i \in [1,n]$ und $\varphi(s'_j) = 0$ für $j \in [1,p]$. $\square$

Alternativer Beweis von Proposition (2.19)(c). Da $s = (s_1,\dots,s_n)$ eine Basis von V ist, gibt es für jedes $v \in V$ genau ein $a \in K^n$ mit $v = \sum_{i\in[1,n]} a_i s_i$. Folglich erhalten wir eine wohldefinierte Abbildung $\varphi\colon V \to W$ mit

$$\varphi(\sum_{i\in[1,n]} a_i s_i) := \sum_{i\in[1,n]} a_i t_i$$

für $a \in K^n$. Wegen

$$\varphi(\sum_{i \in [1,n]} a_i s_i + \sum_{i \in [1,n]} b_i s_i) = \varphi(\sum_{i \in [1,n]} (a_i + b_i) s_i) = \sum_{i \in [1,n]} (a_i + b_i) t_i = \sum_{i \in [1,n]} a_i t_i + \sum_{i \in [1,n]} b_i t_i$$
$$= \varphi(\sum_{i \in [1,n]} a_i s_i) + \varphi(\sum_{i \in [1,n]} b_i s_i)$$

für $a, b \in K^n$ und

$$\varphi(c \sum_{i \in [1,n]} a_i s_i) = \varphi(\sum_{i \in [1,n]} c a_i s_i) = \sum_{i \in [1,n]} c a_i t_i = c \sum_{i \in [1,n]} a_i t_i = c\,\varphi(\sum_{i \in [1,n]} a_i s_i)$$

für $c \in K$, $a \in K^n$ ist φ ein K-Vektorraumhomomorphismus. Schließlich gilt

$$\varphi(s_j) = \varphi(\sum_{i \in [1,n]} \delta_{i,j} s_i) = \sum_{i \in [1,n]} \delta_{i,j} t_i = t_j$$

für $j \in [1, n]$. □

(2.20) Beispiel. Es gibt genau einen $\mathbb{Q}$-Vektorraumhomomorphismus $\varphi\colon \mathbb{Q}^{2 \times 2} \to \mathbb{Q}^{2 \times 1}$ mit

$$\varphi(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ \varphi(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},\ \varphi(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \varphi(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}) = \begin{pmatrix} 2 \\ -2 \end{pmatrix},$$

gegeben durch

$$\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} a + 2d \\ 2a - b - 2d \end{pmatrix}$$

für $a, b, c, d \in \mathbb{Q}$.

Beweis. Es ist

$$(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$

eine Basis von $\mathbb{Q}^{2 \times 2}$. Somit gibt es nach Proposition (2.19)(c) genau einen Homomorphismus $\varphi\colon \mathbb{Q}^{2 \times 2} \to \mathbb{Q}^{2 \times 1}$ mit

$$\varphi(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ \varphi(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},\ \varphi(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \varphi(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}) = \begin{pmatrix} 2 \\ -2 \end{pmatrix},$$

gegeben durch

$$\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \varphi(a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$
$$= a \begin{pmatrix} 1 \\ 2 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \end{pmatrix} + d \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} a + 2d \\ 2a - b - 2d \end{pmatrix}$$

für $a, b, c, d \in \mathbb{Q}$. □

(2.21) Proposition. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben.

(a) Genau dann ist φ surjektiv, wenn $(\varphi(s_1), \dots, \varphi(s_n))$ ein Erzeugendensystem von W ist.

(b) Genau dann ist φ injektiv, wenn $(\varphi(s_1), \dots, \varphi(s_n))$ linear unabhängig in W ist.

(c) Genau dann ist φ ein Isomorphismus, wenn $(\varphi(s_1), \dots, \varphi(s_n))$ eine Basis von W ist.

Beweis. Es sei $t := (\varphi(s_1), \dots, \varphi(s_n))$. Da s eine Basis von V ist, ist $\lambda_s\colon K^n \to V$, $a \mapsto \sum_{i \in [1,n]} a_i s_i$ bijektiv. Nach Bemerkung (2.18) ist $\varphi = \lambda_t \circ \lambda_s^{-1}$ mit $\lambda_t\colon K^n \to W$, $a \mapsto \sum_{i \in [1,n]} a_i t_i$.

(a) Genau dann ist φ surjektiv, wenn λ_t surjektiv ist, d.h. wenn t ein Erzeugendensystem von W ist.

(b) Genau dann ist φ injektiv, wenn λ_t injektiv ist, d.h. wenn t linear unabhängig in W ist.

(c) Nach Bemerkung (2.9) ist φ genau dann ein Isomorphismus, wenn φ bijektiv ist. Dies gilt aber genau dann, wenn λ_t bijektiv ist, d.h. wenn t eine Basis von W ist. □

Rang und Defekt

Nachdem wir in Proposition (2.21) gesehen haben, dass sich Basen gut zur Beschreibung von Vektorraumhomomorphismen eignen, stellen wir nun im endlichdimensionalen Fall einen Zusammenhang zwischen Bild und Kern eines Vektorraumhomomorphismus her.

(2.22) Definition (Rang, Defekt)**.** Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben.

(a) Es sei $\operatorname{Im} \varphi$ endlichdimensional. Der *Rang* von φ über K (oder K-*Rang* von φ oder *Rang* von φ) ist definiert als

$$\operatorname{rk} \varphi = \operatorname{rk}_K \varphi := \dim_K(\operatorname{Im} \varphi).$$

(b) Es sei $\operatorname{Ker} \varphi$ endlichdimensional. Der *Defekt* von φ ist definiert als

$$\operatorname{def} \varphi = \operatorname{def}_K \varphi := \dim_K(\operatorname{Ker} \varphi).$$

(2.23) Beispiel. Es sei der $\mathbb{Q}$-Vektorraumhomomorphismus

$$\varphi\colon \mathbb{Q}^{2\times 2} \to \mathbb{Q}^{1\times 3}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 2a - 2b + c + 5d & -a + b + 3c + d & a - b - c + d \end{pmatrix}$$

gegeben.

(a) Es ist

$$\operatorname{rk}_{\mathbb{Q}} \varphi = 2.$$

(b) Es ist

$$\operatorname{def}_{\mathbb{Q}} \varphi = 2.$$

Beweis.

(a) Nach Beispiel (2.14)(a) ist

$$(\begin{pmatrix} 7 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 2 & 0 \end{pmatrix})$$

eine Basis von $\operatorname{Im} \varphi$ und damit

$$\operatorname{rk} \varphi = \dim(\operatorname{Im} \varphi) = 2.$$

(b) Nach Beispiel (2.14)(b) ist

$$(\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix})$$

eine Basis von $\operatorname{Ker} \varphi$ und damit

$$\operatorname{def} \varphi = \dim(\operatorname{Ker} \varphi) = 2.$$

□

(2.24) Bemerkung. Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben.

(a) Es sei W endlichdimensional. Dann ist $\operatorname{Im} \varphi$ endlichdimensional und es gilt

$$\operatorname{rk}_K \varphi \leq \dim_K W.$$

(b) Es sei V endlichdimensional. Dann ist $\operatorname{Ker} \varphi$ endlichdimensional und es gilt

$$\operatorname{def}_K \varphi \leq \dim_K V.$$

Beweis.

(a) Nach Bemerkung (2.12)(a) ist $\operatorname{Im}\varphi$ ein Untervektorraum des endlichdimensionalen K-Vektorraums W und damit endlichdimensional mit

$$\operatorname{rk}\varphi = \dim(\operatorname{Im}\varphi) \leq \dim W$$

nach Proposition (1.68).

(b) Nach Bemerkung (2.12)(b) ist $\operatorname{Ker}\varphi$ ein Untervektorraum des endlichdimensionalen K-Vektorraums V und damit endlichdimensional mit

$$\operatorname{def}\varphi = \dim(\operatorname{Ker}\varphi) \leq \dim V$$

nach Proposition (1.68). □

(2.25) Bemerkung. Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben.

(a) Es sei W endlichdimensional. Genau dann ist φ surjektiv, wenn

$$\operatorname{rk}_K \varphi = \dim_K W,$$

gilt.

(b) Genau dann ist φ injektiv, wenn $\operatorname{Ker}\varphi$ endlichdimensional ist und

$$\operatorname{def}_K \varphi = 0$$

gilt.

Beweis.

(a) Nach Bemerkung (2.24)(a) ist $\operatorname{Im}\varphi$ endlichdimensional. Da $\operatorname{rk}\varphi = \dim(\operatorname{Im}\varphi)$ ist, gilt nach Proposition (1.68) genau dann $\operatorname{rk}\varphi = \dim W$, wenn $\operatorname{Im}\varphi = W$ ist, d.h. wenn φ surjektiv ist.

(b) Nach Korollar (2.16) ist φ genau dann injektiv, wenn $\operatorname{Ker}\varphi = \{0\}$ ist. Falls $\operatorname{Ker}\varphi$ endlichdimensional ist, so gilt wegen $\operatorname{def}\varphi = \dim(\operatorname{Ker}\varphi)$ genau dann $\operatorname{Ker}\varphi = \{0\}$, wenn $\operatorname{def}\varphi = 0$ ist. Somit ist φ genau dann injektiv, wenn $\operatorname{Ker}\varphi$ endlichdimensional ist und $\operatorname{def}\varphi = 0$ gilt. □

(2.26) Korollar. Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ so gegeben, dass W und $\operatorname{Ker}\varphi$ endlichdimensional sind. Genau dann ist φ ein K-Vektorraumisomorphismus, wenn $\operatorname{def}_K \varphi = 0$ und $\operatorname{rk}_K \varphi = \dim_K W$ ist.

Beweis. Nach Bemerkung (2.9) ist ein Homomorphismus $\varphi\colon V \to W$ genau dann ein Isomorphismus, wenn er bijektiv, also injektiv und surjektiv, ist. Ferner ist φ nach Bemerkung (2.25)(b) genau dann injektiv, wenn $\operatorname{def}\varphi = 0$ ist, und nach Bemerkung (2.25)(a) genau dann surjektiv, wenn $\operatorname{rk}\varphi = \dim W$ ist. Insgesamt ist φ genau dann ein Isomorphismus, wenn $\operatorname{def}\varphi = 0$ und $\operatorname{rk}\varphi = \dim W$ ist. □

(2.27) Proposition. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $r, d \in \mathbb{N}_0$, ein r-Tupel $(s_1, \ldots, s_r)$ in V und ein d-Tupel $(s'_1, \ldots, s'_d)$ in $\operatorname{Ker}\varphi$ gegeben. Wenn zwei der folgenden Bedingungen erfüllt sind, dann auch die dritte.

(a) Es ist $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ eine Basis von V.

(b) Es ist $(\varphi(s_1), \ldots, \varphi(s_r))$ eine Basis von $\operatorname{Im}\varphi$.

(c) Es ist $(s'_1, \ldots, s'_d)$ eine Basis von $\operatorname{Ker}\varphi$.

Beweis. Zunächst gelte Bedingung (a) und Bedingung (b), d.h. es sei $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ eine Basis von V und es sei $(\varphi(s_1), \ldots, \varphi(s_r))$ eine Basis von $\operatorname{Im}\varphi$. Die lineare Unabhängigkeit von $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ in V impliziert dann die lineare Unabhängigkeit von $(s'_1, \ldots, s'_d)$ in $\operatorname{Ker}\varphi$. Um zu zeigen, dass $(s'_1, \ldots, s'_d)$ ein Erzeugendensystem von $\operatorname{Ker}\varphi$ ist, sei $v \in \operatorname{Ker}\varphi$ gegeben. Dann ist insbesondere $v \in V$ und da $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ ein Erzeugendensystem von V ist, gibt es $a \in K^r$ und $a' \in K^d$ mit $v = \sum_{i \in [1,r]} a_i s_i + \sum_{i \in [1,d]} a'_i s'_i$. Da $\operatorname{Ker}\varphi$

nach Bemerkung (2.12)(b) ein Untervektorraum von V ist, folgt aus $v \in \operatorname{Ker} \varphi$ und $s'_i \in \operatorname{Ker} \varphi$ für $i \in [1, d]$ auch $\sum_{i \in [1,r]} a_i s_i = v - \sum_{i \in [1,d]} a'_i s'_i \in \operatorname{Ker} \varphi$. Folglich ist

$$0 = \varphi(\sum_{i \in [1,r]} a_i s_i) = \sum_{i \in [1,r]} a_i \, \varphi(s_i).$$

Die lineare Unabhängigkeit von $(\varphi(s_1), \ldots, \varphi(s_r))$ impliziert $a = 0$, es gilt also $v = \sum_{i \in [1,d]} a'_i s'_i \in \langle s'_1, \ldots, s'_d \rangle$. Somit ist $(s'_1, \ldots, s'_d)$ ein Erzeugendensystem von $\operatorname{Ker} \varphi$. Insgesamt ist $(s'_1, \ldots, s'_d)$ eine Basis von $\operatorname{Ker} \varphi$, d.h. es gilt Bedingung (c).

Als nächstes gelte Bedingung (a) und Bedingung (c), d.h. es sei $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ eine Basis von V und es sei $(s'_1, \ldots, s'_d)$ eine Basis von $\operatorname{Ker} \varphi$. Nach Proposition (2.21)(a) impliziert die Surjektivität der Restriktion $\varphi|^{\operatorname{Im} \varphi} \colon V \to \operatorname{Im} \varphi$, $v \mapsto \varphi(v)$, dass $(\varphi(s_1), \ldots, \varphi(s_r), \varphi(s'_1), \ldots, \varphi(s'_d))$ ein Erzeugendensystem von $\operatorname{Im} \varphi$ ist. Wegen $\varphi(s'_i) = 0$ für $i \in [1, d]$ ist nach Proposition (1.22) dann aber sogar $(\varphi(s_1), \ldots, \varphi(s_r))$ ein Erzeugendensystem von $\operatorname{Im} \varphi$. Um zu zeigen, dass $(\varphi(s_1), \ldots, \varphi(s_r))$ auch linear unabhängig in $\operatorname{Im} \varphi$ ist, sei $a \in K^r$ mit $\sum_{i \in [1,r]} a_i \, \varphi(s_i) = 0$ gegeben. Da φ ein Homomorphismus ist, erhalten wir

$$\varphi(\sum_{i \in [1,r]} a_i s_i) = \sum_{i \in [1,r]} a_i \, \varphi(s_i) = 0,$$

d.h. es ist $\sum_{i \in [1,r]} a_i s_i \in \operatorname{Ker} \varphi$. Da $(s'_1, \ldots, s'_d)$ ein Erzeugendensystem von $\operatorname{Ker} \varphi$ ist, gibt es ein $a' \in K^d$ mit $\sum_{i \in [1,r]} a_i s_i = \sum_{i \in [1,d]} a'_i s'_i$. Es folgt

$$0 = \sum_{i \in [1,r]} a_i s_i - \sum_{i \in [1,d]} a'_i s'_i = \sum_{i \in [1,r]} a_i s_i + \sum_{i \in [1,d]} (-a'_i) s'_i$$

und auf Grund der linearen Unabhängigkeit von $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ somit $a = 0$ und $-a' = 0$. Nach dem Kriterium für lineare Unabhhängigkeit (1.34) ist $(\varphi(s_1), \ldots, \varphi(s_r))$ linear unabhängig. Insgesamt ist $(\varphi(s_1), \ldots, \varphi(s_r))$ ein linear unabhängiges Erzeugendensystem und damit eine Basis von $\operatorname{Im} \varphi$, d.h. es gilt Bedingung (b).

Schließlich gelte Bedingung (b) und Bedingung (c), d.h. es sei $(\varphi(s_1), \ldots, \varphi(s_r))$ eine Basis von $\operatorname{Im} \varphi$ und es sei $(s'_1, \ldots, s'_d)$ eine Basis von $\operatorname{Ker} \varphi$. Um zu zeigen, dass $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ linear unabhängig ist, seien $a \in K^r$, $a' \in K^d$ mit $\sum_{i \in [1,r]} a_i s_i + \sum_{i \in [1,d]} a'_i s'_i = 0$ gegeben. Da φ ein Homomorphismus ist, folgt

$$0 = \varphi(0) = \varphi(\sum_{i \in [1,r]} a_i s_i + \sum_{i \in [1,d]} a'_i s'_i) = \sum_{i \in [1,r]} a_i \, \varphi(s_i) + \sum_{i \in [1,d]} a'_i \, \varphi(s'_i) = \sum_{i \in [1,r]} a_i \, \varphi(s_i)$$

und auf Grund der linearen Unabhängigkeit von $(\varphi(s_1), \ldots, \varphi(s_r))$ in W somit $a = 0$. Dies impliziert aber nun

$$\sum_{i \in [1,d]} a'_i s'_i = \sum_{i \in [1,r]} a_i s_i + \sum_{i \in [1,d]} a'_i s'_i = 0,$$

so dass auf Grund der linearen Unabhängigkeit von $(s'_1, \ldots, s'_d)$ in V auch $a' = 0$ folgt. Nach dem Kriterium für lineare Unabhhängigkeit (1.34) ist $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ linear unabhängig in V. Um zu zeigen, dass $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ ein Erzeugendensystem von V ist, sei $v \in V$ gegeben. Da $(\varphi(s_1), \ldots, \varphi(s_r))$ ein Erzeugendensystem von $\operatorname{Im} \varphi$ ist, gibt es ein $a \in K^r$ mit $\varphi(v) = \sum_{i \in [1,r]} a_i \, \varphi(s_i)$. Nun ist φ ein Homomorphismus, es folgt also

$$\varphi(v - \sum_{i \in [1,r]} a_i s_i) = \varphi(v) - \sum_{i \in [1,r]} a_i \, \varphi(s_i) = 0$$

und damit $v - \sum_{i \in [1,r]} a_i s_i \in \operatorname{Ker} \varphi$. Da außerdem $(s'_1, \ldots, s'_d)$ ein Erzeugendensystem von $\operatorname{Ker} \varphi$ ist, gibt es ein $a' \in K^d$ mit $v - \sum_{i \in [1,r]} a_i s_i = \sum_{i \in [1,d]} a'_i s'_i$. Dies impliziert schließlich

$$v = \sum_{i \in [1,r]} a_i s_i + \sum_{i \in [1,d]} a'_i s'_i \in \langle s_1, \ldots, s_r, s'_1, \ldots, s'_d \rangle.$$

Folglich ist $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ ein Erzeugendensystem von V. Insgesamt ist $(s_1, \ldots, s_r, s'_1, \ldots, s'_d)$ eine Basis von V, d.h. es gilt Bedingung (a). □

(2.28) Korollar (Rangsatz)**.** Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Genau dann ist V endlichdimensional, wenn $\operatorname{Im}\varphi$ und $\operatorname{Ker}\varphi$ endlichdimensional sind. In diesem Fall gibt es $r, d \in \mathbb{N}_0$, ein r-Tupel $(s_1, \dots, s_r)$ und ein d-Tupel $(s'_1, \dots, s'_d)$ in V derart, dass $(s_1, \dots, s_r, s'_1, \dots, s'_d)$ eine Basis von V und $(\varphi(s_1), \dots, \varphi(s_r))$ eine Basis von $\operatorname{Im}\varphi$ und $(s'_1, \dots, s'_d)$ eine Basis von $\operatorname{Ker}\varphi$ ist. Insbesondere gilt

$$\dim_K V = \operatorname{rk}_K \varphi + \operatorname{def}_K \varphi.$$

Beweis. Zunächst sei V endlichdimensional. Nach Bemerkung (2.24)(b) ist dann auch $\operatorname{Ker}\varphi$ endlichdimensional, d.h. es gibt ein $d \in \mathbb{N}_0$ und eine Basis $(s'_1, \dots, s'_d)$ von $\operatorname{Ker}\varphi$. Ferner gibt es nach Korollar (1.55) ein $r \in \mathbb{N}_0$ und ein r-Tupel $(s_1, \dots, s_r)$ in V so, dass $(s_1, \dots, s_r, s'_1, \dots, s'_d)$ eine Basis von V ist. Nach Proposition (2.27) ist $(\varphi(s_1), \dots, \varphi(s_r))$ eine Basis von $\operatorname{Im}\varphi$ und damit $\operatorname{Im}\varphi$ endlichdimensional. Es folgt

$$\dim V = r + d = \dim(\operatorname{Im}\varphi) + \dim(\operatorname{Ker}\varphi) = \operatorname{rk}\varphi + \operatorname{def}\varphi.$$

Umgekehrt seien $\operatorname{Im}\varphi$ und $\operatorname{Ker}\varphi$ endlichdimensional. Da $\operatorname{Im}\varphi$ endlichdimensional ist, gibt es ein $r \in \mathbb{N}_0$ und ein r-Tupel $(s_1, \dots, s_r)$ in V derart, dass $(\varphi(s_1), \dots, \varphi(s_r))$ eine Basis von $\operatorname{Im}\varphi$ ist, und da $\operatorname{Ker}\varphi$ endlichdimensional ist, gibt es ein $d \in \mathbb{N}_0$ und eine Basis $(s'_1, \dots, s'_d)$ von $\operatorname{Ker}\varphi$. Nach Proposition (2.27) ist dann aber $(s_1, \dots, s_r, s'_1, \dots, s'_d)$ eine Basis von V und damit V endlichdimensional. □

(2.29) Korollar. Es sei ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ so gegeben, dass V und W endlichdimensional sind und $\dim_K V = \dim_K W$ gilt. Die folgenden Bedingungen sind äquivalent.

(a) Es ist φ ein K-Vektorraumisomorphismus.

(b) Es ist φ injektiv.

(c) Es ist φ surjektiv.

(d) Es ist $\operatorname{def}_K \varphi = 0$.

(e) Es ist $\operatorname{rk}_K \varphi = \dim_K W$.

Beweis. Nach dem Rangsatz (2.28) gilt

$$\dim W = \dim V = \operatorname{def}\varphi + \operatorname{rk}\varphi.$$

Folglich gilt genau dann $\operatorname{def}\varphi = 0$, wenn $\operatorname{rk}\varphi = \dim W$ ist, d.h. Bedingung (d) und Bedingung (e) sind äquivalent.
Ferner ist $\varphi\colon V \to W$ nach Korollar (2.26) genau dann ein Isomorphismus, wenn $\operatorname{def}\varphi = 0$ und $\operatorname{rk}\varphi = \dim W$ ist. Dies zeigt die Äquivalenz von Bedingung (a), Bedingung (d) und Bedingung (e).
Schließlich ist Bedingung (c) äquivalent zu Bedingung (e) nach Bemerkung (2.25)(a), und es ist Bedingung (b) äquivalent zu Bedingung (d) nach Bemerkung (2.25)(b).
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c), Bedingung (d) und Bedingung (e) äquivalent. □

(2.30) Proposition. Es seien K-Vektorraumhomomorphismen $\varphi\colon V \to W$ und $\psi\colon W \to X$ so gegeben, dass $\operatorname{Im}\varphi$ und $\operatorname{Im}\psi$ endlichdimensional sind. Dann ist auch $\operatorname{Im}(\psi \circ \varphi)$ endlichdimensional und es gilt

$$\operatorname{rk}_K(\psi \circ \varphi) \leq \min(\operatorname{rk}_K \psi, \operatorname{rk}_K \varphi).$$

Beweis. Es ist $\operatorname{Im}(\psi \circ \varphi)$ ein Untervektorraum von $\operatorname{Im}\psi$. Nach Proposition (1.68) ist mit $\operatorname{Im}\psi$ folglich auch $\operatorname{Im}(\psi \circ \varphi)$ endlichdimensional und es gilt

$$\operatorname{rk}(\psi \circ \varphi) = \dim(\operatorname{Im}(\psi \circ \varphi)) \leq \dim(\operatorname{Im}\psi) = \operatorname{rk}\psi.$$

Andererseits ist $\operatorname{Im}(\psi \circ \varphi) = \operatorname{Im}\psi|_{\operatorname{Im}\varphi}$. Da $\operatorname{Im}\varphi$ endlichdimensional ist, folgt nach dem Rangsatz (2.28) auch

$$\operatorname{rk}(\psi \circ \varphi) = \dim(\operatorname{Im}(\psi \circ \varphi)) = \dim(\operatorname{Im}\psi|_{\operatorname{Im}\varphi}) = \operatorname{rk}\psi|_{\operatorname{Im}\varphi} \leq \dim(\operatorname{Im}\varphi) = \operatorname{rk}\varphi.$$

Insgesamt haben wir $\operatorname{rk}(\psi \circ \varphi) \leq \min(\operatorname{rk}\psi, \operatorname{rk}\varphi)$. □

Darstellung von endlichdimensionalen Untervektorräumen als Bild und Kern

Nach Bemerkung (2.12) sind Bild und Kern eines Vektorraumhomomorphismus stets Untervektorräume (des Ziel- bzw. Startvektorraums des Homomorphismus). Wir wollen der Frage nachgehen, ob sich jeder endlichdimensionale Untervektorraum eines (endlichdimensionalen) Vektorraums als Bild bzw. Kern von geeigneten Vektorraumhomomorphismen darstellen lässt.
Hierzu seien ein K-Vektorraum V, ein K-Untervektorraum U von V, $m \in \mathbb{N}_0$ und eine Basis $s = (s_1, \ldots, s_m)$ von U gegeben. Nach Bemerkung (2.18) und Bemerkung (1.33) ist dann

$$\lambda_s \colon K^m \to V,\ a \mapsto \sum_{i \in [1,m]} a_i s_i$$

ein injektiver K-Vektorraumhomomorphismus mit

$$\operatorname{Im} \lambda_s = \langle s_1, \ldots, s_m \rangle = U.$$

Um zu zeigen, dass sich U auch als Kern eines geeigneten Homomorphismus darstellen lässt, sei weiter angenommen, dass V endlichdimensional ist. Nach Korollar (1.55) gibt es dann ein $n \in \mathbb{N}_0$ und ein n-Tupel $(s'_1, \ldots, s'_n)$ in V derart, dass $(s_1, \ldots, s_m, s'_1, \ldots, s'_n)$ eine Basis von V ist. Ferner gibt es nach Proposition (2.19)(c) genau einen K-Vektorraumhomomorphismus $\varphi \colon V \to K^n$ mit $\varphi(s_i) = 0$ für $i \in [1, m]$ und $\varphi(s'_i) = \mathrm{e}_i$ für $i \in [1, n]$, und dieser Homomorphismus ist surjektiv nach Proposition (2.21)(a). Da $(s_1, \ldots, s_m, s'_1, \ldots, s'_n)$ eine Basis von V und $(\varphi(s'_1), \ldots, \varphi(s'_n)) = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ eine Basis von K^n ist, ist nach Proposition (2.27) folglich $(s_1, \ldots, s_m)$ eine Basis von $\operatorname{Ker} \varphi$ und damit

$$\operatorname{Ker} \varphi = \langle s_1, \ldots, s_m \rangle = U.$$

(2.31) Beispiel. Es sei der $\mathbb{Q}$-Untervektorraum

$$U = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle$$

von $\mathbb{Q}^{4 \times 1}$ gegeben.

(a) Die Abbildung

$$\lambda \colon \mathbb{Q}^3 \to \mathbb{Q}^{4 \times 1},\ (a, b, c) \mapsto \begin{pmatrix} a \\ b \\ a + 2b \\ c \end{pmatrix}$$

ist ein $\mathbb{Q}$-Vektorraumhomomorphismus mit $\operatorname{Im} \lambda = U$.

(b) Die Abbildung

$$\varphi \colon \mathbb{Q}^{4 \times 1} \to \mathbb{Q}^1,\ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \begin{pmatrix} -a - 2b + c \end{pmatrix}$$

ist ein $\mathbb{Q}$-Vektorraumhomomorphismus mit $\operatorname{Ker} \varphi = U$.

Beweis.

(a) Nach Bemerkung (2.18) und Bemerkung (1.33) ist

$$\lambda \colon \mathbb{Q}^3 \to \mathbb{Q}^{4 \times 1},\ (a, b, c) \mapsto a \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

ein injektiver Homomorphismus mit

$$\operatorname{Im}\lambda = \langle \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \rangle = U.$$

Für $a, b, c \in \mathbb{Q}$ gilt

$$\lambda(a, b, c) = a \begin{pmatrix}1\\0\\1\\0\end{pmatrix} + b \begin{pmatrix}0\\1\\2\\0\end{pmatrix} + c \begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}a\\b\\a+2b\\c\end{pmatrix}.$$

(b) Da

$$(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix})$$

eine Basis von $\mathbb{Q}^{4\times 1}$ ist, gibt es nach Proposition (2.19)(c) genau einen Homomorphismus $\varphi\colon \mathbb{Q}^{4\times 1} \to \mathbb{Q}^1$ mit

$$\varphi(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}) = 0,\ \varphi(\begin{pmatrix}0\\1\\2\\0\end{pmatrix}) = 0,\ \varphi(\begin{pmatrix}0\\0\\0\\1\end{pmatrix}) = 0,\ \varphi(\begin{pmatrix}0\\0\\1\\0\end{pmatrix}) = \mathrm{e}_1,$$

und da

$$(\varphi(\begin{pmatrix}0\\0\\1\\0\end{pmatrix})) = (\mathrm{e}_1)$$

eine Basis von $\mathbb{Q}^1$ ist, ist φ surjektiv nach Proposition (2.21)(a). Nach Proposition (2.27) ist ferner

$$\operatorname{Ker}\varphi = \langle \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \rangle = U.$$

Für $a, b, c, d \in \mathbb{Q}$ gilt

$$\begin{aligned}\varphi(\begin{pmatrix}a\\b\\c\\d\end{pmatrix}) &= \varphi(a \begin{pmatrix}1\\0\\1\\0\end{pmatrix} + b \begin{pmatrix}0\\1\\2\\0\end{pmatrix} + d \begin{pmatrix}0\\0\\0\\1\end{pmatrix} + (-a-2b+c) \begin{pmatrix}0\\0\\1\\0\end{pmatrix}) \\ &= a\,\varphi(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}) + b\,\varphi(\begin{pmatrix}0\\1\\2\\0\end{pmatrix}) + d\,\varphi(\begin{pmatrix}0\\0\\0\\1\end{pmatrix}) + (-a-2b+c)\,\varphi(\begin{pmatrix}0\\0\\1\\0\end{pmatrix}) \\ &= a \cdot 0 + b \cdot 0 + d \cdot 0 + (-a-2b+c)\mathrm{e}_1 = (-a-2b+c).\end{aligned}$$

□

Klassifikation endlichdimensionaler Vektorräume

In Abschnitt 1 haben wir Vektorräume als Abstraktion der konkreten Beispiele $\mathbb{R}^2$ und $\mathbb{R}^3$ eingeführt. Nun werden wir sehen, dass zumindest alle endlichdimensionalen K-Vektorräume „im Wesentlichen gleich“ dem K-Vektorraum K^n für ein $n \in \mathbb{N}_0$ sind. Hierbei werden wir sehr präzise sagen, was „im Wesentlichen gleich“ bedeutet, wir werden nämlich den in Definition (2.7)(b) eingeführten Isomorphiebegriff verwenden. Diese Isomorphie können wir durch Angabe eines konkreten Isomorphismus realisieren: Für jedes $n \in \mathbb{N}_0$ und jeden n-dimensionalen K-Vektorraum V erhalten wir einen solchen Isomorphismus durch Wahl einer Basis $s = (s_1, \ldots, s_n)$ von V, denn die Abbildung

$$\lambda_s \colon K^n \to V,\, a \mapsto \sum_{i \in [1,n]} a_i s_i$$

ist dann nach Bemerkung (2.18) und Bemerkung (2.9) ein Isomorphismus von K-Vektorräumen.
Wir werden aber sogar noch einen Schritt weitergehen und die Isomorphie von endlichdimensionalen Vektorräumen mit Hilfe der Dimension charakterisieren:

(2.32) Satz. Es seien endlichdimensionale K-Vektorräume V und W gegeben. Genau dann ist

$$V \cong W,$$

wenn

$$\dim_K V = \dim_K W$$

ist.

Beweis. Zunächst gelte $V \cong W$, so dass es einen Isomorphismus $\varphi\colon V \to W$ gibt. Da V endlichdimensional ist, gibt es nach Korollar (1.52) ein $n \in \mathbb{N}_0$ und eine Basis $(s_1, \ldots, s_n)$ von V. Nach Proposition (2.21)(c) ist dann aber $(\varphi(s_1), \ldots, \varphi(s_n))$ eine Basis von W. Insbesondere gilt $\dim V = n = \dim W$.
Nun gelte umgekehrt $\dim V = \dim W$. Da V und W endlichdimensional sind, gibt es nach Korollar (1.52) ein $m \in \mathbb{N}_0$ und eine Basis $(s_1, \ldots, s_m)$ von V sowie ein $n \in \mathbb{N}_0$ und eine Basis $(t_1, \ldots, t_n)$ von W. Es gilt

$$m = \dim V = \dim W = n.$$

Nach Proposition (2.19)(c) gibt es genau einen Homomorphismus $\varphi\colon V \to W$ mit $\varphi(s_i) = t_i$ für alle $i \in [1, n]$, und nach Proposition (2.21)(c) ist φ ein Isomorphismus. Insbesondere gilt $V \cong W$. □

(2.33) Korollar (Klassifikation endlichdimensionaler Vektorräume)**.** Es seien $n \in \mathbb{N}_0$ und ein endlichdimensionaler K-Vektorraum V gegeben. Genau dann ist

$$\dim_K V = n,$$

wenn

$$V \cong K^n$$

ist.

Beweis. Nach Beispiel (1.65)(a) ist $\dim K^n = n$. Somit gilt genau dann $\dim V = n$, wenn $\dim V = \dim K^n$ ist, was nach Satz (2.32) aber äquivalent zu $V \cong K^n$ ist. □

(2.34) Beispiel. Für $m, n \in \mathbb{N}_0$ ist $K^{m \times n} \cong K^{mn}$.

Beweis. Für $m, n \in \mathbb{N}_0$ ist $\dim K^{m \times n} = mn$ nach Beispiel (1.65)(b), so dass aus Korollar (2.33) bereits $K^{m \times n} \cong K^{mn}$ folgt. □

Vektorraum der Homomorphismen

Zum Schluss dieses Abschnitts zeigen wir, dass die Menge der Vektorraumhomomorphismen zwischen zwei Vektorräumen ebenfalls zu einem Vektorraum wird und dass diese Vektorraumstruktur verträglich mit der Komposition von Vektorraumhomomorphismen ist.

(2.35) Bemerkung. Es seien ein K-Vektorraum V und eine Menge X gegeben. Die Menge $\mathrm{Map}(X, V)$ wird ein K-Vektorraum mit Addition gegeben durch

$$(f + g)(x) = f(x) + g(x)$$

für $x \in X$, $f, g \in \mathrm{Map}(X, V)$ und Skalarmultiplikation gegeben durch

$$(af)(x) = af(x)$$

für $x \in X$, $a \in K$, $f \in \mathrm{Map}(X, V)$.

Beweis. Dies lässt sich analog zu Beispiel (1.4)(d) beweisen. Die Details seien dem Leser zur Übung überlassen. □

(2.36) Proposition. Es seien ein Körper K und K-Vektorräume V und W gegeben. Dann ist $\mathrm{Hom}_K(V, W)$ ein K-Untervektorraum von $\mathrm{Map}(V, W)$.

Beweis. Für $\varphi, \psi \in \mathrm{Hom}_K(V, W)$ gilt

$$\begin{aligned}(\varphi + \psi)(v + v') &= \varphi(v + v') + \psi(v + v') = \varphi(v) + \varphi(v') + \psi(v) + \psi(v') = \varphi(v) + \psi(v) + \varphi(v') + \psi(v') \\ &= (\varphi + \psi)(v) + (\varphi + \psi)(v'), \\ (\varphi + \psi)(bv) &= \varphi(bv) + \psi(bv) = b\,\varphi(v) + b\,\psi(v) = b\,(\varphi + \psi)(v)\end{aligned}$$

für alle $b \in K$, $v, v' \in V$ und damit $\varphi + \psi \in \mathrm{Hom}_K(V, W)$ nach dem Kriterium für Vektorraumhomomorphismen (2.2). Ferner ist

$$\begin{aligned}0(v + v') &= 0 = 0 + 0 = 0(v) + 0(v'), \\ 0(bv) &= 0 = b \cdot 0 = b\,0(v)\end{aligned}$$

für alle $b \in K$, $v, v' \in V$ und damit $0 \in \mathrm{Hom}_K(V, W)$ nach dem Kriterium für Vektorraumhomomorphismen (2.2). Für $a \in K$, $\varphi \in \mathrm{Hom}_K(V, W)$ gilt schließlich

$$\begin{aligned}(a\varphi)(v + v') &= a\varphi(v + v') = a(\varphi(v) + \varphi(v')) = a\,\varphi(v) + a\,\varphi(v') = (a\varphi)(v) + (a\varphi)(v'), \\ (a\varphi)(bv) &= a\,\varphi(bv) = ab\,\varphi(v) = ba\,\varphi(v) = b\,(a\varphi)(v)\end{aligned}$$

für alle $b \in K$, $v, v' \in V$ und damit $a\varphi \in \mathrm{Hom}_K(V, W)$ nach dem Kriterium für Vektorraumhomomorphismen (2.2).
Nach dem Untervektorraumkriterium (1.15) ist $\mathrm{Hom}_K(V, W)$ ein Untervektorraum von $\mathrm{Map}(V, W)$. □

(2.37) Proposition. Es seien K-Vektorräume V, W, X gegeben.

(a) Für $\varphi \in \mathrm{Hom}(V, W)$, $\psi, \psi' \in \mathrm{Hom}(W, X)$ gilt

$$(\psi + \psi') \circ \varphi = \psi \circ \varphi + \psi' \circ \varphi.$$

Für $\varphi, \varphi' \in \mathrm{Hom}(V, W)$, $\psi \in \mathrm{Hom}(W, X)$ gilt

$$\psi \circ (\varphi + \varphi') = \psi \circ \varphi + \psi \circ \varphi'.$$

(b) Für $a \in K$, $\varphi \in \mathrm{Hom}(V, W)$, $\psi \in \mathrm{Hom}(W, X)$ gilt

$$(a\psi) \circ \varphi = \psi \circ (a\varphi) = a(\psi \circ \varphi).$$

Beweis.

(a) Für $\varphi \in \mathrm{Hom}(V, W)$, $\psi, \psi' \in \mathrm{Hom}(W, X)$ gilt

$$((\psi + \psi') \circ \varphi)(v) = (\psi + \psi')(\varphi(v)) = \psi(\varphi(v)) + \psi'(\varphi(v)) = (\psi \circ \varphi + \psi' \circ \varphi)(v)$$

für alle $v \in V$ und damit $(\psi + \psi') \circ \varphi = \psi \circ \varphi + \psi' \circ \varphi$. Für $\varphi, \varphi' \in \mathrm{Hom}(V, W)$, $\psi \in \mathrm{Hom}(W, X)$ gilt

$$(\psi \circ (\varphi + \varphi'))(v) = \psi((\varphi + \varphi')(v)) = \psi(\varphi(v) + \varphi'(v)) = \psi(\varphi(v)) + \psi(\varphi'(v)) = (\psi \circ \varphi + \psi \circ \varphi')(v)$$

für alle $v \in V$ und damit $\psi \circ (\varphi + \varphi') = \psi \circ \varphi + \psi \circ \varphi'$.

(b) Für $a \in K$, $\varphi \in \operatorname{Hom}(V, W)$, $\psi \in \operatorname{Hom}(W, X)$ gilt

$$((a\psi) \circ \varphi)(v) = (a\psi)(\varphi(v)) = a\,\psi(\varphi(v)) = (a(\psi \circ \varphi))(v),$$
$$(\psi \circ (a\varphi))(v) = \psi((a\varphi)(v)) = \psi(a\varphi(v)) = a\,\psi(\varphi(v)) = (a(\psi \circ \varphi))(v)$$

für alle $v \in V$ und damit $(a\psi) \circ \varphi = \psi \circ (a\varphi) = a(\psi \circ \varphi)$. □

(2.38) Korollar. Es sei ein K-Vektorraum V gegeben. Die abelsche Gruppe $\operatorname{End}_K(V)$ wird ein Ring mit Multiplikation gegeben durch Komposition. Die Eins von $\operatorname{End}_K(V)$ ist id_V.

3 Matrixkalkül

In Korollar (2.33) haben wir gesehen, dass es über einem gegebenen Körper für jedes $n \in \mathbb{N}_0$ bis auf Isomorphie genau einen n-dimensionalen Vektorraum gibt. Diese Tatsache wollen wir im Folgenden ausnutzen, um einen Kalkül zum Rechnen mit abstrakten Vektorräumen und abstrakten Vektorraumhomomorphismen zu entwickeln. Hierzu werden wir Vektoren als Spalten und Vektorraumhomomorphismen durch beliebige Matrizen kodieren, vgl. Definition (3.6) und Definition (3.8). Die Anwendung eines Homomorphismus auf einen Vektor wird dann einer Multiplikation einer Matrix an eine Spalte entsprechen, vgl. Proposition (3.11), während sich die Komposition von Homomorphismen in eine Matrixmultiplikation übersetzt, vgl. Proposition (3.13)(a).
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein Körper K gegeben.

Spalteninterpretation

Wir beginnen zunächst mit der einfacheren, umgekehrten Richtung und zeigen, dass jede Matrix durch geeignete Interpretation einen Vektorraumhomomorphismus liefert:

(3.1) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Die Abbildung

$$\varphi_A \colon K^{n \times 1} \to K^{m \times 1},\ x \mapsto Ax$$

ist ein K-Vektorraumhomomorphismus.

Beweis. Für $x, x' \in K^{n \times 1}$ ist

$$\varphi_A(x + x') = A(x + x') = Ax + Ax' = \varphi_A(x) + \varphi_A(x').$$

Für $a \in K$, $x \in K^{n \times 1}$ ist

$$\varphi_A(ax) = A(ax) = aAx = a\,\varphi_A(x).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ_A ein Homomorphismus. □

(3.2) Definition (Spalteninterpretation)**.** Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $A \in K^{m \times n}$ heißt

$$\varphi_A \colon K^{n \times 1} \to K^{m \times 1},\ x \mapsto Ax$$

die *Spalteninterpretation* von A.

(3.3) Beispiel. Die Abbildung

$$\varphi \colon \mathbb{Q}^{3 \times 1} \to \mathbb{Q}^{3 \times 1},\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} -2x - y + 4z \\ -3x - y + 5z \\ 2x + 2y - 6z \end{pmatrix}$$

ist ein $\mathbb{Q}$-Vektorraumhomomorphismus.

Beweis. Es sei $A \in \mathbb{Q}^{3 \times 3}$ gegeben durch

$$A = \begin{pmatrix} -2 & -1 & 4 \\ -3 & -1 & 5 \\ 2 & 2 & -6 \end{pmatrix}.$$

Für $x, y, z \in \mathbb{Q}$ gilt dann

$$\varphi(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} -2x - y + 4z \\ -3x - y + 5z \\ 2x + 2y - 6z \end{pmatrix} = \begin{pmatrix} -2 & -1 & 4 \\ -3 & -1 & 5 \\ 2 & 2 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \varphi_A(\begin{pmatrix} x \\ y \\ z \end{pmatrix}),$$

d.h. es ist $\varphi = \varphi_A$. Nach Bemerkung (3.1) ist somit $\varphi = \varphi_A$ ein $\mathbb{Q}$-Vektorraumhomomorphismus. □

(3.4) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Für $x \in K^{n \times 1}$ ist

$$\varphi_A(x) = \sum_{j \in [1,n]} x_j A_{-,j}.$$

Beweis. Für $x \in K^{n \times 1}$ ist

$$\varphi_A(x) = Ax = (\sum_{j \in [1,n]} A_{i,j} x_j)_{i \in [1,m]} = \sum_{j \in [1,n]} x_j (A_{i,j})_{i \in [1,m]} = \sum_{j \in [1,n]} x_j A_{-,j}.$$ □

(3.5) Bemerkung.

(a) Für $m, n, p \in \mathbb{N}_0$, $A \in K^{n \times p}$, $B \in K^{m \times n}$ gilt

$$\varphi_{BA} = \varphi_B \circ \varphi_A.$$

(b) Für $n \in \mathbb{N}_0$ gilt

$$\varphi_{\mathrm{E}_n} = \mathrm{id}_{K^{n \times 1}}.$$

(c) Für $n \in \mathbb{N}_0$, $A \in \mathrm{GL}_n(K)$ ist $\varphi_A \colon K^{n \times 1} \to K^{n \times 1}$ ein K-Vektorraumautomorphismus mit

$$\varphi_A^{-1} = \varphi_{A^{-1}}.$$

(d) Für $m, n \in \mathbb{N}_0$ ist

$$\varphi_- \colon K^{m \times n} \to \mathrm{Hom}_K(K^{n \times 1}, K^{m \times 1}),\ A \mapsto \varphi_A$$

ein K-Vektorraumhomomorphismus.

Beweis.

(a) Da für $x \in K^{p \times 1}$ stets

$$\varphi_{BA}(x) = BAx = \varphi_B(\varphi_A(x)) = (\varphi_B \circ \varphi_A)(x)$$

ist, gilt $\varphi_{BA} = \varphi_B \circ \varphi_A$.

(b) Da für $x \in K^{n \times 1}$ stets

$$\varphi_{\mathrm{E}_n}(x) = \mathrm{E}_n x = x = \mathrm{id}_{K^{n \times 1}}(x)$$

ist, gilt $\varphi_{\mathrm{E}_n} = \mathrm{id}_{K^{n \times 1}}$.

(c) Nach (a), (b) gilt

$$\varphi_{A^{-1}} \circ \varphi_A = \varphi_{A^{-1}A} = \varphi_{\mathrm{E}_n} = \mathrm{id}_{K^{n \times 1}},$$
$$\varphi_A \circ \varphi_{A^{-1}} = \varphi_{AA^{-1}} = \varphi_{\mathrm{E}_n} = \mathrm{id}_{K^{n \times 1}}.$$

Folglich ist φ_A eine invertierbare Abbildung mit $\varphi_A^{-1} = \varphi_{A^{-1}}$. Ferner ist φ_A ein K-Vektorraumhomomorphismus nach Bemerkung (3.1) und damit ein K-Vektorraumisomorphismus nach Bemerkung (2.9).

(d) Für $A, A' \in K^{m \times n}$ gilt

$$\varphi_{A+A'}(x) = (A + A')x = Ax + A'x = \varphi_A(x) + \varphi_{A'}(x) = (\varphi_A + \varphi_{A'})(x)$$

für alle $x \in K^{n \times 1}$ und damit $\varphi_{A+A'} = \varphi_A + \varphi_{A'}$. Für $c \in K$, $A \in K^{m \times n}$ gilt

$$\varphi_{cA}(x) = cAx = c\,\varphi_A(x) = (c\,\varphi_A)(x)$$

für alle $x \in K^{n \times 1}$ und damit $\varphi_{cA} = c\,\varphi_A$. Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ_- ein Homomorphismus. □

Koordinatenspalten und Darstellungsmatrizen

Nachdem wir gesehen haben, dass jede Matrix sich als Vektorraumhomomorphismus interpretieren lässt, wollen wir nun umgekehrt zeigen, dass sich Matrizen als Hilfsmittel zum Rechnen mit allgemeinen Vektorraumhomomorphismen zwischen endlichdimensionalen Vektorräumen nutzen lassen.
Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \ldots, s_n)$ von V gegeben. Nach Bemerkung (2.18) ist die Abbildung

$$\lambda_s \colon K^n \to V,\ a \mapsto \sum_{i \in [1,n]} a_i s_i$$

ein K-Vektorraumhomomorphismus. Ferner ist λ_s bijektiv, da s eine Basis ist. Nach Bemerkung (2.9) ist λ_s also ein K-Vektorraumisomorphismus. Um einen Vektor v des abstrakten Vektorraums V durch einen Vektor des konkreten Vektorraums K^n darzustellen, können wir den inversen Vektorraumisomorphismus $\lambda_s^{-1} \colon V \to K^n$ auf v anwenden; für $v \in V$ ist $\lambda_s^{-1}(v) \in K^n$. Um $\lambda_s^{-1}(v)$ für ein $v \in V$ zu berechnen, müssen wir das eindeutige $a \in K^n$ mit $v = \lambda_s(a) = \sum_{i \in [1,n]} a_i s_i$ bestimmen, denn dann ist $\lambda_s^{-1}(v) = a$.
Durch die Verwendung dieser sich gegenseitig invertierenden Isomorphismen erhalten wir nicht nur eine Korrespondenz zwischen den Vektoren von V und denen von K^n, auch die Art und Weise, wie wir mit den jeweiligen Vektoren rechnen (addieren und mit Skalaren multiplizieren) entspricht sich auf Grund der Homomorphieeigenschaft von λ_s und λ_s^{-1}. Dabei hängt die Art der Darstellung von der Basis s ab; für eine andere Basis $t = (t_1, \ldots, t_n)$ von V ist $\lambda_s^{-1} \neq \lambda_t^{-1}$, d.h. es gibt ein $v \in V$ mit $\lambda_s^{-1}(v) \neq \lambda_t^{-1}(v)$.
Die Vektoren in V lassen sich also bzgl. der Basis s durch konkrete n-Tupel kodieren, indem man von den Vektoren zu den eindeutigen Koeffizienten in der Darstellung als Linearkombinationen von s übergeht. Da wir auch Homomorphismen zwischen endlichdimensionalen Vektorräumen kodieren und insgesamt einen schönen Kalkül entwickeln wollen, bedienen wir uns eines kleinen Tricks und betrachten anstatt der n-Tupel Spaltenmatrizen mit n Einträgen. Präzise ausgedrückt bedeutet dies, dass wir anstatt des Isomorphismus

$$\lambda_s \colon K^n \to V,\ a \mapsto \sum_{i \in [1,n]} a_i s_i$$

den Isomorphismus

$$\lambda_s \circ \lambda_{\mathrm{e}}^{-1} \colon K^{n \times 1} \to V,\ a \mapsto \sum_{i \in [1,n]} a_i s_i$$

betrachten.

$$\begin{tikzcd}
K^n \arrow[r, "\lambda_s", "\cong"'] \arrow[d, "\lambda_{\mathrm{e}}"', "\cong"] & V \\
K^{n \times 1} \arrow[ur, "\cong"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{e}_j \arrow[r, mapsto, "\lambda_s"] \arrow[d, mapsto, "\lambda_{\mathrm{e}}"'] & s_j \\
\mathrm{e}_j \arrow[ur, mapsto] &
\end{tikzcd}$$

(3.6) Definition (Koordinatenspalte)**.** Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \ldots, s_n)$ von V gegeben. Wir bezeichnen mit

$$\kappa_s \colon V \to K^{n \times 1},\ v \mapsto \kappa_s(v)$$

den zu $K^{n \times 1} \to V$, $a \mapsto \sum_{i \in [1,n]} a_i s_i$ inversen K-Vektorraumisomorphismus. Für $v \in V$ heißt $\kappa_s(v)$ die *Koordinatenspalte* (oder der *Koordinatenvektor*) von v zur Basis s.

(3.7) Beispiel.

(a) Die Koordinatenspalte von $a \in K^2$ bzgl. $s = ((1,0),(1,1))$ ist

$$\kappa_s(a) = \begin{pmatrix} a_1 - a_2 \\ a_2 \end{pmatrix}.$$

(b) Es seien $A \in \mathbb{Q}^{2\times 2}$ und eine Basis $s = (s_1, s_2, s_3, s_4)$ von $\mathbb{Q}^{2\times 2}$ gegeben durch

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \; s = (\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

Die Koordinatenspalte von A bzgl. s ist

$$\kappa_s(A) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Beweis.

(a) Es ist

$$a = (a_1, a_2) = (a_1 - a_2)(1, 0) + a_2(1, 1)$$

und damit

$$\kappa_s(a) = \begin{pmatrix} a_1 - a_2 \\ a_2 \end{pmatrix}.$$

(b) Es ist

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = s_1 + s_2 + s_3 + s_4$$

und damit

$$\kappa_s(A) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

□

Als nächstes wollen wir Vektorraumhomomorphismen durch Matrizen kodieren. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Aus Proposition (2.19)(c) wissen wir, dass φ bereits eindeutig durch das n-Tupel $(\varphi(s_1), \dots, \varphi(s_n))$ festgelegt ist. Wie wir gerade gesehen haben, lässt sich jedes $w \in W$ durch seine Koordinatenspalte $\kappa_t(w)$ kodieren; insbesondere ist also $\varphi(s_j)$ für $j \in [1, n]$ durch $\kappa_t(\varphi(s_j))$ festgelegt. Insgesamt wird φ also durch das n-Tupel $(\kappa_t(\varphi(s_1)), \dots, \kappa_t(\varphi(s_n)))$ in $K^{m\times 1}$ bestimmt. Anstelle von Tupeln verwenden wir eine Matrixschreibweise:

(3.8) Definition (Darstellungsmatrix)**.** Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Die *Darstellungsmatrix* (oder *Abbildungsmatrix*) von φ zu den Basen s und t ist gegeben durch

$$\mathrm{M}_{t,s}(\varphi) = \begin{pmatrix} \kappa_t(\varphi(s_1)) & \dots & \kappa_t(\varphi(s_n)) \end{pmatrix}.$$

(3.9) Beispiel.

(a) Die Darstellungsmatrix von $\varphi\colon K^2 \to K^3$, $(x_1, x_2) \mapsto (x_1 + x_2, x_1 - x_2, x_2)$ bzgl. der Standardbasen e^{K^2} und e^{K^3} ist

$$\mathrm{M}_{\mathrm{e}^{K^3}, \mathrm{e}^{K^2}}(\varphi) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

(b) Die Darstellungsmatrix von $\varphi\colon K^{3 \times 2} \to K^{2 \times 3}$, $A \mapsto A^{\mathrm{tr}}$ bzgl. der Basen

$$\begin{aligned} s &= (\mathrm{e}_{1,1}, \mathrm{e}_{1,2}, \mathrm{e}_{2,1}, \mathrm{e}_{2,2}, \mathrm{e}_{3,1}, \mathrm{e}_{3,2}), \\ t &= (\mathrm{e}_{1,1}, \mathrm{e}_{1,2}, \mathrm{e}_{1,3}, \mathrm{e}_{2,1}, \mathrm{e}_{2,2}, \mathrm{e}_{2,3}) \end{aligned}$$

ist

$$\mathrm{M}_{t,s}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Beweis.

(a) Es ist

$$\begin{aligned} \varphi(\mathrm{e}_1) &= (1,1,0) = \mathrm{e}_1 + \mathrm{e}_2 = 1\mathrm{e}_1 + 1\mathrm{e}_2 + 0\mathrm{e}_3, \\ \varphi(\mathrm{e}_2) &= (1,-1,1) = \mathrm{e}_1 - \mathrm{e}_2 + \mathrm{e}_3 = 1\mathrm{e}_1 + (-1)\mathrm{e}_2 + 1\mathrm{e}_3 \end{aligned}$$

und damit

$$\mathrm{M}_{\mathrm{e}^{K^3}, \mathrm{e}^{K^2}}(\varphi) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

(b) Es ist

$$\begin{aligned} \varphi(s_1) &= \varphi(\mathrm{e}_{1,1}) = \mathrm{e}_{1,1}^{\mathrm{tr}} = \mathrm{e}_{1,1} = t_1, \\ \varphi(s_2) &= \varphi(\mathrm{e}_{1,2}) = \mathrm{e}_{1,2}^{\mathrm{tr}} = \mathrm{e}_{2,1} = t_4, \\ \varphi(s_3) &= \varphi(\mathrm{e}_{2,1}) = \mathrm{e}_{2,1}^{\mathrm{tr}} = \mathrm{e}_{1,2} = t_2, \\ \varphi(s_4) &= \varphi(\mathrm{e}_{2,2}) = \mathrm{e}_{2,2}^{\mathrm{tr}} = \mathrm{e}_{2,2} = t_5, \\ \varphi(s_5) &= \varphi(\mathrm{e}_{3,1}) = \mathrm{e}_{3,1}^{\mathrm{tr}} = \mathrm{e}_{1,3} = t_3, \\ \varphi(s_6) &= \varphi(\mathrm{e}_{3,2}) = \mathrm{e}_{3,2}^{\mathrm{tr}} = \mathrm{e}_{2,3} = t_6 \end{aligned}$$

und damit

$$\mathrm{M}_{t,s}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \qquad \square$$

Wir betrachten weitere Beispiele, welche uns verdeutlichen, dass zum einen Darstellungsmatrizen von den gegebenen Basen abhängen und dass zum anderen manche Basen „schönere" Darstellungsmatrizen liefern als andere:

(3.10) Beispiel.

(a) Es sei der $\mathbb{Q}$-Vektorraumhomomorphismus

$$\varphi\colon \mathbb{Q}^4 \to \mathbb{Q}^{2 \times 2},\ x \mapsto \begin{pmatrix} x_1 & x_1 + x_2 \\ x_1 + x_2 + x_3 & x_1 + x_2 + x_3 + x_4 \end{pmatrix}$$

gegeben. Ferner seien Basen $s = (s_1, s_2, s_3, s_4)$ und $s' = (s'_1, s'_2, s'_3, s'_4)$ von $\mathbb{Q}^{2 \times 2}$ gegeben durch

$$s = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}),$$

$$s' = \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Die Darstellungsmatrix von φ bzgl. der Basen e und s ist

$$\mathrm{M}_{s,\mathrm{e}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Die Darstellungsmatrix von φ bzgl. der Basen e und s' ist

$$\mathrm{M}_{s',\mathrm{e}}(\varphi) = \mathrm{E}_4.$$

(b) Es sei der $\mathbb{Q}$-Vektorraumhomomorphismus

$$\varphi\colon \mathbb{Q}^{2\times 2} \to \mathbb{Q}^{1\times 3}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 2a - 2b + c + 5d & -a + b + 3c + d & a - b - c + d \end{pmatrix}$$

gegeben. Ferner seien Basen $s = (s_1, s_2, s_3, s_4)$ und $s' = (s'_1, s'_2, s'_3, s'_4)$ von $\mathbb{Q}^{2\times 2}$ gegeben durch

$$s = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right),$$
$$s' = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix} \right)$$

und es sei eine Basis $t = (t_1, t_2, t_3)$ von $\mathbb{Q}^{1\times 3}$ gegeben durch

$$t = (\begin{pmatrix} 2 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}).$$

Die Darstellungsmatrix von φ bzgl. der Basen s und e ist

$$\mathrm{M}_{\mathrm{e},s}(\varphi) = \begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Die Darstellungsmatrix von φ bzgl. der Basen s' und t ist

$$\mathrm{M}_{t,s'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Beweis.

(a) Es ist

$$\varphi(\mathrm{e}_1) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = s_1 + s_2 + s_3 + s_4,$$
$$\varphi(\mathrm{e}_2) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = s_2 + s_3 + s_4,$$
$$\varphi(\mathrm{e}_3) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = s_3 + s_4,$$
$$\varphi(\mathrm{e}_4) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = s_4$$

und damit

$$\mathrm{M}_{s,\mathrm{e}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Ferner ist

$$\begin{aligned}
\varphi(\mathrm{e}_1) &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = s'_1, \\
\varphi(\mathrm{e}_2) &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = s'_2, \\
\varphi(\mathrm{e}_3) &= \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = s'_3, \\
\varphi(\mathrm{e}_4) &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = s'_4
\end{aligned}$$

und damit

$$\mathrm{M}_{s',\mathrm{e}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathrm{E}_4.$$

(b) Es ist

$$\begin{aligned}
\varphi(s_1) &= \begin{pmatrix} 2 & -1 & 1 \end{pmatrix} = 2\mathrm{e}_1 - \mathrm{e}_2 + \mathrm{e}_3, \\
\varphi(s_2) &= \begin{pmatrix} -2 & 1 & -1 \end{pmatrix} = -2\mathrm{e}_1 + \mathrm{e}_2 - \mathrm{e}_3, \\
\varphi(s_3) &= \begin{pmatrix} 1 & 3 & -1 \end{pmatrix} = \mathrm{e}_1 + 3\mathrm{e}_2 - \mathrm{e}_3, \\
\varphi(s_4) &= \begin{pmatrix} 5 & 1 & 1 \end{pmatrix} = 5\mathrm{e}_1 + \mathrm{e}_2 + \mathrm{e}_3
\end{aligned}$$

und damit

$$\mathrm{M}_{\mathrm{e},s}(\varphi) = \begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Ferner ist

$$\begin{aligned}
\varphi(s'_1) &= \begin{pmatrix} 2 & -1 & 1 \end{pmatrix} = t_1, \\
\varphi(s'_2) &= \begin{pmatrix} 1 & 3 & -1 \end{pmatrix} = t_2, \\
\varphi(s'_3) &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} = 0, \\
\varphi(s'_4) &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} = 0
\end{aligned}$$

und damit

$$\mathrm{M}_{t,s'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

□

Die nachfolgende Proposition zeigt, wieso es nützlich ist, Koordinatenvektoren als Spalten anstatt als Tupel zu schreiben: Die Anwendung eines Vektorraumhomomorphismus übersetzt sich in eine Multiplikation mit der Darstellungsmatrix, vgl. Abbildung 1.

(3.11) Proposition. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Für $v \in V$ gilt

$$\kappa_t(\varphi(v)) = \mathrm{M}_{t,s}(\varphi)\, \kappa_s(v).$$

Beweis. Für $j \in [1, n]$ gilt

$$\begin{aligned}
(\varphi_{\mathrm{M}_{t,s}(\varphi)} \circ \kappa_s)(s_j) &= \varphi_{\mathrm{M}_{t,s}(\varphi)}(\kappa_s(s_j)) = \mathrm{M}_{t,s}(\varphi)\, \kappa_s(s_j) = \mathrm{M}_{t,s}(\varphi)\, \mathrm{e}_j = (\mathrm{M}_{t,s}(\varphi))_{-,j} = \kappa_t(\varphi(s_j)) \\
&= (\kappa_t \circ \varphi)(s_j).
\end{aligned}$$

Da s eine Basis von V ist, folgt $\kappa_t \circ \varphi = \varphi_{\mathrm{M}_{t,s}(\varphi)} \circ \kappa_s$ nach Proposition (2.19)(c) und damit

$$\kappa_t(\varphi(v)) = (\kappa_t \circ \varphi)(v) = (\varphi_{\mathrm{M}_{t,s}(\varphi)} \circ \kappa_s)(v) = \varphi_{\mathrm{M}_{t,s}(\varphi)}(\kappa_s(v)) = \mathrm{M}_{t,s}(\varphi)\, \kappa_s(v)$$

für $v \in V$. □

$$\begin{tikzcd} W \arrow[d, "\kappa_t"', "\cong"] & V \arrow[l, "\varphi"'] \arrow[d, "\kappa_s", "\cong"'] \\ K^{m \times 1} & K^{n \times 1} \arrow[l, "\varphi_{\mathrm{M}_{t,s}(\varphi)}"'] \end{tikzcd}$$

Abbildung 1: Standardinterpretation der Darstellungsmatrix

Alternativer Beweis. Für $j \in [1, n]$ gilt

$$\varphi(s_j) = \sum_{i \in [1,m]} (\kappa_t(\varphi(s_j)))_i \, t_i = \sum_{i \in [1,m]} (\mathrm{M}_{t,s}(\varphi))_{i,j} t_i.$$

Folglich erhalten wir für $v \in V$ stets

$$\begin{aligned} \varphi(v) &= \varphi(\sum_{j \in [1,n]} (\kappa_s(v))_j s_j) = \sum_{j \in [1,n]} (\kappa_s(v))_j \, \varphi(s_j) = \sum_{j \in [1,n]} (\kappa_s(v))_j \sum_{i \in [1,m]} (\mathrm{M}_{t,s}(\varphi))_{i,j} t_i \\ &= \sum_{i \in [1,m]} \sum_{j \in [1,n]} (\mathrm{M}_{t,s}(\varphi))_{i,j} (\kappa_s(v))_j t_i = \sum_{i \in [1,m]} (\mathrm{M}_{t,s}(\varphi) \, \kappa_s(v))_i t_i, \end{aligned}$$

also $(\kappa_t(\varphi(v)))_i = (\mathrm{M}_{t,s}(\varphi) \, \kappa_s(v))_i$ für $i \in [1, m]$ und damit $\kappa_t(\varphi(v)) = \mathrm{M}_{t,s}(\varphi) \, \kappa_s(v)$. □

(3.12) Beispiel.

(a) Es sei der $\mathbb{Q}$-Vektorraumhomomorphismus

$$\varphi \colon \mathbb{Q}^4 \to \mathbb{Q}^{2 \times 2}, \, x \mapsto \begin{pmatrix} x_1 & x_1 + x_2 \\ x_1 + x_2 + x_3 & x_1 + x_2 + x_3 + x_4 \end{pmatrix}$$

gegeben. Ferner seien Basen $s = (s_1, s_2, s_3, s_4)$ und $s' = (s'_1, s'_2, s'_3, s'_4)$ von $\mathbb{Q}^{2 \times 2}$ gegeben durch

$$\begin{aligned} s &= (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}), \\ s' &= (\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}). \end{aligned}$$

Für $x \in \mathbb{Q}^4$ ist

$$\begin{aligned} \kappa_s(\varphi(x)) &= \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \\ x_1 + x_2 + x_3 + x_4 \end{pmatrix}, \\ \kappa_{s'}(\varphi(x)) &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \end{aligned}$$

(b) Es sei der $\mathbb{Q}$-Vektorraumhomomorphismus

$$\varphi \colon \mathbb{Q}^{2 \times 2} \to \mathbb{Q}^{1 \times 3}, \, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 2a - 2b + c + 5d & -a + b + 3c + d & a - b - c + d \end{pmatrix}$$

gegeben. Ferner seien eine Basis $s' = (s'_1, s'_2, s'_3, s'_4)$ von $\mathbb{Q}^{2 \times 2}$ und eine Basis $t = (t_1, t_2, t_3)$ von $\mathbb{Q}^{1 \times 3}$ gegeben durch

$$s' = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}),$$

$$t = (\begin{pmatrix} 2 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}).$$

Für $a, b, c, d \in \mathbb{Q}$ ist

$$\kappa_{\mathrm{e}}(\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix})) = \begin{pmatrix} 2a - 2b + c + 5d \\ -a + b + 3c + d \\ a - b - c + d \end{pmatrix},$$
$$\kappa_t(\varphi(as'_1 + bs'_2 + cs'_3 + ds'_4)) = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}.$$

Beweis.

(a) Nach Beispiel (3.10)(a) ist

$$\mathrm{M}_{s,\mathrm{e}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$
$$\mathrm{M}_{s',\mathrm{e}}(\varphi) = \mathrm{E}_4.$$

Für $x \in \mathbb{Q}^4$ folgt

$$\kappa_s(\varphi(x)) = \mathrm{M}_{s,\mathrm{e}}(\varphi)\, \kappa_{\mathrm{e}}(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \\ x_1 + x_2 + x_3 + x_4 \end{pmatrix},$$
$$\kappa_{s'}(\varphi(x)) = \mathrm{M}_{s',\mathrm{e}}(\varphi)\, \kappa_{\mathrm{e}}(x) = \mathrm{E}_4 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

nach Proposition (3.11).

(b) Es sei eine Basis $s = (s_1, s_2, s_3, s_4)$ von $\mathbb{Q}^{2 \times 2}$ gegeben durch

$$s = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

Nach Beispiel (3.10)(b) ist

$$\mathrm{M}_{\mathrm{e},s}(\varphi) = \begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix},$$
$$\mathrm{M}_{t,s'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Für $a, b, c, d \in \mathbb{Q}$ folgt

$$\kappa_{\mathrm{e}}(\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix})) = \mathrm{M}_{\mathrm{e},s}(\varphi)\, \kappa_s(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 2a - 2b + c + 5d \\ -a + b + 3c + d \\ a - b - c + d \end{pmatrix},$$
$$\kappa_t(\varphi(as'_1 + bs'_2 + cs'_3 + ds'_4)) = \mathrm{M}_{t,s'}(\varphi)\, \kappa_{s'}(as'_1 + bs'_2 + cs'_3 + ds'_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}$$

nach Proposition (3.11). □

Natürlich lässt sich Beispiel (3.12) auch ohne die Verwendung von Proposition (3.11) beweisen, was für die Fälle mit Standardbasen sogar etwas einfacher ist:

Alternativer Beweis von Beispiel (3.12).

(a) Für $x \in \mathbb{Q}^4$ ist

$$\begin{aligned}\varphi(x) &= \begin{pmatrix} x_1 & x_1 + x_2 \\ x_1 + x_2 + x_3 & x_1 + x_2 + x_3 + x_4 \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (x_1 + x_2) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + (x_1 + x_2 + x_3) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + (x_1 + x_2 + x_3 + x_4) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= x_1 s_1 + (x_1 + x_2) s_2 + (x_1 + x_2 + x_3) s_3 + (x_1 + x_2 + x_3 + x_4) s_4\end{aligned}$$

und damit

$$\kappa_s(\varphi(x)) = \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \\ x_1 + x_2 + x_3 + x_4 \end{pmatrix}.$$

Ferner gilt für $x \in \mathbb{Q}^4$ stets

$$\begin{aligned}\varphi(x) &= \begin{pmatrix} x_1 & x_1 + x_2 \\ x_1 + x_2 + x_3 & x_1 + x_2 + x_3 + x_4 \end{pmatrix} = x_1 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} + x_3 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + x_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= x_1 s'_1 + x_2 s'_2 + x_3 s'_3 + x_4 s'_4\end{aligned}$$

und damit

$$\kappa_{s'}(\varphi(x)) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

(b) Für $a, b, c, d \in \mathbb{Q}$ ist

$$\begin{aligned}&\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} 2a - 2b + c + 5d & -a + b + 3c + d & a - b - c + d \end{pmatrix} \\ &= (2a - 2b + c + 5d) \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + (-a + b + 3c + d) \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + (a - b - c + d) \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \\ &= (2a - 2b + c + 5d) \mathrm{e}_1 + (-a + b + 3c + d) \mathrm{e}_2 + (a - b - c + d) \mathrm{e}_3\end{aligned}$$

und damit

$$\kappa_{\mathrm{e}}(\varphi(\begin{pmatrix} a & b \\ c & d \end{pmatrix})) = \begin{pmatrix} 2a - 2b + c + 5d \\ -a + b + 3c + d \\ a - b - c + d \end{pmatrix}.$$

Ferner gilt für $a, b, c, d \in \mathbb{Q}$ stets

$$\begin{aligned}\varphi(a s'_1 + b s'_2 + c s'_3 + d s'_4) &= \varphi(a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + c \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + d \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}) \\ &= \varphi(\begin{pmatrix} a + c - 2d & c \\ b - d & d \end{pmatrix}) \\ &= \big(2(a + c - 2d) - 2c + (b - d) + 5d \quad -(a + c - 2d) + c + 3(b - d) + d \\ &\qquad (a + c - 2d) - c - (b - d) + d\big) \\ &= \begin{pmatrix} 2a + b & -a + 3b & a - b \end{pmatrix} = a \begin{pmatrix} 2 & -1 & 1 \end{pmatrix} + b \begin{pmatrix} 1 & 3 & -1 \end{pmatrix} = a t_1 + b t_2\end{aligned}$$

und damit

$$\kappa_t(\varphi(a s'_1 + b s'_2 + c s'_3 + d s'_4)) = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}.$$

□

$$
\begin{tikzcd}
X \arrow[d, "\kappa_u", "\cong"'] & W \arrow[l, "\psi"'] \arrow[d, "\kappa_t", "\cong"'] & V \arrow[l, "\varphi"'] \arrow[d, "\kappa_s", "\cong"'] \arrow[ll, bend right, "\psi \circ \varphi"'] \\
K^{m \times 1} & K^{n \times 1} \arrow[l, "\varphi_{\mathrm{M}_{u,t}(\psi)}"'] & K^{p \times 1} \arrow[l, "\varphi_{\mathrm{M}_{t,s}(\varphi)}"'] \arrow[ll, bend left, "\varphi_{\mathrm{M}_{u,s}(\psi \circ \varphi)}"]
\end{tikzcd}
$$

Abbildung 2: Multiplikativität der Darstellungsmatrix

Alternativer Beweis von Beispiel (3.7)(b). Es sei $\varphi\colon \mathbb{Q}^4 \to \mathbb{Q}^{2\times 2}$ gegeben durch

$$\varphi(x) = \begin{pmatrix} x_1 & x_1 + x_2 \\ x_1 + x_2 + x_3 & x_1 + x_2 + x_3 + x_4 \end{pmatrix}$$

für $x \in \mathbb{Q}^4$. Dann ist

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1+1 \\ 1+1+1 & 1+1+1+1 \end{pmatrix} = \varphi(1,1,1,1),$$

nach Beispiel (3.12) gilt also

$$\kappa_s(A) = \kappa_s(\varphi(1,1,1,1)) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \qquad \square$$

Als nächstes zeigen wir, dass Darstellungsmatrizen verträglich mit den Operationen auf Vektorraumhomomorphismen bzw. Matrizen sind, vgl. Abbildung 2.

(3.13) Proposition.

(a) Es seien K-Vektorraumhomomorphismen $\varphi\colon V \to W$ und $\psi\colon W \to X$, $m, n, p \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_p)$ von V, eine Basis $t = (t_1, \dots, t_n)$ von W und eine Basis $u = (u_1, \dots, u_m)$ von X gegeben. Dann gilt

$$\mathrm{M}_{u,s}(\psi \circ \varphi) = \mathrm{M}_{u,t}(\psi)\, \mathrm{M}_{t,s}(\varphi).$$

(b) Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Dann gilt

$$\mathrm{M}_{s,s}(\mathrm{id}_V) = \mathrm{E}_n.$$

(c) Es seien ein K-Vektorraumisomorphismus $\varphi\colon V \to W$, $n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_n)$ von W gegeben. Dann ist $\mathrm{M}_{t,s}(\varphi)$ invertierbar und es gilt

$$(\mathrm{M}_{t,s}(\varphi))^{-1} = \mathrm{M}_{s,t}(\varphi^{-1}).$$

Beweis.

(a) Nach Proposition (3.11) gilt

$$\begin{aligned}
\mathrm{M}_{u,s}(\psi \circ \varphi) &= \begin{pmatrix} \kappa_u(\psi(\varphi(s_1))) & \dots & \kappa_u(\psi(\varphi(s_p))) \end{pmatrix} = \begin{pmatrix} \mathrm{M}_{u,t}(\psi)\, \kappa_t(\varphi(s_1)) & \dots & \mathrm{M}_{u,t}(\psi)\, \kappa_t(\varphi(s_p)) \end{pmatrix} \\
&= \mathrm{M}_{u,t}(\psi) \begin{pmatrix} \kappa_t(\varphi(s_1)) & \dots & \kappa_t(\varphi(s_p)) \end{pmatrix} = \mathrm{M}_{u,t}(\psi)\, \mathrm{M}_{t,s}(\varphi).
\end{aligned}$$

(b) Es gilt

$$\mathrm{M}_{s,s}(\mathrm{id}_V) = \begin{pmatrix} \kappa_s(\mathrm{id}_V(s_1)) & \dots & \kappa_s(\mathrm{id}_V(s_n)) \end{pmatrix} = \begin{pmatrix} \kappa_s(s_1) & \dots & \kappa_s(s_n) \end{pmatrix} = \begin{pmatrix} \mathrm{e}_1 & \dots & \mathrm{e}_n \end{pmatrix} = \mathrm{E}_n.$$

(c) Nach (a), (b) gilt

$$\begin{aligned}
\mathrm{M}_{s,t}(\varphi^{-1})\,\mathrm{M}_{t,s}(\varphi) &= \mathrm{M}_{s,s}(\varphi^{-1}\circ\varphi) = \mathrm{M}_{s,s}(\mathrm{id}_V) = \mathrm{E}_n,\\
\mathrm{M}_{t,s}(\varphi)\,\mathrm{M}_{s,t}(\varphi^{-1}) &= \mathrm{M}_{t,t}(\varphi\circ\varphi^{-1}) = \mathrm{M}_{t,t}(\mathrm{id}_W) = \mathrm{E}_n.
\end{aligned}$$

Folglich ist $\mathrm{M}_{t,s}(\varphi)$ invertierbar mit $(\mathrm{M}_{t,s}(\varphi))^{-1} = \mathrm{M}_{s,t}(\varphi^{-1})$. □

(3.14) Proposition. Es seien K-Vektorräume V und W, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \ldots, s_n)$ von V und eine Basis $t = (t_1, \ldots, t_m)$ von W gegeben. Dann ist

$$\mathrm{M}_{t,s}\colon \mathrm{Hom}_K(V, W) \to K^{m\times n},\ \varphi \mapsto \mathrm{M}_{t,s}(\varphi)$$

ein K-Vektorraumisomorphismus.

Beweis. Für $\varphi, \varphi' \in \mathrm{Hom}(V, W)$ gilt

$$\begin{aligned}
\mathrm{M}_{t,s}(\varphi+\varphi') &= \begin{pmatrix}\kappa_t((\varphi+\varphi')(s_1)) & \ldots & \kappa_t((\varphi+\varphi')(s_n))\end{pmatrix}\\
&= \begin{pmatrix}\kappa_t(\varphi(s_1)+\varphi'(s_1)) & \ldots & \kappa_t(\varphi(s_n)+\varphi'(s_n))\end{pmatrix}\\
&= \begin{pmatrix}\kappa_t(\varphi(s_1))+\kappa_t(\varphi'(s_1)) & \ldots & \kappa_t(\varphi(s_n))+\kappa_t(\varphi'(s_n))\end{pmatrix}\\
&= \begin{pmatrix}\kappa_t(\varphi(s_1)) & \ldots & \kappa_t(\varphi(s_n))\end{pmatrix} + \begin{pmatrix}\kappa_t(\varphi'(s_1)) & \ldots & \kappa_t(\varphi'(s_n))\end{pmatrix} = \mathrm{M}_{t,s}(\varphi) + \mathrm{M}_{t,s}(\varphi').
\end{aligned}$$

Für $c \in K$, $\varphi \in \mathrm{Hom}(V, W)$ gilt

$$\begin{aligned}
\mathrm{M}_{t,s}(c\varphi) &= \begin{pmatrix}\kappa_t((c\varphi)(s_1)) & \ldots & \kappa_t((c\varphi)(s_n))\end{pmatrix} = \begin{pmatrix}\kappa_t(c\,\varphi(s_1)) & \ldots & \kappa_t(c\,\varphi(s_n))\end{pmatrix}\\
&= \begin{pmatrix}c\,\kappa_t(\varphi(s_1)) & \ldots & c\,\kappa_t(\varphi(s_n))\end{pmatrix} = c\begin{pmatrix}\kappa_t(\varphi(s_1)) & \ldots & \kappa_t(\varphi(s_n))\end{pmatrix} = c\,\mathrm{M}_{t,s}(\varphi).
\end{aligned}$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist $\mathrm{M}_{t,s}$ ein Homomorphismus.
Da s eine Basis von V ist, ist

$$\Phi_1\colon \mathrm{Hom}(V, W) \to W^n,\ \varphi \mapsto (\varphi(s_1), \ldots, \varphi(s_n))$$

nach Proposition (2.19)(c) eine Bijektion. Da t eine Basis von W ist, ist $\kappa_t\colon W \to K^m$ eine wohldefinierte Bijektion und damit

$$\Phi_2\colon W^n \to (K^m)^n,\ (w_1, \ldots, w_n) \mapsto (\kappa_t(w_1), \ldots, \kappa_t(w_n))$$

eine Bijektion. Schließlich ist

$$\Phi_3\colon (K^m)^n \to K^{m\times n},\ (y_1, \ldots, y_n) \mapsto \begin{pmatrix}y_1 & \ldots & y_n\end{pmatrix}$$

eine Bijektion.
Für $\varphi \in \mathrm{Hom}(V, W)$ gilt

$$(\Phi_3 \circ \Phi_2 \circ \Phi_1)(\varphi) = \Phi_3(\Phi_2(\Phi_1(\varphi))) = \begin{pmatrix}\kappa_t(\varphi(s_1)) & \ldots & \kappa_t(\varphi(s_n))\end{pmatrix} = \mathrm{M}_{t,s}(\varphi),$$

d.h. es ist $\mathrm{M}_{t,s} = \Phi_3 \circ \Phi_2 \circ \Phi_1$ nach dem Gleichheitskriterium für Abbildungen (A.35). Folglich ist $\mathrm{M}_{t,s}$ ein bijektiver Homomorphismus und damit ein Isomorphismus nach Bemerkung (2.9). □

Als nächstes untersuchen wir, was beim Wechsel von Basen passiert. Es stellt sich beispielsweise die Frage, wie die Koordinatenspalten bzgl. einer Basis in die Koordinatenspalten bzgl. einer anderen Basis umgerechnet werden können. Für die Antwort werden wir nachfolgenden Begriff benötigen.

(3.15) Definition (Basiswechselmatrix)**.** Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$ und Basen $s = (s_1, \ldots, s_n)$ und $s' = (s'_1, \ldots, s'_n)$ von V gegeben. Die Darstellungsmatrix $\mathrm{M}_{s,s'}(\mathrm{id}_V)$ heißt *Basiswechselmatrix* (oder *Transformationsmatrix*) von s nach s'.

Für einen K-Vektorraum V, $n \in \mathbb{N}_0$ und Basen $s = (s_1, \ldots, s_n)$ und $s' = (s'_1, \ldots, s'_n)$ von V stehen in der Basiswechselmatrix von s nach s' also die Koeffizienten der Einträge der Basis s' in der Darstellung als Linearkombinationen in der Basis s; es ist

$$\mathrm{M}_{s,s'}(\mathrm{id}_V) = \begin{pmatrix}\kappa_s(s'_1) & \ldots & \kappa_s(s'_n)\end{pmatrix}.$$

$$
\begin{tikzcd}
V & & V \arrow[ll, "\mathrm{id}_V"', "="] \\
K^{n\times 1} \arrow[from=u, "\kappa_{s'}"', "\cong"] & & K^{n\times 1} \arrow[ll, "\varphi_{(\mathrm{M}_{s,s'}(\mathrm{id}_V))^{-1}}"', "\cong"] \arrow[from=u, "\cong"', "\kappa_s"]
\end{tikzcd}
\qquad
\begin{tikzcd}
W & W \arrow[l, "\mathrm{id}_W"', "="] & V \arrow[l, "\varphi"'] & V \arrow[l, "\mathrm{id}_V"', "="] \arrow[lll, bend right, "\varphi"'] \\
K^{m\times 1} \arrow[from=u, "\cong"', "\kappa_{t'}"] & K^{m\times 1} \arrow[l, "\varphi_{(\mathrm{M}_{t,t'}(\mathrm{id}_W))^{-1}}"', "\cong"] \arrow[from=u, "\cong"', "\kappa_t"] & K^{n\times 1} \arrow[l, "\varphi_{\mathrm{M}_{t,s}(\varphi)}"'] \arrow[from=u, "\cong"', "\kappa_s"] & K^{n\times 1} \arrow[l, "\varphi_{\mathrm{M}_{s,s'}(\mathrm{id}_V)}"', "\cong"] \arrow[from=u, "\cong"', "\kappa_{s'}"] \arrow[lll, bend left, "\varphi_{\mathrm{M}_{t',s'}(\varphi)}"]
\end{tikzcd}
$$

Abbildung 3: Basiswechsel

(3.16) Beispiel. Die Basiswechselmatrix zu den Basen $\mathrm{e} = ((1,0),(0,1))$ und $s = ((1,0),(1,1))$ von K^2 ist

$$\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{K^2}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Beweis. Es ist

$$s_1 = (1,0) = \mathrm{e}_1,$$
$$s_2 = (1,1) = \mathrm{e}_1 + \mathrm{e}_2$$

und damit

$$\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{K^2}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$ □

Die Basiswechselmatrix zum Basiswechsel von einer Basis s nach einer Basis s' eines endlichdimensionalen Vektorraums V ist per Definition die Darstellungsmatrix von $\mathrm{id}_V \colon V \to V$ zu den Basen s' (für den Startvektorraum V) und s (für den Zielvektorraum V). Der Grund, warum die Reihenfolge in der Bezeichnung so gewählt ist, liegt in der praktischen Verwendung: Üblicherweise (zum Beispiel auf einem Computer) gibt man alle Vektoren eines (abstrakten) endlichdimensionalen Vektorraums V als Koordinatenspalten bzgl. einer gewählten Basis s an. Möchte man nun die Vektoren durch eine andere Basis s' darstellen, also einen Basiswechsel von der Basis s zur Basis s' durchführen, so hat man insbesondere die Basis s' durch Koordinatenspalten bzgl. der Basis s vorliegen – man hat also gerade die Basiswechselmatrix $\mathrm{M}_{s,s'}(\mathrm{id}_V)$ zum Basiswechsel von s nach s' gegeben. Zum Umrechnen der Koordinatenspalten wird dann entweder die Inverse dieser Basiswechselmatrix verwendet oder ein lineares Gleichungssystem gelöst, siehe Proposition (3.18)(a).
Ein typisches Beispiel ist der Fall, dass $V = K^{n\times 1}$ für ein $n \in \mathbb{N}_0$ und $s = \mathrm{e} = (\mathrm{e}_1, \dots, \mathrm{e}_n)$ ist. Für eine beliebige Basis s' von $K^{n\times 1}$ ist die Basiswechselmatrix zum Basiswechsel von $s = \mathrm{e}$ nach s' dann durch

$$\mathrm{M}_{s,s'}(\mathrm{id}_{K^{n\times 1}}) = \mathrm{M}_{\mathrm{e},s'}(\mathrm{id}_{K^{n\times 1}}) = \begin{pmatrix} \kappa_{\mathrm{e}}(s'_1) & \dots & \kappa_{\mathrm{e}}(s'_n) \end{pmatrix} = \begin{pmatrix} s'_1 & \dots & s'_n \end{pmatrix}$$

gegeben.

(3.17) Bemerkung. Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$ und Basen $s = (s_1, \dots, s_n)$ und $s' = (s'_1, \dots, s'_n)$ von V gegeben. Die Basiswechselmatrix $\mathrm{M}_{s,s'}(\mathrm{id}_V)$ ist invertierbar mit

$$(\mathrm{M}_{s,s'}(\mathrm{id}_V))^{-1} = \mathrm{M}_{s',s}(\mathrm{id}_V)$$

Beweis. Dies folgt aus Proposition (3.13)(c). □

(3.18) Proposition (Basiswechselformeln)**.**

(a) Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$ und Basen $s = (s_1, \dots, s_n)$ und $s' = (s'_1, \dots, s'_n)$ von V gegeben. Für $v \in V$ gilt

$$\kappa_{s'}(v) = (\mathrm{M}_{s,s'}(\mathrm{id}_V))^{-1}\kappa_s(v).$$

(b) Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, Basen $s = (s_1, \dots, s_n)$ und $s' = (s'_1, \dots, s'_n)$ von V und Basen $t = (t_1, \dots, t_m)$ und $t' = (t'_1, \dots, t'_m)$ von W gegeben. Dann gilt

$$\mathrm{M}_{t',s'}(\varphi) = (\mathrm{M}_{t,t'}(\mathrm{id}_W))^{-1}\,\mathrm{M}_{t,s}(\varphi)\,\mathrm{M}_{s,s'}(\mathrm{id}_V).$$

Beweis.

(a) Nach Proposition (3.11) und Bemerkung (3.17) gilt

$$\kappa_{s'}(v) = \kappa_{s'}(\mathrm{id}_V(v)) = \mathrm{M}_{s',s}(\mathrm{id}_V)\,\kappa_s(v) = (\mathrm{M}_{s,s'}(\mathrm{id}_V))^{-1}\kappa_s(v).$$

(b) Nach Proposition (3.13)(a) und Bemerkung (3.17) gilt

$$\begin{aligned}\mathrm{M}_{t',s'}(\varphi) &= \mathrm{M}_{t',s'}(\mathrm{id}_W \circ \varphi \circ \mathrm{id}_V) = \mathrm{M}_{t',t}(\mathrm{id}_W)\,\mathrm{M}_{t,s}(\varphi)\,\mathrm{M}_{s,s'}(\mathrm{id}_V)\\ &= (\mathrm{M}_{t,t'}(\mathrm{id}_W))^{-1}\,\mathrm{M}_{t,s}(\varphi)\,\mathrm{M}_{s,s'}(\mathrm{id}_V).\end{aligned}$$ □

(3.19) Beispiel. Es seien $A \in \mathbb{R}^{2\times 2}$ und eine Basis $s = (s_1, s_2)$ von $\mathbb{R}^{2\times 1}$ gegeben durch

$$A = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix},\ s = (\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}).$$

Dann ist

$$\mathrm{M}_{s,s}(\varphi_A) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Beweis. Die Basiswechselmatrix von e nach s ist gegeben durch

$$\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{\mathbb{R}^{2\times 1}}) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix},$$

ihre Inverse durch

$$(\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{\mathbb{R}^{2\times 1}}))^{-1} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

Nach der Basiswechselformel (3.18)(b) erhalten wir

$$\begin{aligned}\mathrm{M}_{s,s}(\varphi_A) &= (\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{\mathbb{R}^{2\times 1}}))^{-1}\,\mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)\,\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{\mathbb{R}^{2\times 1}}) = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\\ &= \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\frac{1}{5}\begin{pmatrix} 5 & 10 \\ 10 & -5 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\end{aligned}$$ □

Spaltenraum

Im Folgenden werden die in Abschnitt 2 eingeführten Konzepte auf die Spalteninterpretation angewandt. Wir beginnen mit dem Bild, welches nach Bemerkung (2.12)(a) stets ein Untervektorraum des Zielvektorraums ist.

(3.20) Definition (Spaltenraum)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m\times n}$ gegeben. Der K-Untervektorraum

$$\mathrm{Col}(A) := \mathrm{Im}\,\varphi_A$$

von $K^{m\times 1}$ heißt *Spaltenraum* von A.

(3.21) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m\times n}$ gegeben. Dann ist

$$\mathrm{Col}(A) = \{Ax \mid x \in K^{n\times 1}\} = \langle A_{-,1}, \dots, A_{-,n}\rangle.$$

Beweis. Es ist

$$\operatorname{Col}(A) = \operatorname{Im} \varphi_A = \{\varphi_A(x) \mid x \in K^{n\times 1}\} = \{Ax \mid x \in K^{n\times 1}\}$$

Nach Bemerkung (3.4) ist ferner $Ax = \sum_{j\in[1,n]} x_j A_{-,j}$ für $x \in K^{n\times 1}$ und damit

$$\operatorname{Col}(A) = \{Ax \mid x \in K^{n\times 1}\} = \{\sum_{j\in[1,n]} x_j A_{-,j} \mid x \in K^{n\times 1}\} = \langle A_{-,1}, \dots, A_{-,n}\rangle. \qquad \square$$

(3.22) Beispiel. Es sei $A \in \mathbb{Q}^{3\times 3}$ gegeben durch

$$A = \begin{pmatrix} -2 & -1 & 4 \\ -3 & -1 & 5 \\ 2 & 2 & -6 \end{pmatrix}.$$

Dann ist

$$\operatorname{Col}(A) = \mathbb{Q}\begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} + \mathbb{Q}\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Beweis. Wir formen A^{tr} durch elementare Zeilenoperationen um:

$$\begin{pmatrix} -2 & -3 & 2 \\ -1 & -1 & 2 \\ 4 & 5 & -6 \end{pmatrix} \xmapsto{\text{add}_{3,2,4} \circ \text{add}_{1,2,-2}} \begin{pmatrix} 0 & -1 & -2 \\ -1 & -1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \xmapsto{\text{add}_{3,1,1} \circ \text{add}_{2,1,-1}} \begin{pmatrix} 0 & -1 & -2 \\ -1 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xmapsto{\text{mul}_{2,-1} \circ \text{mul}_{1,-1}} \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & -4 \\ 0 & 0 & 0 \end{pmatrix} \xmapsto{\text{sw}_{1,2}} \begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Nach Bemerkung (3.21), Korollar (1.23) und Proposition (1.22) ist

$$\operatorname{Col}(A) = \langle \begin{pmatrix} -2 \\ -3 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ -6 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \rangle = \mathbb{Q}\begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} + \mathbb{Q}\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}. \qquad \square$$

Mit Hilfe eines Spaltenraums lässt sich das Bild eines beliebigen Vektorraumhomomorphismus zwischen endlichdimensionalen Vektorräumen bestimmen:

(3.23) Bemerkung. Es seien ein K-Vektorraumhomomorphismus $\varphi \colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Dann ist

$$\kappa_t(\operatorname{Im} \varphi) = \operatorname{Col}(\mathrm{M}_{t,s}(\varphi))$$

und damit

$$\operatorname{Im} \varphi = \{\sum_{i\in[1,m]} b_i t_i \mid b \in \operatorname{Col}(\mathrm{M}_{t,s}(\varphi))\}.$$

Beweis. Da $\kappa_s \colon V \to K^{n\times 1}$ eine Bijektion ist, gilt

$$\begin{aligned} \kappa_t(\operatorname{Im} \varphi) &= \kappa_t(\{\varphi(v) \mid v \in V\}) = \{\kappa_t(\varphi(v)) \mid v \in V\} = \{\mathrm{M}_{t,s}(\varphi)\, \kappa_s(v) \mid v \in V\} = \{\mathrm{M}_{t,s}(\varphi)\, x \mid x \in K^{n\times 1}\} \\ &= \operatorname{Col}(\mathrm{M}_{t,s}(\varphi)) \end{aligned}$$

nach Proposition (3.11). Da ferner $\kappa_t \colon W \to K^{m\times 1}$ eine Bijektion ist, folgt

$$\operatorname{Im} \varphi = \kappa_t^{-1}(\operatorname{Col}(\mathrm{M}_{t,s}(\varphi))) = \{\kappa_t^{-1}(b) \mid b \in \operatorname{Col}(\mathrm{M}_{t,s}(\varphi))\} = \{\sum_{i\in[1,m]} b_i t_i \mid b \in \operatorname{Col}(\mathrm{M}_{t,s}(\varphi))\}. \qquad \square$$

Alternativer Beweis von Beispiel (2.14)(a). Es sei eine Basis $s = (s_1, s_2, s_3, s_4)$ von $\mathbb{Q}^{2\times 2}$ gegeben durch

$$s = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

Nach Beispiel (3.10)(b) ist

$$\mathrm{M}_{\mathrm{e},s}(\varphi) = \begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Wir formen $\mathrm{M}_{\mathrm{e},s}(\varphi)^{\mathrm{tr}}$ mittels elementarer Zeilenoperationen um:

$$\begin{pmatrix} 2 & -1 & 1 \\ -2 & 1 & -1 \\ 1 & 3 & -1 \\ 5 & 1 & 1 \end{pmatrix} \xmapsto{\mathrm{add}_{4,1,-1} \circ \mathrm{add}_{3,1,1} \circ \mathrm{add}_{2,1,1}} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 0 & 0 \\ 3 & 2 & 0 \\ 3 & 2 & 0 \end{pmatrix} \xmapsto{\mathrm{mul}_{1,2}} \begin{pmatrix} 4 & -2 & 2 \\ 0 & 0 & 0 \\ 3 & 2 & 0 \\ 3 & 2 & 0 \end{pmatrix}$$

$$\xmapsto{\mathrm{add}_{4,3,-1} \circ \mathrm{add}_{1,3,1}} \begin{pmatrix} 7 & 0 & 2 \\ 0 & 0 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Nach Bemerkung (3.21), Korollar (1.23) und Proposition (1.22) ist

$$\mathrm{Col}(\mathrm{M}_{\mathrm{e},s}(\varphi)) = \langle \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix} \rangle = \langle \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \rangle = \mathbb{Q} \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

und nach Bemerkung (3.23) folgt

$$\begin{aligned} \mathrm{Im}\,\varphi &= \kappa_{\mathrm{e}}^{-1}(\mathrm{Col}(\mathrm{M}_{\mathrm{e},s}(\varphi))) = \kappa_{\mathrm{e}}^{-1}(\mathbb{Q} \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}) = \mathbb{Q}\,\kappa_{\mathrm{e}}^{-1}(\begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix}) + \mathbb{Q}\,\kappa_{\mathrm{e}}^{-1}(\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}) \\ &= \mathbb{Q} \begin{pmatrix} 7 & 0 & 2 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 3 & 2 & 0 \end{pmatrix}. \end{aligned}$$

□

(3.24) Bemerkung. Es seien $m, n, p \in \mathbb{N}_0$ und $A \in K^{n\times p}$, $B \in K^{m\times n}$ gegeben. Dann ist $\mathrm{Col}(BA)$ ein K-Untervektorraum von $\mathrm{Col}(B)$.

Beweis. Nach Bemerkung (3.5)(a) ist

$$\mathrm{Col}(BA) = \mathrm{Im}\,\varphi_{BA} = \mathrm{Im}(\varphi_B \circ \varphi_A) \subseteq \mathrm{Im}\,\varphi_B = \mathrm{Col}(B).$$

Da sowohl $\mathrm{Col}(BA)$ als auch $\mathrm{Col}(B)$ Untervektorräume von $K^{m\times 1}$ sind, ist folglich $\mathrm{Col}(BA)$ ein Untervektorraum von $\mathrm{Col}(B)$. □

(3.25) Korollar. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m\times n}$, $P \in \mathrm{GL}_n(K)$ gegeben. Dann ist

$$\mathrm{Col}(AP) = \mathrm{Col}(A).$$

Beweis. Nach Bemerkung (3.24) gilt $\mathrm{Col}(AP) \leq \mathrm{Col}(A)$ und $\mathrm{Col}(A) = \mathrm{Col}(APP^{-1}) \leq \mathrm{Col}(AP)$, also $\mathrm{Col}(AP) = \mathrm{Col}(A)$. □

(3.26) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m\times n}$, $P \in \mathrm{GL}_m(K)$ gegeben. Dann ist

$$\varphi_P|_{\mathrm{Col}(A)}^{\mathrm{Col}(PA)} \colon \mathrm{Col}(A) \to \mathrm{Col}(PA),\ y \mapsto Py$$

ein wohldefinierter K-Vektorraumisomorphismus.

Beweis. Auf Grund der Invertierbarkeit von P ist $\varphi_P \colon K^{m\times 1} \to K^{m\times 1}$, $y \mapsto Py$ ein Isomorphismus. Folglich ist

$$\varphi_P(\mathrm{Col}(A)) = \varphi_P(\mathrm{Im}\,\varphi_A) = \mathrm{Im}(\varphi_P \circ \varphi_A) = \mathrm{Im}\,\varphi_{PA} = \mathrm{Col}(PA),$$

so dass $\varphi_P \colon K^{m\times 1} \to K^{m\times 1}$ zu einem surjektiven Homomorphismus $\varphi_P|_{\mathrm{Col}(A)}^{\mathrm{Col}(PA)} \colon \mathrm{Col}(A) \to \mathrm{Col}(PA)$ einschränkt. Die Injektivität von φ_P impliziert ferner die Injektivität von $\varphi_P|_{\mathrm{Col}(A)}^{\mathrm{Col}(PA)}$. Insgesamt ist $\varphi_P|_{\mathrm{Col}(A)}^{\mathrm{Col}(PA)}$ bijektiv und daher ein Isomorphismus nach Bemerkung (2.9). □

Rang einer Matrix

Als nächstes studieren wir den Rang der Spalteninterpretation.

(3.27) Definition (Rang)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Der *Rang* (oder *Spaltenrang*) von A ist definiert als

$$\operatorname{rk} A = \operatorname{rk}_K A := \operatorname{rk}_K \varphi_A.$$

Da der Rang eines Homomorphismus per Definition die Dimension seines Bildes ist, ergibt sich folgender Zusammenhang:

(3.28) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Dann ist

$$\operatorname{rk}_K A = \dim_K \operatorname{Col}(A).$$

Beweis. Es ist

$$\operatorname{rk} A = \operatorname{rk} \varphi_A = \dim(\operatorname{Im} \varphi_A) = \dim \operatorname{Col}(A). \qquad \square$$

(3.29) Bemerkung. Es seien ein K-Vektorraumhomomorphismus $\varphi \colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Dann ist

$$\operatorname{rk}_K \varphi = \operatorname{rk}_K \mathrm{M}_{t,s}(\varphi).$$

Beweis. Nach Bemerkung (3.23) ist $\kappa_t(\operatorname{Im} \varphi) = \operatorname{Col}(\mathrm{M}_{t,s}(\varphi))$. Da $\kappa_t \colon W \to K^{m \times 1}$ ein Isomorphismus ist, folgt

$$\operatorname{rk} \varphi = \dim(\operatorname{Im} \varphi) = \dim \kappa_t(\operatorname{Im} \varphi) = \dim \operatorname{Col}(\mathrm{M}_{t,s}(\varphi)) = \operatorname{rk} \mathrm{M}_{t,s}(\varphi)$$

nach Satz (2.32) und Bemerkung (3.28). $\square$

(3.30) Beispiel. Es sei $A \in \mathbb{Q}^{3 \times 3}$ gegeben durch

$$A = \begin{pmatrix} -2 & -1 & 4 \\ -3 & -1 & 5 \\ 2 & 2 & -6 \end{pmatrix}.$$

Dann ist

$$\operatorname{rk}_{\mathbb{Q}} A = 2.$$

Beweis. Nach Beispiel (3.22) ist

$$\operatorname{Col}(A) = \langle \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \rangle$$

und nach Bemerkung (3.28) somit

$$\operatorname{rk} A = \dim \operatorname{Col}(A) = 2. \qquad \square$$

(3.31) Bemerkung. Für $n \in \mathbb{N}_0$ ist

$$\mathrm{GL}_n(K) = \{A \in K^{n \times n} \mid \operatorname{rk}_K A = n\}.$$

Beweis. Nach Korollar (2.29) gilt für $n \in \mathbb{N}_0$ stets

$$\begin{aligned} \mathrm{GL}_n(K) &= \{A \in K^{n \times n} \mid A \text{ ist invertierbar}\} = \{A \in K^{n \times n} \mid \varphi_A \text{ ist ein Isomorphismus}\} \\ &= \{A \in K^{n \times n} \mid \operatorname{rk} \varphi_A = \dim K^{n \times 1}\} = \{A \in K^{n \times n} \mid \operatorname{rk} A = n\}. \end{aligned} \qquad \square$$

(3.32) Korollar. Es seien ein K-Vektorraum V, $n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und $P \in \mathrm{GL}_n(K)$ gegeben. Ferner sei $s' := (\kappa_s^{-1}(P_{-,1}), \dots, \kappa_s^{-1}(P_{-,n}))$. Dann ist s' eine Basis von V mit

$$\mathrm{M}_{s,s'}(\mathrm{id}_V) = P.$$

Beweis. Nach Bemerkung (3.28) und Bemerkung (3.31) ist $\dim \mathrm{Col}(P) = \mathrm{rk}\, P = n$. Folglich ist $(P_{-,1}, \dots, P_{-,n})$ eine Basis von $K^{n \times 1}$. Da $\kappa_s^{-1} \colon K^{n \times 1} \to V$ ein Isomorphismus ist, ist $s' = (\kappa_s^{-1}(P_{-,1}), \dots, \kappa_s^{-1}(P_{-,n}))$ nach Proposition (2.21)(c) eine Basis von V. Schließlich gilt

$$\mathrm{M}_{s,s'}(\mathrm{id}_V) = \begin{pmatrix} \kappa_s(s'_1) & \dots & \kappa_s(s'_n) \end{pmatrix} = P. \qquad \square$$

(3.33) Bemerkung. Es seien $m, n, p \in \mathbb{N}_0$ gegeben. Für $A \in K^{n \times p}$, $B \in K^{m \times n}$ gilt

$$\mathrm{rk}_K(BA) \leq \min(\mathrm{rk}_K\, B, \mathrm{rk}_K\, A).$$

Beweis. Nach Proposition (2.30) gilt

$$\mathrm{rk}(BA) = \mathrm{rk}\, \varphi_{BA} = \mathrm{rk}(\varphi_B \circ \varphi_A) \leq \min(\mathrm{rk}\, \varphi_B, \mathrm{rk}\, \varphi_A) = \min(\mathrm{rk}\, B, \mathrm{rk}\, A). \qquad \square$$

(3.34) Korollar. Für $n \in \mathbb{N}_0$ ist

$$\begin{aligned} \mathrm{GL}_n(K) &= \{A \in K^{n \times n} \mid \text{es gibt ein } B \in K^{n \times n} \text{ mit } BA = \mathrm{E}_n\} \\ &= \{A \in K^{n \times n} \mid \text{es gibt ein } B \in K^{n \times n} \text{ mit } AB = \mathrm{E}_n\}. \end{aligned}$$

Beweis. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Wenn $A \in \mathrm{GL}_n(K)$ ist, gilt $A^{-1}A = \mathrm{E}_n$. Es existiere also umgekehrt ein $B \in K^{n \times n}$ mit $BA = \mathrm{E}_n$. Nach Bemerkung (3.33) folgt

$$n = \mathrm{rk}\, \mathrm{E}_n = \mathrm{rk}(BA) \leq \mathrm{rk}\, A.$$

Da aber stets auch $\mathrm{rk}\, A \leq n$ ist, impliziert dies bereits $\mathrm{rk}\, A = n$ und damit $A \in \mathrm{GL}_n(K)$ nach Bemerkung (3.31). Insgesamt gilt $\mathrm{GL}_n(K) = \{A \in K^{n \times n} \mid \text{es gibt ein } B \in K^{n \times n} \text{ mit } BA = \mathrm{E}_n\}$ für $n \in \mathbb{N}_0$.
Die Gleichung $\mathrm{GL}_n(K) = \{A \in K^{n \times n} \mid \text{es gibt ein } B \in K^{n \times n} \text{ mit } AB = \mathrm{E}_n\}$ für $n \in \mathbb{N}_0$ lässt sich analog zeigen. $\square$

Lineare Gleichungssysteme

Als nächstes zeigen wir einen Zusammenhang zwischen Spalteninterpretationen von Matrizen und Lösungsmengen linearer Gleichungssysteme auf.

(3.35) Bemerkung. Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Dann ist

$$\mathrm{Sol}(A, b) = \varphi_A^{-1}(\{b\}).$$

Beweis. Es ist

$$\mathrm{Sol}(A, b) = \{x \in K^{n \times 1} \mid Ax = b\} = \{x \in K^{n \times 1} \mid \varphi_A(x) = b\} = \varphi_A^{-1}(\{b\}). \qquad \square$$

(3.36) Korollar. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Dann ist

$$\mathrm{Sol}(A, 0) = \mathrm{Ker}\, \varphi_A.$$

Insbesondere ist $\mathrm{Sol}(A, 0)$ ein endlichdimensionaler K-Untervektorraum von $K^{n \times 1}$ der Dimension

$$\dim_K \mathrm{Sol}(A, 0) = n - \mathrm{rk}_K\, A.$$

Beweis. Nach Bemerkung (3.35) ist

$$\mathrm{Sol}(A, 0) = \varphi_A^{-1}(\{0\}) = \mathrm{Ker}\, \varphi_A.$$

Folglich ist $\mathrm{Sol}(A, 0)$ nach Bemerkung (2.12)(b) und dem Rangsatz (2.28) ein endlichdimensionaler Untervektorraum von $K^{n \times 1}$ mit

$$\dim \mathrm{Sol}(A, 0) = \dim(\mathrm{Ker}\, \varphi_A) = \mathrm{def}\, \varphi_A = \dim K^{n \times 1} - \mathrm{rk}\, \varphi_A = n - \mathrm{rk}\, A. \qquad \square$$

(3.37) Korollar. Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$, $b \in K^{m \times 1}$ und $x \in \mathrm{Sol}(A, b)$ gegeben. Dann ist

$$\mathrm{Sol}(A, 0) \to \mathrm{Sol}(A, b),\ u \mapsto x + u$$

eine wohldefinierte Bijektion. Insbesondere ist

$$\mathrm{Sol}(A, b) = x + \mathrm{Sol}(A, 0).$$

Beweis. Nach Proposition (2.15) ist $\mathrm{Ker}\,\varphi_A \to \varphi_A^{-1}(\{\varphi_A(x)\})$, $u \mapsto x + u$ eine wohldefinierte Bijektion und $\varphi_A^{-1}(\{\varphi_A(x)\}) = x + \mathrm{Ker}\,\varphi_A$. Nun ist aber $\mathrm{Ker}\,\varphi_A = \mathrm{Sol}(A, 0)$ nach Korollar (3.36) sowie $\varphi_A^{-1}(\{\varphi_A(x)\}) = \varphi_A^{-1}(\{b\}) = \mathrm{Sol}(A, b)$ nach Bemerkung (3.35). Folglich ist

$$\mathrm{Sol}(A, 0) \to \mathrm{Sol}(A, b),\ u \mapsto x + u$$

eine wohldefinierte Bijektion und

$$\mathrm{Sol}(A, b) = \varphi_A^{-1}(\{\varphi_A(x)\}) = x + \mathrm{Ker}\,\varphi_A = x + \mathrm{Sol}(A, 0). \qquad \square$$

Mit Hilfe der Lösungsmenge eines homogenen linearen Gleichungssystems lässt sich der Kern eines beliebigen Vektorraumhomomorphismus zwischen endlichdimensionalen Vektorräumen bestimmen:

(3.38) Bemerkung. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $w \in W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Dann ist

$$\kappa_s(\varphi^{-1}(\{w\})) = \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(w))$$

und damit

$$\varphi^{-1}(\{w\}) = \{\sum_{j \in [1,n]} a_j s_j \mid a \in \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(w))\}.$$

Beweis. Da $\kappa_t\colon W \to K^{m \times 1}$ ein injektiver Homomorphismus und $\kappa_s\colon V \to K^{n \times 1}$ eine Bijektion ist, gilt

$$\begin{aligned}\kappa_s(\varphi^{-1}(\{w\})) &= \kappa_s(\{v \in V \mid \varphi(v) = w\}) = \{\kappa_s(v) \mid v \in V \text{ mit } \varphi(v) = w\} \\ &= \{\kappa_s(v) \mid v \in V \text{ mit } \kappa_t(\varphi(v)) = \kappa_t(w)\} = \{\kappa_s(v) \mid v \in V \text{ mit } \mathrm{M}_{t,s}(\varphi)\,\kappa_s(v) = \kappa_t(w)\} \\ &= \{x \in K^{n \times 1} \mid \mathrm{M}_{t,s}(\varphi)x = \kappa_t(w)\} = \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(w))\end{aligned}$$

nach Proposition (3.11) und damit

$$\begin{aligned}\varphi^{-1}(\{w\}) &= \kappa_s^{-1}(\mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(w))) = \{\kappa_s^{-1}(a) \mid a \in \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(w))\} \\ &= \{\sum_{j \in [1,n]} a_j s_j \mid a \in \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(w))\}. \qquad \square\end{aligned}$$

(3.39) Korollar. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Dann ist

$$\kappa_s(\mathrm{Ker}\,\varphi) = \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), 0)$$

und

$$\mathrm{Ker}\,\varphi = \{\sum_{j \in [1,n]} a_j s_j \mid a \in \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), 0)\}.$$

Beweis. Nach Bemerkung (3.38) ist

$$\kappa_s(\mathrm{Ker}\,\varphi) = \kappa_s(\varphi^{-1}(\{0\})) = \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(0)) = \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), 0)$$

und

$$\mathrm{Ker}\,\varphi = \varphi^{-1}(\{0\}) = \{\sum_{j \in [1,n]} a_j s_j \mid a \in \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), \kappa_t(0))\} = \{\sum_{j \in [1,n]} a_j s_j \mid a \in \mathrm{Sol}(\mathrm{M}_{t,s}(\varphi), 0)\}. \qquad \square$$

(3.40) Korollar. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W gegeben. Dann ist

$$\operatorname{def}_K \varphi = \dim_K \operatorname{Sol}(\mathrm{M}_{t,s}(\varphi), 0).$$

Beweis. Nach Bemerkung (3.39) ist $\kappa_s(\operatorname{Ker} \varphi) = \operatorname{Sol}(\mathrm{M}_{t,s}(\varphi), 0)$. Da $\kappa_s\colon V \to K^{n \times 1}$ ein Isomorphismus ist, folgt

$$\operatorname{def} \varphi = \dim(\operatorname{Ker} \varphi) = \dim \kappa_s(\operatorname{Ker} \varphi) = \dim \operatorname{Sol}(\mathrm{M}_{t,s}(\varphi), 0)$$

nach Satz (2.32). □

Alternativer Beweis von Beispiel (2.14)(b). Es sei eine Basis $s = (s_1, s_2, s_3, s_4)$ von $\mathbb{Q}^{2 \times 2}$ gegeben durch

$$s = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

Nach Beispiel (3.10)(b) ist

$$\mathrm{M}_{\mathrm{e},s}(\varphi) = \begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Wir formen $\mathrm{M}_{\mathrm{e},s}(\varphi)$ mittels elementarer Zeilenoperationen um:

$$\begin{pmatrix} 2 & -2 & 1 & 5 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \xmapsto{\mathrm{add}_{2,3,1} \circ \mathrm{add}_{1,3,-2}} \begin{pmatrix} 0 & 0 & 3 & 3 \\ 0 & 0 & 2 & 2 \\ 1 & -1 & -1 & 1 \end{pmatrix} \xmapsto{\mathrm{mul}_{2,\frac{1}{2}}} \begin{pmatrix} 0 & 0 & 3 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$
$$\xmapsto{\mathrm{add}_{3,2,1} \circ \mathrm{add}_{1,2,-3}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 2 \end{pmatrix} \xmapsto{\mathrm{sw}_{1,3}} \begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Nach Proposition (A.163) ist

$$\operatorname{Sol}(\mathrm{M}_{\mathrm{e},s}(\varphi), 0) = \operatorname{Sol}(\begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, 0) = \mathbb{Q} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

und nach Bemerkung (3.39) folgt

$$\operatorname{Ker} \varphi = \kappa_s^{-1}(\operatorname{Sol}(\mathrm{M}_{\mathrm{e},s}(\varphi), 0)) = \kappa_s^{-1}(\mathbb{Q} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}) = \mathbb{Q}\, \kappa_s^{-1}(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}) + \mathbb{Q}\, \kappa_s^{-1}(\begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix})$$
$$= \mathbb{Q} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}.$$ □

Da die Lösungsmenge eines linearen Gleichungssystems eine Faser der Spalteninterpretation der zugehörigen Koeffizientenmatrix ist, führt das Studium dieses Homomorphismus zu konzeptionellen Kriterien für die Lösbarkeit eines linearen Gleichungssystems. Vgl. Proposition (A.159).

(3.41) Proposition (Lösbarkeitskriterium für lineare Gleichungssysteme)**.** Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ besitzt eine Lösung.

(b) Es ist $b \in \operatorname{Col}(A)$.

(c) Es ist $\operatorname{Col}(\begin{pmatrix} A \mid b \end{pmatrix}) = \operatorname{Col}(A)$.

(d) Es ist $\operatorname{rk}_K \begin{pmatrix} A \mid b \end{pmatrix} = \operatorname{rk}_K A$.

Beweis. Nach Bemerkung (3.35) ist $\mathrm{Sol}(A, b) = \varphi_A^{-1}(\{b\})$. Folglich besitzt das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ genau dann eine Lösung, wenn $\varphi_A^{-1}(\{b\}) \neq \emptyset$ ist. Dies bedeutet aber gerade, dass es ein $x \in K^{n \times 1}$ mit $\varphi_A(x) = b$ gibt, d.h. dass $b \in \operatorname{Im} \varphi_A = \mathrm{Col}(A)$ ist. Dies zeigt die Äquivalenz von Bedingung (a) und Bedingung (b).
Nach Bemerkung (3.21) ist $\mathrm{Col}(A) = \langle A_{-,1}, \dots, A_{-,n} \rangle$ und $\mathrm{Col}(\begin{pmatrix} A \mid b \end{pmatrix}) = \langle A_{-,1}, \dots, A_{-,n}, b \rangle$. Somit gilt genau dann $b \in \mathrm{Col}(A)$, wenn b eine Linearkombination von $(A_{-,1}, \dots, A_{-,n})$ ist. Dies ist nach Proposition (1.22) wiederum äquivalent zu $\langle A_{-,1}, \dots, A_{-,n} \rangle = \langle A_{-,1}, \dots, A_{-,n}, b \rangle$, d.h. zu $\mathrm{Col}(A) = \mathrm{Col}(\begin{pmatrix} A \mid b \end{pmatrix})$. Dies zeigt die Äquivalenz von Bedingung (b) und Bedingung (c).
Es gilt stets

$$\mathrm{Col}(A) = \langle A_{-,1}, \dots, A_{-,n} \rangle \leq \langle A_{-,1}, \dots, A_{-,n}, b \rangle = \mathrm{Col}(\begin{pmatrix} A \mid b \end{pmatrix}).$$

Nach Proposotion (1.68) gilt also genau dann $\mathrm{Col}(\begin{pmatrix} A \mid b \end{pmatrix}) = \mathrm{Col}(A)$, wenn $\dim \mathrm{Col}(\begin{pmatrix} A \mid b \end{pmatrix}) = \dim \mathrm{Col}(A)$ ist, d.h. wenn $\mathrm{rk} \begin{pmatrix} A \mid b \end{pmatrix} = \mathrm{rk}\, A$ ist. Dies zeigt die Äquivalenz von Bedingung (c) und Bedingung (d).
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c) und Bedingung (d) äquivalent. □

(3.42) Bemerkung (Eindeutigkeitskriterium für Lösungen linearer Gleichungssysteme)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Für jedes $b \in K^{m \times 1}$ hat das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ höchstens eine Lösung.

(b) Das homogene lineare Gleichungssystem zur Koeffizientenmatrix A hat nur die Lösung 0.

(c) Es gibt ein $b \in K^{m \times 1}$ derart, dass das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ genau eine Lösung hat.

(d) Es ist $\mathrm{rk}_K\, A = n$.

Beweis. Nach Bemerkung (3.35) ist $\mathrm{Sol}(A, b) = \varphi_A^{-1}(\{b\})$ für $b \in K^{m \times 1}$. Somit gilt genau dann Bedingung (a), d.h. für jedes $b \in K^{m \times 1}$ hat das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ genau dann höchstens eine Lösung, wenn jede Faser von $\varphi_A \colon K^{n \times 1} \to K^{m \times 1}$ höchstens ein Element besitzt, nach Satz (A.49)(a) also genau dann, wenn φ_A injektiv ist.
Ferner ist $\mathrm{Sol}(A, 0) = \operatorname{Ker} \varphi_A$ nach Korollar (3.36). Folglich gilt genau dann Bedingung (b), d.h. das homogene lineare Gleichungssystem zur Koeffizientenmatrix A hat genau dann nur die Lösung 0, wenn $\operatorname{Ker} \varphi_A = \{0\}$ ist.
Schließlich gilt genau dann Bedingung (c), d.h. es gibt genau dann ein $b \in K^{m \times 1}$ derart, dass das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ genau eine Lösung hat, wenn es ein $b \in K^{m \times 1}$ derart gibt, dass $\varphi_A^{-1}(\{b\})$ einelementig ist.
Die Äquivalenz von Bedingung (a), Bedingung (b) und Bedingung (c) folgt daher aus Korollar (2.16).
Ferner gilt genau dann $\mathrm{Sol}(A, 0) = \{0\}$, wenn $\dim \mathrm{Sol}(A, 0) = 0$ ist. Nach Korollar (3.36) ist jedoch $\dim \mathrm{Sol}(A, 0) = n - \mathrm{rk}\, A$, so dass $\dim \mathrm{Sol}(A, 0) = 0$ äquivalent zu $\mathrm{rk}\, A = n$ ist. Dies zeigt die Äquivalenz von Bedingung (b) und Bedingung (d).
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c) und Bedingung (d) äquivalent. □

Anwendung: Interpolationen

Als Anwendung des Matrixkalküls studieren wir das sogenannte Interpolationsproblem: Zu gegebenen Paaren (a_1, b_1), …, (a_m, b_m) wird eine Abbildung f derart gesucht, dass $\{a_1, \dots, a_m\}$ eine (in der Regel echte) Teilmenge von $\mathrm{Source}\, f$ ist und $f(a_i) = b_i$ für $i \in [1, m]$ gilt. Hierbei interessiert man sich üblicherweise nicht für alle möglichen Abbildungen, sondern nur für solche mit gewissen Zusatzeigenschaften, welche aus dem Anwendungsproblem naheliegend sind. So wäre etwa Stetigkeit oder Differenzierbarkeit eine sinnvolle Forderung. Oft betrachtet man sogar noch speziellere Teilmengen von Abbildungen wie etwa Polynomfunktionen oder trigonometrische Funktionen.
Zum Studium von linearen Gleichungssystemen hat es sich als fruchtbar erwiesen, mit der Spalteninterpretation einen geeigneten Vektorraumhomomorphismus zu betrachten und die aus Abschnitt 2 bekannten Konzepte und Resultate anzuwenden. Bei Interpolationen gehen wir analog vor und betrachten die folgende naheliegende Abbildung:

(3.43) Notation. Es seien eine Menge X, $m \in \mathbb{N}_0$ und $a \in X^m$ gegeben. Wir schreiben

$$\varepsilon_a \colon \mathrm{Map}(X, K) \to K^m,\ f \mapsto (f(a_1), \dots, f(a_m)).$$

(3.44) Bemerkung. Es seien eine Menge X, $m \in \mathbb{N}_0$ und $a \in X^m$ gegeben. Die Abbildung

$$\varepsilon_a \colon \mathrm{Map}(X, K) \to K^m,\ f \mapsto (f(a_1), \dots, f(a_m))$$

ist ein K-Vektorraumhomomorphismus.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.45) Definition (Interpolation)**.** Es seien eine Menge X, eine Teilmenge S von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$ und $a \in X^m$, $b \in K^m$ gegeben. Die *Menge der Interpolationen* zu (a, b) in S ist definiert als

$$\mathrm{Intpol}_S(a, b) := (\varepsilon_a|_S)^{-1}(\{b\}).$$

Ein Element von $\mathrm{Intpol}_S(a, b)$ wird *Interpolation* (oder *Interpolationsfunktion*) zu (a, b) in S genannt.

(3.46) Bemerkung. Es seien eine Menge X, eine Teilmenge S von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$ und $a \in X^m$, $b \in K^m$ gegeben. Die Menge der Interpolationen zu (a, b) in S ist gegeben durch

$$\mathrm{Intpol}_S(a, b) = \{f \in S \mid f(a_i) = b_i \text{ für } i \in [1, m]\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.47) Beispiel. Es sind $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 - 3x + 2$ und $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^3 - 2x^2 - x + 2$ Interpolationen zu $((0, 1, 2), (2, 0, 0))$ in $\{\mathbb{R} \to \mathbb{R},\ x \mapsto f(x) \mid f \in \mathbb{R}[X]\}$.

Beweis. Wegen

$$\begin{aligned}
&0^2 - 3 \cdot 0 + 2 = 2,\\
&1^2 - 3 \cdot 1 + 2 = 0,\\
&2^2 - 3 \cdot 2 + 2 = 0
\end{aligned}$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 - 3x + 2$ eine Interpolation zu $((0, 1, 2), (2, 0, 0))$ in $\{\mathbb{R} \to \mathbb{R},\ x \mapsto f(x) \mid f \in \mathbb{R}[X]\}$.
Wegen

$$\begin{aligned}
&0^3 - 2 \cdot 0^2 - 0 + 2 = 2,\\
&1^3 - 2 \cdot 1^2 - 1 + 2 = 0,\\
&2^3 - 2 \cdot 2^2 - 2 + 2 = 0
\end{aligned}$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^3 - 2x^2 - x + 2$ eine Interpolation zu $((0, 1, 2), (2, 0, 0))$ in $\{\mathbb{R} \to \mathbb{R},\ x \mapsto f(x) \mid f \in \mathbb{R}[X]\}$. □

(3.48) Beispiel. Es sei $S \subseteq \mathrm{Map}(\mathbb{R}, \mathbb{R})$ gegeben durch

$$S = \{x \mapsto x + 1,\ x \mapsto x^2 + x + 2,\ x \mapsto 2^x,\ x \mapsto 2^{x+1},\ x \mapsto \sin(x)\}$$

Dann ist

$$\mathrm{Intpol}_S((0, 1, 2), (2, 4, 8)) = \{x \mapsto x^2 + x + 2,\ x \mapsto 2^{x+1}\}.$$

Beweis. Wegen

$$1 + 1 = 2 \neq 4$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto x + 1$ keine Interpolation zu $((0, 1, 2), (2, 4, 8))$ in S.
Wegen

$$\begin{aligned}
&0^2 + 0 + 2 = 2,\\
&1^2 + 1 + 2 = 4,
\end{aligned}$$

$$2^2 + 2 + 2 = 8$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 + x + 2$ eine Interpolation zu $((0, 1, 2), (2, 4, 8))$ in S.
Wegen

$$2^0 = 1 \neq 2$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto 2^x$ keine Interpolation zu $((0, 1, 2), (2, 4, 8))$ in S.
Wegen

$$2^{0+1} = 2,$$
$$2^{1+1} = 4,$$
$$2^{2+1} = 8$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto 2^{x+1}$ eine Interpolation zu $((0, 1, 2), (2, 4, 8))$ in S.
Wegen

$$\sin(0) = 0 \neq 2$$

ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto \sin(x)$ keine Interpolation zu $((0, 1, 2), (2, 4, 8))$ in S.
Insgesamt ist

$$\mathrm{Intpol}_S((0, 1, 2), (2, 4, 8)) = \{x \mapsto x^2 + x + 2,\ x \mapsto 2^{x+1}\}.$$ □

(3.49) Bemerkung. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$ und $a \in X^m$ gegeben. Dann ist

$$\mathrm{Intpol}_U(a, 0) = \mathrm{Ker}\, \varepsilon_a|_U.$$

Insbesondere ist $\mathrm{Intpol}_U(a, 0)$ ein K-Untervektorraum von U.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.50) Bemerkung. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$, $a \in X^m$, $b \in K^m$ und $f \in \mathrm{Intpol}_U(a, b)$ gegeben. Dann ist

$$\mathrm{Intpol}_U(a, 0) \to \mathrm{Intpol}_U(a, b),\ u \mapsto f + u$$

eine wohldefinierte Bijektion. Insbesondere ist

$$\mathrm{Intpol}_U(a, b) = f + \mathrm{Intpol}_U(a, 0).$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Ist der Untervektorraum, in welchem wir Interpolationen betrachten, endlichdimensional, so lassen sich Eigenschaften von Interpolationen aus Eigenschaften von Darstellungsmatrizen herleiten, wie wir im Folgenden sehen werden. Sind Interpolationen eindeutig, so ist die Endlichdimensionalität eine notwendige Eigenschaft, siehe das Eindeutigkeitskriterium für Interpolationen (3.55).

(3.51) Bemerkung. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \ldots, s_n)$ von U und $a \in X^m$ gegeben. Die Darstellungsmatrix von $\varepsilon_a|_U \colon U \to K^m$ zu den Basen s und e ist gegeben durch

$$\mathrm{M}_{\mathrm{e},s}(\varepsilon_a|_U) = \begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.52) Korollar. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von U und $a \in X^m$ gegeben. Dann ist $\mathrm{Intpol}_U(a, 0)$ endlichdimensional mit

$$\dim_K \mathrm{Intpol}_U(a, 0) = n - \mathrm{rk}_K \begin{pmatrix} s_1(a_1) & \dots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \dots & s_n(a_m) \end{pmatrix}$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.53) Korollar. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von U, $a \in X^m$ und $b \in K^m$ gegeben. Dann gilt

$$\kappa_s(\mathrm{Intpol}_U(a, b)) = \mathrm{Sol}(\begin{pmatrix} s_1(a_1) & \dots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \dots & s_n(a_m) \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})$$

und

$$\mathrm{Intpol}_U(a, b) = \{ \sum_{j \in [1,n]} c_j s_j \mid c \in \mathrm{Sol}(\begin{pmatrix} s_1(a_1) & \dots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \dots & s_n(a_m) \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.54) Korollar (Interpolationskriterium)**.** Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von U, $a \in X^m$ und $b \in K^m$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es gibt eine Interpolation zu (a, b) in U.

(b) Das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix

$$\left(\begin{array}{ccc|c} s_1(a_1) & \dots & s_n(a_1) & b_1 \\ \vdots & & \vdots & \vdots \\ s_1(a_m) & \dots & s_n(a_m) & b_m \end{array}\right)$$

besitzt eine Lösung.

(c) Es ist b eine Linearkombination von $((s_1(a_1), \dots, s_1(a_m)), \dots, (s_n(a_1), \dots, s_n(a_m)))$.

(d) Es ist

$$\begin{aligned} &\langle (s_1(a_1), \dots, s_1(a_m)), \dots, (s_n(a_1), \dots, s_n(a_m)), b \rangle \\ &= \langle (s_1(a_1), \dots, s_1(a_m)), \dots, (s_n(a_1), \dots, s_n(a_m)) \rangle. \end{aligned}$$

(e) Es ist

$$\mathrm{rk}_K \left(\begin{array}{ccc|c} s_1(a_1) & \dots & s_n(a_1) & b_1 \\ \vdots & & \vdots & \vdots \\ s_1(a_m) & \dots & s_n(a_m) & b_m \end{array}\right) = \mathrm{rk}_K \begin{pmatrix} s_1(a_1) & \dots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \dots & s_n(a_m) \end{pmatrix}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.55) Proposition (Eindeutigkeitskriterium für Interpolationen)**.** Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$ und $a \in X^m$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Für jedes $b \in K^m$ gibt es höchstens eine Interpolation zu (a, b) in U.

(b) Für jedes $f \in U$ gilt: Wenn $a_1, \ldots, a_m$ Nullstellen ([2]) von f sind, dann ist $f = 0$.

(c) Es gibt ein $b \in K^m$ derart, dass es genau eine Interpolation zu (a, b) in U gibt.

(d) Der K-Untervektorraum U ist endlichdimensional und für jedes $n \in \mathbb{N}_0$ und jede Basis $s = (s_1, \ldots, s_n)$ von U ist

$$\operatorname{rk}_K \begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix} = n.$$

(e) Es gibt ein $n \in [0, m]$ und eine Basis $s = (s_1, \ldots, s_n)$ von U mit

$$\operatorname{rk}_K \begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix} = n.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(3.56) Korollar. Es seien eine Menge X, ein K-Untervektorraum U von $\operatorname{Map}(X, K)$, $m \in \mathbb{N}_0$ und $a \in X^m$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Für jedes $b \in K^m$ gibt es genau eine Interpolation zu (a, b) in U.

(b) (i) Für jedes $f \in U$ gilt: Wenn $a_1, \ldots, a_m$ Nullstellen von f sind, dann ist $f = 0$.
 (ii) Es ist U endlichdimensional mit $\dim_K U = m$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Als Illustration wenden wir unsere Theorie auf Polynomfunktionen an:

(3.57) Notation. Für $n \in \mathbb{N}_0$ sei

$$\operatorname{Pol}_{<n}(K, K) := \{K \to K,\, x \mapsto f(x) \mid f \in K[X]_{<n}\}.$$

(3.58) Definition (Vandermondematrix)**.** Es seien $m, n \in \mathbb{N}_0$ und ein m-Tupel $a = (a_1, \ldots, a_m)$ in K gegeben. Wir nennen

$$\mathrm{V}(a) = \mathrm{V}_{<n}(a) = \mathrm{V}_{<n}(a_1, \ldots, a_m) := (a_i^{j-1})_{i \in [1,m], j \in [1,n]} = \begin{pmatrix} 1 & a_1 & \ldots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_m & \ldots & a_m^{n-1} \end{pmatrix}$$

die *Vandermondematrix* zu a.

(3.59) Beispiel (Polynominterpolation mittels Vandermondematrix)**.** Es sei $n \in \mathbb{N}_0$ gegeben und es habe K mindestens n Elemente.

(a) Es sei $s = (s_1, \ldots, s_n)$ in $\operatorname{Pol}_{<n}(K, K)$ gegeben durch

$$s_j \colon K \to K,\, x \mapsto x^{j-1}$$

für $j \in [1, n]$.

 (i) Das n-Tupel s ist eine Basis von $\operatorname{Pol}_{<n}(K, K)$.

 (ii) Für $m \in \mathbb{N}_0$, $a \in K^m$ ist die Darstellungsmatrix von $\varepsilon_a|_{\operatorname{Pol}_{<n}(K,K)} \colon \operatorname{Pol}_{<n}(K, K) \to K^m$ zu den Basen s und e gegeben durch

$$\mathrm{M}_{\mathrm{e},s}(\varepsilon_a|_{\operatorname{Pol}_{<n}(K,K)}) = \mathrm{V}_{<n}(a).$$

[2] Ein $c \in K$ wird *Nullstelle* von f genannt, falls $f(c) = 0$ ist.

(b) Für $m \in \mathbb{N}_0$, $a, b \in K^m$ ist

$$\mathrm{Intpol}_{\mathrm{Pol}_{<n}(K,K)}(a,b) = \{K \to K,\, x \mapsto \sum_{j \in [1,n]} c_j x^{j-1} \mid c \in \mathrm{Sol}(\mathrm{V}_{<n}(a), \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})\}.$$

(c) Es seien $m \in \mathbb{N}_0$ und $a, b \in K^m$ so gegeben, dass die Einträge von a verschieden sind. Die folgenden Bedingungen sind äquivalent.

 (i) Es gibt eine Interpolation zu (a, b) in $\mathrm{Pol}_{<n}(K, K)$.

 (ii) Das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix

$$\left(\begin{array}{c|c} & b_1 \\ \mathrm{V}_{<n}(a) & \vdots \\ & b_m \end{array}\right)$$

 besitzt eine Lösung.

 (iii) Es ist b eine Linearkombination von $((a_1^0, \ldots, a_m^0), \ldots, (a_1^{n-1}, \ldots, a_m^{n-1}))$.

 (iv) Es ist $\langle (a_1^0, \ldots, a_m^0), \ldots, (a_1^{n-1}, \ldots, a_m^{n-1}), b \rangle = \langle (a_1^0, \ldots, a_m^0), \ldots, (a_1^{n-1}, \ldots, a_m^{n-1}) \rangle$.

 (v) Es ist

$$\mathrm{rk}_K \left(\begin{array}{c|c} & b_1 \\ \mathrm{V}_{<n}(a) & \vdots \\ & b_m \end{array}\right) = \mathrm{rk}_K \mathrm{V}_{<n}(a).$$

(d) Es seien $m \in \mathbb{N}_0$ und $a \in K^m$ so gegeben, dass die Einträge von a verschieden sind. Die folgenden Bedingungen sind äquivalent.

 (i) Für jedes $b \in K^m$ gibt es höchtens eine Interpolation zu (a, b) in $\mathrm{Pol}_{<n}(K, K)$.

 (ii) Für jedes $f \in \mathrm{Pol}_{<n}(K, K)$ gilt: Wenn $a_1, \ldots, a_m$ Nullstellen von f sind, dann ist $f = 0$.

 (iii) Es gibt ein $b \in K^m$ derart, dass es genau eine Interpolation zu (a, b) in $\mathrm{Pol}_{<n}(K, K)$ gibt.

 (iv) Es ist

$$\mathrm{rk}_K \mathrm{V}_{<n}(a) = n.$$

 (v) Es ist $m \geq n$.

(e) Es seien $m \in \mathbb{N}_0$ und $a \in K^m$ so gegeben, dass die Einträge von a verschieden sind. Für jedes $b \in K^m$ gibt es genau dann genau eine Interpolation zu (a, b) in $\mathrm{Pol}_{<n}(K, K)$, wenn $m = n$ ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Äquivalenz und Ähnlichkeit von Matrizen

Die Basiswechselformel (3.18)(b) motiviert folgende Terminologie, siehe Bemerkung (3.63).

(3.60) Definition (Äquivalenz von Matrizen, Ähnlichkeit von Matrizen)**.**

(a) Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $A, B \in K^{m \times n}$ sagen wir, dass A *äquivalent* zu B ist, wenn es $P \in \mathrm{GL}_n(K)$, $Q \in \mathrm{GL}_m(K)$ mit

$$Q^{-1} A P = B$$

gibt.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Für $A, B \in K^{n \times n}$ sagen wir, dass A *ähnlich* zu B ist, wenn es ein $P \in \mathrm{GL}_n(K)$ mit

$$P^{-1}AP = B$$

gibt.

(3.61) Beispiel.

(a) Es seien $A, B \in \mathbb{Q}^{2 \times 3}$ gegeben durch

$$A = \begin{pmatrix} 15 & 13 & -10 \\ 15 & 11 & -9 \end{pmatrix}, \, B = \begin{pmatrix} 2 & -1 & 2 \\ 1 & 0 & -2 \end{pmatrix}.$$

Dann ist A äquivalent zu B.

(b) Es seien $A, B \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} -2 & 9 \\ -1 & 4 \end{pmatrix}, \, B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Dann ist A ähnlich zu B.

Beweis.

(a) Es seien $P \in \mathbb{Q}^{3 \times 3}$, $Q \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$P = \begin{pmatrix} -1 & 1 & 3 \\ -1 & 2 & 5 \\ -3 & 4 & 12 \end{pmatrix}, \, Q = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}.$$

Dann ist $P \in \mathrm{GL}_3(\mathbb{Q})$, $Q \in \mathrm{GL}_2(\mathbb{Q})$ mit

$$P^{-1} = \begin{pmatrix} -4 & 0 & 1 \\ 3 & 3 & -2 \\ -2 & -1 & 1 \end{pmatrix}, \, Q^{-1} = \begin{pmatrix} 3 & -4 \\ 1 & -1 \end{pmatrix}$$

und es gilt

$$Q^{-1}AP = \begin{pmatrix} 3 & -4 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 15 & 13 & -10 \\ 15 & 11 & -9 \end{pmatrix} \begin{pmatrix} -1 & 1 & 3 \\ -1 & 2 & 5 \\ -3 & 4 & 12 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 2 \\ 1 & 0 & -2 \end{pmatrix} = B.$$

Folglich ist A äquivalent zu B.

(b) Es seien $P \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$P = \begin{pmatrix} 3 & -4 \\ 1 & -1 \end{pmatrix}.$$

Dann ist $P \in \mathrm{GL}_2(\mathbb{Q})$ mit

$$P^{-1} = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}$$

und es gilt

$$P^{-1}AP = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} -2 & 9 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 3 & -4 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = B.$$

Folglich ist A ähnlich zu B. □

(3.62) Bemerkung.

(a) Es seien $m, n \in \mathbb{N}_0$ gegeben. Äquivalenz von Matrizen ist eine Äquivalenzrelation auf $K^{m \times n}$.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Ähnlichkeit von Matrizen ist eine Äquivalenzrelation auf $K^{n \times n}$.

Beweis.

(a) Es seien $A, B, C \in K^{m \times n}$ so gegeben, dass A äquivalent zu B und B äquivalent zu C ist. Dann gibt es $P, R \in \mathrm{GL}_n(K)$, $Q, S \in \mathrm{GL}_m(K)$ mit $Q^{-1}AP = B$ und $S^{-1}BR = C$. Es folgt

$$(QS)^{-1}APR = S^{-1}Q^{-1}APR = S^{-1}BR = C,$$

und da $PR \in \mathrm{GL}_n(K)$ und $QS \in \mathrm{GL}_m(K)$ gilt, ist somit A äquivalent zu C. Folglich ist Äquivalenz von Matrizen transitiv.

Für alle $A \in K^{m \times n}$ gilt $\mathrm{E}_m^{-1} A \mathrm{E}_n = A$, und da $\mathrm{E}_n \in \mathrm{GL}_n(K)$ und $\mathrm{E}_m \in \mathrm{GL}_m(K)$ gilt, ist somit A äquivalent zu A. Folglich ist Äquivalenz von Matrizen reflexiv.

Es seien $A, B \in K^{m \times n}$ so gegeben, dass A äquivalent zu B ist. Dann gibt es $P \in \mathrm{GL}_n(K)$, $Q \in \mathrm{GL}_m(K)$ mit $Q^{-1}AP = B$. Es folgt

$$A = QBP^{-1} = (Q^{-1})^{-1}BP^{-1},$$

und da $P^{-1} \in \mathrm{GL}_n(K)$ und $Q^{-1} \in \mathrm{GL}_m(K)$ gilt, ist somit B äquivalent zu A. Folglich ist Äquivalenz von Matrizen symmetrisch.

Insgesamt ist Äquivalenz von Matrizen eine Äquivalenzrelation auf $K^{m \times n}$.

(b) Es seien $A, B, C \in K^{n \times n}$ so gegeben, dass A ähnlich zu B und B ähnlich zu C ist. Dann gibt es $P, Q \in \mathrm{GL}_n(K)$ mit $P^{-1}AP = B$ und $Q^{-1}BQ = C$. Es folgt

$$(PQ)^{-1}APQ = Q^{-1}P^{-1}APQ = Q^{-1}BQ = C,$$

und da $PQ \in \mathrm{GL}_n(K)$ gilt, ist somit A ähnlich zu C. Folglich ist Ähnlichkeit von Matrizen transitiv.

Für alle $A \in K^{n \times n}$ gilt $\mathrm{E}_n^{-1} A \mathrm{E}_n = A$, und da $\mathrm{E}_n \in \mathrm{GL}_n(K)$ gilt, ist somit A ähnlich zu A. Folglich ist Ähnlichkeit von Matrizen reflexiv.

Es seien $A, B \in K^{n \times n}$ so gegeben, dass A ähnlich zu B ist. Dann gibt es ein $P \in \mathrm{GL}_n(K)$ mit $P^{-1}AP = B$. Es folgt

$$A = PBP^{-1} = (P^{-1})^{-1}BP^{-1},$$

und da $P^{-1} \in \mathrm{GL}_n(K)$ gilt, ist somit B ähnlich zu A. Folglich ist Ähnlichkeit von Matrizen symmetrisch.

Insgesamt ist Ähnlichkeit von Matrizen eine Äquivalenzrelation auf $K^{n \times n}$. □

(3.63) Bemerkung.

(a) Es seien ein K-Vektorraumhomomorphismus $\varphi \colon V \to W$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \ldots, s_n)$ von V und eine Basis $t = (t_1, \ldots, t_m)$ von W gegeben. Die Äquivalenzklasse von $\mathrm{M}_{t,s}(\varphi)$ bzgl. Äquivalenz von Matrizen ist gegeben durch

$$[\mathrm{M}_{t,s}(\varphi)] = \{\mathrm{M}_{t',s'}(\varphi) \mid s' \text{ ist eine Basis von } V,\ t' \text{ ist eine Basis von } W\}.$$

(b) Es seien ein K-Vektorraumendomorphismus $\varphi \colon V \to V$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \ldots, s_n)$ von V gegeben. Die Äquivalenzklasse von $\mathrm{M}_{s,s}(\varphi)$ bzgl. Ähnlichkeit von Matrizen ist gegeben durch

$$[\mathrm{M}_{s,s}(\varphi)] = \{\mathrm{M}_{s',s'}(\varphi) \mid s' \text{ ist eine Basis von } V\}.$$

Beweis.

(a) Zunächst sei $B \in [\mathrm{M}_{t,s}(\varphi)]$, d.h. es sei $B \in K^{m \times n}$ äquivalent zu $\mathrm{M}_{t,s}(\varphi)$. Dann gibt es $P \in \mathrm{GL}_n(K)$ und $Q \in \mathrm{GL}_m(K)$ mit $B = Q^{-1} \mathrm{M}_{t,s}(\varphi) P$. Nach Korollar (3.32) ist $s' := (\kappa_s^{-1}(P_{-,1}), \dots, \kappa_s^{-1}(P_{-,n}))$ eine Basis von V mit $\mathrm{M}_{s,s'}(\mathrm{id}_V) = P$ und $t' := (\kappa_t^{-1}(Q_{-,1}), \dots, \kappa_t^{-1}(Q_{-,m}))$ eine Basis von W mit $\mathrm{M}_{t,t'}(\mathrm{id}_W) = Q$. Mit der Basiswechselformel (3.18)(b) erhalten wir

$$B = Q^{-1} \mathrm{M}_{t,s}(\varphi) P = (\mathrm{M}_{t,t'}(\mathrm{id}_W))^{-1} \mathrm{M}_{t,s}(\varphi) \mathrm{M}_{s,s'}(\mathrm{id}_V) = \mathrm{M}_{t',s'}(\varphi).$$

Sind umgekehrt eine Basis s' von V und eine Basis t' von W gegeben, so ist $\mathrm{M}_{t,s}(\varphi)$ äquivalent zu

$$(\mathrm{M}_{t,t'}(\mathrm{id}_W))^{-1} \mathrm{M}_{t,s}(\varphi) \mathrm{M}_{s,s'}(\mathrm{id}_V) = \mathrm{M}_{t',s'}(\varphi)$$

nach der Basiswechselformel (3.18)(b).

(b) Zunächst sei $B \in [\mathrm{M}_{s,s}(\varphi)]$, d.h. es sei $B \in K^{n \times n}$ ähnlich zu $\mathrm{M}_{s,s}(\varphi)$. Dann gibt es ein $P \in \mathrm{GL}_n(K)$ mit $B = P^{-1} \mathrm{M}_{s,s}(\varphi) P$. Nach Korollar (3.32) ist $s' = (\kappa_s^{-1}(P_{-,1}), \dots, \kappa_s^{-1}(P_{-,n}))$ eine Basis von V mit $\mathrm{M}_{s,s'}(\mathrm{id}_V) = P$. Mit der Basiswechselformel (3.18)(b) erhalten wir

$$B = P^{-1} \mathrm{M}_{s,s}(\varphi) P = (\mathrm{M}_{s,s'}(\mathrm{id}_V))^{-1} \mathrm{M}_{s,s}(\varphi) \mathrm{M}_{s,s'}(\mathrm{id}_V) = \mathrm{M}_{s',s'}(\varphi).$$

Ist umgekehrt eine Basis s' von V gegeben, so ist $\mathrm{M}_{s,s}(\varphi)$ ähnlich zu

$$(\mathrm{M}_{s,s'}(\mathrm{id}_V))^{-1} \mathrm{M}_{s,s}(\varphi) \mathrm{M}_{s,s'}(\mathrm{id}_V) = \mathrm{M}_{s',s'}(\varphi)$$

nach der Basiswechselformel (3.18)(b). □

Unser nächstes Ziel ist es, Matrizen bis auf Äquivalenz zu charakterisieren. Wir gehen also der Frage nach, wie für $m, n \in \mathbb{N}_0$ die Elemente der Quotientenmenge von $K^{m \times n}$ modulo Äquivalenz von Matrizen aussehen.

Klassifikation von Matrizen modulo Äquivalenz

Mit Hilfe des Rangs von Matrizen können wir nun Matrizen bis auf Äquivalenz charakterisieren. Zunächst definieren wir Normalformen für Matrizen von gegebenem Rang:

(3.64) Definition (Quasieinheitsmatrix)**.** Es seien $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$ gegeben. Die Matrix $\mathrm{Q}_r \in K^{m \times n}$ gegeben durch

$$\mathrm{Q}_r = \begin{pmatrix} \mathrm{E}_r & 0 \\ 0 & 0 \end{pmatrix}$$

heißt *Quasieinheitsmatrix.*

(3.65) Beispiel. Es ist $\mathrm{Q}_2 \in \mathbb{Q}^{3 \times 4}$ gegeben durch

$$\mathrm{Q}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(3.66) Bemerkung. Es seien $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$ gegeben. Der Rang von $\mathrm{Q}_r \in K^{m \times n}$ ist

$$\mathrm{rk}_K \mathrm{Q}_r = r.$$

(3.67) Bemerkung. Es seien ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$, eine Basis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W so gegeben, dass $(s_{r+1}, \dots, s_n)$ eine Basis von $\mathrm{Ker}\,\varphi$ und $t_j = \varphi(s_j)$ für $j \in [1, r]$ ist. Dann ist

$$\mathrm{M}_{t,s}(\varphi) = \mathrm{Q}_r.$$

Beweis. Für $j \in [1, n]$ gilt

$$\kappa_t(\varphi(s_j)) = \begin{cases} \mathrm{e}_j, & \text{falls } j \in [1, r], \\ 0, & \text{falls } j \in [r+1, n]. \end{cases}$$

Es folgt

$$\begin{aligned} \mathrm{M}_{t,s}(\varphi) &= \begin{pmatrix} \kappa_t(\varphi(s_1)) & \dots & \kappa_t(\varphi(s_r)) & \kappa_t(\varphi(s_{r+1})) & \dots & \kappa_t(\varphi(s_n)) \end{pmatrix} = \begin{pmatrix} \mathrm{e}_1 & \dots & \mathrm{e}_r & 0 & \dots & 0 \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{E}_r & 0 \\ 0 & 0 \end{pmatrix} = \mathrm{Q}_r. \end{aligned}$$

□

(3.68) Korollar. Es seien $m, n \in \mathbb{N}_0$ und ein $A \in K^{m \times n}$ gegeben. Dann ist A äquivalent zu $\mathrm{Q}_{\mathrm{rk}_K A}$.

Beweis. Da $K^{n\times 1}$ endlichdimensional mit $\dim K^{n\times 1} = n$ ist, gibt es nach dem Rangsatz (2.28) ein $r \in \mathbb{N}_0$ mit $r \leq n$ und eine Basis $s = (s_1, \dots, s_n)$ von $K^{n\times 1}$ derart, dass $(\varphi(s_1), \dots, \varphi(s_r))$ eine Basis von $\operatorname{Im} \varphi_A$ und $(s_{r+1}, \dots, s_n)$ eine Basis von $\operatorname{Ker} \varphi_A$ ist. Insbesondere ist $\operatorname{Im} \varphi_A$ endlichdimensional und es gilt

$$\operatorname{rk} A = \operatorname{rk} \varphi_A = \dim(\operatorname{Im} \varphi_A) = r.$$

Da auch $K^{m\times 1}$ endlichdimensional ist, gibt es nach Korollar (1.55) eine Basis $t = (t_1, \dots, t_m)$ von $K^{m\times 1}$ mit $t_j = \varphi_A(s_j)$ für $j \in [1, r]$.
Nach Bemerkung (3.63)(a) ist $A = \mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)$ äquivalent zu $\mathrm{M}_{t,s}(\varphi_A)$ und nach Bemerkung (3.67) gilt

$$\mathrm{M}_{t,s}(\varphi_A) = \mathrm{Q}_r = \mathrm{Q}_{\mathrm{rk}\, A}.$$

□

(3.69) Satz. Es seien $m, n \in \mathbb{N}_0$ und $A, B \in K^{m\times n}$ gegeben. Genau dann ist A äquivalent zu B, wenn

$$\operatorname{rk}_K A = \operatorname{rk}_K B$$

ist.

Beweis. Zunächst sei A äquivalent zu B, so dass es $P \in \mathrm{GL}_n(K)$, $Q \in \mathrm{GL}_m(K)$ mit $Q^{-1}AP = B$ gibt. Nach Korollar (3.25) und Bemerkung (3.26) ist

$$\operatorname{Col}(A) = \operatorname{Col}(AP) \cong \operatorname{Col}(Q^{-1}AP) = \operatorname{Col}(B),$$

also

$$\operatorname{rk} A = \dim \operatorname{Col}(A) = \dim \operatorname{Col}(B) = \operatorname{rk} B$$

nach Satz (2.32).
Nun sei umgekehrt $\operatorname{rk} A = \operatorname{rk} B$. Nach Korollar (3.68) ist A äquivalent zu $\mathrm{Q}_{\mathrm{rk}\, A}$ und B äquivalent zu $\mathrm{Q}_{\mathrm{rk}\, B}$. Da aber $\operatorname{rk} A = \operatorname{rk} B$ ist, folgt $\mathrm{Q}_{\mathrm{rk}\, A} = \mathrm{Q}_{\mathrm{rk}\, B}$, und da Äquivalenz von Matrizen nach Bemerkung (3.62)(a) eine Äquivalenzrelation auf $K^{m\times n}$ ist, ist A folglich äquivalent zu B. □

(3.70) Korollar (Klassifikation von Matrizen modulo Äquivalenz)**.** Es seien $m, n, r \in \mathbb{N}_0$ und $A \in K^{m\times n}$ gegeben. Genau dann ist

$$\operatorname{rk}_K A = r,$$

wenn A äquivalent zu Q_r ist.

Beweis. Es ist $\operatorname{rk} \mathrm{Q}_r = r$. Somit gilt genau dann $\operatorname{rk} A = r$, wenn $\operatorname{rk} A = \operatorname{rk} \mathrm{Q}_r$ ist, was nach Satz (3.69) aber genau dann gilt, wenn A äquivalent zu Q_r ist. □

(3.71) Korollar. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m\times n}$ gegeben. Dann ist

$$\operatorname{rk}_K A = \operatorname{rk}_K \psi_A,$$

wobei $\psi_A \colon K^{1\times m} \to K^{1\times n}$, $x \mapsto xA$.

Beweisskizze. Es ist $\operatorname{rk} A = \operatorname{rk} \varphi_A = \dim(\operatorname{Im} \varphi_A)$ mit $\varphi_A \colon K^{n \times 1} \to K^{m \times 1}$, also der Multiplikation von A an Spalten. Die Abbildung ψ_A ist die Multiplikation von A an Zeilen, die *Zeileninterpretation* von A. Jede Aussage, welche wir in diesem Abschnitt für φ_A bewiesen haben, besitzt ein Analogon für Zeilen statt Spalten. Das Analogon zu Korollar (3.70) besagt: Für $r \in \mathbb{N}_0$ ist genau dann $\operatorname{rk} \psi_A = r$, wenn A äquivalent zu Q_r ist. Nach Korollar (3.68) ist aber A äquivalent zu $\mathrm{Q}_{\operatorname{rk} A}$, so dass das Analogon zu Korollar (3.70) bereits

$$\operatorname{rk} \psi_A = \operatorname{rk} A$$

impliziert. □

(3.72) Korollar. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Dann ist

$$\operatorname{rk}_K A^{\mathrm{tr}} = \operatorname{rk}_K A.$$

Darstellung von Spaltenräumen als Lösungsräume

Es seien $m, n \in \mathbb{N}_0$ und $B \in K^{m \times n}$ gegeben. Wenn wir ein homogenes lineares Gleichungssystem zur Koeffizientenmatrix B lösen, so bestimmen wir eine Basis $(t_1, \dots, t_p)$ des Untervektorraums $\operatorname{Sol}(B, 0)$ von $K^{n \times 1}$, wobei $p = n - \operatorname{rk} B$, vgl. Korollar (3.36). Schreiben wir diese Basis als Spalten in eine Matrix A, so gilt

$$\operatorname{Sol}(B, 0) = \langle t_1, \dots, t_p \rangle = \langle A_{-,1}, \dots, A_{-,p} \rangle = \operatorname{Col}(A)$$

nach Bemerkung (3.21), d.h. wir haben den Lösungsraum als Spaltenraum geschrieben.
Wir wollen nun das umgekehrte Problem betrachten, d.h. einen Spaltenraum als Lösungsraum schreiben. Hierzu seien $n, p \in \mathbb{N}_0$ und $A \in K^{n \times p}$ gegeben. Da der Spaltenraum das Bild und der Lösungsraum der Kern eines Homomorphismus sind, können wir das in Abschnitt 2 unter Verwendung von Proposition (2.27) beschriebene Verfahren verwenden: Zunächst bestimmen wir eine Basis von $\operatorname{Col}(A)$ und ergänzen diese zu einer Basis von $K^{n \times 1}$, d.h. wir bestimmen $r \in \mathbb{N}_0$ und eine Basis $t = (t_1, \dots, t_n)$ von $K^{n \times 1}$ so, dass $(t_1, \dots, t_r)$ eine Basis von $\operatorname{Col}(A)$ ist. Es bezeichne $\psi \colon K^{n \times 1} \to K^{m \times 1}$ den eindeutigen K-Vektorraumhomomorphismus mit $\psi(t_j) = 0$ für $j \in [1, r]$ und $\psi(t_{r+i}) = \mathrm{e}_i$ für $i \in [1, m]$, wobei $m = n - r$. Dann gilt also

$$\mathrm{M}_{\mathrm{e},t}(\psi) = \begin{pmatrix} 0 & \mathrm{E}_m \end{pmatrix}.$$

Für $B \in K^{m \times n}$ gegeben durch $B = \mathrm{M}_{\mathrm{e},\mathrm{e}}(\psi)$ gilt nun $\psi = \varphi_B$ nach Proposition (3.11) und damit

$$\operatorname{Sol}(B, 0) = \operatorname{Ker} \varphi_B = \operatorname{Ker} \psi = \langle t_1, \dots, t_r \rangle = \operatorname{Col}(A)$$

nach Korollar (3.36) und Proposition (2.27). Nach der Basiswechselformel (3.18)(b) und Bemerkung (3.17) ist

$$B = \mathrm{M}_{\mathrm{e},\mathrm{e}}(\psi) = \mathrm{M}_{\mathrm{e},t}(\psi)\, \mathrm{M}_{t,\mathrm{e}}(\mathrm{id}_{K^{n \times 1}}) = \mathrm{M}_{\mathrm{e},t}(\psi)\, (\mathrm{M}_{\mathrm{e},t}(\mathrm{id}_{K^{n \times 1}}))^{-1} = \begin{pmatrix} 0 & \mathrm{E}_m \end{pmatrix} \begin{pmatrix} t_1 & \dots & t_n \end{pmatrix}^{-1}.$$

Dieses Verfahren ist etwas länglich, es liefert allerdings auch etwas mehr als zunächst gewünscht: Es gilt nicht nur $\operatorname{Sol}(B, 0) = \operatorname{Col}(A)$ sondern zusätzlich auch noch $B \begin{pmatrix} t_{r+1} & \dots & t_n \end{pmatrix} = \mathrm{E}_m$. Da wir t erst im Laufe des Verfahrens konstruiert haben, ergibt sich hierdurch eine vergleichsweise hohe Flexibilität hinsichtlich der Konstruktion von B.
Wir stellen nun ein alternatives Verfahren vor, welches in der Regel etwas schneller durchführbar ist. Grundlage dieses Verfahrens ist die Transposition von Matrizen, welche doppelt angewandt wird.

(3.73) Proposition. Es seien $m, n, p \in \mathbb{N}_0$, $A \in K^{n \times p}$ und $B \in K^{m \times n}$ gegeben. Genau dann gilt

$$\operatorname{Sol}(B, 0) = \operatorname{Col}(A),$$

wenn

$$\operatorname{Sol}(A^{\mathrm{tr}}, 0) = \operatorname{Col}(B^{\mathrm{tr}})$$

gilt.

Beweis. Zunächst gelte $\operatorname{Sol}(B,0) = \operatorname{Col}(A)$. Für $k \in [1,p]$ ist dann $A_{-,k} \in \operatorname{Col}(A) = \operatorname{Sol}(B,0)$, es gilt also

$$BA = B\begin{pmatrix} A_{-,1} & \dots & A_{-,p} \end{pmatrix} = \begin{pmatrix} BA_{-,1} & \dots & BA_{-,p} \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 \end{pmatrix} = 0.$$

Nach Proposition (A.148)(a) ist

$$\begin{pmatrix} A^{\mathrm{tr}}(B^{\mathrm{tr}})_{-,1} & \dots & A^{\mathrm{tr}}(B^{\mathrm{tr}})_{-,m} \end{pmatrix} = A^{\mathrm{tr}}B^{\mathrm{tr}} = (BA)^{\mathrm{tr}} = 0^{\mathrm{tr}} = 0,$$

also $(B^{\mathrm{tr}})_{-,i} \in \operatorname{Sol}(A^{\mathrm{tr}},0)$ für $i \in [1,m]$. Da $\operatorname{Sol}(A^{\mathrm{tr}},0)$ ein Untervektorraum von $K^{n\times 1}$ ist, folgt $\operatorname{Col}(B^{\mathrm{tr}}) = \langle (B^{\mathrm{tr}})_{-,1},\dots,(B^{\mathrm{tr}})_{-,m}\rangle \subseteq \operatorname{Sol}(A^{\mathrm{tr}},0)$ nach Bemerkung (3.21), und da $\operatorname{Col}(B^{\mathrm{tr}})$ ein Untervektorraum von $K^{n\times 1}$ ist, folgt $\operatorname{Col}(B^{\mathrm{tr}}) \leq \operatorname{Sol}(A^{\mathrm{tr}},0)$. Nach Korollar (3.36), Korollar (3.72) und Bemerkung (3.28) gilt schließlich

$$\begin{aligned} \dim \operatorname{Sol}(A^{\mathrm{tr}},0) &= n - \operatorname{rk} A^{\mathrm{tr}} = n - \operatorname{rk} A = n - \dim \operatorname{Col}(A) = n - \dim \operatorname{Sol}(B,0) = \operatorname{rk} B = \operatorname{rk} B^{\mathrm{tr}} \\ &= \dim \operatorname{Col}(B^{\mathrm{tr}}), \end{aligned}$$

so dass sich nach Proposition (1.68) bereits $\operatorname{Sol}(A^{\mathrm{tr}},0) = \operatorname{Col}(B^{\mathrm{tr}})$ ergibt.
Gilt umgekehrt $\operatorname{Sol}(A^{\mathrm{tr}},0) = \operatorname{Col}(B^{\mathrm{tr}})$, so auch

$$\operatorname{Sol}(B,0) = \operatorname{Sol}((B^{\mathrm{tr}})^{\mathrm{tr}},0) = \operatorname{Col}((A^{\mathrm{tr}})^{\mathrm{tr}}) = \operatorname{Col}(A).$$

□

Algorithmisch lässt sich Proposition (3.73) wie folgt formulieren:

(3.74) Algorithmus.

- Eingabe: $A \in K^{n\times p}$ für gewisse $n, p \in \mathbb{N}_0$
- Ausgabe: $B \in K^{m\times n}$ für ein $m \in \mathbb{N}_0$ mit $\operatorname{Sol}(B,0) = \operatorname{Col}(A)$ und $\operatorname{rk} B = m$
- Verfahren:

```
function matrixsle(A)
    bestimme Basis (s_1, ..., s_m) von Sol(A^tr, 0);
    return (s_1 ... s_m)^tr;
end function;
```

(3.75) Beispiel. Es sei $A \in \mathbb{Q}^{5\times 2}$ gegeben durch

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

Für $B \in \mathbb{Q}^{3\times 5}$ gegeben durch

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

gilt

$$\operatorname{Col}(A) = \operatorname{Sol}(B,0).$$

Beweis. Nach Bemerkung (3.21) gilt

$$\begin{aligned} \operatorname{Sol}(A^{\mathrm{tr}},0) &= \operatorname{Sol}(\begin{pmatrix} 1 & -2 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}) = \langle \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \rangle = \operatorname{Col}(\begin{pmatrix} 2 & 1 & -3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}) \\ &= \operatorname{Col}(B^{\mathrm{tr}}), \end{aligned}$$

so dass aus Proposition (3.73) bereits $\operatorname{Col}(A) = \operatorname{Sol}(B,0)$ folgt.

□

4 Lineare Kodierungstheorie

In diesem Abschnitt behandeln wir die Grundlagen der linearen Kodierungstheorie, einem Anwendungsgebiet der linearen Algebra. Die Kodierungstheorie beschäftigt sich mit sogenannten fehlererkennenden und fehlerkorrigierenden Codes. Dabei ist ein solcher Code eine Art Sprache, welche zur Übermittlung von Nachrichten über störungsanfällige Kanäle verwandt wird. Das Ziel ist dabei nicht die Verschlüsselung der zu sendenden Nachricht (damit beschäftigt sich die Kryptographie), sondern das Erkennen und ggf. Korrigieren von Fehlern, welche durch Störungen des Kanals hervorgerufen werden können.
Zur Einführung betrachten wir ein Beispiel: Hierbei soll eine Information aus einer Menge an acht verschiedenen Informationen übertragen werden. Wenn wir davon ausgehen, dass sowohl Sender als auch Empfänger alle möglichen Informationen kennen, so können wir die Menge der *Informationswörter* durch $\mathbb{F}_2^3$ modellieren. Die Informationswörter sind also Ausdrücke der Form (x_1, x_2, x_3), wobei $x_1, x_2, x_3 \in \mathbb{F}_2 = \{0, 1\}$, Wörter bestehend aus drei *Bits*. Der Einfachheit wegen schreiben wir in dieser Einführung $x_1x_2x_3$ anstatt (x_1, x_2, x_3). Wir haben also folgende Informationswörter, wobei jedes Wort genau einer Information entspricht:

000, 001, 010, 011, 100, 101, 110, 111

Nehmen wir nun beispielsweise an, der Sender übermittelt das Informationswort 101. Wir wollen davon ausgehen, dass die Übertragung möglicherweise nicht störungsfrei verläuft, so dass die Nachricht bei der Übermittlung verändert werden kann. Unter der Annahme, dass eine Störung bei der Übermittlung die Nachricht an genau einer Stelle verändert, erhält der Empfänger ein Wort, welches nicht der zu übermittelnden Information entspricht, nämlich eines der Wörter 001, 111 oder 100. Schlimmer noch, da sowohl 001 als auch 111 als auch 100 selbst jeweils ein mögliches Informationswort ist, erkennt der Empfänger nicht einmal, dass bei der Übermittlung ein Fehler passiert ist, er erhält schlicht eine falsche Information.
Um dem entgegen zu wirken, kann der Sender seiner Information *redundante Bits*, also zur Übermittlung der Information überflüssige Bits, hinzufügen. Beispielsweise könnte er seinen Informationswörtern jeweils ein einzelnes *Kontrollbit* so hinzufügen, dass die Summe aller Bits (in $\mathbb{F}_2$) gleich 0 wird. Hierdurch würde sich die Länge der möglichen zu übermittelnden Nachrichten auf 4 vergrößern. Konkret ergäben sich die folgenden Wörter:

0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111

Da diese Wörter nun neben der eigentlich zu übermittelnden Information noch Redundanzen erhalten, sprechen wir von *Codewörtern*, den Übergang von einem Informationswort zu einem Codewort nennt man *Kodierung.* Ein *Code* ist die Menge aller Codewörter, also einfach eine (in der Praxis echte) Teilmenge von $\mathbb{F}_2^n$.
Anstatt das Informationswort 101 direkt zu übermitteln, könnte der Sender dieses also zuerst mittels Hinzufügen eines Kontrollbits kodieren und stattdessen 1010 verschicken. Würde nun bei der Übermittlung genau ein Bit verändert, so würde der Empfänger eines der Wörter 0010, 1110, 1000 oder 1011 erhalten. Da keines dieser Wörter ein Codewort ist, würde der Empfänger nun also zumindest erkennen, dass ein Fehler passiert ist, wenn er auch nicht weiß, an welcher Stelle der Fehler passiert ist. Würden hingegen bei der Übertragung zwei Fehler passieren, so würde der Empfänger wieder ein Codewort erhalten, beispielsweise 1100, und die Übermittlungsfehler somit nicht bemerken. Wir sagen, dass der Code (nur) *einen Fehler erkennt.* Für die Anwendungen der Kodierungstheorie muss also sichergestellt sein, dass im Verhältnis zur Länge der zu übermittelnden Codewörter *nicht zu viele* Fehler passieren.
Eine andere Möglichkeit, den Informationswörtern redundante Bits hinzuzufügen, wäre etwa, die Information zu wiederholen. Der Sender könnte beispielsweise jedes Informationswort doppelt senden und würde somit Codewörter der Länge 6 erhalten:

000000, 001001, 010010, 011011, 100100, 101101, 110110, 111111

Würde der Sender nun das Codewort 101101 statt des Informationsworts 101 senden, so würde der Empfänger manchmal zwei Fehler erkennen, etwa beim Empfang des Worts 111001, manchmal jedoch auch nicht, etwa beim Empfang des Worts 111111.
Als nächstes wollen wir die Möglichkeit betrachten, dass der Sender jedes Informationswort sogar dreimal sendet, so dass die Codewörter die Länge 9 erhalten:

000000000, 001001001, 010010010, 011011011, 100100100, 101101101, 110110110, 111111111

Anstatt des Informationsworts 101 würde der Sender nun also das Codewort 101101101 verschicken. Wir wollen weiter annehmen, dass bei der Übertragung höchstens ein Fehler passiert, so dass der Empfänger beispielsweise

das Wort 111101101 erhält. In diesem Fall würde der Empfänger nicht nur sofort erkennen, dass ein Fehler aufgetreten sein muss, er könnte unter der Annahme, dass höchstens ein Fehler passiert ist, sogar Rückschlüsse auf das gesendete Codewort machen: Der Fehler muss an der zweiten Stelle passiert sein, da jede Veränderung an den anderen Stellen des empfangenen Worts nicht zu einem Codewort führt. Oder anders ausgedrückt: Jedes andere Codewort unterscheidet sich an mindestens zwei Stellen vom empfangenen Wort. Wir sagen, dass der Code *einen Fehler korrigiert*. Sowohl den Übergang von einem empfangenen Wort zu einem Codewort als auch den Übergang von einem Codewort zum zugeordneten Informationswort nennt man *Dekodierung*, wobei man in der Kodierungstheorie vor allem den ersten Aspekt studiert.
Der Preis dafür, dass wir bei diesem Code einen Fehler korrigieren können, liegt natürlich in der Länge der verwendeten Codewörter. Das Ziel der Kodierungstheorie ist es, solche Codes zu finden, bei denen die Länge der Codewörter im Verhältnis zur Anzahl der möglichen zu korrigierenden Fehler klein ist (oder umgekehrt, bei der die Anzahl der möglichen zu korrigierenden Fehler im Verhältnis zur Länge der Codewörter möglichst groß ist).
Anstatt nur Wörter mit Einträgen in $\mathbb{F}_2$ zu betrachten, können wir auch allgemeinere Codes definieren. Dabei können wir im Prinzip jede Menge als „Zeichenvorrat" wählen. Wir werden unseren Codes stets einen Körper K zu Grunde legen. Typischerweise wird K endlich sein, was für unsere Theorie jedoch nur an wenigen Stellen eine Rolle spielen wird und deshalb erst einmal nicht angenommen wird. Eine größere Rolle spielt hingegen die Struktur, welche ein Körper mit sich bringt: Die Lineare Algebra kommt in der Kodierungstheorie bei der effizienten Handhabung von Codes ins Spiel – anstatt beliebiger Teilmengen können wir Untervektorräume von K^n für ein $n \in \mathbb{N}_0$ betrachten und so die bisherige Theorie anwenden. Da wir im Laufe des Matrixkalküls den Standpunkt eingenommen haben, mit Spaltenmatrizen statt Tupeln zu arbeiten, werden wir der Bequemlichkeit wegen diesen Standpunkt auch weiterhin annehmen und lineare Codes als Untervektorräume von $K^{n \times 1}$ (statt K^n) definieren. ([3])
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, seien stets ein Körper K und eine nicht-negative ganze Zahl n gegeben.

Blockcodes

Die Wörter eines Codes werden alle stets dieselbe Länge n haben. Aus diesem Grund definieren wir:

(4.1) Definition (Blockcode)**.** Ein *Blockcode* über K der Länge n ist eine nicht-leere Teilmenge C von $K^{n \times 1}$. Ein Element von C wird *Codewort* von C genannt.
Ein *Blockcode* über K ist ein Blockcode über K der Länge m für ein $m \in \mathbb{N}_0$.

(4.2) Beispiel.

(a) Die Teilmenge C von $\mathbb{F}_2^{3 \times 1}$ gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \}$$

ist ein Blockcode über $\mathbb{F}_2$ der Länge 3.

(b) Die Teilmenge C von $\mathbb{F}_2^{6 \times 1}$ gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \}$$

ist ein Blockcode über $\mathbb{F}_2$ der Länge 6.

[3] In der Literatur zur Kodierungstheorie wird meist mit Zeilenmatrizen gearbeitet, d.h. lineare Codes sind hier Untervektorräume von $K^{1 \times n}$ für ein $n \in \mathbb{N}_0$.

(c) Die Teilmenge C von $\mathbb{F}_2^{5\times 1}$ gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \}$$

ist ein Blockcode über $\mathbb{F}_2$ der Länge 5.

Hamming-Gewicht und Hamming-Abstand

In obiger Einführung haben wir gesehen, dass die Erkennung und Korrektur von Fehlern davon abhängt, wie „weit“ ein empfangenes Wort von einem Codewort entfernt ist. Dabei ist die Weite im Sinne der Anzahl der bei der Übertragung aufgetretenen Fehler zu verstehen. Mit anderen Worten, die Weite entspricht der „Größe“ der Differenz von empfangenem Wort und Codewort, des *Fehlerworts*. Dabei sehen wir ein Fehlerwort als umso größer an, je mehr Fehler passiert sind.
Wir werden diesen Sachverhalt nun präzisieren und eine entsprechende Terminologie einführen.

(4.3) Definition (Hamming-Gewicht)**.** Für $a \in K^{n\times 1}$ heißt

$$\mathrm{wt}(a) := |\{i \in [1,n] \mid a_i \neq 0\}|$$

das *Hamming-Gewicht* von a.

Das Hamming-Gewicht eines Worts entspricht also gerade dem Maß für die Größe dieses Worts, wenn wir es als Fehlerwort ansehen.

(4.4) Beispiel. In $\mathbb{F}_2^{4\times 1}$ ist

$$\mathrm{wt}(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}) = 2.$$

Beweis. Es ist

$$\mathrm{wt}(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}) = |\{2,3\}| = 2.$$ □

(4.5) Proposition.

(a) *Positive Definitheit.* Für $a \in K^{n\times 1}$ gilt

$$\mathrm{wt}(a) \geq 0$$

und genau dann $\mathrm{wt}(a) = 0$, wenn $a = 0$ ist.

(b) *Symmetrie.* Für $a \in K^{n\times 1}$ gilt

$$\mathrm{wt}(-a) = \mathrm{wt}(a).$$

(c) *Dreiecksungleichung.* Für $a, b \in K^{n\times 1}$ gilt

$$\mathrm{wt}(a+b) \leq \mathrm{wt}(a) + \mathrm{wt}(b).$$

Beweis.

(a) Es sei $a \in K^{n \times 1}$ gegeben. Dann ist $\mathrm{wt}(a) = |\{i \in [1,n] \mid a_i \neq 0\}| \geq 0$. Ferner gilt genau dann $\mathrm{wt}(a) = 0$, wenn $\{i \in [1,n] \mid a_i \neq 0\} = \emptyset$ ist, was aber gerade bedeutet, dass $a_i = 0$ für alle $i \in [1,n]$ ist, d.h. dass $a = 0$ ist.

(b) Für $a \in K^{n \times 1}$ gilt

$$\mathrm{wt}(-a) = |\{i \in [1,n] \mid (-a)_i \neq 0\}| = |\{i \in [1,n] \mid -a_i \neq 0\}| = |\{i \in [1,n] \mid a_i \neq 0\}| = \mathrm{wt}(a).$$

(c) Für $a, b \in K^{n \times 1}$ gilt

$$\begin{aligned} \{i \in [1,n] \mid (a+b)_i \neq 0\} &= \{i \in [1,n] \mid a_i + b_i \neq 0\} \subseteq \{i \in [1,n] \mid a_i \neq 0 \text{ oder } b_i \neq 0\} \\ &= \{i \in [1,n] \mid a_i \neq 0\} \cup \{i \in [1,n] \mid b_i \neq 0\} \end{aligned}$$

und damit

$$\begin{aligned} \mathrm{wt}(a+b) &= |\{i \in [1,n] \mid (a+b)_i \neq 0\}| \leq |\{i \in [1,n] \mid a_i \neq 0\} \cup \{i \in [1,n] \mid b_i \neq 0\}| \\ &\leq |\{i \in [1,n] \mid a_i \neq 0\}| + |\{i \in [1,n] \mid b_i \neq 0\}| = \mathrm{wt}(a) + \mathrm{wt}(b). \end{aligned}$$ □

(4.6) Definition (Hamming-Abstand)**.** Für $a, b \in K^{n \times 1}$ heißt

$$\mathrm{d}(a,b) := \mathrm{wt}(b-a)$$

der *Hamming-Abstand* (oder *Hamming-Distanz*) zwischen a und b.

(4.7) Bemerkung. Für $a, b \in K^{n \times 1}$ ist

$$\mathrm{d}(a,b) = |\{i \in [1,n] \mid a_i \neq b_i\}|.$$

Beweis. Für $a, b \in K^{n \times 1}$ ist

$$\mathrm{d}(a,b) = \mathrm{wt}(b-a) = |\{i \in [1,n] \mid (b-a)_i \neq 0\}| = |\{i \in [1,n] \mid b_i - a_i \neq 0\}| = |\{i \in [1,n] \mid a_i \neq b_i\}|.$$ □

Der Hamming-Abstand zwischen einem empfangenen Wort und einem Codewort entspricht also gerade einem Maß für die Anzahl der bei der Übertragung passierten Fehler.

(4.8) Beispiel. In $\mathbb{F}_2^{4 \times 1}$ ist

$$\mathrm{d}(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}) = 2.$$

Beweis. Es ist

$$\mathrm{d}(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}) = |\{2,3\}| = 2.$$ □

(4.9) Proposition.

(a) *Positive Definitheit.* Für $a, b \in K^{n \times 1}$ gilt

$$\mathrm{d}(a,b) \geq 0$$

und genau dann $\mathrm{d}(a,b) = 0$, wenn $a = b$ ist.

(b) *Symmetrie.* Für $a, b \in K^{n \times 1}$ gilt

$$\mathrm{d}(b,a) = \mathrm{d}(a,b).$$

(c) *Dreiecksungleichung.* Für $a, b, c \in K^{n\times 1}$ gilt

$$d(a, c) \leq d(a, b) + d(b, c).$$

(d) *Translationsinvarianz.* Für $a, b, c \in K^{n\times 1}$ gilt

$$d(a + c, b + c) = d(a, b).$$

Beweis.

(a) Auf Grund der positiven Definitheit des Hamming-Gewichts (4.5)(a) gilt für $a, b \in K^{n\times 1}$ stets $d(a, b) = \mathrm{wt}(b - a) \geq 0$ sowie genau dann $d(a, b) = 0$, wenn $b - a = 0$ ist, d.h. wenn $a = b$ gilt.

(b) Auf Grund der Symmetrie des Hamming-Gewichts (4.5)(b) gilt für $a, b \in K^{n\times 1}$ stets

$$d(b, a) = \mathrm{wt}(a - b) = \mathrm{wt}(-(b - a)) = \mathrm{wt}(b - a) = d(a, b).$$

(c) Auf Grund der Dreiecksungleichung für das Hamming-Gewicht (4.5)(c) gilt für $a, b, c \in K^{n\times 1}$ stets

$$d(a, c) = \mathrm{wt}(c - a) = \mathrm{wt}(c - b + b - a) \leq \mathrm{wt}(c - b) + \mathrm{wt}(b - a) = d(b, c) + d(a, b) = d(a, b) + d(b, c).$$

(d) Für $a, b, c \in K^{n\times 1}$ gilt

$$d(a + c, b + c) = \mathrm{wt}((b + c) - (a + c)) = \mathrm{wt}(b - a) = d(a, b). \qquad \square$$

Wenn wir davon ausgehen, dass bei einer Übertragung eines Codeworts c weniger als r Fehler passieren, so müssen wir also all diejenigen Wörter in Betracht ziehen, deren Hamming-Abstand zu c strikt kleiner als r ist. In Analogie zum üblichen Abstandsbegriff (etwa im $\mathbb{R}^2$ oder $\mathbb{R}^3$) verwenden wir folgende Terminologie:

(4.10) Definition (Kugel)**.** Es seien $r \in \mathbb{R}_{\geq 0}$ und $a \in K^{n\times 1}$ gegeben. Wir nennen

$$\mathrm{B}_r(a) := \{x \in K^{n\times 1} \mid d(a, x) < r\}$$

die (*offene*) *Kugel* vom *Radius* r um den *Mittelpunkt* a in $K^{n\times 1}$ bzgl. des Hamming-Abstands.

Man könnte anstatt der offenen Kugel auch die *abgeschlossene Kugel* betrachten, welche für $r \in \mathbb{R}_{\geq 0}$, $a \in K^{n\times 1}$ durch

$$\overline{\mathrm{B}}_r(a) := \{x \in K^{n\times 1} \mid d(a, x) \leq r\}$$

gegeben ist. Da der Hamming-Abstand nur Werte in $\mathbb{N}_0$ annimmt, gilt $\overline{\mathrm{B}}_r(a) = \mathrm{B}_{r+1}(a)$ für $r \in \mathbb{R}_{\geq 0}$, $a \in K^{n\times 1}$.

(4.11) Beispiel. In $\mathbb{F}_2^{4\times 1}$ ist

$$\mathrm{B}_3(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}) = \{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\\1\end{pmatrix}\}$$

Minimalabstand eines Blockcodes

Mit Hilfe des Hamming-Abstands können wir nun den folgenden für die Dekodierung relevanten Begriff einführen.

(4.12) Definition (Minimalabstand)**.** Es sei ein Blockcode C über K der Länge n gegeben. Der *Minimalabstand* (oder die *Minimaldistanz*) von C ist definiert als

$$d(C) := \begin{cases} \min\{d(c, c') \mid c, c' \in C \text{ mit } c \neq c'\}, & \text{falls } |C| > 1, \\ 0, & \text{falls } |C| = 1. \end{cases}$$

Der Minimalabstand gibt also den minimalen Abstand zwischen zwei Codewörtern eines Blockcodes an.

(4.13) Beispiel.

(a) Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 3 gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \}.$$

Dann ist $\mathrm{d}(C) = 2$.

(b) Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 6 gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \}.$$

Dann ist $\mathrm{d}(C) = 3$.

(c) Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 5 gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \}.$$

Dann ist $\mathrm{d}(C) = 3$.

Beweis.

(a) Es ist

$$\mathrm{d}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) = 2,\ \mathrm{d}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}) = 2,\ \mathrm{d}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}) = 2,$$
$$\mathrm{d}(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}) = 2,\ \mathrm{d}(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}) = 2,\ \mathrm{d}(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}) = 2.$$

und damit

$$\mathrm{d}(C) = \min\{2\} = 2.$$

(b) Es ist

$$\mathrm{d}(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}) = 3,\ \mathrm{d}(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}) = 3,\ \mathrm{d}(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}) = 6,$$
$$\mathrm{d}(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}) = 6,\ \mathrm{d}(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}) = 3,\ \mathrm{d}(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}) = 3.$$

und damit

$$\mathrm{d}(C) = \min\{3, 6\} = 3.$$

(c) Es ist

$$\mathrm{d}(\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}) = 3,\ \mathrm{d}(\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}) = 3,\ \mathrm{d}(\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}) = 4,$$
$$\mathrm{d}(\begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}) = 4,\ \mathrm{d}(\begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}) = 3,\ \mathrm{d}(\begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}) = 3.$$

und damit

$$\mathrm{d}(C) = \min\{3, 4\} = 3.$$ □

Wir werden nun sehen, dass die Größe des Minimalabstands für die Anzahl der erkennbaren bzw. korrigierbaren Fehler eines Blockcodes verantwortlich ist.

(4.14) Bemerkung. Es sei ein Blockcode C über K der Länge n gegeben.

(a) Für $c \in C$ ist $\mathrm{B}_{\mathrm{d}(C)}(c) \cap C = \{c\}$.

(b) Für $c, c' \in C$ mit $c \neq c'$ sind $\mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c)$ und $\mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c')$ disjunkt.

Beweis.

(a) Für $c \in C$ und $c' \in \mathrm{B}_{\mathrm{d}(C)}(c) \cap C$ gilt $\mathrm{d}(c, c') < \mathrm{d}(C)$ und damit $c = c'$.

(b) Es seien $c, c' \in C$ so gegeben, dass $\mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c)$ und $\mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c')$ nicht disjunkt sind, d.h. so, dass $\mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c) \cap \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c') \neq \emptyset$ ist. Ferner sei $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c) \cap \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c')$ gegeben. Dann ist $\mathrm{d}(c, x) < \frac{\mathrm{d}(C)}{2}$ und $\mathrm{d}(c', x) < \frac{\mathrm{d}(C)}{2}$, auf Grund der Dreiecksungleichung (4.9)(c) und der Symmetrie des Hamming-Abstands (4.9)(b) also

$$\mathrm{d}(c, c') \leq \mathrm{d}(c, x) + \mathrm{d}(x, c') = \mathrm{d}(c, x) + \mathrm{d}(c', x) < \frac{\mathrm{d}(C)}{2} + \frac{\mathrm{d}(C)}{2} = \mathrm{d}(C)$$

und damit $c = c'$. □

(4.15) Definition (nächster Nachbar)**.** Es seien ein Blockcode C über K der Länge n und ein $x \in K^{n \times 1}$ gegeben. Ein *nächster Nachbar* von x in C ist ein $c \in C$ mit

$$\mathrm{d}(c, x) = \min\{\mathrm{d}(c', x) \mid c' \in C\}.$$

(4.16) Beispiel. Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 5 gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}\}$$

und es sei $x \in \mathbb{F}_2^{5 \times 1}$ gegeben durch

$$x = \begin{pmatrix}0\\1\\1\\1\\1\end{pmatrix}.$$

Dann ist

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

ein nächster Nachbar von x in C.

Beweis. Es seien $c_1, c_2, c_3, c_4 \in \mathbb{F}_2^{5 \times 1}$ gegeben durch

$$c_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ c_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix},\ c_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix},\ c_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

so dass $C = \{c_1, c_2, c_3, c_4\}$ gilt. Dann ist

$$\begin{aligned} \mathrm{d}(c_1, x) &= 4, \\ \mathrm{d}(c_2, x) &= 3, \\ \mathrm{d}(c_3, x) &= 1, \\ \mathrm{d}(c_4, x) &= 2. \end{aligned}$$

Folglich ist

$$\mathrm{d}(c_3, x) = \min \{\mathrm{d}(c_1, x), \mathrm{d}(c_2, x), \mathrm{d}(c_3, x), \mathrm{d}(c_4, x)\},$$

d.h. c_3 ist ein nächster Nachbar von x in C. □

(4.17) Bemerkung. Es seien ein Blockcode C über K der Länge n und ein Codewort c von C gegeben. Für alle $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c)$ ist c der eindeutige nächste Nachbar von x in C.

Beweis. Es sei $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c)$ gegeben. Für $c' \in C$ mit $\mathrm{d}(c', x) \leq \mathrm{d}(c, x)$ gilt dann

$$\mathrm{d}(c', x) \leq \mathrm{d}(c, x) < \frac{\mathrm{d}(C)}{2},$$

d.h. es ist $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c')$ und damit $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c) \cap \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c')$. Nach Bemerkung (4.14)(b) folgt $c = c'$. Folglich gilt einerseits $\mathrm{d}(c, x) \leq \mathrm{d}(c', x)$ für alle $c' \in C$ und damit $\mathrm{d}(c, x) = \min \{\mathrm{d}(c', x) \mid c' \in C\}$, d.h. c ist ein nächster Nachbar von x in C. Andererseits gilt für jeden beliebigen nächsten Nachbarn c' von x in C insbesondere $\mathrm{d}(c', x) \leq \mathrm{d}(c, x)$ und damit $c = c'$. Insgesamt ist c der eindeutige nächste Nachbar von x in C. □

Unter der Annahme, dass bei der Übertragung nicht zu viele Fehler passieren, nämlich strikt weniger als die Hälfte des Minimalabstands, können wir nach Bemerkung (4.17) also ein empfangenes Wort zum gesendeten Codewort dekodieren.

(4.18) Anwendung (Nächster-Nachbar-Dekodierung)**.**

- Initialisierung:
  - Wähle einen Blockcode C über K der Länge n so, dass bei der Übertragung eines Codeworts (mit hoher Wahrscheinlichkeit) weniger als $\frac{\mathrm{d}(C)}{2}$ Fehler passieren.
- Dekodierung eines empfangenen Worts $x \in K^{n \times 1}$:
  - Bestimme den nächsten Nachbarn c von x.
  - Dekodiere x zu c.

- Beispiel:
  - Wir wählen den Blockcode $C = \{c_1, c_2, c_3, c_4\}$ über $\mathbb{F}_2$ der Länge 5 gegeben durch
  $$c_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, c_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, c_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, c_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$
  - Das empfangene Wort $x \in \mathbb{F}_2^{5 \times 1}$ sei gegeben durch
  $$x = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$
  Nach Beispiel (4.16) ist c_3 der nächste Nachbar von x in C. Es wird x zu c_3 dekodiert.

Lineare Codes

Um allgemeine Blockcodes beschreiben zu können, müssen wir die einzelnen Codewörter angeben. Da dies bei vielen Codewörtern in der Praxis zu aufwändig ist, führen wir nun spezielle Blockcodes ein, bei denen mit Hilfe der linearen Algebra einfachere Beschreibungen möglich sind.

(4.19) Definition (linearer Code)**.**

(a) Ein *linearer Code* über K der Länge n ist ein Blockcode C über K der Länge n so, dass C ein K-Untervektorraum von $K^{n \times 1}$ ist.

Ein *linearer Code* über K ist ein linearer Code über K der Länge m für ein $m \in \mathbb{N}_0$.

(b) Es seien $k \in [0, n]$ und ein linearer Code C über K gegeben. Wir sagen, dass C ein *linearer* $[n, k]$*-Code* ist, wenn C ein linearer Code über K der Länge n mit $\dim_K C = k$ ist.

(c) Es seien $d, k \in [0, n]$ und ein linearer Code C über K gegeben. Wir sagen, dass C ein *linearer* $[n, k, d]$*-Code* ist (oder dass C die *Parameter* $[n, k, d]$ hat), wenn C ein linearer $[n, k]$-Code über K mit $\mathrm{d}(C) = d$ ist.

Lineare Codes werden beispielsweise bei der Compact Disc (CD) eingesetzt; dort verwendet man einen linearen $[28, 24, 5]$-Code und einen linearen $[32, 28, 5]$-Code über einem Körper mit $2^8 = 256$ Elementen [15, Abschn. 3].

(4.20) Beispiel.

(a) Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 3 gegeben durch
$$C = \{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\}.$$
Dann ist C ein linearer $[3, 2, 2]$-Code über $\mathbb{F}_2$.

(b) Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 6 gegeben durch
$$C = \{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\}.$$
Dann ist C ein linearer $[6, 2, 3]$-Code über $\mathbb{F}_2$.

(c) Es sei C der Blockcode über $\mathbb{F}_2$ der Länge 5 gegeben durch

$$C = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \}.$$

Dann ist C ein linearer $[5, 2, 3]$-Code über $\mathbb{F}_2$.

Beweis.

(a) Es ist

$$C = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rangle.$$

Folglich ist C ein Untervektorraum von $\mathbb{F}_2^{3 \times 1}$ der Dimension $\dim C = 2$. Nach Beispiel (4.13)(a) ist der Minimalabstand von C durch $\mathrm{d}(C) = 2$ gegeben. Insgesamt ist C ein linearer $[3, 2, 2]$-Code über $\mathbb{F}_2$.

(b) Es ist

$$C = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \rangle.$$

Folglich ist C ein Untervektorraum von $\mathbb{F}_2^{6 \times 1}$ der Dimension $\dim C = 2$. Nach Beispiel (4.13)(b) ist der Minimalabstand von C durch $\mathrm{d}(C) = 3$ gegeben. Insgesamt ist C ein linearer $[6, 2, 3]$-Code über $\mathbb{F}_2$.

(c) Es ist

$$C = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \rangle.$$

Folglich ist C ein Untervektorraum von $\mathbb{F}_2^{5 \times 1}$ der Dimension $\dim C = 2$. Nach Beispiel (4.13)(c) ist der Minimalabstand von C durch $\mathrm{d}(C) = 3$ gegeben. Insgesamt ist C ein linearer $[5, 2, 3]$-Code über $\mathbb{F}_2$. □

(4.21) Bemerkung. Es sei ein linearer Code C der Länge n über K gegeben. Dann ist

$$\mathrm{d}(C) = \begin{cases} \min \{\mathrm{wt}(c) \mid c \in C \setminus \{0\}\}, & \text{falls } C \neq \{0\}, \\ 0, & \text{falls } C = \{0\}. \end{cases}$$

Beweis. Es gelte $C \neq \{0\}$. Da C ein Untervektorraum von $K^{n \times 1}$ ist, gilt $\{c' - c \mid c, c' \in C,\, c \neq c'\} = C \setminus \{0\}$. Folglich ist

$$\begin{aligned} \mathrm{d}(C) &= \min \{\mathrm{d}(c, c') \mid c, c' \in C,\, c \neq c'\} = \min \{\mathrm{wt}(c' - c) \mid c, c' \in C,\, c \neq c'\} \\ &= \min \{\mathrm{wt}(c) \mid c \in C \setminus \{0\}\}. \end{aligned}$$

□

Alternativer Beweis von Beispiel (4.13).

(a) Nach Beispiel (4.20)(a) ist C ein linearer Code. Wegen

$$\mathrm{wt}(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) = 2,\ \mathrm{wt}(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}) = 2,\ \mathrm{wt}(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}) = 2,$$

gilt nach Bemerkung (4.21) also

$$\mathrm{d}(C) = \min \{2\} = 2.$$

(b) Nach Beispiel (4.20)(b) ist C ein linearer Code. Wegen

$$\mathrm{wt}(\begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}) = 3,\ \mathrm{wt}(\begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix}) = 3,\ \mathrm{wt}(\begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix}) = 6,$$

gilt nach Bemerkung (4.21) also

$$\mathrm{d}(C) = \min\{3, 6\} = 3.$$

(c) Nach Beispiel (4.20)(c) ist C ein linearer Code. Wegen

$$\mathrm{wt}(\begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}) = 3,\ \mathrm{wt}(\begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}) = 3,\ \mathrm{wt}(\begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}) = 4,$$

gilt nach Bemerkung (4.21) also

$$\mathrm{d}(C) = \min\{3, 4\} = 3.$$

□

Erzeugermatrizen

Mit Hilfe von Basen lassen sich endlichdimensionale Untervektorräume und damit insbesondere lineare Codes der Länge n als Untervektorräume von $K^{n\times 1}$ kompakt beschreiben. Wir benutzen das Konzept des Spaltenraums, siehe Definition (3.20), also die Beschreibung als Bild der Spalteninterpretation $\varphi_A \colon K^{k\times 1} \to K^{n\times 1}$ für gewisse $k \in \mathbb{N}_0$, $A \in K^{n\times k}$.

(4.22) Definition (Erzeugermatrix)**.** Es seien $k \in [0, n]$ und ein linearer $[n, k]$-Code C über K gegeben. Eine *Erzeugermatrix* (oder *Generatormatrix*) für C ist ein $A \in K^{n\times k}$ mit

$$C = \mathrm{Col}(A).$$

Da $K^{n\times 1}$ endlichdimensional ist, ist jeder lineare Code nach Proposition (1.68) endlichdimensional und besitzt demnach eine Basis. Nach Bemerkung (3.21) gibt es daher für jeden linearen Code eine Erzeugermatrix.

(4.23) Beispiel.

(a) Es seien ein linearer Code C über $\mathbb{F}_2$ der Länge 3 und $A \in \mathbb{F}_2^{3\times 2}$ gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\},\ A = \begin{pmatrix}1 & 0\\0 & 1\\1 & 1\end{pmatrix}.$$

Dann ist A eine Erzeugermatrix für C.

(b) Es seien ein linearer Code C über $\mathbb{F}_2$ der Länge 6 und $A \in \mathbb{F}_2^{6\times 2}$ gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix}\},\ A = \begin{pmatrix}1 & 0\\0 & 1\\1 & 0\\0 & 1\\1 & 0\\0 & 1\end{pmatrix}.$$

Dann ist A eine Erzeugermatrix für C.

(c) Es seien ein linearer Code C über $\mathbb{F}_2$ der Länge 5 und $A \in \mathbb{F}_2^{5\times 2}$ gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}\},\ A = \begin{pmatrix}1&0\\0&1\\1&1\\1&0\\0&1\end{pmatrix}.$$

Dann ist A eine Erzeugermatrix für C.

Beweis.

(a) Nach Bemerkung (3.21) ist

$$\mathrm{Col}(A) = \langle \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix} \rangle = \{a \begin{pmatrix}1\\0\\1\end{pmatrix} + b \begin{pmatrix}0\\1\\1\end{pmatrix} \mid a, b \in \mathbb{F}_2\} = \{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\} = C.$$

(b) Nach Bemerkung (3.21) ist

$$\mathrm{Col}(A) = \langle \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix} \rangle = \{a \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix} + b \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix} \mid a, b \in \mathbb{F}_2\} = \{\begin{pmatrix}0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix}\} = C.$$

(c) Nach Bemerkung (3.21) ist

$$\mathrm{Col}(A) = \langle \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix} \rangle = \{a \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix} + b \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix} \mid a, b \in \mathbb{F}_2\} = \{\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}\} = C. \qquad \square$$

(4.24) Bemerkung. Es seien $k \in [0, n]$, ein linearer $[n, k]$-Code C über K und eine Erzeugermatrix A für C gegeben. Dann gilt

$$\mathrm{rk}_K\, A = k.$$

Beweis. Nach Bemerkung (3.28) gilt

$$\mathrm{rk}\, A = \dim \mathrm{Col}(A) = \dim C = k. \qquad \square$$

Neben der effizienten Beschreibung eines linearen Codes lassen sich Erzeugermatrizen auch zur Umwandlung von Informationswörtern in Codewörter und umgekehrt, also zum Kodieren und Dekodieren, benutzen: Es seien ein linearer $[n, k]$-Code C über K und eine Erzeugermatrix A für C gegeben. Nach dem Rangsatz (2.28) und Bemerkung (4.24) ist

$$\mathrm{def}\, \varphi_A = \dim K^{k\times 1} - \mathrm{rk}\, \varphi_A = k - \mathrm{rk}\, A = 0$$

und damit $\varphi_A \colon K^{k\times 1} \to K^{n\times 1}$, $w \mapsto Aw$ injektiv nach Bemerkung (2.25)(b). Wegen $C = \mathrm{Col}(A) = \mathrm{Im}\, \varphi_A$ ist daher

$$\varphi_A|^C \colon K^{k\times 1} \to C,\ w \mapsto Aw$$

eine Bijektion (und sogar ein Isomorphismus nach Bemerkung (3.1) und Bemerkung (2.9)).
Wir betonen, dass wir unter Kodierung hierbei den Übergang von einem Informationswort zu einem Codewort bezeichnen, während wir unter Dekodierung den Übergang von einem Codewort zu einem Informationswort meinen. Diese gegenseitige Zuordnung ist abhängig von einer gewählten Erzeugermatrix, weswegen wir diese bei der Kodierung bzw. Dekodierung mit angeben müssen. Unter Dekodierung ist hier hingegen *nicht* die Dekodierung eines empfangenen Wortes zu einem Codewort im Sinne der Nächster-Nachbar-Dekodierung (4.18) gemeint, diese ist unabhängig von der gewählten Erzeugermatrix.

(4.25) Anwendung (Umwandlung von Informationswörtern und Codewörtern)**.**

- Setup: Es seien $k \in [0, n]$, ein linearer $[n, k]$-Code C über K und eine Erzeugermatrix A für C gegeben.
- Kodierung eines Informationsworts $w \in K^{k \times 1}$ zu einem Codewort bzgl. A:
  - Berechne Aw.
- Dekodierung eines Codeworts $c \in C$ zu einem Informationswort bzgl. A:
  - Berechne $w \in K^{k \times 1}$ mit $Aw = c$.
- Beispiel:
  - Es sei C der lineare Code über $\mathbb{F}_2$ mit Erzeugermatrix $A \in \mathbb{F}_2^{5 \times 2}$ gegeben durch
  $$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$
  - Die Kodierung des Informationsworts
  $$w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
  bzgl. der Erzeugermatrix A ergibt
  $$Aw = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$
  - Die Dekodierung des Codeworts
  $$c = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$
  bzgl. der Erzeugermatrix A ergibt
  $$\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
  denn es ist
  $$A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = c.$$

Kontrollmatrizen

Um zu testen, ob ein Wort der Länge n ein Codewort eines durch eine Erzeugermatrix gegebenen Codes ist, müssen wir überprüfen, ob sich dieses Wort als Linearkombination der Spalten der Erzeugermatrix schreiben lässt – mit anderen Worten, wir müssen ein lineares Gleichungssystem auf dessen Lösbarkeit testen. Im Folgenden werden wir ein alternatives Verfahren zur Erkennung von Codewörtern mittels sogenannter Kontrollmatrizen angeben. Während man mit einer Erzeugermatrix eine Beschreibung eines linearen Codes als Spaltenraum, nach Definition (3.20) also als Bild einer Spalteninterpretation, erhält, liefert eine Kontrollmatrix eine Beschreibung als Lösungsraum eines homogenen linearen Gleichungssystems, nach Korollar (3.36) also als Kern einer Spalteninterpretation.

(4.26) Definition (Kontrollmatrix)**.** Es seien $k \in [0, n]$ und ein linearer $[n, k]$-Code C über K gegeben. Eine *Kontrollmatrix* für C ist ein $B \in K^{(n-k)\times n}$ mit

$$C = \operatorname{Sol}(B, 0).$$

(4.27) Beispiel.

(a) Es seien ein linearer Code C über $\mathbb{F}_2$ der Länge 3 und $B \in \mathbb{F}_2^{1\times 3}$ gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\},\ B = \begin{pmatrix}1 & 1 & 1\end{pmatrix}.$$

Dann ist B eine Kontrollmatrix für C.

(b) Es seien ein linearer Code C über $\mathbb{F}_2$ der Länge 6 und $B \in \mathbb{F}_2^{4\times 6}$ gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix}\},\ B = \begin{pmatrix}1 & 0 & 0 & 0 & 1 & 0\\0 & 1 & 0 & 0 & 0 & 1\\0 & 0 & 1 & 0 & 1 & 0\\0 & 0 & 0 & 1 & 0 & 1\end{pmatrix}.$$

Dann ist B eine Kontrollmatrix für C.

(c) Es seien ein linearer Code C über $\mathbb{F}_2$ der Länge 5 und $B \in \mathbb{F}_2^{3\times 5}$ gegeben durch

$$C = \{\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\\1\end{pmatrix}\},\ B = \begin{pmatrix}1 & 0 & 0 & 1 & 0\\0 & 1 & 0 & 0 & 1\\0 & 0 & 1 & 1 & 1\end{pmatrix}.$$

Dann ist B eine Kontrollmatrix für C.

Beweis.

(a) Es ist

$$\operatorname{Sol}(B, 0) = \langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix} \rangle = \{a \begin{pmatrix}1\\1\\0\end{pmatrix} + b \begin{pmatrix}1\\0\\1\end{pmatrix} \mid a, b \in \mathbb{F}_2\} = \{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}\} = C.$$

(b) Es ist

$$\operatorname{Sol}(B, 0) = \langle \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix} \rangle = \{a \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix} + b \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix} \mid a, b \in \mathbb{F}_2\} = \{\begin{pmatrix}0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix}\} = C.$$

(c) Es ist

$$\mathrm{Sol}(B,0) = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \rangle = \{a \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \mid a, b \in \mathbb{F}_2\} = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \} = C. \qquad \square$$

Eine Kontrollmatrix für einen durch eine Erzeugermatrix gegebenen linearen Code lässt sich mit Hilfe des in Proposition (3.73) vorgestellten Verfahrens berechnen:

Alternativer Beweis von Beispiel (4.27)(a). Nach Beispiel (4.23)(a) ist $A \in \mathbb{F}_2^{3 \times 2}$ gegeben durch

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

eine Erzeugermatrix von C. Wegen

$$\mathrm{Sol}(A^{\mathrm{tr}}, 0) = \mathrm{Sol}(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}) = \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle = \mathrm{Col}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}) = \mathrm{Col}(B^{\mathrm{tr}})$$

gilt $\mathrm{Sol}(B,0) = \mathrm{Col}(A) = C$ nach Proposition (3.73). Folglich ist B eine Kontrollmatrix von C. $\square$

(4.28) Bemerkung. Es seien $k \in [0,n]$, ein linearer $[n,k]$-Code C über K und eine Kontrollmatrix B für C gegeben. Dann gilt

$$\mathrm{rk}_K\, B = n - k.$$

Beweis. Nach Korollar (3.36) gilt

$$\mathrm{rk}\, B = n - \dim \mathrm{Sol}(B,0) = n - \dim C = n - k. \qquad \square$$

(4.29) Anwendung (Codewort-Test)**.**

- Setup: Es seien $k \in [0,n]$, ein linearer $[n,k]$-Code C über K und eine Kontrollmatrix B für C gegeben.
- Test, ob ein empfangenes Wort $x \in K^{n \times 1}$ ein Codewort von C ist:
  - Berechne Bx.
  - Wenn $Bx = 0$ ist, dann ist x ein Codewort von C, ansonsten nicht.
- Beispiel:
  - Es sei C der lineare Code über $\mathbb{F}_2$ mit Kontrollmatrix $B \in \mathbb{F}_2^{3 \times 5}$ gegeben durch
  $$B = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$
  - Das empfangene Wort $x_1 \in \mathbb{F}_2^{5 \times 1}$ gegeben durch
  $$x_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$
  ist ein Codewort von C, denn es gilt
  $$Bx_1 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

– Das empfangene Wort $x_2 \in \mathbb{F}_2^{5\times 1}$ gegeben durch

$$x_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

ist kein Codewort von C, denn es gilt

$$Bx_2 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \neq 0.$$

Ein weiterer Vorteil von Kontrollmatrizen ergibt sich bei der Bestimmung des Minimalabstands:

(4.30) Bemerkung. Es seien $k \in [1, n]$, ein linearer $[n, k]$-Code C über K und eine Kontrollmatrix B für C gegeben. Dann ist

$$\begin{aligned} \mathrm{d}(C) &= \min \{r \in \mathbb{N} \mid \text{es gibt } j_1, \ldots, j_r \in [1, n] \text{ mit } j_1 < \ldots < j_r \text{ so, dass } (B_{-,j_1}, \ldots, B_{-,j_r}) \text{ linear} \\ &\qquad \text{abhängig ist}\} \\ &= \max \{r \in \mathbb{N} \mid \text{für alle } j_1, \ldots, j_{r-1} \in [1, n] \text{ mit } j_1 < \ldots < j_{r-1} \text{ ist } (B_{-,j_1}, \ldots, B_{-,j_{r-1}}) \text{ linear} \\ &\qquad \text{unabhängig}\}. \end{aligned}$$

Beweis. Für $c \in K^{n\times 1}$ gilt genau dann $c \in C$, wenn $Bc = 0$ ist, d.h. wenn $\sum_{j\in[1,n]} c_j B_{-,j} = 0$ ist. Also gilt

$$\begin{aligned} \mathrm{d}(C) &= \min \{\mathrm{wt}(c) \mid c \in C \setminus \{0\}\} = \min \{\mathrm{wt}(c) \mid c \in K^{n\times 1} \setminus \{0\} \text{ mit } \sum_{j\in[1,n]} c_j B_{-,j} = 0\} \\ &= \min \{|\{j \in [1, n] \mid c_j \neq 0\}| \mid c \in K^{n\times 1} \setminus \{0\} \text{ mit } \sum_{\substack{j\in[1,n] \\ c_j\neq 0}} c_j B_{-,j} = 0\} \\ &= \min \{|J| \mid J \subseteq [1, n] \text{ so, dass } (B_{-,j})_{j\in J} \text{ linear abhängig ist}\} \\ &= \min \{r \in \mathbb{N}_0 \mid \text{es gibt } j_1, \ldots, j_r \in [1, n] \text{ mit } j_1 < \ldots < j_r \text{ so, dass } (B_{-,j_1}, \ldots, B_{-,j_r}) \text{ linear} \\ &\qquad \text{abhängig ist}\}. \end{aligned}$$

□

Alternativer Beweis von Beispiel (4.13).

(a) Nach Beispiel (4.20)(a) ist C ein linearer Code und nach Beispiel (4.27)(a) ist $B \in \mathbb{F}_2^{1\times 3}$ gegeben durch

$$B = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

eine Kontrollmatrix für C. Für alle $j \in [1, 3]$ ist $(B_{-,j})$ linear unabhängig. Ferner ist $(B_{-,1}, B_{-,2})$ linear abhängig. Nach Bemerkung (4.30) ist somit

$$\mathrm{d}(C) = 2.$$

(b) Nach Beispiel (4.20)(b) ist C ein linearer Code und nach Beispiel (4.27)(b) ist $B \in \mathbb{F}_2^{4\times 6}$ gegeben durch

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

eine Kontrollmatrix für C. Für alle $j_1, j_2 \in [1, 6]$ mit $j_1 < j_2$ ist (B_{-,j_1}, B_{-,j_2}) linear unabhängig. Ferner ist $(B_{-,1}, B_{-,3}, B_{-,5})$ linear abhängig. Nach Bemerkung (4.30) ist somit

$$\mathrm{d}(C) = 3.$$

(c) Nach Beispiel (4.20)(c) ist C ein linearer Code und nach Beispiel (4.27)(c) ist $B \in \mathbb{F}_2^{3\times 5}$ gegeben durch

$$B = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

eine Kontrollmatrix für C. Für alle $j_1, j_2 \in [1, 5]$ mit $j_1 < j_2$ ist (B_{-,j_1}, B_{-,j_2}) linear unabhängig. Ferner ist $(B_{-,1}, B_{-,3}, B_{-,4})$ linear abhängig. Nach Bemerkung (4.30) ist somit

$$\mathrm{d}(C) = 3.$$

□

Syndromdekodierung

Wenn wir ein empfangenes Wort zu einem nächstgelegenen Nachbarn eines Blockcodes dekodieren möchten, müssen wir die Abstände zwischen dem empfangenen Wort und allen Codewörtern berechnen. Mit Hilfe von Kontrollmatrizen wollen wir nun eine Variante dieses Verfahrens für lineare Codes entwickeln.
Es seien $k \in [0, n]$, ein linearer $[n, k]$-Code C über K und eine Kontrollmatrix B für C gegeben. Nach Bemerkung (4.28) ist $\operatorname{rk} \varphi_B = \operatorname{rk} B = n - k$ und damit $\varphi_B \colon K^{n\times 1} \to K^{(n-k)\times 1}$, $x \mapsto Bx$ surjektiv nach Bemerkung (2.25)(a). Ferner ist $C = \mathrm{Sol}(B, 0) = \{x \in K^{n\times 1} \mid Bx = 0\}$, d.h. für $x \in K^{n\times 1}$ gilt genau dann $x \in C$, wenn $Bx = 0$ ist. Mit anderen Worten: Ist $x \in K^{n\times 1}$ ein empfangenes Wort so, dass $Bx \neq 0$ ist, so wissen wir, dass bei der Übertragung ein Fehler passiert ist. Die Idee ist nun, von dem Wert Bx einen Rückschluss auf den Fehler zu machen und dies zur Dekodierung zu verwenden.

(4.31) Definition (Syndrom)**.** Es seien $k \in [0, n]$, ein linearer $[n, k]$-Code C über K und eine Kontrollmatrix B für C gegeben. Für $x \in K^{n\times 1}$ heißt $Bx \in K^{(n-k)\times 1}$ das *Syndrom* von x bzgl. B.

(4.32) Beispiel. Es sei C der lineare Code C über $\mathbb{F}_2$ mit Kontrollmatrix $B \in \mathbb{F}_2^{3\times 5}$ gegeben durch

$$B = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

und es sei $x \in \mathbb{F}_2^{5\times 1}$ gegeben durch

$$x = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Das Syndrom von x bzgl. B ist

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

Beweis. Das Syndrom von x bzgl. B ist

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

□

(4.33) Bemerkung. Es seien ein linearer Code C über K der Länge n und eine Kontrollmatrix B für C gegeben. Ferner sei $x \in K^{n\times 1}$ gegeben und es sei y das Syndrom von x. Für $x' \in K^{n\times 1}$ gilt: Genau dann hat x' das Syndrom y, wenn $x' - x$ ein Codewort von C ist.

Beweis. Es sei $x' \in K^{n\times 1}$ gegeben. Genau dann hat x' das Syndrom y, wenn $Bx' = y$ ist, d.h. wenn $x' \in \mathrm{Sol}(B, y)$ ist. Nach Korollar (3.37) gilt aber

$$\mathrm{Sol}(B, y) = x + \mathrm{Sol}(B, 0) = x + C,$$

es hat also x' genau dann das Syndrom y, wenn $x' \in x + C$ ist, d.h. wenn $x' - x$ ein Codewort von C ist. □

Sind bei der Übertragung eines Codeworts c eines linearen Codes C über K der Länge n Fehler passiert, so gilt $x = c + e$ und damit $c = x - e$ für ein *Fehlerwort* $e \in K^{n \times 1}$. Nach Bemerkung (4.33) hat ein empfangenes Wort somit stets das gleiche Syndrom wie das zugehörige Fehlerwort. Da wir annehmen, dass nicht zu viele Fehler passiert sind, suchen wir bei der Dekodierung eines empfangenen Worts x ein Fehlerwort e mit dem gleichen Syndrom und von möglichst kleinem Gewicht.

(4.34) Definition (Anführer)**.** Es seien $k \in [0, n]$, ein linearer $[n, k]$-Code C über K, eine Kontrollmatrix B für C und ein $y \in K^{(n-k) \times 1}$ gegeben. Ein *Anführer* für y bzgl. B ist ein $e \in K^{n \times 1}$ mit Syndrom y und

$$\mathrm{wt}(e) = \min \{\mathrm{wt}(x) \mid x \in K^{n \times 1} \text{ so, dass } x \text{ das Syndrom } y \text{ hat}\}.$$

(4.35) Beispiel. Es sei C der lineare Code C über $\mathbb{F}_2$ mit Kontrollmatrix $B \in \mathbb{F}_2^{3 \times 5}$ gegeben durch

$$B = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

und es sei $y \in \mathbb{F}_2^{3 \times 1}$ gegeben durch

$$y = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Dann ist

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

ein Anführer für y bzgl. B.

Beweis. Es sei $e \in K^{5 \times 1}$ gegeben durch

$$e = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Dann ist

$$Be = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = y,$$

d.h. e hat das Syndrom y bzgl. B. Ferner ist

$$\mathrm{wt}(e) = |\{4\}| = 1.$$

Da 0 das Syndrom 0 hat, gilt auf Grund der positiven Definitheit des Hamming-Gewichts (4.5)(a) also notwendigerweise

$$\mathrm{wt}(e) = \min \{\mathrm{wt}(x) \mid x \in K^{n \times 1} \text{ so, dass } x \text{ das Syndrom } y \text{ hat}\},$$

d.h. e ist ein Anführer für y bzgl. B. □

(4.36) Proposition. Es seien ein linearer Code C über K der Länge n und eine Kontrollmatrix B für C gegeben. Ferner seien $x \in K^{n \times 1}$ und ein Codewort c von C gegeben und es sei y das Syndrom von x bzgl. B. Genau dann ist c ein nächster Nachbar von x in C, wenn $x - c$ ein Anführer für y ist.

Beweis. Genau dann ist c ein nächster Nachbar von x in C, wenn

$$\mathrm{d}(c, x) = \min \{\mathrm{d}(c', x) \mid c' \in C\}$$

ist. Für $c' \in C$ gilt $\mathrm{d}(c', x) = \mathrm{wt}(x - c')$, es ist also c genau dann ein nächster Nachbar von x in C, wenn

$$\mathrm{wt}(x - c) = \min \{\mathrm{wt}(x - c') \mid c' \in C\}$$

gilt. Nach Bemerkung (4.33) ist aber $\{x - c' \mid c' \in C\} = \{x' \in K^{n\times 1} \mid x' \text{ hat das Syndrom } y\}$, so dass c also genau dann ein nächster Nachbar von x in C ist, wenn

$$\mathrm{wt}(x - c) = \min \{\mathrm{wt}(x') \mid x' \in K^{n\times 1} \text{ so, dass } x' \text{ das Syndrom } y \text{ hat}\}$$

gilt, d.h. wenn $x - c$ ein Anführer für y ist. □

(4.37) Korollar. Es seien $k \in \mathbb{N}_0$, ein linearer $[n, k]$-Code C über K, eine Kontrollmatrix B für C und ein Codewort c von C gegeben. Für alle $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c)$ ist $x - c$ der eindeutige Anführer für das Syndrom von x bzgl. B.

Beweis. Es sei $x \in \mathrm{B}_{\frac{\mathrm{d}(C)}{2}}(c)$ gegeben. Nach Bemerkung (4.17) ist c der eindeutige nächste Nachbar von x in C und nach Proposition (4.36) ist damit $x - c$ der eindeutige Anführer für das Syndrom von x bzgl. B. □

(4.38) Anwendung (Syndromdekodierung)**.**

- Initialisierung:
  - Wähle $k \in [0, n]$ und einen linearen $[n, k]$-Code C über K so, dass bei der Übertragung eines Codeworts (mit hoher Wahrscheinlichkeit) weniger als $\frac{\mathrm{d}(C)}{2}$ Fehler passieren.
  - Bestimme eine Kontrollmatrix B für C.
  - Bestimme für jedes $y \in K^{(n-k)\times 1}$ einen Anführer e_y und lege ein Wörterbuch an.
- Dekodierung eines empfangenen Worts $x \in K^{n\times 1}$:
  - Berechne das Syndrom Bx von x bzgl. B.
  - Bestimme das Fehlerwort e_{Bx} durch Nachschauen im Wörterbuch.
  - Dekodiere x zu

    $$x - e_{Bx}.$$
- Beispiel:
  - Wir wählen den linearen Code C über $\mathbb{F}_2$ mit Kontrollmatrix $B \in \mathbb{F}_2^{3\times 5}$ gegeben durch

    $$B = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$
  - Wir bestimmen für jedes $y \in \mathbb{F}_2^{3\times 1}$ einen Anführer:

Syndrom	$\begin{pmatrix}0\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\1\end{pmatrix}$
Anführer	$\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\1\end{pmatrix}$	$\begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\\0\end{pmatrix}$

– Das empfangene Wort $x \in \mathbb{F}_2^{5 \times 1}$ sei gegeben durch

$$x = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Dann ist das Syndrom von x bzgl. B gegeben durch

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

Der Anführer e_{Bx} zum Syndrom von x bzgl. B ist gegeben durch

$$e_{Bx} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Es wird x dekodiert zu

$$x - e_{Bx} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Dass wir im Beispiel in (4.38) für zwei Syndrome mehrere Anführer haben, ist nicht verwunderlich – nach Beispiel (4.27)(c) und Beispiel (4.13)(c) hat dieser Code den Minimalabstand 3, der Code korrigiert also weniger als $\frac{3}{2}$ Fehler, d.h. er korrigiert 1 Fehler. Die Anführer für die genannten Syndrome haben aber allesamt Gewicht 2, so dass bei diesen Syndromen keine eindeutige Dekodierung mehr gewährleistet ist. Oder positiv formuliert: Unter der Annahme, dass wir den Code so gewählt haben, dass höchstens 1 Fehler passiert, treten diese Syndrome bei den empfangenen Wörtern nicht auf.

5 Determinante

In diesem Abschnitt führen wir die Determinante von quadratischen Matrizen ein. Hauptanwendung der Determinante ist die Charakterisierung invertierbarer Matrizen in Satz (5.14): Eine quadratische Matrix mit Einträgen in einem kommutativen, unitären Ring ist genau dann invertierbar, wenn ihre Determinante invertierbar ist. Ist der Ring ein Körper, so ist dies gleichbedeutend damit, dass die Determinante von Null verschieden ist.
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, seien stets ein kommutativer, unitärer Ring R und eine nicht-negative ganze Zahl n gegeben.

Begriffsbildung

Um die Determinante definieren zu können, benötigen wir einige Grundlagen zur symmetrischen Gruppe, siehe Definition (A.98) und Definition (A.103). Wir definieren die Determinante über die sogenannte Leibniz-Formel:

(5.1) Definition (Determinante)**.** Die *Determinante* auf $R^{n \times n}$ ist die Abbildung $\det \colon R^{n \times n} \to R$ gegeben durch

$$\det A = \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} A_{\pi(j),j}$$

für $A \in R^{n \times n}$.

(5.2) Beispiel.

(a) Die Determinante auf $R^{0 \times 0}$ ist gegeben durch

$$\det A = \det \big(\big) = 1$$

für $A \in R^{0 \times 0}$.

(b) Die Determinante auf $R^{1 \times 1}$ ist gegeben durch

$$\det A = \det \begin{pmatrix} A_{1,1} \end{pmatrix} = A_{1,1}$$

für $A \in R^{1 \times 1}$.

(c) Die Determinante auf $R^{2 \times 2}$ ist gegeben durch

$$\det A = \det \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} = A_{1,1} A_{2,2} - A_{1,2} A_{2,1}$$

für $A \in R^{2 \times 2}$.

(d) Die Determinante auf $R^{3 \times 3}$ ist gegeben durch

$$\begin{aligned} \det A &= \det \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix} \\ &= A_{1,1} A_{2,2} A_{3,3} + A_{1,2} A_{2,3} A_{3,1} + A_{1,3} A_{2,1} A_{3,2} - A_{1,1} A_{2,3} A_{3,2} - A_{1,2} A_{2,1} A_{3,3} - A_{1,3} A_{2,2} A_{3,1} \end{aligned}$$

für $A \in R^{3 \times 3}$.

Beweis.

(a) Es ist $\mathrm{S}_0 = \{(\;)\}$, also

$$\det A = \sum_{\pi \in \mathrm{S}_0} (\operatorname{sgn} \pi) \prod_{j \in [1,0]} A_{\pi(j),j} = (\operatorname{sgn} (\;)) 1 = 1$$

für $A \in R^{0 \times 0}$.

(b) Es ist $\mathrm{S}_1 = \{(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})\}$, also

$$\det A = \sum_{\pi \in \mathrm{S}_1} (\operatorname{sgn} \pi) \prod_{j \in [1,1]} A_{\pi(j),j} = (\operatorname{sgn} (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})) A_{1,1} = A_{1,1}$$

für $A \in R^{1 \times 1}$.

(c) Es ist $\mathrm{S}_2 = \{(\begin{smallmatrix} 1 & 2 \\ 1 & 2 \end{smallmatrix}), (\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix})\}$, also

$$\det A = \sum_{\pi \in \mathrm{S}_2} (\operatorname{sgn} \pi) \prod_{j \in [1,2]} A_{\pi(j),j} = (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 \\ 1 & 2 \end{smallmatrix})) A_{1,1} A_{2,2} + (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix})) A_{2,1} A_{1,2} = A_{1,1} A_{2,2} - A_{1,2} A_{2,1}$$

für $A \in R^{2 \times 2}$.

(d) Es ist $\mathrm{S}_3 = \{(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}), (\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}), (\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}), (\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}), (\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}), (\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix})\}$ und damit

$$\begin{aligned} \det A &= \sum_{\pi \in \mathrm{S}_3} (\operatorname{sgn} \pi) \prod_{j \in [1,3]} A_{\pi(j),j} \\ &= (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix})) A_{1,1} A_{2,2} A_{3,3} + (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix})) A_{2,1} A_{1,2} A_{3,3} + (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix})) A_{3,1} A_{2,2} A_{1,3} \\ &\qquad + (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix})) A_{1,1} A_{3,2} A_{2,3} + (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix})) A_{2,1} A_{3,2} A_{1,3} + (\operatorname{sgn} (\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix})) A_{3,1} A_{1,2} A_{2,3} \\ &= A_{1,1} A_{2,2} A_{3,3} - A_{2,1} A_{1,2} A_{3,3} - A_{3,1} A_{2,2} A_{1,3} - A_{1,1} A_{3,2} A_{2,3} + A_{2,1} A_{3,2} A_{1,3} + A_{3,1} A_{1,2} A_{2,3} \\ &= A_{1,1} A_{2,2} A_{3,3} + A_{1,2} A_{2,3} A_{3,1} + A_{1,3} A_{2,1} A_{3,2} - A_{1,1} A_{2,3} A_{3,2} - A_{1,2} A_{2,1} A_{3,3} - A_{1,3} A_{2,2} A_{3,1} \end{aligned}$$

für $A \in R^{3 \times 3}$. □

(5.3) Beispiel.

(a) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

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

Dann ist

$$\det A = 7.$$

(b) Es sei $A \in \mathbb{R}^{3 \times 3}$ gegeben durch

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

Dann ist

$$\det A = -12.$$

Beweis.

(a) Nach Beispiel (5.2)(c) ist

$$\det A = \det \begin{pmatrix} 2 & -3 \\ 1 & 2 \end{pmatrix} = 2 \cdot 2 - (-3) \cdot 1 = 7.$$

(b) Nach Beispiel (5.2)(d) ist

$$\begin{aligned} \det A &= \det \begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} \\ &= 1 \cdot 3 \cdot 1 + (-1) \cdot (-1) \cdot 2 + 2 \cdot 4 \cdot (-1) - 1 \cdot (-1) \cdot (-1) - (-1) \cdot 4 \cdot 1 - 2 \cdot 3 \cdot 2 \\ &= 3 + 2 - 8 - 1 + 4 - 12 = -12. \end{aligned}$$

□

Die Leibniz-Formel für die Determinante ist asymmetrisch bzgl. der Rolle von Zeilen und Spalten einer Matrix. Die folgende Proposition zeigt, dass dies kein Charakteristikum der Determinante ist.

(5.4) Proposition. Für $A \in R^{n \times n}$ gilt

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

Beweis. Für $\pi \in \mathrm{S}_n$ ist $\operatorname{sgn} \pi = \operatorname{sgn} \pi^{-1}$ nach Korollar (A.105). Da $\mathrm{S}_n \to \mathrm{S}_n$, $\pi \mapsto \pi^{-1}$ eine Bijektion ist, folgt

$$\begin{aligned} \det A^{\mathrm{tr}} &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} A^{\mathrm{tr}}_{\pi(j),j} = \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} A_{j,\pi(j)} = \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi^{-1}) \prod_{j \in [1,n]} A_{\pi^{-1}(j),j} \\ &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} A_{\pi(j),j} = \det A \end{aligned}$$

für $A \in R^{n \times n}$. □

Determinante als normierte alternierende Multilinearform

Im Folgenden werden wir eine Matrix oft als ein Tupel ihrer Spalten auffassen. Nach Proposition (5.4) könnten wir auch die Zeilen einer Matrix betrachten und würden analoge Aussagen erhalten. Insbesondere gibt uns dann Korollar (5.6) eine Antwort auf die Frage, wie sich die Determinante ändert, wenn wir auf eine Matrix eine elementare Zeilenoperation anwenden.

(5.5) Proposition. Die Determinante auf $R^{n \times n}$ ist eine *normierte, alternierende Multilinearform* (*bzgl. der Spalten*) auf $R^{n \times n}$, d.h. es gelten folgende Eigenschaften.

- *Multilinearität in den Spalten.* Für $k \in [1, n]$ und jede Familie $(x_j)_{j \in [1,n] \setminus \{k\}}$ in $R^{n \times 1}$ ist

  $$R^{n \times 1} \to R,\, y \mapsto \det \begin{pmatrix} x_1 & \dots & x_{k-1} & y & x_{k+1} & \dots & x_n \end{pmatrix}$$

  linear ([4]), d.h. für $y, z \in R^{n \times 1}$ gilt

  $$\begin{aligned} &\det \begin{pmatrix} x_1 & \dots & x_{k-1} & y+z & x_{k+1} & \dots & x_n \end{pmatrix} \\ &= \det \begin{pmatrix} x_1 & \dots & x_{k-1} & y & x_{k+1} & \dots & x_n \end{pmatrix} + \det \begin{pmatrix} x_1 & \dots & x_{k-1} & z & x_{k+1} & \dots & x_n \end{pmatrix}, \end{aligned}$$

  und für $a \in R$, $y \in R^{n \times 1}$ gilt

  $$\det \begin{pmatrix} x_1 & \dots & x_{k-1} & ay & x_{k+1} & \dots & x_n \end{pmatrix} = a \det \begin{pmatrix} x_1 & \dots & x_{k-1} & y & x_{k+1} & \dots & x_n \end{pmatrix}.$$

- *Alternierend in den Spalten.* Für jedes n-Tupel $(x_1, \dots, x_n)$ in $R^{n \times 1}$ und $k, l \in [1, n]$ mit $k \neq l$, $x_k = x_l$ ist

  $$\det \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} = 0.$$

- *Normiertheit.* Es ist

  $$\det \mathrm{E}_n = 1.$$

Beweis. Für die Multilinearität seien $k \in [1, n]$ und eine Familie $(x_j)_{j \in [1,n] \setminus \{k\}}$ in $R^{n \times 1}$ gegeben. Dann gilt

$$\begin{aligned} &\det \begin{pmatrix} x_1 & \dots & x_{k-1} & y+z & x_{k+1} & \dots & x_n \end{pmatrix} \\ &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot (y+z)_{\pi(k)} \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)} \\ &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot (y_{\pi(k)} + z_{\pi(k)}) \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)} \\ &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot y_{\pi(k)} \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)} \\ &\qquad + \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot z_{\pi(k)} \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)} \\ &= \det \begin{pmatrix} x_1 & \dots & x_{k-1} & y & x_{k+1} & \dots & x_n \end{pmatrix} + \det \begin{pmatrix} x_1 & \dots & x_{k-1} & z & x_{k+1} & \dots & x_n \end{pmatrix} \end{aligned}$$

für $y, z \in R^{n \times 1}$ sowie

$$\begin{aligned} &\det \begin{pmatrix} x_1 & \dots & x_{k-1} & by & x_{k+1} & \dots & x_n \end{pmatrix} = \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) (\prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot (by)_{\pi(k)} \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)}) \\ &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot b y_{\pi(k)} \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)} \\ &= b \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,k-1]} (x_j)_{\pi(j)} \cdot y_{\pi(k)} \cdot \prod_{j \in [k+1,n]} (x_j)_{\pi(j)} = b \det \begin{pmatrix} x_1 & \dots & x_{k-1} & y & x_{k+1} & \dots & x_n \end{pmatrix} \end{aligned}$$

für $b \in R$, $y \in R^{n \times 1}$. Folglich ist

$$R^{n \times 1} \to R,\, y \mapsto \det \begin{pmatrix} x_1 & \dots & x_{k-1} & y & x_{k+1} & \dots & x_n \end{pmatrix}$$

linear. Da $k \in [1, n]$ und die Familie $(x_j)_{j \in [1,n] \setminus \{k\}}$ in $R^{n \times 1}$ beliebig gewählt waren, ist det eine Multilinearform auf $R^{n \times n}$.

[4] Falls R ein Körper ist, so haben wir also den üblichen Begriff eines R-Vektorraumhomomorphismus. Falls R kein Körper ist, so würde man von einem *R-Modulhomomorphismus* sprechen.

Als nächstes seien ein n-Tupel $(x_1, \dots, x_n)$ in $R^{n \times 1}$ und $k, l \in [1, n]$ mit $k \neq l$ und $x_k = x_l$ gegeben. Wir definieren $\tau \in \mathrm{S}_n$ durch $\tau := (k, l)$. Dann ist

$$\begin{aligned}\prod_{j \in [1,n]} (x_j)_{\pi(\tau(j))} &= (\prod_{j \in [1,n] \setminus \{k,l\}} (x_j)_{\pi(\tau(j))})(x_k)_{\pi(\tau(k))}(x_l)_{\pi(\tau(l))} = (\prod_{j \in [1,n] \setminus \{k,l\}} (x_j)_{\pi(j)})(x_k)_{\pi(l)}(x_l)_{\pi(k)} \\ &= (\prod_{j \in [1,n] \setminus \{k,l\}} (x_j)_{\pi(j)})(x_l)_{\pi(l)}(x_k)_{\pi(k)} = \prod_{j \in [1,n]} (x_j)_{\pi(j)}.\end{aligned}$$

Nach dem Produktsatz (A.104) gilt für $\pi \in \mathrm{S}_n$ stets

$$\operatorname{sgn}(\pi \circ \tau) = (\operatorname{sgn} \pi)(\operatorname{sgn} \tau) = (\operatorname{sgn} \pi)(-1) = -\operatorname{sgn} \pi.$$

Folglich ist

$$\{\pi \in \mathrm{S}_n \mid \operatorname{sgn} \pi = 1\} \to \{\sigma \in \mathrm{S}_n \mid \operatorname{sgn} \sigma = -1\},\ \pi \mapsto \pi \circ \tau$$

eine wohldefinierte Bijektion. Da R kommutativ ist, erhalten wir

$$\begin{aligned}\det \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} (x_j)_{\pi(j)} \\ &= \sum_{\substack{\pi \in \mathrm{S}_n \\ \operatorname{sgn} \pi = 1}} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} (x_j)_{\pi(j)} + \sum_{\substack{\sigma \in \mathrm{S}_n \\ \operatorname{sgn} \sigma = -1}} (\operatorname{sgn} \sigma) \prod_{j \in [1,n]} (x_j)_{\sigma(j)} \\ &= \sum_{\substack{\pi \in \mathrm{S}_n \\ \operatorname{sgn} \pi = 1}} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} (x_j)_{\pi(j)} + \sum_{\substack{\pi \in \mathrm{S}_n \\ \operatorname{sgn} \pi = 1}} \operatorname{sgn}(\pi \circ \tau) \prod_{j \in [1,n]} (x_j)_{\pi(\tau(j))} \\ &= \sum_{\substack{\pi \in \mathrm{S}_n \\ \operatorname{sgn} \pi = 1}} \prod_{j \in [1,n]} (x_j)_{\pi(j)} + \sum_{\substack{\pi \in \mathrm{S}_n \\ \operatorname{sgn} \pi = 1}} (-1) \prod_{j \in [1,n]} (x_j)_{\pi(\tau(j))} \\ &= \sum_{\substack{\pi \in \mathrm{S}_n \\ \operatorname{sgn} \pi = 1}} (\prod_{j \in [1,n]} (x_j)_{\pi(j)} - \prod_{j \in [1,n]} (x_j)_{\pi(\tau(j))}) = 0.\end{aligned}$$

Folglich ist det alternierend.
Schließlich gilt

$$\prod_{j \in [1,n]} (\mathrm{E}_n)_{\pi(j),j} = \prod_{j \in [1,n]} \delta_{\pi(j),j} = \left.\begin{cases} 1, & \text{falls } \pi(j) = j \text{ für alle } j \in [1,n], \\ 0, & \text{falls } \pi(j) \neq j \text{ für ein } j \in [1,n], \end{cases}\right\} = \begin{cases} 1, & \text{falls } \pi = \mathrm{id}_{[1,n]}, \\ 0, & \text{falls } \pi \neq \mathrm{id}_{[1,n]}, \end{cases}$$

und damit

$$\det \mathrm{E}_n = \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} (\mathrm{E}_n)_{\pi(j),j} = \operatorname{sgn} \mathrm{id}_{[1,n]} = 1.$$

□

Es lässt sich zeigen, dass die Eigenschaften aus Proposition (5.5) die Determinante eindeutig charakterisieren: Ist $d \colon R^{n \times n} \to R$ eine beliebige normierte, alternierende Multilinearform auf $R^{n \times n}$, so gilt bereits $d = \det$.

(5.6) Korollar. Es sei ein n-Tupel $(x_1, \dots, x_n)$ in $R^{n \times 1}$ gegeben.

(a) Für $k, l \in [1, n]$ mit $k \neq l$ gilt

$$\det \begin{pmatrix} x_1 & \dots & x_{k-1} & x_l & x_{k+1} & \dots & x_{l-1} & x_k & x_{l+1} & \dots & x_n \end{pmatrix} = -\det \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

(b) Für $k, l \in [1, n]$ mit $k \neq l$ und $a \in R$ gilt

$$\det \begin{pmatrix} x_1 & \dots & x_{k-1} & x_k + a x_l & x_{k+1} & \dots & x_n \end{pmatrix} = \det \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

(c) Für $k \in [1, n]$ und $a \in R^{\times}$ gilt

$$\det \begin{pmatrix} x_1 & \dots & x_{k-1} & a x_k & x_{k+1} & \dots & x_n \end{pmatrix} = a \det \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

Beweis.

(a) Da die Determinante auf $R^{n \times n}$ nach Proposition (5.5) eine alternierende Multilinearform ist, gilt

$$\begin{aligned}
0 &= \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_k + x_l & x_{k+1} & \dots & x_{l-1} & x_k + x_l & x_{l+1} & \dots & x_n \end{pmatrix} \\
&= \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_k & x_{k+1} & \dots & x_{l-1} & x_k & x_{l+1} & \dots & x_n \end{pmatrix} \\
&\quad + \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_k & x_{k+1} & \dots & x_{l-1} & x_l & x_{l+1} & \dots & x_n \end{pmatrix} \\
&\quad + \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_l & x_{k+1} & \dots & x_{l-1} & x_k & x_{l+1} & \dots & x_n \end{pmatrix} \\
&\quad + \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_l & x_{k+1} & \dots & x_{l-1} & x_l & x_{l+1} & \dots & x_n \end{pmatrix} \\
&= \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} + \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_l & x_{k+1} & \dots & x_{l-1} & x_k & x_{l+1} & \dots & x_n \end{pmatrix}
\end{aligned}$$

und damit

$$\det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_l & x_{k+1} & \dots & x_{l-1} & x_k & x_{l+1} & \dots & x_n \end{pmatrix} = -\det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

(b) Da die Determinante auf $R^{n \times n}$ nach Proposition (5.5) eine alternierende Multilinearform ist, gilt für $k, l \in [1, n]$ mit $k \neq l$ und $a \in R$ stets

$$\begin{aligned}
&\det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_k + a x_l & x_{k+1} & \dots & x_n \end{pmatrix} \\
&= \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_k & x_{k+1} & \dots & x_n \end{pmatrix} + a \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_l & x_{k+1} & \dots & x_n \end{pmatrix} \\
&= \det\begin{pmatrix} x_1 & \dots & x_{k-1} & x_k & x_{k+1} & \dots & x_n \end{pmatrix}.
\end{aligned}$$

(c) Da die Determinante auf $R^{n \times n}$ nach Proposition (5.5) eine alternierende Multilinearform ist, gilt für $k \in [1, n]$ und $a \in R^{\times}$ stets

$$\det\begin{pmatrix} x_1 & \dots & x_{k-1} & a x_k & x_{k+1} & \dots & x_n \end{pmatrix} = a \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

□

(5.7) Korollar. Es sei ein n-Tupel $(x_1, \dots, x_n)$ in $R^{n \times 1}$ gegeben. Für $\pi \in \mathrm{S}_n$ ist

$$\det\begin{pmatrix} x_{\pi(1)} & \dots & x_{\pi(n)} \end{pmatrix} = (\operatorname{sgn} \pi) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

Beweis. Wir zeigen durch Induktion nach $l \in \mathbb{N}_0$: Für $\pi \in \mathrm{S}_n$ so, dass π ein Kompositum von l Transpositionen ist, gilt

$$\det\begin{pmatrix} x_{\pi(1)} & \dots & x_{\pi(n)} \end{pmatrix} = (\operatorname{sgn} \pi) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.$$

Für $l = 0$ gilt $\pi = \mathrm{id}_{[1,n]}$, also

$$\begin{aligned}
\det\begin{pmatrix} x_{\pi(1)} & \dots & x_{\pi(n)} \end{pmatrix} &= \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} = (\operatorname{sgn} \mathrm{id}_{[1,n]}) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \\
&= (\operatorname{sgn} \pi) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.
\end{aligned}$$

Es sei also $l \in \mathbb{N}$ gegeben und es sei angenommen, dass

$$\det\begin{pmatrix} x_{\pi'(1)} & \dots & x_{\pi'(n)} \end{pmatrix} = (\operatorname{sgn} \pi') \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$$

für $\pi' \in \mathrm{S}_n$ so, dass π' ein Kompositum von $l-1$ Transpositionen ist. Ferner sei $\pi \in \mathrm{S}_n$ gegeben und es sei angenommen, dass π ein Kompositum von l Transpositionen ist. Dann ist $\pi = \pi' \circ \tau$ für ein $\pi' \in \mathrm{S}_n$, welches ein Kompositum von $l-1$ Transpositionen ist, und eine Transposition $\tau \in \mathrm{S}_n$. Unter Ausnutzung der Induktionsvoraussetzung, Korollar (5.6)(a) und dem Produktsatz (A.104), erhalten wir

$$\begin{aligned}
\det\begin{pmatrix} x_{\pi(1)} & \dots & x_{\pi(n)} \end{pmatrix} &= \det\begin{pmatrix} x_{\pi'(\tau(1))} & \dots & x_{\pi'(\tau(n))} \end{pmatrix} = (\operatorname{sgn} \pi') \det\begin{pmatrix} x_{\tau(1)} & \dots & x_{\tau(n)} \end{pmatrix} \\
&= (\operatorname{sgn} \pi') \left(-\det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}\right) = (\operatorname{sgn} \pi')(\operatorname{sgn} \tau) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \\
&= \operatorname{sgn}(\pi' \circ \tau) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} = (\operatorname{sgn} \pi) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}.
\end{aligned}$$

Nach dem Induktionsprinzip gilt somit

$$\det\begin{pmatrix} x_{\pi(1)} & \dots & x_{\pi(n)} \end{pmatrix} = (\operatorname{sgn} \pi) \det\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$$

für alle $\pi \in \mathrm{S}_n$.

□

Mit Hilfe von Korollar (5.6) und Proposition (5.4) lässt sich die Berechnung der Determinante vereinfachen, indem wir zunächst elementare Zeilenoperationen anwenden. Hierbei müssen wir beachten, dass Vertauschungs- und Multiplikationsoperationen die Determinante verändern; diese Operationen müssen wir also dokumentieren.

Alternativer Beweis von Beispiel (5.3)(b). Zunächst wenden wir Additionsoperationen auf die Zeilen von A an:

$$\begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} \xmapsto{\mathrm{add}_{3,1,-2} \circ \mathrm{add}_{2,1,-4}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 7 & -9 \\ 0 & 1 & -3 \end{pmatrix} \xmapsto{\mathrm{add}_{2,3,-7}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix}$$

Nach Proposition (5.4), Korollar (5.6) und Beispiel (5.2)(d) ist also

$$\begin{aligned} \det A &= \det \begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix} \\ &= 1 \cdot 0 \cdot (-3) + (-1) \cdot 12 \cdot 0 + 2 \cdot 0 \cdot 1 - 1 \cdot 12 \cdot 1 - (-1) \cdot 0 \cdot (-3) - 2 \cdot 0 \cdot 0 = -12. \end{aligned}$$

□

Der Produktsatz

Als nächstes zeigen wir die Multiplikativität der Determinante.

(5.8) Proposition (Produktsatz)**.**

(a) Für $A, B \in R^{n \times n}$ gilt
$$\det(BA) = (\det B)(\det A).$$

(b) Es ist
$$\det \mathrm{E}_n = 1.$$

(c) Für $A \in \mathrm{GL}_n(R)$ ist $\det A \in R^\times$ mit
$$(\det A)^{-1} = \det A^{-1}.$$

Beweis.

(a) Es seien $A, B \in R^{n \times n}$ gegeben. Da die Determinante nach Proposition (5.5) multilinear ist, haben wir
$$\begin{aligned} \det(BA) &= \det \begin{pmatrix} BA_{-,1} & \dots & BA_{-,n} \end{pmatrix} = \det \begin{pmatrix} \sum_{j \in [1,n]} A_{j,1} B_{-,j} & \dots & \sum_{j \in [1,n]} A_{j,n} B_{-,j} \end{pmatrix} \\ &= \det \begin{pmatrix} \sum_{j_1 \in [1,n]} A_{j_1,1} B_{-,j_1} & \dots & \sum_{j_n \in [1,n]} A_{j_n,n} B_{-,j_n} \end{pmatrix} \\ &= \sum_{(j_1,\dots,j_n) \in [1,n]^n} \big(\prod_{k \in [1,n]} A_{j_k,k} \big) \det \begin{pmatrix} B_{-,j_1} & \dots & B_{-,j_n} \end{pmatrix}. \end{aligned}$$

Ferner ist die Determinante nach Proposition (5.5) alternierend, es gilt für $(j_1, \dots, j_n) \in [1,n]^n$ also $\det \begin{pmatrix} B_{-,j_1} & \dots & B_{-,j_n} \end{pmatrix} = 0$, sofern $[1,n] \to [1,n]$, $k \mapsto j_k$ nicht bijektiv ist. Nach Korollar (5.7) folgt
$$\begin{aligned} \det(BA) &= \sum_{(j_1,\dots,j_n) \in [1,n]^n} \big(\prod_{k \in [1,n]} A_{j_k,k} \big) \det \begin{pmatrix} B_{-,j_1} & \dots & B_{-,j_n} \end{pmatrix} \\ &= \sum_{\pi \in \mathrm{S}_n} \big(\prod_{k \in [1,n]} A_{\pi(k),k} \big) \det \begin{pmatrix} B_{-,\pi(1)} & \dots & B_{-,\pi(n)} \end{pmatrix} \\ &= \sum_{\pi \in \mathrm{S}_n} \big(\prod_{k \in [1,n]} A_{\pi(k),k} \big) (\operatorname{sgn} \pi) \det \begin{pmatrix} B_{-,1} & \dots & B_{-,n} \end{pmatrix} \\ &= (\det B) \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \big(\prod_{k \in [1,n]} A_{\pi(k),k} \big) = (\det B)(\det A). \end{aligned}$$

(b) Dies gilt nach Proposition (5.5).

(c) Es sei $A \in \mathrm{GL}_n(R)$ gegeben. Nach (a) gilt dann
$$1 = \det \mathrm{E}_n = \det(A^{-1} A) = (\det A^{-1})(\det A).$$
Wegen der Kommutativität von R ist folglich $\det A \in R^\times$ mit $(\det A)^{-1} = \det A^{-1}$. □

Die Determinante eines Vektorraumendomorphismus

Schließlich führen wir den Begriff der Determinante eines Endomorphismus auf den der Determinante einer Matrix zurück.

(5.9) Bemerkung. Es seien ein Körper K, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und Basen $s = (s_1, \dots, s_n)$ und $s' = (s'_1, \dots, s'_n)$ von V gegeben. Dann gilt

$$\det \mathrm{M}_{s,s}(\varphi) = \det \mathrm{M}_{s',s'}(\varphi).$$

Beweis. Nach Bemerkung (3.63)(b) ist $\mathrm{M}_{s,s}(\varphi)$ ähnlich zu $\mathrm{M}_{s',s'}(\varphi)$, so dass es ein $P \in \mathrm{GL}_n(K)$ mit

$$P^{-1} \,\mathrm{M}_{s,s}(\varphi)\, P = \mathrm{M}_{s',s'}(\varphi)$$

gibt. Wir erhalten

$$\det \mathrm{M}_{s',s'}(\varphi) = \det(P^{-1} \,\mathrm{M}_{s,s}(\varphi)\, P) = (\det P)^{-1}(\det \mathrm{M}_{s,s}(\varphi))(\det P) = \det \mathrm{M}_{s,s}(\varphi)$$

nach Proposition (5.8)(a), (c) und der Kommutativität von K. □

(5.10) Definition (Determinante)**.** Es seien ein Körper K und ein endlichdimensionaler K-Vektorraum V gegeben. Die *Determinante* auf $\mathrm{End}(V)$ ist die Abbildung $\det\colon \mathrm{End}(V) \to K$ gegeben durch

$$\det \varphi = \det \mathrm{M}_{s,s}(\varphi),$$

für $\varphi \in \mathrm{End}(V)$, wobei s eine beliebige Basis von V bezeichne.

Laplace-Entwicklung

Die Berechnung von Determinanten größerer Matrizen ist rekursiv mit der sogenannten Laplace-Entwicklung möglich. Zur Formulierung benötigen wir den folgenden Begriff:

(5.11) Definition (Minor)**.** Es seien $A \in R^{n \times n}$ sowie $k, l \in [1, n]$ gegeben. Der *Minor* von A an der Stelle (k, l) ist definiert als

$$\mathrm{Minor}_{k,l}(A) := \det \begin{pmatrix} A_{1,1} & \dots & A_{1,l-1} & A_{1,l+1} & \dots & A_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{k-1,1} & \dots & A_{k-1,l-1} & A_{k-1,l+1} & \dots & A_{k-1,n} \\ A_{k+1,1} & \dots & A_{k+1,l-1} & A_{k+1,l+1} & \dots & A_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,l-1} & A_{n,l+1} & \dots & A_{n,n} \end{pmatrix}$$

(5.12) Beispiel. Es sei $A \in \mathbb{Q}^{3 \times 3}$ gegeben durch

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

Dann ist

$$\mathrm{Minor}_{2,1}(A) = 1.$$

Beweis. Es ist

$$\mathrm{Minor}_{2,1}(A) = \det \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix} = (-1) \cdot 1 - 2 \cdot (-1) = 1.$$ □

(5.13) Satz (Laplace-Entwicklung)**.** Es sei $A \in R^{n \times n}$ gegeben.

(a) *Entwicklung nach der l-ten Spalte.* Für $l \in [1, n]$ gilt

$$\det A = \sum_{k \in [1,n]} (-1)^{k+l} A_{k,l} \mathrm{Minor}_{k,l}(A).$$

(b) *Entwicklung nach der k-ten Zeile.* Für $k \in [1, n]$ gilt

$$\det A = \sum_{l \in [1,n]} (-1)^{k+l} A_{k,l} \mathrm{Minor}_{k,l}(A).$$

Beweis.

(a) Es sei $l \in [1, n]$ gegeben.
Zunächst sei $A_{-,l} = \mathrm{e}_k$ für ein $k \in [1, n]$. Wir definieren $B \in R^{(n-1)\times(n-1)}$ durch

$$B := \begin{pmatrix} A_{1,1} & \dots & A_{1,l-1} & A_{1,l+1} & \dots & A_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{k-1,1} & \dots & A_{k-1,l-1} & A_{k-1,l+1} & \dots & A_{k-1,n} \\ A_{k+1,1} & \dots & A_{k+1,l-1} & A_{k+1,l+1} & \dots & A_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,l-1} & A_{n,l+1} & \dots & A_{n,n} \end{pmatrix},$$

so dass

$$B_{i,j} = \begin{cases} A_{i,j}, & \text{falls } i \in [1, k-1],\ j \in [1, l-1], \\ A_{i+1,j}, & \text{falls } i \in [k, n-1],\ j \in [1, l-1], \\ A_{i,j+1}, & \text{falls } i \in [1, k-1],\ j \in [l, n-1], \\ A_{i+1,j+1}, & \text{falls } i \in [k, n-1],\ j \in [l, n-1], \end{cases}$$

für $i, j \in [1, n]$ gilt. Für $\pi \in \mathrm{S}_n$ gilt $A_{\pi(l),l} = (\mathrm{e}_k)_{\pi(l)} = \delta_{\pi(l),k}$, so dass wir

$$\begin{aligned} \det A &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \prod_{j \in [1,n]} A_{\pi(j),j} = \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \Big(\prod_{j \in [1,l-1]} A_{\pi(j),j} \Big) A_{\pi(l),l} \Big(\prod_{j \in [l+1,n]} A_{\pi(j),j} \Big) \\ &= \sum_{\pi \in \mathrm{S}_n} (\operatorname{sgn} \pi) \Big(\prod_{j \in [1,l-1]} A_{\pi(j),j} \Big) \delta_{\pi(l),k} \Big(\prod_{j \in [l+1,n]} A_{\pi(j),j} \Big) \\ &= \sum_{\substack{\pi \in \mathrm{S}_n \\ \pi(l)=k}} (\operatorname{sgn} \pi) \Big(\prod_{j \in [1,l-1]} A_{\pi(j),j} \Big) \Big(\prod_{j \in [l+1,n]} A_{\pi(j),j} \Big) \end{aligned}$$

erhalten. Nun ist

$$\begin{aligned} f \colon \{\sigma \in \mathrm{S}_n \mid \sigma(n) = n\} &\to \{\pi \in \mathrm{S}_n \mid \pi(l) = k\}, \\ \sigma &\mapsto \left(\begin{smallmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n-1 & n \\ 1 & \dots & k-1 & k+1 & k+2 & \dots & n & k \end{smallmatrix}\right) \circ \sigma \circ \left(\begin{smallmatrix} 1 & \dots & l-1 & l & l+1 & \dots & n-1 & n \\ 1 & \dots & l-1 & l+1 & l+2 & \dots & n & l \end{smallmatrix}\right)^{-1} \end{aligned}$$

eine Bijektion. Für $\sigma \in \mathrm{S}_n$ mit $\sigma(n) = n$ gilt

$$\begin{aligned} \operatorname{sgn} f(\sigma) &= \operatorname{sgn}\Big(\left(\begin{smallmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n-1 & n \\ 1 & \dots & k-1 & k+1 & k+2 & \dots & n & k \end{smallmatrix}\right) \circ \sigma \circ \left(\begin{smallmatrix} 1 & \dots & l-1 & l & l+1 & \dots & n-1 & n \\ 1 & \dots & l-1 & l+1 & l+2 & \dots & n & l \end{smallmatrix}\right)^{-1}\Big) \\ &= \Big(\operatorname{sgn} \left(\begin{smallmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n-1 & n \\ 1 & \dots & k-1 & k+1 & k+2 & \dots & n & k \end{smallmatrix}\right)\Big)(\operatorname{sgn} \sigma)\Big(\operatorname{sgn} \left(\begin{smallmatrix} 1 & \dots & l-1 & l & l+1 & \dots & n-1 & n \\ 1 & \dots & l-1 & l+1 & l+2 & \dots & n & l \end{smallmatrix}\right)\Big) \\ &= (-1)^{n-k} (\operatorname{sgn} \sigma) (-1)^{n-l} = (-1)^{k+l} (\operatorname{sgn} \sigma) \end{aligned}$$

sowie

$$(f(\sigma))(j) = \begin{cases} \sigma(j), & \text{falls } j \in [1, l-1],\ \sigma(j) \in [1, k-1], \\ \sigma(j) + 1, & \text{falls } j \in [1, l-1],\ \sigma(j) \in [k, n-1], \\ k, & \text{falls } j = l, \\ \sigma(j-1), & \text{falls } j \in [l+1, n],\ \sigma(j-1) \in [1, k-1], \\ \sigma(j-1) + 1, & \text{falls } j \in [l+1, n],\ \sigma(j-1) \in [k, n-1], \end{cases}$$

für $j \in [1, n]$. Es folgt

$$\begin{aligned}
\det A &= \sum_{\substack{\pi \in \mathrm{S}_n \\ \pi(l) = k}} (\operatorname{sgn} \pi)(\prod_{j \in [1, l-1]} A_{\pi(j), j})(\prod_{j \in [l+1, n]} A_{\pi(j), j}) \\
&= \sum_{\substack{\sigma \in \mathrm{S}_n \\ \sigma(n) = n}} (\operatorname{sgn} f(\sigma))(\prod_{j \in [1, l-1]} A_{(f(\sigma))(j), j})(\prod_{j \in [l+1, n]} A_{(f(\sigma))(j), j}) \\
&= \sum_{\substack{\sigma \in \mathrm{S}_n \\ \sigma(n) = n}} (-1)^{k+l} (\operatorname{sgn} \sigma)(\prod_{\substack{j \in [1, l-1] \\ \sigma(j) \in [1, k-1]}} A_{\sigma(j), j})(\prod_{\substack{j \in [1, l-1] \\ \sigma(j) \in [k, n-1]}} A_{\sigma(j)+1, j})(\prod_{\substack{j \in [l+1, n] \\ \sigma(j-1) \in [1, k-1]}} A_{\sigma(j-1), j}) \\
&\quad (\prod_{\substack{j \in [l+1, n] \\ \sigma(j-1) \in [k, n-1]}} A_{\sigma(j-1)+1, j}) \\
&= (-1)^{k+l} \sum_{\substack{\sigma \in \mathrm{S}_n \\ \sigma(n) = n}} (\operatorname{sgn} \sigma)(\prod_{\substack{j \in [1, l-1] \\ \sigma(j) \in [1, k-1]}} A_{\sigma(j), j})(\prod_{\substack{j \in [1, l-1] \\ \sigma(j) \in [k, n-1]}} A_{\sigma(j)+1, j})(\prod_{\substack{j \in [l, n-1] \\ \sigma(j) \in [1, k-1]}} A_{\sigma(j), j+1}) \\
&\quad (\prod_{\substack{j \in [l, n-1] \\ \sigma(j) \in [k, n-1]}} A_{\sigma(j)+1, j+1}) \\
&= (-1)^{k+l} \sum_{\sigma \in \mathrm{S}_{n-1}} (\operatorname{sgn} \sigma)(\prod_{\substack{j \in [1, l-1] \\ \sigma(j) \in [1, k-1]}} B_{\sigma(j), j})(\prod_{\substack{j \in [1, l-1] \\ \sigma(j) \in [k, n-1]}} B_{\sigma(j), j})(\prod_{\substack{j \in [l, n-1] \\ \sigma(j) \in [1, k-1]}} B_{\sigma(j), j}) \\
&\quad (\prod_{\substack{j \in [l, n-1] \\ \sigma(j) \in [k, n-1]}} B_{\sigma(j), j}) \\
&= (-1)^{k+l} \sum_{\sigma \in \mathrm{S}_{n-1}} (\operatorname{sgn} \sigma)(\prod_{j \in [1, n-1]} B_{\sigma(j), j}) = (-1)^{k+l} \det B = (-1)^{k+l} \operatorname{Minor}_{k,l}(A).
\end{aligned}$$

Nun sei $A \in R^{n \times n}$ beliebig. Dann folgt

$$\begin{aligned}
\det A &= \det \begin{pmatrix} A_{-,1} & \dots & A_{-,l-1} & A_{-,l} & A_{-,l+1} & \dots & A_{-,n} \end{pmatrix} \\
&= \det \begin{pmatrix} A_{-,1} & \dots & A_{-,l-1} & \sum_{k \in [1,n]} A_{k,l} \mathrm{e}_k & A_{-,l+1} & \dots & A_{-,n} \end{pmatrix} \\
&= \sum_{k \in [1,n]} A_{k,l} \det \begin{pmatrix} A_{-,1} & \dots & A_{-,l-1} & \mathrm{e}_k & A_{-,l+1} & \dots & A_{-,n} \end{pmatrix} \\
&= \sum_{k \in [1,n]} A_{k,l} (-1)^{k+l} \operatorname{Minor}_{k,l}(\begin{pmatrix} A_{-,1} & \dots & A_{-,l-1} & \mathrm{e}_k & A_{-,l+1} & \dots & A_{-,n} \end{pmatrix}) \\
&= \sum_{k \in [1,n]} (-1)^{k+l} A_{k,l} \operatorname{Minor}_{k,l}(A).
\end{aligned}$$

(b) Es sei $k \in [1, n]$ gegeben. Nach Proposition (5.4) gilt

$$\operatorname{Minor}_{l,k}(A^{\mathrm{tr}}) = \det \begin{pmatrix} A^{\mathrm{tr}}_{1,1} & \dots & A^{\mathrm{tr}}_{1,k-1} & A^{\mathrm{tr}}_{1,k+1} & \dots & A^{\mathrm{tr}}_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A^{\mathrm{tr}}_{l-1,1} & \dots & A^{\mathrm{tr}}_{l-1,k-1} & A^{\mathrm{tr}}_{l-1,k+1} & \dots & A^{\mathrm{tr}}_{l-1,n} \\ A^{\mathrm{tr}}_{l+1,1} & \dots & A^{\mathrm{tr}}_{l+1,k-1} & A^{\mathrm{tr}}_{l+1,k+1} & \dots & A^{\mathrm{tr}}_{l+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A^{\mathrm{tr}}_{n,1} & \dots & A^{\mathrm{tr}}_{n,k-1} & A^{\mathrm{tr}}_{n,k+1} & \dots & A^{\mathrm{tr}}_{n,n} \end{pmatrix}$$

$$= \det \begin{pmatrix} A^{\mathrm{tr}}_{1,1} & \dots & A^{\mathrm{tr}}_{1,k-1} & A^{\mathrm{tr}}_{1,k+1} & \dots & A^{\mathrm{tr}}_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A^{\mathrm{tr}}_{l-1,1} & \dots & A^{\mathrm{tr}}_{l-1,k-1} & A^{\mathrm{tr}}_{l-1,k+1} & \dots & A^{\mathrm{tr}}_{l-1,n} \\ A^{\mathrm{tr}}_{l+1,1} & \dots & A^{\mathrm{tr}}_{l+1,k-1} & A^{\mathrm{tr}}_{l+1,k+1} & \dots & A^{\mathrm{tr}}_{l+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A^{\mathrm{tr}}_{n,1} & \dots & A^{\mathrm{tr}}_{n,k-1} & A^{\mathrm{tr}}_{n,k+1} & \dots & A^{\mathrm{tr}}_{n,n} \end{pmatrix}^{\mathrm{tr}}$$

$$= \det \begin{pmatrix} A^{\mathrm{tr}}_{1,1} & \dots & A^{\mathrm{tr}}_{l-1,1} & A^{\mathrm{tr}}_{l+1,1} & \dots & A^{\mathrm{tr}}_{n,1} \\ \vdots & & \vdots & \vdots & & \vdots \\ A^{\mathrm{tr}}_{1,k-1} & \dots & A^{\mathrm{tr}}_{l-1,k-1} & A^{\mathrm{tr}}_{l+1,k-1} & \dots & A^{\mathrm{tr}}_{n,k-1} \\ A^{\mathrm{tr}}_{1,k+1} & \dots & A^{\mathrm{tr}}_{l-1,k+1} & A^{\mathrm{tr}}_{l+1,k+1} & \dots & A^{\mathrm{tr}}_{n,k+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ A^{\mathrm{tr}}_{1,n} & \dots & A^{\mathrm{tr}}_{l-1,n} & A^{\mathrm{tr}}_{l+1,n} & \dots & A^{\mathrm{tr}}_{n,n} \end{pmatrix}$$

$$= \det \begin{pmatrix} A_{1,1} & \dots & A_{1,l-1} & A_{1,l+1} & \dots & A_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{k-1,1} & \dots & A_{k-1,l-1} & A_{k-1,l+1} & \dots & A_{k-1,n} \\ A_{k+1,1} & \dots & A_{k+1,l-1} & A_{k+1,l+1} & \dots & A_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,l-1} & A_{n,l+1} & \dots & A_{n,n} \end{pmatrix} = \mathrm{Minor}_{k,l}(A)$$

für $l \in [1, n]$ und damit

$$\det A = \det A^{\mathrm{tr}} = \sum_{l \in [1,n]} (-1)^{l+k} A^{\mathrm{tr}}_{l,k} \mathrm{Minor}_{l,k}(A^{\mathrm{tr}}) = \sum_{l \in [1,n]} (-1)^{k+l} A_{k,l} \mathrm{Minor}_{k,l}(A)$$

nach (a). □

Alternativer Beweis von Beispiel (5.3)(b). Eine Laplace-Entwicklung (5.13)(a) nach der ersten Spalte liefert

$$\det A = \det \begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} = 1 \det \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix} - 4 \det \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix} + 2 \det \begin{pmatrix} -1 & 2 \\ 3 & -1 \end{pmatrix}$$
$$= 1(3 \cdot 1 - (-1)(-1)) - 4((-1) \cdot 1 - 2(-1)) + 2((-1)(-1) - 2 \cdot 3) = 1 \cdot 2 - 4 \cdot 1 + 2(-5) = -12.$$ □

Etwas weniger chaotisch gestaltet sich die Rechnung, wenn wir die Matrix zunächst durch elementare Zeilenoperationen vereinfachen. Hierbei gilt zu beachten, dass Vertauschungs- und Multiplikationsoperatoren die Determinante verändern, siehe Proposition (5.4) und Korollar (5.6).

Alternativer Beweis von Beispiel (5.3)(b). Zunächst wenden wir Additionsoperationen auf die Zeilen von A an:

$$\begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} \xmapsto{\mathrm{add}_{3,1,-2} \circ \mathrm{add}_{2,1,-4}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 7 & -9 \\ 0 & 1 & -3 \end{pmatrix} \xmapsto{\mathrm{add}_{2,3,-7}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix}$$

Nach Proposition (5.4) und Korollar (5.6) gilt

$$\det A = \det \begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix}.$$

Eine Laplace-Entwicklung (5.13)(a) liefert

$$\det A = \det \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix} = \det \begin{pmatrix} 0 & 12 \\ 1 & -3 \end{pmatrix} = -\det \begin{pmatrix} 12 \end{pmatrix} = -12.$$ □

De facto lässt sich die Determinante einer Matrix in Zeilenstufenform ablesen, siehe Korollar (5.16).

(5.14) Satz. Es ist

$$\mathrm{GL}_n(R) = \{A \in R^{n \times n} \mid \det A \in R^\times\}.$$

Beweis. Es sei $A \in R^{n \times n}$ gegeben. Wenn $A \in \mathrm{GL}_n(R)$ ist, dann ist $\det A \in R^\times$ nach Proposition (5.8)(c). Umgekehrt sei also $\det A \in R^\times$ und es sei $B \in R^{n \times n}$ gegeben durch

$$B = ((-1)^{i+j} \mathrm{Minor}_{j,i}(A))_{i,j \in [1,n]}.$$

Für $i, k \in [1, n]$ liefert eine Laplace-Entwicklung (5.13)(a) nach der i-ten Spalte

$$\begin{aligned}(BA)_{i,k} &= \sum_{j \in [1,n]} B_{i,j} A_{j,k} = \sum_{j \in [1,n]} A_{j,k} \, (-1)^{i+j} \mathrm{Minor}_{j,i}(A) \\ &= \sum_{j \in [1,n]} A_{j,k} \, (-1)^{i+j} \mathrm{Minor}_{j,i}(\begin{pmatrix} A_{-,1} & \dots & A_{-,i-1} & A_{-,k} & A_{-,i+1} & \dots & A_{-,n} \end{pmatrix}) \\ &= \det \begin{pmatrix} A_{-,1} & \dots & A_{-,i-1} & A_{-,k} & A_{-,i+1} & \dots & A_{-,n} \end{pmatrix} = \delta_{i,k}(\det A) = ((\det A) \, \mathrm{E}_n)_{i,k}.\end{aligned}$$

Folglich gilt $BA = (\det A)\mathrm{E}_n$ und damit $(\det A)^{-1} BA = \mathrm{E}_n$.
Analog lässt sich $A \cdot (\det A)^{-1} B = \mathrm{E}_n$ zeigen. Insgesamt ist $A \in \mathrm{GL}_n(R)$ mit $A^{-1} = (\det A)^{-1} B$. □

Der Kästchensatz

Als nächstes entwickeln wir einige praktische Rechentechniken für Determinanten von speziellen Matrizen.

(5.15) Proposition (Kästchensatz für Determinanten)**.** Es sei $m \in \mathbb{N}_0$ gegeben. Für $A \in R^{m \times m}$, $B \in R^{m \times n}$, $C \in R^{n \times n}$ gilt

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = (\det A)(\det C).$$

Beweis. Wir führen Induktion nach m.
Zunächst sei $m = 0$ und es seien $A \in R^{m \times m}$, $B \in R^{m \times n}$, $C \in R^{n \times n}$. Nach Beispiel (5.2)(a) ist $\det A = 1$ und damit

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det C = (\det A)(\det C).$$

Nun sei $m \in \mathbb{N}$ so gegeben, dass für $A' \in R^{(m-1) \times (m-1)}$, $B' \in R^{(m-1) \times n}$, $C \in R^{n \times n}$ stets

$$\det \begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix} = (\det A')(\det C)$$

gilt. Ferner seien $A \in R^{m \times m}$, $B \in R^{m \times n}$, $C \in R^{n \times n}$ gegeben. Nach der Induktionsvoraussetzung ist

$$\mathrm{Minor}_{k,1}(\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}) = \det \begin{pmatrix} A_{1,2} & \dots & A_{1,m} & B_{1,1} & \dots & B_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{k-1,2} & \dots & A_{k-1,m} & B_{k-1,1} & \dots & B_{k-1,n} \\ A_{k+1,2} & \dots & A_{k+1,m} & B_{k+1,1} & \dots & B_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{m,2} & \dots & A_{m,m} & B_{m,1} & \dots & B_{m,n} \\ 0 & \dots & 0 & C_{1,1} & \dots & C_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & C_{n,1} & \dots & C_{n,n} \end{pmatrix}$$

$$= (\det \begin{pmatrix} A_{1,2} & \dots & A_{1,m} \\ \vdots & & \vdots \\ A_{k-1,2} & \dots & A_{k-1,m} \\ A_{k+1,2} & \dots & A_{k+1,m} \\ \vdots & & \vdots \\ A_{m,2} & \dots & A_{m,m} \end{pmatrix})(\det \begin{pmatrix} C_{1,1} & \dots & C_{1,n} \\ \vdots & & \vdots \\ C_{n,1} & \dots & C_{n,n} \end{pmatrix}) = \operatorname{Minor}_{k,1}(A)(\det C)$$

für $k \in [1, m]$ und

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}_{k,1} = 0$$

für $k \in [m+1, m+n]$. Eine Laplace-Entwicklung (5.13)(a) nach der ersten Spalte liefert

$$\begin{aligned} \det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} &= \sum_{k \in [1,m+n]} (-1)^{k+1} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}_{k,1} \operatorname{Minor}_{k,1}(\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}) \\ &= \sum_{k \in [1,m]} (-1)^{k+1} A_{k,1} \operatorname{Minor}_{k,1}(A)(\det C) = (\sum_{k \in [1,m]} (-1)^{k+1} A_{k,1} \operatorname{Minor}_{k,1}(A))(\det C) \\ &= (\det A)(\det C). \end{aligned}$$

Nach dem Induktionsprinzip gilt

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = (\det A)(\det C)$$

für alle $A \in R^{m \times m}$, $B \in R^{m \times n}$, $C \in R^{n \times n}$, $m \in \mathbb{N}_0$. □

(5.16) Korollar. Für jede obere Dreiecksmatrix $A \in R^{n \times n}$ gilt

$$\det A = \prod_{j \in [1,n]} A_{j,j}.$$

Beweis. Dies folgt aus dem Kästchensatz für Determinanten (5.15) und Induktion. □

Nach Korollar (5.16) lassen sich Determinanten von Matrizen in Zeilenstufenform ablesen.

(5.17) Beispiel. Es sei $A \in \mathbb{Z}^{4 \times 4}$ gegeben durch

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

Dann ist

$$\det A = -4.$$

Beweis. Zunächst formen wir A mittels Vertauschungsoperationen auf den Zeilen um:

$$\begin{pmatrix} 3 & 4 & -2 & -4 \\ 0 & 0 & 1 & 4 \\ 2 & 2 & -3 & -2 \\ 0 & 0 & 1 & 2 \end{pmatrix} \xmapsto{\mathrm{sw}_{2,3}} \begin{pmatrix} 3 & 4 & -2 & -4 \\ 2 & 2 & -3 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Nach Proposition (5.4), Korollar (5.6) und dem Kästchensatz (5.15) gilt

$$\det A = \det \begin{pmatrix} 3 & 4 & -2 & -4 \\ 0 & 0 & 1 & 4 \\ 2 & 2 & -3 & -2 \\ 0 & 0 & 1 & 2 \end{pmatrix} = -\det \begin{pmatrix} 3 & 4 & -2 & -4 \\ 2 & 2 & -3 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 2 \end{pmatrix} = -(\det \begin{pmatrix} 3 & 4 \\ 2 & 2 \end{pmatrix})(\det \begin{pmatrix} 1 & 4 \\ 1 & 2 \end{pmatrix})$$
$$= -(3 \cdot 2 - 4 \cdot 2)(1 \cdot 2 - 4 \cdot 1) = -4.$$ □

Alternativer Beweis von Beispiel (5.3)(b). Zunächst wenden wir Additionsoperationen auf die Zeilen von A an:

$$\begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} \xmapsto{\mathrm{add}_{3,1,-2} \circ \mathrm{add}_{2,1,-4}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 7 & -9 \\ 0 & 1 & -3 \end{pmatrix} \xmapsto{\mathrm{add}_{2,3,-7}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix}$$

Nach Proposition (5.4), Korollar (5.6) und Korollar (5.16) gilt

$$\det A = \det \begin{pmatrix} 1 & -1 & 2 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 12 \\ 0 & 1 & -3 \end{pmatrix} = -\det \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 12 \end{pmatrix} = -(1 \cdot 1 \cdot 12) = -12. \qquad \square$$

Zusätzliche Konzepte

Die im Beweis von Satz (5.14) verwendete Matrix B ist unter folgendem Namen bekannt:

(5.18) Definition (Adjunkte)**.** Für $A \in R^{n \times n}$ heißt $\mathrm{Adj}(A) \in R^{n \times n}$ gegeben durch

$$\mathrm{Adj}(A) = ((-1)^{i+j} \mathrm{Minor}_{j,i}(A))_{i,j \in [1,n]}$$

die *Adjunkte* von A (oder die *komplementäre Matrix* zu A).

Die Laplace-Entwicklung lässt sich kompakt in folgender Formel zusammenfassen:

(5.19) Satz. Für $A \in R^{n \times n}$ gilt

$$\mathrm{Adj}(A)\, A = A\, \mathrm{Adj}(A) = (\det A) \mathrm{E}_n.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Die Adjunkte einer invertierbaren Matrix steht in engem Zusammenhang zur Inversen:

(5.20) Korollar. Für $A \in \mathrm{GL}_n(R)$ ist

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

Beweis. Dies sei dem Leser zur Übung überlassen. □

6 Eigenwerttheorie

Die Darstellungsmatrix eines Endomorphismus $\varphi\colon V \to V$ eines endlichdimensionalen Vektorraums V ist abhängig von einer gewählten Basis $s = (s_1, \dots, s_n)$ von V. Dabei sind einige Basen zum Rechnen besser geeignet als andere:
Es sei etwa $\varphi\colon \mathbb{R}[X]_{<3} \to \mathbb{R}[X]_{<3}$ gegeben durch

$$\varphi(a_0 + a_1 X + a_2 X^2) = (-a_0 - 3a_1 + 3a_2) + (2a_0 + 5a_1 - 4a_2)X + (2a_0 + 3a_1 - 2a_2)X^2$$

für $a \in \mathbb{R}^{[0,2]}$. Dann ist die Darstellungsmatrix von φ bzgl. der Basis $s = (1, X, X^2)$ von $\mathbb{R}[X]_{<3}$ gegeben durch

$$\mathrm{M}_{s,s}(\varphi) = \begin{pmatrix} -1 & -3 & 3 \\ 2 & 5 & -4 \\ 2 & 3 & -2 \end{pmatrix}.$$

Wollen wir nun die Darstellungsmatrix einer Potenz, etwa von φ^3 ermitteln, so haben wir folgende Rechnung durchzuführen: Nach Proposition (3.13)(a) ist

$$\mathrm{M}_{s,s}(\varphi^3) = (\mathrm{M}_{s,s}(\varphi))^3 = \begin{pmatrix} -1 & -3 & 3 \\ 2 & 5 & -4 \\ 2 & 3 & -2 \end{pmatrix}^3 = \begin{pmatrix} -1 & -3 & 3 \\ 2 & 5 & -4 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} -1 & -3 & 3 \\ 2 & 5 & -4 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} -1 & -3 & 3 \\ 2 & 5 & -4 \\ 2 & 3 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & -3 & 3 \\ 2 & 5 & -4 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} 1 & -3 & 3 \\ 0 & 7 & -6 \\ 0 & 3 & -2 \end{pmatrix} = \begin{pmatrix} -1 & -9 & 9 \\ 2 & 17 & -16 \\ 2 & 9 & -8 \end{pmatrix}.$$

Für die Basis $s' = (s'_1, s'_2, s'_3)$ von $\mathbb{R}[X]_{<3}$ gegeben durch

$$s' = (-1 + X + X^2, X + X^2, -1 + 2X + X^2)$$

gilt hingegen

$$\mathrm{M}_{s',s'}(\varphi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Für einige Zwecke sind die Basis s' und die Darstellungsmatrix $\mathrm{M}_{s',s'}(\varphi)$ besser geeignet als die Basis s und die Darstellungsmatrix $\mathrm{M}_{s,s}(\varphi)$. So vereinfacht sich etwa die Berechnung von Potenzen von (der Darstellungsmatrix von) φ:

$$\mathrm{M}_{s',s'}(\varphi^3) = (\mathrm{M}_{s',s'}(\varphi))^3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}^3 = \begin{pmatrix} (-1)^3 & 0 & 0 \\ 0 & 1^3 & 0 \\ 0 & 0 & 2^3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Unser Bestreben ist es, nach Basen zu suchen, bzgl. derer die Darstellungsmatrix eines Vektorraumendomorphismus eine besonders „schöne“, d.h. für das Rechnen praktische, Gestalt hat. Die „schönste“ vorstellbare Form ist dabei eine Diagonalmatrix wie in unserem gerade betrachteten einführenden Beispiel. Leider gibt es nicht immer eine Basis derart, dass wir eine solche Gestalt erreichen können, siehe Beispiel (6.51). Wir werden Kriterien für die Existenz einer solchen Basis herleiten, siehe Satz (6.52) und Korollar (6.53).
Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein Körper K gegeben.

Eigenräume

Es seien $n \in \mathbb{N}_0$, ein n-dimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Wenn es eine Basis $s = (s_1, \ldots, s_n)$ von V so gibt, dass

$$\mathrm{M}_{s,s}(\varphi) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

für gewisse $a_j \in K$ für $j \in [1, n]$ ist, so bedeutet dies gerade, dass für $j \in [1, n]$ stets

$$\varphi(s_j) = a_j s_j$$

gilt. Dies führt uns auf die Begriffe Eigenwert und Eigenvektor, welche wir in diesem Abschnitt einführen wollen.

(6.1) Definition (Eigenwert)**.** Es sei $a \in K$ gegeben.

(a) Es sei ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Wir setzen

$$\mathrm{Eig}_a(\varphi) := \mathrm{Ker}(\varphi - a\,\mathrm{id}_V).$$

Wenn $\mathrm{Eig}_a(\varphi) \neq \{0\}$ ist, so nennen wir a einen *Eigenwert* von φ sowie $\mathrm{Eig}_a(\varphi)$ den *Eigenraum* von φ zum Eigenwert a und jedes $v \in \mathrm{Eig}_a(\varphi) \setminus \{0\}$ einen *Eigenvektor* von φ zum Eigenwert a.

(b) Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Wir setzen

$$\mathrm{Eig}_a(A) := \mathrm{Eig}_a(\varphi_A).$$

Wenn $\mathrm{Eig}_a(A) \neq \{0\}$ ist, so nennen wir a einen *Eigenwert* von A sowie $\mathrm{Eig}_a(A)$ den *Eigenraum* von A zum Eigenwert a und jedes $x \in \mathrm{Eig}_a(A) \setminus \{0\}$ eine *Eigenspalte* (oder *Eigenvektor*) von A zum Eigenwert a.

Nach Bemerkung (2.35) und Proposition (2.36) sind Addition und Skalarmultiplikation auf $\mathrm{End}_K(V)$ für einen K-Vektorraum V werteweise erklärt, d.h. für einen K-Vektorraumhomomorphismus $\varphi\colon V \to V$ ist $\varphi - a\,\mathrm{id}_V\colon V \to V$ gegeben durch

$$(\varphi - a\,\mathrm{id}_V)(v) = \varphi(v) - a\,\mathrm{id}_V(v) = \varphi(v) - av$$

für $v \in V$.

(6.2) Bemerkung. Es sei $a \in K$ gegeben.

(a) Es sei ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Dann ist

$$\mathrm{Eig}_a(\varphi) = \{v \in V \mid \varphi(v) = av\}.$$

Genau dann ist a ein Eigenwert von φ, wenn es ein $v \in V \setminus \{0\}$ mit $\varphi(v) = av$ gibt, und in diesem Fall ist v ein Eigenvektor von φ zum Eigenwert a.

(b) Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Dann ist

$$\mathrm{Eig}_a(A) = \{x \in K^{n \times 1} \mid Ax = ax\}.$$

Genau dann ist a ein Eigenwert von A, wenn es ein $x \in K^{n \times 1} \setminus \{0\}$ mit $Ax = ax$ gibt, und in diesem Fall ist x ein Eigenvektor von A zum Eigenwert a.

Beweis.

(a) Es ist

$$\begin{aligned}\mathrm{Eig}_a(\varphi) &= \mathrm{Ker}(\varphi - a\,\mathrm{id}_V) = \{v \in V \mid (\varphi - a\,\mathrm{id}_V)(v) = 0\} = \{v \in V \mid \varphi(v) - a\,\mathrm{id}_V(v) = 0\} \\ &= \{v \in V \mid \varphi(v) - av = 0\} = \{v \in V \mid \varphi(v) = av\}.\end{aligned}$$

(b) Nach (a) ist

$$\mathrm{Eig}_a(A) = \mathrm{Eig}_a(\varphi_A) = \{x \in K^{n \times 1} \mid \varphi_A(x) = ax\} = \{x \in K^{n \times 1} \mid Ax = ax\}. \qquad \square$$

Wir betonen, dass per Definition der Nullvektor kein Eigenvektor eines Vektorraumendomorphismus $\varphi\colon V \to V$ ist, obwohl $\varphi(0) = 0 = a \cdot 0$ für alle $a \in K$ gilt.

(6.3) Beispiel. Es seien $A \in \mathbb{Q}^{2 \times 2}$, $x, y \in \mathbb{Q}^{2 \times 1}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},\ x = \begin{pmatrix} 1 \\ -1 \end{pmatrix},\ y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Dann ist x ein Eigenvektor von A zum Eigenwert 1 und y ein Eigenvektor von A zum Eigenwert 3.

Beweis. Wir haben

$$\begin{aligned}Ax &= \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = x, \\ Ay &= \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3y.\end{aligned} \qquad \square$$

In Beispiel (6.3) konnten wir für gewisse Skalare und Vektoren nachrechnen, dass es sich um Eigenwerte bzw. Eigenvektoren einer Matrix handelt. Es stellt sich jedoch die Frage, wie man die Eigenwerte und Eigenvektoren zu einem gegebenen Endomorphismus bzw. zu einer gegebenen Matrix bestimmen kann. Zunächst zeigen wir, dass sich die Berechnung der Eigenräume auf das Lösen eines homogenen linearen Gleichungssystems zurückführen lässt.

(6.4) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n \times n}$ und $a \in K$ gegeben. Dann ist

$$\mathrm{Eig}_a(A) = \mathrm{Sol}(A - a\mathrm{E}_n, 0).$$

Beweis. Nach Bemerkung (3.5)(b), (d) und Korollar (3.36) gilt

$$\operatorname{Eig}_a(A) = \operatorname{Eig}_a(\varphi_A) = \operatorname{Ker}(\varphi_A - a\,\mathrm{id}_{K^{n\times 1}}) = \operatorname{Ker}(\varphi_A - a\varphi_{\mathrm{E}_n}) = \operatorname{Ker}(\varphi_{A-a\mathrm{E}_n}) = \operatorname{Sol}(A - a\mathrm{E}_n, 0). \qquad \square$$

(6.5) Korollar. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$, eine Basis $s = (s_1, \ldots, s_n)$ von V und ein $a \in K$ gegeben. Dann ist

$$\kappa_s(\operatorname{Eig}_a(\varphi)) = \operatorname{Eig}_a(\mathrm{M}_{s,s}(\varphi))$$

und

$$\operatorname{Eig}_a(\varphi) = \{\sum_{j\in[1,n]} b_j s_j \mid b \in \operatorname{Eig}_a(\mathrm{M}_{s,s}(\varphi))\}.$$

Genau dann ist a ein Eigenwert von φ, wenn a ein Eigenwert von $\mathrm{M}_{s,s}(\varphi)$ ist, und in diesem Fall ist ein $v \in V$ genau dann ein Eigenvektor von φ zum Eigenwert a, wenn $\kappa_s(v)$ ein Eigenvektor von $\mathrm{M}_{s,s}(\varphi)$ zum Eigenwert a ist.

Beweis. Nach Bemerkung (3.39), Proposition (3.14), Proposition (3.13)(b) und Bemerkung (6.4) ist

$$\begin{aligned}\kappa_s(\operatorname{Eig}_a(\varphi)) &= \kappa_s(\operatorname{Ker}(\varphi - a\,\mathrm{id}_V)) = \operatorname{Sol}(\mathrm{M}_{s,s}(\varphi - a\,\mathrm{id}_V), 0) = \operatorname{Sol}(\mathrm{M}_{s,s}(\varphi) - a\,\mathrm{M}_{s,s}(\mathrm{id}_V), 0) \\ &= \operatorname{Sol}(\mathrm{M}_{s,s}(\varphi) - a\mathrm{E}_n, 0) = \operatorname{Eig}_a(\mathrm{M}_{s,s}(\varphi))\end{aligned}$$

und

$$\operatorname{Eig}_a(\varphi) = \operatorname{Ker}(\varphi - a\,\mathrm{id}_V) = \{\sum_{j\in[1,n]} b_j s_j \mid b \in \operatorname{Sol}(\mathrm{M}_{s,s}(\varphi - a\,\mathrm{id}_V), 0)\} = \{\sum_{j\in[1,n]} b_j s_j \mid b \in \operatorname{Eig}_a(\mathrm{M}_{s,s}(\varphi))\}.$$

Da κ_s ein Isomorphismus ist, ist genau dann $\operatorname{Eig}_a(\varphi) \neq \{0\}$, wenn $\operatorname{Eig}_a(\mathrm{M}_{s,s}(\varphi)) \neq \{0\}$ ist, d.h. genau dann ist a ein Eigenwert von φ, wenn a ein Eigenwert von $\mathrm{M}_{s,s}(\varphi)$ ist, und in diesem Fall ist ein $v \in V$ genau dann ein Eigenvektor von φ zum Eigenwert a, wenn $\kappa_s(v)$ ein Eigenvektor von $\mathrm{M}_{s,s}(\varphi)$ zum Eigenwert a ist. $\square$

Nach Korollar (6.5) lässt sich das Problem der Bestimmung von Eigenwerten und Eigenräumen eines Endomorphismus durch Übergang zu einer Darstellungsmatrix in ein Problem für Matrizen übersetzen.

(6.6) Proposition. Es sei ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Für alle $n \in \mathbb{N}_0$ und jedes n-Tupel $(a_1, \ldots, a_n)$ in K mit verschiedenen Einträgen ist $(\operatorname{Eig}_{a_1}(\varphi), \ldots, \operatorname{Eig}_{a_n}(\varphi))$ unabhängig.

Beweis. Wir führen Induktion nach n, wobei für $n = 0$ nichts zu zeigen ist. Es sei also $n \in \mathbb{N}$ so gegeben, dass $(\operatorname{Eig}_{a_1}(\varphi), \ldots, \operatorname{Eig}_{a_{n-1}}(\varphi))$ unabhängig. Um zu zeigen, dass $(\operatorname{Eig}_{a_1}(\varphi), \ldots, \operatorname{Eig}_{a_n}(\varphi))$ unabhängig ist, sei $v \in \operatorname{Eig}_{a_n}(\varphi) \cap \sum_{i\in[1,n-1]} \operatorname{Eig}_{a_i}(\varphi)$ gegeben. Dann gibt es $v_i \in \operatorname{Eig}_{a_i}(\varphi)$ für $i \in [1, n-1]$ mit $v = \sum_{i\in[1,n-1]} v_i$. Es folgt

$$\begin{aligned}0 &= (\varphi - a_n \mathrm{id}_V)(v) = (\varphi - a_n \mathrm{id}_V)(\sum_{i\in[1,n-1]} v_i) = \sum_{i\in[1,n-1]} (\varphi - a_n \mathrm{id}_V)(v_i) \\ &= \sum_{i\in[1,n-1]} (\varphi(v_i) - a_n \mathrm{id}_V(v_i)) = \sum_{i\in[1,n-1]} (a_i v_i - a_n v_i) = \sum_{i\in[1,n-1]} (a_i - a_n) v_i.\end{aligned}$$

Die Unabhängigkeit von $(\operatorname{Eig}_{a_1}(\varphi), \ldots, \operatorname{Eig}_{a_{n-1}}(\varphi))$ impliziert $(a_i - a_n)v_i = 0$ für $i \in [1, n-1]$. Wegen $a_i \neq a_n$ für $i \in [1, n-1]$ folgt $v_i = 0$ für $i \in [1, n-1]$ und damit $v = \sum_{i\in[1,n-1]} v_i = 0$. Nach Proposition (1.78) ist $(\operatorname{Eig}_{a_1}(\varphi), \ldots, \operatorname{Eig}_{a_n}(\varphi))$ unabhängig.
Nach dem Induktionsprinzip ist für $n \in \mathbb{N}_0$ und jedes n-Tupel $(a_1, \ldots, a_n)$ in K mit verschiedenen Einträgen das Tupel $(\operatorname{Eig}_{a_1}(\varphi), \ldots, \operatorname{Eig}_{a_n}(\varphi))$ unabhängig. $\square$

(6.7) Korollar. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$, ein Tupel $(a_1, \ldots, a_n)$ aus verschiedenen Eigenwerten von φ und ein n-Tupel $(s_1, \ldots, s_n)$ in V so gegeben, dass s_i für $i \in [1, n]$ ein Eigenvektor von φ zum Eigenwert a_i ist. Dann ist $(s_1, \ldots, s_n)$ linear unabhängig.

Beweis. Dies folgt aus Proposition (6.6) und Proposition (1.79). $\square$

Charakteristisches Polynom

Zur Berechnung der Eigenwerte eines Endomorphismus bzw. einer quadratischen Matrix führen wir den Begriff des charakteristischen Polynoms ein, wobei wir mit der Variante für Matrizen beginnen.

(6.8) Definition (charakteristisches Polynom)**.** Es seien ein $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Das *charakteristische Polynom* von A ist das Polynom χ_A mit Koeffizienten in K gegeben durch

$$\chi_A = \det(X\mathrm{E}_n - A).$$

In Definition (6.8) kommt uns zu Gute, dass wir in Abschnitt 5 Determinanten von Matrizen mit Einträgen in kommutativen Ringen studiert haben: für $A \in K^{n \times n}$, $n \in \mathbb{N}_0$ ist $X\mathrm{E}_n - A \in K[X]^{n \times n}$ und $K[X]$ ist ein kommutativer Ring, aber kein Körper.

(6.9) Beispiel.

(a) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Das charakteristische Polynom von A ist

$$\chi_A = X^2 - 4X + 3 = (X - 1)(X - 3).$$

(b) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

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

Das charakteristische Polynom von A ist

$$\chi_A = X^2.$$

Beweis.

(a) Es ist

$$\begin{aligned} \chi_A &= \det(X\mathrm{E}_2 - A) = \det(X \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}) = \det \begin{pmatrix} X - 2 & -1 \\ -1 & X - 2 \end{pmatrix} = (X - 2)^2 - (-1)^2 \\ &= X^2 - 4X + 3 = (X - 1)(X - 3). \end{aligned}$$

(b) Es ist

$$\chi_A = \det(X\mathrm{E}_2 - A) = \det(X \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) = \det \begin{pmatrix} X & -1 \\ 0 & X \end{pmatrix} = X^2. \qquad \square$$

(6.10) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Dann ist

$$\chi_{A^{\mathrm{tr}}} = \chi_A.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Wir werden den Begriff des charakteristischen Polynoms eines Endomorphismus durch Übergang zu einer Darstellungsmatrix auf den des charakteristischen Polynoms einer Matrix zurückführen. Hierzu müssen wir zeigen, dass das charakteristische Polynom einer Darstellungsmatrix eines Endomorphismus unabhängig von der für die Darstellungsmatrix gewählten Basis ist.

(6.11) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und Basen $s = (s_1, \dots, s_n)$ und $s' = (s'_1, \dots, s'_n)$ von V gegeben. Dann gilt

$$\chi_{\mathrm{M}_{s,s}(\varphi)} = \chi_{\mathrm{M}_{s',s'}(\varphi)}.$$

Beweis. Nach Bemerkung (3.63)(b) ist $\mathrm{M}_{s,s}(\varphi)$ ähnlich zu $\mathrm{M}_{s',s'}(\varphi)$, so dass es ein $P \in \mathrm{GL}_n(K)$ mit

$$P^{-1}\,\mathrm{M}_{s,s}(\varphi)\,P = \mathrm{M}_{s',s'}(\varphi)$$

gibt. Wir erhalten

$$P^{-1}(X\mathrm{E}_n - \mathrm{M}_{s,s}(\varphi))P = XP^{-1}\mathrm{E}_nP - P^{-1}\,\mathrm{M}_{s,s}(\varphi)\,P = X\mathrm{E}_n - \mathrm{M}_{s',s'}(\varphi)$$

und nach Proposition (5.8)(a), (c) folgt

$$\begin{aligned}\chi_{\mathrm{M}_{s',s'}(\varphi)} &= \det(X\mathrm{E}_n - \mathrm{M}_{s',s'}(\varphi)) = \det(P^{-1}(X\mathrm{E}_n - \mathrm{M}_{s,s}(\varphi))P) \\ &= (\det P)^{-1}\det(X\mathrm{E}_n - \mathrm{M}_{s,s}(\varphi))\,(\det P) = \det(X\mathrm{E}_n - \mathrm{M}_{s,s}(\varphi)) = \chi_{\mathrm{M}_{s,s}(\varphi)}.\end{aligned}$$

□

(6.12) Definition (charakteristisches Polynom)**.** Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Das *charakteristische Polynom* von φ ist das Polynom χ_φ mit Koeffizienten in K gegeben durch

$$\chi_\varphi = \chi_{\mathrm{M}_{s,s}(\varphi)}$$

für eine beliebige Basis s von V.

Als nächstes werden wir sehen, dass sich die Eigenwerte eines Endomorphismus bzw. einer quadratischen Matrix mit Hilfe des charakteristischen Polynoms berechnen lassen.

(6.13) Proposition. Es seien $n \in \mathbb{N}_0$, $A \in K^{n\times n}$ und $a \in K$ gegeben. Die folgenden Bedingungen sind äquivalent:

(a) Es ist a ein Eigenwert von A.

(b) Es ist $\mathrm{rk}_K(A - a\,\mathrm{E}_n) < n$.

(c) Es ist $\det(A - a\,\mathrm{E}_n) = 0$.

(d) Es ist a eine Nullstelle des charakteristischen Polynoms von A.

Beweis. Genau dann ist a ein Eigenwert von A, wenn $\mathrm{Eig}_a(A) \neq \{0\}$ ist. Nach Bemerkung (6.4) ist $\mathrm{Eig}_a(A) = \mathrm{Sol}(A - a\mathrm{E}_n, 0)$ und damit

$$\dim \mathrm{Eig}_a(A) = \dim \mathrm{Sol}(A - a\mathrm{E}_n, 0) = n - \mathrm{rk}(A - a\mathrm{E}_n)$$

nach Korollar (3.36). Folglich ist a genau dann ein Eigenwert von A, wenn $\mathrm{rk}(A - a\mathrm{E}_n) < n$ ist, d.h. Bedingung (a) und Bedingung (b) sind äquivalent.
Nach Bemerkung (3.31) und Satz (5.14) ist genau dann $\mathrm{rk}_K(A - a\,\mathrm{E}_n) < n$, wenn $\det(A - a\,\mathrm{E}_n) = 0$ ist, d.h. auch Bedingung (b) und Bedingung (c) sind äquivalent.
Schließlich gilt

$$\chi_A = \sum_{\pi\in\mathrm{S}_n}(\operatorname{sgn}\pi)\prod_{j\in[1,n]}(X\mathrm{E}_n - A)_{\pi(j),j} = \sum_{\pi\in\mathrm{S}_n}(\operatorname{sgn}\pi)\prod_{j\in[1,n]}(\delta_{\pi(j),j}X - A_{\pi(j),j})$$

und damit

$$\begin{aligned}\chi_A(a) &= \sum_{\pi\in\mathrm{S}_n}(\operatorname{sgn}\pi)\prod_{j\in[1,n]}(\delta_{\pi(j),j}a - A_{\pi(j),j}) = \sum_{\pi\in\mathrm{S}_n}(\operatorname{sgn}\pi)\prod_{j\in[1,n]}(a\mathrm{E}_n - A)_{\pi(j),j} = \det(a\mathrm{E}_n - A) \\ &= (-1)^n\det(A - a\mathrm{E}_n),\end{aligned}$$

so dass genau dann $\det(A - a\mathrm{E}_n) = 0$ ist, wenn a eine Nullstelle von χ_A ist. Folglich sind auch Bedingung (c) und Bedingung (d) äquivalent.
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c) und Bedingung (d) äquivalent. □

(6.14) Korollar. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$ gegeben. Die Eigenwerte von A^{tr} sind genau die Eigenwerte von A.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Zur Berechung von Eigenwerten und Eigenräumen/Eigenvektoren haben wir nun also folgende Strategie: Zuerst berechnen wir das charakteristische Polynom als Determinante einer Matrix mit polynomiellen Einträgen. Danach bestimmen wir die Nullstellen des charakteristischen Polynoms, diese bilden genau die Eigenwerte. Der Eigenraum zu einem Eigenwert ergibt sich dann schließlich als Lösungsraum eines homogenen linearen Gleichungssystems.

(6.15) Beispiel.

(a) Es sei $A \in \mathbb{Q}^{2\times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Dann sind die Eigenwerte von A gerade 1 und 3 und es gilt

$$\operatorname{Eig}_1(A) = \mathbb{Q}\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$
$$\operatorname{Eig}_3(A) = \mathbb{Q}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(b) Es sei $A \in \mathbb{Q}^{2\times 2}$ gegeben durch

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

Dann ist 0 der einzige Eigenwert von A und es gilt

$$\operatorname{Eig}_0(A) = \mathbb{Q}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Beweis.

(a) Nach Beispiel (6.9)(a) ist das charakteristische Polynom von A gegeben durch

$$\chi_A = X^2 - 4X + 3 = (X-1)(X-3),$$

so dass nach Proposition (6.13) die Eigenwerte von A daher durch 1 und 3 gegeben sind. Mit Bemerkung (6.4) erhalten wir

$$\operatorname{Eig}_1(A) = \operatorname{Sol}(A - \mathrm{E}_2, 0) = \operatorname{Sol}(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, 0) = \operatorname{Sol}(\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, 0) = \mathbb{Q}\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$
$$\operatorname{Eig}_3(A) = \operatorname{Sol}(A - 3\,\mathrm{E}_2, 0) = \operatorname{Sol}(\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, 0) = \operatorname{Sol}(\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, 0) = \mathbb{Q}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(b) Nach Beispiel (6.9)(b) ist das charakteristische Polynom von A gegeben durch

$$\chi_A = X^2,$$

so dass 0 nach Proposition (6.13) der einzige Eigenwert von A ist. Mit Bemerkung (6.4) erhalten wir

$$\operatorname{Eig}_0(A) = \operatorname{Sol}(A - 0\mathrm{E}_2, 0) = \operatorname{Sol}(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, 0) = \mathbb{Q}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$ □

Geometrische und algebraische Vielfachheit

Als nächstes wollen wir für einen Endomorphismus $\varphi\colon V \to V$ ausdrücken, wie stark dieser durch einen gegebenen Eigenwert bestimmt ist. Gilt für ein $a \in K$ etwa $\mathrm{Eig}_a(\varphi) = V$, so ist $\varphi(v) = av$ für alle $v \in V$ und damit $\varphi = a\,\mathrm{id}_V$. Gilt hingegen $\mathrm{Eig}_a(\varphi) = \{0\}$, so ist a kein Eigenwert von φ, es gibt also kein $v \in V \setminus \{0\}$ mit $\varphi(v) = av$. Im ersten Fall bestimmt a den Endomorphismus φ vollständig, im zweiten Fall gar nicht. Ist a ein Eigenwert von φ und $\mathrm{Eig}_a(\varphi)$ ein echter Untervektorraum von V, so liegt die Wichtigkeit von a für den Endomorphismus φ irgendwo zwischen diesen beiden Extremfällen. Wir wollen nun versuchen, diese Wichtigkeit zu messen, und führen hierzu gleich zwei Vielfachheiten für Eigenwerte ein.

Es seien $f \in K[X]\setminus\{0\}$ und $a \in K$ gegeben. Die Vielfachheit von a als Nullstelle von f ist das eindeutige $m \in \mathbb{N}_0$, für welches es ein $g \in K[X]$ mit $f = (X - a)^m g$ und $g(a) \neq 0$ gibt, siehe Definition (A.129). Wir notieren die Vielfachheit von a als Nullstelle von f als $\mathrm{m}_a(f) := m$.

(6.16) Definition (geometrische Vielfachheit, algebraische Vielfachheit)**.** Es sei $a \in K$ gegeben.

(a) Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben.

   (i) Die *geometrische Vielfachheit* (oder *geometrische Multiplizität*) von a als Eigenwert von φ ist definiert als

   $$\mathrm{g}_a(\varphi) := \dim_K \mathrm{Eig}_a(\varphi).$$

   (ii) Die *algebraische Vielfachheit* (oder *algebraische Multiplizität*) von a als Eigenwert von φ ist definiert als

   $$\mathrm{m}_a(\varphi) := \mathrm{m}_a(\chi_\varphi).$$

(b) Es seien $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$ gegeben.

   (i) Die *geometrische Vielfachheit* (oder *geometrische Multiplizität*) von a als Eigenwert von A ist definiert als

   $$\mathrm{g}_a(A) := \mathrm{g}_a(\varphi_A).$$

   (ii) Die *algebraische Vielfachheit* (oder *algebraische Multiplizität*) von a als Eigenwert von A ist definiert als

   $$\mathrm{m}_a(A) := \mathrm{m}_a(\varphi_A).$$

(6.17) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n\times n}$ und $a \in K$ gegeben.

(a) Es ist

$$\mathrm{g}_a(A) = \dim_K \mathrm{Eig}_a(A).$$

(b) Es ist

$$\mathrm{m}_a(A) = \mathrm{m}_a(\chi_A).$$

Beweis.

(a) Es ist

$$\mathrm{g}_a(A) = \mathrm{g}_a(\varphi_A) = \dim \mathrm{Eig}_a(\varphi_A) = \dim \mathrm{Eig}_a(A).$$

(b) Es ist

$$\mathrm{m}_a(A) = \mathrm{m}_a(\varphi_A) = \mathrm{m}_a(\chi_{\varphi_A}) = \mathrm{m}_a(\chi_{\mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)}) = \mathrm{m}_a(\chi_A).$$ □

(6.18) Beispiel.

(a) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch
$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$
Die geometrischen Vielfachheiten von 1 und 3 als Eigenwerte von A sind gegeben durch
$$\mathrm{g}_1(A) = 1,$$
$$\mathrm{g}_3(A) = 1.$$
Die algebraischen Vielfachheiten von 1 und 3 als Eigenwerte von A sind gegeben durch
$$\mathrm{m}_1(A) = 1,$$
$$\mathrm{m}_3(A) = 1.$$

(b) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
Die geometrische Vielfachheit von 0 als Eigenwert von A ist gegeben durch
$$\mathrm{g}_0(A) = 1.$$
Die algebraische Vielfachheit von 0 als Eigenwert von A ist gegeben durch
$$\mathrm{m}_0(A) = 2.$$

Beweis.

(a) Nach Beispiel (6.15)(a) ist
$$\mathrm{Eig}_1(A) = \mathbb{Q} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$
$$\mathrm{Eig}_3(A) = \mathbb{Q} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
und damit
$$\mathrm{g}_1(A) = \dim \mathrm{Eig}_1(A) = 1,$$
$$\mathrm{g}_3(A) = \dim \mathrm{Eig}_3(A) = 1$$
nach Bemerkung (6.17)(a). Nach Beispiel (6.9)(a) ist ferner $\chi_A = (X-1)(X-3)$ und damit
$$\mathrm{m}_1(A) = 1,$$
$$\mathrm{m}_3(A) = 1$$
nach Bemerkung (6.17)(b).

(b) Nach Beispiel (6.15)(b) ist
$$\mathrm{Eig}_0(A) = \mathbb{Q} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
und damit
$$\mathrm{g}_0(A) = 1$$
nach Bemerkung (6.17)(a). Nach Beispiel (6.9)(b) ist ferner $\chi_A = X^2$ und damit
$$\mathrm{m}_0(A) = 2$$
nach Bemerkung (6.17)(b). $\square$

(6.19) Proposition. Es seien ein endlichdimensionaler K-Vektorraum V, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ und ein $a \in K$ gegeben. Dann gilt

$$\mathrm{g}_a(\varphi) \leq \mathrm{m}_a(\varphi).$$

Beweis. Es seien $n := \dim V$ und $m := \mathrm{g}_a(\varphi) = \dim \mathrm{Eig}_a(\varphi)$. Da $\mathrm{Eig}_a(\varphi)$ ein K-Untervektorraum von V ist, gibt es nach Korollar (1.52) und Korollar (1.55) eine Basis $s = (s_1, \dots, s_n)$ derart, dass $(s_1, \dots, s_m)$ eine Basis von $\mathrm{Eig}_a(\varphi)$ ist. Folglich ist

$$\begin{aligned}\mathrm{M}_{s,s}(\varphi) &= \begin{pmatrix}\kappa_s(\varphi(s_1)) & \dots & \kappa_s(\varphi(s_m)) & \kappa_s(\varphi(s_{m+1})) & \dots & \kappa_s(\varphi(s_n))\end{pmatrix} \\ &= \begin{pmatrix}\kappa_s(as_1) & \dots & \kappa_s(as_m) & \kappa_s(\varphi(s_{m+1})) & \dots & \kappa_s(\varphi(s_n))\end{pmatrix} \\ &= \begin{pmatrix}a\mathrm{e}_1 & \dots & a\mathrm{e}_m & \kappa_s(\varphi(s_{m+1})) & \dots & \kappa_s(\varphi(s_n))\end{pmatrix} = \begin{pmatrix}a\mathrm{E}_m & B \\ 0 & C\end{pmatrix}\end{aligned}$$

für gewisse $B \in K^{m \times (n-m)}$, $C \in K^{(n-m)\times(n-m)}$. Nach dem Kästchensatz für Determinanten (5.15) erhalten wir

$$\begin{aligned}\chi_\varphi = \chi_{\mathrm{M}_{s,s}(\varphi)} &= \det(X\mathrm{E}_n - \mathrm{M}_{s,s}(\varphi)) = \det(X\begin{pmatrix}\mathrm{E}_m & 0 \\ 0 & \mathrm{E}_{n-m}\end{pmatrix} - \begin{pmatrix}a\mathrm{E}_m & B \\ 0 & C\end{pmatrix}) \\ &= \det\begin{pmatrix}(X-a)\mathrm{E}_m & -B \\ 0 & X\mathrm{E}_{n-m} - C\end{pmatrix} = \det((X-a)\mathrm{E}_m)\det(X\mathrm{E}_{n-m} - C) = (X-a)^m\,\chi_C.\end{aligned}$$

Folglich ist

$$\mathrm{m}_a(\varphi) = \mathrm{m}_a(\chi_\varphi) = m + \mathrm{m}_a(\chi_C) \geq m = \mathrm{g}_a(\varphi).$$ □

Begleitmatrix

Das charakteristische Polynom einer quadratischen Matrix ist ein normiertes Polynom. Als nächstes wollen wir sehen, dass jedes solche Polynom als charakteristisches Polynom vorkommt.

(6.20) Definition (Begleitmatrix)**.** Es seien $f \in K[X]$, $n \in \mathbb{N}_0$ und $a \in K^{[0,n-1]}$ mit

$$f = X^n - \sum_{i \in [0,n-1]} a_i X^i$$

gegeben. Dann heißt $\mathrm{C}(f) \in K^{n \times n}$ gegeben durch

$$\mathrm{C}(f) = \begin{pmatrix}0 & & 0 & a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & a_{n-2} \\ 0 & & 1 & a_{n-1}\end{pmatrix}$$

die *Begleitmatrix* zu f.

(6.21) Beispiel. Über $\mathbb{Q}$ ist

$$\mathrm{C}(X^3 + 2X^2 - 3X + 1) = \begin{pmatrix}0 & 0 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & -2\end{pmatrix}.$$

(6.22) Proposition. Es sei ein normiertes $f \in K[X] \setminus \{0\}$ gegeben. Das charakteristische Polynom von $\mathrm{C}(f)$ ist gegeben durch

$$\chi_{\mathrm{C}(f)} = f.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.23) Korollar. Es sei ein normiertes $f \in K[X] \setminus \{0\}$ gegeben. Die Eigenwerte von $\mathrm{C}(f)$ sind genau die Nullstellen von f.

Beweis. Nach Proposition (6.13) sind die Eigenwerte von $\mathrm{C}(f)$ genau die Nullstellen von $\chi_{\mathrm{C}(f)}$ und nach Proposition (6.22) ist $\chi_{\mathrm{C}(f)} = f$. Folglich sind die Eigenwerte von $\mathrm{C}(f)$ genau die Nullstellen von f. □

Eigenwerte stochastischer Matrizen

Um eine Anwendung der Eigenwerttheorie, den PageRank, definieren zu können, müssen wir Eigenwerte sogenannter stochastischer Matrizen studieren. Bevor wir stochastische Matrizen einführen können, benötigen wir zunächst folgenden Begriff:

(6.24) Definition (komponentenweise Ordnung)**.** Es sei eine Menge I gegeben.

(a) Ein $x \in \mathbb{R}^I$ heißt *komponentenweise nicht-negativ*, falls x_i für alle $i \in I$ nicht-negativ ist.

(b) Ein $x \in \mathbb{R}^I$ heißt *komponentenweise nicht-positiv*, falls x_i für alle $i \in I$ nicht-positiv ist.

(c) Ein $x \in \mathbb{R}^I$ heißt *komponentenweise positiv*, falls x_i für alle $i \in I$ positiv ist.

(d) Ein $x \in \mathbb{R}^I$ heißt *komponentenweise negativ*, falls x_i für alle $i \in I$ negativ ist.

(6.25) Beispiel.

(a) Es sei $x \in \mathbb{R}^{3 \times 1}$ gegeben durch

$$x = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Dann ist x komponentenweise nicht-negativ.

(b) Es sei $A \in \mathbb{R}^{2 \times 2}$ gegeben durch

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

Dann ist A komponentenweise positiv.

Beweis.

(a) Wegen $x_1 = 3 \geq 0$, $x_2 = 0 \geq 0$, $x_3 = 1 \geq 0$ ist x komponentenweise nicht-negativ.

(b) Wegen $A_{1,1} = 2 > 0$, $A_{1,2} = 2 > 0$, $A_{2,1} = 1 > 0$, $A_{2,2} = 3 > 0$ ist A komponentenweise positiv. □

(6.26) Proposition. Es seien eine Menge I und ein $\mathbb{R}$-Untervektorraum U von $\mathbb{R}^I$ gegeben. Wenn jeder Vektor in U entweder komponentenweise nicht-negativ oder komponentenweise nicht-positiv ist, dann ist U entweder 0-dimensional oder 1-dimensional.

Beweis. Es sei U nicht 0-dimensional und nicht 1-dimensional. Wir wollen zeigen, dass es dann einen Vektor in U gibt, der weder komponentenweise nicht-negativ noch komponentenweise nicht-positiv ist. Da U weder 0-dimensional noch 1-dimensional ist, gibt es ein linear unabhängiges Paar (s, t) in U. Insbesondere gibt es $i, j \in I$ mit $i \neq j$ so, dass $((s_i, s_j), (t_i, t_j))$ linear unabhängig in $\mathbb{R}^2$ ist. Da mit (s, t) auch $(-s, t)$, $(s, -t)$ und $(-s, -t)$ linear unabhängig sind, können wir o.B.d.A. annehmen, dass s und t beide komponentenweise nicht-negativ oder beide komponentenweise nicht-positiv sind. Wegen der linearen Unabhängigkeit von $((s_i, s_j), (t_i, t_j))$ in $\mathbb{R}^2$ ist ferner $(s_i, s_j) \neq 0$ und $(t_i, t_j) \neq 0$, so dass $s_i + s_j \neq 0$ und $t_i + t_j \neq 0$ folgt.
Für $x \in \langle s, t \rangle \subseteq U$ gegeben durch

$$x = (t_i + t_j)s + (-s_i - s_j)t$$

gilt nun

$$\begin{aligned} x_i + x_j &= (t_i + t_j)s_i + (-s_i - s_j)t_i + (t_i + t_j)s_j + (-s_i - s_j)t_j = (t_i + t_j)(s_i + s_j) - (s_i + s_j)(t_i + t_j) \\ &= 0. \end{aligned}$$

Da wegen $(t_i + t_j, -s_i - s_j) \neq (0, 0)$ und der linearen Unabhängigkeit von $((s_i, s_j), (t_i, t_j))$ aber $(x_i, x_j) \neq 0$ gilt, haben x_i und x_j somit verschiedene Vorzeichen. Folglich ist x weder komponentenweise nicht-negativ noch komponentenweise nicht-positiv.
Im Umkehrschluss folgt: Wenn jeder Vektor in U entweder komponentenweise nicht-negativ oder komponentenweise nicht-positiv ist, dann ist U entweder 0-dimensional oder 1-dimensional. □

(6.27) Definition (stochastische Matrix)**.** Es seien $m, n \in \mathbb{N}_0$ gegeben. Ein $A \in \mathbb{R}^{m \times n}$ heißt *stochastisch* (genauer *spaltenstochastisch*, oder *Übergangsmatrix* oder *Markovmatrix*), falls A komponentenweise nicht-negativ ist und für $j \in [1, n]$ stets

$$\sum_{i \in [1,m]} A_{i,j} = 1$$

gilt.

(6.28) Beispiel.

(a) Es sei $A \in \mathbb{R}^{4 \times 4}$ gegeben durch

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

Dann ist A stochastisch.

(b) Es sei $x \in \mathbb{R}^{4 \times 1}$ gegeben durch

$$x = \begin{pmatrix} \frac{12}{31} \\ \frac{4}{31} \\ \frac{9}{31} \\ \frac{6}{31} \end{pmatrix}.$$

Dann ist x stochastisch.

Beweis.

(a) Für $i, j \in [1, 4]$ ist $A_{i,j} \geq 0$, d.h. A ist komponentenweise nicht-negativ. Wegen

$$\sum_{i \in [1,4]} A_{i,1} = 0 + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1,$$
$$\sum_{i \in [1,4]} A_{i,2} = 0 + 0 + \frac{1}{2} + \frac{1}{2} = 1,$$
$$\sum_{i \in [1,4]} A_{i,3} = 1 + 0 + 0 + 0 = 1,$$
$$\sum_{i \in [1,4]} A_{i,4} = \frac{1}{2} + 0 + \frac{1}{2} + 0 = 1$$

ist A stochastisch.

(b) Für $i \in [1, 4]$ ist $x_i \geq 0$, d.h. x ist komponentenweise nicht-negativ. Wegen

$$\sum_{i \in [1,4]} x_i = \frac{12}{31} + \frac{4}{31} + \frac{9}{31} + \frac{6}{31} = 1$$

ist x stochastisch. $\square$

Wir schreiben $[0, 1]_{\mathbb{R}} = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ und $[0, 1)_{\mathbb{R}} = \{x \in \mathbb{R} \mid 0 \leq x < 1\}$.

(6.29) Bemerkung. Es seien $m, n \in \mathbb{N}_0$, $c \in [0, 1]_{\mathbb{R}}$, $A, B \in \mathbb{R}^{m \times n}$ gegeben. Wenn A und B stochastisch sind, dann ist auch

$$cA + (1 - c)B$$

stochastisch.

Beweis. Es seien A und B stochastisch, d.h. es seien A und B komponentenweise nicht-negativ und es gelte für $j \in [1,n]$ stets $\sum_{i\in[1,m]} A_{i,j} = 1$ und $\sum_{i\in[1,m]} B_{i,j} = 1$. Für $i \in [1,m]$, $j \in [1,n]$ ist dann auch

$$(cA + (1-c)B)_{i,j} = cA_{i,j} + (1-c)B_{i,j} \geq 0,$$

d.h. $cA + (1-c)B$ ist komponentenweise nicht-negativ, und für $j \in [1,n]$ gilt

$$\begin{aligned}\sum_{i\in[1,m]} (cA + (1-c)B)_{i,j} &= \sum_{i\in[1,m]} (cA_{i,j} + (1-c)B_{i,j}) = c \sum_{i\in[1,m]} A_{i,j} + (1-c) \sum_{i\in[1,m]} B_{i,j} \\ &= c \cdot 1 + (1-c) \cdot 1 = 1.\end{aligned}$$

Insgesamt ist $cA + (1-c)B$ stochastisch. □

(6.30) Proposition. Es seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$ gegeben. Wenn A stochastisch ist, dann ist 1 ein Eigenwert von A. ([5])

Beweis. Es sei A stochastisch. Für $s \in \mathbb{R}^{n\times 1}$ gegeben durch

$$s = \sum_{i\in[1,n]} \mathrm{e}_i = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

gilt dann

$$\begin{aligned} s^{\mathrm{tr}} A &= (\sum_{i\in[1,n]} \mathrm{e}_i)^{\mathrm{tr}} A = (\sum_{i\in[1,n]} \mathrm{e}_i^{\mathrm{tr}}) A = \sum_{i\in[1,n]} \mathrm{e}_i^{\mathrm{tr}} A = \sum_{i\in[1,n]} A_{i,-} = \sum_{i\in[1,n]} \begin{pmatrix} A_{i,1} & \dots & A_{i,n} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i\in[1,n]} A_{i,1} & \dots & \sum_{i\in[1,n]} A_{i,n} \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix} = s^{\mathrm{tr}} \end{aligned}$$

und damit $A^{\mathrm{tr}} s = (s^{\mathrm{tr}} A)^{\mathrm{tr}} = s$. Folglich ist 1 ein Eigenwert von A^{tr}. Da die Eigenwerte von A^{tr} jedoch genau die Eigenwerte von A sind, ist 1 somit auch ein Eigenwert von A. □

(6.31) Satz (Satz von Perron/Frobenius für komponentenweise positive stochastische Matrizen)**.** Es seien $n \in \mathbb{N}$ und ein komponentenweise positives, stochastisches $A \in \mathbb{R}^{n\times n}$ gegeben. Dann gibt es genau einen stochastischen Eigenvektor von A zum Eigenwert 1, und dieser ist komponentenweise positiv.

Beweis. Nach Proposition (6.30) ist 1 ein Eigenwert von A. Zunächst sei ein beliebiger Eigenvektor x von A zum Eigenwert 1 gegeben. Wegen $x = Ax$ und der komponentenweisen Nicht-Negativität von A folgt

$$|x_i| = |\sum_{j\in[1,n]} A_{i,j} x_j| \leq \sum_{j\in[1,n]} |A_{i,j} x_j| = \sum_{j\in[1,n]} |A_{i,j}|\,|x_j| = \sum_{j\in[1,n]} A_{i,j} |x_j|$$

für $i \in [1,n]$. Da A stochastisch ist, gilt andererseits aber auch

$$\sum_{i\in[1,n]} \sum_{j\in[1,n]} A_{i,j}|x_j| = \sum_{j\in[1,n]} \sum_{i\in[1,n]} A_{i,j}|x_j| = \sum_{j\in[1,n]} (\sum_{i\in[1,n]} A_{i,j})|x_j| = \sum_{j\in[1,n]} |x_j| = \sum_{i\in[1,n]} |x_i|.$$

Folglich gilt sogar $|x_i| = \sum_{j\in[1,n]} A_{i,j}|x_j|$ für $i \in [1,n]$. Dann ist x aber entweder komponentenweise nicht-negativ oder komponentenweise nicht-positiv. Wegen der komponentenweisen Positivität von A und $x \neq 0$ folgt

$$x_i = \sum_{j\in[1,n]} A_{i,j} x_j \neq 0$$

für $i \in [1,n]$, d.h. x ist sogar entweder komponentenweise positiv oder komponentenweise negativ. Somit ist $\frac{1}{\sum_{j\in[1,n]} x_j} x$ ein komponentenweise positiver, stochastischer Eigenvektor von A zum Eigenwert 1. Da x aber als beliebiger Eigenvektor von A zum Eigenwert 1 gegeben war, gilt

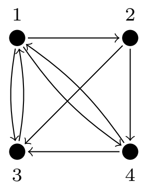
$$\dim \mathrm{Eig}_1(A) = 1$$

nach Proposition (6.26), so dass $\frac{1}{\sum_{j\in[1,n]} x_j} x$ der einzige stochastische Eigenvektor zum Eigenwert 1 ist. □

[5] Für $n = 0$ stimmt die Aussage nicht, da in diesem Fall $\mathbb{R}^{n\times 1} = \{0\}$ ist.

Anwendung: PageRank

Um bei einer Suchanfrage die relevanten Suchergebnisse in der Ergebnisliste möglichst weit oben erscheinen zu lassen, ordnet Google die Webseiten des Internets nach ihrer Wichtigkeit. Die Grundlage dieser Sortierung bildet der sogenannte *PageRank*, welchen wir im Folgenden als Anwendung der Eigenwerttheorie studieren wollen.
Zur Modellierung des Internets benutzen wir einen endlichen schlichten gerichteten Graphen G. Hierbei entspricht jede Ecke von G einer Webseite (im Folgenden kurz Seite) des Internets. Wir wollen der Einfachheit halber davon ausgehen, dass die Eckenmenge von G durch $\mathrm{V}(G) = [1, n]$ für ein $n \in \mathbb{N}_0$ gegeben ist. Für Ecken j und i in G entspricht eine Kante von j nach i einem Link der Seite j auf die Seite i. Zur Illustration betrachten wir folgendes Beispiel eines schlichten gerichteten Graphen.



Für eine Ecke j in G bezeichnet $\Gamma_+(j) = \{i \in \mathrm{V}(G) \mid \text{es gibt eine Kante von } j \text{ nach } i \text{ in } G\}$ die *Nachfolgerschaft* von j in G und $\Gamma_-(j) = \{i \in \mathrm{V}(G) \mid \text{es gibt eine Kante von } i \text{ nach } j \text{ in } G\}$ die *Vorgängerschaft* von j in G. Ferner bezeichnet $\deg_+(j) = |\Gamma_+(j)|$ den *Außengrad* von j in G und $\deg_-(j) = |\Gamma_-(j)|$ den *Innengrad* von j in G. Im Beispiel ist etwa $\Gamma_+(1) = \{2, 3, 4\}$ und $\Gamma_-(1) = \{3, 4\}$, und damit $\deg_+(1) = 3$ und $\deg_-(1) = 2$.
Im Folgenden wollen wir die Wichtigkeit einer Seite i des Internets, also einer Ecke i im gerichteten Graphen G, durch eine nicht-negative reelle Zahl x_i modellieren, die umso größer ist, je wichtiger die Seite i ist.
Es stellt sich nun die Frage, was eine gute Definition für die Wichtigkeit x_i einer Seite i ist. Unsere grundlegende Annahme wird dabei sein, dass das Internet an Hand der Links selbst die Wichtigkeit seiner Seiten bestimmt.
Ein erster naiver Ansatz könnte wie folgt lauten: Die Wichtigkeit einer Seite i entspricht gerade der Anzahl der Links auf diese Seite. Oder mit anderen Worten: Jeder Link, welcher auf die Seite i verweist, erhöht die Wichtigkeit der Seite um den Wert 1. Stellen wir uns eine Verlinkung von einer Seite j auf die Seite i als eine Art „Wahlvorgang“ vor, so würde dies bedeuten, dass die Seite j der Seite i gerade eine Stimme verleiht.
In unserem mathematischen Modell für das Internet würde das also bedeuten: Für jede Ecke i in G gilt

$$x_i = \deg_-(i) = |\Gamma_-(i)| = \sum_{j \in \Gamma_-(i)} 1.$$

Unser Beispiel bekäme demnach die Wichtigkeiten

$$x_1 = 2,\ x_2 = 1,\ x_3 = 3,\ x_4 = 2$$

zugeordnet, so dass sich das folgende Ranking ergeben würde:

$$x_3 > x_1 = x_4 > x_2$$

Problematisch an diesem Ansatz ist die Tatsache, dass unwichtige Seiten die gleiche Stimmgewalt wie wichtige Seiten haben. Von einem sinnvollen Ranking erwarten wir aber, dass ein Link von einer wichtigen Seite j die Wichtigkeit einer Seite i stärker erhöht als ein Link von einer unwichtigen Seite j'.
Im Beispiel scheint etwa Seite 3 wichtig zu sein, da jede andere Seite auf diese Seite verlinkt. Die Seiten 1 und 4 erhalten beide jeweils zwei Links von anderen Seiten, aber nur die Seite 1 erhält einen Link von der wichtigen Seite 3. Daher erwarten wir, dass die Wichtigkeit von Seite 1 größer als die Wichtigkeit von Seite 4 ist, entgegen unserem bisherigen Ansatz.
Dieses Problem können wir umgehen, indem wir die definierende Formel für die Wichtigkeit x_i der Seite i so modifizieren, dass die Stimmen für die Seite i durch die Wichtigkeit der für die Seite i stimmenden Seiten gewichtet werden. Wir erhalten den Ansatz

$$x_i = \sum_{j \in \Gamma_-(i)} x_j$$

für jede Ecke i in G.

In unserem Beispiel ergäbe sich nun

$$\begin{aligned}
x_1 &= x_3 + x_4,\\
x_2 &= x_1,\\
x_3 &= x_1 + x_2 + x_4,\\
x_4 &= x_1 + x_2.
\end{aligned}$$

Durch Umformen erhielten wir ein homogenes lineares Gleichungssystem mit der eindeutigen Lösung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Dass wir nun erst recht kein vernünftiges Ranking mehr erhalten, liegt unter anderem an einem anderen Problem, welches wir bisher noch nicht beachtet haben: In unserem jetzigen Modell kann eine Seite einen großen Einfluss auf das Ranking ausüben, indem es viele Seiten verlinkt. Somit kann eine Seite, unabhängig von seiner eigenen Wichtigkeit, eine nur durch die Gesamtzahl aller Seiten begrenzte, aber ansonsten beliebig hohe Stimmgewalt haben.
Dies können wir vermeiden, indem wir jeder Seite nur genau eine Stimme zuteilen. Verlinkt eine Seite auf mehrere andere Seiten, so wird diese Stimme unter allen verlinkten Seiten gleichmäßig verteilt – natürlich weiterhin gewichtet mit der Wichtigkeit der verlinkenden Seite.
Unser mathematisches Modell für die Wichtigkeit einer Seite i lautet nun also wie folgt: Für jede Ecke i in G gilt

$$x_i = \sum_{j \in \Gamma_-(i)} \frac{1}{\deg_+(j)} x_j.$$

Für unser Beispiel ergibt sich hiermit

$$\begin{aligned}
x_1 &= x_3 + \frac{1}{2}x_4,\\
x_2 &= \frac{1}{3}x_1,\\
x_3 &= \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_4,\\
x_4 &= \frac{1}{3}x_1 + \frac{1}{2}x_2.
\end{aligned}$$

Wir erhalten die bis auf skalare Vielfache eindeutige Lösung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \\ 9 \\ 6 \end{pmatrix}.$$

Als Ranking ergibt sich:

$$x_1 > x_3 > x_4 > x_2$$

Nimmt man eine andere Lösung mit komponentenweise positiven Einträgen, und nur diese lassen sich für unsere Anwendung sinnvoll interpretieren, so verändert dies das Ranking nicht.
Wir erkennen, dass die Wichtigkeit der Seite 3 die Wichtigkeit der Seite 1 nun so stark beeinflusst, dass diese Seite sogar an die erste Stelle des Rankings rutscht. Dies liegt daran, dass Seite 3 einzig und allein auf Seite 1 verlinkt und daher ihre gesamte Stimmgewalt dieser Seite zukommen lässt, während alle andere Seiten ihre Stimme unter mehreren verlinkten Seiten aufteilen.
Um die Wichtigkeiten mit Mitteln der linearen Algebra berechnen zu können, formulieren wir die definierende Formel zunächst um: Für jede Ecke i in G gilt

$$x_i = \sum_{j \in \Gamma_-(i)} \frac{1}{\deg_+(j)} x_j = \sum_{\substack{j \in [1,n] \\ i \in \Gamma_+(j)}} \frac{1}{\deg_+(j)} x_j = \sum_{\substack{j \in [1,n] \\ i \in \Gamma_+(j)}} \frac{1}{\deg_+(j)} x_j + \sum_{\substack{j \in [1,n] \\ i \notin \Gamma_+(j)}} 0 x_j.$$

Setzen wir für alle Ecken i und j in G nun

$$L_{i,j} := \begin{cases} \frac{1}{\deg_+(j)}, & \text{falls } i \in \Gamma_+(j), \\ 0, & \text{falls } i \notin \Gamma_+(j), \end{cases}$$

so erhalten wir also

$$x_i = \sum_{j \in [1,n]} L_{i,j} x_j$$

für jede Ecke i in G. Diese Gleichungen lassen sich mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, L = \begin{pmatrix} L_{1,1} & \dots & L_{1,n} \\ \vdots & & \vdots \\ L_{n,1} & \dots & L_{n,n} \end{pmatrix}$$

zusammenfassen als

$$x = Lx,$$

so dass die Bestimmung von x der Bestimmung eines komponentenweise nicht-negativen Eigenvektors von L zum Eigenwert 1 entspricht.

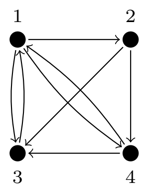
(6.32) Definition (Linkmatrix)**.** Es seien $n \in \mathbb{N}_0$ und ein schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ gegeben. Die Matrix $L \in \mathbb{R}^{n \times n}$ gegeben durch

$$L_{i,j} = \begin{cases} \frac{1}{\deg_+(j)}, & \text{falls } i \in \Gamma_+(j), \\ 0, & \text{falls } i \notin \Gamma_+(j), \end{cases}$$

für $i, j \in [1, n]$ heißt *Linkmatrix* (oder *normierte Adjazenzmatrix*) von G.

(6.33) Beispiel.

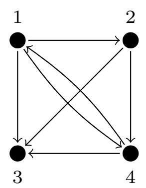
(a) Es sei G der folgende schlichte gerichtete Graph.



Dann ist die Linkmatrix von G gegeben durch

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

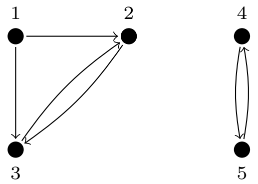
(b) Es sei G der folgende schlichte gerichtete Graph.



Dann ist die Linkmatrix von G gegeben durch

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

(c) Es sei G der folgende schlichte gerichtete Graph.



Dann ist die Linkmatrix von G gegeben durch

$$\left(\begin{array}{ccc|cc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

(6.34) Bemerkung. Es seien $n \in \mathbb{N}_0$ und ein schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ gegeben und es bezeichne L die Linkmatrix von G. Für $j \in [1, n]$ gilt

$$\sum_{i \in [1,n]} L_{i,j} = \begin{cases} 1, & \text{falls } \deg_+(j) > 0, \\ 0, & \text{falls } \deg_+(j) = 0. \end{cases}$$

Beweis. Für $j \in [1, n]$ gilt

$$\sum_{i \in [1,n]} L_{i,j} = \sum_{i \in \Gamma_+(j)} L_{i,j} + \sum_{i \in [1,n] \setminus \Gamma_+(j)} L_{i,j} = \sum_{i \in \Gamma_+(j)} \frac{1}{\deg_+(j)} + \sum_{i \in [1,n] \setminus \Gamma_+(j)} 0 = \begin{cases} 1, & \text{falls } \deg_+(j) > 0, \\ 0, & \text{falls } \deg_+(j) = 0. \end{cases} \qquad \square$$

(6.35) Korollar. Es seien $n \in \mathbb{N}_0$ und ein schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ gegeben und es bezeichne L die Linkmatrix von G. Genau dann ist L stochastisch, wenn $\deg_+(j) > 0$ für alle $j \in [1, n]$ ist.

Mit unserer Interpretation des Internets als schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ für ein $n \in \mathbb{N}_0$ lässt sich die Linkmatrix L wie folgt interpretieren: Ein Surfer klickt sich „zufällig“ durch die Seiten des Internets. Befindet er sich auf einer Seite j, so gibt der Eintrag $L_{i,j}$ der Linkmatrix an, mit welcher Wahrscheinlichkeit er von der Seite j auf die Seite i wechselt. Diese Wahrscheinlichkeit ist 0, sofern es keinen Link von der Seite j zur Seite i gibt, und ansonsten gleich $\frac{1}{\deg_+(j)}$, so dass der Surfer auf jede verlinkte Seite mit der gleichen Wahrscheinlichkeit wechselt.

Dabei lässt sich die Linkmatrix zur Bestimmung eines Rankings nur bei einem „hinreichend guten“ Internet verwenden. Durch die Zufallssurfer-Interpretation erkennen wir, dass wir vor einem Problem stehen, falls eine Webseite j überhaupt keinen Link zu einer anderen Seite hat – in diesem Fall könnte der Surfer nach Betreten der Seite j zu keiner anderen Seite mehr wechseln.

Dieses Problem schlägt sich auch auf theoretischer Seite nieder: Die j-te Spalte der Linkmatrix ist gleich 0 und damit nicht stochastisch. In Beispiel (6.33)(b) haben wir etwa die Linkmatrix $L \in \mathbb{R}^{4 \times 4}$ gegeben durch

$$L = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix},$$

und es gilt $\operatorname{rk}(L - \mathrm{E}_4) = 4$. Folglich ist 1 nach Proposition (6.13) in diesem Fall kein Eigenwert von L.

Wir korrigieren diesen Defekt des Internets, indem wir statt der Linkmatrix L die Matrix $\tilde{L}$ verwenden, welche aus L wie folgt entsteht: Jede Spalte von L, welche gleich 0 ist, wird durch $\frac{1}{n} \sum_{i \in [1,n]} \mathrm{e}_i$ ersetzt, also den stochastischen Vektor, in welchem jeder Eintrag gleich $\frac{1}{n}$ ist. In Beispiel (6.33)(b) ergibt sich etwa

$$\tilde{L} = \begin{pmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & 0 \end{pmatrix}.$$

Durch diese Prozedur erhalten wir offenbar eine stochastische Matrix $\tilde{L}$, welche nach Proposition (6.30) den Eigenwert 1 hat.
Diese Modifikation an der Linkmatrix lässt folgende Interpretation zu: Befindet sich der Zufallssurfer auf einer Seite ohne Links, so verharrt er nicht bis auf alle Zeiten auf dieser Seite, sondern wechselt mit der gleichen Wahrscheinlichkeit auf eine beliebige Seite des Internets.
Ein weiteres Problem, welches im „realen Internet“ auftaucht, lässt sich in Beispiel (6.33)(c) beobachten. Hier ist die Linkmatrix L durch

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

gegeben, sie ist stochastisch und hat nach Proposition (6.30) somit den Eigenwert 1. Allerdings ist

$$(y, z) = (\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix})$$

ein linear unabhängiges Paar von Eigenvektoren von L zum Eigenwert 1, so dass prinzipiell jede Linearkombination $ay + bz$ mit $(a, b) \in (\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}) \setminus \{(0, 0)\}$ eine interpretierbare Lösung des Eigenwertproblems ergibt. Dies liegt daran, dass der gerichtete Graph G in Beispiel (6.33)(c) nicht zusammenhängend ist, er besteht aus zwei Komponenten. Hierdurch lassen sich die Seiten von verschiedenen Komponenten nicht vergleichen.
In der Praxis wird dieses Problem dadurch behoben, dass man die (modifizierte) Linkmatrix mit einem sogenannten Dämpfungsfaktor $d \in [0, 1)_{\mathbb{R}}$ gewichtet, und mit dem Restgewicht $1 - d$ die Matrix $\frac{1}{n} \sum_{i,j \in [1,n]} \mathrm{e}_{i,j}$ eingehen lässt, also die stochastische Matrix, in welcher jeder Eintrag gleich $\frac{1}{n}$ ist. In Beispiel (6.33)(c) ergibt sich hierdurch die Matrix

$$P_d = d \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} + (1 - d) \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix},$$

welche nach Bemerkung (6.29) stochastisch und wegen $d < 1$ sogar komponentenweise positiv ist. Nach dem Satz von Perron/Frobenius (6.31) besitzt P_d genau einen stochastischen Eigenvektor zum Eigenwert 1.
Auch die durch den Dämpfungsfaktor modifizierte Matrix lässt sich wieder an Hand des Zufallssurfer-Modells interpretieren: Anstatt auf jeder Seite zufällig auf einen Link zu klicken, macht der Surfer dies nur mit einer Wahrscheinlichkeit d, mit einer Restwahrscheinlichkeit $1 - d$ wechselt er auf eine beliebige andere Seite des Internets.

(6.36) Definition (Googlematrix)**.** Es seien $n \in \mathbb{N}$ und ein schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ gegeben. Ferner sei $S \in \mathbb{R}^{n \times n}$ gegeben durch

$$S = \sum_{i,j \in [1,n]} \mathrm{e}_{i,j}$$

und es sei $D \in \mathbb{R}^{n \times n}$ gegeben durch

$$D_{i,j} = \begin{cases} 0, & \text{falls } \deg_+(j) > 0, \\ 1, & \text{falls } \deg_+(j) = 0, \end{cases}$$

für $i, j \in [1, n]$. Schließlich bezeichne L die Linkmatrix von G. Für $d \in [0, 1]_{\mathbb{R}}$ heißt $P_d \in \mathbb{R}^{n \times n}$ gegeben durch

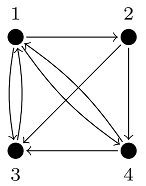
$$P_d = d(L + \frac{1}{n}D) + (1 - d)\frac{1}{n}S$$

die *Googlematrix* von G zum Dämpfungsfaktor d.

Als Dämpfungsfaktor verwendet Google üblicherweise den Wert $d = 0.85$.

(6.37) Beispiel.

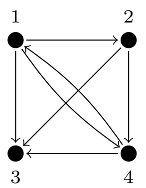
(a) Es sei G der folgende schlichte gerichtete Graph.



Dann ist die Googlematrix von G zum Dämpfungsfaktor 0.85 gegeben durch

$$\begin{pmatrix} \frac{3}{80} & \frac{3}{80} & \frac{71}{80} & \frac{37}{80} \\ \frac{77}{240} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{3}{80} & \frac{37}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{3}{80} & \frac{3}{80} \end{pmatrix} \approx \begin{pmatrix} 0.038 & 0.038 & 0.888 & 0.463 \\ 0.321 & 0.038 & 0.038 & 0.038 \\ 0.321 & 0.463 & 0.038 & 0.463 \\ 0.321 & 0.463 & 0.038 & 0.038 \end{pmatrix}.$$

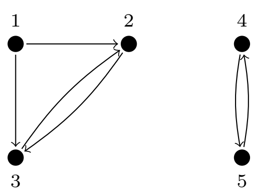
(b) Es sei G der folgende schlichte gerichtete Graph.



Dann ist die Googlematrix von G zum Dämpfungsfaktor 0.85 gegeben durch

$$\begin{pmatrix} \frac{3}{80} & \frac{3}{80} & \frac{1}{4} & \frac{37}{80} \\ \frac{77}{240} & \frac{3}{80} & \frac{1}{4} & \frac{3}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{1}{4} & \frac{37}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{1}{4} & \frac{3}{80} \end{pmatrix} \approx \begin{pmatrix} 0.038 & 0.038 & 0.250 & 0.463 \\ 0.321 & 0.038 & 0.250 & 0.038 \\ 0.321 & 0.463 & 0.250 & 0.463 \\ 0.321 & 0.463 & 0.250 & 0.038 \end{pmatrix}.$$

(c) Es sei G der folgende schlichte gerichtete Graph.



Dann ist die Googlematrix von G zum Dämpfungsfaktor 0.85 gegeben durch

$$\begin{pmatrix} \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} \\ \frac{91}{200} & \frac{3}{100} & \frac{22}{25} & \frac{3}{100} & \frac{3}{100} \\ \frac{91}{200} & \frac{22}{25} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} \\ \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{22}{25} \\ \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{22}{25} & \frac{3}{100} \end{pmatrix} = \begin{pmatrix} 0.030 & 0.030 & 0.030 & 0.030 & 0.030 \\ 0.455 & 0.030 & 0.880 & 0.030 & 0.030 \\ 0.455 & 0.880 & 0.030 & 0.030 & 0.030 \\ 0.030 & 0.030 & 0.030 & 0.030 & 0.880 \\ 0.030 & 0.030 & 0.030 & 0.880 & 0.030 \end{pmatrix}.$$

Beweis.

(a) Nach Beispiel (6.33)(a) ist die Linkmatrix von G gegeben durch

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

Somit ist die Googlematrix von G zum Dämpfungsfaktor 0.85 gegeben durch

$$0.85(\begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}) + 0.15 \cdot \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{3}{80} & \frac{3}{80} & \frac{71}{80} & \frac{37}{80} \\ \frac{77}{240} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{3}{80} & \frac{37}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{3}{80} & \frac{3}{80} \end{pmatrix} \approx \begin{pmatrix} 0.038 & 0.038 & 0.888 & 0.463 \\ 0.321 & 0.038 & 0.038 & 0.038 \\ 0.321 & 0.463 & 0.038 & 0.463 \\ 0.321 & 0.463 & 0.038 & 0.038 \end{pmatrix}.$$

(b) Nach Beispiel (6.33)(b) ist die Linkmatrix von G gegeben durch

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

Somit ist die Googlematrix von G zum Dämpfungsfaktor 0.85 gegeben durch

$$0.85(\begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}) + 0.15 \cdot \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{3}{80} & \frac{3}{80} & \frac{1}{4} & \frac{37}{80} \\ \frac{77}{240} & \frac{3}{80} & \frac{1}{4} & \frac{3}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{1}{4} & \frac{37}{80} \\ \frac{77}{240} & \frac{37}{80} & \frac{1}{4} & \frac{3}{80} \end{pmatrix} \approx \begin{pmatrix} 0.038 & 0.038 & 0.250 & 0.463 \\ 0.321 & 0.038 & 0.250 & 0.038 \\ 0.321 & 0.463 & 0.250 & 0.463 \\ 0.321 & 0.463 & 0.250 & 0.038 \end{pmatrix}.$$

(c) Nach Beispiel (6.33)(c) ist die Linkmatrix von G gegeben durch

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

Somit ist die Googlematrix von G zum Dämpfungsfaktor 0.85 gegeben durch

$$0.85(\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}) + 0.15 \cdot \frac{1}{5}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} \\ \frac{91}{200} & \frac{3}{100} & \frac{22}{25} & \frac{3}{100} & \frac{3}{100} \\ \frac{91}{200} & \frac{22}{25} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} \\ \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{22}{25} \\ \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \frac{22}{25} & \frac{3}{100} \end{pmatrix} = \begin{pmatrix} 0.030 & 0.030 & 0.030 & 0.030 & 0.030 \\ 0.455 & 0.030 & 0.880 & 0.030 & 0.030 \\ 0.455 & 0.880 & 0.030 & 0.030 & 0.030 \\ 0.030 & 0.030 & 0.030 & 0.030 & 0.880 \\ 0.030 & 0.030 & 0.030 & 0.880 & 0.030 \end{pmatrix}. \qquad \square$$

(6.38) Bemerkung. Es seien $n \in \mathbb{N}$, ein schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ und $d \in [0, 1)_{\mathbb{R}}$ gegeben. Die Googlematrix P_d von G zum Dämpfungsfaktor d ist stochastisch und komponentenweise positiv.

Beweis. Es sei $S \in \mathbb{R}^{n \times n}$ gegeben durch

$$S = \sum_{i,j \in [1,n]} \mathrm{e}_{i,j}$$

und es sei $D \in \mathbb{R}^{n \times n}$ gegeben durch

$$D_{i,j} = \begin{cases} 1, & \text{falls } \deg_+(j) = 0, \\ 0, & \text{falls } \deg_+(j) > 0, \end{cases}$$

für $i, j \in [1, n]$. Schließlich bezeichne L die Linkmatrix von G.
Für $i, j \in [1, n]$ gilt dann

$$(L + \frac{1}{n}D)_{i,j} = \left. \begin{cases} \frac{1}{n}, & \text{falls } \deg_+(j) = 0, \\ \frac{1}{\deg_+(j)}, & \text{falls } \deg_+(j) > 0,\ i \in \Gamma_+^G(j), \\ 0, & \text{falls } \deg_+(j) > 0,\ i \notin \Gamma_+^G(j) \end{cases} \right\} \geq 0,$$

d.h. $L + \frac{1}{n}D$ ist komponentenweise nicht-negativ, und für $j \in [1, n]$ gilt

$$\sum_{i \in [1,n]} (L + \frac{1}{n}D)_{i,j} = \left. \begin{cases} \sum_{i \in [1,n]} \frac{1}{n}, & \text{falls } \deg_+(j) = 0, \\ \sum_{i \in \Gamma_+^G(j)} \frac{1}{\deg_+(j)} + \sum_{i \in [1,n] \setminus \Gamma_+^G(j)} 0, & \text{falls } \deg_+(j) > 0 \end{cases} \right\} = 1.$$

Folglich ist $L + \frac{1}{n}D$ stochastisch.
Für $i, j \in [1, n]$ gilt ferner

$$(\frac{1}{n}S)_{i,j} = \frac{1}{n} > 0,$$

d.h. $\frac{1}{n}S$ ist komponentenweise positiv, und für $j \in [1, n]$ gilt

$$\sum_{i \in [1,n]} (\frac{1}{n}S)_{i,j} = \sum_{i \in [1,n]} \frac{1}{n} = 1.$$

Folglich ist $\frac{1}{n}S$ stochastisch.
Nach Bemerkung (6.29) ist mit $L + \frac{1}{n}D$ und $\frac{1}{n}S$ aber auch

$$P_d = d(L + \frac{1}{n}D) + (1 - d)\frac{1}{n}S$$

stochastisch. Wegen $d \in [0, 1)_{\mathbb{R}}$ ist schließlich $d \geq 0$ und $1 - d > 0$, und da $L + \frac{1}{n}D$ komponentenweise nicht-negativ und $\frac{1}{n}S$ komponentweise positiv ist, folgt

$$(P_d)_{i,j} = (d(L + \frac{1}{n}D) + (1 - d)\frac{1}{n}S)_{i,j} = d(L + \frac{1}{n}D)_{i,j} + (1 - d)(\frac{1}{n}S)_{i,j} > 0$$

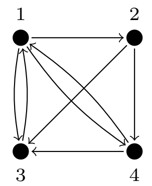
für $i, j \in [1, n]$, d.h. P_d ist komponentenweise positiv. $\square$

Da die Googlematrix eines schlichten gerichteten Graphen nach Bemerkung (6.38) stochastisch und komponentweise positiv ist, hat sie nach dem Satz von Perron/Frobenius (6.31) genau einen stochastischen Eigenvektor zum Eigenwert 1, und dieser ist komponentenweise positiv. An Hand dieses stochastischen Eigenvektors wird das Ranking der Webseiten des Internets erstellt.

(6.39) Definition (PageRank; S. Brin, L. Page; 1998)**.** Es seien $n \in \mathbb{N}$, ein schlichter gerichteter Graph G mit $\mathrm{V}(G) = [1, n]$ und $d \in [0, 1)_{\mathbb{R}}$ gegeben. Ferner bezeichne P_d die Googlematrix von G zum Dämpfungsfaktor d. Der *PageRank* von G zum Dämpfungsfaktor d ist der eindeutige stochastische Eigenvektor von P_d zum Eigenwert 1.

(6.40) Beispiel.

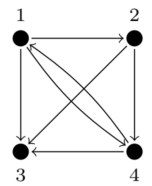
(a) Es sei G der folgende schlichte gerichtete Graph.



Der PageRank von G zum Dämpfungsfaktor 0.85 ist gegeben durch

$$\begin{pmatrix} \frac{319839}{868772} \\ \frac{30800}{217193} \\ \frac{250173}{868772} \\ \frac{43890}{217193} \end{pmatrix} \approx \begin{pmatrix} 0.368 \\ 0.142 \\ 0.288 \\ 0.202 \end{pmatrix}.$$

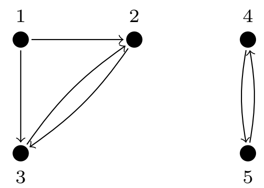
(b) Es sei G der folgende schlichte gerichtete Graph.



Der PageRank von G zum Dämpfungsfaktor 0.85 ist gegeben durch

$$\begin{pmatrix} \frac{22020}{100439} \\ \frac{17600}{100439} \\ \frac{35739}{100439} \\ \frac{25080}{100439} \end{pmatrix} \approx \begin{pmatrix} 0.219 \\ 0.175 \\ 0.356 \\ 0.250 \end{pmatrix}.$$

(c) Es sei G der folgende schlichte gerichtete Graph.



Der PageRank von G zum Dämpfungsfaktor 0.85 ist gegeben durch

$$\begin{pmatrix} \frac{3}{100} \\ \frac{57}{200} \\ \frac{57}{200} \\ \frac{1}{5} \\ \frac{1}{5} \end{pmatrix} = \begin{pmatrix} 0.030 \\ 0.285 \\ 0.285 \\ 0.200 \\ 0.200 \end{pmatrix}.$$

Der gerichtete Graph für das reale Internet besteht aus vielen Millionen Ecken und Kanten und ergibt demnach eine Googlematrix aus vielen Millionen Zeilen und Spalten. Für ein solches Setup lässt sich der PageRank nicht mehr in vertretbarer Zeit mit Hilfe der Gaußelimination berechnen, so dass Approximationsverfahren wie die sogenannte Potenzmethode verwandt werden müssen.

Diagonalisierbarkeit

Wie wir in der Einleitung gesehen haben, gibt es Endomorphismen, welche eine Diagonalmatrix als (eine) Darstellungsmatrix haben. Da dies nicht auf alle Endomorphismen zutrifft, wollen wir ein Kriterium herleiten, welches diese Situation charakterisiert. Durch Analyse einer solchen Darstellungsmatrix haben wir die Begriffe eines Eigenwerts und eines Eigenvektors motiviert. Die Eigenvektoren waren hierbei gerade die Einträge einer Basis. Dies führt uns zu folgendem Begriff, welcher uns bei der genannten Charakterisierung behilflich sein wird.

(6.41) Definition (Eigenbasis)**.** Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Wir sagen, dass s eine *Eigenbasis* von V bzgl. φ ist, wenn s_i für jedes $i \in [1, n]$ ein Eigenvektor von φ ist.

(b) Es sei $A \in K^{n \times n}$ gegeben. Eine *Eigenbasis* von $K^{n \times 1}$ bzgl. A ist eine Eigenbasis von $K^{n \times 1}$ bzgl. φ_A.

(6.42) Beispiel. Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Dann ist

$$(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

eine Eigenbasis von $\mathbb{Q}^{2 \times 1}$ bzgl. A.

Beweis. Dies folgt aus Beispiel (6.3). □

(6.43) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Genau dann ist s eine Eigenbasis von V bzgl. φ, wenn $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix ist.

Beweis. Genau dann ist s eine Eigenbasis von V bzgl. φ, wenn für jedes $j \in [1, n]$ ein $a_j \in K$ mit $\varphi(s_j) = a_j s_j$ existiert. Wegen $\kappa_s(\varphi(s_j)) = (\mathrm{M}_{s,s}(\varphi))_{-,j}$ und $\kappa_s(a_j s_j) = a_j \mathrm{e}_j$ für $a_j \in K$, $j \in [1, n]$ und da $\kappa_s\colon V \to K^{n \times 1}$ eine Bijektion ist, ist dies aber dazu äquivalent, dass es für jedes $j \in [1, n]$ ein $a_j \in K$ mit $(\mathrm{M}_{s,s}(\varphi))_{-,j} = a_j \mathrm{e}_j$ gibt, d.h. dazu, dass $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix ist. □

(6.44) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und ein n-Tupel $s' = (s'_1, \dots, s'_n)$ in V gegeben. Genau dann ist s' eine Eigenbasis von V bzgl. φ, wenn $(\kappa_s(s'_1), \dots, \kappa_s(s'_n))$ eine Eigenbasis von $K^{n \times 1}$ bzgl. $\mathrm{M}_{s,s}(\varphi)$ ist.

Beweis. Da $\kappa_s\colon V \to K^{n \times 1}$ ein Isomorphismus ist, ist $s' = (s'_1, \dots, s'_n)$ nach Proposition (2.21)(c) genau dann eine Basis von V, wenn $(\kappa_s(s'_1), \dots, \kappa_s(s'_n))$ eine Basis von $K^{n \times 1}$ ist. Für jedes $i \in [1, n]$ ist ferner s'_i nach Korollar (6.5) genau dann ein Eigenvektor von φ, wenn $\kappa_s(s'_i)$ ein Eigenvektor von $\mathrm{M}_{s,s}(\varphi)$ ist. Insgesamt ist s' genau dann eine Eigenbasis von V bzgl. φ, wenn $(\kappa_s(s'_1), \dots, \kappa_s(s'_n))$ eine Eigenbasis von $K^{n \times 1}$ bzgl. $\mathrm{M}_{s,s}(\varphi)$ ist. □

(6.45) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n \times n}$, ein n-Tupel $(a_1, \dots, a_n)$ in K und eine Basis $s = (s_1, \dots, s_n)$ von $K^{n \times 1}$ gegeben und es sei $P \in \mathrm{GL}_n(K)$ gegeben durch

$$P = \begin{pmatrix} s_1 & \dots & s_n \end{pmatrix}.$$

Genau dann ist für jedes $i \in [1, n]$ der Vektor s_i ein Eigenvektor von A zum Eigenwert a_i, wenn

$$P^{-1} A P = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

gilt.

Beweis. Zunächst ist

$$\begin{aligned} P^{-1}AP &= (\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{K^{n\times 1}}))^{-1}\,\mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)\,\mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{K^{n\times 1}}) = \mathrm{M}_{s,s}(\varphi_A) = \begin{pmatrix}\kappa_s(\varphi_A(s_1)) & \dots & \kappa_s(\varphi_A(s_n))\end{pmatrix} \\ &= \begin{pmatrix}\kappa_s(As_1) & \dots & \kappa_s(As_n)\end{pmatrix}. \end{aligned}$$

Die Abbildung $\kappa_s\colon V \to K^{n\times 1}$ ist ein Isomorphismus mit $\kappa_s(s_i) = \mathrm{e}_i$ für $i \in [1,n]$. Demnach gilt für $i \in [1,n]$ genau dann $As_i = a_i s_i$, wenn $\kappa_s(As_i) = a_i \mathrm{e}_i$ ist. Folglich ist genau dann s_i für jedes $i \in [1,n]$ ein Eigenvektor von A zum Eigenwert a_i, wenn

$$P^{-1}AP = \begin{pmatrix}a_1\mathrm{e}_1 & \dots & a_n\mathrm{e}_n\end{pmatrix} = \begin{pmatrix}a_1 & & 0\\ & \ddots & \\ 0 & & a_n\end{pmatrix}$$

gilt. □

(6.46) Definition (diagonalisierbar)**.**

(a) Es sei ein endlichdimensionaler K-Vektorraum V gegeben. Ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ heißt *diagonalisierbar*, falls es eine Eigenbasis von V bzgl. φ gibt.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Ein $A \in K^{n\times n}$ heißt *diagonalisierbar* (über K), falls $\varphi_A\colon K^{n\times 1} \to K^{n\times 1}$ diagonalisierbar ist.

(6.47) Beispiel. Es sei $A \in \mathbb{Q}^{2\times 2}$ gegeben durch

$$A = \begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}.$$

Dann ist A diagonalisierbar.

Beweis. Nach Beispiel (6.42) ist

$$(\begin{pmatrix}1\\ -1\end{pmatrix}, \begin{pmatrix}1\\ 1\end{pmatrix})$$

eine Eigenbasis von $\mathbb{Q}^{2\times 1}$ bzgl. A. Somit ist A diagonalisierbar. □

(6.48) Bemerkung. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Genau dann ist φ diagonalisierbar, wenn es eine Basis s von V derart gibt, dass $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix ist.

Beweis. Genau dann ist φ diagonalisierbar, wenn es eine Eigenbasis von V bzgl. φ gibt. Eine Basis s von V ist nach Bemerkung (6.43) genau dann eine Eigenbasis von V bzgl. φ, wenn $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix ist. Insgesamt ist φ genau dann diagonalisierbar, wenn es eine Basis s von V derart gibt, dass $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix ist. □

(6.49) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Genau dann ist φ diagonalisierbar, wenn $\mathrm{M}_{s,s}(\varphi)$ diagonalisierbar ist.

Beweis. Nach Bemerkung (6.44) ist ein n-Tupel $s' = (s'_1, \dots, s'_n)$ in V genau dann eine Eigenbasis von V bzgl. φ, wenn $(\kappa_s(s'_1), \dots, \kappa_s(s'_n))$ eine Eigenbasis von $K^{n\times 1}$ bzgl. $\mathrm{M}_{s,s}(\varphi)$ ist. Da $\kappa_s\colon V \to K^{n\times 1}$ eine Bijektion ist, gibt es folglich genau dann eine Eigenbasis von V bzgl. φ, wenn es eine Eigenbasis von $K^{n\times 1}$ bzgl. $\mathrm{M}_{s,s}(\varphi)$ gibt. Somit ist φ genau dann diagonalisierbar, wenn $\mathrm{M}_{s,s}(\varphi)$ diagonalisierbar ist. □

(6.50) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Die Matrix A ist diagonalisierbar.

(b) Es gibt eine Eigenbasis von $K^{n\times 1}$ bzgl. A.

(c) Die Matrix A ist ähnlich zu einer Diagonalmatrix.

Beweis. Wir zeigen zuerst die Äquivalenz von Bedingung (a) und Bedingung (b), danach die Äquivalenz von Bedingung (a) und Bedingung (c).
Genau dann ist A diagonalisierbar, wenn φ_A diagonalisierbar ist, d.h. wenn es eine Eigenbasis von $K^{n \times 1}$ bzgl. φ_A gibt. Da die Eigenbasen von $K^{n \times 1}$ bzgl. A gerade die Eigenbasen von $K^{n \times 1}$ bzgl. φ_A sind, ist folglich A genau dann diagonalisierbar, wenn es eine Eigenbasis von $K^{n \times 1}$ bzgl. A gibt. Somit sind Bedingung (a) und Bedingung (b) äquivalent.
Genau dann ist A diagonalisierbar, wenn φ_A diagonalisierbar ist, nach Bemerkung (6.49) also genau dann, wenn es eine Basis s von $K^{n \times 1}$ derart gibt, dass $\mathrm{M}_{s,s}(\varphi_A)$ eine Diagonalmatrix ist. Nach Bemerkung (3.63)(b) gilt dies aber genau dann, wenn $A = \mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)$ ähnlich zu einer Diagonalmatrix ist. Somit sind Bedingung (a) und Bedingung (c) äquivalent.
Insgesamt sind Bedingung (a), Bedingung (b) und Bedingung (c) äquivalent. □

(6.51) Beispiel. Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

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

Dann ist A nicht diagonalisierbar.

Beweis. Angenommen, A ist diagonalisierbar. Dann ist A nach Bemerkung (6.50) ähnlich zu einer Diagonalmatrix D, d.h. es gibt eine invertierbare Matrix $P \in \mathbb{Q}^{2 \times 2}$ mit $P^{-1}AP = D$. Da ähnliche Matrizen das gleiche charakteristische Polynom haben, ist 0 als einziger Eigenwert von A der einzige Eigenwert von D. Da D eine Diagonalmatrix ist, folgt $D = 0$ und damit $A = PDP^{-1} = P0P^{-1} = 0$. Da dies ein Widerspruch ist, war unsere Annahme falsch, d.h. A ist nicht diagonalisierbar. □

Wir geben eine Charakterisierung der Diagonalisierbarkeit mittels der zuvor eingeführten Vielfachheiten.

(6.52) Satz. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Der Endomorphismus φ ist diagonalisierbar.

(b) Es ist

$$V = \dot{\sum_{\substack{a \in K \\ a \text{ ist Eigenwert von } \varphi}}} \mathrm{Eig}_a(\varphi).$$

(c) Es ist

$$\dim_K V = \sum_{a \in K} \mathrm{g}_a(\varphi).$$

(d) Das charakteristische Polynom χ_φ zerfällt in Linearfaktoren und für jeden Eigenwert a von φ gilt

$$\mathrm{g}_a(\varphi) = \mathrm{m}_a(\varphi).$$

Beweis. Es seien $a_1, \ldots, a_k$ für ein $k \in \mathbb{N}_0$ die verschiedenen Eigenwerte von φ.
Genau dann ist φ diagonalisierbar, wenn es eine Eigenbasis von V bzgl. φ gibt. Nach Korollar (1.80) ist dies äquivalent dazu, dass

$$V = \dot{\sum_{i \in [1,k]}} \mathrm{Eig}_{a_i}(\varphi)$$

ist. Somit sind Bedingung (a) und Bedingung (b) äquivalent.
Nach Proposition (6.6) ist $(\mathrm{Eig}_{a_1}(\varphi), \ldots, \mathrm{Eig}_{a_k}(\varphi))$ unabhängig, es gilt also stets

$$\sum_{i \in [1,k]} \mathrm{Eig}_{a_i}(\varphi) = \dot{\sum_{i \in [1,k]}} \mathrm{Eig}_{a_i}(\varphi).$$

Da $\dot{\sum}_{i \in [1,k]} \operatorname{Eig}_{a_i}(\varphi)$ ein Untervektorraum von V ist, gilt nach Proposition (1.68) genau dann

$$V = \dot{\sum_{i \in [1,k]}} \operatorname{Eig}_{a_i}(\varphi),$$

wenn

$$\dim V = \dim \dot{\sum_{i \in [1,k]}} \operatorname{Eig}_{a_i}(\varphi)$$

ist. Nach Korollar (1.81) ist jedoch

$$\dim \dot{\sum_{i \in [1,k]}} \operatorname{Eig}_{a_i}(\varphi) = \sum_{i \in [1,k]} \dim \operatorname{Eig}_{a_i}(\varphi) = \sum_{i \in [1,k]} \mathrm{g}_{a_i}(\varphi) = \sum_{a \in K} \mathrm{g}_a(\varphi).$$

Somit sind Bedingung (b) und Bedingung (c) äquivalent.
Nach Proposition (6.19) gilt stets $\mathrm{g}_{a_i}(\varphi) \leq \mathrm{m}_{a_i}(\varphi)$ für $i \in [1,k]$ und damit

$$\sum_{i \in [1,k]} \mathrm{g}_{a_i}(\varphi) \leq \sum_{i \in [1,k]} \mathrm{m}_{a_i}(\varphi) = \sum_{i \in [1,k]} \mathrm{m}_{a_i}(\chi_\varphi) \leq \deg \chi_\varphi = \dim V.$$

Folglich gilt genau dann $\dim V = \sum_{a \in K} \mathrm{g}_a(\varphi) = \sum_{i \in [1,k]} \mathrm{g}_{a_i}(\varphi)$, wenn sowohl $\sum_{i \in [1,k]} \mathrm{g}_{a_i}(\varphi) = \sum_{i \in [1,k]} \mathrm{m}_{a_i}(\varphi)$ als auch $\sum_{i \in [1,k]} \mathrm{m}_{a_i}(\chi_\varphi) = \deg \chi_\varphi$ gilt. Da für $i \in [1,k]$ stets $\mathrm{g}_{a_i}(\varphi) \leq \mathrm{m}_{a_i}(\varphi)$ gilt, ist die Gleichung $\sum_{i \in [1,k]} \mathrm{g}_{a_i}(\varphi) = \sum_{i \in [1,k]} \mathrm{m}_{a_i}(\varphi)$ äquivalent zu den Gleichungen $\mathrm{g}_{a_i}(\varphi) = \mathrm{m}_{a_i}(\varphi)$ für $i \in [1,k]$. Ferner gilt genau dann $\sum_{i \in [1,k]} \mathrm{m}_{a_i}(\chi_\varphi) = \deg \chi_\varphi$, wenn χ_φ in Linearfaktoren zerfällt. Somit sind Bedingung (c) und Bedingung (d) äquivalent.
Insgesamt sind Bedingung (a), Bedingung (b), Bedingung (c) und Bedingung (d) äquivalent. □

(6.53) Korollar. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Wenn die Anzahl der Eigenwerte von φ gleich $\dim V$ ist, dann ist φ diagonalisierbar.

Beweis. Es sei $n := \dim V$ und es seien $a_1, \ldots, a_n$ die Eigenwerte von φ. Dann gilt $\chi_\varphi = \prod_{i \in [1,n]}(X - a_i)$ und damit $\mathrm{m}_{a_i}(\varphi) = 1$ für $i \in [1,n]$. Nach Proposition (6.19) gilt ferner $\mathrm{g}_{a_i}(\varphi) \leq \mathrm{m}_{a_i}(\varphi)$ und wegen $\operatorname{Eig}_{a_i}(\varphi) \neq \{0\}$ ist außerdem $\mathrm{g}_{a_i}(\varphi) = \dim \operatorname{Eig}_{a_i}(\varphi) \geq 1 = \mathrm{m}_{a_i}(\varphi)$, insgesamt also $\mathrm{g}_{a_i}(\varphi) = \mathrm{m}_{a_i}(\varphi)$ für $i \in [1,n]$. Nach Satz (6.52) ist φ diagonalisierbar. □

(6.54) Korollar. Es sei ein normiertes $f \in K[X] \setminus \{0\}$ gegeben. Wenn es $\deg f$ verschiedene Nullstellen von f gibt, dann ist $\mathrm{C}(f)$ diagonalisierbar.

Beweis. Es sei $n := \deg f$, so dass $\mathrm{C}(f) \in K^{n \times n}$ ist. Nach Korollar (6.23) sind die Nullstellen von f gerade die Eigenwerte von $\mathrm{C}(f)$. Wenn es $\deg f$ verschiedene Nullstellen von f gibt, so hat $\mathrm{C}(f)$ also $\deg f = n = \dim K^{n \times 1}$ verschiedene Eigenwerte und ist damit nach Korollar (6.53) diagonalisierbar. □

Alternativer Beweis von Beispiel (6.47). Nach Beispiel (6.15) sind 1 und 3 Eigenwerte von A, so dass A insbesondere 2 verschiedene Eigenwerte hat. Nach Korollar (6.53) ist A diagonalisierbar. □

Alternativer Beweis von Beispiel (6.51). Nach Beispiel (6.18)(b) ist $\mathrm{g}_0(A) = 1 < 2 = \mathrm{m}_0(A)$. Folglich ist A nach Satz (6.52) nicht diagonalisierbar. □

Polynome diagonalisierbarer Matrizen

Zu Beginn dieses Abschnitts haben wir an Hand eines Beispiels gesehen, dass sich Potenzen von Diagonalmatrizen vergleichsweise leicht berechnen lassen. Im Folgenden wollen wir eine Methode zur Berechnung von Potenzen und sogar von Polynomen von diagonalisierbaren Matrizen angeben.
Hierbei werden in ein Polynom Matrizen in folgendem Sinne „eingesetzt“. Für Details siehe Definition (B.18).

(6.55) Notation. Es seien $f \in K[X]$ und $a \in K^{(\mathbb{N}_0)}$ mit $f = \sum_{i \in \mathbb{N}_0} a_i X^i$ gegeben.

(a) Für einen K-Vektorraumendomorphismus $\varphi \colon V \to V$ schreiben wir

$$f(\varphi) := \sum_{i \in \mathbb{N}_0} a_i \varphi^i.$$

(b) Für $n \in \mathbb{N}_0$, $A \in K^{n \times n}$ schreiben wir

$$f(A) := \sum_{i \in \mathbb{N}_0} a_i A^i.$$

(6.56) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n \times n}$, $P \in \mathrm{GL}_n(K)$ und $a_i \in K$ für $i \in [1, n]$ mit

$$P^{-1}AP = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

gegeben. Für $f \in K[X]$ gilt dann

$$f(A) = P \begin{pmatrix} f(a_1) & & 0 \\ & \ddots & \\ 0 & & f(a_n) \end{pmatrix} P^{-1}.$$

Beweis. Zunächst zeigen wir

$$A^k = P \begin{pmatrix} a_1^k & & 0 \\ & \ddots & \\ 0 & & a_n^k \end{pmatrix} P^{-1}$$

für $k \in \mathbb{N}_0$. Hierzu führen wir Induktion nach k. Für $k = 0$ gilt

$$A^k = A^0 = \mathrm{E}_n = P\mathrm{E}_n P^{-1} = P \begin{pmatrix} a_1^0 & & 0 \\ & \ddots & \\ 0 & & a_n^0 \end{pmatrix} P^{-1} = P \begin{pmatrix} a_1^k & & 0 \\ & \ddots & \\ 0 & & a_n^k \end{pmatrix} P^{-1}.$$

Es sei also $k \in \mathbb{N}$ mit

$$A^{k-1} = P \begin{pmatrix} a_1^{k-1} & & 0 \\ & \ddots & \\ 0 & & a_n^{k-1} \end{pmatrix} P^{-1}$$

gegeben. Dann gilt auch

$$\begin{aligned} A^k = A^{k-1}A &= P \begin{pmatrix} a_1^{k-1} & & 0 \\ & \ddots & \\ 0 & & a_n^{k-1} \end{pmatrix} P^{-1} P \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} P^{-1} \\ &= P \begin{pmatrix} a_1^{k-1} & & 0 \\ & \ddots & \\ 0 & & a_n^{k-1} \end{pmatrix} \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} P^{-1} = P \begin{pmatrix} a_1^k & & 0 \\ & \ddots & \\ 0 & & a_n^k \end{pmatrix} P^{-1}. \end{aligned}$$

Nach dem Induktionsprinzip gilt

$$A^k = P \begin{pmatrix} a_1^k & & 0 \\ & \ddots & \\ 0 & & a_n^k \end{pmatrix} P^{-1}$$

für alle $k \in \mathbb{N}_0$.
Nun seien $f \in K[X]$ und $b \in K^{(\mathbb{N}_0)}$ mit $f = \sum_{i \in \mathbb{N}_0} b_i X^i$ gegeben. Dann gilt

$$\begin{aligned} f(A) &= \sum_{i \in \mathbb{N}_0} b_i A^i = \sum_{i \in \mathbb{N}_0} b_i P \begin{pmatrix} a_1^i & & 0 \\ & \ddots & \\ 0 & & a_n^i \end{pmatrix} P^{-1} = P(\sum_{i \in \mathbb{N}_0} b_i \begin{pmatrix} a_1^i & & 0 \\ & \ddots & \\ 0 & & a_n^i \end{pmatrix})P^{-1} \\ &= P(\begin{pmatrix} \sum_{i \in \mathbb{N}_0} b_i a_1^i & & 0 \\ & \ddots & \\ 0 & & \sum_{i \in \mathbb{N}_0} b_i a_n^i \end{pmatrix})P^{-1} = P \begin{pmatrix} f(a_1) & & 0 \\ & \ddots & \\ 0 & & f(a_n) \end{pmatrix} P^{-1}. \end{aligned}$$

□

(6.57) Beispiel. Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Für $k \in \mathbb{N}_0$ gilt

$$A^k = \frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix}.$$

Beweis. Es sei $P \in \mathrm{GL}_2(\mathbb{Q})$ gegeben durch

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Nach Beispiel (6.3) ist $P_{-,1}$ ein Eigenvektor von A zum Eigenwert 1 und $P_{-,2}$ ein Eigenvektor von A zum Eigenwert 3. Ferner ist $(P_{-,1}, P_{-,2})$ linear unabhängig nach Korollar (6.7), also $P \in \mathrm{GL}_2(\mathbb{Q})$, und nach Bemerkung (6.45) gilt

$$P^{-1} A P = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

Nach Bemerkung (6.56) folgt

$$\begin{aligned} A^k &= P \begin{pmatrix} 1^k & 0 \\ 0 & 3^k \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3^k \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 3^k & 3^k \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix} \end{aligned}$$

für $k \in \mathbb{N}_0$. □

Anwendung: Homogene autonome lineare Rekursionsgleichungen

Als nächste Anwendung der Eigenwerttheorie wollen wir sogenannte homogene autonome lineare Rekursionsgleichungen studieren. Betrachten wir zum Einstieg die sogenannte Fibonacci-Folge $f = (f_k)_{k \in \mathbb{N}_0} \in \mathbb{R}^{\mathbb{N}_0}$, welche rekursiv durch

$$f_k = \begin{cases} 0, & \text{für } k = 0, \\ 1, & \text{für } k = 1, \\ f_{k-2} + f_{k-1}, & \text{für } k \in \mathbb{N}_0 \text{ mit } k \geq 2 \end{cases}$$

gegeben ist. Es ist also $f_0 = 0$, $f_1 = 1$, $f_2 = f_0 + f_1 = 1$, $f_3 = f_1 + f_2 = 2$, $f_4 = f_2 + f_3 = 3$, usw. Möchte man nun nur einen Wert der Fibonacci-Folge an einer großen Stelle wissen, etwa f_{10000}, so muss man mit dieser rekursiven Definition erst die Werte f_i für $i \in \mathbb{N}_0$ mit $i \leq 9999$ berechnen. Es stellt sich daher die Frage, ob es nicht eine „geschlossene Formel" für diese Folge gibt, welche einem die direkte Berechnung von f_{10000} ermöglicht.

Es sei $n \in \mathbb{N}_0$ gegeben. Eine *homogene autonome lineare Rekursionsgleichung* vom Grad n in den Unbekannten x_k für $k \in \mathbb{N}_0$ über K ist durch

$$x_{k+n} = a_0 x_k + a_1 x_{k+1} + \ldots + a_{n-1} x_{k+n-1} = \sum_{i \in [0,n-1]} a_i x_{k+i}$$

für $k \in \mathbb{N}_0$, wobei $a_i \in K$ für $i \in [0, n-1]$, gegeben. Genau genommen handelt es sich bei einer homogenen autonomen linearen Rekursionsgleichung also nicht um eine Gleichung, sondern um unendlich viele Gleichungen, welche alle den gleichen Aufbau haben:

$$\begin{array}{rcccccccc} x_n & = & a_0 x_0 & + & a_1 x_1 & + & \ldots & + & a_{n-1} x_{n-1}, \\ x_{n+1} & = & a_0 x_1 & + & a_1 x_2 & + & \ldots & + & a_{n-1} x_n, \\ x_{n+2} & = & a_0 x_2 & + & a_1 x_3 & + & \ldots & + & a_{n-1} x_{n+1}, \\ & \vdots & & & & & & & \end{array}$$

Wenn wir eine solche homogene autonome lineare Rekursionsgleichung lösen wollen, so suchen wir die Menge aller $x = (x_k)_{k \in \mathbb{N}_0} \in K^{\mathbb{N}_0}$, welche alle diese Gleichungen erfüllen.
Während sich lineare Gleichungssysteme mit Hilfe von Matrizen ausdrücken lassen, siehe Definition (A.155), eignen sich normierte Polynome zur Kodierung homogener autonomer linearer Rekursionsgleichungen. Für eine homogene autonome lineare Rekursionsgleichung wie oben nennen wir

$$X^n - \sum_{i \in [0,n-1]} a_i X^i$$

das *charakteristische Polynom* der homogenen autonomen linearen Rekursionsgleichung. Wir werden in Lemma (6.73) sehen, dass in der Tat dieses Polynom und nicht nur dessen Koeffizienten eine tragende Rolle bei der Lösung der homogenen autonomen Rekursionsgleichung eine Rolle spielt.

(6.58) Bemerkung. Die Abbildung

$$\sigma \colon K^{\mathbb{N}_0} \to K^{\mathbb{N}_0},\ x \mapsto (x_{k+1})_{k \in \mathbb{N}_0}$$

ist ein K-Vektorraumendomorphismus von $K^{\mathbb{N}_0}$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.59) Definition (Shift)**.** Der K-Vektorraumendomorphismus

$$\sigma \colon K^{\mathbb{N}_0} \to K^{\mathbb{N}_0},\ x \mapsto (x_{k+1})_{k \in \mathbb{N}_0}$$

wird *Shift* (oder *Shiftendomorphismus*) von $K^{\mathbb{N}_0}$ genannt.

Ähnlich wie bei homogenen linearen Gleichungssystemen, vgl. Korollar (3.36), und bei Interpolationen, vgl. Bemerkung (3.49), ist der Lösungsraum einer homogenen autonomen linearen Rekursionsgleichung der Kern eines geeigneten Vektorraumhomomorphismus und damit insbesondere ein Untervektorraum:

(6.60) Definition (Lösung einer homogenen autonomen linearen Rekursionsgleichung)**.** Es sei ein normiertes $f \in K[X] \setminus \{0\}$ gegeben. Die *Lösungsmenge der homogenen autonomen linearen Rekursionsgleichung* (oder der *Lösungsraum der homogenen autonomen linearen Rekursionsgleichung*) zum charakterischen Polynom f ist definiert als

$$\mathrm{SolRec}(f) := \mathrm{Ker}\, f(\sigma).$$

Ein Element von $\mathrm{SolRec}(f)$ wird *Lösung der homogenen autonomen linearen Rekursionsgleichung* zum charakteristischen Polynom f genannt.

(6.61) Bemerkung. Es seien $n \in \mathbb{N}_0$, $f \in K[X]$ und $a \in K^{[0,n-1]}$ mit $f = X^n - \sum_{i \in [0,n-1]} a_i X^i$ gegeben. Der Lösungsraum der homogenen autonomen linearen Rekursionsgleichung zum charakteristischen Polynom f ist gegeben durch

$$\mathrm{SolRec}(f) = \{x \in K^{\mathbb{N}_0} \mid x_{k+n} = \sum_{i \in [0,n-1]} a_i x_{k+i} \text{ für } k \in \mathbb{N}_0\}.$$

Beweis. Es ist

$$\begin{aligned}
\mathrm{SolRec}(f) &= \operatorname{Ker} f(\sigma) = \{x \in K^{\mathbb{N}_0} \mid (f(\sigma))(x) = 0\} = \{x \in K^{\mathbb{N}_0} \mid (\sigma^n - \sum_{i \in [0,n-1]} a_i \sigma^i)(x) = 0\} \\
&= \{x \in K^{\mathbb{N}_0} \mid \sigma^n(x) - \sum_{i \in [0,n-1]} a_i \sigma^i(x) = 0\} \\
&= \{x \in K^{\mathbb{N}_0} \mid (x_{k+n})_{k \in \mathbb{N}_0} - \sum_{i \in [0,n-1]} a_i (x_{k+i})_{k \in \mathbb{N}_0} = 0\} \\
&= \{x \in K^{\mathbb{N}_0} \mid (x_{k+n} - \sum_{i \in [0,n-1]} a_i x_{k+i})_{k \in \mathbb{N}_0} = 0\} \\
&= \{x \in K^{\mathbb{N}_0} \mid x_{k+n} - \sum_{i \in [0,n-1]} a_i x_{k+i} = 0 \text{ für } k \in \mathbb{N}_0\} \\
&= \{x \in K^{\mathbb{N}_0} \mid x_{k+n} = \sum_{i \in [0,n-1]} a_i x_{k+i} \text{ für } k \in \mathbb{N}_0\}.
\end{aligned}$$

□

(6.62) Beispiel. Es sei $a \in K$ gegeben. Für $x \in K^{\mathbb{N}_0}$ gilt genau dann

$$x_{k+1} = a x_k$$

für $k \in \mathbb{N}_0$, wenn

$$x_k = a^k x_0$$

für $k \in \mathbb{N}_0$ ist. Der Lösungsraum der homogenen autonomen linearen Rekursionsgleichung zum charakteristischen Polynom $X - a$ ist gegeben durch

$$\mathrm{SolRec}(X - a) = K(a^k)_{k \in \mathbb{N}_0}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

In Korollar (6.74) werden wir sehen, dass der Lösungsraum einer homogenen autonomen linearen Rekursionsgleichung endlichdimensional ist und die Dimension gerade dem Grad ihres charakteristischen Polynoms entspricht.

Die Fibonacci-Folge $f = (f_k)_{k \in \mathbb{N}_0}$ in $\mathbb{R}$ erfüllt die homogene autonome lineare Rekursionsgleichung

$$f_{k+2} = f_k + f_{k+1}$$

für $k \in \mathbb{N}_0$. Der Trick zum Lösen einer solchen Rekursionsgleichung besteht darin, dass wir die Gleichung zu einem ganzen Gleichungssystem „aufblasen“: Statt f betrachten wir die Folge $g = (g_k)_{k \in \mathbb{N}_0}$ in $\mathbb{R}^{2 \times 1}$ gegeben durch

$$g_k = \begin{pmatrix} f_k \\ f_{k+1} \end{pmatrix}$$

für $k \in \mathbb{N}_0$. Diese Folge mit Einträgen in $\mathbb{R}^{2 \times 1}$ erfüllt

$$g_{k+1} = \begin{pmatrix} f_{k+1} \\ f_{k+2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} f_k \\ f_{k+1} \end{pmatrix} = \mathrm{C}(X^2 - X - 1)^{\mathrm{tr}} \, g_k$$

für $k \in \mathbb{N}_0$. Wir erkennen also, dass das Studium einer homogenen autonomen Rekursionsgleichung vom Grad 2 zum Studium eines sogenannten homogenen autonomen linearen Rekursionsgleichungssystems von Grad 1 führt. Hierbei entspricht das Polynom $X^2 - X - 1$ gerade dem charakteristischen Polynom der die Fibonacci-Folge definierenden homogenen autonomen Rekursionsgleichung sowie dem charakteristischen Polynom der Matrix $\mathrm{C}(X^2 - X - 1)^{\mathrm{tr}}$, siehe Proposition (6.22).

Es sei $n \in \mathbb{N}_0$ gegeben. Ein *homogenes autonomes lineares Rekursionsgleichungssystem* vom Grad 1 in den Unbekannten x_k für $k \in \mathbb{N}_0$ über K ist durch

$$x_{k+1} = A x_k$$

für $k \in \mathbb{N}_0$, wobei $A \in K^{n \times n}$, gegeben. Wenn wir ein solches homogenes autonomes lineares Rekursionsgleichungssystem lösen wollen, so suchen wir die Menge aller $x = (x_k)_{k \in \mathbb{N}_0} \in (K^{n \times 1})^{\mathbb{N}_0}$, welche alle diese Gleichungen erfüllen.
Wir werden sehen, dass sich die Lösungen von homogenen autonomen linearen Rekursionsgleichungssystemen vom Grad 1 als direkte Verallgemeinerungen der Lösungen von homogenen autonomen linearen Rekursionsgleichungen vom Grad 1 ergeben, siehe Beispiel (6.62) und Proposition (6.67). Da homogene autonome lineare Rekursionsgleichungen wie oben angedeutet in Zusammenhang zu homogenen autonomen linearen Rekursionsgleichungssystemen vom Grad 1 stehen, ergeben sich hieraus auch die Lösungen von homogenen autonomen linearen Rekursionsgleichungen beliebigen Grades als Spezialfälle, siehe Korollar (6.74). Der präzise Zusammenhang zwischen homogenen autonomen linearen Rekursionsgleichungen und homogenen autonomen linearen Rekursionsgleichungssystemen vom Grad 1 findet sich in Lemma (6.73).

(6.63) Notation. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Wir schreiben

$$\tau_A \colon (K^{n \times 1})^{\mathbb{N}_0} \to (K^{n \times 1})^{\mathbb{N}_0},\ x \mapsto (x_{k+1} - A x_k)_{k \in \mathbb{N}_0}.$$

(6.64) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Die Abbildung

$$\tau_A \colon (K^{n \times 1})^{\mathbb{N}_0} \to (K^{n \times 1})^{\mathbb{N}_0},\ x \mapsto (x_{k+1} - A x_k)_{k \in \mathbb{N}_0}$$

ist ein K-Vektorraumendomorphismus von $(K^{n \times 1})^{\mathbb{N}_0}$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.65) Definition (Lösung eines homogenen autonomen linearen Rekursionsgleichungssystems)**.** Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Die *Lösungsmenge des homogenen autonomen linearen Rekursionsgleichungssystems* (oder der *Lösungsraum des homogenen autonomen linearen Rekursionsgleichungssystems*) zur Koeffizientenmatrix A ist definiert als

$$\operatorname{SolRec}(A) := \operatorname{Ker} \tau_A.$$

Ein Element von $\operatorname{SolRec}(A)$ wird *Lösung des homogenen autonomen linearen Rekursionsgleichungssystems* zur Koeffizientenmatrix A genannt.

(6.66) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Der Lösungsraum des homogenen autonomen linearen Rekursionsgleichungssystems zur Koeffizientenmatrix A ist gegeben durch

$$\operatorname{SolRec}(A) = \{x \in (K^{n \times 1})^{\mathbb{N}_0} \mid x_{k+1} = A x_k \text{ für } k \in \mathbb{N}_0\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Als nächstes geben wir eine Beschreibung des Lösungsraums eines beliebigen homogenen autonomen linearen Rekursionsgleichungssystems vom Grad 1 als Verallgemeinerung zu Beispiel (6.62) an.

(6.67) Proposition. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Die Abbildungen

$$\operatorname{SolRec}(A) \to K^{n \times 1},\ x \mapsto x_0,$$
$$K^{n \times 1} \to \operatorname{SolRec}(A),\ c \mapsto (A^k c)_{k \in \mathbb{N}_0}$$

sind wohldefinierte, sich gegenseitig invertierende K-Vektorraumisomorphismen. Insbesondere ist der Lösungsraum des homogenen autonomen linearen Rekursionsgleichungssystems zur Koeffizientenmatrix A gegeben durch

$$\operatorname{SolRec}(A) = \{(A^k c)_{k \in \mathbb{N}_0} \mid c \in K^{n \times 1}\}.$$

Beweis. Es sei $\varphi \colon \operatorname{SolRec}(A) \to K^{n \times 1}$, $x \mapsto x_0$. Für $x, y \in (K^{n \times 1})^{\mathbb{N}_0}$ gilt

$$\varphi(x + y) = \varphi((x_k + y_k)_{k \in \mathbb{N}_0}) = x_0 + y_0 = \varphi(x) + \varphi(y).$$

Für $a \in K$, $x \in (K^{n \times 1})^{\mathbb{N}_0}$ gilt

$$\varphi(ax) = \varphi((a x_k)_{k \in \mathbb{N}_0}) = a x_0 = a\,\varphi(x).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein Homomorphismus.
Da es nach dem Rekursionssatz [12, Prop. (9.5)] für jedes $c \in K^{n \times 1}$ genau ein $x \in \mathrm{SolRec}(A)$ mit $x_0 = c$ gibt, ist φ bijektiv und damit ein Isomorphismus nach Bemerkung (2.9).
Für $c \in K^{n \times 1}$ gilt $A^{k+1}c = AA^k c$ für alle $k \in \mathbb{N}_0$ und damit $(A^k c)_{k \in \mathbb{N}_0} \in \mathrm{SolRec}(A)$. Somit erhalten wir eine wohldefinierte Abbildung

$$\psi\colon K^{n \times 1} \to \mathrm{SolRec}(A),\ c \mapsto (A^k c)_{k \in \mathbb{N}_0}.$$

Wegen

$$\varphi(\psi(c)) = \varphi((A^k c)_{k \in \mathbb{N}_0}) = A^0 c = c$$

für $c \in K^{n \times 1}$ gilt $\varphi \circ \psi = \mathrm{id}_{K^{n \times 1}}$ und damit $\varphi^{-1} = \psi$. Insbesondere ist ψ nach Bemerkung (2.9) ebenfalls ein Isomorphismus. □

(6.68) Beispiel. Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Für $x \in (\mathbb{Q}^{2 \times 1})^{\mathbb{N}_0}$ gilt genau dann

$$x_{k+1} = Ax_k$$

für $k \in \mathbb{N}_0$, wenn

$$x_k = \frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix} x_0$$

für $k \in \mathbb{N}_0$ ist. Der Lösungsraum des homogenen autonomen linearen Rekursionsgleichungssystems zur Koeffizientenmatrix A ist gegeben durch

$$\mathrm{SolRec}(A) = \mathbb{Q}\,(\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0} + \mathbb{Q}\,(3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0}.$$

Beweis. Es sei $x \in (\mathbb{Q}^{2 \times 1})^{\mathbb{N}_0}$ gegeben. Nach Proposition (6.67) gilt genau dann $x \in \mathrm{SolRec}(A)$, wenn $x = (A^k x_0)_{k \in \mathbb{N}_0}$ ist. Da nach Beispiel (6.57) für $k \in \mathbb{N}_0$ stets

$$A^k = \frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix}$$

ist, gilt somit genau dann $x_{k+1} = Ax_k$ für $k \in \mathbb{N}_0$, wenn

$$x_k = \frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix} x_0$$

für $k \in \mathbb{N}_0$ ist.
Nach Korollar (1.23)(b), (c) gilt insbesondere

$$\begin{aligned}
\mathrm{SolRec}(A) &= \{(A^k c)_{k \in \mathbb{N}_0} \mid c \in \mathbb{Q}^{2 \times 1}\} = \{(\frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix} c)_{k \in \mathbb{N}_0} \mid c \in \mathbb{Q}^{2 \times 1}\} \\
&= \{(c_1 \frac{1}{2} \begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix} + c_2 \frac{1}{2} \begin{pmatrix} -1 + 3^k \\ 1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} \mid c \in \mathbb{Q}^{2 \times 1}\} \\
&= \{c_1 \frac{1}{2} (\begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} + c_2 \frac{1}{2} (\begin{pmatrix} -1 + 3^k \\ 1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} \mid c \in \mathbb{Q}^{2 \times 1}\} \\
&= \langle \frac{1}{2} (\begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0}, \frac{1}{2} (\begin{pmatrix} -1 + 3^k \\ 1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} \rangle \\
&= \langle \frac{1}{2} (\begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0}, \frac{1}{2} (\begin{pmatrix} -1 + 3^k \\ 1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} + \frac{1}{2} (\begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} \rangle \\
&= \langle (\begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0}, (\begin{pmatrix} 3^k \\ 3^k \end{pmatrix})_{k \in \mathbb{N}_0} \rangle = \langle (\begin{pmatrix} 1 + 3^k \\ -1 + 3^k \end{pmatrix})_{k \in \mathbb{N}_0} - (\begin{pmatrix} 3^k \\ 3^k \end{pmatrix})_{k \in \mathbb{N}_0}, (\begin{pmatrix} 3^k \\ 3^k \end{pmatrix})_{k \in \mathbb{N}_0} \rangle \\
&= \langle (\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0}, (3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0} \rangle = \mathbb{Q}\,(\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0} + \mathbb{Q}\,(3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0}.
\end{aligned}$$

□

Im Beweis zu Beispiel (6.68) haben wir gesehen, dass wir die zunächst erhaltene Basis

$$(\frac{1}{2}(\begin{pmatrix} 1+3^k \\ -1+3^k \end{pmatrix})_{k \in \mathbb{N}_0}, \frac{1}{2}(\begin{pmatrix} -1+3^k \\ 1+3^k \end{pmatrix})_{k \in \mathbb{N}_0})$$

noch etwas umformen mussten, um auf die angegebene Basis

$$((\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0}, (3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0})$$

von etwas einfacherer Gestalt zu kommen. Dies liegt daran, dass die Basis des Lösungsraums eines homogenen autonomen linearen Rekursionsgleichungssystems im Allgemeinen nicht eindeutig ist. Wir wollen nun zeigen, dass die Angabe einer „schönen" Basis wie in Beispiel (6.68) möglich ist, sofern die Koeffizientenmatrix des homogenen autonomen linearen Rekursionsgleichungssystems diagonalisierbar ist, siehe Bemerkung (6.72). Zunächst betrachten wir beliebige Basen des Lösungsraums:

(6.69) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n \times n}$ und eine Basis $(s_1, \dots, s_n)$ von $K^{n \times 1}$ gegeben. Dann ist

$$t := ((A^k s_1)_{k \in \mathbb{N}_0}, \dots (A^k s_n)_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$ und für $x \in \mathrm{SolRec}(A)$ ist die Koordinatenspalte $\kappa_t(x)$ von x zur Basis t gerade die eindeutige Lösung des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\left(\begin{array}{ccc|c} s_1 & \dots & s_n & x_0 \end{array}\right)$.

Beweis. Nach Proposition (6.67) sind $\varphi\colon \mathrm{SolRec}(A) \to K^{n \times 1}$, $x \mapsto x_0$ und $\psi\colon K^{n \times 1} \to \mathrm{SolRec}(A)$, $c \mapsto (A^k c)_{k \in \mathbb{N}_0}$ wohldefinierte, sich gegenseitig invertierende K-Vektorraumisomorphismen, so dass nach Proposition (2.21) das n-Tupel

$$t = (\psi(s_1), \dots, \psi(s_n)) = ((A^k s_1)_{k \in \mathbb{N}_0}, \dots (A^k s_n)_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$ ist.
Die Darstellungsmatrix von ψ zu den Basen s und t ist $\mathrm{M}_{t,s}(\psi) = \mathrm{E}_n$. Für $x \in \mathrm{SolRec}(A)$ gilt daher

$$\begin{aligned} \begin{pmatrix} s_1 & \dots & s_n \end{pmatrix} \kappa_t(x) &= \mathrm{M}_{\mathrm{e},s}(\mathrm{id}_{K^{n\times 1}})(\mathrm{M}_{t,s}(\psi))^{-1} \kappa_t(x) = \mathrm{M}_{\mathrm{e},t}(\mathrm{id}_{K^{n\times 1}} \circ \psi^{-1})\, \kappa_t(x) = \mathrm{M}_{\mathrm{e},t}(\varphi)\, \kappa_t(x) \\ &= \kappa_{\mathrm{e}}(\varphi(x)) = x_0 \end{aligned}$$

nach Proposition (3.13)(c), (a) und Proposition (3.11), d.h. $\kappa_t(x) \in \mathrm{Sol}(\begin{pmatrix} s_1 & \dots & s_n \end{pmatrix}, x_0)$. ([6]) □

(6.70) Korollar. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Dann ist

$$t := (((A^k)_{-,1})_{k \in \mathbb{N}_0}, \dots ((A^k)_{-,n})_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$ und für $x \in \mathrm{SolRec}(A)$ ist die Koordinatenspalte von x zur Basis t gegeben durch

$$\kappa_t(x) = x_0.$$

Beweis. Nach Bemerkung (6.69) ist

$$((A^k \mathrm{e}_1)_{k \in \mathbb{N}_0}, \dots (A^k \mathrm{e}_n)_{k \in \mathbb{N}_0}) = (((A^k)_{-,1})_{k \in \mathbb{N}_0}, \dots ((A^k)_{-,n})_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$ und für $x \in \mathrm{SolRec}(A)$ gilt $\{x_0\} = \mathrm{Sol}(\mathrm{E}_n, x_0) = \mathrm{Sol}(\begin{pmatrix} \mathrm{e}_1 & \dots & \mathrm{e}_n \end{pmatrix}, x_0) = \{\kappa_t(x)\}$, also $\kappa_t(x) = x_0$. □

Als nächstes zeigen wir, dass Eigenwerte und Eigenvektoren der Koeffizientenmatrix eines homogenen autonomen linearen Rekursionsgleichungssystems zu Lösungen von besonders einfacher Gestalt führen.

(6.71) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Für jeden Eigenwert a von A und jeden Eigenvektor c von A zum Eigenwert a ist

$$(a^k c)_{k \in \mathbb{N}_0}$$

eine Lösung des homogenen autonomen linearen Rekursionsgleichungssystems zur Koeffizientenmatrix A.

[6] Da $(s_1, \dots, s_n)$ eine Basis von $K^{n \times 1}$ ist, gilt $\mathrm{rk} \begin{pmatrix} s_1 & \dots & s_n \end{pmatrix} = n$. Das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\left(\begin{array}{ccc|c} s_1 & \dots & s_n & x_0 \end{array}\right)$ hat nach Bemerkung (3.42) also nur $\kappa_t(x)$ als Lösung.

Beweis. Für jeden Eigenwert a von A und jeden Eigenvektor c von A zum Eigenwert a ist

$$(a^k c)_{k \in \mathbb{N}_0} = (A^k c)_{k \in \mathbb{N}_0} \in \mathrm{SolRec}(A)$$

nach Proposition (6.67). □

(6.72) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n \times n}$, ein n-Tupel $(a_1, \dots, a_n)$ in K und eine Basis $(s_1, \dots, s_n)$ von $K^{n \times 1}$ so gegeben, dass für jedes $i \in [1, n]$ der Vektor s_i ein Eigenvektor von A zum Eigenwert a_i ist. Dann ist

$$t := ((a_1^k s_1)_{k \in \mathbb{N}_0}, \dots (a_n^k s_n)_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$ und für $x \in \mathrm{SolRec}(A)$ ist die Koordinatenspalte $\kappa_t(x)$ von x zur Basis t gerade die eindeutige Lösung des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\left(\begin{array}{ccc|c} s_1 & \dots & s_n & x_0 \end{array}\right)$.

Beweis. Nach Bemerkung (6.69) ist

$$((A^k s_1)_{k \in \mathbb{N}_0}, \dots (A^k s_n)_{k \in \mathbb{N}_0}) = ((a_1^k s_1)_{k \in \mathbb{N}_0}, \dots (a_n^k s_n)_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$ und für $x \in \mathrm{SolRec}(A)$ gilt $\mathrm{Sol}(\begin{pmatrix} s_1 & \dots & s_n \end{pmatrix}, x_0) = \{\kappa_t(x)\}$. □

Mit Hilfe der Erkenntnisse aus Bemerkung (6.72) lässt sich der Beweis von Beispiel (6.68) ohne die Vorarbeit des Potenzierens der Koeffizientenmatrix aus Beispiel (6.57) führen.

Alternativer Beweis von Beispiel (6.68). Es sei (s_1, s_2) in $\mathbb{Q}^{2 \times 1}$ gegeben durch

$$s_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix},\ s_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Nach Beispiel (6.3) ist dann s_1 ein Eigenvektor von A zum Eigenwert 1 und s_2 ein Eigenvektor von A zum Eigenwert 3. Ferner ist (s_1, s_2) nach Korollar (6.7) linear unabhängig und damit eine Basis von $\mathbb{Q}^{2 \times 1}$. Nach Bemerkung (6.72) ist somit

$$t := ((1^k \begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0}, (3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0}) = ((\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0}, (3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(A)$, so dass insbesondere

$$\mathrm{SolRec}(A) = \mathbb{Q}\,(1^k \begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0} + \mathbb{Q}\,(3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0} = \mathbb{Q}\,(\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0} + \mathbb{Q}\,(3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0}$$

gilt. Für $x \in \mathrm{SolRec}(A)$ ist ferner $\mathrm{Sol}(\begin{pmatrix} s_1 & s_2 \end{pmatrix}, x_0) = \{\kappa_t(x)\}$, also

$$\kappa_t(x) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^{-1} x_0 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x_0.$$

Folglich gilt für $x \in (\mathbb{Q}^{2 \times 1})^{\mathbb{N}_0}$ genau dann $x_{k+1} = A x_k$ für $k \in \mathbb{N}_0$, wenn

$$x = (\frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x_0)_1 \, (\begin{pmatrix} 1 \\ -1 \end{pmatrix})_{k \in \mathbb{N}_0} + (\frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x_0)_2 \, (3^k \begin{pmatrix} 1 \\ 1 \end{pmatrix})_{k \in \mathbb{N}_0} = (\begin{pmatrix} 1 & 3^k \\ -1 & 3^k \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x_0)_{k \in \mathbb{N}_0}$$

ist, d.h. wenn

$$x_k = \begin{pmatrix} 1 & 3^k \\ -1 & 3^k \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x_0 = \frac{1}{2} \begin{pmatrix} 1 + 3^k & -1 + 3^k \\ -1 + 3^k & 1 + 3^k \end{pmatrix} x_0$$

für $k \in \mathbb{N}_0$ gilt. □

Nachdem wir einen Überblick über den Lösungsraum eines beliebigen homogenen autonomen linearen Rekursionsgleichungssystems vom Grad 1 gewonnen haben, geben wir nun den Zusammenhang zwischen homogenen autonomen linearen Rekursionsgleichungen beliebigen Grades und eben solchen homogenen autonomen linearen Rekursionsgleichungssystemen vom Grad 1 an.

(6.73) Lemma. Es seien $n \in \mathbb{N}$ und ein normiertes $f \in K[X] \setminus \{0\}$ mit $\deg f = n$ gegeben. Dann sind

$$\mathrm{SolRec}(f) \to \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}),\ x \mapsto (\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0},$$

$$\mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}) \to \mathrm{SolRec}(f),\ y \mapsto (y_{k,1})_{k \in \mathbb{N}_0}$$

wohldefinierte, sich gegenseitig invertierende K-Vektorraumisomorphismen. ([7])

Beweis. Es sei $a \in K^{[0,n-1]}$ mit $f = X^n - \sum_{i \in [0,n-1]} a_i X^i$ gegeben. Für $y \in (K^{n \times 1})^{\mathbb{N}_0}$ gilt dann

$$\begin{aligned}
\mathrm{C}(f)^{\mathrm{tr}} y_k &= \begin{pmatrix} 0 & & 0 & a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & a_{n-2} \\ 0 & & 1 & a_{n-1} \end{pmatrix}^{\mathrm{tr}} \begin{pmatrix} y_{k,1} \\ \vdots \\ y_{k,n-1} \\ y_{k,n} \end{pmatrix} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \\ a_0 & \dots & a_{n-2} & a_{n-1} \end{pmatrix} \begin{pmatrix} y_{k,1} \\ \vdots \\ y_{k,n-1} \\ y_{k,n} \end{pmatrix} \\
&= \begin{pmatrix} y_{k,2} \\ \vdots \\ y_{k,n} \\ \sum_{i \in [0,n-1]} a_i y_{k,i+1} \end{pmatrix}
\end{aligned}$$

für $k \in \mathbb{N}_0$.
Insbesondere gilt für $x \in \mathrm{SolRec}(f)$ stets $x_{k+n} = \sum_{i \in [0,n-1]} a_i x_{k+i}$ für $k \in \mathbb{N}_0$, also auch

$$\begin{pmatrix} x_{k+1} \\ \vdots \\ x_{k+n-1} \\ x_{k+n} \end{pmatrix} = \begin{pmatrix} x_{k+1} \\ \vdots \\ x_{k+n-1} \\ \sum_{i \in [0,n-1]} a_i x_{k+i} \end{pmatrix} = \mathrm{C}(f)^{\mathrm{tr}} \begin{pmatrix} x_k \\ \vdots \\ x_{k+n-2} \\ x_{k+n-1} \end{pmatrix}$$

für $k \in \mathbb{N}_0$, d.h. es ist

$$(\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} \in \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}).$$

Somit haben wir eine wohldefinierte Abbildung

$$\varphi\colon \mathrm{SolRec}(f) \to \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}),\ x \mapsto (\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0}.$$

Für $x, y \in \mathrm{SolRec}(f)$ gilt

$$\begin{aligned}
\varphi(x+y) &= \varphi((x_k + y_k)_{k \in \mathbb{N}_0}) = (\begin{pmatrix} x_k + y_k \\ \vdots \\ x_{k+n-1} + y_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} = (\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix} + \begin{pmatrix} y_k \\ \vdots \\ y_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} \\
&= (\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} + (\begin{pmatrix} y_k \\ \vdots \\ y_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} = \varphi(x) + \varphi(y).
\end{aligned}$$

[7] Im Fall $n = 0$ sind $\mathrm{SolRec}(f) \to \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}})$, $x \mapsto (())_{k \in \mathbb{N}_0}$ und $\mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}) \to \mathrm{SolRec}(f)$, $y \mapsto (0)_{k \in \mathbb{N}_0}$ wohldefinierte, sich gegenseitig invertierende K-Vektorraumisomorphismen.

Für $a \in K$, $x \in \operatorname{SolRec}(f)$ gilt

$$\varphi(ax) = \varphi((ax_k)_{k \in \mathbb{N}_0}) = (\begin{pmatrix} ax_k \\ \vdots \\ ax_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} = (a \begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} = a\,(\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0} = a\,\varphi(x).$$

Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ ein Homomorphismus.
Für $y \in \operatorname{SolRec}(\mathrm{C}(f)^{\mathrm{tr}})$ gilt ferner

$$\begin{pmatrix} y_{k+1,1} \\ \vdots \\ y_{k+1,n} \end{pmatrix} = y_{k+1} = \mathrm{C}(f)^{\mathrm{tr}} y_k = \begin{pmatrix} y_{k,2} \\ \vdots \\ y_{k,n} \\ \sum_{i \in [0,n-1]} a_i y_{k,i+1} \end{pmatrix}$$

für $k \in \mathbb{N}_0$, also auch

$$y_{k+n,1} = y_{k+1,n} = \sum_{i \in [0,n-1]} a_i y_{k,i+1} = \sum_{i \in [0,n-1]} a_i y_{k+i,1}$$

für $k \in \mathbb{N}_0$, d.h. es ist $(y_{k,1})_{k \in \mathbb{N}_0} \in \operatorname{SolRec}(f)$. Somit haben wir eine wohldefinierte Abbildung

$$\psi\colon \operatorname{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}) \to \operatorname{SolRec}(f),\ y \mapsto (y_{k,1})_{k \in \mathbb{N}_0}.$$

Wir erhalten

$$\psi(\varphi(x)) = \psi((\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0}) = (x_k)_{k \in \mathbb{N}_0} = x$$

für $x \in \operatorname{SolRec}(f)$ sowie

$$\varphi(\psi(y)) = \varphi((y_{k,1})_{k \in \mathbb{N}_0}) = (\begin{pmatrix} y_{k,1} \\ \vdots \\ y_{k+n-1,1} \end{pmatrix})_{k \in \mathbb{N}_0} = (\begin{pmatrix} y_{k,1} \\ \vdots \\ y_{k,n} \end{pmatrix})_{k \in \mathbb{N}_0} = (y_k)_{k \in \mathbb{N}_0} = y$$

für $y \in \operatorname{SolRec}(\mathrm{C}(f)^{\mathrm{tr}})$. Somit gilt $\psi \circ \varphi = \mathrm{id}_{\operatorname{SolRec}(f)}$ und $\varphi \circ \psi = \mathrm{id}_{\operatorname{SolRec}(A)}$, d.h. φ und ψ sind sich gegenseitig invertierende Abbildungen. Da φ ein Homomorphismus ist, sind φ und $\psi = \varphi^{-1}$ nach Bemerkung (2.9) sogar Isomorphismen. □

Unsere Erkenntnisse über die Lösungsmenge eines linearen Rekursionsgleichungssystems vom Grad 1 liefern nun folgende Beschreibung des Lösungsraums einer beliebigen homogenen autonomen linearen Rekursionsgleichung:

(6.74) Korollar. Es seien $n \in \mathbb{N}$ und ein normiertes $f \in K[X] \setminus \{0\}$ mit $\deg f = n$ gegeben. Die Abbildungen

$$\operatorname{SolRec}(f) \to K^{n \times 1},\ x \mapsto \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix},$$

$$K^{n \times 1} \to \operatorname{SolRec}(f),\ c \mapsto (((\mathrm{C}(f)^{\mathrm{tr}})^k\, c)_1)_{k \in \mathbb{N}_0}$$

sind wohldefinierte, sich gegenseitig invertierende Vektorraumisomorphismen. Insbesondere ist der Lösungsraum der homogenen autonomen linearen Rekursionsgleichung zum charakteristischen Polynom f gegeben durch

$$\operatorname{SolRec}(f) = \{(((\mathrm{C}(f)^{\mathrm{tr}})^k c)_1)_{k \in \mathbb{N}_0} \mid c \in K^{n \times 1}\}.$$

Beweis. Nach Lemma (6.73) ist

$$\varphi\colon \operatorname{SolRec}(f) \to \operatorname{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}),\ x \mapsto (\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0}$$

ein wohldefinierter Isomorphismus mit Inverser

$$\varphi^{-1} \colon \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}) \to \mathrm{SolRec}(f),\, y \mapsto (y_{k,1})_{k \in \mathbb{N}_0}$$

und nach Proposition (6.67) ist

$$\psi \colon \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}) \to K^{n \times 1},\, y \mapsto y_0$$

ein Isomorphismus mit Inverser

$$\psi^{-1} \colon K^{n \times 1} \to \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}),\, c \mapsto ((\mathrm{C}(f)^{\mathrm{tr}})^k\, c)_{k \in \mathbb{N}_0}.$$

Folglich ist auch das Kompositum

$$\psi \circ \varphi \colon \mathrm{SolRec}(f) \to K^{n \times 1},\, x \mapsto \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix}$$

ein wohldefinierter Isomorphismus mit Inverser

$$(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1} \colon K^{n \times 1} \to \mathrm{SolRec}(f),\, c \mapsto (((\mathrm{C}(f)^{\mathrm{tr}})^k\, c)_1)_{k \in \mathbb{N}_0}. \qquad \square$$

Mit Hilfe der Eigenwerttheorie lässt sich eine homogene autonome lineare Rekursionsgleichung nach Korollar (6.74) also lösen, indem man die Potenzen der Transponierten der Begleitmatrix ihres charakteristischen Polynoms berechnet. Hierzu benutzen wir die in Bemerkung (6.56) angegebene Methode, sofern die Transponierte der Begleitmatrix diagonalisierbar ist. Wenn die Begleitmatrix nicht diagonalisierbar ist, lassen sich ähnliche Methoden verwenden, die allerdings im Allgemeinen etwas tiefergehende Resultate der Eigenwerttheorie (die sogenannte Jordanform) verwenden, welche wir im Rahmen dieser Vorlesung nicht studieren werden.
Als Illustration leiten wir im folgenden Beispiel eine geschlossene Formel für die Lösungen einer beliebigen nicht-ausgearteten homogenen autonomen linearen Rekursionsgleichung vom Grad 2 her. Nicht-ausgeartet bedeutet hierbei, dass unser Setup gerade so gewählt ist, dass die Transponierte der Begleitmatrix genau zwei verschiedene Eigenwerte besitzt und daher diagonalisierbar wird.

(6.75) Beispiel. Es seien $f \in K[X]$, $a_0, a_1 \in K$ mit $f = X^2 - a_1 X - a_0$ und Nullstellen b_1, b_2 von f mit $b_1 \neq b_2$ gegeben. Für $x \in K^{\mathbb{N}_0}$ gilt genau dann

$$x_{k+2} = a_0 x_k + a_1 x_{k+1}$$

für $k \in \mathbb{N}_0$, wenn

$$x_k = \frac{b_1^k b_2 - b_2^k b_1}{b_2 - b_1} x_0 + \frac{b_2^k - b_1^k}{b_2 - b_1} x_1$$

für $k \in \mathbb{N}_0$ ist. Der Lösungsraum der homogenen autonomen linearen Rekursionsgleichung zum charakteristischen Polynom f ist gegeben durch

$$\mathrm{SolRec}(f) = K(b_1^k)_{k \in \mathbb{N}_0} + K(b_2^k)_{k \in \mathbb{N}_0}.$$

Beweis. Die Eigenwerte von $\mathrm{C}(f)^{\mathrm{tr}}$ sind nach Korollar (6.14) gerade die Eigenwerte von $\mathrm{C}(f)$, nach Korollar (6.23) also gerade die Nullstellen b_1 und b_2 von f. Da $\mathrm{C}(f)^{\mathrm{tr}}$ wegen $b_1 \neq b_2$ somit 2 verschiedene Eigenwerte hat, ist $\mathrm{C}(f)^{\mathrm{tr}}$ nach Korollar (6.53) diagonalisierbar. Nach Bemerkung (6.4) ist der Eigenraum von $\mathrm{C}(f)^{\mathrm{tr}}$ zum Eigenwert b_i für $i \in \{1, 2\}$ gegeben durch

$$\mathrm{Eig}_{b_i}(\mathrm{C}(f)^{\mathrm{tr}}) = \mathrm{Sol}(\mathrm{C}(f)^{\mathrm{tr}} - b_i \mathrm{E}_2, 0) = \mathrm{Sol}(\begin{pmatrix} -b_i & 1 \\ a_0 & a_1 - b_i \end{pmatrix}, 0) = K \begin{pmatrix} 1 \\ b_i \end{pmatrix}.$$

Für $P \in \mathrm{GL}_2(K)$ gegeben durch

$$P = \begin{pmatrix} 1 & 1 \\ b_1 & b_2 \end{pmatrix}$$

gilt nach Bemerkung (6.45) folglich

$$P^{-1}\,\mathrm{C}(f)^{\mathrm{tr}}\,P = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}.$$

Nach Bemerkung (6.56) folgt

$$\begin{aligned}(\mathrm{C}(f)^{\mathrm{tr}})^k &= P \begin{pmatrix} b_1^k & 0 \\ 0 & b_2^k \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 1 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} b_1^k & 0 \\ 0 & b_2^k \end{pmatrix} \cdot \frac{1}{b_2 - b_1} \begin{pmatrix} b_2 & -1 \\ -b_1 & 1 \end{pmatrix} \\ &= \frac{1}{b_2 - b_1} \begin{pmatrix} 1 & 1 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} b_1^k b_2 & -b_1^k \\ -b_2^k b_1 & b_2^k \end{pmatrix} = \frac{1}{b_2 - b_1} \begin{pmatrix} b_1^k b_2 - b_2^k b_1 & b_2^k - b_1^k \\ b_1^{k+1} b_2 - b_2^{k+1} b_1 & b_2^{k+1} - b_1^{k+1} \end{pmatrix} \\ &= \frac{1}{b_2 - b_1} \begin{pmatrix} b_1^k b_2 - b_2^k b_1 & b_2^k - b_1^k \\ b_1^{k+1} b_2 - b_2^{k+1} b_1 & b_2^{k+1} - b_1^{k+1} \end{pmatrix}\end{aligned}$$

für $k \in \mathbb{N}_0$.
Nun sei $x \in K^{\mathbb{N}_0}$ gegeben. Nach Korollar (6.74) gilt genau dann $x \in \mathrm{SolRec}(f)$, wenn

$$x = (((\mathrm{C}(f)^{\mathrm{tr}})^k \begin{pmatrix} x_0 \\ x_1 \end{pmatrix})_1)_{k \in \mathbb{N}_0}$$

ist. Somit gilt genau dann $x_{k+2} = a_0 x_k + a_1 x_{k+1}$ für $k \in \mathbb{N}_0$, wenn

$$x_k = (\frac{1}{b_2 - b_1} \begin{pmatrix} b_1^k b_2 - b_2^k b_1 & b_2^k - b_1^k \\ b_1^{k+1} b_2 - b_2^{k+1} b_1 & b_2^{k+1} - b_1^{k+1} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix})_1 = \frac{b_1^k b_2 - b_2^k b_1}{b_2 - b_1} x_0 + \frac{b_2^k - b_1^k}{b_2 - b_1} x_1$$

für $k \in \mathbb{N}_0$ ist.
Nach Korollar (1.23)(b), (c) gilt insbesondere

$$\begin{aligned}\mathrm{SolRec}(f) &= \{(((\mathrm{C}(f)^{\mathrm{tr}})^k c)_1)_{k \in \mathbb{N}_0} \mid c \in K^{2 \times 1}\} = \{(\frac{b_1^k b_2 - b_2^k b_1}{b_2 - b_1} c_1 + \frac{b_2^k - b_1^k}{b_2 - b_1} c_2)_{k \in \mathbb{N}_0} \mid c \in K^{2 \times 1}\} \\ &= \{c_1 \frac{1}{b_2 - b_1} (b_1^k b_2 - b_2^k b_1)_{k \in \mathbb{N}_0} + c_2 \frac{1}{b_2 - b_1} (b_2^k - b_1^k)_{k \in \mathbb{N}_0} \mid c \in K^{2 \times 1}\} \\ &= \langle \frac{1}{b_2 - b_1} (b_1^k b_2 - b_2^k b_1)_{k \in \mathbb{N}_0}, \frac{1}{b_2 - b_1} (b_2^k - b_1^k)_{k \in \mathbb{N}_0} \rangle \\ &= \langle \frac{1}{b_2 - b_1} (b_1^k b_2 - b_2^k b_1)_{k \in \mathbb{N}_0} + b_1 \frac{1}{b_2 - b_1} (b_2^k - b_1^k)_{k \in \mathbb{N}_0}, \frac{1}{b_2 - b_1} (b_2^k - b_1^k)_{k \in \mathbb{N}_0} \rangle \\ &= \langle \frac{1}{b_2 - b_1} (b_1^k b_2 - b_1^{k+1})_{k \in \mathbb{N}_0}, \frac{1}{b_2 - b_1} (b_2^k - b_1^k)_{k \in \mathbb{N}_0} \rangle = \langle (b_1^k)_{k \in \mathbb{N}_0}, \frac{1}{b_2 - b_1} (b_2^k - b_1^k)_{k \in \mathbb{N}_0} \rangle \\ &= \langle (b_1^k)_{k \in \mathbb{N}_0}, (b_2^k - b_1^k)_{k \in \mathbb{N}_0} \rangle = \langle (b_1^k)_{k \in \mathbb{N}_0}, (b_2^k - b_1^k)_{k \in \mathbb{N}_0} + (b_1^k)_{k \in \mathbb{N}_0} \rangle = \langle (b_1^k)_{k \in \mathbb{N}_0}, (b_2^k)_{k \in \mathbb{N}_0} \rangle. \qquad \square\end{aligned}$$

Als Anwendung von Beispiel (6.75) können wir nun insbesondere eine geschlossene Formel für die Fibonacci-Folge bestimmen.

(6.76) Beispiel (Binets Formel für die Fibonacci-Folge)**.** Es sei $f \in \mathbb{R}^{\mathbb{N}_0}$ gegeben durch

$$f_k = \begin{cases} 0, & \text{für } k = 0, \\ 1, & \text{für } k = 1, \\ f_{k-2} + f_{k-1}, & \text{für } k \in \mathbb{N}_0 \text{ mit } k \geq 2. \end{cases}$$

Dann gilt

$$f_k = \frac{1}{\sqrt{5}}((\frac{1 + \sqrt{5}}{2})^k - (\frac{1 - \sqrt{5}}{2})^k)$$

für $k \in \mathbb{N}_0$.

Beweis. Die Nullstellen a_1 und a_2 von $X^2 - X - 1$ sind gegeben durch

$$a_1 = \frac{1+\sqrt{5}}{2},\ a_2 = \frac{1-\sqrt{5}}{2}.$$

Nach Beispiel (6.75) folgt

$$f_k = \frac{a_1^k a_2 - a_2^k a_1}{a_2 - a_1} f_0 + \frac{a_2^k - a_1^k}{a_2 - a_1} f_1 = \frac{a_2^k - a_1^k}{a_2 - a_1} = \frac{(\frac{1-\sqrt{5}}{2})^k - (\frac{1+\sqrt{5}}{2})^k}{\frac{1-\sqrt{5}}{2} - \frac{1+\sqrt{5}}{2}} = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^k - (\frac{1-\sqrt{5}}{2})^k)$$

für $k \in \mathbb{N}_0$. □

Das zunächst Erstaunliche an Binets Formel (6.76) ist die Verwendung der irrationalen Zahlen $\frac{1+\sqrt{5}}{2}$ und $\frac{1-\sqrt{5}}{2}$; es ist dieser Formel nicht anzusehen, dass die Fibonacci-Folge lediglich ganzzahlige Einträge besitzt.
Im Beweis zu Beispiel (6.75) haben wir die Basis

$$(\frac{1}{b_2 - b_1}(b_1^k b_2 - b_2^k b_1)_{k \in \mathbb{N}_0}, \frac{1}{b_2 - b_1}(b_2^k - b_1^k)_{k \in \mathbb{N}_0})$$

umgeformt, um die etwas „schönere“ Basis

$$((b_1^k)_{k \in \mathbb{N}_0}, (b_2^k)_{k \in \mathbb{N}_0})$$

zu erhalten. Dass dies möglich war, liegt wie bei allgemeinen (nicht-ausgearteten) homogenen autonomen linearen Rekursionsgleichungssystemen an der Diagonalisierbarkeit der Koeffizientenmatrix, in diesem Fall also an der Diagonalisierbarkeit von $\mathrm{C}(f)^{\mathrm{tr}}$, siehe Korollar (6.54).
Um ein solches Resultat für homogene autonome lineare Rekursionsgleichungen von beliebigem Grad zu erhalten, müssen wir die Eigenräume von $\mathrm{C}(f)^{\mathrm{tr}}$ studieren. Wir verwenden die in Definition (3.58) eingeführte Vandermondematrix:

(6.77) Bemerkung. Es sei ein normiertes $f \in K[X] \setminus \{0\}$ mit $\deg f \geq 1$ gegeben. Für jede Nullstelle a von f ist $\mathrm{V}(a)^{\mathrm{tr}}$ ein Eigenvektor von $\mathrm{C}(f)^{\mathrm{tr}}$ zum Eigenwert a.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.78) Korollar. Es seien $n \in \mathbb{N}$ und ein normiertes $f \in K[X] \setminus \{0\}$ mit $\deg f = n$ gegeben. Für jede Nullstelle a von f ist

$$(a^k)_{k \in \mathbb{N}_0}$$

eine Lösung der homogenen autonomen linearen Rekursionsgleichung zum charakteristischen Polynom f.

Beweis. Es sei eine Nullstelle a von f gegeben. Nach Bemerkung (6.77) ist dann $\mathrm{V}(a)^{\mathrm{tr}}$ ein Eigenvektor von $\mathrm{C}(f)^{\mathrm{tr}}$ zum Eigenwert a. Folglich ist $(a^k\,\mathrm{V}(a)^{\mathrm{tr}})_{k \in \mathbb{N}_0} \in \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}})$ nach Bemerkung (6.71) und damit

$$(a^k)_{k \in \mathbb{N}_0} = ((a^k\,\mathrm{V}(a)^{\mathrm{tr}})_1)_{k \in \mathbb{N}_0} \in \mathrm{SolRec}(f)$$

nach Lemma (6.73). □

(6.79) Korollar. Es seien $n \in \mathbb{N}$, ein normiertes $f \in K[X] \setminus \{0\}$ mit $\deg f = n$ und ein n-Tupel $(a_1, \ldots, a_n)$ in K von verschiedenen Nullstellen von f gegeben. Dann ist

$$s := ((a_1^k)_{k \in \mathbb{N}_0}, \ldots, (a_n^k)_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(f)$ und für $x \in \mathrm{SolRec}(f)$ ist die Koordinatenspalte $\kappa_s(x)$ von x zur Basis s gerade die eindeutige Lösung des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix

$$\left(\begin{array}{c|c} & x_0 \\ \mathrm{V}_{<n}(a_1, \ldots, a_n)^{\mathrm{tr}} & \vdots \\ & x_{n-1} \end{array}\right) = \left(\begin{array}{ccc|c} a_1^0 & \ldots & a_n^0 & x_0 \\ \vdots & & \vdots & \vdots \\ a_1^{n-1} & \ldots & a_n^{n-1} & x_{n-1} \end{array}\right).$$

Beweis. Nach Bemerkung (6.77) ist $\mathrm{V}(a_i)^{\mathrm{tr}}$ für $i \in [1, n]$ ein Eigenvektor von $\mathrm{C}(f)^{\mathrm{tr}}$ zum Eigenwert a_i. Da $(a_1, \dots, a_n)$ ein n-Tupel aus verschiedenen Nullstellen von f ist, ist $(\mathrm{V}(a_1)^{\mathrm{tr}}, \dots, \mathrm{V}(a_n)^{\mathrm{tr}})$ nach Korollar (6.7) linear unabhängig und damit nach Beispiel (b) und Korollar (1.67) eine Basis von $K^{n \times 1}$. Nach Bemerkung (6.72) ist

$$t := ((a_1^k \mathrm{V}(a_1)^{\mathrm{tr}})_{k \in \mathbb{N}_0}, \dots (a_n^k \mathrm{V}(a_n)^{\mathrm{tr}})_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}})$ und für $y \in \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}})$ gilt

$$\mathrm{Sol}(\mathrm{V}(a_1, \dots, a_n)^{\mathrm{tr}}, y_0) = \mathrm{Sol}(\begin{pmatrix} \mathrm{V}(a_1)^{\mathrm{tr}} & \dots & \mathrm{V}(a_n)^{\mathrm{tr}} \end{pmatrix}, y_0) = \{\kappa_t(y)\}.$$

Da

$$\mathrm{SolRec}(f) \to \mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}),\ x \mapsto (\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0},$$
$$\mathrm{SolRec}(\mathrm{C}(f)^{\mathrm{tr}}) \to \mathrm{SolRec}(f),\ y \mapsto (y_{k,1})_{k \in \mathbb{N}_0}$$

nach Lemma (6.73) wohldefinierte, sich gegenseitig invertierende Isomorphismen sind, ist

$$s = ((a_1^k)_{k \in \mathbb{N}_0}, \dots, (a_n^k)_{k \in \mathbb{N}_0}) = (((a_1^k \mathrm{V}(a_1)^{\mathrm{tr}})_1)_{k \in \mathbb{N}_0}, \dots ((a_n^k \mathrm{V}(a_n)^{\mathrm{tr}})_1)_{k \in \mathbb{N}_0})$$

nach Proposition (2.21) eine Basis von $\mathrm{SolRec}(f)$ und für $x \in \mathrm{SolRec}(f)$ gilt

$$\mathrm{Sol}(\mathrm{V}(a_1, \dots, a_n)^{\mathrm{tr}}, \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix}) = \{\kappa_t((\begin{pmatrix} x_k \\ \vdots \\ x_{k+n-1} \end{pmatrix})_{k \in \mathbb{N}_0})\} = \{\kappa_s(x)\}.$$ □

Alternativer Beweis von Beispiel (6.75). Nach Korollar (6.79) ist

$$s := ((b_1^k)_{k \in \mathbb{N}_0}, (b_2^k)_{k \in \mathbb{N}_0})$$

eine Basis von $\mathrm{SolRec}(f)$, so dass insbesondere

$$\mathrm{SolRec}(f) = K(b_1^k)_{k \in \mathbb{N}_0} + K(b_2^k)_{k \in \mathbb{N}_0}$$

gilt. Für $x \in \mathrm{SolRec}(f)$ ist ferner $\mathrm{Sol}(\mathrm{V}(b_1, b_2)^{\mathrm{tr}}, \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}) = \{\kappa_s(x)\}$, also

$$\kappa_s(x) = \begin{pmatrix} 1 & 1 \\ b_1 & b_2 \end{pmatrix}^{-1} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \frac{1}{b_2 - b_1} \begin{pmatrix} b_2 & -1 \\ -b_1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \frac{1}{b_2 - b_1} \begin{pmatrix} b_2 x_0 - x_1 \\ -b_1 x_0 + x_1 \end{pmatrix}.$$

Folglich gilt für $x \in K^{\mathbb{N}_0}$ genau dann $x_{k+2} = a_0 x_k + a_1 x_{k+1}$ für $k \in \mathbb{N}_0$, wenn

$$x = \frac{b_2 x_0 - x_1}{b_2 - b_1} (b_1^k)_{k \in \mathbb{N}_0} + \frac{-b_1 x_0 + x_1}{b_2 - b_1} (b_2^k)_{k \in \mathbb{N}_0} = (\frac{b_2 x_0 - x_1}{b_2 - b_1} b_1^k + \frac{-b_1 x_0 + x_1}{b_2 - b_1} b_2^k)_{k \in \mathbb{N}_0}$$

ist, d.h. wenn

$$x_k = \frac{b_2 x_0 - x_1}{b_2 - b_1} b_1^k + \frac{-b_1 x_0 + x_1}{b_2 - b_1} b_2^k = \frac{b_1^k b_2 - b_2^k b_1}{b_2 - b_1} x_0 + \frac{b_2^k - b_1^k}{b_2 - b_1} x_1$$

für $k \in \mathbb{N}_0$ gilt. □

Invariante Untervektorräume

Zum Abschluss dieses Abschnitts wollen wir den Satz von Cayley/Hamilton (6.84) herleiten. Hierzu benötigen wir den nachfolgenden Begriff:

(6.80) Definition (invarianter Untervektorraum)**.**

(a) Es sei ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Ein K-Untervektorraum U von V heißt *invariant* unter φ (oder *φ-invariant*), wenn

$$\varphi(U) \subseteq U$$

gilt.

(b) Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Ein K-Untervektorraum U von $K^{n \times 1}$ heißt *invariant* unter A (oder *A-invariant*), wenn U invariant unter φ_A ist.

(6.81) Beispiel. Es seien $a_1, a_2, a_3, a_4, a_5 \in K$ und $A \in K^{3 \times 3}$ mit

$$A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & a_3 \\ 0 & a_4 & a_5 \end{pmatrix}$$

gegeben. Dann sind $\langle \mathrm{e}_1 \rangle$ und $\langle \mathrm{e}_2, \mathrm{e}_3 \rangle$ invariant unter A.

Beweis. Es ist

$$A\mathrm{e}_1 = a_1 \mathrm{e}_1 \in \langle \mathrm{e}_1 \rangle$$

und damit $A(b_1 \mathrm{e}_1) = b_1(A\mathrm{e}_1) \in \langle \mathrm{e}_1 \rangle$ für alle $b_1 \in K$. Ferner ist

$$\begin{aligned} A\mathrm{e}_2 &= a_2 \mathrm{e}_2 + a_4 \mathrm{e}_3 \in \langle \mathrm{e}_2, \mathrm{e}_3 \rangle, \\ A\mathrm{e}_3 &= a_3 \mathrm{e}_2 + a_5 \mathrm{e}_3 \in \langle \mathrm{e}_2, \mathrm{e}_3 \rangle \end{aligned}$$

und damit $A(b_2 \mathrm{e}_2 + b_3 \mathrm{e}_3) = b_2(A\mathrm{e}_2) + b_3(A\mathrm{e}_3)$ für alle $b_2, b_3 \in K$. Folglich sind $\langle \mathrm{e}_1 \rangle$ und $\langle \mathrm{e}_2, \mathrm{e}_3 \rangle$ invariant unter A. □

Für einen Endomorphismus $\varphi\colon V \to V$ und einen Untervektorraum U von V lässt sich stets die Einschränkung $\varphi|_U\colon U \to V$, $u \mapsto \varphi(u)$ bilden. Ist U zusätzlich invariant unter φ, so können wir auch den Zielvektorraum einschränken, d.h. die Abbildung $\varphi|_U^U\colon U \to U$, $u \mapsto \varphi(u)$ ist wohldefiniert.

(6.82) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, ein K-Untervektorraum U von V, $m, n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V derart gegeben, dass $s' = (s_1, \dots, s_m)$ eine Basis von U ist. Genau dann ist U invariant unter φ, wenn es $A \in K^{m \times m}$, $B \in K^{m \times (n-m)}$, $C \in K^{(n-m) \times (n-m)}$ mit

$$\mathrm{M}_{s,s}(\varphi) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

gibt. In diesem Fall gilt

$$\mathrm{M}_{s',s'}(\varphi|_U^U) = A.$$

Beweis. Zunächst sei $U = \langle s_1, \dots, s_m \rangle$ invariant unter φ. Dann gilt insbesondere $\varphi(s_j) \in U = \langle s_1, \dots, s_m \rangle$ und damit $\kappa_s(\varphi(s_j)) \in \langle \mathrm{e}_1, \dots, \mathrm{e}_m \rangle$ für $j \in [1, m]$. Folglich gibt es $A \in K^{m \times m}$, $B \in K^{m \times (n-m)}$, $C \in K^{(n-m) \times (n-m)}$ mit

$$\mathrm{M}_{s,s}(\varphi) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}.$$

Ferner gilt

$$A = \begin{pmatrix} \kappa_{s'}(\varphi(s_1)) & \dots & \kappa_{s'}(\varphi(s_m)) \end{pmatrix} = \begin{pmatrix} \kappa_{s'}(\varphi|_U^U(s_1)) & \dots & \kappa_{s'}(\varphi|_U^U(s_m)) \end{pmatrix} = \mathrm{M}_{s',s'}(\varphi|_U^U).$$

Umgekehrt sei nun

$$\mathrm{M}_{s,s}(\varphi) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

für gewisse $A \in K^{m \times m}$, $B \in K^{m \times (n-m)}$, $C \in K^{(n-m) \times (n-m)}$. Für $j \in [1, m]$ ist dann $\kappa_s(\varphi(s_j)) \in \langle e_1, \ldots, e_m \rangle$ und damit $\varphi(s_j) \in \langle s_1, \ldots, s_m \rangle = U$. Dies impliziert aber bereits

$$\varphi(\sum_{j \in [1,m]} a_j s_j) = \sum_{j \in [1,m]} a_j \varphi(s_j) \in U$$

für alle $a \in K^m$ und damit $\varphi(U) \subseteq U$. Folglich ist U invariant unter φ. □

(6.83) Korollar. Es seien ein endlichdimensionaler K-Vektorraum V, ein K-Vektorraumendomorphismus $\varphi \colon V \to V$ und ein φ-invarianter K-Untervektorraum U von V gegeben. Dann gilt

$$\chi_{\varphi|_U^U} \mid \chi_\varphi.$$

Beweis. Nach Korollar (1.52) und Korollar (1.55) gibt es eine Basis $s = (s_1, \ldots, s_n)$ von V derart, dass $s' = (s_1, \ldots, s_m)$ eine Basis von U ist. Da U invariant unter φ ist, gilt

$$\mathrm{M}_{s,s}(\varphi) = \begin{pmatrix} \mathrm{M}_{s',s'}(\varphi|_U^U) & B \\ 0 & C \end{pmatrix}$$

für gewisse $B \in K^{m \times (n-m)}$, $C \in K^{(n-m) \times (n-m)}$ nach Bemerkung (6.82). Nach dem Kästchensatz für Determinanten (5.15) folgt

$$\begin{aligned}
\chi_\varphi &= \chi_{\mathrm{M}_{s,s}(\varphi)} = \det(X\mathrm{E}_n - \mathrm{M}_{s,s}(\varphi)) = \det(X \begin{pmatrix} \mathrm{E}_m & 0 \\ 0 & \mathrm{E}_{n-m} \end{pmatrix} - \begin{pmatrix} \mathrm{M}_{s',s'}(\varphi|_U^U) & B \\ 0 & C \end{pmatrix}) \\
&= \det \begin{pmatrix} X\mathrm{E}_m - \mathrm{M}_{s',s'}(\varphi|_U^U) & -B \\ 0 & X\mathrm{E}_{n-m} - C \end{pmatrix} = \det(X\mathrm{E}_m - \mathrm{M}_{s',s'}(\varphi|_U^U)) \det(X\mathrm{E}_{n-m} - C) \\
&= \chi_{\mathrm{M}_{s',s'}(\varphi|_U^U)} \chi_C = \chi_{\varphi|_U^U} \chi_C
\end{aligned}$$

und damit $\chi_{\varphi|_U^U} \mid \chi_\varphi$. □

Alternativer Beweis von Proposition (6.19). Es lässt sich zeigen, dass $\mathrm{Eig}_a(\varphi)$ ein φ-invarianter K-Untervektorraum von V ist; der Beweis sei dem Leser zur Übung überlassen. Die Einschränkung

$$\varphi|_{\mathrm{Eig}_a(\varphi)}^{\mathrm{Eig}_a(\varphi)} \colon \mathrm{Eig}_a(\varphi) \to \mathrm{Eig}_a(\varphi)$$

ist gegeben durch

$$(\varphi|_{\mathrm{Eig}_a(\varphi)}^{\mathrm{Eig}_a(\varphi)})(v) = \varphi(v) = av$$

für $v \in \mathrm{Eig}_a(\varphi)$, d.h. es ist $\varphi|_{\mathrm{Eig}_a(\varphi)}^{\mathrm{Eig}_a(\varphi)} = a \, \mathrm{id}_{\mathrm{Eig}_a(\varphi)}$. Nach Korollar (6.83) gilt

$$\chi_{\varphi|_{\mathrm{Eig}_a(\varphi)}^{\mathrm{Eig}_a(\varphi)}} \mid \chi_\varphi,$$

es folgt also

$$\mathrm{m}_a(\varphi) = \mathrm{m}_a(\chi_\varphi) \geq \mathrm{m}_a(\chi_{\varphi|_{\mathrm{Eig}_a(\varphi)}^{\mathrm{Eig}_a(\varphi)}}) = \mathrm{m}_a(a \, \mathrm{id}_{\mathrm{Eig}_a(\varphi)}) = \mathrm{m}_a((X - a)^{\mathrm{g}_a(\varphi)}) = \mathrm{g}_a(\varphi).$$ □

Der Satz von Cayley/Hamilton

Bei der Formulierung des Satzes von Cayley/Hamilton werden in das charakteristische Polynom Endomorphismen und Matrizen eingesetzt, siehe Notation (6.55).

(6.84) Satz (Satz von Cayley/Hamilton)**.**

(a) Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi \colon V \to V$ gegeben. Dann ist

$$\chi_\varphi(\varphi) = 0.$$

(b) Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Dann ist

$$\chi_A(A) = 0.$$

Beweis.

(a) Es sei $v \in V$ beliebig gegeben. Dann ist $(v, \varphi(v), \dots, \varphi^{\dim V}(v))$ linear abhängig in V. Es sei

$$m := \min \{n \in \mathbb{N}_0 \mid (v, \varphi(v), \dots, \varphi^n(v)) \text{ ist linear abhängig}\}$$

und es sei $U := \langle v, \varphi(v), \dots, \varphi^{m-1}(v) \rangle$. Dann ist $s := (v, \varphi(v), \dots, \varphi^{m-1}(v))$ linear unabhängig in V und damit eine Basis von U. Wegen der linearen Unabhängigkeit von $(v, \varphi(v), \dots, \varphi^{m-1}(v))$ und der linearen Abhängigkeit von $(v, \varphi(v), \dots, \varphi^m(v))$ ist ferner $\varphi^m(v) \in \langle v, \varphi(v), \dots, \varphi^{m-1}(v) \rangle = U$ nach Proposition (1.37) und damit U ein φ-invarianter Untervektorraum von V.

Nun sei $a \in K^{[0,m-1]}$ mit $\varphi^m(v) = \sum_{i \in [0,m-1]} a_i \varphi^i(v)$ gegeben und es sei $\mu \in K[X]$ gegeben durch $\mu = X^m - \sum_{i \in [0,m-1]} a_i X^i$, so dass

$$(\mu(\varphi))(v) = (\varphi^m - \sum_{i \in [0,m-1]} a_i \varphi^i)(v) = \varphi^m(v) - \sum_{i \in [0,m-1]} a_i \varphi^i(v) = 0$$

gilt. Dann ist

$$\mathrm{M}_{s,s}(\varphi|_U^U) = \begin{pmatrix} 0 & & 0 & a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & a_{n-2} \\ 0 & & 1 & a_{n-1} \end{pmatrix} = \mathrm{C}(\mu)$$

und somit

$$\chi_{\varphi|_U^U} = \chi_{\mathrm{M}_{s,s}(\varphi|_U^U)} = \chi_{\mathrm{C}(\mu)} = \mu$$

nach Proposition (6.22). Folglich gilt $\mu = \chi_{\varphi|_U^U} \mid \chi_\varphi$ nach Korollar (6.83), d.h. es gibt ein $q \in K[X]$ mit $\chi_\varphi = q\mu$. Wir erhalten $\chi_\varphi(\varphi) = (q\mu)(\varphi) = q(\varphi) \circ \mu(\varphi)$ und da $q(\varphi)$ ein Endomorphismus von V ist, folgt

$$(\chi_\varphi(\varphi))(v) = (q(\varphi) \circ \mu(\varphi))(v) = (q(\varphi))((\mu(\varphi))(v)) = (q(\varphi))(0) = 0.$$

Da $v \in V$ beliebig war, gilt also $\chi_\varphi(\varphi) = 0$ nach dem Gleichheitskriterium für Abbildungen (A.35).

(b) Es ist $\chi_A = \chi_{\mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)} = \chi_{\varphi_A}$, also $\chi_A(\varphi_A) = \chi_{\varphi_A}(\varphi_A) = 0$ nach (a) und damit

$$\chi_A(A) = \chi_A(\mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)) = \mathrm{M}_{\mathrm{e},\mathrm{e}}(\chi_A(\varphi_A)) = \mathrm{M}_{\mathrm{e},\mathrm{e}}(0) = 0.$$ □

Zusätzliche Konzepte

Ein weiterer Zugang zur Eigenwerttheorie lässt sich mit Hilfe des Konzepts des Minimalpolynoms begehen, siehe Definition (B.33). Wir fassen im Folgenden die hierfür relevanten Ergebnisse aus Anhang B zusammen:

Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Nach Beispiel (B.4)(b), Proposition (3.14) und Satz (2.32) ist $\mathrm{End}_K(V)$ eine endlichdimensionale K-Algebra. Folglich gibt es nach Korollar (B.36) und Korollar (B.32) genau ein normiertes $\mu \in K[X] \setminus \{0\}$ derart, dass $\mu(\varphi) = 0$ ist und dass für $f \in K[X]$ mit $f(\varphi) = 0$ stets $\mu \mid f$ folgt. Das Polynom μ wird *Minimalpolynom* von φ (in $\mathrm{End}_K(V)$ über K) genannt und als $\mu_\varphi = \mu$ notiert. Nach Proposition (B.31) ist

$$\deg \mu_\varphi = \min \{\deg f \mid f \in K[X] \setminus \{0\} \text{ mit } f(\varphi) = 0\}.$$

Analog für Matrizen: Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Nach Beispiel (B.4)(c) und Beispiel (1.65)(b) ist $K^{n \times n}$ eine endlichdimensionale K-Algebra. Folglich gibt es Korollar (B.36) und Korollar (B.32) genau ein normiertes $\mu \in K[X] \setminus \{0\}$ derart, dass $\mu(A) = 0$ ist und dass für $f \in K[X]$ mit $f(A) = 0$ stets $\mu \mid f$ folgt. Das Polynom μ wird *Minimalpolynom* von A (in $K^{n \times n}$ über K) genannt und als $\mu_A = \mu$ notiert. Nach Proposition (B.31) ist

$$\deg \mu_A = \min \{\deg f \mid f \in K[X] \setminus \{0\} \text{ mit } f(A) = 0\}.$$

Eine Spezialisierung von Korollar (B.36) führt zu einer naiven Berechnungsmethode für das Minimalpolynom:

(6.85) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Ferner sei $m \in \mathbb{N}_0$ so gegeben, dass $(\mathrm{E}_n, A, \dots, A^{m-1})$ linear unabhängig ist, und es sei $a \in K^{[0,m-1]}$ mit

$$A^m = \sum_{i \in [0,m-1]} a_i A^i$$

gegeben. Dann ist

$$\mu_A = X^m - \sum_{i \in [0,m-1]} a_i X^i.$$

(6.86) Beispiel.

(a) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Das charakteristische Polynom von A ist

$$\mu_A = X^2 - 4X + 3 = (X-1)(X-3).$$

(b) Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

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

Das Minimalpolynom von A ist

$$\mu_A = X^2.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Es gibt eine Methode zur Berechnung des Minimalpolynoms, welche konzeptionell etwas komplizierter ist, bei welcher sich jedoch insbesondere für größere Matrizen der Rechenaufwand erheblich verringern lässt. Diese Methode verwendet die Technik aus dem Beweis des Satzes von Cayley/Hamilton (6.84), wo eine „lokale“ Variante des Minimalpolynoms studiert wird.
Für das Minimalpolynom gelten die üblichen Zusammenhänge zwischen Endomorphismen und Matrizen:

(6.87) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Dann ist

$$\mu_A = \mu_{\varphi_A}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.88) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Dann ist

$$\mu_\varphi = \mu_{\mathrm{M}_{s,s}(\varphi)}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Nach Korollar (B.37), Proposition (3.14) und Satz (2.32) gilt

$$\deg \mu_\varphi \leq \dim_K \operatorname{End}_K(V) = (\dim_K V)^2.$$

Mit Hilfe des Satzes von Cayley/Hamilton (6.84)(a) können wir diese Abschätzung sogar noch verschärfen:

(6.89) Bemerkung. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Dann ist

$$\deg \mu_\varphi \leq \dim_K V.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Das Minimalpolynom einer Begleitmatrix lässt sich wie das charakteristische Polynom ablesen:

(6.90) Proposition. Es sei ein normiertes $f \in K[X] \setminus \{0\}$ gegeben. Das Minimalpolynom von $\mathrm{C}(f)$ ist gegeben durch

$$\mu_{\mathrm{C}(f)} = f.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Die Bedeutung des Minimalpolynoms für die Eigenwerttheorie liegt darin begründet, dass sich auch mit Hilfe des Minimalpolynoms Eigenwerte bestimmen sowie Diagonalisierbarkeit testen lassen:

(6.91) Proposition. Es seien ein endlichdimensionaler K-Vektorraum V, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ und $a \in K$ gegeben. Genau dann ist a ein Eigenwert von φ, wenn a eine Nullstelle des Minimalpolynoms von φ ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.92) Satz. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Genau dann ist φ diagonalisierbar, wenn das Minimalpolynom μ_φ in verschiedene Linearfaktoren zerfällt.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Eine Abschwächung des Diagonalisierbarkeitskonzepts ist das Konzept der Trigonalisierbarkeit, welches wir nun betrachten wollen:

(6.93) Beispiel. Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

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

und es sei $s = (s_1, s_2)$ gegeben durch

$$s = (\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}).$$

Dann ist s eine Basis von $\mathbb{Q}^{2 \times 1}$ derart, dass $\langle s_1 \rangle$ und $\langle s_1, s_2 \rangle$ invariant unter A sind.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.94) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Genau dann ist für jedes $i \in [1, n]$ die lineare Hülle $\langle s_1, \dots, s_i \rangle$ invariant unter φ, wenn $\mathrm{M}_{s,s}(\varphi)$ eine obere Dreiecksmatrix ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.95) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$, eine Basis $s = (s_1, \dots, s_n)$ von V und ein n-Tupel $s' = (s'_1, \dots, s'_n)$ in V gegeben. Für $i \in [1, n]$ gilt: Genau dann ist $\langle s'_1, \dots, s'_i \rangle$ invariant unter φ, wenn $(\kappa_s(s'_1), \dots, \kappa_s(s'_i))$ invariant unter $\mathrm{M}_{s,s}(\varphi)$ ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.96) Bemerkung. Es seien $n \in \mathbb{N}_0$, $A \in K^{n \times n}$, ein n-Tupel $(a_1, \dots, a_n)$ in K und eine Basis $s = (s_1, \dots, s_n)$ von $K^{n \times 1}$ gegeben und es sei $P \in \mathrm{GL}_n(K)$ gegeben durch

$$P = \begin{pmatrix} s_1 & \dots & s_n \end{pmatrix}.$$

Genau dann ist für jedes $i \in [1, n]$ die lineare Hülle $\langle s_1, \dots, s_i \rangle$ invariant unter φ, wenn $P^{-1}AP$ eine obere Dreiecksmatrix ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.97) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Eigenbasis $s = (s_1, \dots, s_n)$ von V gegeben. Dann ist für jedes $i \in [1, n]$ die lineare Hülle $\langle s_1, \dots, s_i \rangle$ invariant unter φ.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.98) Definition (trigonalisierbar)**.**

(a) Es sei ein endlichdimensionaler K-Vektorraum V gegeben. Ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ heißt *trigonalisierbar* (oder *triangulierbar*), falls es ein $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V so gibt, dass für jedes $i \in [1, n]$ die lineare Hülle $\langle s_1, \dots, s_i \rangle$ invariant unter φ ist.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Ein $A \in K^{n \times n}$ heißt *trigonalisierbar* (oder *triangulierbar*) (über K), falls $\varphi_A\colon K^{n \times 1} \to K^{n \times 1}$ trigonalisierbar ist.

(6.99) Beispiel. Es sei $A \in \mathbb{Q}^{2 \times 2}$ gegeben durch

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

Dann ist A trigonalisierbar.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.100) Bemerkung. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Genau dann ist φ trigonalisierbar, wenn es eine Basis s von V derart gibt, dass $\mathrm{M}_{s,s}(\varphi)$ eine obere Dreiecksmatrix ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.101) Bemerkung. Es seien ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Basis $s = (s_1, \dots, s_n)$ von V gegeben. Genau dann ist φ trigonalisierbar, wenn $\mathrm{M}_{s,s}(\varphi)$ trigonalisierbar ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.102) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Die Matrix A ist trigonalisierbar.

(b) Es gibt eine Basis $s = (s_1, \dots, s_n)$ von V so, dass für jedes $i \in [1, n]$ die lineare Hülle $\langle s_1, \dots, s_i \rangle$ invariant unter A ist.

(c) Die Matrix A ist ähnlich zu einer oberen Dreiecksmatrix.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Wir geben eine Charakterisierung der Trigonalisierbarkeit. Man vergleiche die Charakterisierung der Diagonalisierbarkeit in Satz (6.52).

(6.103) Satz. Es seien ein endlichdimensionaler K-Vektorraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Genau dann ist φ trigonalisierbar, wenn das charakteristische Polynom χ_φ in Linearfaktoren zerfällt.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(6.104) Beispiel. Es sei $A \in \mathbb{R}^{2 \times 2}$ gegeben durch

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

Dann ist A nicht trigonalisierbar.

Beweis. Dies sei dem Leser zur Übung überlassen. □

7 Skalarprodukträume

Zum Abschluss studieren wir Skalarprodukte auf Vektorräumen über dem Körper der reellen Zahlen $\mathbb{R}$ und dem Körper der komplexen Zahlen $\mathbb{C}$. Mit Hilfe von Skalarprodukten lassen sich Längen definieren, siehe Definition (7.8), aber auch Orthogonalität, siehe Definition (7.21).
Im Folgenden bezeichnen wir bis zum Ende des Abschnitts mit K entweder den Körper der reellen Zahlen $\mathbb{R}$ oder den Körper der komplexen Zahlen $\mathbb{C}$.

Begriffsbildung

Wir beginnen mit der axiomatischen Definition eines Skalarprodukts, genauer gesagt mit der einer Struktur bestehend aus einem Vektorraum und einem Skalarprodukt.

(7.1) Definition (Skalarproduktraum)**.** Ein *K-Skalarproduktraum* (oder *K-Vektorraum mit Skalarprodukt*) besteht aus einem K-Vektorraum V zusammen mit einer Abbildung $\langle -, = \rangle \colon V \times V \to K$, $(v, w) \mapsto \langle v, w \rangle$, genannt *Skalarprodukt*, so, dass $\langle -, = \rangle$ eine *positiv definite, hermitesche Sesquilinearform* ist, d.h. derart, dass folgende Axiome gelten.

- *Sesquilinearität.* Für $v \in V$ ist $\langle v, - \rangle \colon V \to K$, $w \mapsto \langle v, w \rangle$ ein K-Vektorraumhomomorphismus. Für $w \in V$ ist $\langle -, w \rangle \colon V \to K$, $v \mapsto \langle v, w \rangle$ eine *K-semilineare Abbildung*, d.h. ein Homomorphismus abelscher Gruppen und für $a \in K$, $v \in V$ ist $\langle av, w \rangle = \overline{a} \langle v, w \rangle$.
- *Hermitizität.* Für $v, w \in V$ gilt $\langle w, v \rangle = \overline{\langle v, w \rangle}$.
- *Positive Definitheit.* Für $v \in V$ gilt $\langle v, v \rangle \geq 0$ und genau dann $\langle v, v \rangle = 0$, wenn $v = 0$ ist.

Explizit bedeutet die Sesquilinearität des Skalarprodukts eines K-Skalarproduktraums V:

- *Linearität in der zweiten Komponente.* Für $v, w, w' \in V$ ist $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$. Für $a \in K$, $v, w \in V$ ist $\langle v, aw \rangle = a \langle v, w \rangle$.
- *Semilinearität in der ersten Komponente.* Für $v, v', w \in V$ ist $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$. Für $a \in K$, $v, w \in V$ ist $\langle av, w \rangle = \overline{a} \langle v, w \rangle$.

(7.2) Beispiel. Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Es wird K^n ein Skalarproduktraum, wobei das Skalarprodukt durch

$$\langle x, y \rangle = \sum_{i \in [1,n]} \overline{x_i} y_i$$

für $x, y \in K^n$ gegeben ist.

(b) Es wird $K^{n \times 1}$ ein Skalarproduktraum, wobei das Skalarprodukt durch

$$\langle x, y \rangle = x^{\mathrm{ad}} y = \sum_{i \in [1,n]} \overline{x_i} y_i$$

für $x, y \in K^{n \times 1}$ gegeben ist.

In Beispiel (7.2)(b) haben wir $K^{1 \times 1}$ mit K identifiziert.

(7.3) Konvention. Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Sofern nicht anders erwähnt, betrachten wir ab jetzt K^n als Skalarproduktraum wie in Beispiel (7.2)(a). Wir nennen $\langle -, = \rangle$ in diesem Fall auch das *Standardskalarprodukt* auf K^n.

(b) Sofern nicht anders erwähnt, betrachten wir ab jetzt $K^{n \times 1}$ als Skalarproduktraum wie in Beispiel (7.2)(b). Wir nennen $\langle -, = \rangle$ in diesem Fall auch das *Standardskalarprodukt* auf $K^{n \times 1}$.

(7.4) Definition (euklidischer Vektorraum, unitärer Vektorraum)**.**

(a) Ein *euklidischer Vektorraum* ist ein endlichdimensionaler $\mathbb{R}$-Skalarproduktraum.

(b) Ein *unitärer Vektorraum* ist ein endlichdimensionaler $\mathbb{C}$-Skalarproduktraum.

(7.5) Beispiel. In $K^{2\times 1}$ gilt

$$\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \end{pmatrix} \rangle = -1.$$

Beweis. Es ist

$$\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \end{pmatrix} \rangle = 2 \cdot 1 + 1 \cdot (-3) = -1.$$

□

(7.6) Lemma (Cauchy-Schwarz-Ungleichung)**.** Es seien ein K-Skalarproduktraum und $v, w \in V$ gegeben. Dann gilt

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle \langle w, w\rangle.$$

Genau dann gilt

$$|\langle v, w\rangle|^2 = \langle v, v\rangle \langle w, w\rangle,$$

wenn (v, w) linear abhängig in V ist.

Beweis. Zunächst sei (v, w) linear abhängig in V, so dass es ein $x \in V$ und $a, b \in K$ mit $v = ax$ und $w = bx$ gibt. Dann gilt

$$\begin{aligned} |\langle v, w\rangle|^2 &= |\langle ax, bx\rangle|^2 = |\bar{a}b\langle x, x\rangle|^2 = |\bar{a}|^2\,|b|^2\,|\langle x, x\rangle|^2 = \bar{a}ab\bar{b}\langle x, x\rangle^2 = \bar{a}a\langle x, x\rangle\bar{b}b\langle x, x\rangle = \langle ax, ax\rangle\langle bx, bx\rangle \\ &= \langle v, v\rangle\langle w, w\rangle \end{aligned}$$

und damit $|\langle v, w\rangle|^2 = \langle v, v\rangle\langle w, w\rangle$.
Im Folgenden sei (v, w) linear unabhängig in V. Dann ist insbesondere $v \neq 0$, wegen der positiven Definitheit des Skalarprodukts also $\langle v, v\rangle > 0$. Ferner ist $w - \frac{\langle v, w\rangle}{\langle v, v\rangle} v \neq 0$, wegen der positiven Definitheit des Skalarprodukts also

$$\begin{aligned} 0 &< \langle w - \frac{\langle v, w\rangle}{\langle v, v\rangle} v, w - \frac{\langle v, w\rangle}{\langle v, v\rangle} v\rangle = \langle w, w\rangle - \frac{\langle v, w\rangle}{\langle v, v\rangle}\langle w, v\rangle - \frac{\overline{\langle v, w\rangle}}{\langle v, v\rangle}\langle v, w\rangle + \frac{\overline{\langle v, w\rangle}}{\langle v, v\rangle}\frac{\langle v, w\rangle}{\langle v, v\rangle}\langle v, v\rangle \\ &= \langle w, w\rangle - \frac{\langle v, w\rangle}{\langle v, v\rangle}\overline{\langle v, w\rangle} = \langle w, w\rangle - \frac{|\langle v, w\rangle|^2}{\langle v, v\rangle} \end{aligned}$$

und damit $|\langle v, w\rangle|^2 < \langle v, v\rangle\langle w, w\rangle$. □

Mit nachfolgendem Resultat ist es möglich, eine Gleichheit von Vektoren in einem Skalarproduktraum mit Hilfe des Skalarproduktes zu überprüfen. Diese Eigenschaft, welche auf der positiven Definitheit des Skalarprodukts beruht, wird *Nichtausartung* des Skalarprodukts genannt.

(7.7) Lemma. Es sei ein K-Skalarproduktraum V gegeben.

(a) Für $v, v' \in V$ gilt: Genau dann gilt für $w \in V$ stets $\langle v, w\rangle = \langle v', w\rangle$, wenn $v = v'$ ist.

(b) Für $w, w' \in V$ gilt: Genau dann gilt für $v \in V$ stets $\langle v, w\rangle = \langle v, w'\rangle$, wenn $w = w'$ ist.

Beweis.

(a) Es seien $v, v' \in V$ gegeben. Zunächst gelte $\langle v, w\rangle = \langle v', w\rangle$ für $w \in V$. Dann ist

$$0 = \langle v, w\rangle - \langle v', w\rangle = \langle v - v', w\rangle$$

für $w \in W$, also insbesondere $\langle v - v', v - v'\rangle = 0$. Mit der positiven Definitheit des Skalarprodukts folgt $v - v' = 0$, also $v = v'$. Umgekehrt, wenn $v = v'$ ist, also gilt für $w \in V$ auch stets $\langle v, w\rangle = \langle v', w\rangle$.

(b) Es seien $w, w' \in V$ gegeben. Auf Grund der Hermitizität des Skalarprodukts gilt genau dann für $v \in V$ stets $\langle v, w\rangle = \langle v, w'\rangle$, wenn für $v \in V$ stets

$$\langle w, v\rangle = \overline{\langle v, w\rangle} = \overline{\langle v, w'\rangle} = \langle w', v\rangle$$

gilt, was nach (a) äquivalent zu $w = w'$ ist. □

Skalarprodukträume als normierte Vektorräume

Wir werden in Bemerkung (7.10) sehen, dass sich mit Hilfe des Skalarprodukts Längen definieren lassen. Zunächst wollen wir die Axiomatik eines K-Vektorraums mit einem Längenmaß einführen:

(7.8) Definition (normierter Vektorraum)**.** Ein *normierter Vektorraum über* K (oder *normierter* K*-Vektorraum*) besteht aus einem K-Vektorraum V zusammen mit einer Abbildung $\|-\| \colon V \to \mathbb{R}$, $v \mapsto \|v\|$, genannt *Norm*, derart, dass folgende Axiome gelten.

- *Positive Definitheit.* Für $v \in V$ gilt

  $$\|v\| \geq 0$$

  und genau dann $\|v\| = 0$, wenn $v = 0$ ist.
- *Absolut-Homogenität.* Für $a \in K$, $v \in V$ gilt

  $$\|av\| = |a|\, \|v\|.$$
- *Dreiecksungleichung.* Für $v, w \in V$ gilt

  $$\|v + w\| \leq \|v\| + \|w\|.$$

(7.9) Beispiel. Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Es wird $\mathbb{R}^{n \times 1}$ ein normierter $\mathbb{R}$-Vektorraum, wobei die Norm durch

$$\|x\| = \|x\|_1 = \sum_{i \in [1,n]} |x_i|$$

für $x \in \mathbb{R}^{n \times 1}$ gegeben ist.

(b) Es wird $\mathbb{R}^{n \times 1}$ ein normierter $\mathbb{R}$-Vektorraum, wobei die Norm durch

$$\|x\| = \|x\|_\infty = \max \{|x_i| \mid i \in [1, n]\}$$

für $x \in \mathbb{R}^{n \times 1}$ gegeben ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Sofern es nicht zu Zweideutigkeiten kommen kann, werden wir im Folgenden ohne weiteren Kommentar einen Skalarproduktraum stets als einen normierten Vektorraum wie folgt ansehen.

(7.10) Bemerkung. Es sei ein K-Skalarproduktraum V gegeben. Der unterliegende K-Vektorraum von V wird zu einem normierten K-Vektorraum mit Norm gegeben durch

$$\|v\| = \sqrt{\langle v, v \rangle}$$

für $v \in V$.

Beweis. Da das Skalarprodukt positiv definit ist, erhalten wir eine wohldefinierte Abbildung $\|-\| \colon V \to \mathbb{R}$, $v \mapsto \sqrt{\langle v, v \rangle}$. Für $v \in V$ gilt $\|v\| = \sqrt{\langle v, v \rangle} \geq 0$ sowie genau dann $\|v\| = \sqrt{\langle v, v \rangle} = 0$, wenn $\langle v, v \rangle = 0$ ist, d.h. wenn $v = 0$ ist. Ferner ist

$$\|av\|^2 = \langle av, av \rangle = \bar{a}a\langle v, v \rangle = |a|^2\, \|v\|^2 = (|a|\, \|v\|)^2$$

und damit $\|av\| = |a|\, \|v\|$ für $a \in \mathbb{R}$, $v \in V$. Nach der Cauchy-Schwarz-Ungleichung (7.6) gilt für $v, w \in V$ schließlich

$$(\operatorname{Re} \langle v, w \rangle)^2 \leq |\langle v, w \rangle|^2 \leq \langle v, v \rangle \langle w, w \rangle = \|v\|^2\, \|w\|^2,$$

also

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle = \langle v, v \rangle + \langle v, w \rangle + \overline{\langle v, w \rangle} + \langle w, w \rangle \\ &= \|v\|^2 + 2 \operatorname{Re} \langle v, w \rangle + \|w\|^2 \leq \|v\|^2 + 2\|v\|\, \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2 \end{aligned}$$

und damit $\|v + w\| \leq \|v\| + \|w\|$. Insgesamt wird V zusammen mit $\|-\|$ zu einem normierten K-Vektorraum. □

Da wir $K^{n \times 1}$ für $n \in \mathbb{N}_0$ nach Konvention (7.3)(b) als Skalarproduktraum bzgl. des Standardskalarprodukts auffassen, wird nach Bemerkung (7.10) insbesondere auch $K^{n \times 1}$ zu einem normierten Vektorraum. Die Norm ist dann wie folgt gegeben:

(7.11) Beispiel. Es sei $n \in \mathbb{N}_0$ gegeben. Für $x \in K^{n \times 1}$ gilt

$$\|x\| = \sqrt{\sum_{i \in [1,n]} |x_i|^2}.$$

Beweis. Für $x \in K^{n \times 1}$ gilt

$$\|x\|^2 = \langle x, x \rangle = \sum_{i \in [1,n]} \overline{x_i} x_i = \sum_{i \in [1,n]} |x_i|^2$$

und damit

$$\|x\| = \sqrt{\sum_{i \in [1,n]} |x_i|^2}.$$

□

(7.12) Beispiel. In $K^{2 \times 1}$ gilt

$$\|\begin{pmatrix} 2 \\ -1 \end{pmatrix}\| = \sqrt{5}.$$

Beweis. Es ist

$$\|\begin{pmatrix} 2 \\ -1 \end{pmatrix}\| = \sqrt{2^2 + (-1)^2} = \sqrt{5}.$$

□

(7.13) Definition (normierter Vektor)**.** Es sei ein normierter K-Vektorraum V gegeben. Ein Vektor v in V heißt *normiert*, wenn

$$\|v\| = 1$$

ist.

(7.14) Beispiel. In $K^{2 \times 1}$ ist

$$\begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}$$

normiert.

Beweis. Es ist

$$\|\begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}\| = \sqrt{(\frac{\sqrt{3}}{2})^2 + (\frac{1}{2})^2} = \sqrt{\frac{3}{4} + \frac{1}{4}} = 1.$$

□

(7.15) Bemerkung. Es sei ein normierter K-Vektorraum V gegeben. Für $v \in V \setminus \{0\}$ ist $\frac{1}{\|v\|} v$ normiert.

Beweis. Wegen der Absoluthomogenität der Norm gilt für $v \in V \setminus \{0\}$ stets

$$\|\frac{1}{\|v\|} v\| = |\frac{1}{\|v\|}| \, \|v\| = \frac{1}{\|v\|} \|v\| = 1,$$

d.h. $\frac{1}{\|v\|} v$ ist normiert. □

Wenn die Norm eines normierten Vektorraums von einem Skalarprodukt induziert wird, gilt die folgende Eigenschaft.

(7.16) Proposition (Parallelogrammgesetz)**.** Es sei ein K-Skalarproduktraum V gegeben. Für $v, w \in V$ gilt

$$\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Es lässt sich zeigen, dass jede Norm eines normierten K-Vektorraums, in welchem das Parallelogrammgesetz gilt, von einem Skalarprodukt induziert wird. Das Skalarprodukt lässt sich durch die Norm wie folgt ausdrücken:

(7.17) Lemma (Polarisationsformeln)**.** Es sei ein K-Skalarproduktraum V gegeben.

(a) Wenn $K = \mathbb{R}$ ist, so gilt für $v, w \in V$ stets

$$\langle v, w\rangle = \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2) = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2).$$

(b) Wenn $K = \mathbb{C}$ ist, so gilt für $v, w \in V$ stets

$$\langle v, w\rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2) + \mathrm{i}\frac{1}{4}(-\|v+\mathrm{i}w\|^2 + \|v-\mathrm{i}w\|^2).$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Winkel

Als nächstes wollen wir mit Hilfe des Skalarprodukts Winkel in einem $\mathbb{R}$-Skalarproduktraum V definieren.

(7.18) Bemerkung. Es seien ein $\mathbb{R}$-Skalarproduktraum V und $v, w \in V \setminus \{0\}$ gegeben. Dann gibt es genau ein $\alpha \in [0, \pi]_{\mathbb{R}}$ mit

$$\cos\alpha = \frac{\langle v, w\rangle}{\|v\|\,\|w\|}.$$

Beweis. Nach der Cauchy-Schwarz-Ungleichung (7.6) gilt $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, also $|\frac{\langle v,w\rangle}{\|v\|\,\|w\|}| \leq 1$ und damit

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|\,\|w\|} \leq 1.$$

Da die Abbildung $\cos\colon [0, \pi]_{\mathbb{R}} \to [-1, 1]_{\mathbb{R}}$, $\alpha \mapsto \cos\alpha$ eine (streng monoton fallende) Bijektion ist, gibt es folglich genau ein $\alpha \in [0, \pi]_{\mathbb{R}}$ mit $\cos\alpha = \frac{\langle v,w\rangle}{\|v\|\,\|w\|}$. □

(7.19) Definition (Winkel)**.** Es seien ein $\mathbb{R}$-Skalarproduktraum V und $v, w \in V \setminus \{0\}$ gegeben. Der *Winkel* zwischen v und w ist definiert als

$$\angle(v, w) := \arccos(\frac{\langle v, w\rangle}{\|v\|\,\|w\|}).$$

(7.20) Beispiel. In $\mathbb{R}^{2\times 1}$ gilt

$$\angle(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}) = \frac{5}{12}\pi.$$

Beweis. Es ist

$$\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \rangle = 1 \cdot (-1) + 1 \cdot \sqrt{3} = -1 + \sqrt{3},$$

$$\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \| = \sqrt{1^2 + 1^2} = \sqrt{2},$$

$$\| \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \| = \sqrt{(-1)^2 + (\sqrt{3})^2} = 2$$

und damit

$$\angle(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}) = \arccos(\frac{\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \rangle}{\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \| \, \| \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \|}) = \arccos(\frac{-1 + \sqrt{3}}{\sqrt{2} \cdot 2}) = \frac{5}{12}\pi.$$ □

Orthogonalität

Im Folgenden beschäftigen wir uns mit Orthogonalität, einer Relation, welche sich mit Hilfe des Skalarprodukts auf jedem Skalarproduktraum definieren lässt.

(7.21) Definition (Orthogonalität)**.** Es sei ein K-Skalarproduktraum V gegeben. Für $v, w \in V$ sagen wir, dass v *orthogonal* zu w ist, geschrieben $v \perp w$, wenn

$$\langle v, w \rangle = 0$$

gilt.

(7.22) Beispiel. In $K^{2 \times 1}$ gilt

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \perp \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Beweis. Es ist

$$\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \rangle = 1 \cdot (-2) + 2 \cdot 1 = 0$$

und damit

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \perp \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$ □

(7.23) Bemerkung (Satz des Pythagoras)**.** Es seien ein K-Skalarproduktraum V und $v, w \in V$ mit $v \perp w$ gegeben. Dann gilt

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Beweis. Wegen $v \perp w$ ist $\langle v, w \rangle = 0$. Folglich gilt

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle = \|v\|^2 + \langle v, w \rangle + \overline{\langle v, w \rangle} + \|w\|^2 \\ &= \|v\|^2 + \|w\|^2. \end{aligned}$$ □

Orthogonalraum

Diejenigen Vektoren, welche zu einer gegebenen Menge von Vektoren orthogonal sind, bilden stets einen Untervektorraum:

(7.24) Bemerkung. Es seien ein K-Skalarproduktraum V und eine Teilmenge S von V gegeben. Dann ist

$$S^\perp = \{v \in V \mid s \perp v \text{ für } s \in S\}$$

ein K-Untervektorraum von V.

Beweis. Zunächst seien $v, w \in S^\perp$ gegeben. Für $s \in S$ gilt dann $s \perp v$ und $s \perp w$, d.h. $\langle s, v \rangle = \langle s, w \rangle = 0$. Dann gilt für $s \in S$ aber auch

$$\langle s, v + w \rangle = \langle s, v \rangle + \langle s, w \rangle = 0 + 0 = 0,$$

d.h. $s \perp v + w$. Folglich ist auch $v + w \in S^\perp$.
Da für $s \in S$ stets $\langle 0, s \rangle = 0$ und damit $0 \perp s$ gilt, ist $0 \in S^\perp$.
Schließlich seien $a \in K$, $v \in S^\perp$ gegeben. Für $s \in S$ gilt dann $v \perp s$, d.h. $\langle s, v \rangle = 0$. Dann gilt für $s \in S$ aber auch

$$\langle s, av \rangle = a \langle s, v \rangle = a \cdot 0 = 0,$$

d.h. $av \perp s$. Folglich ist auch $av \in S^\perp$.
Nach dem Untervektorraumkriterium (1.15) ist $S^\perp$ ein Untervektorraum von V. □

(7.25) Definition (Orthogonalraum)**.** Es seien ein K-Skalarproduktraum V und eine Teilmenge S von V gegeben. Der K-Untervektorraum

$$S^\perp = \{v \in V \mid s \perp v \text{ für } s \in S\}$$

von V aus Bemerkung (7.24) heißt *Orthogonalraum* zu S in V.

(7.26) Beispiel. Es sei ein K-Untervektorraum U von $K^{3\times 1}$ gegeben durch

$$U = K\mathrm{e}_1 + K\mathrm{e}_2.$$

Dann ist

$$U^\perp = K\mathrm{e}_3.$$

Beweis. Wegen

$$U = K\mathrm{e}_1 + K\mathrm{e}_2 = \{a\mathrm{e}_1 + b\mathrm{e}_2 \mid a, b \in K\} = \{\begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \mid a, b \in K\}$$

ist

$$\begin{aligned} U^\perp &= \{x \in K^{3\times 1} \mid u \perp x \text{ für } u \in U\} = \{x \in K^{3\times 1} \mid \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \perp \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ für } a, b \in K\} \\ &= \{x \in K^{3\times 1} \mid \langle \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rangle = 0 \text{ für } a, b \in K\} = \{x \in K^{3\times 1} \mid ax_1 + bx_2 + 0x_3 = 0 \text{ für } a, b \in K\} \\ &= \{x \in K^{3\times 1} \mid ax_1 + bx_2 = 0 \text{ für } a, b \in K\} = \{x \in K^{3\times 1} \mid x_1 = 0,\, x_2 = 0\} = \{\begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} \mid c \in K\} \\ &= K\mathrm{e}_3. \end{aligned}$$

□

(7.27) Proposition. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \ldots, s_n)$ in V gegeben. Dann gilt

$$\langle s_1, \ldots, s_n \rangle^\perp = \{s_1, \ldots, s_n\}^\perp.$$

Beweis. Für $v \in V$ ist $\{v\}^\perp$ nach Bemerkung (7.24) ein Untervektorraum von V, es gilt also genau dann $\{s_1, \ldots, s_n\} \subseteq \{v\}^\perp$, wenn $\langle s_1, \ldots, s_n \rangle \subseteq \{v\}^\perp$ gilt. Folglich ist

$$\begin{aligned} \{s_1, \ldots, s_n\}^\perp &= \{v \in V \mid s_i \perp v \text{ für } i \in [1, n]\} = \{v \in V \mid v \perp s_i \text{ für } i \in [1, n]\} \\ &= \{v \in V \mid \{s_1, \ldots, s_n\} \subseteq \{v\}^\perp\} = \{v \in V \mid \langle s_1, \ldots, s_n \rangle \subseteq \{v\}^\perp\} \\ &= \{v \in V \mid v \perp u \text{ für } u \in \langle s_1, \ldots, s_n \rangle\} = \{v \in V \mid u \perp v \text{ für } u \in \langle s_1, \ldots, s_n \rangle\} \\ &= \langle s_1, \ldots, s_n \rangle^\perp. \end{aligned}$$

□

Alternativer Beweis von Beispiel (7.26). Nach Proposition (7.27) gilt

$$\begin{aligned} U^\perp &= \langle \mathrm{e}_1, \mathrm{e}_2 \rangle^\perp = \{\mathrm{e}_1, \mathrm{e}_2\}^\perp = \{x \in K^{3\times 1} \mid \mathrm{e}_1 \perp x,\, \mathrm{e}_2 \perp x\} \\ &= \{x \in K^{3\times 1} \mid \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rangle = 0,\, \langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rangle = 0\} = \{x \in K^{3\times 1} \mid x_1 = 0,\, x_2 = 0\} \\ &= \{\begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} \mid c \in K\} = K\mathrm{e}_3. \end{aligned}$$

□

Die positive Definitheit hat folgende Eigenschaft des Orthogonalraums zur Folge:

(7.28) Bemerkung. Es seien ein K-Skalarproduktraum V und ein K-Untervektorraum U von V gegeben. Dann ist $(U, U^\perp)$ unabhängig.

Beweis. Für $v \in U \cap U^\perp$ gilt $\langle v, v \rangle = 0$, also $v = 0$ wegen der positiven Definitheit des Skalarprodukts. Folglich ist $U \cap U^\perp = \{0\}$ und damit $(U, U^\perp)$ unabhängig nach Proposition (1.78). □

Es sei ein K-Untervektorraum U von V gegeben. Ein Untervektorraumkomplement von U in V ist ein K-Untervektorraum U' von V mit $V = U \dotplus U'$, siehe Definition (1.86), d.h. so, dass $V = U + U'$ und (U, U') unabhängig ist. In dieser Situation gibt es eine Projektion $\mathrm{pr}_1 \colon V \to U$ von V auf U sowie eine Projektion $\mathrm{pr}_2 \colon V \to U'$ von V auf U', welche durch die Eigenschaft

$$v = \mathrm{pr}_1(v) + \mathrm{pr}_2(v)$$

charakterisiert ist, siehe Definition (1.84). Die Projektion von V auf U hängt hierbei sowohl von U als auch vom gegebenen Komplement U' ab, siehe Beispiel (1.85).
Wir wollen nun die Situation studieren, in welcher $U^\perp$ ein Komplement von U ist. Nach Bemerkung (7.28) ist $(U, U^\perp)$ stets unabhängig, so dass $U^\perp$ genau dann ein Komplement von U ist, wenn $V = U + U^\perp$ ist. Da der Orthogonalraum $U^\perp$ durch U festgelegt ist, hängt in dieser Situation die Projektion von V auf U allein von U ab. Dies gibt Anlass zu folgender Terminologie:

(7.29) Definition (Orthogonalprojektion)**.** Es seien ein K-Skalarproduktraum V und ein K-Untervektorraum U von V mit $V = U \dotplus U^\perp$ gegeben. Die Projektion von V auf U bzgl. $(U, U^\perp)$ heißt *Orthogonalprojektion* von V auf U und wird als $\mathrm{pr}_U \colon V \to U$ notiert.

In Korollar (7.47) werden wir sehen, dass der Orthogonalraum eines endlichdimensionalen Untervektorraums stets ein Komplement zu diesem ist; in diesem Fall ist also diese Prämisse von Definition (7.29) automatisch erfüllt.

(7.30) Beispiel. Es seien ein K-Untervektorraum U von $K^{3\times 1}$ und $x \in K^{3\times 1}$ gegeben durch

$$U = K\mathrm{e}_1 + K\mathrm{e}_2,\ x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}.$$

Dann ist $K^{3\times 1} = U \dotplus U^\perp$ und es gilt

$$\mathrm{pr}_U(x) = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix},$$
$$\mathrm{pr}_{U^\perp}(x) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Beweis. Nach Beispiel (7.26) ist $U^\perp = K\mathrm{e}_3$ und damit $K^{3\times 1} = U \dotplus U^\perp$. Wegen

$$x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

und

$$\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \in U,\ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \in U^\perp$$

gilt

$$\mathrm{pr}_U(x) = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix},$$
$$\mathrm{pr}_{U^\perp}(x) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$ □

In Proposition (7.41) werden wir eine Möglichkeit herleiten, die Orthogonalprojektion mit Hilfe von Orthogonalbasen und dem Skalarprodukt zu berechnen.

Orthogonalbasen

Als nächstes werden wir Tupel betrachten, deren Einträge zueinander orthogonal sind. In Korollar (7.46) werden wir sehen, dass jeder Skalarproduktraum eine Basis besitzt, deren Einträge jeweils normiert und zueinander orthogonal sind.

(7.31) Definition (Orthogonalität, Orthonormalität)**.** Es seien ein K-Skalarproduktraum V und $n \in \mathbb{N}_0$ gegeben.

(a) Ein n-Tupel $(s_1, \dots, s_n)$ in V heißt *orthogonal* (oder ein *Orthogonalsystem*), falls für $i, j \in [1, n]$ mit $i \neq j$ stets $s_i \perp s_j$ ist.

(b) Ein n-Tupel $(s_1, \dots, s_n)$ in V heißt *orthonormal* (oder *orthonormiert* oder ein *Orthonormalsystem*), falls es orthogonal und s_i für jedes $i \in [1, n]$ normiert ist.

(7.32) Beispiel.

(a) In $K^{4\times 1}$ ist das Tripel

$$(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix})$$

orthogonal.

(b) In $K^{4\times 1}$ ist das Tripel

$$(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix})$$

orthonormal.

Beweis.

(a) Wegen

$$\begin{aligned}
\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \rangle &= 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 = 0, \\
\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \rangle &= 1 \cdot 1 + 0 \cdot 0 + 1 \cdot (-1) + 0 \cdot 0 = 0, \\
\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \rangle &= 0 \cdot 1 + 1 \cdot 0 + 0 \cdot (-1) + 0 \cdot 0 = 0
\end{aligned}$$

ist

$$(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix})$$

orthogonal in $K^{4\times 1}$.

(b) Nach (a) ist

$$\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\right)$$

und nach Bemerkung (7.24) somit auch

$$\left(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\right)$$

orthogonal in $K^{4\times 1}$. Wegen

$$\|\begin{pmatrix}1\\0\\1\\0\end{pmatrix}\|^2 = 1^2 + 0^2 + 1^2 + 0^2 = 2,$$

$$\|\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\|^2 = 0^2 + 1^2 + 0^2 + 0^2 = 1,$$

$$\|\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\|^2 = 1^2 + 0^2 + (-1)^2 + 0^2 = 2$$

ist

$$\left(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\right)$$

nach Bemerkung (7.15) somit orthonormal in $K^{4\times 1}$. □

(7.33) Proposition. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein orthogonales n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Ferner seien $v, w \in V$ und $a, b \in K^n$ mit $v = \sum_{i\in[1,n]} a_i s_i$ und $w = \sum_{i\in[1,n]} b_i s_i$ gegeben. Dann ist

$$\langle v, w\rangle = \sum_{i\in[1,n]} \overline{a_i}\langle s_i, s_i\rangle b_i.$$

Beweis. Es ist

$$\langle v, w\rangle = \langle \sum_{i\in[1,n]} a_i s_i, \sum_{j\in[1,n]} b_j s_j\rangle = \sum_{i\in[1,n]}\sum_{j\in[1,n]} \overline{a_i} b_j \langle s_i, s_j\rangle = \sum_{i\in[1,n]} \overline{a_i}\langle s_i, s_i\rangle b_i.$$ □

(7.34) Korollar. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein orthonormales n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Ferner seien $v, w \in V$ und $a, b \in K^n$ mit $v = \sum_{i\in[1,n]} a_i s_i$ und $w = \sum_{i\in[1,n]} b_i s_i$ gegeben. Dann ist

$$\langle v, w\rangle = \langle a, b\rangle.$$

Beweis. Nach Proposition (7.33) ist

$$\langle v, w\rangle = \sum_{i\in[1,n]} \overline{a_i}\langle s_i, s_i\rangle b_i = \sum_{i\in[1,n]} \overline{a_i} b_i = \langle a, b\rangle.$$ □

Die folgende Korollar zeigt, dass sich die Koeffizienten von Linearkombinationen von orthogonalen Tupeln vergleichsweise einfach mit Hilfe des Skalarproduktes berechnen lassen.

(7.35) Korollar. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein orthogonales n-Tupel $(s_1, \dots, s_n)$ in V mit $s_i \neq 0$ für $i \in [1, n]$ gegeben. Ferner seien $v \in V$ und $a \in K^n$ mit $v = \sum_{i \in [1,n]} a_i s_i$ gegeben. Für $i \in [1, n]$ gilt dann

$$a_i = \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle}.$$

Beweis. Für $i \in [1, n]$ gilt

$$\langle s_i, v \rangle = \sum_{j \in [1,n]} \overline{\delta_{i,j}} \langle s_j, s_j \rangle a_j = \langle s_i, s_i \rangle a_i$$

nach Proposition (7.33) und damit

$$a_i = \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle}.$$ □

(7.36) Korollar. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein orthogonales n-Tupel $(s_1, \dots, s_n)$ in V mit $s_i \neq 0$ für $i \in [1, n]$ gegeben. Dann ist s linear unabhängig.

Beweis. Für $a \in K^n$ mit $\sum_{i \in [1,n]} a_i s_i = 0$ gilt nach Korollar (7.35) stets

$$a_i = \frac{\langle s_i, 0 \rangle}{\langle s_i, s_i \rangle} = 0$$

für $i \in [1, n]$. Nach dem Kriterium für lineare Unabhängigkeit (1.34) ist $(s_1, \dots, s_n)$ linear unabhängig. □

(7.37) Definition. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben.

(a) Wir sagen, dass $(s_1, \dots, s_n)$ eine *Orthogonalbasis* von V ist, falls $(s_1, \dots, s_n)$ orthogonal und eine Basis von V ist.

(b) Wir sagen, dass $(s_1, \dots, s_n)$ eine *Orthonormalbasis* von V ist, falls $(s_1, \dots, s_n)$ orthonormal und eine Basis von V ist.

(7.38) Beispiel.

(a) Das Quadrupel

$$(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix})$$

ist eine Orthogonalbasis von $K^{4 \times 1}$.

(b) Das Quadrupel

$$(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix})$$

ist eine Orthonormalbasis von $K^{4 \times 1}$.

Beweis.

(a) Nach Beispiel (7.32)(a) ist

$$\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\right)$$

orthogonal in $K^{4\times 1}$. Wegen

$$\langle\begin{pmatrix}a\\b\\c\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\rangle = a\cdot 0 + b\cdot 0 + c\cdot 0 + 0\cdot 1 = 0$$

für $a, b, c \in K$ ist somit auch

$$\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right)$$

orthogonal in $K^{4\times 1}$. Als Basis von $K^{4\times 1}$ ist

$$\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right)$$

daher eine Orthogonalbasis von $K^{4\times 1}$.

(b) Nach (a) ist

$$\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right)$$

und nach Bemerkung (7.24) und Korollar (1.23)(c) somit auch

$$\left(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right)$$

eine Orthogonalbasis von $K^{4\times 1}$. Ferner sind die Einträge von

$$\left(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\right)$$

nach Beispiel (7.32)(b) normiert. Wegen

$$\|\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\|^2 = 0^2 + 0^2 + 0^2 + 1^2 = 1$$

ist schließlich

$$(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix})$$

eine Orthonormalbasis von $K^{4\times 1}$. □

(7.39) Bemerkung. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein n-Tupel $(s_1, \dots, s_n)$ in V gegeben.

(a) Es gelte $s_i \neq 0$ für $i \in [1, n]$. Die folgenden Bedingungen sind äquivalent.

(i) Das n-Tupel $(s_1, \dots, s_n)$ ist eine Orthogonalbasis von V.

(ii) Das n-Tupel $(s_1, \dots, s_n)$ ist orthogonal in V und ein Erzeugendensystem von V.

(iii) Das n-Tupel $(s_1, \dots, s_n)$ ist orthogonal in V und es gilt $n = \dim_K V$.

(b) Die folgenden Bedingungen sind äquivalent.

(i) Das n-Tupel $(s_1, \dots, s_n)$ ist eine Orthonormalbasis von V.

(ii) Das n-Tupel $(s_1, \dots, s_n)$ ist orthonormal in V und ein Erzeugendensystem von V.

(iii) Das n-Tupel $(s_1, \dots, s_n)$ ist orthonormal in V und $n = \dim_K V$.

Beweis.

(a) Genau dann ist $(s_1, \dots, s_n)$ eine Orthogonalbasis von V, wenn $(s_1, \dots, s_n)$ orthogonal und eine Basis von V ist.

Es sei $(s_1, \dots, s_n)$ orthogonal. Dann ist $(s_1, \dots, s_n)$ linear unabhängig nach Korollar (7.36). Folglich ist $(s_1, \dots, s_n)$ genau dann eine Basis von V, wenn $(s_1, \dots, s_n)$ ein Erzeugendensystem von V ist. Ferner ist $(s_1, \dots, s_n)$ nach Korollar (1.67) genau dann eine Basis von V, wenn $n = \dim V$ ist. Folglich sind Bedingung (i), Bedingung (ii) und Bedingung (iii) äquivalent.

(b) Dies folgt aus (a). □

(7.40) Bemerkung (Parsevalscher Entwicklungssatz)**.** Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und eine Orthogonalbasis $(s_1, \dots, s_n)$ von V gegeben. Für $v \in V$ gilt

$$v = \sum_{i \in [1,n]} \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle} s_i.$$

Beweis. Es sei $v \in V$ gegeben. Da $(s_1, \dots, s_n)$ eine Basis von V ist, gibt es ein $a \in K^n$ mit

$$v = \sum_{i \in [1,n]} a_i s_i.$$

Nach Korollar (7.35) folgt $a_i = \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle}$ für $i \in [1, n]$ und damit

$$v = \sum_{i \in [1,n]} \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle} s_i.$$ □

(7.41) Proposition. Es seien ein K-Skalarproduktraum V, ein K-Untervektorraum U von V, $n \in \mathbb{N}_0$ und eine Orthogonalbasis $(s_1, \dots, s_n)$ von U gegeben. Dann ist $V = U \dotplus U^\perp$. Für $v \in V$ gilt

$$\mathrm{pr}_U(v) = \sum_{i \in [1,n]} \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle} s_i,$$

$$\mathrm{pr}_{U^\perp}(v) = v - \sum_{i \in [1,n]} \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle} s_i.$$

Beweis. Es sei $v \in V$ gegeben. Für $j \in [1, n]$ gilt

$$\langle s_j, v - \sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} s_i\rangle = \langle s_j, v\rangle - \sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} \langle s_j, s_i\rangle = \langle s_j, v\rangle - \frac{\langle s_j, v\rangle}{\langle s_j, s_j\rangle} \langle s_j, s_j\rangle = 0,$$

also $s_j \perp v - \sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} s_i$. Nach Proposition (7.27) ist daher

$$v - \sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} s_i \in \{s_1, \dots, s_n\}^\perp = \langle s_1, \dots, s_n\rangle^\perp = U^\perp.$$

Ferner ist offenbar $\sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} s_i \in U$ und damit

$$v = \sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} s_i + (v - \sum_{i \in [1,n]} \frac{\langle s_i, v\rangle}{\langle s_i, s_i\rangle} s_i) \in U + U^\perp.$$

Da $v \in V$ beliebig war, gilt somit $V = U + U^\perp$. Nach Bemerkung (7.28) ist ferner $(U, U^\perp)$ unabhängig und damit $V = U \dot{+} U^\perp$. □

Um mit Hilfe von Proposition (7.41) Orthogonalprojektionen berechnen zu können, benötigen wir also eine Möglichkeit zur Berechnung von Orthogonalbasen. Ein solches Verfahren liefert die Gram-Schmidt-Orthogonalisierung (7.42), welche wir gleich studieren werden. Eine Illustration zu Proposition (7.41) folgt dann in Beispiel (7.44).
Proposition (7.41) liefert folgende Formulierung des entscheidenden Schritts im Beweis der Cauchy-Schwarz-Ungleichung (7.6):

$$\begin{aligned} 0 &< \langle \mathrm{pr}_{(Kv)^\perp}(w), \mathrm{pr}_{(Kv)^\perp}(w)\rangle = \langle w - \mathrm{pr}_{Kv}(w), \mathrm{pr}_{(Kv)^\perp}(w)\rangle = \langle w, \mathrm{pr}_{(Kv)^\perp}(w)\rangle - \langle \mathrm{pr}_{Kv}(w), \mathrm{pr}_{(Kv)^\perp}(w)\rangle \\ &= \langle w, w - \frac{\langle v, w\rangle}{\langle v, v\rangle} v\rangle = \langle w, w\rangle - \frac{\langle v, w\rangle}{\langle v, v\rangle} \langle w, v\rangle = \langle w, w\rangle - \frac{\langle v, w\rangle}{\langle v, v\rangle} \overline{\langle v, w\rangle} = \langle w, w\rangle - \frac{|\langle v, w\rangle|^2}{\langle v, v\rangle} \end{aligned}$$

Gram-Schmidt-Orthogonalisierung

Wir zeigen nun, wie sich linear unabhängige Tupel schrittweise „orthogonalisieren“ lassen. Insbesondere liefert dies eine Methode, um ausgehend von einer beliebigen Basis eine Orthogonalbasis und unter anschließender Anwendung von Bemerkung (7.15) auch eine Orthonormalbasis zu konstruieren.
Das entscheidende Hilfsmittel hierbei ist Proposition (7.41), welche uns erlaubt, Vektoren auf beliebige endlichdimensionale Untervektorräume und deren Orthogonalräume zu projizieren. Jedes linear unabhängige n-Tupel $(s_1, \dots, s_n)$ in einem K-Skalarproduktraum V, wobei $n \in \mathbb{N}_0$, liefert ineinander enthaltene Untervektorräume

$$\langle s_1\rangle < \langle s_1, s_2\rangle < \ldots < \langle s_1, \dots, s_n\rangle.$$

Haben wir eine Orthogonalbasis $(t_1, \dots, t_{j-1})$ von $\langle s_1, \dots, s_{j-1}\rangle$ gegeben, so liefert die Projektion von s_j für $j \in [0, n]$ auf $\langle s_1, \dots, s_{j-1}\rangle^\perp$ einen Vektor t_j derart, dass $(t_1, \dots, t_j)$ eine Orthogonalbasis von $\langle s_1, \dots, s_j\rangle$ ist. Nach rekursiver Anwendung haben wir eine Orthogonalbasis von $\langle s_1, \dots, s_n\rangle$ konstruiert.

(7.42) Proposition (Gram-Schmidt-Orthogonalisierung)**.** Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$ und ein linear unabhängiges n-Tupel $(s_1, \dots, s_n)$ in V gegeben. Ferner sei $(t_1, \dots, t_n)$ gegeben durch

$$t_j = \mathrm{pr}_{\langle s_1, \dots, s_{j-1}\rangle^\perp}(s_j)$$

für $j \in [1, n]$. Für jedes $j \in [0, n]$ ist dann $(t_1, \dots, t_j)$ eine Orthogonalbasis von $\langle s_1, \dots, s_j\rangle$.

Beweis. Wir führen Induktion nach j, wobei für $j = 0$ nichts zu zeigen ist. Es sei also $j \in [1, n]$ derart gegeben, dass $(t_1, \dots, t_{j-1})$ eine Orthogonalbasis von $\langle s_1, \dots, s_{j-1}\rangle$ ist. Wegen der linearen Unabhängigkeit von $(s_1, \dots, s_j)$ ist $s_j \notin \langle s_1, \dots, s_{j-1}\rangle$ nach Proposition (1.37) und damit

$$t_j = \mathrm{pr}_{\langle s_1, \dots, s_{j-1}\rangle^\perp}(s_j) = s_j - \mathrm{pr}_{\langle s_1, \dots, s_{j-1}\rangle}(s_j) \neq 0.$$

Ferner ist $t_j = \mathrm{pr}_{\langle s_1,\ldots,s_{j-1}\rangle^\perp}(s_j) = \mathrm{pr}_{\langle t_1,\ldots,t_{j-1}\rangle^\perp}(s_j) \perp t_i$ für $i \in [1, j-1]$ und damit $(t_1, \ldots, t_j)$ orthogonal. Nach Korollar (1.23) ist schließlich

$$\langle t_1, \ldots, t_j\rangle = \langle s_1, \ldots, s_{j-1}, s_j - \mathrm{pr}_{\langle s_1,\ldots,s_{j-1}\rangle}(s_j)\rangle = \langle s_1, \ldots, s_j\rangle,$$

d.h. $(t_1, \ldots, t_j)$ ist ein Erzeugendensystem von $\langle s_1, \ldots, s_j\rangle$. Insgesamt ist $(t_1, \ldots, t_j)$ eine Orthogonalbasis von $\langle s_1, \ldots, s_j\rangle$ nach Bemerkung (7.39)(a).
Nach dem Induktionsprinzip ist $(t_1, \ldots, t_j)$ für jedes $j \in [0, n]$ eine Orthogonalbasis von $\langle s_1, \ldots, s_j\rangle$. □

Nach Proposition (7.41) ist das n-Tupel $(t_1, \ldots, t_n)$ in der Formulierung der Gram-Schmidt-Orthogonalisierung (7.42) rekursiv gegeben durch

$$t_j = \mathrm{pr}_{\langle s_1,\ldots,s_{j-1}\rangle^\perp}(s_j) = \mathrm{pr}_{\langle t_1,\ldots,t_{j-1}\rangle^\perp}(s_j) = s_j - \sum_{i\in[1,j-1]} \frac{\langle t_i, s_j\rangle}{\langle t_i, t_i\rangle} t_i$$

für $j \in [1, n]$.

(7.43) Beispiel. Es seien $s = (s_1, s_2, s_3)$ und $t = (t_1, t_2, t_3)$ in $K^{3\times 1}$ gegeben durch

$$s = (\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}),\ t = (\frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}).$$

Dann ist t eine Orthonormalbasis von $K^{3\times 1}$ und es gilt $\langle s_1\rangle = \langle t_1\rangle$, $\langle s_1, s_2\rangle = \langle t_1, t_2\rangle$, $\langle s_1, s_2, s_3\rangle = \langle t_1, t_2, t_3\rangle$.

Beweis. Nach der Gram-Schmidt-Orthogonalisierung (7.42) ist (t'_1, t'_2, t'_3) gegeben durch $t'_j = \mathrm{pr}_{\langle s_1,\ldots,s_{j-1}\rangle^\perp}(s_j)$ für $j \in [1, 3]$ eine Orthogonalbasis von $\langle s_1, s_2, s_3\rangle$ und es gilt $\langle s_1\rangle = \langle t'_1\rangle$, $\langle s_1, s_2\rangle = \langle t'_1, t'_2\rangle$, $\langle s_1, s_2, s_3\rangle = \langle t'_1, t'_2, t'_3\rangle$.
Wir berechnen (t'_1, t'_2, t'_3) mit Proposition (7.41):

$$\begin{aligned}
t'_1 &= \mathrm{pr}_{\langle\,\rangle^\perp}(s_1) = s_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \\
\langle t'_1, t'_1\rangle &= 1^2 + 2^2 + 2^2 = 9, \\
\langle t'_1, s_2\rangle &= 1 \cdot 1 + 2 \cdot 0 + 2 \cdot 1 = 3, \\
t'_2 &= \mathrm{pr}_{\langle s_1\rangle^\perp}(s_2) = s_2 - \frac{\langle t'_1, s_2\rangle}{\langle t'_1, t'_1\rangle} t'_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{3}{9}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \\
\langle t'_2, t'_2\rangle &= \frac{2^2 + (-2)^2 + 1^2}{3^2} = 1, \\
\langle t'_1, s_3\rangle &= 1 \cdot 0 + 2 \cdot (-1) + 2 \cdot 1 = 0, \\
\langle t'_2, s_3\rangle &= \frac{1}{3}(2 \cdot 0 + (-2) \cdot (-1) + 1 \cdot 1) = 1, \\
t'_3 &= \mathrm{pr}_{\langle s_1,s_2\rangle^\perp}(s_3) = s_3 - \frac{\langle t'_1, s_3\rangle}{\langle t'_1, t'_1\rangle} t'_1 - \frac{\langle t'_2, s_3\rangle}{\langle t'_2, t'_2\rangle} t'_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - \frac{0}{9}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{1}{1} \cdot \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \\
\langle t'_3, t'_3\rangle &= \frac{(-2)^2 + (-1)^2 + 2^2}{3^2} = 1
\end{aligned}$$

Nach Bemerkung (7.15) ist folglich

$$t = (\frac{1}{\|t'_1\|} t'_1, \frac{1}{\|t'_2\|} t'_2, \frac{1}{\|t'_3\|} t'_3) = (\frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix})$$

eine Orthonormalbasis von $K^{3\times 1}$. □

Mit Hilfe einer Berechnungsmöglichkeit von Orthogonalbasen und Proposition (7.41) können wir nun ein Beispiel zur Berechnung von Orthogonalprojektionen auf endlichdimensionale Untervektorräume angeben:

(7.44) Beispiel. Es seien ein K-Untervektorraum U von $K^{3\times 1}$ und $x \in K^{3\times 1}$ gegeben durch

$$U = K\begin{pmatrix}1\\2\\2\end{pmatrix} + K\begin{pmatrix}1\\0\\1\end{pmatrix},\ x = \begin{pmatrix}1\\1\\1\end{pmatrix}.$$

Dann gilt

$$\mathrm{pr}_U(x) = \frac{1}{9}\begin{pmatrix}7\\8\\11\end{pmatrix},$$
$$\mathrm{pr}_{U^\perp}(x) = \frac{1}{9}\begin{pmatrix}2\\1\\-2\end{pmatrix}.$$

Beweis. Nach Beispiel (7.43) ist $s = (s_1, s_2)$ in $K^{3\times 1}$ gegeben durch

$$s = (\frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix}, \frac{1}{3}\begin{pmatrix}2\\-2\\1\end{pmatrix})$$

eine Orthonormalbasis von U. Folglich gilt

$$\mathrm{pr}_U(x) = \langle s_1, x\rangle s_1 + \langle s_2, x\rangle s_2 = \frac{1\cdot 1 + 2\cdot 1 + 2\cdot 1}{9}\begin{pmatrix}1\\2\\2\end{pmatrix} + \frac{2\cdot 1 + (-2)\cdot 1 + 1\cdot 1}{9}\begin{pmatrix}2\\-2\\1\end{pmatrix} = \frac{1}{9}\begin{pmatrix}7\\8\\11\end{pmatrix},$$
$$\mathrm{pr}_{U^\perp}(x) = x - \mathrm{pr}_U(x) = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1}{9}\begin{pmatrix}7\\8\\11\end{pmatrix} = \frac{1}{9}\begin{pmatrix}2\\1\\-2\end{pmatrix}$$

nach Proposition (7.41). □

(7.45) Proposition (Ergänzungssatz für Orthogonalbasen)**.** Es seien ein endlichdimensionaler K-Skalarproduktraum V, $m \in \mathbb{N}_0$ und ein orthogonales m-Tupel $(s_1, \dots, s_m)$ in V mit $s_i \neq 0$ für $i \in [1, m]$ gegeben. Dann gibt es ein $n \in \mathbb{N}_0$ und ein orthogonales n-Tupel $(t_1, \dots, t_n)$ in V so, dass $(s_1, \dots, s_m, t_1, \dots, t_n)$ eine Orthogonalbasis von V ist.

Beweis. Nach Korollar (7.36) ist $(s_1, \dots, s_m)$ linear unabhängig in V. Da V endlich erzeugt ist, gibt es nach dem Basisergänzungssatz (1.53) ein n-Tupel $(u_1, \dots, u_n)$ in V so, dass $(s_1, \dots, s_m, u_1, \dots, u_n)$ eine Basis von V ist. Insbesondere ist $(s_1, \dots, s_m, u_1, \dots, u_n)$ linear unabhängig in V. Es sei $(t_1, \dots, t_n)$ in V gegeben durch

$$t_j = \mathrm{pr}_{\langle s_1, \dots, s_m, u_1, \dots, u_{j-1}\rangle^\perp}(u_j)$$

für $j \in [1, n]$. Wegen der Orthogonalität von $(s_1, \dots, s_m)$ ist $(s_1, \dots, s_m, t_1, \dots, t_n)$ nach der Gram-Schmidt-Orthogonalisierung (7.42) eine Orthogonalbasis von $\langle s_1, \dots, s_m, u_1, \dots, u_n\rangle = V$. □

Der Ergänzungssatz für Orthogonalbasen (7.45) besagt also, dass jede linear unabhängige Teilmenge eines endlichdimensionalen Skalarproduktraums zu einer Orthogonalbasis dieses Vektorraums ergänzt werden kann.

(7.46) Korollar. Es sei ein endlichdimensionaler K-Skalarproduktraum V gegeben. Dann gibt es eine Orthonormalbasis von V.

Beweis. Da das 0-Tupel ( ) orthogonal in V ist, gibt es nach dem Ergänzungssatz für Orthogonalbasen (7.45) ein $n \in \mathbb{N}_0$ und eine Orthonormalbasis $(s_1, \dots, s_n)$ von V. □

(7.47) Korollar. Es seien ein K-Skalarproduktraum V und ein endlichdimensionaler K-Untervektorraum U von V gegeben. Dann ist $U^\perp$ ein Untervektorraumkomplement von U in V.

Beweis. Nach Korollar (7.46) gibt es ein $n \in \mathbb{N}_0$ und eine Orthonormalbasis $(s_1, \ldots, s_n)$ von U, so dass $U^\perp$ nach Proposition (7.41) ein Komplement von U in V ist. □

(7.48) Korollar. Es seien ein endlichdimensionaler K-Skalarproduktraum V, ein K-Untervektorraum U von V, $m \in \mathbb{N}_0$ und eine Orthogonalbasis $(s_1, \ldots, s_m)$ von U gegeben. Dann gibt es ein $n \in \mathbb{N}_0$ und ein orthogonales n-Tupel $(t_1, \ldots, t_n)$ in V so, dass $(s_1, \ldots, s_m, t_1, \ldots, t_n)$ eine Orthogonalbasis von V ist.

Beweis. Als Orthogonalbasis von U ist $(s_1, \ldots, s_m)$ insbesondere orthogonal in U und damit auch in V. Nach dem Ergänzungssatz für Orthogonalbasen (7.45) gibt es ein $n \in \mathbb{N}_0$ und ein orthogonales n-Tupel $(t_1, \ldots, t_n)$ in V so, dass $(s_1, \ldots, s_m, t_1, \ldots, t_n)$ eine Orthogonalbasis von V ist. □

Der Ergänzungssatz für Orthogonalbasen (7.45) liefert eine Methode zur Berechnung einer Orthogonalbasis des Orthogonalraums eines Untervektorraums in einem endlichdimensionalen Vektorraum:

(7.49) Bemerkung. Es seien ein endlichdimensionaler K-Skalarproduktraum V, ein K-Untervektorraum U von V, $m, n \in \mathbb{N}_0$, eine Orthogonalbasis $(s_1, \ldots, s_m)$ von U und ein n-Tupel $(t_1, \ldots, t_n)$ in V derart gegeben, dass $(s_1, \ldots, s_m, t_1, \ldots, t_n)$ eine Orthogonalbasis von V ist. Dann ist $(t_1, \ldots, t_n)$ eine Orthogonalbasis von $U^\perp$.

Beweis. Da U endlichdimensional ist, gilt $V = U \dot{+} U^\perp$ nach Korollar (7.47) und damit

$$\dim U^\perp = \dim V - \dim U = m + n - m = n$$

nach Korollar (1.81). Wegen der Orthogonalität von $(s_1, \ldots, s_m, t_1, \ldots, t_n)$ ist ferner $(t_1, \ldots, t_n)$ ein orthogonales n-Tupel in $\langle s_1, \ldots, s_m \rangle^\perp = U^\perp$. Insgesamt ist $(t_1, \ldots, t_n)$ nach Bemerkung (7.39)(a) eine Orthogonalbasis von $U^\perp$. □

Alternativer Beweis von Bemerkung (7.49). Nach dem Parsevalschen Entwicklungssatz (7.40) gilt für $v \in V$ stets

$$v = \sum_{i \in [1,m]} \frac{\langle s_i, v \rangle}{\langle s_i, s_i \rangle} s_i + \sum_{j \in [1,n]} \frac{\langle t_j, v \rangle}{\langle t_j, t_j \rangle} t_j,$$

so dass nach Proposition (7.27) nun

$$\begin{aligned} U^\perp &= \langle s_1, \ldots, s_m \rangle^\perp = \{s_1, \ldots, s_m\}^\perp = \{v \in V \mid s_i \perp v \text{ für } i \in [1,m]\} \\ &= \{v \in V \mid \langle s_i, v \rangle = 0 \text{ für } i \in [1,m]\} = \langle t_1, \ldots, t_n \rangle \end{aligned}$$

folgt. Somit ist $(t_1, \ldots, t_n)$ ein Erzeugendensystem von $U^\perp$. Wegen der Orthogonalität von $(s_1, \ldots, s_m, t_1, \ldots, t_n)$ in V ist $(t_1, \ldots, t_n)$ orthogonal in $U^\perp$ und folglich nach Bemerkung (7.39)(a) eine Orthogonalbasis von $U^\perp$. □

Alternativer Beweis von Beispiel (7.44). Nach Beispiel (7.43) ist $s = (s_1, s_2, s_3)$ in $K^{3 \times 1}$ gegeben durch

$$s = (\frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \frac{1}{3} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix})$$

eine Orthonormalbasis von $K^{3 \times 1}$ und (s_1, s_2) eine Orthonormalbasis von U. Folglich ist (s_3) nach Bemerkung (7.49) eine Orthonormalbasis von $U^\perp$. Nach Proposition (7.41) folgt

$$\begin{aligned} \mathrm{pr}_{U^\perp}(x) &= \langle s_3, x \rangle s_3 = \frac{(-2) \cdot 1 + (-1) \cdot 1 + 2 \cdot 1}{9} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} = -\frac{1}{9} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}, \\ \mathrm{pr}_U(x) &= x - \mathrm{pr}_{U^\perp}(x) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{9} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 7 \\ 8 \\ 11 \end{pmatrix}. \end{aligned}$$

□

QR-Zerlegung

In Matrixform lässt sich die Gram-Schmidt-Orthogonalisierung (7.42) wie folgt formulieren.

(7.50) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ mit $\operatorname{rk} A = n$ gegeben. Ferner seien n-Tupel $(t'_1, \dots, t'_n)$ und $(t_1, \dots, t_n)$ in $K^{m \times 1}$ gegeben durch

$$\begin{aligned} t'_j &= \operatorname{pr}_{\langle A_{-,1}, \dots, A_{-,j-1} \rangle^\perp}(A_{-,j}), \\ t_j &:= \frac{1}{\|t'_j\|} t'_j \end{aligned}$$

für $j \in [1, n]$ und es seien $Q \in K^{m \times n}$, $R \in K^{n \times n}$ gegeben durch

$$Q = \begin{pmatrix} t_1 & \dots & t_n \end{pmatrix}, \; R = \begin{pmatrix} \langle t_1, A_{-,1} \rangle & \dots & \langle t_1, A_{-,n} \rangle \\ & \ddots & \vdots \\ 0 & & \langle t_n, A_{-,n} \rangle \end{pmatrix}.$$

Dann ist $R \in \operatorname{GL}_n(K)$ und es gilt

$$A = QR.$$

Beweis. Nach der Gram-Schmidt-Orthogonalisierung (7.42) ist $(t'_1, \dots, t'_j)$ für $j \in [1, n]$ eine Orthogonalbasis von $\langle A_{-,1}, \dots, A_{-,j} \rangle$. Somit gilt

$$\begin{aligned} \|t'_j\|^2 &= \langle t'_j, t'_j \rangle = \langle t'_j, \operatorname{pr}_{\langle A_{-,1}, \dots, A_{-,j-1} \rangle^\perp}(A_{-,j}) \rangle = \langle t'_j, A_{-,j} - \operatorname{pr}_{\langle A_{-,1}, \dots, A_{-,j-1} \rangle}(A_{-,j}) \rangle \\ &= \langle t'_j, A_{-,j} - \operatorname{pr}_{\langle t'_1, \dots, t'_{j-1} \rangle}(A_{-,j}) \rangle = \langle t'_j, A_{-,j} \rangle - \langle t'_j, \operatorname{pr}_{\langle t'_1, \dots, t'_{j-1} \rangle}(A_{-,j}) \rangle = \langle t'_j, A_{-,j} \rangle \end{aligned}$$

für $j \in [1, n]$ nach Proposition (7.41) und damit

$$\det R = \prod_{j \in [1,n]} \langle t_j, A_{-,j} \rangle = \prod_{j \in [1,n]} \langle \frac{1}{\|t'_j\|} t'_j, A_{-,j} \rangle = \prod_{j \in [1,n]} \frac{1}{\|t'_j\|} \langle t'_j, A_{-,j} \rangle = \prod_{j \in [1,n]} \|t'_j\| > 0$$

nach Korollar (5.16). Insbesondere ist $R \in \operatorname{GL}_n(K)$ nach Satz (5.14). Ferner gilt

$$t'_j = \|t'_j\| \, t_j = \frac{1}{\|t'_j\|} \langle t'_j, A_{-,j} \rangle t_j = \langle \frac{1}{\|t'_j\|} t'_j, A_{-,j} \rangle t_j = \langle t_j, A_{-,j} \rangle t_j$$

und damit

$$\begin{aligned} A_{-,j} &= \operatorname{pr}_{\langle t'_1, \dots, t'_{j-1} \rangle}(A_{-,j}) + \operatorname{pr}_{\langle t'_1, \dots, t'_{j-1} \rangle^\perp}(A_{-,j}) = \sum_{i \in [1,j-1]} \langle t_i, A_{-,j} \rangle t_i + t'_j = \sum_{i \in [1,j]} \langle t_i, A_{-,j} \rangle t_i \\ &= \sum_{i \in [1,j]} \langle t_i, A_{-,j} \rangle t_i + \sum_{i \in [j+1,n]} 0 \, t_i = \begin{pmatrix} t_1 & \dots & t_j & t_{j+1} & \dots & t_n \end{pmatrix} \begin{pmatrix} \langle t_1, A_{-,j} \rangle \\ \vdots \\ \langle t_j, A_{-,j} \rangle \\ 0 \\ \vdots \\ 0 \end{pmatrix} = QR_{-,j} \end{aligned}$$

für $j \in [1, n]$, also

$$A = \begin{pmatrix} A_{-,1} & \dots & A_{-,n} \end{pmatrix} = \begin{pmatrix} QR_{-,1} & \dots & QR_{-,n} \end{pmatrix} = Q \begin{pmatrix} R_{-,1} & \dots & R_{-,n} \end{pmatrix} = QR.$$

□

(7.51) Definition (QR-Zerlegung)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ mit $\operatorname{rk}_K A = n$ gegeben. Das Paar (Q, R) mit $Q \in K^{m \times n}$ und $R \in K^{n \times n}$ gegeben wie in Bemerkung (7.50) wird *QR-Zerlegung* von A genannt.

(7.52) Beispiel. Es seien $A, Q, R \in K^{3 \times 3}$ gegeben durch

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \\ 2 & 1 & 1 \end{pmatrix}, \; Q = \frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ 2 & -2 & -1 \\ 2 & 1 & 2 \end{pmatrix}, \; R = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dann ist $R \in \operatorname{GL}_3(K)$ und es gilt $A = QR$.

Beweis. Dies folgt aus Beispiel (7.43) und Bemerkung (7.50). □

Adjunktion von Vektorraumhomomorphismen

Als nächstes geben wir der Adjunktion von Matrizen, siehe Definition (A.149), einen konzeptionellen Hintergrund:

(7.53) Definition (Adjungierbarkeit von Vektorraumhomomorphismen)**.** Es seien K-Skalarprodukträume V und W und ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben.

(a) Eine *Adjungierte* (oder ein *adjungierter K-Vektorraumhomomorphismus*) zu φ ist ein K-Vektorraumhomomorphismus $\psi\colon W \to V$ derart, dass für $v \in V$, $w \in W$ stets

$$\langle \varphi(v), w\rangle = \langle v, \psi(w)\rangle$$

gilt.

(b) Der K-Vektorraumhomomorphismus φ heißt *adjungierbar*, falls es eine Adjungierte zu φ gibt.

In Proposition (7.69) werden wir sehen, dass jeder Homomorphismus zwischen endlichdimensionalen K-Vektorräumen adjungierbar ist.

(7.54) Beispiel. Es seien

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}^3,\ x \mapsto (x_1 + 2x_2, x_1 - x_2, -2x_1 + x_2),$$
$$\psi\colon \mathbb{R}^3 \to \mathbb{R}^2,\ y \mapsto (y_1 + y_2 - 2y_3, 2y_1 - y_2 + y_3).$$

Dann ist ψ ein zu φ adjungierter K-Vektorraumhomomorphismus.

Beweis. Für $x \in \mathbb{R}^2$, $y \in \mathbb{R}^3$ gilt

$$\begin{aligned}\langle \varphi(x), y\rangle &= \langle (x_1 + 2x_2, x_1 - x_2, -2x_1 + x_2), y\rangle = (x_1 + 2x_2)y_1 + (x_1 - x_2)y_2 + (-2x_1 + x_2)y_3 \\ &= x_1y_1 + 2x_2y_1 + x_1y_2 - x_2y_2 - 2x_1y_3 + x_2y_3 = x_1y_1 + x_1y_2 - 2x_1y_3 + 2x_2y_1 - x_2y_2 + x_2y_3 \\ &= x_1(y_1 + y_2 - 2y_3) + x_2(2y_1 - y_2 + y_3) = \langle x, (y_1 + y_2 - 2y_3, 2y_1 - y_2 + y_3)\rangle = \langle x, \psi(y)\rangle,\end{aligned}$$

d.h. ψ ist adjungiert zu φ. □

So wie die Inverse einer invertierbaren Abbildung ist die Adjungierten eines adjungierbaren Endomorphismus stets eindeutig:

(7.55) Bemerkung. Es seien K-Skalarprodukträume V und W und ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann gibt es höchstens eine Adjungierte zu φ.

Beweis. Es seien $\psi\colon W \to V$ und $\psi'\colon W \to V$ zu φ adjungierte Homomorphismen. Für $v \in V$, $w \in W$ gilt dann

$$\langle v, \psi(w)\rangle = \langle \varphi(v), w\rangle = \langle v, \psi'(w)\rangle,$$

also $\psi(w) = \psi'(w)$ für $w \in W$ nach Lemma (7.7)(b) und damit $\psi = \psi'$ nach dem Gleichheitskriterium für Abbildungen (A.35). □

(7.56) Notation. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Die Adjungierte zu φ notieren wir als $\varphi^{\mathrm{ad}}\colon W \to V$.

(7.57) Proposition.

(a) Es seien K-Skalarprodukträume V, W und X und adjungierbare K-Vektorraumhomomorphismen $\varphi\colon V \to W$ und $\psi\colon W \to X$ gegeben. Dann ist auch $\psi \circ \varphi\colon V \to X$ adjungierbar mit

$$(\psi \circ \varphi)^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ \psi^{\mathrm{ad}}.$$

(b) Es sei ein K-Skalarproduktraum V gegeben. Die Identität $\mathrm{id}_V\colon V \to V$ ist ein adjungierbarer K-Vektorraumhomomorphismus mit

$$\mathrm{id}_V^{\mathrm{ad}} = \mathrm{id}_V.$$

(c) Es seien K-Skalarprodukträume V und W und adjungierbare K-Vektorraumhomomorphismen φ, $\varphi'\colon V \to W$ gegeben. Dann ist auch $\varphi + \varphi'\colon V \to W$ adjungierbar mit

$$(\varphi + \varphi')^{\mathrm{ad}} = \varphi^{\mathrm{ad}} + \varphi'^{\mathrm{ad}}.$$

(d) Es seien $a \in K$, K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist auch $a\varphi\colon V \to W$ adjungierbar mit

$$(a\varphi)^{\mathrm{ad}} = \bar{a}\varphi^{\mathrm{ad}}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.58) Bemerkung. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist auch $\varphi^{\mathrm{ad}}\colon W \to V$ adjungierbar mit

$$(\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.59) Korollar. Es seien K-Skalarprodukträume V und W und ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Genau dann ist φ adjungierbar, wenn für $v \in V$, $w \in W$ stets

$$\langle w, \varphi(v)\rangle = \langle \varphi^{\mathrm{ad}}(w), v\rangle$$

gilt.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.60) Bemerkung. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Für jede Teilmenge S von V ist

$$\varphi(S)^{\perp} = (\varphi^{\mathrm{ad}})^{-1}(S^{\perp}).$$

Beweis. Für jede Teilmenge S von V ist

$$\begin{aligned}\varphi(S)^{\perp} &= \{w \in W \mid t \perp w \text{ für } t \in \varphi(S)\} = \{w \in W \mid \varphi(s) \perp w \text{ für } s \in S\}\\ &= \{w \in W \mid \langle \varphi(s), w\rangle = 0 \text{ für } s \in S\} = \{w \in W \mid \langle s, \varphi^{\mathrm{ad}}(w)\rangle = 0 \text{ für } s \in S\}\\ &= \{w \in W \mid \varphi^{\mathrm{ad}}(w) \in S^{\perp}\} = (\varphi^{\mathrm{ad}})^{-1}(S^{\perp}).\end{aligned}$$ □

(7.61) Korollar. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist

$$(\operatorname{Im}\varphi)^{\perp} = \operatorname{Ker}\varphi^{\mathrm{ad}}.$$

Beweis. Nach Bemerkung (7.60) ist

$$(\operatorname{Im}\varphi)^{\perp} = \varphi(V)^{\perp} = (\varphi^{\mathrm{ad}})^{-1}(V^{\perp}) = (\varphi^{\mathrm{ad}})^{-1}(\{0\}) = \operatorname{Ker}\varphi^{\mathrm{ad}}.$$ □

(7.62) Proposition. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Für $w \in W$ ist

$$\varphi^{-1}(\{\mathrm{pr}_{\operatorname{Im}\varphi}(w)\}) = (\varphi^{\mathrm{ad}} \circ \varphi)^{-1}(\{\varphi^{\mathrm{ad}}(w)\}).$$

Beweis. Wir zeigen zuerst, dass $\operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi) = \operatorname{Ker}\varphi$ ist, und danach, dass für $w \in W$ stets $\varphi^{-1}(\{\mathrm{pr}_{\operatorname{Im}\varphi}(w)\}) = (\varphi^{\mathrm{ad}} \circ \varphi)^{-1}(\varphi^{\mathrm{ad}}(w))$ ist.
Für $v \in \operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi)$ gilt $\varphi^{\mathrm{ad}}(\varphi(v)) = (\varphi^{\mathrm{ad}} \circ \varphi)(v) = 0$ und damit

$$\langle \varphi(v), \varphi(v)\rangle = \langle v, \varphi^{\mathrm{ad}}(\varphi(v))\rangle = \langle v, 0\rangle = 0,$$

also $\varphi(v) = 0$ wegen der positiven Definitheit des Skalarprodukts und damit $v \in \operatorname{Ker} \varphi$. Umgekehrt gilt für $v \in \operatorname{Ker} \varphi$ stets $\varphi(v) = 0$, also auch

$$(\varphi^{\mathrm{ad}} \circ \varphi)(v) = \varphi^{\mathrm{ad}}(\varphi(v)) = \varphi^{\mathrm{ad}}(0) = 0$$

und damit $v \in \operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi)$. Insgesamt ist in der Tat $\operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi) = \operatorname{Ker} \varphi$.
Wegen $\mathrm{pr}_{\operatorname{Im} \varphi}(w) \in \operatorname{Im} \varphi$ gibt es ein $v \in V$ mit $\mathrm{pr}_{\operatorname{Im} \varphi}(w) = \varphi(v)$. Nach Korollar (7.61) ist $(\operatorname{Im} \varphi)^{\perp} = \operatorname{Ker} \varphi^{\mathrm{ad}}$, also insbesondere $\varphi^{\mathrm{ad}}(\mathrm{pr}_{(\operatorname{Im} \varphi)^{\perp}}(w)) = 0$ und damit

$$\begin{aligned}\varphi^{\mathrm{ad}}(w) &= \varphi^{\mathrm{ad}}(\mathrm{pr}_{\operatorname{Im} \varphi}(w) + \mathrm{pr}_{(\operatorname{Im} \varphi)^{\perp}}(w)) = \varphi^{\mathrm{ad}}(\mathrm{pr}_{\operatorname{Im} \varphi}(w)) + \varphi^{\mathrm{ad}}(\mathrm{pr}_{(\operatorname{Im} \varphi)^{\perp}}(w)) = \varphi^{\mathrm{ad}}(\mathrm{pr}_{\operatorname{Im} \varphi}(w)) \\ &= \varphi^{\mathrm{ad}}(\varphi(v)).\end{aligned}$$

Nach Proposition (2.15) erhalten wir

$$\begin{aligned}\varphi^{-1}(\{\mathrm{pr}_{\operatorname{Im} \varphi}(w)\}) &= \varphi^{-1}(\{\varphi(v)\}) = v + \operatorname{Ker} \varphi = v + \operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi) = (\varphi^{\mathrm{ad}} \circ \varphi)^{-1}(\{(\varphi^{\mathrm{ad}} \circ \varphi)(v)\}) \\ &= (\varphi^{\mathrm{ad}} \circ \varphi)^{-1}(\{\varphi^{\mathrm{ad}}(\varphi(v))\}) = (\varphi^{\mathrm{ad}} \circ \varphi)^{-1}(\{\varphi^{\mathrm{ad}}(w)\}).\end{aligned}$$ □

(7.63) Korollar. Es seien ein endlichdimensionaler K-Skalarproduktraum V, ein K-Skalarproduktraum W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist

$$\operatorname{rk}_K(\varphi^{\mathrm{ad}} \circ \varphi) = \operatorname{rk}_K \varphi.$$

Beweis. Nach Proposition (7.62) ist $\operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi) = \operatorname{Ker} \varphi$, also insbesondere

$$\operatorname{def}(\varphi^{\mathrm{ad}} \circ \varphi) = \dim \operatorname{Ker}(\varphi^{\mathrm{ad}} \circ \varphi) = \dim(\operatorname{Ker} \varphi) = \operatorname{def} \varphi.$$

Mit dem Rangsatz (2.28) folgt

$$\operatorname{rk}(\varphi^{\mathrm{ad}} \circ \varphi) = \dim V - \operatorname{def}(\varphi^{\mathrm{ad}} \circ \varphi) = \dim V - \operatorname{def} \varphi = \operatorname{rk} \varphi.$$ □

(7.64) Korollar. Es seien ein endlichdimensionaler K-Skalarproduktraum V, ein K-Skalarproduktraum W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Genau dann ist $\varphi^{\mathrm{ad}} \circ \varphi$ ein K-Vektorraumautomorphismus von V, wenn

$$\operatorname{rk}_K \varphi = \dim_K V$$

ist.

Beweis. Nach Korollar (2.29) ist genau dann $\varphi^{\mathrm{ad}} \circ \varphi \in \mathrm{GL}(V)$, wenn $\operatorname{rk}(\varphi^{\mathrm{ad}} \circ \varphi) = \dim V$ ist, und nach Korollar (7.63) gilt stets $\operatorname{rk}(\varphi^{\mathrm{ad}} \circ \varphi) = \operatorname{rk} \varphi$. Insgesamt ist somit genau dann $\varphi^{\mathrm{ad}} \circ \varphi \in \mathrm{GL}(V)$, wenn $\operatorname{rk} \varphi = \dim V$ ist. □

(7.65) Bemerkung. Es seien ein K-Skalarproduktraum V, ein adjungierbarer K-Vektorraumendomorphismus $\varphi\colon V \to V$ und ein K-Untervektorraum U von V gegeben. Wenn U invariant unter φ ist, dann ist $U^{\perp}$ invariant unter φ^{ad}.

Beweis. Wenn U invariant unter φ ist, so gilt $\varphi(U) \subseteq U$, nach Bemerkung (7.60) also auch

$$U^{\perp} \subseteq \varphi(U)^{\perp} = (\varphi^{\mathrm{ad}})^{-1}(U^{\perp})$$

und damit $\varphi^{\mathrm{ad}}(U^{\perp}) \subseteq U^{\perp}$, d.h. $U^{\perp}$ ist invariant unter φ^{ad}. □

(7.66) Korollar. Es seien ein K-Skalarproduktraum V, ein adjungierbarer K-Vektorraumendomorphismus $\varphi\colon V \to V$ und ein K-Untervektorraum U von V mit $V = U \dot{+} U^{\perp}$ gegeben. Genau dann ist U invariant unter φ, wenn $U^{\perp}$ invariant unter φ^{ad} ist.

Beweis. Wenn U invariant unter φ ist, dann ist $U^{\perp}$ invariant unter φ^{ad} nach Bemerkung (7.65). Wegen $V = U \dot{+} U^{\perp}$ gilt jedoch $(U^{\perp})^{\perp} = U$, so dass umgekehrt aus der Invarianz von $U^{\perp}$ unter φ^{ad} auch die Invarianz von $(U^{\perp})^{\perp} = U$ unter $(\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi$ folgt. □

(7.67) Bemerkung. Es seien ein K-Skalarproduktraum V, ein adjungierbarer K-Vektorraumendomorphismus $\varphi\colon V \to V$ und ein K-Untervektorraum U von V so gegeben, dass U invariant unter φ und unter φ^{ad} ist. Dann ist die Restriktion $\varphi|_U^U\colon U \to U$ adjungierbar mit

$$(\varphi|_U^U)^{\mathrm{ad}} = \varphi^{\mathrm{ad}}|_U^U.$$

Beweis. Für $u, u' \in U$ gilt

$$\langle \varphi|_U^U(u), u' \rangle = \langle \varphi(u), u' \rangle = \langle u, \varphi^{\mathrm{ad}}(u') \rangle = \langle u, \varphi^{\mathrm{ad}}|_U^U(u') \rangle.$$

Folglich ist $\varphi|_U^U$ adjungierbar mit $(\varphi|_U^U)^{\mathrm{ad}} = \varphi^{\mathrm{ad}}|_U^U$. □

Nachdem wir bisher Eigenschaften adjungierbarer Homomorphismen studiert haben, wollen wir nun die Verbindung zwischen Adjunktion von Homomorphismen und Matrizen herstellen sowie zeigen, dass jeder Homomorphismus zwischen endlichdimensionalen K-Vektorräumen adjungierbar ist.

(7.68) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Die Spalteninterpretation $\upvarphi_A\colon K^{n \times 1} \to K^{m \times 1}$ ist adjungierbar mit

$$\upvarphi_A^{\mathrm{ad}} = \upvarphi_{A^{\mathrm{ad}}}$$

Beweis. Für $x \in K^{n \times 1}$, $y \in K^{m \times 1}$ gilt

$$\langle \upvarphi_A(x), y \rangle = \langle Ax, y \rangle = (Ax)^{\mathrm{ad}} y = x^{\mathrm{ad}} A^{\mathrm{ad}} y = \langle x, A^{\mathrm{ad}} y \rangle = \langle x, \upvarphi_{A^{\mathrm{ad}}}(y) \rangle.$$

Folglich ist $\upvarphi_A$ adjungierbar mit $\upvarphi_A^{\mathrm{ad}} = \upvarphi_{A^{\mathrm{ad}}}$. □

(7.69) Proposition. Es seien endlichdimensionale K-Skalarprodukträume V und W und ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist φ adjungierbar und für $m, n \in \mathbb{N}_0$, jede Orthonormalbasis $s = (s_1, \dots, s_n)$ von V und jede Orthonormalbasis $t = (t_1, \dots, t_m)$ von W gilt

$$\mathrm{M}_{s,t}(\varphi^{\mathrm{ad}}) = \mathrm{M}_{t,s}(\varphi)^{\mathrm{ad}}.$$

Beweis. Nach Bemerkung (1.63) und Korollar (7.46) gibt es $m, n \in \mathbb{N}_0$, eine Orthonormalbasis $s = (s_1, \dots, s_n)$ von V und eine Orthonormalbasis $t = (t_1, \dots, t_m)$ von W. Ferner ist $\mathrm{M}_{s,t}\colon \mathrm{Hom}_K(W, V) \to K^{n \times m}$ nach Proposition (3.14) insbesondere surjektiv, so dass es einen K-Vektorraumhomomorphismus $\psi\colon W \to V$ mit $\mathrm{M}_{t,s}(\varphi)^{\mathrm{ad}} = \mathrm{M}_{s,t}(\psi)$ gibt. Mit Korollar (7.34), Proposition (3.11) und Bemerkung (7.68) erhalten wir

$$\begin{aligned}\langle \varphi(v), w \rangle &= \langle \upkappa_t(\varphi(v)), \upkappa_t(w) \rangle = \langle \mathrm{M}_{t,s}(\varphi)\, \upkappa_s(v), \upkappa_t(w) \rangle = \langle \upkappa_s(v), \mathrm{M}_{t,s}(\varphi)^{\mathrm{ad}}\, \upkappa_t(w) \rangle = \langle \upkappa_s(v), \mathrm{M}_{s,t}(\psi)\, \upkappa_t(w) \rangle \\ &= \langle \upkappa_s(v), \upkappa_s(\psi(w)) \rangle = \langle v, \psi(w) \rangle\end{aligned}$$

für $v \in V$, $w \in W$, d.h. φ ist adjungierbar mit $\varphi^{\mathrm{ad}} = \psi$.
Nun seien umgekehrt $m, n \in \mathbb{N}_0$, eine beliebige Orthonormalbasis $s = (s_1, \dots, s_n)$ von V und eine beliebige Orthonormalbasis $t = (t_1, \dots, t_m)$ von W gegeben. Für $v, w \in V$ gilt dann

$$\begin{aligned}\langle \mathrm{M}_{t,s}(\varphi)\, \upkappa_s(v), \upkappa_t(w) \rangle &= \langle \upkappa_t(\varphi(v)), \upkappa_t(w) \rangle = \langle \varphi(v), w \rangle = \langle v, \varphi^{\mathrm{ad}}(w) \rangle = \langle \upkappa_s(v), \upkappa_s(\psi(w)) \rangle \\ &= \langle \upkappa_s(v), \mathrm{M}_{s,t}(\varphi^{\mathrm{ad}})\, \upkappa_t(w) \rangle\end{aligned}$$

nach Proposition (3.11) und Korollar (7.34). Da $\upkappa_s\colon V \to K^{n \times 1}$ und $\upkappa_t\colon W \to K^{m \times 1}$ Bijektionen sind, erhalten wir $\langle \mathrm{M}_{t,s}(\varphi)\, x, y \rangle = \langle x, \mathrm{M}_{s,t}(\varphi^{\mathrm{ad}})\, y \rangle$ für $x \in K^{n \times 1}$, $y \in K^{m \times 1}$. Nach Bemerkung (7.68) ist also

$$\upvarphi_{\mathrm{M}_{s,t}(\varphi^{\mathrm{ad}})} = \upvarphi_{\mathrm{M}_{t,s}(\varphi)}^{\mathrm{ad}} = \upvarphi_{\mathrm{M}_{t,s}(\varphi)^{\mathrm{ad}}}.$$

Da jedoch auch $\upvarphi_-\colon K^{m \times n} \to \mathrm{Hom}_K(K^{n \times 1}, K^{m \times 1})$ eine Bijektion ist, impliziert dies bereits

$$\mathrm{M}_{s,t}(\varphi^{\mathrm{ad}}) = \mathrm{M}_{t,s}(\varphi)^{\mathrm{ad}}.$$ □

Beste Approximation

Als nächstes studieren wir den Begriff der besten Approximation, welcher zu Anwendungen etwa bei Suchmaschinen (Term-Dokument-Matrizen, siehe Definition (7.78)), in der Ausgleichsrechnung („Methode der kleinsten Fehlerquadrate“, siehe Definition (7.84) und Definition (7.90)) oder auch in der Datenkompression führt.

(7.70) Definition (beste Approximation)**.** Es seien ein normierter K-Vektorraum V, eine Teilmenge S von V und ein $v \in V$ gegeben. Die *Menge der besten Approximationen* (oder *Menge der Proxima*) von v in S ist definiert als

$$\mathrm{Prox}_S(v) = \{s \in S \mid \|v - s\| = \min\{\|v - s'\| \mid s' \in S\}\}.$$

Ein Element von $\mathrm{Prox}_S(v)$ wird *beste Approximation* (oder *Proximum*) von v in S genannt.

Setzt man in Definition (7.70) für $s' \in S$ stets $\mathrm{d}(s', v) := \|v - s'\|$, so erkennt man, dass der Begriff einer besten Approximation in Analogie zum Begriff des nächsten Nachbarn (4.15) definiert ist.

(7.71) Beispiel.

(a) Es seien eine Teilmenge S von $K^{3\times 1}$ und $x \in K^{3\times 1}$ gegeben durch

$$S = \{\mathrm{e}_1, \mathrm{e}_2\},\ x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}.$$

Dann ist e_2 eine beste Approximation von x in S.

(b) Es seien ein K-Untervektorraum U von $K^{3\times 1}$ sowie $x \in K^{3\times 1}$, $u \in U$ gegeben durch

$$U = K\mathrm{e}_1 + K\mathrm{e}_2,\ x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix},\ u = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

Dann ist u eine beste Approximation von x in U.

Beweis.

(a) Wegen

$$\|x - \mathrm{e}_1\| = \| \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \| = \| \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} \| = \sqrt{1^2 + 3^2 + (-1)^2} = \sqrt{11},$$
$$\|x - \mathrm{e}_2\| = \| \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \| = \| \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix} \| = \sqrt{2^2 + 2^2 + (-1)^2} = 3$$

ist $\|x - \mathrm{e}_2\| = 3 < \sqrt{11} = \|x - \mathrm{e}_1\|$ und damit

$$\min\{\|x - s\| \mid s \in S\} = \min\{\|x - \mathrm{e}_1\|, \|x - \mathrm{e}_2\|\} = \|x - \mathrm{e}_2\|,$$

d.h. e_2 ist eine beste Approximation von x in S.

(b) Für $u' \in U$ gibt es $a, b \in K$ mit $u' = a\mathrm{e}_1 + b\mathrm{e}_2$, es gilt also

$$\|x - u'\|^2 = \| \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \|^2 = \| \begin{pmatrix} 2 - a \\ 3 - b \\ -1 \end{pmatrix} \|^2 = |2 - a|^2 + |3 - b|^2 + (-1)^2.$$

Da für $a, b \in K$ stets $|2 - a|^2 \geq 0$ und $|3 - b|^2 \geq 0$ sowie genau dann $|2 - a|^2 = 0$ und $|3 - b|^2 = 0$ gilt, wenn $a = 2$ und $b = 3$ ist, folgt

$$\|x - u'\|^2 = |2 - a|^2 + |3 - b|^2 + (-1)^2 \geq |2 - 2|^2 + |3 - 3|^2 + (-1)^2 = \|x - u\|^2$$

für $u' \in U$. Wegen $u \in U$ ist folglich $\|x - u\| = \min\{\|x - u'\| \mid u' \in U\}$, d.h. u ist eine beste Approximation von x in U. □

(7.72) Bemerkung. Es seien ein normierter K-Vektorraum V und eine Teilmenge S von V gegeben. Dann ist

$$S = \{v \in V \mid v \in \operatorname{Prox}_S(v)\} = \{v \in V \mid \operatorname{Prox}_S(v) = \{v\}\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Zuerst behandeln wir den Fall einer besten Approximation in einer endlichen Menge von normierten Vektoren:

(7.73) Bemerkung. Es seien ein K-Skalarproduktraum V, $s, t, v \in V$ gegeben und es seien s und t normiert. Genau dann ist $\|v - s\| \leq \|v - t\|$, wenn $\operatorname{Re}\langle s, v\rangle \geq \operatorname{Re}\langle t, v\rangle$ ist.

Beweis. Da s normiert ist, gilt

$$\begin{aligned}\|v - s\|^2 &= \langle v - s, v - s\rangle = \langle v, v\rangle - \langle v, s\rangle - \langle s, v\rangle + \langle s, s\rangle = \|v\|^2 - \langle v, s\rangle - \overline{\langle v, s\rangle} + \|s\|^2 \\ &= \|v\|^2 - 2\operatorname{Re}\langle s, v\rangle + 1,\end{aligned}$$

und da t normiert ist, gilt analog

$$\|v - t\|^2 = \|v\|^2 - 2\operatorname{Re}\langle t, v\rangle + 1.$$

Somit gilt genau dann $\|v - s\| \leq \|v - t\|$, wenn

$$\|v\|^2 - 2\operatorname{Re}\langle s, v\rangle + 1 = \|v - s\|^2 \leq \|v - t\|^2 = \|v\|^2 - 2\operatorname{Re}\langle t, v\rangle + 1$$

gilt, also genau dann, wenn $\operatorname{Re}\langle s, v\rangle \geq \operatorname{Re}\langle t, v\rangle$ ist. □

(7.74) Korollar. Es seien ein K-Skalarproduktraum V und eine endliche Teilmenge S von V derart gegeben, dass jedes Element von S normiert ist. Ferner sei $v \in V$ gegeben. Dann ist

$$\operatorname{Prox}_S(v) = \{s \in S \mid \operatorname{Re}\langle s, v\rangle = \max\{\operatorname{Re}\langle s', v\rangle \mid s' \in S\}\}.$$

Alternativer Beweis von Beispiel (7.71)(a). Wegen

$$\begin{aligned}\langle \mathrm{e}_1, x\rangle &= \langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\3\\-1\end{pmatrix}\rangle = 1 \cdot 2 + 0 \cdot 3 + 0 \cdot (-1) = 2, \\ \langle \mathrm{e}_2, x\rangle &= \langle \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}2\\3\\-1\end{pmatrix}\rangle = 0 \cdot 2 + 1 \cdot 3 + 0 \cdot (-1) = 3\end{aligned}$$

ist $\langle \mathrm{e}_2, x\rangle = 3 > 2 = \langle \mathrm{e}_1, x\rangle$ und damit

$$\max\{\langle s, x\rangle \mid s \in S\} = \max\{\langle \mathrm{e}_1, x\rangle, \langle \mathrm{e}_2, x\rangle\} = \langle \mathrm{e}_2, x\rangle.$$

Nach Korollar (7.74) ist e_2 eine beste Approximation von x in S. □

Als nächstes studieren wir beste Approximationen in einem Untervektorraum, welcher eine Orthogonalprojektion zulässt.

(7.75) Proposition. Es seien ein K-Skalarproduktraum V, ein K-Untervektorraum U von V mit $V = U \dot{+} U^\perp$ und ein $v \in V$ gegeben. Dann ist $\operatorname{pr}_U(v)$ die eindeutige beste Approximation von v in U.

Beweis. Wegen $\operatorname{pr}_{U^\perp}(v) \in U^\perp$ gilt für $u \in U$ stets $\operatorname{pr}_U(v) - u \perp \operatorname{pr}_{U^\perp}(v) = v - \operatorname{pr}_U(v)$ und damit

$$\|v - u\|^2 = \|v - \operatorname{pr}_U(v) + \operatorname{pr}_U(v) - u\|^2 = \|v - \operatorname{pr}_U(v)\|^2 + \|\operatorname{pr}_U(v) - u\|^2$$

nach dem Satz des Pythagoras (7.23).
Folglich erhalten wir einerseits $\|v - \operatorname{pr}_U(v)\| \leq \|v - u\|$ für alle $u \in U$ und damit

$$\|v - \operatorname{pr}_U(v)\| = \min\{\|v - u\| \mid u \in U\},$$

d.h. $\operatorname{pr}_U(v)$ ist eine beste Approximation von v in U. Andererseits gilt für jede beliebige beste Approximation u von v in U insbesondere $\|v - u\| \leq \|v - \operatorname{pr}_U(v)\|$, also

$$\|\operatorname{pr}_U(v) - u\|^2 = \|v - u\|^2 - \|v - \operatorname{pr}_U(v)\|^2 \leq 0$$

und damit $u = \operatorname{pr}_U(v)$. Insgesamt ist $\operatorname{pr}_U(v)$ die eindeutige beste Approximation von v in U. □

(7.76) Beispiel. Es seien ein K-Untervektorraum U von $K^{3\times 1}$ und $x \in K^{3\times 1}$ gegeben durch

$$U = K \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} + K \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \, x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die eindeutige beste Approximation von x in U ist

$$\frac{1}{9} \begin{pmatrix} 7 \\ 8 \\ 11 \end{pmatrix}.$$

Beweis. Nach Proposition (7.75) und Beispiel (7.44) ist

$$\mathrm{pr}_U(x) = \frac{1}{9} \begin{pmatrix} 7 \\ 8 \\ 11 \end{pmatrix}$$

die eindeutige beste Approximation von x in U. □

Alternativer Beweis von Beispiel (7.71)(b). Nach Proposition (7.75) ist $\mathrm{pr}_U(x)$ die eindeutige beste Approximation von x in U und nach Beispiel (7.30) gilt

$$\mathrm{pr}_U(x) = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = u.$$ □

Wir erläutern, wie sich der Begriff der besten Approximation für Anwendungen nutzen lässt: Es sei eine Abbildung $f \colon X \to Y$ und ein $y \in Y$ gegeben. Die Elemente der Faser

$$f^{-1}(\{y\}) = \{x \in X \mid f(x) = y\}$$

lassen sich dann als *(exakte) Lösungen* zu den gegebenen Daten (f, y) auffassen. Bspw. ist für $m, n \in \mathbb{N}_0$, $A \in K^{m\times n}$, $b \in K^{m\times 1}$ die Lösungsmenge des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ nach Bemerkung (3.35) gegeben durch

$$\mathrm{Sol}(A, b) = \varphi_A^{-1}(\{b\}).$$

Ein weiteres Beispiel liefern Interpolationen, siehe Definition (3.45).
Genau dann gibt es eine (exakte) Lösung zu (f, y), wenn $y \in \operatorname{Im} f$ ist, vgl. bspw. das Lösbarkeitskriterium für lineare Gleichungssysteme (3.41).
Nun seien ein normierter K-Vektorraum W und eine Abbildung $f \colon X \to W$ gegeben. Nach Bemerkung (7.72) ist

$$\operatorname{Im} f = \{w \in W \mid \mathrm{Prox}_{\operatorname{Im} f}(w) = \{w\}\}.$$

Folglich gibt es für $w \in W$ genau dann eine (exakte) Lösung zu (f, w), wenn $\mathrm{Prox}_{\operatorname{Im} f}(w) = \{w\}$ ist, und in diesem Fall ist die Menge der (exakten) Lösungen zu (f, w) gegeben durch

$$f^{-1}(\{w\}) = f^{-1}(\mathrm{Prox}_{\operatorname{Im} f}(w)).$$

Für $w \in W$ werden daher im Allgemeinen die Elemente von $f^{-1}(\mathrm{Prox}_{\operatorname{Im} f}(w))$ als *(beste) Approximationslösungen* zu (f, w) aufgefasst. Vgl. Bemerkung (7.81), Definition (7.84) und Definition (7.90).

(7.77) Bemerkung. Es seien ein normierter K-Vektorraum W, eine Abbildung $f \colon X \to W$ und ein $w \in W$ gegeben. Dann ist

$$\begin{aligned} f^{-1}(\mathrm{Prox}_{\operatorname{Im} f}(w)) &= \{x \in X \mid f(x) \text{ ist eine beste Approximation von } w \text{ in } \operatorname{Im} f\} \\ &= \{x \in X \mid \|w - f(x)\| = \min\{\|w - f(x')\| \mid x' \in X\}\} \\ &= \{x \in X \mid \|w - f(x)\|^2 = \min\{\|w - f(x')\|^2 \mid x' \in X\}\}. \end{aligned}$$

Beweis. Da $\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $x \mapsto x^2$ eine streng monoton steigende Bijektion ist, gilt

$$\begin{aligned}
f^{-1}(\operatorname{Prox}_{\operatorname{Im} f}(w)) &= \{x \in X \mid f(x) \in \operatorname{Prox}_{\operatorname{Im} f}(\{w\})\} \\
&= \{x \in X \mid f(x) \text{ ist eine beste Approximation von } w \text{ in } \operatorname{Im} f\} \\
&= \{x \in X \mid \|w - f(x)\| = \min\{\|w - f(x')\| \mid x' \in X\}\} \\
&= \{x \in X \mid \|w - f(x)\|^2 = \min\{\|w - f(x')\|^2 \mid x' \in X\}\}.
\end{aligned}$$

□

Anwendung: Term-Dokument-Matrizen

Eine Anwendung der besten Approximation tritt in einem möglichen Modell zur Arbeitsweise von Suchmaschinen auf. Hierzu werden die Häufigkeiten der möglichen zu suchenden Terme in der zu durchsuchenden Dokumenten in einer sogenannten Term-Dokument-Matrix organisiert. Hierbei entsprechen die Terme den Zeilen (weswegen wir der Einfachheit halber annehmen, dass diese durch die Zeilenindizes gegeben sind) und die Dokumente den Spalten der Term-Dokument-Matrix. Damit längere Dokumente nicht schon auf Grund der Länge von der Suchmaschine eine Bevorzugung erhalten, führt man noch einen Normierungsprozess der Spalten durch und geht zur sogenannten normierten Term-Dokument-Matrix über. Dieser Vorgang ist also analog zu dem bei der Linkmatrix, siehe Definition (6.32), der Unterschied besteht in der verwendeten Norm: Während bei der Linkmatrix die Summe der Einträge der jeweiligen Spalte normiert wird, vgl. Beispiel (7.9)(a), verwenden wir hier die durch das Skalarprodukt induzierte Norm.
Es sei $m \in \mathbb{N}_0$ gegeben. Ein *String* in $[1, m]$ ist ein k-Tupel $(i_1, \ldots, i_k)$ in $[1, m]$ für ein $k \in \mathbb{N}_0$, siehe [12, Def. (10.2)].

(7.78) Definition (Term-Dokument-Matrix)**.** Es seien $m, n \in \mathbb{N}_0$ und ein n-Tupel $(D_1, \ldots, D_n)$ von Strings in $[1, m]$ gegeben.

(a) Für $j \in [1, n]$ sei l_j die Länge von D_j. Die Matrix $A' \in \mathbb{R}^{m \times n}$ gegeben durch

$$A'_{i,j} = |\{k \in [1, l_j] \mid (D_j)_k = i\}|$$

für $i \in [1, m]$, $j \in [1, n]$ heißt *Term-Dokument-Matrix* von $(D_1, \ldots, D_n)$.

(b) Es sei A' die Term-Dokument-Matrix von $(D_1, \ldots, D_n)$. Die Matrix $A \in \mathbb{R}^{m \times n}$ gegeben durch

$$A_{-,j} = \frac{1}{\|A'_{-,j}\|} A'_{-,j}$$

für $j \in [1, n]$ heißt *normierte Term-Dokument-Matrix* von $(D_1, \ldots, D_n)$.

Die Spalten der Term-Dokument-Matrix sind also *Häufigkeitsvektoren* der Dokumente, vgl. [12, Def. (18.21)].

(7.79) Beispiel. Es sei $(D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8)$ in $[1, 6]$ gegeben durch

$$\begin{aligned}
&D_1 = (1, 2, 3, 1, 6, 2), \quad D_2 = (4, 2, 1, 2), \quad D_3 = (3, 6, 6, 3), \quad D_4 = (1, 5, 2, 2, 5), \\
&D_5 = (3, 5, 1, 2, 1), \quad D_6 = (1, 4, 2, 4, 4, 2), \quad D_7 = (3, 5), \quad D_8 = (5, 2, 4, 2, 6).
\end{aligned}$$

(a) Die Term-Dokument-Matrix A' zu $(D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8)$ ist gegeben durch

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

(b) Die normierte Term-Dokument-Matrix A zu $(D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8)$ ist gegeben durch

$$A = \begin{pmatrix} \frac{2}{\sqrt{10}} & \frac{1}{\sqrt{6}} & 0 & \frac{1}{3} & \frac{2}{\sqrt{7}} & \frac{1}{\sqrt{14}} & 0 & 0 \\ \frac{2}{\sqrt{10}} & \frac{2}{\sqrt{6}} & 0 & \frac{2}{3} & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{14}} & 0 & \frac{2}{\sqrt{7}} \\ \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{7}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 & 0 & 0 & \frac{3}{\sqrt{14}} & 0 & \frac{1}{\sqrt{7}} \\ 0 & 0 & 0 & \frac{2}{3} & \frac{1}{\sqrt{7}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{7}} \\ \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.632 & 0.408 & 0.000 & 0.333 & 0.756 & 0.267 & 0.000 & 0.000 \\ 0.632 & 0.816 & 0.000 & 0.667 & 0.378 & 0.535 & 0.000 & 0.756 \\ 0.316 & 0.000 & 0.707 & 0.000 & 0.378 & 0.000 & 0.707 & 0.000 \\ 0.000 & 0.408 & 0.000 & 0.000 & 0.000 & 0.802 & 0.000 & 0.378 \\ 0.000 & 0.000 & 0.000 & 0.667 & 0.378 & 0.000 & 0.707 & 0.378 \\ 0.316 & 0.000 & 0.707 & 0.000 & 0.000 & 0.000 & 0.000 & 0.378 \end{pmatrix}.$$

Der Vergleich einer Suchanfrage mit den Spalten der normierten Term-Dokument-Matrix besteht nun in der Ermittlung einer besten Approximation des sogenannten Indikatorvektors zu den gesuchten Termen in der Menge der Spalten.
Es seien $m \in \mathbb{N}_0$ und eine Teilmenge I von $[1, m]$ gegeben. Der *Indikatorvektor* von I ist $\chi_I \in \mathbb{R}^{m \times 1}$ gegeben durch

$$\chi_{I,i} = (\chi_I)_i = \begin{cases} 1 & \text{für } i \in I, \\ 0 & \text{für } i \in [1, m] \setminus I, \end{cases}$$

vgl. [12, Def. (3.34)].

(7.80) Definition (Suchanfrage)**.** Es seien $m, n \in \mathbb{N}_0$, ein n-Tupel $(D_1, \ldots, D_n)$ von Strings in $[1, m]$ und eine Teilmenge I von $[1, m]$ gegeben und es sei A die normierte Term-Dokument-Matrix von $(D_1, \ldots, D_n)$. Für $j \in [1, n]$ sagen wir, dass D_j am *Besten* zur *Suchanfrage I passt*, wenn $A_{-,j}$ eine beste Approximation des Indikatorvektors χ_I in $\{A_{-,1}, \ldots, A_{-,n}\}$ ist.

(7.81) Bemerkung. Es seien $m, n \in \mathbb{N}_0$, ein n-Tupel $(D_1, \ldots, D_n)$ von Strings in $[1, m]$ und eine Teilmenge I von $[1, m]$ gegeben und es sei A die normierte Term-Dokument-Matrix von $(D_1, \ldots, D_n)$ und

$$f \colon [1, n] \to K^{m \times 1},\ j \mapsto A_{-,j}.$$

Für $j \in [1, n]$ gilt: Genau dann passt D_j am Besten zur Suchanfrage I, wenn $j \in f^{-1}(\mathrm{Prox}_{\mathrm{Im}\, f}(\chi_I))$ ist.

(7.82) Bemerkung. Es seien $m, n \in \mathbb{N}_0$, ein n-Tupel $(D_1, \ldots, D_n)$ von Strings in $[1, m]$, eine Teilmenge I von $[1, m]$ und ein $j \in [1, n]$ gegeben und es sei A die normierte Term-Dokument-Matrix von $(D_1, \ldots, D_n)$. Genau dann passt D_j am Besten zur Suchanfrage I, wenn

$$\sum_{i \in I} A_{i,j} = \max\{\sum_{i \in I} A_{i,j'} \mid j' \in [1, n]\}$$

ist.

Beweis. Genau dann passt D_j am Besten zur Suchanfrage I, wenn $A_{-,j}$ eine beste Approximation von χ_I in $\{A_{-,j'} \mid j' \in [1, n]\}$ ist. Nach Korollar (7.74) ist dies äquivalent zu

$$\langle A_{-,j}, \chi_I \rangle = \max\{\langle A_{-,j'}, \chi_I \rangle \mid j' \in [1, n]\}.$$

Da für $j' \in [1, n]$ jedoch

$$\langle A_{-,j'}, \chi_I \rangle = \sum_{i \in [1,m]} A_{i,j'} \chi_{I,i} = \sum_{i \in I} A_{i,j'}$$

gilt, passt D_j folglich genau dann am Besten zur Suchanfrage I, wenn

$$\sum_{i \in I} A_{i,j} = \max\{\sum_{i \in I} A_{i,j'} \mid j' \in [1, n]\}$$

ist. □

(7.83) Beispiel. Es sei $(D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8)$ in $[1, 6]$ gegeben durch

$$\begin{array}{llll} D_1 = (1, 2, 3, 1, 6, 2), & D_2 = (4, 2, 1, 2), & D_3 = (3, 6, 6, 3), & D_4 = (1, 5, 2, 2, 5), \\ D_5 = (3, 5, 1, 2, 1), & D_6 = (1, 4, 2, 4, 4, 2), & D_7 = (3, 5), & D_8 = (5, 2, 4, 2, 6). \end{array}$$

(a) Zur Suchanfrage $\{3, 4\}$ passt am Besten D_6.

(b) Zur Suchanfrage $\{1, 3, 4\}$ passt am Besten D_5.

Beweis. Die normierte Term-Dokument-Matrix A zu $(D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8)$ ist gegeben durch

$$A = \begin{pmatrix} \frac{2}{\sqrt{10}} & \frac{1}{\sqrt{6}} & 0 & \frac{1}{3} & \frac{2}{\sqrt{7}} & \frac{1}{\sqrt{14}} & 0 & 0 \\ \frac{2}{\sqrt{10}} & \frac{2}{\sqrt{6}} & 0 & \frac{2}{3} & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{14}} & 0 & \frac{2}{\sqrt{7}} \\ \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{7}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 & 0 & 0 & \frac{3}{\sqrt{14}} & 0 & \frac{1}{\sqrt{7}} \\ 0 & 0 & 0 & \frac{2}{3} & \frac{1}{\sqrt{7}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{7}} \\ \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} \end{pmatrix}.$$

(a) Wegen

$$\begin{aligned} \sum_{i\in\{3,4\}} A_{i,1} &= \frac{1}{\sqrt{10}} \approx 0.316, \\ \sum_{i\in\{3,4\}} A_{i,2} &= \frac{1}{\sqrt{6}} \approx 0.408, \\ \sum_{i\in\{3,4\}} A_{i,3} &= \frac{1}{\sqrt{2}} \approx 0.707, \\ \sum_{i\in\{3,4\}} A_{i,4} &= 0 \approx 0.000, \\ \sum_{i\in\{3,4\}} A_{i,5} &= \frac{1}{\sqrt{7}} \approx 0.378, \\ \sum_{i\in\{3,4\}} A_{i,6} &= \frac{3}{\sqrt{14}} \approx 0.802, \\ \sum_{i\in\{3,4\}} A_{i,7} &= \frac{1}{\sqrt{2}} \approx 0.707, \\ \sum_{i\in\{3,4\}} A_{i,8} &= \frac{1}{\sqrt{7}} \approx 0.378 \end{aligned}$$

ist

$$\max\{\sum_{i\in\{3,4\}} A_{i,j'} \mid j' \in [1, n]\} = \max\{\frac{1}{\sqrt{10}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{7}}, \frac{3}{\sqrt{14}}, \frac{1}{\sqrt{7}}\} = \frac{3}{\sqrt{14}} = \sum_{i\in\{3,4\}} A_{i,6}.$$

Nach Bemerkung (7.82) passt D_6 am Besten zur Suchanfrage $\{3, 4\}$.

(b) Wegen

$$\begin{aligned} \sum_{i\in\{1,3,4\}} A_{i,1} &= \frac{3}{\sqrt{10}} \approx 0.949, \\ \sum_{i\in\{1,3,4\}} A_{i,2} &= \frac{2}{\sqrt{6}} \approx 0.816, \\ \sum_{i\in\{1,3,4\}} A_{i,3} &= \frac{1}{\sqrt{2}} \approx 0.707, \\ \sum_{i\in\{1,3,4\}} A_{i,4} &= \frac{1}{3} \approx 0.333, \\ \sum_{i\in\{1,3,4\}} A_{i,5} &= \frac{3}{\sqrt{7}} \approx 1.134, \end{aligned}$$

$$\sum_{i \in \{1,3,4\}} A_{i,6} = \frac{4}{\sqrt{14}} \approx 1.069,$$
$$\sum_{i \in \{1,3,4\}} A_{i,7} = \frac{1}{\sqrt{2}} \approx 0.707,$$
$$\sum_{i \in \{1,3,4\}} A_{i,8} = \frac{1}{\sqrt{7}} \approx 0.378$$

ist

$$\max\{\sum_{i \in \{1,3,4\}} A_{i,j'} \mid j' \in [1,n]\} = \max\{\frac{3}{\sqrt{10}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{2}}, \frac{1}{3}, \frac{3}{\sqrt{7}}, \frac{4}{\sqrt{14}}, \frac{1}{\sqrt{7}}\} = \frac{3}{\sqrt{7}} = \sum_{i \in \{1,3,4\}} A_{i,5}.$$

Nach Bemerkung (7.82) passt D_5 am Besten zur Suchanfrage $\{1, 3, 4\}$. □

Anwendung: Beste Näherungslösungen für lineare Gleichungssysteme

Nach dem Lösbarkeitskriterium für lineare Gleichungssysteme (3.41) ist für gegebene $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$, $b \in K^{m \times 1}$ das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ genau dann lösbar, wenn $b \in \mathrm{Col}(A)$ ist. Erhalten wir die Daten eines solchen linearen Gleichungssystems durch experimentelle Messungen, so können kleine Mess- oder Rundungsfehler in der Praxis zur Unlösbarkeit des linearen Gleichungssystems führen. In diesem Fall sind wir an bestmöglichen Spalten $x \in K^{n \times 1}$ mit $Ax \approx b$ interessiert, in folgendem präzisen Sinn:

(7.84) Definition (beste Näherungslösung eines linearen Gleichungssystems)**.** Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Die *Menge der besten Näherungslösungen des linearen Gleichungssystems* zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ ist definiert als

$$\mathrm{ApprSol}(A, b) = \varphi_A^{-1}(\mathrm{Prox}_{\mathrm{Col}(A)}(b)).$$

Ein Element von $\mathrm{ApprSol}(A, b)$ wird *beste Näherungslösung des linearen Gleichungssystems* zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ genannt.

Zu gegebenen $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$, $b \in K^{m \times 1}$ ist eine beste Näherungslösung des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ also ein solches Element von $K^{n \times 1}$, dessen Bild unter der Spalteninterpretation

$$\varphi_A \colon K^{n \times 1} \to K^{m \times 1},\ x \mapsto Ax$$

optimal im Sinne einer besten Approximation von b in $\mathrm{Col}(A)$ ist.
Die definierende Bedingung einer besten Approximation bzgl. der vom Standardskalarprodukt induzierten Norm liefert nachfolgende explizite Beschreibung einer besten Näherungslösung eines linearen Gleichungssystems als eine Spalte, welche die „Fehlerquadratsumme“ minimiert. Bei der Bestimmung einer solchen besten Näherungslösung spricht man daher auch von der *Methode der kleinsten Fehlerquadrate*.

(7.85) Bemerkung. Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Die Menge der besten Näherungslösungen des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ ist gegeben durch

$$\begin{aligned} \mathrm{ApprSol}(A, b) &= \{x \in K^{n \times 1} \mid Ax \text{ ist eine beste Approximation von } b \text{ in } \mathrm{Col}(A)\} \\ &= \{x \in K^{n \times 1} \mid \sum_{i \in [1,m]} |b_i - Ax_i|^2 = \min\{\sum_{i \in [1,m]} |b_i - Ax'_i|^2 \mid x' \in K^{n \times 1}\}\}. \end{aligned}$$

Beweis. Nach Bemerkung (7.77) ist

$$\begin{aligned} \mathrm{ApprSol}(A, b) = \varphi_A^{-1}(\mathrm{Prox}_{\mathrm{Col}(A)}(b)) &= \{x \in K^{n \times 1} \mid \varphi_A(x) \text{ ist eine beste Approximation von } b \text{ in } \mathrm{Im}\,\varphi_A\} \\ &= \{x \in K^{n \times 1} \mid Ax \text{ ist eine beste Approximation von } b \text{ in } \mathrm{Col}(A)\} \end{aligned}$$

und nach Bemerkung (7.77) und Beispiel (7.11) ist

$$\begin{aligned}\operatorname{ApprSol}(A,b) &= \varphi_A^{-1}(\operatorname{Prox}_{\operatorname{Col}(A)}(b)) = \{x \in K^{n\times 1} \mid \|b - \varphi_A(x)\|^2 = \min\{\|b - \varphi_A(x')\|^2 \mid x' \in K^{n\times 1}\}\} \\ &= \{x \in K^{n\times 1} \mid \|b - Ax\|^2 = \min\{\|b - Ax'\|^2 \mid x' \in K^{n\times 1}\}\} \\ &= \{x \in K^{n\times 1} \mid \sum_{i\in[1,m]} |(b - Ax)_i|^2 = \min\{\sum_{i\in[1,m]} |(b - Ax')_i|^2 \mid x' \in K^{n\times 1}\}\} \\ &= \{x \in K^{n\times 1} \mid \sum_{i\in[1,m]} |b_i - Ax_i|^2 = \min\{\sum_{i\in[1,m]} |b_i - Ax'_i|^2 \mid x' \in K^{n\times 1}\}\}.\end{aligned}$$ □

(7.86) Korollar. Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m\times n}$ und $b \in K^{m\times 1}$ gegeben. Dann ist

$$\operatorname{ApprSol}(A,b) = \operatorname{Sol}(A, \operatorname{pr}_{\operatorname{Col}(A)}(b)) = \operatorname{Sol}(A^{\mathrm{ad}}A, A^{\mathrm{ad}}b).$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.87) Beispiel.

(a) Es seien $A \in \mathbb{R}^{3\times 2}$ und $b \in \mathbb{R}^{3\times 1}$ gegeben durch

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},\ b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die Menge der besten Näherungslösungen des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\left(A \mid b\right)$ ist gegeben durch

$$\operatorname{ApprSol}(A,b) = \{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\}.$$

(b) Es seien $A \in \mathbb{R}^{3\times 2}$ und $b \in \mathbb{R}^{3\times 1}$ gegeben durch

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 2 & 1 \end{pmatrix},\ b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die Menge der besten Näherungslösungen des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\left(A \mid b\right)$ ist gegeben durch

$$\operatorname{ApprSol}(A,b) = \{\frac{1}{9}\begin{pmatrix} 4 \\ 3 \end{pmatrix}\}.$$

Beweis.

(a) Wegen $b = \mathrm{e}_3 \in (\mathbb{R}\mathrm{e}_1 \dot{+} \mathbb{R}\mathrm{e}_2)^\perp = \operatorname{Col}(A)^\perp$ ist $\operatorname{pr}_{\operatorname{Col}(A)}(b) = 0$ und damit

$$\operatorname{ApprSol}(A,b) = \operatorname{Sol}(A, \operatorname{pr}_{\operatorname{Col}(A)}(b)) = \operatorname{Sol}(A, 0) = \{0\}$$

nach Korollar (7.86).

(b) Nach Beispiel (7.44) ist

$$\operatorname{pr}_{\operatorname{Col}(A)}(b) = \frac{1}{9}\begin{pmatrix} 7 \\ 8 \\ 11 \end{pmatrix}.$$

Wir wenden elementare Zeilenoperationen auf die erweiterte Koeffizientenmatrix $\left(A \mid \operatorname{pr}_{\operatorname{Col}(A)}(b)\right)$ an:

$$\left(\begin{array}{cc|c} 1 & 1 & \frac{7}{9} \\ 2 & 0 & \frac{8}{9} \\ 2 & 1 & \frac{11}{9} \end{array}\right) \xmapsto{\mathrm{add}_{3,1,-1}} \left(\begin{array}{cc|c} 1 & 1 & \frac{7}{9} \\ 2 & 0 & \frac{8}{9} \\ 1 & 0 & \frac{4}{9} \end{array}\right) \xmapsto{\mathrm{add}_{2,3,-2} \circ \mathrm{add}_{1,3,-1}} \left(\begin{array}{cc|c} 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \\ 1 & 0 & \frac{4}{9} \end{array}\right) \xmapsto{\mathrm{sw}_{2,3} \circ \mathrm{sw}_{1,3}} \left(\begin{array}{cc|c} 1 & 0 & \frac{4}{9} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{array}\right)$$

Nach Korollar (7.86), Proposition (A.163) und Proposition (A.159) ist

$$\operatorname{ApprSol}(A,b) = \operatorname{Sol}(A, \operatorname{pr}_{\operatorname{Col}(A)}(b)) = \{\frac{1}{9}\begin{pmatrix} 4 \\ 3 \end{pmatrix}\}.$$ □

Alternativer Beweis von Beispiel (7.87).

(a) Nach Korollar (7.86)

$$\operatorname{ApprSol}(A,b) = \operatorname{Sol}(A^{\mathrm{tr}}A, A^{\mathrm{tr}}b) = \operatorname{Sol}(\mathrm{E}_2, 0) = \{0\}.$$

(b) Es ist

$$A^{\mathrm{tr}}A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 3 \\ 3 & 2 \end{pmatrix},$$

$$A^{\mathrm{tr}}b = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}.$$

Wir wenden elementare Zeilenoperationen auf die erweiterte Koeffizientenmatrix $\big(A^{\mathrm{tr}}A \mid A^{\mathrm{tr}}b\big)$ an:

$$\left(\begin{array}{cc|c} 9 & 3 & 5 \\ 3 & 2 & 2 \end{array}\right) \xmapsto{\mathrm{add}_{1,2,-3}} \left(\begin{array}{cc|c} 0 & -3 & -1 \\ 3 & 2 & 2 \end{array}\right) \xmapsto{\mathrm{sw}_{1,2}} \left(\begin{array}{cc|c} 3 & 2 & 2 \\ 0 & -3 & -1 \end{array}\right) \xmapsto{\mathrm{mul}_{2,-\frac{1}{3}}} \left(\begin{array}{cc|c} 3 & 2 & 2 \\ 0 & 1 & \frac{1}{3} \end{array}\right)$$
$$\xmapsto{\mathrm{add}_{1,2,-2}} \left(\begin{array}{cc|c} 3 & 0 & \frac{4}{3} \\ 0 & 1 & \frac{1}{3} \end{array}\right) \xmapsto{\mathrm{mul}_{1,\frac{1}{3}}} \left(\begin{array}{cc|c} 1 & 0 & \frac{4}{9} \\ 0 & 1 & \frac{1}{3} \end{array}\right)$$

Nach Korollar (7.86), Proposition (A.163) und Proposition (A.159) ist

$$\operatorname{ApprSol}(A,b) = \operatorname{Sol}(A^{\mathrm{tr}}A, A^{\mathrm{tr}}b) = \{\frac{1}{9}\begin{pmatrix} 4 \\ 3 \end{pmatrix}\}.$$ □

(7.88) Bemerkung. Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Dann ist

$$\operatorname{Sol}(A,b) = \begin{cases} \operatorname{ApprSol}(A,b), & \text{falls } b \in \operatorname{Col}(A), \\ \emptyset, & \text{falls } b \notin \operatorname{Col}(A). \end{cases}$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.89) Proposition (Eindeutigkeitskriterium für beste Näherungslösungen linearer Gleichungssysteme)**.** Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Genau dann hat das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix $\big(A \mid b\big)$ genau eine beste Näherungslösung, wenn

$$\operatorname{rk}_K A = n$$

ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Anwendung: Ausgleichsfunktionen

Als dritte Anwendung zur besten Approximation betrachten wir eine Methode zur Ermittlung einer Ausgleichsfunktion an eine Menge vorgegebener Paare von „Messwerten“ (a_1, b_1), …, (a_m, b_m). Hierzu nehmen wir im Allgemeinen an, dass wir bereits eine Vorstellung von der „Art“ der Funktion haben, welche wir an die gemessenen Werte anschmiegen wollen, das heißt wir geben uns eine (in der Regel echte) Teilmenge von $\operatorname{Map}(K, K)$ vor. Die Messwerte organisieren wir in einem Paar von Tupeln.

(7.90) Definition (Ausgleichsfunktion)**.** Es seien eine Menge X, eine Teilmenge S von $\operatorname{Map}(X, K)$, $m \in \mathbb{N}_0$, $a \in X^m$ und $b \in K^m$ gegeben. Die *Menge der Ausgleichsfunktionen* zu (a, b) in S ist definiert als

$$\operatorname{Fit}_S(a,b) = (\varepsilon_a|_S)^{-1}(\operatorname{Prox}_{\varepsilon_a(S)}(b)).$$

Ein Element von $\operatorname{Fit}_S(a, b)$ wird *Ausgleichsfunktion* zu (a, b) in S genannt.

Zu einer gegebenen Menge X, einer gegebenen Teilmenge S von $\mathrm{Map}(X, K)$ und gegebenen $m \in \mathbb{N}_0$, $a \in X^m$, $b \in K^m$ ist eine Ausgleichsfunktion zu (a, b) also ein solches Element von S, dessen Bild unter dem Homomorphismus

$$\varepsilon_a \colon \mathrm{Map}(X, K) \to K^m,\ f' \mapsto (f'(a_1), \dots, f'(a_m))$$

optimal im Sinne einer besten Approximation von b in $\varepsilon_a(S)$ ist.
Die definierende Bedingung einer besten Approximation bzgl. der vom Standardskalarprodukt induzierten Norm liefert folgende explizite Beschreibung einer Ausgleichsfunktion als Funktion, welche die „Fehlerquadratsumme“ minimiert:

(7.91) Bemerkung. Es seien eine Menge X, eine Teilmenge S von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$, $a \in X^m$ und $b \in K^m$ gegeben. Die Menge der Ausgleichsfunktionen zu (a, b) in S ist gegeben durch

$$\begin{aligned}&\mathrm{Fit}_S(a, b)\\ &= \{f \in S \mid (f(a_1), \dots, f(a_m)) \text{ ist eine beste Approximation von } b \text{ in } \{(f'(a_1), \dots, f'(a_m)) \mid f' \in S\}\}\\ &= \{f \in S \mid \sum_{i \in [1,m]} |b_i - f(a_i)|^2 = \min\{\sum_{i \in [1,m]} |b_i - f'(a_i)|^2 \mid f' \in S\}\}.\end{aligned}$$

Beweis. Nach Bemerkung (7.77) ist

$$\begin{aligned}&\mathrm{Fit}_S(a, b) = (\varepsilon_a|_S)^{-1}(\mathrm{Prox}_{\varepsilon_a(S)}(b)) = \{f \in S \mid \varepsilon_a(f) \text{ ist eine beste Approximation von } b \text{ in } \mathrm{Im}\,\varepsilon_a\}\\ &= \{f \in S \mid (f(a_1), \dots, f(a_m)) \text{ ist eine beste Approximation von } b \text{ in } \{(f'(a_1), \dots, f'(a_m)) \mid f' \in S\}\}\end{aligned}$$

und nach Bemerkung (7.77) und Beispiel (7.11) ist

$$\begin{aligned}&\mathrm{Fit}_S(a, b) = (\varepsilon_a|_S)^{-1}(\mathrm{Prox}_{\varepsilon_a(S)}(b)) = \{f \in S \mid \|b - \varepsilon_a(f)\|^2 = \min\{\|b - \varepsilon_a(f')\|^2 \mid f' \in S\}\}\\ &= \{f \in S \mid \|(b_1 - f(a_1), \dots, b_m - f(a_m))\|^2 = \min\{\|(b_1 - f'(a_1), \dots, b_m - f'(a_m))\|^2 \mid f' \in S\}\}\\ &= \{f \in S \mid \sum_{i \in [1,m]} |b_i - f(a_i)|^2 = \min\{\sum_{i \in [1,m]} |b_i - f'(a_i)|^2 \mid f' \in S\}\}.\end{aligned}$$

□

(7.92) Bemerkung. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m \in \mathbb{N}_0$, $a \in X^m$ und $b \in K^m$ gegeben und es sei

$$U' := \{(f'(a_1), \dots, f'(a_m)) \mid f' \in U\}.$$

Dann ist

$$\mathrm{Fit}_U(a, b) = \{f \in U \mid (f(a_1), \dots, f(a_m)) = \mathrm{pr}_{U'}(b)\}.$$

Beweis. Nach Bemerkung (3.44) ist

$$\varepsilon_a \colon \mathrm{Map}(X, K) \to K^m,\ f \mapsto (f(a_1), \dots, f(a_m))$$

ein Homomorphismus und damit $U' = \{(f'(a_1), \dots, f'(a_m)) \mid f' \in U\} = \varepsilon_a(U) = \mathrm{Im}\,\varepsilon_a|_U$ nach Bemerkung (2.12)(a) ein Untervektorraum von K^m. Nach Proposition (7.75) ist daher $\mathrm{pr}_{U'}(b)$ die eindeutige beste Approximation von b in U', so dass mit Bemerkung (7.91) bereits

$$\mathrm{Fit}_U(a, b) = \{f \in U \mid (f(a_1), \dots, f(a_m)) = \mathrm{pr}_{U'}(b)\}$$

folgt. □

(7.93) Beispiel. Es seien $a, b \in \mathbb{R}^5$ gegeben durch

$$a = (-2, -1, 0, 1, 2),\ b = (2, 1, 1, 1, 0).$$

(a) Die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<1}(\mathbb{R}, \mathbb{R})$ ist

$$\mathbb{R} \to \mathbb{R},\ x \mapsto 1.$$

(b) Die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<2}(\mathbb{R}, \mathbb{R})$ ist

$$\mathbb{R} \to \mathbb{R},\ x \mapsto 1 - \frac{2}{5}x.$$

(c) Die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<3}(\mathbb{R}, \mathbb{R})$ ist

$$\mathbb{R} \to \mathbb{R},\ x \mapsto 1 - \frac{2}{5}x.$$

(d) Die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<4}(\mathbb{R}, \mathbb{R})$ ist

$$\mathbb{R} \to \mathbb{R},\ x \mapsto 1 + \frac{1}{6}x - \frac{1}{6}x^3.$$

Beweis. Nach Bemerkung (3.44) ist $\varepsilon_a \colon \mathrm{Map}(\mathbb{R}, \mathbb{R}) \to \mathbb{R}^5$ gegeben durch

$$\varepsilon_a(f) = (f(a_1), f(a_2), f(a_3), f(a_4), f(a_5)) = (f(-2), f(-1), f(0), f(1), f(2))$$

für $f \in \mathrm{Map}(\mathbb{R}, \mathbb{R})$ ein Homomorphismus.
Für $k \in \mathbb{N}_0$ sei $s_k \colon \mathbb{R} \to \mathbb{R},\ x \mapsto x^k$ und

$$t_k := \varepsilon_a(s_k) = (s_k(-2), s_k(-1), s_k(0), s_k(1), s_k(2)) = ((-2)^k, (-1)^k, 0^k, 1^k, 2^k)$$

sowie $U'_k := \varepsilon_a(\mathrm{Pol}_{<k}(\mathbb{R}, \mathbb{R}))$. Dann gilt

$$\begin{aligned} U'_k &= \varepsilon_a(\mathrm{Pol}_{<k}(\mathbb{R}, \mathbb{R})) = \varepsilon_a(\langle s_0, \ldots, s_{k-1}\rangle) = \langle \varepsilon_a(s_0), \ldots, \varepsilon_a(s_{k-1})\rangle = \langle t_0, \ldots, t_{k-1}\rangle \\ &= \langle ((-2)^0, (-1)^0, 0^0, 1^0, 2^0), \ldots, ((-2)^{k-1}, (-1)^{k-1}, 0^{k-1}, 1^{k-1}, 2^{k-1})\rangle \end{aligned}$$

für $k \in \mathbb{N}_0$. Für $k \in \mathbb{N}_0$ ist $\varepsilon_a|_{\mathrm{Pol}_{<k}(\mathbb{R},\mathbb{R})}^{U'_k} \colon U_k \to U'_k$ ein wohldefinierter Isomorphismus; die Details seien dem Leser zur Übung überlassen.
Wir berechnen mit der Gram-Schmidt-Orthogonalisierung (7.42) eine Orthonormalbasis von U'_k für $k \in [1, 4]$:

$$\begin{aligned}
u'_0 &:= t_0 = (1, 1, 1, 1, 1), \\
\langle u'_0, u'_0\rangle &= \langle (1, 1, 1, 1, 1), (1, 1, 1, 1, 1)\rangle = 1^2 + 1^2 + 1^2 + 1^2 + 1^2 = 5, \\
u_0 &:= \frac{1}{\|u'_0\|} u'_0 = \frac{1}{\sqrt{5}} t_0 = \frac{1}{\sqrt{5}}(1, 1, 1, 1, 1), \\
\langle u'_0, t_1\rangle &= \langle (1, 1, 1, 1, 1), (-2, -1, 0, 1, 2)\rangle = 1 \cdot (-2) + 1 \cdot (-1) + 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 2 = 0, \\
u'_1 &:= t_1 - \frac{\langle u'_0, t_1\rangle}{\langle u'_0, u'_0\rangle} u'_0 = t_1 - \frac{0}{5} t_0 = t_1 = (-2, -1, 0, 1, 2), \\
\langle u'_1, u'_1\rangle &= \langle (-2, -1, 0, 1, 2), (-2, -1, 0, 1, 2)\rangle = (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 = 10, \\
u_1 &:= \frac{1}{\|u'_1\|} u'_1 = \frac{1}{\sqrt{10}} t_1 = \frac{1}{\sqrt{10}}(-2, -1, 0, 1, 2), \\
\langle u'_0, t_2\rangle &= \langle (1, 1, 1, 1, 1), (4, 1, 0, 1, 4)\rangle = 1 \cdot 4 + 1 \cdot 1 + 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 4 = 10, \\
\langle u'_1, t_2\rangle &= \langle (-2, -1, 0, 1, 2), (4, 1, 0, 1, 4)\rangle = (-2) \cdot 4 + (-1) \cdot 1 + 0 \cdot 0 + 1 \cdot 1 + 2 \cdot 4 = 0, \\
u'_2 &:= t_2 - \frac{\langle u'_0, t_2\rangle}{\langle u'_0, u'_0\rangle} u'_0 - \frac{\langle u'_1, t_2\rangle}{\langle u'_1, u'_1\rangle} u'_1 = t_2 - \frac{10}{5} t_0 - \frac{0}{10} t_1 = t_2 - 2t_0 = (4, 1, 0, 1, 4) - 2(1, 1, 1, 1, 1) \\
&= (2, -1, -2, -1, 2), \\
\langle u'_2, u'_2\rangle &= \langle (2, -1, -2, -1, 2), (2, -1, -2, -1, 2)\rangle = 2^2 + (-1)^2 + (-2)^2 + (-1)^2 + 2^2 = 14, \\
u_2 &:= \frac{1}{\|u'_2\|} u'_2 = \frac{1}{\sqrt{14}}(t_2 - 2t_0) = \frac{1}{\sqrt{14}} t_2 - \frac{2}{\sqrt{14}} t_0 = \frac{1}{\sqrt{14}}(2, -1, -2, -1, 2), \\
\langle u'_0, t_3\rangle &= \langle (1, 1, 1, 1, 1), (-8, -1, 0, 1, 8)\rangle = 1 \cdot (-8) + 1 \cdot (-1) + 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 8 = 0, \\
\langle u'_1, t_3\rangle &= \langle (-2, -1, 0, 1, 2), (-8, -1, 0, 1, 8)\rangle = (-2) \cdot (-8) + (-1) \cdot (-1) + 0 \cdot 0 + 1 \cdot 1 + 2 \cdot 8 = 34, \\
\langle u'_2, t_3\rangle &= \langle (2, -1, -2, -1, 2), (-8, -1, 0, 1, 8)\rangle = 2 \cdot (-8) + (-1) \cdot (-1) + (-2) \cdot 0 + (-1) \cdot 1 + 2 \cdot 8 = 0,
\end{aligned}$$

$$u'_3 := t_3 - \frac{\langle u'_0, t_3\rangle}{\langle u'_0, u'_0\rangle} u'_0 - \frac{\langle u'_1, t_3\rangle}{\langle u'_1, u'_1\rangle} u'_1 - \frac{\langle u'_2, t_3\rangle}{\langle u'_2, u'_2\rangle} u'_2 = t_3 - \frac{0}{5} t_0 - \frac{34}{10} t_1 - \frac{0}{14}(t_2 - 2t_0) = t_3 - \frac{17}{5} t_1$$
$$= (-8, -1, 0, 1, 8) - \frac{17}{5}(-2, -1, 0, 1, 2) = \frac{6}{5}(-1, 2, 0, -2, 1),$$
$$\langle u'_3, u'_3\rangle = \langle \frac{6}{5}(-1, 2, 0, -2, 1), \frac{6}{5}(-1, 2, 0, -2, 1)\rangle = (\frac{6}{5})^2((-1)^2 + 2^2 + 0^2 + (-2)^2 + 1^2) = (\frac{6}{5})^2 \cdot 10,$$
$$u_3 := \frac{1}{\|u'_3\|} u'_3 = \frac{1}{\frac{6}{5}\sqrt{10}}(t_3 - \frac{17}{5} t_1) = \frac{5}{6\sqrt{10}} t_3 - \frac{17}{6\sqrt{10}} t_1 = \frac{1}{\sqrt{10}}(-1, 2, 0, -2, 1)$$

Nach der Gram-Schmidt-Orthogonalisierung (7.42) und Bemerkung (7.15) ist $(u_0, \ldots, u_{k-1})$ für $k \in [1, 4]$ eine Orthonormalbasis von U'_k.

(a) Wegen

$$\langle u_0, b\rangle = \langle \frac{1}{\sqrt{5}}(1, 1, 1, 1, 1), (2, 1, 1, 1, 0)\rangle = \frac{1}{\sqrt{5}}(1 \cdot 2 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 0) = \frac{5}{\sqrt{5}}$$

ist

$$\mathrm{pr}_{U'_1}(b) = \langle u_0, b\rangle u_0 = \frac{5}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}} t_0 = t_0.$$

Nach Bemerkung (7.92) ist daher

$$(\varepsilon_a|^{U'_1}_{\mathrm{Pol}_{<1}(\mathbb{R},\mathbb{R})})^{-1}(\mathrm{pr}_{U'_1}(b)) = (\varepsilon_a|^{U'_1}_{\mathrm{Pol}_{<1}(\mathbb{R},\mathbb{R})})^{-1}(t_0) = s_0$$

die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<1}(\mathbb{R}, \mathbb{R})$.

(b) Wegen

$$\langle u_1, b\rangle = \langle \frac{1}{\sqrt{10}}(-2, -1, 0, 1, 2), (2, 1, 1, 1, 0)\rangle = \frac{1}{\sqrt{10}}((-2) \cdot 2 + (-1) \cdot 1 + 0 \cdot 1 + 1 \cdot 1 + 2 \cdot 0) = -\frac{4}{\sqrt{10}}$$

ist

$$\langle u_1, b\rangle u_1 = -\frac{4}{\sqrt{10}} \cdot \frac{1}{\sqrt{10}} t_1 = -\frac{2}{5} t_1$$

und damit

$$\mathrm{pr}_{U'_2}(b) = \langle u_0, b\rangle u_0 + \langle u_1, b\rangle u_1 = t_0 - \frac{2}{5} t_1.$$

Nach Bemerkung (7.92) ist daher

$$(\varepsilon_a|^{U'_2}_{\mathrm{Pol}_{<2}(\mathbb{R},\mathbb{R})})^{-1}(\mathrm{pr}_{U'_2}(b)) = (\varepsilon_a|^{U'_2}_{\mathrm{Pol}_{<2}(\mathbb{R},\mathbb{R})})^{-1}(t_0 - \frac{2}{5} t_1) = s_0 - \frac{2}{5} s_1$$

die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<2}(\mathbb{R}, \mathbb{R})$.

(c) Wegen

$$\langle u_2, b\rangle = \langle \frac{1}{\sqrt{14}}(2, -1, -2, -1, 2), (2, 1, 1, 1, 0)\rangle = \frac{1}{\sqrt{14}}(2 \cdot 2 + (-1) \cdot 1 + (-2) \cdot 1 + (-1) \cdot 1 + 2 \cdot 0) = 0$$

ist

$$\langle u_2, b\rangle u_2 = 0$$

und damit

$$\mathrm{pr}_{U'_3}(b) = \langle u_0, b\rangle u_0 + \langle u_1, b\rangle u_1 + \langle u_2, b\rangle u_2 = t_0 - \frac{2}{5} t_1.$$

Nach Bemerkung (7.92) ist daher

$$(\varepsilon_a|^{U'_3}_{\mathrm{Pol}_{<3}(\mathbb{R},\mathbb{R})})^{-1}(\mathrm{pr}_{U'_3}(b)) = (\varepsilon_a|^{U'_3}_{\mathrm{Pol}_{<3}(\mathbb{R},\mathbb{R})})^{-1}(t_0 - \frac{2}{5} t_1) = s_0 - \frac{2}{5} s_1$$

die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<3}(\mathbb{R}, \mathbb{R})$.

(d) Wegen

$$\langle u_3, b\rangle = \langle \frac{1}{\sqrt{10}}(-1,2,0,-2,1),(2,1,1,1,0)\rangle = \frac{1}{\sqrt{10}}((-1)\cdot 2 + 2\cdot 1 + 0\cdot 1 + (-2)\cdot 1 + 1\cdot 0) = -\frac{2}{\sqrt{10}}$$

ist

$$\langle u_3, b\rangle u_3 = -\frac{2}{\sqrt{10}}(\frac{5}{6\sqrt{10}}t_3 - \frac{17}{6\sqrt{10}}t_1) = -\frac{1}{6}t_3 + \frac{17}{30}t_1$$

und damit

$$\mathrm{pr}_{U'_4}(b) = \langle u_0, b\rangle u_0 + \langle u_1, b\rangle u_1 + \langle u_2, b\rangle u_2 + \langle u_3, b\rangle u_3 = t_0 - \frac{2}{5}t_1 - \frac{1}{6}t_3 + \frac{17}{30}t_1 = t_0 + \frac{1}{6}t_1 - \frac{1}{6}t_3.$$

Nach Bemerkung (7.92) ist daher

$$(\varepsilon_a|^{U'_4}_{\mathrm{Pol}_{<4}(\mathbb{R},\mathbb{R})})^{-1}(\mathrm{pr}_{U'_4}(b)) = (\varepsilon_a|^{U'_4}_{\mathrm{Pol}_{<4}(\mathbb{R},\mathbb{R})})^{-1}(t_0 + \frac{1}{6}t_1 - \frac{1}{6}t_3) = s_0 + \frac{1}{6}s_1 - \frac{1}{6}s_3$$

die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<4}(\mathbb{R},\mathbb{R})$. □

So wie sich Interpolationen in einem endlichdimensionalen Vektorraum nach Basiswahl mittels Lösungen von linearen Gleichungssystemen zu einer zugehörigen Darstellungsmatrix berechnen lassen, siehe Korollar (3.53), erhalten wir Ausgleichsfunktionen auf analoge Weise mittels besten Näherungslösungen:

(7.94) Bemerkung. Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \ldots, s_n)$ von U, $a \in X^m$ und $b \in K^m$ gegeben. Dann gilt

$$\kappa_s(\mathrm{Fit}_U(a,b)) = \mathrm{ApprSol}(\begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})$$

und

$$\mathrm{Fit}_U(a,b) = \{\sum_{j\in[1,n]} c_j s_j \mid c \in \mathrm{ApprSol}(\begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.95) Korollar (Eindeutigkeitskriterium für Ausgleichsfunktionen)**.** Es seien eine Menge X, ein K-Untervektorraum U von $\mathrm{Map}(X, K)$, $m, n \in \mathbb{N}_0$, eine Basis $s = (s_1, \ldots, s_n)$ von U, $a \in X^m$ und $b \in K^m$ gegeben. Genau dann gibt es genau eine Ausgleichsfunktion zu (a, b) in U, wenn

$$\mathrm{rk}_K \begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix} = n$$

ist.

Beweis. Nach Bemerkung (7.94) ist

$$\mathrm{Fit}_U(a,b) = \{\sum_{j\in[1,n]} c_j s_j \mid c \in \mathrm{ApprSol}(\begin{pmatrix} s_1(a_1) & \ldots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})\}.$$

Somit gibt es genau dann genau eine Ausgleichsfunktion zu (a, b) in U, wenn das lineare Gleichungssystem zur erweiterten Koeffizientenmatrix

$$\left(\begin{array}{ccc|c} s_1(a_1) & \ldots & s_n(a_1) & b_1 \\ \vdots & & \vdots & \vdots \\ s_1(a_m) & \ldots & s_n(a_m) & b_n \end{array}\right)$$

genau eine beste Näherungslösung hat. Dies ist nach dem Eindeutigkeitskriterium für beste Näherungslösungen linearer Gleichungssysteme (7.89) genau dann der Fall, wenn

$$\operatorname{rk}\begin{pmatrix} s_1(a_1) & \dots & s_n(a_1) \\ \vdots & & \vdots \\ s_1(a_m) & \dots & s_n(a_m) \end{pmatrix} = n$$

ist. □

Als Illustration wenden wir unsere Theorie auf Polynomfunktionen an:

(7.96) Beispiel. Es seien $m, n \in \mathbb{N}_0$, $a, b \in K^m$ gegeben.

(a) Es ist

$$\operatorname{Fit}_{\operatorname{Pol}_{<n}(K,K)}(a,b) = \{K \to K,\, x \mapsto \sum_{j \in [1,n]} c_j x^{j-1} \mid c \in \operatorname{ApprSol}(\operatorname{V}_{<n}(a), \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix})\}.$$

(b) Es sei angenommen, dass die Einträge von a verschieden sind. Die folgenden Bedingungen sind äquivalent.

(i) Es gibt genau eine Ausgleichsfunktion zu (a, b) in $\operatorname{Pol}_{<n}(K, K)$.

(ii) Es ist

$$\operatorname{rk}_K \operatorname{V}_{<n}(a) = n.$$

(iii) Es ist $m \geq n$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Alternativer Beweis von Beispiel (7.93). Für $n \in [1, 4]$ ist

$$\operatorname{Fit}_{\operatorname{Pol}_{<n}(\mathbb{R},\mathbb{R})}(a,b) = \{\mathbb{R} \to \mathbb{R},\, x \mapsto \sum_{j \in [1,n]} c_j x^{j-1} \mid c \in \operatorname{ApprSol}(\begin{pmatrix} (-2)^0 & \dots & (-2)^{n-1} \\ (-1)^0 & \dots & (-1)^{n-1} \\ 0^0 & \dots & 0^{n-1} \\ 1^0 & \dots & 1^{n-1} \\ 2^0 & \dots & 2^{n-1} \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix})\}.$$

nach Beispiel (7.96)(a). Um die angegebene Menge der besten Näherungslösungen zu bestimmen, verwenden wir jeweils Korollar (7.86).

(a) Es ist

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix}.$$

Wir wenden elementare Zeilenoperationen auf die erweiterte Koeffizientenmatrix

$$\begin{pmatrix} 5 \mid 5 \end{pmatrix}$$

an:

$$\left(5 \mid 5\right) \xmapsto{\mathrm{mul}_{1,\frac{1}{5}}} \left(1 \mid 1\right)$$

Nach Korollar (7.86), Proposition (A.163) und Proposition (A.159) ist

$$\mathrm{ApprSol}(\begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix}) = \mathrm{Sol}(\begin{pmatrix}5\end{pmatrix}, \begin{pmatrix}5\end{pmatrix}) = \{\begin{pmatrix}1\end{pmatrix}\}.$$

Folglich ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto 1$ die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<1}(\mathbb{R}, \mathbb{R})$.

(b) Es ist

$$\begin{pmatrix}1 & 1 & 1 & 1 & 1\\-2 & -1 & 0 & 1 & 2\end{pmatrix}\begin{pmatrix}1 & -2\\1 & -1\\1 & 0\\1 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}5 & 0\\0 & 10\end{pmatrix},$$

$$\begin{pmatrix}1 & 1 & 1 & 1 & 1\\-2 & -1 & 0 & 1 & 2\end{pmatrix}\begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix} = \begin{pmatrix}5\\-4\end{pmatrix}.$$

Wir wenden elementare Zeilenoperationen auf die erweiterte Koeffizientenmatrix

$$\left(\begin{array}{cc|c}5 & 0 & 5\\0 & 10 & -4\end{array}\right)$$

an:

$$\left(\begin{array}{cc|c}5 & 0 & 5\\0 & 10 & -4\end{array}\right) \xmapsto{\mathrm{mul}_{2,\frac{1}{10}} \circ \mathrm{mul}_{1,\frac{1}{5}}} \left(\begin{array}{cc|c}1 & 0 & 1\\0 & 1 & -\frac{2}{5}\end{array}\right)$$

Nach Korollar (7.86), Proposition (A.163) und Proposition (A.159) ist

$$\mathrm{ApprSol}(\begin{pmatrix}1 & -2\\1 & -1\\1 & 0\\1 & 1\\1 & 2\end{pmatrix}, \begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix}) = \mathrm{Sol}(\begin{pmatrix}5 & 0\\0 & 10\end{pmatrix}, \begin{pmatrix}5\\-4\end{pmatrix}) = \{\begin{pmatrix}1\\-\frac{2}{5}\end{pmatrix}\}.$$

Folglich ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto 1 - \frac{2}{5}x$ die eindeutige Ausgleichsfunktion zu (a, b) in $\mathrm{Pol}_{<2}(\mathbb{R}, \mathbb{R})$.

(c) Es ist

$$\begin{pmatrix}1 & 1 & 1 & 1 & 1\\-2 & -1 & 0 & 1 & 2\\4 & 1 & 0 & 1 & 4\end{pmatrix}\begin{pmatrix}1 & -2 & 4\\1 & -1 & 1\\1 & 0 & 0\\1 & 1 & 1\\1 & 2 & 4\end{pmatrix} = \begin{pmatrix}5 & 0 & 10\\0 & 10 & 0\\10 & 0 & 34\end{pmatrix},$$

$$\begin{pmatrix}1 & 1 & 1 & 1 & 1\\-2 & -1 & 0 & 1 & 2\\4 & 1 & 0 & 1 & 4\end{pmatrix}\begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix} = \begin{pmatrix}5\\-4\\10\end{pmatrix}.$$

Wir wenden elementare Zeilenoperationen auf die erweiterte Koeffizientenmatrix

$$\left(\begin{array}{ccc|c} 5 & 0 & 10 & 5 \\ 0 & 10 & 0 & -4 \\ 10 & 0 & 34 & 10 \end{array}\right)$$

an:

$$\left(\begin{array}{ccc|c} 5 & 0 & 10 & 5 \\ 0 & 10 & 0 & -4 \\ 10 & 0 & 34 & 10 \end{array}\right) \xmapsto{\text{add}_{3,1,-2}} \left(\begin{array}{ccc|c} 5 & 0 & 10 & 5 \\ 0 & 10 & 0 & -4 \\ 0 & 0 & 14 & 0 \end{array}\right) \xmapsto{\text{mul}_{3,\frac{1}{14}}} \left(\begin{array}{ccc|c} 5 & 0 & 10 & 5 \\ 0 & 10 & 0 & -4 \\ 0 & 0 & 1 & 0 \end{array}\right)$$
$$\xmapsto{\text{add}_{1,3,-10}} \left(\begin{array}{ccc|c} 5 & 0 & 0 & 5 \\ 0 & 10 & 0 & -4 \\ 0 & 0 & 1 & 0 \end{array}\right) \xmapsto{\text{mul}_{1,\frac{1}{5}}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{2}{5} \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Nach Korollar (7.86), Proposition (A.163) und Proposition (A.159) ist

$$\operatorname{ApprSol}(\begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}) = \operatorname{Sol}(\begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}, \begin{pmatrix} 5 \\ -4 \\ 10 \end{pmatrix}) = \{\begin{pmatrix} 1 \\ -\frac{2}{5} \\ 0 \end{pmatrix}\}.$$

Folglich ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto 1 - \frac{2}{5}x$ die eindeutige Ausgleichsfunktion zu (a, b) in $\operatorname{Pol}_{<3}(\mathbb{R}, \mathbb{R})$.

(d) Es ist

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \\ -8 & -1 & 0 & 1 & 8 \end{pmatrix} \begin{pmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 10 & 0 \\ 0 & 10 & 0 & 34 \\ 10 & 0 & 34 & 0 \\ 0 & 34 & 0 & 130 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \\ -8 & -1 & 0 & 1 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ 10 \\ -16 \end{pmatrix}.$$

Wir wenden elementare Zeilenoperationen auf die erweiterte Koeffizientenmatrix

$$\left(\begin{array}{cccc|c} 5 & 0 & 10 & 0 & 5 \\ 0 & 10 & 0 & 34 & -4 \\ 10 & 0 & 34 & 0 & 10 \\ 0 & 34 & 0 & 130 & -16 \end{array}\right)$$

an:

$$\left(\begin{array}{cccc|c} 5 & 0 & 10 & 0 & 5 \\ 0 & 10 & 0 & 34 & -4 \\ 10 & 0 & 34 & 0 & 10 \\ 0 & 34 & 0 & 130 & -16 \end{array}\right) \xmapsto{\text{add}_{3,1,-2}} \left(\begin{array}{cccc|c} 5 & 0 & 10 & 0 & 5 \\ 0 & 10 & 0 & 34 & -4 \\ 0 & 0 & 14 & 0 & 0 \\ 0 & 34 & 0 & 130 & -16 \end{array}\right) \xmapsto{\text{add}_{4,2,-3}} \left(\begin{array}{cccc|c} 5 & 0 & 10 & 0 & 5 \\ 0 & 10 & 0 & 34 & -4 \\ 0 & 0 & 14 & 0 & 0 \\ 0 & 4 & 0 & 28 & -4 \end{array}\right)$$
$$\xmapsto{\text{mul}_{4,\frac{1}{4}} \circ \text{mul}_{3,\frac{1}{14}} \circ \text{mul}_{1,\frac{1}{5}}} \left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 1 \\ 0 & 10 & 0 & 34 & -4 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 7 & -1 \end{array}\right) \xmapsto{\text{add}_{2,4,-10} \circ \text{add}_{1,3,-2}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -36 & 6 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 7 & -1 \end{array}\right)$$
$$\xmapsto{\text{mul}_{2,-\frac{1}{36}}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -\frac{1}{6} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 7 & -1 \end{array}\right) \xmapsto{\text{add}_{4,2,-7}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -\frac{1}{6} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{6} \end{array}\right) \xmapsto{\text{sw}_{1,3}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & \frac{1}{6} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{6} \end{array}\right)$$

Nach Korollar (7.86), Proposition (A.163) und Proposition (A.159) ist

$$\operatorname{ApprSol}(\begin{pmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}) = \operatorname{Sol}(\begin{pmatrix} 5 & 0 & 10 & 0 \\ 0 & 10 & 0 & 34 \\ 10 & 0 & 34 & 0 \\ 0 & 34 & 0 & 130 \end{pmatrix}, \begin{pmatrix} 5 \\ -4 \\ 10 \\ -16 \end{pmatrix}) = \{\begin{pmatrix} 1 \\ \frac{1}{6} \\ 0 \\ -\frac{1}{6} \end{pmatrix}\}.$$

Folglich ist $\mathbb{R} \to \mathbb{R}$, $x \mapsto 1 + \frac{1}{6}x - \frac{1}{6}x^3$ die eindeutige Ausgleichsfunktion zu (a, b) in $\operatorname{Pol}_{<4}(\mathbb{R}, \mathbb{R})$. □

Unitäre Matrizen

Nach Bemerkung (3.31) entsprechen invertierbare $(n \times n)$-Matrizen mit Einträgen in K für $n \in \mathbb{N}_0$ genau den Matrizen, deren Spalten eine Basis von $K^{n\times 1}$ ergeben, bzw. nach Proposition (3.13)(c) und Korollar (3.32) den Basiswechselmatrizen zu beliebigen n-dimensionalen K-Vektorräumen. Wir betrachten nun unitäre Matrizen, welche die analoge Rolle für Orthonormalbasen in K-Skalarproduktträumen innehaben.

(7.97) Definition (unitäre Matrix)**.** Es sei $n \in \mathbb{N}_0$ gegeben. Die *unitäre Gruppe* vom Grad n über K ist definiert als

$$\mathrm{U}_n(K) := \{A \in \mathrm{GL}_n(K) \mid A^{-1} = A^{\mathrm{ad}}\}.$$

Ein Element von $\mathrm{U}_n(K)$ wird *unitäre $(n \times n)$-Matrix* über K genannt.

(7.98) Beispiel. Es sei $A \in K^{2\times 2}$ gegeben durch

$$A = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}.$$

Dann ist A eine unitäre (2×2)-Matrix über K.

Beweis. Wegen

$$A^{\mathrm{ad}}A = (\frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix})^{\mathrm{ad}} \cdot \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = \mathrm{E}_2$$

ist $A \in \mathrm{U}_2(K)$ nach Korollar (3.34). □

(7.99) Bemerkung. Es sei $n \in \mathbb{N}_0$ gegeben. Dann ist

$$\mathrm{U}_n(K) = \{A \in K^{n\times n} \mid (A_{-,1}, \ldots, A_{-,n}) \text{ ist eine Orthonormalbasis von } K^{n\times 1}\}.$$

Beweis. Nach Korollar (3.34) ist A genau dann unitär, wenn $A^{\mathrm{ad}}A = \mathrm{E}_n$ ist, d.h. wenn für $i, j \in [1, n]$ stets $(A^{\mathrm{ad}}A)_{i,j} = \delta_{i,j}$ ist. Für $i, j \in [1, n]$ ist jedoch

$$(A^{\mathrm{ad}}A)_{i,j} = A^{\mathrm{ad}}_{i,-} A_{-,j} = (A_{-,i})^{\mathrm{ad}} A_{-,j} = \langle A_{-,i}, A_{-,j}\rangle,$$

d.h. $(A^{\mathrm{ad}}A)_{i,j} = \delta_{i,j}$ ist äquivalent zu $\langle A_{-,i}, A_{-,j}\rangle = \delta_{i,j}$. Schließlich gilt genau dann $\langle A_{-,i}, A_{-,j}\rangle = \delta_{i,j}$ für $i, j \in [1, n]$, wenn $(A_{-,1}, \ldots, A_{-,n})$ eine Orthonormalbasis von $K^{n\times 1}$ ist. Insgesamt ist A genau dann unitär, wenn $(A_{-,1}, \ldots, A_{-,n})$ eine Orthonormalbasis von $K^{n\times 1}$ ist. □

(7.100) Proposition. Es seien ein K-Skalarproduktraum V, $n \in \mathbb{N}_0$, eine Orthonormalbasis $s = (s_1, \ldots, s_n)$ von V und $P \in \mathrm{U}_n(K)$ gegeben. Ferner sei $s' := (\kappa_s^{-1}(P_{-,1}), \ldots, \kappa_s^{-1}(P_{-,n}))$. Dann ist s' eine Orthonormalbasis von V mit

$$\mathrm{M}_{s,s'}(\mathrm{id}_V) = P.$$

Beweis. Nach Korollar (3.32) ist s' eine Basis von V mit $\mathrm{M}_{s,s'}(\mathrm{id}_V) = P$. Da s eine Orthonormalbasis von V ist, gilt außerdem

$$\langle s'_i, s'_j\rangle = \langle \kappa_s(s'_i), \kappa_s(s'_j)\rangle = \delta_{i,j}$$

für $i, j \in [1, n]$ nach Korollar (7.34) und Bemerkung (7.99), d.h. auch s' ist eine Orthonormalbasis von V. □

Unitäre Äquivalenz und unitäre Ähnlichkeit von Matrizen

So wie die Basiswechselformel (3.18)(b) zu den Begriffen der Äquivalenz von Matrizen und der Ähnlichkeit von Matrizen, siehe Definition (3.60), führt, motiviert sie in analoger Weise auch die folgenden analogen Begriffe für unitäre Matrizen, welche nach Proposition (7.100) gerade den Basiswechselmatrizen zu Orthonormalbasen entsprechen.

(7.101) Definition (unitäre Äquivalenz von Matrizen, unitäre Ähnlichkeit von Matrizen)**.**

(a) Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $A, B \in K^{m \times n}$ sagen wir, dass A *unitär äquivalent* zu B ist, wenn es $P \in \mathrm{U}_n(K)$, $Q \in \mathrm{U}_m(K)$ mit

$$Q^{-1}AP = B$$

gibt.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Für $A, B \in K^{n \times n}$ sagen wir, dass A *unitär ähnlich* zu B ist, wenn es ein $P \in \mathrm{U}_n(K)$ mit

$$P^{-1}AP = B$$

gibt.

(7.102) Beispiel.

(a) Es seien $A, B \in \mathbb{R}^{2 \times 3}$ gegeben durch

$$A = \frac{1}{6}\begin{pmatrix} 6+\sqrt{3} & 3+5\sqrt{3} & -6-\sqrt{3} \\ 11-2\sqrt{3} & 7-\sqrt{3} & 1+2\sqrt{3} \end{pmatrix},\ B = \begin{pmatrix} \sqrt{3} & 0 & -\sqrt{3} \\ 0 & 1 & \sqrt{3} \end{pmatrix}.$$

Dann ist A unitär äquivalent zu B.

(b) Es seien $A, B \in \mathbb{R}^{2 \times 2}$ gegeben durch

$$A = \frac{1}{4}\begin{pmatrix} 4-\sqrt{3} & -3 \\ 1 & 4+\sqrt{3} \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Dann ist A unitär ähnlich zu B.

Beweis.

(a) Es seien $P \in \mathbb{R}^{3 \times 3}$, $Q \in \mathbb{R}^{2 \times 2}$ gegeben durch

$$P = \frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ 2 & -2 & -1 \\ 2 & 1 & 2 \end{pmatrix},\ Q = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}.$$

Nach Beispiel (7.43) und Bemerkung (7.99) ist $P \in \mathrm{U}_3(\mathbb{R})$ und nach Beispiel (7.98) ist $Q \in \mathrm{U}_2(\mathbb{R})$ und es gilt

$$\begin{aligned} Q^{-1}AP = Q^{\mathrm{ad}}AP &= \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix} \cdot \frac{1}{6}\begin{pmatrix} 6+\sqrt{3} & 3+5\sqrt{3} & -6-\sqrt{3} \\ 11-2\sqrt{3} & 7-\sqrt{3} & 1+2\sqrt{3} \end{pmatrix} \cdot \frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ 2 & -2 & -1 \\ 2 & 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{3} & 0 & -\sqrt{3} \\ 0 & 1 & \sqrt{3} \end{pmatrix} = B. \end{aligned}$$

Folglich ist A unitär äquivalent zu B.

(b) Es seien $P \in \mathbb{Q}^{2 \times 2}$ gegeben durch

$$P = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}.$$

Nach Beispiel (7.98) ist $P \in \mathrm{U}_2(\mathbb{R})$ und es gilt

$$P^{-1}AP = P^{\mathrm{ad}}AP = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix} \cdot \frac{1}{4}\begin{pmatrix} 4-\sqrt{3} & -3 \\ 1 & 4+\sqrt{3} \end{pmatrix} \cdot \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = B.$$

Folglich ist A unitär ähnlich zu B. □

(7.103) Bemerkung.

(a) Es seien $m, n \in \mathbb{N}_0$ gegeben. Unitäre Äquivalenz von Matrizen ist eine Äquivalenzrelation auf $K^{m \times n}$.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Unitäre Ähnlichkeit von Matrizen ist eine Äquivalenzrelation auf $K^{n \times n}$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.104) Bemerkung.

(a) Es seien K-Skalarprodukträume V und W, ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n \in \mathbb{N}_0$, eine Orthonormalbasis $s = (s_1, \dots, s_n)$ von V und eine Orthonormalbasis $t = (t_1, \dots, t_m)$ von W gegeben. Die Äquivalenzklasse von $\mathrm{M}_{t,s}(\varphi)$ bzgl. unitärer Äquivalenz von Matrizen ist gegeben durch

$$[\mathrm{M}_{t,s}(\varphi)] = \{\mathrm{M}_{t',s'}(\varphi) \mid s' \text{ ist eine Orthonormalbasis von } V,\ t' \text{ ist eine Orthonormalbasis von } W\}.$$

(b) Es seien ein K-Skalarproduktraum V, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Orthonormalbasis $s = (s_1, \dots, s_n)$ von V gegeben. Die Äquivalenzklasse von $\mathrm{M}_{s,s}(\varphi)$ bzgl. unitärer Ähnlichkeit von Matrizen ist gegeben durch

$$[\mathrm{M}_{s,s}(\varphi)] = \{\mathrm{M}_{s',s'}(\varphi) \mid s' \text{ ist eine Orthonormalbasis von } V\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Unitäre Diagonalisierbarkeit

Der zum Begriff der Diagonalisierbarkeit, siehe Definition (6.46), analoge Begriff für Skalarprodukträume und Orthonormalbasen ist der folgende:

(7.105) Definition (unitär diagonalisierbar)**.**

(a) Es sei ein endlichdimensionaler K-Skalarproduktraum V gegeben. Ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ heißt *unitär diagonalisierbar*, falls es eine orthonormale Eigenbasis von V bzgl. φ gibt.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Ein $A \in K^{n \times n}$ heißt *unitär diagonalisierbar* (über K), falls $\varphi_A\colon K^{n \times 1} \to K^{n \times 1}$ unitär diagonalisierbar ist.

(7.106) Beispiel. Es sei $A \in K^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Dann ist A unitär diagonalisierbar.

Beweis. Nach Beispiel (6.42) ist

$$s := (\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

eine Eigenbasis von $\mathbb{Q}^{2 \times 1}$ bzgl. A. Wegen

$$\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle = 1 \cdot 1 + (-1) \cdot 1 = 0$$

ist s orthogonal und damit

$$(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

nach Bemerkung (7.15), Korollar (1.23)(c) und Korollar (1.38)(c) eine Orthonormalbasis von $K^{n \times 1}$ bzgl. A. Somit ist A unitär diagonalisierbar. □

(7.107) Bemerkung. Es seien ein endlichdimensionaler K-Skalarproduktraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Genau dann ist φ unitär diagonalisierbar, wenn es eine Orthonormalbasis s von V derart gibt, dass $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.108) Bemerkung. Es seien ein K-Skalarproduktraum V, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Orthonormalbasis $s = (s_1, \dots, s_n)$ von V gegeben. Genau dann ist φ unitär diagonalisierbar, wenn $\mathrm{M}_{s,s}(\varphi)$ unitär diagonalisierbar ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.109) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n \times n}$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Die Matrix A ist unitär diagonalisierbar.

(b) Es gibt eine orthonormale Eigenbasis von $K^{n \times 1}$ bzgl. A.

(c) Die Matrix A ist unitär ähnlich zu einer Diagonalmatrix.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Normale Endomorphismen

Der Spektralsatz (7.120) liefert eine Charakterisierung von unitärer Diagonalisierbarkeit mittels sogenannter normaler Endomorphismen bzw. normaler Matrizen:

(7.110) Definition (normaler Vektorraumendomorphismus, normale Matrix)**.**

(a) Es sei ein endlichdimensioner K-Skalarproduktraum V gegeben. Die *Menge der normalen K-Vektorraumendomorphismen* von V ist definiert als

$$\mathrm{End}_{\mathrm{norm}}(V) = \mathrm{End}_{K,\mathrm{norm}}(V) := \{\varphi \in \mathrm{End}_K(V) \mid \varphi^{\mathrm{ad}} \circ \varphi = \varphi \circ \varphi^{\mathrm{ad}}\}.$$

Ein Element von $\mathrm{End}_{K,\mathrm{norm}}(V)$ wird *normaler K-Vektorraumendomorphismus* von V genannt.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Die *Menge der normalen $(n \times n)$-Matrizen* über K ist definiert als

$$K^{n \times n}_{\mathrm{norm}} := \{A \in K^{n \times n} \mid \varphi_A \in \mathrm{End}_{K,\mathrm{norm}}(K^{n \times 1})\}.$$

Ein Element von $K^{n \times n}_{\mathrm{norm}}$ wird *normale $(n \times n)$-Matrix* über K genannt.

(7.111) Bemerkung. Es sei $n \in \mathbb{N}_0$ gegeben. Dann ist

$$K^{n \times n}_{\mathrm{norm}} = \{A \in K^{n \times n} \mid A^{\mathrm{ad}} A = A A^{\mathrm{ad}}\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.112) Beispiel.

(a) Es sei $A \in K^{2 \times 2}$ gegeben durch

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

Dann ist A normal.

(b) Es sei $A \in \mathbb{C}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix}.$$

Dann ist A normal.

Beweis.

(a) Wegen

$$\begin{aligned} A^{\mathrm{ad}} A &= \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}^{\mathrm{ad}} \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}^{\mathrm{ad}} = A A^{\mathrm{ad}} \end{aligned}$$

ist A normal nach Bemerkung (7.111).

(b) Wegen

$$\begin{aligned} A^{\mathrm{ad}} A &= \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix}^{\mathrm{ad}} \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{\mathrm{i}} & \overline{1} \end{pmatrix} \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -\mathrm{i} & 1 \end{pmatrix} \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1+\mathrm{i} \\ 1-\mathrm{i} & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -\mathrm{i} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{\mathrm{i}} & \overline{1} \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & \mathrm{i} \\ 1 & 1 \end{pmatrix}^{\mathrm{ad}} = A A^{\mathrm{ad}} \end{aligned}$$

ist A normal nach Bemerkung (7.111). □

(7.113) Bemerkung. Es seien $n \in \mathbb{N}_0$ und $A \in \mathrm{U}_n(K)$ gegeben. Dann ist A normal.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.114) Proposition. Es seien ein endlichdimensioner K-Skalarproduktraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Der K-Vektorraumendomorphismus $\varphi\colon V \to V$ ist normal.

(b) Für $v, w \in V$ gilt

$$\langle \varphi^{\mathrm{ad}}(v), \varphi^{\mathrm{ad}}(w) \rangle = \langle \varphi(v), \varphi(w) \rangle.$$

(c) Für $v \in V$ gilt

$$\|\varphi^{\mathrm{ad}}(v)\| = \|\varphi(v)\|.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.115) Bemerkung. Es seien ein endlichdimensioner K-Skalarproduktraum V und ein normaler K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Für $a \in K$ ist $\varphi - a \,\mathrm{id}_V$ normal.

Beweis. Für $a \in K$ ist

$$(\varphi - a \,\mathrm{id}_V)^{\mathrm{ad}} = \varphi^{\mathrm{ad}} - (a \,\mathrm{id}_V)^{\mathrm{ad}} = \varphi^{\mathrm{ad}} - \overline{a} \,\mathrm{id}_V^{\mathrm{ad}} = \varphi^{\mathrm{ad}} - \overline{a} \,\mathrm{id}_V$$

nach Proposition (7.57)(c), (d), (b), also

$$\begin{aligned} (\varphi - a \,\mathrm{id}_V) \circ (\varphi - a \,\mathrm{id}_V)^{\mathrm{ad}} &= (\varphi - a \,\mathrm{id}_V) \circ (\varphi^{\mathrm{ad}} - \overline{a} \,\mathrm{id}_V) = \varphi \circ \varphi^{\mathrm{ad}} - \overline{a}\varphi - a\varphi^{\mathrm{ad}} + a\overline{a} \,\mathrm{id}_V \\ &= \varphi^{\mathrm{ad}} \circ \varphi - a\varphi^{\mathrm{ad}} - \overline{a}\varphi + \overline{a}a \,\mathrm{id}_V = (\varphi^{\mathrm{ad}} - \overline{a} \,\mathrm{id}_V) \circ (\varphi - a \,\mathrm{id}_V) \\ &= (\varphi - a \,\mathrm{id}_V)^{\mathrm{ad}} \circ (\varphi - a \,\mathrm{id}_V). \end{aligned}$$

□

(7.116) Proposition. Es seien ein endlichdimensioner K-Skalarproduktraum V, ein normaler K-Vektorraumendomorphismus $\varphi\colon V \to V$ und $a \in K$ gegeben. Dann ist

$$\mathrm{Eig}_{\overline{a}}(\varphi^{\mathrm{ad}}) = \mathrm{Eig}_a(\varphi).$$

Genau dann ist a ein Eigenwert von φ, wenn $\overline{a}$ ein Eigenwert von φ^{ad} ist.

Beweis. Für jeden normalen Endomorphismus $\psi\colon V \to V$ ist

$$\begin{aligned}\operatorname{Ker}\varphi^{\mathrm{ad}} &= \{v \in V \mid \varphi^{\mathrm{ad}}(v) = 0\} = \{v \in V \mid \|\varphi^{\mathrm{ad}}(v)\| = 0\} = \{v \in V \mid \|\varphi(v)\| = 0\} = \{v \in V \mid \varphi(v) = 0\} \\ &= \operatorname{Ker}\varphi\end{aligned}$$

nach Proposition (7.114). Mit φ ist nach Bemerkung (7.115) aber auch $\varphi - a\,\mathrm{id}_V$ normal, es gilt also

$$\mathrm{Eig}_{\overline{a}}(\varphi^{\mathrm{ad}}) = \operatorname{Ker}(\varphi^{\mathrm{ad}} - \overline{a}\,\mathrm{id}_V) = \operatorname{Ker}(\varphi - a\,\mathrm{id}_V)^{\mathrm{ad}} = \operatorname{Ker}(\varphi - a\,\mathrm{id}_V) = \mathrm{Eig}_a(\varphi)$$

nach Proposition (7.57)(c), (d), (b).
Insbesondere ist genau dann $\mathrm{Eig}_a(\varphi) \neq \{0\}$, wenn $\mathrm{Eig}_{\overline{a}}(\varphi^{\mathrm{ad}}) \neq \{0\}$ ist, d.h. genau dann ist a ein Eigenwert von φ, wenn $\overline{a}$ ein Eigenwert von φ^{ad} ist. □

(7.117) Korollar. Es seien ein endlichdimensioner K-Skalarproduktraum V, ein normaler K-Vektorraum-endomorphismus $\varphi\colon V \to V$ und $a, b \in K$ mit $a \neq b$ gegeben. Für $v \in \mathrm{Eig}_a(\varphi)$ und $w \in \mathrm{Eig}_b(\varphi)$ gilt $v \perp w$.

Beweis. Es seien $v \in \mathrm{Eig}_a(\varphi)$ und $w \in \mathrm{Eig}_b(\varphi)$ gegeben. Da φ normal ist, gilt $\mathrm{Eig}_{\overline{a}}(\varphi^{\mathrm{ad}}) = \mathrm{Eig}_a(\varphi)$ nach Proposition (7.116) und damit

$$a\langle v, w\rangle = \langle \overline{a}v, w\rangle = \langle \varphi^{\mathrm{ad}}(v), w\rangle = \langle v, \varphi(w)\rangle = \langle v, bw\rangle = b\langle v, w\rangle$$

nach Korollar (7.59). Wegen $a \neq b$ impliziert dies bereits $\langle v, w\rangle = 0$ nach Proposition (A.91), d.h. es gilt $v \perp w$. □

(7.118) Bemerkung. Es seien ein endlichdimensionaler K-Skalarproduktraum V, ein normaler K-Vektor-raumendomorphismus $\varphi\colon V \to V$ und ein K-Untervektorraum U von V so gegeben, dass U und $U^{\perp}$ invariant unter φ sind. Dann ist die Restriktion $\varphi|_U^U\colon U \to U$ normal.

Beweis. Nach Bemerkung (7.67) ist mit φ auch $\varphi|_U^U$ adjungierbar mit $(\varphi|_U^U)^{\mathrm{ad}} = \varphi^{\mathrm{ad}}|_U^U$. Wir erhalten

$$\varphi|_U^U \circ (\varphi|_U^U)^{\mathrm{ad}} = \varphi|_U^U \circ \varphi^{\mathrm{ad}}|_U^U = (\varphi \circ \varphi^{\mathrm{ad}})|_U^U = (\varphi^{\mathrm{ad}} \circ \varphi)|_U^U = \varphi^{\mathrm{ad}}|_U^U \circ \varphi|_U^U = (\varphi|_U^U)^{\mathrm{ad}} \circ \varphi|_U^U.$$

Folglich ist $\varphi|_U^U$ normal. □

(7.119) Bemerkung. Es seien ein K-Skalarproduktraum V, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Orthonormalbasis $s = (s_1, \ldots, s_n)$ von V gegeben. Genau dann ist φ normal, wenn $\mathrm{M}_{s,s}(\varphi)$ normal ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.120) Satz (Spektralsatz)**.** Es seien ein endlichdimensionaler K-Skalarproduktraum V und ein K-Vektor-raumendomorphismus $\varphi\colon V \to V$ gegeben. Genau dann ist φ unitär diagonalisierbar, wenn χ_φ in Linearfaktoren zerfällt und φ normal ist.

Beweis. Zunächst sei φ unitär diagonalisierbar. Dann ist φ insbesondere diagonalisierbar, so dass χ_φ nach Satz (6.52) in Linearfaktoren zerfällt. Nach Bemerkung (7.107) gibt es außerdem eine Orthonormalbasis s von V derart, dass $\mathrm{M}_{s,s}(\varphi)$ eine Diagonalmatrix von V ist. Dann ist aber $\mathrm{M}_{s,s}(\varphi^{\mathrm{ad}}) = \mathrm{M}_{s,s}(\varphi)^{\mathrm{ad}}$ nach Proposition (7.69) ebenfalls eine Diagonalmatrix. Folglich gilt $\mathrm{M}_{s,s}(\varphi)^{\mathrm{ad}}\,\mathrm{M}_{s,s}(\varphi) = \mathrm{M}_{s,s}(\varphi)\,\mathrm{M}_{s,s}(\varphi)^{\mathrm{ad}}$, d.h. $\mathrm{M}_{s,s}(\varphi)$ ist normal. Dann ist aber auch φ normal nach Bemerkung (7.119).
Umgekehrt zerfalle χ_φ in Linearfaktoren zerfällt und es sei φ normal. Um zu zeigen, dass φ unitär diagonalisierbar ist, führen wir Induktion nach $n := \dim V$, wobei für $V = \{0\}$ nichts zu zeigen ist. Es sei also angenommen, dass $V \neq \{0\}$ ist und dass für jeden endlichdimensionalen K-Vektorraum U mit $\dim U = n - 1$ und für jeden Endomorphismus $\psi\colon U \to U$ derart, dass χ_ψ in Linearfaktoren zerfällt und ψ normal ist, gilt, dass ψ unitär diagonalisierbar ist.
Da χ_φ in Linearfaktoren zerfällt, hat φ einen Eigenwert a und einen zugehörigen Eigenvektor v. Dann ist Kv invariant unter φ. Für $u \in (Kv)^{\perp}$ gilt $\langle v, u\rangle = 0$ und damit

$$\langle v, \varphi(u)\rangle = \langle \varphi^{\mathrm{ad}}(v), u\rangle = \langle \overline{a}v, u\rangle = a\langle v, u\rangle = 0$$

nach Korollar (7.59), also $\varphi(u) \in \{v\}^{\perp} = (Kv)^{\perp}$ nach Proposition (7.27). Folglich ist auch $(Kv)^{\perp}$ invariant unter φ. Nach Bemerkung (7.118) ist daher $\varphi|_{(Kv)^{\perp}}^{(Kv)^{\perp}}\colon (Kv)^{\perp} \to (Kv)^{\perp}$ ein normaler Endomorphismus. Da $\chi_{\varphi|_{(Kv)^{\perp}}^{(Kv)^{\perp}}}$

nach Korollar (6.83) ein Teiler von χ_φ ist, zerfällt $\chi_{\varphi|_{(Kv)^\perp}^{(Kv)^\perp}}$ ferner ebenfalls in Linearfaktoren. Schließlich ist $\dim (Kv)^\perp = \dim V - \dim Kv = n - 1$ nach Proposition (7.41) und Korollar (1.81). Nach Induktionsvoraussetzung ist daher $\varphi|_{(Kv)^\perp}^{(Kv)^\perp}$ unitär diagonalisierbar, d.h. es existiert eine orthonormale Eigenbasis $(s_1, \dots, s_{n-1})$ von $(Kv)^\perp$ bzgl. $\varphi|_{(Kv)^\perp}^{(Kv)^\perp}$. Dann ist aber $(s_1, \dots, s_{n-1}, \frac{1}{\|v\|} v)$ eine orthonormale Eigenbasis von V bzgl. φ, d.h. auch φ ist unitär diagonalisierbar. □

Hermitesche Endomorphismen

Neben den unitären Matrizen liefern hermitesche Matrizen im folgenden Sinn eine wichtige Beispielquelle für normale Matrizen, siehe Bemerkung (7.124). In der Konsequenz können wir einen Spektralsatz für hermitesche Endomorphismen zwischen endlichdimensionalen Skalarprodukträumen beweisen, siehe Proposition (7.127).

(7.121) Definition (hermitescher Vektorraumendomorphismus, hermitesche Matrix)**.**

(a) Es sei ein K-Skalarproduktraum V gegeben. Die *Menge der hermiteschen K-Vektorraumendomorphismen* von V ist definiert als

$$\mathrm{End}_{\mathrm{herm}}(V) = \mathrm{End}_{K,\mathrm{herm}}(V) := \{\varphi \in \mathrm{End}_K(V) \mid \varphi \text{ ist adjungierbar und } \varphi^{\mathrm{ad}} = \varphi\}.$$

Ein Element von $\mathrm{End}_{K,\mathrm{herm}}(V)$ wird *hermitescher* (oder *selbstadjungierter*) *K-Vektorraumendomorphismus* von V genannt.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Die *Menge der hermiteschen $(n \times n)$-Matrizen* über K ist definiert als

$$K^{n \times n}_{\mathrm{herm}} := \{A \in K^{n \times n} \mid \varphi_A \in \mathrm{End}_{K,\mathrm{herm}}(K^{n \times 1})\}.$$

Ein Element von $K^{n \times n}_{\mathrm{herm}}$ wird *hermitesche* (oder *selbstadjungierte*) *$(n \times n)$-Matrix* über K genannt.

(7.122) Bemerkung. Es sei $n \in \mathbb{N}_0$ gegeben. Dann ist

$$K^{n \times n}_{\mathrm{herm}} = \{A \in K^{n \times n} \mid A^{\mathrm{ad}} = A\}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.123) Beispiel. Es sei $A \in \mathbb{C}^{2 \times 2}$ gegeben durch

$$A = \begin{pmatrix} 2 & 1 + \mathrm{i} \\ 1 - \mathrm{i} & 2 \end{pmatrix}.$$

Dann ist A hermitesch.

Beweis. Wegen

$$A^{\mathrm{ad}} = \begin{pmatrix} 2 & 1 + \mathrm{i} \\ 1 - \mathrm{i} & 2 \end{pmatrix}^{\mathrm{ad}} = \begin{pmatrix} \overline{2} & \overline{1 - \mathrm{i}} \\ \overline{1 + \mathrm{i}} & \overline{2} \end{pmatrix} = \begin{pmatrix} 2 & 1 + \mathrm{i} \\ 1 - \mathrm{i} & 2 \end{pmatrix} = A$$

ist A hermitesch nach Bemerkung (7.122). □

(7.124) Bemerkung. Es seien ein endlichdimensionaler K-Skalarproduktraum V und ein K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Wenn φ hermitesch ist, dann ist φ normal.

Beweis. Wenn φ hermitesch ist, dann ist φ adjungierbar und es gilt $\varphi^{\mathrm{ad}} = \varphi$, also insbesondere $\varphi^{\mathrm{ad}} \circ \varphi = \varphi \circ \varphi^{\mathrm{ad}}$, d.h. φ ist normal. □

(7.125) Proposition. Es seien ein K-Skalarproduktraum V und ein hermitescher K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Jeder Eigenwert von φ ist reell.

Beweis. Es sei ein Eigenwert a von φ gegeben. Da φ nach Bemerkung (7.124) normal ist, gilt

$$\mathrm{Eig}_{\overline{a}}(\varphi) = \mathrm{Eig}_{\overline{a}}(\varphi^{\mathrm{ad}}) = \mathrm{Eig}_a(\varphi) \neq \{0\}$$

nach Proposition (7.116), d.h. es gibt ein $v \in V \setminus \{0\}$ mit $\overline{a} v = \varphi(v) = av$. Dies impliziert aber bereits $\overline{a} = a$ nach Korollar (1.11). □

(7.126) Bemerkung. Es seien ein K-Skalarproduktraum V, ein K-Vektorraumendomorphismus $\varphi\colon V \to V$, $n \in \mathbb{N}_0$ und eine Orthonormalbasis $s = (s_1, \ldots, s_n)$ von V gegeben. Genau dann ist φ hermitesch, wenn $\mathrm{M}_{s,s}(\varphi)$ hermitesch ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.127) Proposition (Spektralsatz für hermitesche Endomorphismen)**.** Es sei ein endlichdimensionaler K-Skalarproduktraum V gegeben. Jeder hermitesche K-Vektorraumendomorphismus von V ist unitär diagonalisierbar.

Beweis. Es sei ein hermitescher K-Vektorraumendomorphismus $\varphi\colon V \to V$ gegeben. Wenn $K = \mathbb{C}$ ist, so zerfällt χ_φ nach dem Fundamentalsatz der Algebra und [12, Prop. (12.48)] in Linearfaktoren. Im Folgenden gelte daher $K = \mathbb{R}$. Nach Korollar (7.46) gibt es $n \in \mathbb{N}_0$ und eine Orthonormalbasis $s = (s_1, \ldots, s_n)$ von V. Nach Bemerkung (7.126) ist mit φ auch $\mathrm{M}_{s,s}(\varphi) \in \mathbb{R}^{n\times n}$ hermitesch. Da auch $\mathrm{M}_{s,s}(\varphi) \in \mathbb{C}^{n\times n}$ ist, zerfällt $\chi_{\mathrm{M}_{s,s}(\varphi)}$ über $\mathbb{C}$ in Linearfaktoren. Da nach Proposition (7.125) jedoch alle Eigenwerte von $\mathrm{M}_{s,s}(\varphi)$ reell sind, zerfällt folglich auch $\chi_\varphi = \chi_{\mathrm{M}_{s,s}(\varphi)}$ über $\mathbb{R}$ in Linearfaktoren. Da außerdem φ nach Bemerkung (7.126) normal ist, ist φ nach dem Spektralsatz (7.120) unitär diagonalisierbar. □

Singulärwertzerlegung

Zum Abschluss dieses Abschnitts liefern wir mit Hilfe der sogenannten Singulärwertzerlegung, siehe Korollar (7.138), eine Klassifikation von Matrizen modulo unitärer Äquivalenz, siehe Korollar (7.140). Die Singulärwertzerlegung hat eine wichtige Bedeutung in der numerischen linearen Algebra und führt in diesem Rahmen zu Anwendungen.

(7.128) Bemerkung. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist $\varphi^{\mathrm{ad}} \circ \varphi\colon V \to V$ hermitesch.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.129) Korollar. Es seien ein endlichdimensionaler K-Skalarproduktraum V, ein K-Skalarproduktraum W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Dann ist $\varphi^{\mathrm{ad}} \circ \varphi\colon V \to V$ unitär diagonalisierbar.

Beweis. Das Kompositum $\varphi^{\mathrm{ad}} \circ \varphi\colon V \to V$ ist hermitesch nach Bemerkung (7.128) und damit unitär diagonalisierbar nach Proposition (7.127). □

(7.130) Bemerkung. Es seien K-Skalarprodukträume V und W, ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $a \in K$ und ein normierter Eigenvektor v von $\varphi^{\mathrm{ad}} \circ \varphi\colon V \to V$ zum Eigenwert a gegeben. Dann ist

$$a = \|\varphi(v)\|^2.$$

Beweis. Es ist

$$\|\varphi(v)\|^2 = \langle \varphi(v), \varphi(v)\rangle = \langle v, \varphi^{\mathrm{ad}}(\varphi(v))\rangle = \langle v, av\rangle = a\langle v, v\rangle = a\|v\|^2 = a.$$ □

(7.131) Korollar. Es seien K-Skalarprodukträume V und W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Jeder Eigenwert von $\varphi^{\mathrm{ad}} \circ \varphi$ ist eine nicht-negative reelle Zahl.

Beweis. Es sei ein Eigenwert a von $\varphi^{\mathrm{ad}} \circ \varphi$ und ein zugehöriger Eigenvektor $v \in V \setminus \{0\}$ gegeben. Nach Bemerkung (7.15) ist dann $\frac{1}{\|v\|}v$ ein normierter Eigenvektor von $\varphi^{\mathrm{ad}} \circ \varphi\colon V \to V$ zum Eigenwert a. Nach Bemerkung (7.130) folgt

$$a = \|\varphi(v)\|^2 \geq 0.$$ □

(7.132) Definition (Singulärwerttupel)**.**

(a) Es seien $n \in \mathbb{N}_0$, ein n-dimensionaler K-Skalarproduktraum V, ein K-Skalarproduktraum W und ein adjungierbarer K-Vektorraumhomomorphismus $\varphi\colon V \to W$ gegeben. Ferner seien $r \in \mathbb{N}_0$ und ein n-Tupel $(a_1, \ldots, a_n)$ in $\mathbb{R}$ derart gegeben, dass

$$a_1 \geq \ldots \geq a_r > 0 = a_{r+1} = \ldots = a_n$$

gilt und

$$\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

eine Darstellungsmatrix von $\varphi^{\mathrm{ad}} \circ \varphi$ ist. Dann heißt

$$\sigma(\varphi) = (\sigma_1(\varphi), \ldots, \sigma_r(\varphi)) := (\sqrt{a_1}, \ldots, \sqrt{a_r})$$

das *Singulärwerttupel* von φ. Für $j \in [1, r]$ wird $\sigma_j(\varphi)$ der *j-te Singulärwert* von φ genannt.

(b) Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Dann heißt

$$\sigma(A) = (\sigma_1(A), \ldots, \sigma_r(A)) := \sigma(\varphi_A)$$

das *Singulärwerttupel* von A. Für $j \in [1, r]$ wird $\sigma_j(A)$ der *j-te Singulärwert* von A genannt.

(7.133) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben. Ferner seien $r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$ und ein n-Tupel $(a_1, \ldots, a_n)$ in $\mathbb{R}$ derart gegeben, dass

$$a_1 \geq \ldots \geq a_r > 0 = a_{r+1} = \ldots = a_n$$

gilt und $A^{\mathrm{ad}} A$ ähnlich zu

$$\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

ist. Dann ist

$$\sigma(A) = (\sqrt{a_1}, \ldots, \sqrt{a_r}).$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.134) Notation. Es seien $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$ und ein r-Tupel $\sigma = (\sigma_1, \ldots, \sigma_r)$ in $\mathbb{R}_{>0}$ gegeben. Wir notieren als Σ_σ die $(m \times n)$-Matrix mit Einträgen in $\mathbb{R}$ gegeben durch

$$\Sigma_\sigma = \Sigma_{\sigma_1, \ldots, \sigma_r} = \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right).$$

(7.135) Beispiel. Es ist $\Sigma_{\sqrt{2},1} \in \mathbb{R}^{3 \times 4}$ gegeben durch

$$\Sigma_{\sqrt{2},1} = \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(7.136) Bemerkung. Es seien $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$ und ein r-Tupel $\sigma = (\sigma_1, \ldots, \sigma_r)$ in $\mathbb{R}_{>0}$ gegeben. Das Singulärwerttupel von $\Sigma_\sigma \in K^{m \times n}$ ist

$$\sigma(\Sigma_\sigma) = \sigma.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(7.137) Bemerkung. Es seien K-Skalarprodukträume V und W, ein K-Vektorraumhomomorphismus $\varphi\colon V \to W$, $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$, ein n-Tupel $(a_1, \dots, a_n)$ in $\mathbb{R}$ mit

$$a_1 \geq \ldots \geq a_r > 0 = a_{r+1} = \ldots = a_n,$$

eine Orthonormalbasis $s = (s_1, \dots, s_n)$ von V und eine Basis $t = (t_1, \dots, t_m)$ von W so gegeben, dass s_j für $j \in [1, n]$ ein Eigenvektor von $\varphi^{\mathrm{ad}} \circ \varphi\colon V \to V$ zum Eigenwert a_j und $t_j = \frac{1}{\sqrt{a_j}} \varphi(s_j)$ für $j \in [1, r]$ ist. Dann ist

$$\mathrm{M}_{t,s}(\varphi) = \Sigma_{\sqrt{a_1}, \dots, \sqrt{a_r}}.$$

Beweis. Für $j \in [1, n]$ gilt $\varphi(s_j) = \sqrt{a_j} t_j$, also

$$\kappa_t(\varphi(s_j)) = \begin{cases} \sqrt{a_j} \mathrm{e}_j, & \text{falls } j \in [1, r], \\ 0, & \text{falls } j \in [r+1, n]. \end{cases}$$

Es folgt

$$\mathrm{M}_{t,s}(\varphi) = \begin{pmatrix} \kappa_t(\varphi(s_1)) & \ldots & \kappa_t(\varphi(s_r)) & \kappa_t(\varphi(s_{r+1})) & \ldots & \kappa_t(\varphi(s_n)) \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{a_1} \mathrm{e}_1 & \ldots & \sqrt{a_r} \mathrm{e}_r & 0 & \ldots & 0 \end{pmatrix} = \left(\begin{array}{ccc|c} \sqrt{a_1} & & 0 & \\ & \ddots & & 0 \\ 0 & & \sqrt{a_r} & \\ \hline & 0 & & 0 \end{array}\right) = \Sigma_{\sqrt{a_1}, \dots, \sqrt{a_r}}. \qquad \square$$

(7.138) Korollar (Singulärwertzerlegung)**.** Es seien $m, n \in \mathbb{N}_0$ und ein $A \in K^{m \times n}$ gegeben. Dann ist A unitär äquivalent zu $\Sigma_{\sigma(A)}$.

Beweis. Nach Korollar (7.129) ist $A^{\mathrm{ad}} A$ unitär diagonalisierbar, d.h. es gibt $r \in \mathbb{N}_0$ und ein n-Tupel $(a_1, \dots, a_n)$ in $\mathbb{R}$ mit

$$a_1 \geq \ldots \geq a_r > 0 = a_{r+1} = \ldots = a_n$$

und eine Orthonormalbasis $s = (s_1, \dots, s_n)$ von $K^{n \times 1}$ so, dass s_j für $j \in [1, n]$ ein Eigenvektor von $A^{\mathrm{ad}} A$ zum Eigenwert a_j ist. Dann ist $(s_{r+1}, \dots, s_n)$ eine Basis von $\mathrm{Eig}_0(A^{\mathrm{ad}} A) = \mathrm{Sol}(A^{\mathrm{ad}} A, 0) = \mathrm{Sol}(A, 0)$ nach Bemerkung (6.4) und Proposition (7.62), und damit $(A s_1, \dots, A s_r)$ eine Basis von $\mathrm{Col}(A)$ nach Proposition (2.27). Für $i, j \in [1, r]$ gilt

$$\langle A s_i, A s_j \rangle = \langle s_i, A^{\mathrm{ad}} A s_j \rangle = \langle s_i, a_j s_j \rangle = a_j \langle s_i, s_j \rangle = a_j \delta_{i,j}.$$

Nach Korollar (1.23)(c), Korollar (1.38)(c) und Bemerkung (7.15) ist daher

$$\left(\frac{1}{\sqrt{a_1}} A s_1, \dots, \frac{1}{\sqrt{a_r}} A s_r\right) = \left(\frac{1}{\|A s_1\|} A s_1, \dots, \frac{1}{\|A s_r\|} A s_r\right)$$

eine Orthonormalbasis von $\mathrm{Col}(A)$, so dass es nach Korollar (7.48) und Bemerkung (7.15) eine Orthonormalbasis $t = (t_1, \dots, t_m)$ von $K^{m \times 1}$ mit $t_j = \frac{1}{\sqrt{a_j}} \varphi(s_j)$ für $j \in [1, r]$ gibt.
Nach Bemerkung (7.104)(a) ist $A = \mathrm{M}_{\mathrm{e},\mathrm{e}}(\varphi_A)$ äquivalent zu $\mathrm{M}_{t,s}(\varphi_A)$ und nach Bemerkung (7.137) und Bemerkung (7.133) gilt

$$\mathrm{M}_{t,s}(\varphi_A) = \Sigma_{\sqrt{a_1}, \dots, \sqrt{a_r}} = \Sigma_{\sigma(A)}. \qquad \square$$

Ausgeschrieben bedeutet die Singulärwertzerlegung (7.138), dass es für $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ stets $P \in \mathrm{U}_n(K)$ und $Q \in \mathrm{U}_m(K)$ mit $P^{-1} A Q = \Sigma_{\sigma(A)}$ gibt, also mit

$$A = Q \Sigma_{\sigma(A)} P^{-1} = Q \Sigma_{\sigma(A)} P^{\mathrm{ad}}.$$

(7.139) Satz. Es seien $m, n \in \mathbb{N}_0$ und $A, B \in K^{m \times n}$ gegeben. Genau dann ist A unitär äquivalent zu B, wenn

$$\sigma(A) = \sigma(B)$$

ist.

Beweis. Zunächst sei A unitär äquivalent zu B, so dass es $P \in \mathrm{U}_n(K)$, $Q \in \mathrm{U}_m(K)$ mit $Q^{-1}AP = B$ gibt. Dann ist

$$P^{-1}A^{\mathrm{ad}}AP = P^{\mathrm{ad}}A^{\mathrm{ad}}(Q^{-1})^{\mathrm{ad}}Q^{-1}AP = (Q^{-1}AP)^{\mathrm{ad}}\,Q^{-1}AP = B^{\mathrm{ad}}B,$$

d.h. $A^{\mathrm{ad}}A$ und $B^{\mathrm{ad}}B$ sind unitär ähnlich. Nach Korollar (7.129) ist $A^{\mathrm{ad}}A$ unitär diagonalisierbar und nach Korollar (7.131) ist jeder Eigenwert von $A^{\mathrm{ad}}A$ eine nicht-negative reelle Zahl. Folglich gibt es $r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$ und ein n-Tupel $(a_1, \dots, a_n)$ in $\mathbb{R}$ mit

$$a_1 \geq \ldots \geq a_r > 0 = a_{r+1} = \ldots = a_n,$$

und derart, dass $A^{\mathrm{ad}}A$ und damit auch $B^{\mathrm{ad}}B$ unitär ähnlich zu

$$\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

ist. Wir erhalten

$$\sigma(A) = (\sqrt{a_1}, \dots, \sqrt{a_r}) = \sigma(B).$$

Nun sei umgekehrt $\sigma(A) = \sigma(B)$. Nach der Singulärwertzerlegung (7.138) ist A unitär äquivalent zu $\Sigma_{\sigma(A)}$ und B unitär äquivalent zu $\Sigma_{\sigma(B)}$. Da aber $\sigma(A) = \sigma(B)$ ist, folgt $\Sigma_{\sigma(A)} = \Sigma_{\sigma(B)}$, und da unitäre Äquivalenz von Matrizen nach Bemerkung (7.103)(a) eine Äquivalenzrelation auf $K^{m \times n}$ ist, ist A folglich unitär äquivalent zu B. □

(7.140) Korollar (Klassifikation von Matrizen modulo unitärer Äquivalenz)**.** Es seien $m, n, r \in \mathbb{N}_0$ mit $r \leq \min(m, n)$, ein r-Tupel $\sigma = (\sigma_1, \dots, \sigma_r)$ in $\mathbb{R}_{>0}$ und $A \in K^{m \times n}$ gegeben. Genau dann ist

$$\sigma(A) = \sigma,$$

wenn A unitär äquivalent zu Σ_σ ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

A Grundlagen

Wir wiederholen einige Grundlagen aus der Vorlesung *Diskrete Strukturen* [12].
Obwohl wir auch Beispiele geben, ist diese Zusammenfassung sehr knapp gehalten. Es werden nur wenige Eigenschaften genannt und fast keine Beweise gegeben. Die Zusammenfassung eignet sich daher *nicht* zum erstmaligen Lernen der Begriffe und sollte ggf. durch weitere Literatur ergänzt werden. Ihr Ziel ist es, eine Orientierung hinsichtlich des für die Lineare Algebra notwendigen Vorwissens zu geben.

Mengen

Die durch Anführungsstriche markierten Wörter in diesem Abschnitt werden nicht genauer präzisiert.

(A.1) Vorstellung (Menge; Cantor)**.**

(a) Unter einer *Menge* verstehen wir eine „Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen“.

(b) Es sei eine Menge X gegeben. Diejenigen Objekte, welche durch X zusammengefasst werden, bezeichnen wir als *Elemente* von X. Ist ein Objekt x ein Element von X, so schreiben wir $x \in X$, anderenfalls $x \notin X$.

(c) Mengen X und Y sind *gleich*, geschrieben $X = Y$, falls sie die gleichen Elemente enthalten, d.h. falls für jedes Objekt x genau dann $x \in X$ gilt, wenn $x \in Y$ gilt.

Im Folgenden werden wir einige Notationen zur Beschreibung von Mengen angeben. In aller Regel erfolgt eine solche Beschreibung durch die Angabe einer „Eigenschaft", welche die Elemente einer Menge erfüllen, oder durch eine einfache „Aufzählung" ihrer Elemente. Letzteres wird vor allem bei einer Menge mit „endlich" vielen Elementen gemacht – etwas unpräzise aber auch bei „unendlich" vielen Elementen, sofern aus dem Kontext klar ist (bzw. klar sein sollte), welche Objekte aufgezählt werden.

(A.2) Notation.

(a) Es seien eine Menge X und eine Eigenschaft φ so gegeben, dass für jedes Objekt x genau dann $x \in X$ gilt, wenn x die Eigenschaft φ erfüllt. Wir schreiben

$$\{x \mid x \text{ erfüllt } \varphi\} := X.$$

(b) Es sei eine Menge X gegeben. Für eine Eigenschaft φ schreiben wir

$$\{x \in X \mid x \text{ erfüllt } \varphi\} := \{x \mid x \in X \text{ und } x \text{ erfüllt } \varphi\}.$$

(c) Es seien Objekte a_1, …, a_n gegeben. Wir schreiben

$$\{a_1, \dots, a_n\} := \{x \mid x = a_1 \text{ oder } \dots \text{ oder } x = a_n\}.$$

(d) Für jede natürliche Zahl i sei ein Objekt a_i gegeben. Wir schreiben

$$\{a_1, a_2, a_3, \dots\} := \{x \mid \text{es gibt eine natürliche Zahl } i \text{ mit } x = a_i\}.$$

(A.3) Beispiel. Wir wollen davon ausgehen, dass wir wissen, was die folgenden Mengen sind.

(a) Die *Menge der natürlichen Zahlen* ist durch

$$\mathbb{N} = \{x \mid x = 1 \text{ oder es gibt eine natürliche Zahl } y \text{ mit } x = y + 1\} = \{1, 2, 3, \dots\}$$

gegeben. Die *Menge der natürlichen Zahlen mit Null* ist durch

$$\mathbb{N}_0 = \{x \mid x \in \mathbb{N} \text{ oder } x = 0\}$$

gegeben.

(b) Die *Menge der ganzen Zahlen* ist durch

$$\mathbb{Z} = \{x \mid x \in \mathbb{N} \text{ oder } x = 0 \text{ oder } -x \in \mathbb{N}\}$$

gegeben.

(c) Die *Menge der rationalen Zahlen* ist durch

$$\mathbb{Q} = \{x \mid \text{es gibt } p, q \in \mathbb{Z} \text{ mit } q \neq 0 \text{ und } x = \frac{p}{q}\}$$

gegeben.

(d) Die *Menge der reellen Zahlen* wird als $\mathbb{R}$ notiert.

(A.4) Beispiel.

(a) Es ist $\{1, 3, 17\}$ eine Menge.

(b) Es sind $\{1\}$, $\{\{1\}\}$ und $\{1, \{1\}\}$ Mengen.

(c) Es ist $\{x \mid x \text{ ist eine Primzahl}\}$ eine Menge.

(d) Es ist $\{x \in \mathbb{Z} \mid x \text{ ist gerade}\}$ eine Menge.

(A.5) Beispiel.

(a) Es ist $\{1, 3, 17\} = \{3, 1, 17\} = \{1, 3, 17, 1\}$.

(b) Es ist $\{1\} = \{1, 1, 1\} \neq \{1, 2\}$.

(c) Es ist $\{1\} \neq \{\{1\}\}$ und $\{1\} \neq \{1, \{1\}\}$ und $\{\{1\}\} \neq \{1, \{1\}\}$.

(d) Es ist $\{x \in \mathbb{R} \mid x^3 + 2x = 3x^2\} = \{0, 1, 2\}$.

(A.6) Definition (leere Menge)**.** Die Menge, welche keine Elemente enthält, heißt *leere Menge* und wird als $\emptyset$ notiert.

(A.7) Notation. Für $a, b \in \mathbb{Z}$ mit $a \leq b + 1$ schreiben wir

$$[a, b] := \{x \in \mathbb{Z} \mid a \leq x \leq b\}.$$

(A.8) Definition (Teilmenge)**.** Es sei eine Menge X gegeben. Eine *Teilmenge* von X ist eine Menge U derart, dass aus $u \in U$ stets $u \in X$ folgt.
Eine Teilmenge U von X heißt *echt* (oder *strikt*), falls $U \neq X$ gilt.
Es sei eine Menge U gegeben. Ist U eine Teilmenge von X, so schreiben wir $U \subseteq X$. Ist U keine Teilmenge von X, so schreiben wir $U \nsubseteq X$. Ist U eine echte Teilmenge von X, so schreiben wir $U \subset X$.

(A.9) Beispiel.

(a) Es ist $\{2, 3, 4, 7\}$ eine Teilmenge von $\{1, 2, 3, 4, 5, 6, 7\}$.

(b) Es ist $\{0, 3, 4, 7\}$ keine Teilmenge von $\{1, 2, 3, 4, 5, 6, 7\}$.

(c) Es gilt $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$, d.h. es ist $\mathbb{N}$ eine Teilmenge von $\mathbb{Z}$, es ist $\mathbb{Z}$ eine Teilmenge von $\mathbb{Q}$ und es ist $\mathbb{Q}$ eine Teilmenge von $\mathbb{R}$.

(A.10) Bemerkung (Gleichheitskriterium für Mengen)**.** Es seien Mengen X und Y gegeben. Genau dann ist $X = Y$, wenn $X \subseteq Y$ und $Y \subseteq X$ gilt.

Beweis. Siehe [12, Bem. (2.13)]. □

(A.11) Vorstellung (Familie)**.**

(a) Es sei eine Menge I gegeben. Eine *Familie* über I ist ein „Ausdruck“ der Form $x = (x_i)_{i \in I}$ für gewisse Objekte x_i für $i \in I$. Die Menge I wird *Indexmenge* von x genannt. Ein Element von I wird *Index* (oder *Stelle*) von x genannt. Für $i \in I$ wird x_i die *Komponente* (oder der *Eintrag*) von x an der Stelle i genannt.

(b) Es seien Mengen I und J, eine Familie $x = (x_i)_{i \in I}$ über I und eine Familie $y = (y_j)_{j \in J}$ über J gegeben. Die Familien $x = (x_i)_{i \in I}$ und $y = (y_j)_{j \in J}$ sind *gleich*, geschrieben $x = y$, wenn $I = J$ und für $i \in I$ stets $x_i = y_i$ gilt.

(A.12) Beispiel.

(a) Es ist $x = (x_i)_{i \in \{1,2,3\}}$ gegeben durch $x_1 = 3$, $x_2 = -\frac{2}{5}$, $x_3 = \sqrt{7}$ eine Familie über $\{1, 2, 3\}$.

(b) Es ist $x = (x_i)_{i \in \mathbb{N}_0}$ gegeben durch $x_i = i^2 + 1$ für $i \in \mathbb{N}_0$ eine Familie über $\mathbb{N}_0$.

(c) Es ist $(i^2 - 1)_{i \in \mathbb{Z}}$ eine Familie über $\mathbb{Z}$.

(d) Es sei $I := \{\emptyset, \{1\}, \{1, 2\}\}$. Dann ist $x = (x_i)_{i \in I}$ gegeben durch $x_\emptyset = 0$, $x_{\{1\}} = \{\{1\}\}$, $x_{\{1,2\}} = \{3\}$ eine Familie über I.

(e) Es gibt keine Familie $x = (x_i)_{i \in \{2,5\}}$ über $\{2, 5\}$ mit $x_2 = 0$ und $x_2 = 1$.

(A.13) Definition (Familie)**.** Es seien eine Menge X und eine Menge I gegeben. Die *Menge der Familien* in X über I ist definiert als

$$X^I := \{x \mid x \text{ ist eine Familie über } I \text{ mit } x_i \in X \text{ für } i \in I\}.$$

Ein Element von X^I wird eine *Familie* in X über I genannt.

(A.14) Definition (Tupel)**.** Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Ein *n-Tupel* ist eine Familie über $[1, n]$. ([8])
Für ein n-Tupel x schreiben wir auch

$$(x_1, \dots, x_n) := x.$$

(b) Es sei eine Menge X gegeben. Die *Menge der n-Tupel* in X ist definiert als

$$X^n := X^{[1,n]}.$$

Ein Element von X^n wird *n-Tupel* in X genannt.

(A.15) Definition (leeres Tupel, Paar, Tripel, Quadrupel, Quintupel)**.**

(a) Das 0-Tupel wird auch das *leere Tupel* genannt.

(b) Ein 1-Tupel wird auch *Single* genannt.

(c) Ein 2-Tupel wird auch *Paar* genannt.

(d) Ein 3-Tupel wird auch *Tripel* genannt.

(e) Ein 4-Tupel wird auch *Quadrupel* genannt.

(f) Ein 5-Tupel wird auch *Quintupel* genannt.

(A.16) Beispiel.

(a) Es ist $(1, 2, 4)$ ein Tripel.

(b) Es ist $x = (x_i)_{i \in [1,9]}$ gegeben durch $x_i = i - i^2$ für $i \in [1, 9]$ ein 9-Tupel.

(c) Es ist $(\{1\}, \{2\})$ ein Paar.

(d) Es gibt kein Quadrupel x in $\mathbb{Z}$ mit $x_2 = 0$ und $x_2 = 1$.

(A.17) Beispiel.

(a) Es sei ein Quadrupel x gegeben durch

$$x = (i^3)_{i \in [1,4]}.$$

Dann ist $x = (1, 8, 27, 64)$.

(b) Es ist $(3, \{4\}) \neq (\{4\}, 3)$.

(c) Es ist $(1, 2, 3) \neq (1, 2, 3, 2)$.

(A.18) Definition (Differenz)**.** Es seien Mengen X und Y gegeben. Die Menge

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

heißt *Differenz* (oder *Mengendifferenz*) von X und Y.

[8] In manchen Texten werden auch Familien über $[0, n-1]$ als n-Tupel bezeichnet. Allgemeiner: Für $k \in \mathbb{Z}$ wird eine Familie über $[k+1, k+n]$ auch ein *durch* $[k+1, k+n]$ *indiziertes n-Tupel* genannt.

(A.19) Beispiel. Es ist

$$\{1,2,3\} \setminus \{2,4\} = \{1,3\},$$
$$\{2,4\} \setminus \{1,2,3\} = \{4\}.$$

(A.20) Definition (Schnitt, Vereinigung)**.**

(a) (i) Es seien eine nicht-leere Menge I und eine Familie von Mengen $X = (X_i)_{i \in I}$ über I ([9]) gegeben. Die Menge

$$\bigcap X = \bigcap_{i \in I} X_i := \{x \mid x \in X_i \text{ für } i \in I\}$$

heißt *Schnitt* (oder *Durchschnitt*) von X.

(ii) Es seien $n \in \mathbb{N}$ und ein n-Tupel von Mengen $(X_1, \dots, X_n)$ gegeben. Wir schreiben auch

$$X_1 \cap \ldots \cap X_n := \bigcap X.$$

(iii) Es seien Mengen X und Y gegeben. Der Schnitt $X \cap Y$ von (X, Y) wird auch *Schnitt* (oder *Durchschnitt*) von X und Y genannt.

(b) (i) Es seien eine Menge I und eine Familie von Mengen $X = (X_i)_{i \in I}$ über I gegeben. Die Menge

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

heißt *Vereinigung* von X.

(ii) Es seien $n \in \mathbb{N}_0$ und ein n-Tupel von Mengen $(X_1, \dots, X_n)$ gegeben. Wir schreiben auch

$$X_1 \cup \ldots \cup X_n := \bigcup X.$$

(iii) Es seien Mengen X und Y gegeben. Die Vereinigung $X \cup Y$ von (X, Y) wird auch *Vereinigung* von X und Y genannt.

(A.21) Beispiel.

(a) (i) Es ist

$$\{1,2,3\} \cap \{2,4\} = \{2\}.$$

(ii) Es ist

$$\bigcap_{i \in \mathbb{N}} \{qi \mid q \in \mathbb{Z}\} = \{0\}.$$

(b) (i) Es ist

$$\{1,2,3\} \cup \{2,4\} = \{1,2,3,4\}.$$

(ii) Es ist

$$\bigcup_{i \in \mathbb{N}} \{qi \mid q \in \mathbb{Z}\} = \mathbb{Z}.$$

(A.22) Definition (Disjunktheit)**.**

(a) Es seien eine Menge I und eine Familie von Mengen $X = (X_i)_{i \in I}$ über I gegeben. Wir sagen, dass X *disjunkt* ist, falls für $i, j \in I$ mit $i \neq j$ stets $X_i \cap X_j = \emptyset$ gilt.

[9] Wir nehmen also an, dass X_i für $i \in I$ stets eine Menge ist.

(b) Es seien Mengen X und Y gegeben. Wir sagen, dass X und Y *disjunkt* sind, falls (X, Y) disjunkt ist.

(A.23) Beispiel. Es seien $X := \{1, 2, 4\}$, $Y := \{3, 6\}$, $Z := \{5, 6\}$.

(a) Die Menge X und Y sind disjunkt.

(b) Die Menge X und Z sind disjunkt.

(c) Die Mengen Y und Z sind nicht disjunkt.

(A.24) Definition ((innere) disjunkte Vereinigung)**.**

(a) Es seien eine Menge I und eine Familie von Mengen $X = (X_i)_{i \in I}$ über I gegeben. Wenn X disjunkt ist, so sagen wir, dass $\bigcup X$ eine (*innere*) *disjunkte Vereinigung* von X ist, und schreiben

$$\dot{\bigcup} X = \dot{\bigcup_{i \in I}} X_i := \bigcup X.$$

(b) Es seien $n \in \mathbb{N}_0$ und ein disjunktes n-Tupel von Mengen $(X_1, \dots, X_n)$ gegeben. Wir schreiben auch

$$X_1 \mathbin{\dot{\cup}} \cdots \mathbin{\dot{\cup}} X_n := \dot{\bigcup} X.$$

(c) Es seien disjunkte Mengen X und Y gegeben. Die disjunkte Vereinigung $X \mathbin{\dot{\cup}} Y$ von (X, Y) wird auch (*innere*) *disjunkte Vereinigung* von X und Y genannt.

(A.25) Beispiel. Es ist

$$\mathbb{N} = \{n \in \mathbb{N} \mid n \text{ ist gerade}\} \mathbin{\dot{\cup}} \{n \in \mathbb{N} \mid n \text{ ist ungerade}\}.$$

(A.26) Definition (kartesisches Produkt)**.**

(a) Es seien $n \in \mathbb{N}_0$ und ein n-Tupel von Mengen $(X_1, \dots, X_n)$ gegeben. Die Menge

$$X_1 \times \ldots \times X_n := \{x \mid x = (x_1, \dots, x_n) \text{ ist ein } n\text{-Tupel mit } x_i \in X_i \text{ für } i \in [1, n]\}$$

heißt *kartesisches Produkt* von $(X_1, \dots, X_n)$.

(b) Es seien Mengen X und Y gegeben. Das kartesische Produkt $X \times Y$ von (X, Y) wird auch *kartesisches Produkt* von X und Y genannt.

(A.27) Beispiel. Es ist

$$\{1, 2\} \times \{3, 4, 5\} = \{(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5)\}.$$

(A.28) Vorstellung ((un)endliche Menge)**.** Eine Menge X heißt *endlich*, wenn sie endlich viele Elemente besitzt, und ansonsten *unendlich*.

(A.29) Beispiel.

(a) Die Menge $\{1, 3, 17\}$ ist endlich.

(b) Die Mengen $\mathbb{N}$ und $\{x \in \mathbb{N} \mid x \text{ gerade}\}$ sind unendlich.

(c) Die leere Menge ist endlich.

(d) Die Menge $\{x \in \mathbb{R} \mid x^3 + 2x = 3x^2\}$ ist endlich.

(A.30) Vorstellung (Kardinalität)**.** Es sei eine endliche Menge X gegeben. Die *Kardinalität* (oder *Mächtigkeit*) von X ist die Anzahl der Elemente von X und wird mit $|X|$ bezeichnet.

(A.31) Beispiel.

(a) Es ist $|\{1, 3, 17\}| = 3$.

(b) Es ist $|\{1, 1, 1\}| = 1$.

(c) Es ist $|\{\{1\}\}| = 1$.

(d) Es ist $|\{1, \{1\}\}| = 2$.

Abbildungen

(A.32) Definition (Abbildung)**.**

(a) Eine *Abbildung* (oder *Funktion*) besteht aus Mengen X und Y zusammen mit einer Teilmenge f von $X \times Y$ so, dass es für jedes $x \in X$ genau ein $y \in Y$ mit $(x, y) \in f$ gibt. Unter Missbrauch der Notation bezeichnen wir sowohl die besagte Abbildung als auch die Teilmenge von $X \times Y$ mit f. Die Menge X wird *Startmenge* (oder *Definitionsbereich*) von f genannt. Ein Element von X wird *Argument* von f genannt. Die Menge Y wird *Zielmenge* (oder *Wertebereich*) von f genannt. Ein Element von Y wird *Zielwert* (oder *Zielelement*) von f genannt.

Für eine Abbildung f mit Startmenge X und Zielmenge Y schreiben wir $\operatorname{Source} f := X$ und $\operatorname{Target} f := Y$. Für $x \in X$ heißt das Element $y \in Y$ mit $(x, y) \in f$ das *Bild* (oder *Bildelement*) von x unter f, wir schreiben $f(x) := y$. Für $y \in Y$, $x \in X$ mit $y = f(x)$ wird x ein *Urbild* (oder *Urbildelement*) von y unter f genannt.

(b) Es seien Mengen X und Y gegeben. Die *Menge der Abbildungen* von X nach Y ist definiert als

$$\operatorname{Map}(X, Y) := \{f \mid f \text{ ist eine Abbildung mit } \operatorname{Source} f = X \text{ und } \operatorname{Target} f = Y\}.$$

Ein Element von $\operatorname{Map}(X, Y)$ wird *Abbildung* von X nach Y genannt; wir schreiben $f \colon X \to Y$ sowie $f \colon X \to Y$, $x \mapsto f(x)$ für $f \in \operatorname{Map}(X, Y)$.

(A.33) Beispiel.

(a) Es ist $\{1, 2, 3\} \to \{4, 5, 6\}$, $1 \mapsto 4$, $2 \mapsto 5$, $3 \mapsto 4$ eine Abbildung.

(b) Es ist $\mathbb{Z} \to \mathbb{Q}$, $x \mapsto 2x^2$ eine Abbildung.

(c) Es gibt keine Abbildung $f \colon \mathbb{N} \to \mathbb{N}$ mit $f(x) = \sqrt{x}$ für $x \in \mathbb{N}$.

(A.34) Beispiel. Es ist

$$\operatorname{Map}(\{1, 2\}, \{3, 4\}) = \{(1 \mapsto 3, 2 \mapsto 3), (1 \mapsto 3, 2 \mapsto 4), (1 \mapsto 4, 2 \mapsto 3), (1 \mapsto 4, 2 \mapsto 4)\},$$

wobei wir etwa $(1 \mapsto 3, 2 \mapsto 3)$ als Kurzschreibweise für die Abbildung $\{1, 2\} \to \{3, 4\}$, $1 \mapsto 3$, $2 \mapsto 3$ verwendet haben.

(A.35) Bemerkung (Gleichheitskriterium für Abbildungen)**.** Es seien Abbildungen $f \colon X \to Y$ und $f' \colon X' \to Y'$ gegeben. Genau dann gilt $f = f'$, wenn $X = X'$, $Y = Y'$ und $f(x) = f'(x)$ in Y für alle $x \in X$ ist.

Beweis. Siehe [12, Bem. (3.6)]. □

(A.36) Konvention. Es seien Mengen X und I gegeben.

(a) Es sei eine Familie x in X über I gegeben. Unter Missbrauch der Notation bezeichnen wir die Abbildung $I \to X$, $i \mapsto x_i$ wieder als x.

(b) Es sei eine Abbildung $f \colon I \to X$ gegeben. Unter Missbrauch der Notation bezeichnen wir die Familie $(f(i))_{i \in I}$ wieder als f.

(A.37) Definition (Kompositum)**.** Es seien Abbildungen $f \colon X \to Y$ und $g \colon Y \to Z$ gegeben. Die Abbildung

$$g \circ f \colon X \to Z,\ x \mapsto g(f(x))$$

heißt *Kompositum* von f und g.

(A.38) Beispiel. Es seien $f \colon \mathbb{N} \to \mathbb{Z}$, $x \mapsto x + 1$ und $g \colon \mathbb{Z} \to \mathbb{Q}$, $y \mapsto 2y^2$. Dann ist $g \circ f \colon \mathbb{N} \to \mathbb{Q}$, $x \mapsto 2(x + 1)^2$.

(A.39) Definition (identische Abbildung)**.** Es sei eine Menge X gegeben. Die Abbildung

$$\mathrm{id} = \mathrm{id}_X \colon X \to X,\ x \mapsto x$$

heißt *Identität* (oder *identische Abbildung*) auf X.

(A.40) Beispiel. Die Identität auf $\{1, 2, 3\}$ ist gegeben durch

$$\mathrm{id}_{\{1,2,3\}} \colon \{1, 2, 3\} \to \{1, 2, 3\},\ 1 \mapsto 1,\ 2 \mapsto 2,\ 3 \mapsto 3.$$

(A.41) Definition (Invertierbarkeit von Abbildungen)**.** Es sei eine Abbildung $f \colon X \to Y$ gegeben.

(a) Eine *Inverse* (oder *inverse Abbildung* oder *Umkehrabbildung*) zu f ist eine Abbildung $g \colon Y \to X$ derart, dass $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$ gilt.

(b) Die Abbildung f heißt *invertierbar*, falls es eine Inverse zu f gibt.

(A.42) Beispiel. Es seien $\mathbb{Q}_{>0} := \{x \in \mathbb{Q} \mid x > 0\}$ und $\mathbb{Q}_{<0} := \{x \in \mathbb{Q} \mid x < 0\}$. Es seien $f \colon \mathbb{Q}_{>0} \to \mathbb{Q}_{<0}$, $x \mapsto -x$ und $g \colon \mathbb{Q}_{<0} \to \mathbb{Q}_{>0}$, $y \mapsto -y$. Dann ist g eine zu f inverse Abbildung.

(A.43) Bemerkung. Es sei eine Abbildung $f \colon X \to Y$ gegeben. Dann gibt es höchstens eine Inverse zu f.

Beweis. Siehe [12, Bem. (3.19)]. □

(A.44) Notation. Die zu einer invertierbaren Abbildung $f \colon X \to Y$ gegebene inverse Abbildung notieren wir als $f^{-1} \colon Y \to X$.

(A.45) Definition (injektiv, surjektiv)**.** Es sei eine Abbildung $f \colon X \to Y$ gegeben.

(a) Wir sagen, dass f *injektiv* (oder eine *Injektion*) ist, falls für alle $x, x' \in X$ mit $f(x) = f(x')$ in Y stets $x = x'$ in X folgt.

(b) Wir sagen, dass f *surjektiv* (oder eine *Surjektion*) ist, falls für jedes $y \in Y$ ein $x \in X$ mit $y = f(x)$ existiert.

(c) Wir sagen, dass f *bijektiv* (oder eine *Bijektion*) ist, falls f injektiv und surjektiv ist.

(A.46) Beispiel.

(a) Die Abbildung $f \colon \{1, 2, 3\} \to \{4, 5\}$, $1 \mapsto 4$, $2 \mapsto 4$, $3 \mapsto 5$ ist surjektiv, aber nicht injektiv.

(b) Die Abbildung $f \colon \{1, 2\} \to \{4, 5, 6\}$, $1 \mapsto 4$, $2 \mapsto 5$ ist injektiv, aber nicht surjektiv.

(c) Die Abbildung $f \colon \{1, 2, 3\} \to \{4, 5, 6\}$, $1 \mapsto 5$, $2 \mapsto 6$, $3 \mapsto 4$ ist bijektiv.

(d) Die Abbildung $f \colon \{1, 2, 3\} \to \{4, 5, 6\}$, $1 \mapsto 5$, $2 \mapsto 6$, $3 \mapsto 5$ ist weder injektiv noch surjektiv.

(A.47) Definition (Bild, Urbild)**.** Es sei eine Abbildung $f \colon X \to Y$ gegeben.

(a) Für eine Teilmenge U von X heißt

$$f(U) := \{f(u) \mid u \in U\} := \{y \in Y \mid \text{es gibt ein } u \in U \text{ mit } y = f(u)\}$$

das *Bild* von U unter f. Ferner heißt $\operatorname{Im} f := f(X)$ das *Bild* von f.

(b) Für eine Teilmenge V von Y heißt

$$f^{-1}(V) := \{x \in X \mid f(x) \in V\}$$

das *Urbild* von V unter f. Für $y \in Y$ heißt $f^{-1}(\{y\})$ die *Faser* von f über y.

Die Notation von Urbild und Faser setzt nicht die Existenz einer inversen Abbildung voraus – wir haben es mit einer mehrdeutigen Schreibweise zu tun.

(A.48) Beispiel. Es sei $f \colon \{1, 2, 3, 4\} \to \{5, 6, 7, 8\}$, $1 \mapsto 5$, $2 \mapsto 7$, $3 \mapsto 5$, $4 \mapsto 8$. Dann ist $f(\{1, 2, 3\}) = \{5, 7\}$, $\operatorname{Im} f = \{5, 7, 8\}$, $f^{-1}(\{5, 8\}) = \{1, 3, 4\}$, $f^{-1}(\{5\}) = \{1, 3\}$.

(A.49) Satz. Es sei eine Abbildung $f \colon X \to Y$ gegeben.

(a) Die folgenden Aussagen sind äquivalent.

(i) Die Abbildung f ist injektiv.

(ii) Jede Faser von f besitzt höchstens ein Element.
(iii) Es ist $X = \emptyset$ oder es gibt eine Abbildung $g\colon Y \to X$ mit $g \circ f = \mathrm{id}_X$.

(b) Die folgenden Aussagen sind äquivalent.

(i) Die Abbildung f ist surjektiv.
(ii) Jede Faser von f besitzt mindestens ein Element.
(iii) Es gibt eine Abbildung $g\colon Y \to X$ mit $f \circ g = \mathrm{id}_Y$.

(c) Die folgenden Aussagen sind äquivalent.

(i) Die Abbildung f ist bijektiv.
(ii) Jede Faser von f besitzt genau ein Element.
(iii) Die Abbildung f ist invertierbar.

Beweis. Siehe [12, Satz (3.29)]. □

(A.50) Definition (Restriktion)**.** Es seien eine Abbildung $f\colon X \to Y$, eine Teilmenge U von X und eine Teilmenge V von Y mit $f(U) \subseteq V$ gegeben. Die Abbildung

$$f|_U^V\colon U \to V,\, u \mapsto f(u)$$

wird *Restriktion* (oder *Einschränkung*) von f bzgl. U und V genannt.
Für $U \subseteq X$ setzen wir

$$f|_U := f|_U^Y.$$

Für $V \subseteq Y$ mit $\operatorname{Im} f \subseteq V$ setzen wir

$$f|^V := f|_X^V.$$

(A.51) Beispiel. Es sei $f\colon \{2,3,5,7,11\} \to \{0,1,2,3\}$, $2 \mapsto 2$, $3 \mapsto 3$, $5 \mapsto 1$, $7 \mapsto 3$, $11 \mapsto 3$.

(a) Es ist

$$f|_{\{3,7,11\}}^{\{0,1,3\}}\colon \{3,7,11\} \to \{0,1,3\},\ 3 \mapsto 3,\ 7 \mapsto 3,\ 11 \mapsto 3.$$

(b) Es ist

$$f|_{\{2,7,11\}}\colon \{2,7,11\} \to \{0,1,2,3\},\ 2 \mapsto 2,\ 7 \mapsto 3,\ 11 \mapsto 3.$$

(c) Es ist

$$f|^{\{1,2,3\}}\colon \{2,3,5,7,11\} \to \{1,2,3\},\ 2 \mapsto 2,\ 3 \mapsto 3,\ 5 \mapsto 1,\ 7 \mapsto 3,\ 11 \mapsto 3.$$

Äquivalenzrelationen und Quotientenmengen

(A.52) Definition (Relation)**.** Es sei eine Menge X gegeben. Eine *Relation* (genauer *binäre Relation*) auf X ist eine Teilmenge von $X \times X$.
Es seien eine Relation r auf X und $x, y \in X$ gegeben. Falls $(x, y) \in r$ ist, so sagen wir, dass x bzgl. r in *Relation* zu y *steht* und schreiben $x \mathrel{r} y$.

(A.53) Beispiel.

(a) Für $m, n \in \mathbb{N}$ gelte $m < n$, falls es ein $p \in \mathbb{N}$ mit $n = p + m$ gibt. Dann ist $<$ eine Relation auf $\mathbb{N}$. Als Teilmenge von $\mathbb{N} \times \mathbb{N}$ ist $<$ gegeben durch

$$< \,= \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid \text{es gibt ein } p \in \mathbb{N} \text{ mit } n = p + m\}.$$

(b) Es sei eine Menge X gegeben. Dann ist $=$ eine Relation auf X. Wir nennen $=$ die *Gleichheitsrelation* (oder *Gleichheit*) auf X. Als Teilmenge von $X \times X$ ist $=$ gegeben durch

$$= \; = \{(x, x) \mid x \in X\}.$$

(A.54) Definition (Transitivität, Reflexivität, Symmetrie)**.** Es seien eine Menge X und eine Relation r auf X gegeben.

(a) Wir sagen, dass r *transitiv* ist, falls für $x, y, z \in X$ aus $x \; r \; y$ und $y \; r \; z$ stets $x \; r \; z$ folgt.

(b) Wir sagen, dass r *reflexiv* ist, falls für $x \in X$ stets $x \; r \; x$ gilt.

(c) Wir sagen, dass r *symmetrisch* ist, falls für $x, y \in X$ aus $x \; r \; y$ stets $y \; r \; x$ folgt.

(A.55) Beispiel. Die Relation $<$ auf $\mathbb{N}$ ist transitiv, aber nicht reflexiv und nicht symmetrisch.

(A.56) Definition (Äquivalenzrelation)**.** Es sei eine Menge X gegeben. Eine *Äquivalenzrelation* auf X ist eine Relation auf X, welche transitiv, reflexiv und symmetrisch ist.

(A.57) Beispiel.

(a) Für $x, y \in \mathbb{R}$ gelte genau dann $x \; c \; y$, wenn $x = y$ oder $x = -y$ ist. Dann ist c eine Äquivalenzrelation auf $\mathbb{R}$.

(b) Für $x, y \in \mathbb{Z}$ gelte genau dann $x \equiv_2 y$, wenn x und y entweder beide gerade oder beide ungerade sind. Dann ist $\equiv_2$ eine Äquivalenzrelation auf $\mathbb{Z}$.

(c) Die Relation $\{(1, 1), (2, 2), (3, 3), (4, 4), (1, 2), (2, 1), (1, 4), (4, 1), (2, 4), (4, 2)\}$ auf $\{1, 2, 3, 4\}$ ist eine Äquivalenzrelation.

(d) Für jede Menge X ist die Gleichheitsrelation $=$ auf X eine Äquivalenzrelation auf X.

(A.58) Definition (Äquivalenzklasse)**.** Es seien eine Menge X und eine Äquivalenzrelation c auf X gegeben. Für $x \in X$ heißt

$$[x] = [x]_c := \{\tilde{x} \in X \mid \tilde{x} \; c \; x\}$$

die *Äquivalenzklasse* von x in X bzgl. c, und es heißt x ein *Repräsentant* von $[x]_c$.

(A.59) Beispiel.

(a) Für $x, y \in \mathbb{R}$ gelte genau dann $x \; c \; y$, wenn $x = y$ oder $x = -y$ ist. Dann ist c eine Äquivalenzrelation auf $\mathbb{R}$. Für $x \in \mathbb{R}$ ist $[x]_c = \{x, -x\}$.

(b) Für $x, y \in \mathbb{Z}$ gelte genau dann $x \equiv_2 y$, wenn x und y entweder beide gerade oder beide ungerade sind. Dann ist $\equiv_2$ eine Äquivalenzrelation auf $\mathbb{Z}$. Es ist $[0]_{\equiv_2} = 2\mathbb{Z} = \{2q \mid q \in \mathbb{Z}\}$ und $[1]_{\equiv_2} = 2\mathbb{Z}{+}1 = \{2q{+}1 \mid q \in \mathbb{Z}\}$.

(c) Es sei $c := \{(1, 1), (2, 2), (3, 3), (4, 4), (1, 2), (2, 1), (1, 4), (4, 1), (2, 4), (4, 2)\}$. Dann ist c eine Äquivalenzrelation auf $\{1, 2, 3, 4\}$. Es ist $[1]_c = [2]_c = [4]_c = \{1, 2, 4\}$ und $[3]_c = \{3\}$.

(d) Es sei eine Menge X gegeben. Dann ist die Gleichheitsrelation $=$ eine Äquivalenzrelation auf X. Für $x \in X$ ist $[x]_= = \{x\}$.

(A.60) Definition (Quotientenmenge)**.** Es seien eine Menge X und eine Äquivalenzrelation c auf X gegeben. Die Menge

$$X/c := \{[x]_c \mid x \in X\}$$

heißt *Quotientenmenge* (oder *Quotient*) von X modulo c. Die Abbildung

$$\mathrm{quo} = \mathrm{quo}^{X/c} \colon X \to X/c, \, x \mapsto [x]_c$$

wird *Quotientenabbildung* von X/c genannt.

(A.61) Beispiel. Es sei $c := \{(1, 1), (2, 2), (3, 3), (4, 4), (1, 2), (2, 1), (1, 4), (4, 1), (2, 4), (4, 2)\}$. Dann ist

$$\{1, 2, 3, 4\}/c = \{[1]_c, [3]_c\}.$$

Verknüpfungen

(A.62) Definition (Verknüpfung)**.** Es sei eine Menge X gegeben. Eine *Verknüpfung* (oder *binäre algebraische Operation*) ist eine Abbildung $m \colon X \times X \to X$. Für $(x, y) \in X \times X$ schreiben wir $x \, m \, y := m(x, y)$.

(A.63) Beispiel.

(a) Auf $\mathbb{N}$ haben wir die Verknüpfungen $(m, n) \mapsto m + n$ und $(m, n) \mapsto m \cdot n$.

(b) Auf $\mathbb{N}_0$ haben wir die Verknüpfungen $(m, n) \mapsto m + n$ und $(m, n) \mapsto m \cdot n$.

(c) Auf $\mathbb{Z}$ haben wir die Verknüpfungen $(x, y) \mapsto x + y$ und $(x, y) \mapsto x - y$ und $(x, y) \mapsto x \cdot y$.

(d) Auf $\mathbb{Q}$ haben wir die Verknüpfungen $(x, y) \mapsto x + y$ und $(x, y) \mapsto x - y$ und $(x, y) \mapsto x \cdot y$.

(A.64) Definition (Assoziativität, Kommutativität)**.** Es seien eine Menge X und eine Verknüpfung m auf X gegeben.

(a) Wir sagen, dass m *assoziativ* ist, wenn für $x, y, z \in X$ stets

$$x \, m \, (y \, m \, z) = (x \, m \, y) \, m \, z$$

gilt.

(b) Wir sagen, dass m *kommutativ* ist, wenn für $x, y \in X$ stets

$$x \, m \, y = y \, m \, x$$

gilt.

(A.65) Beispiel. Die Verknüpfungen $(m, n) \mapsto m+n$ und $(m, n) \mapsto m \cdot n$ auf $\mathbb{N}$ sind assoziativ und kommutativ.

(A.66) Definition (neutrales Element)**.** Es seien eine Menge X und eine Verknüpfung m auf X gegeben. Ein *neutrales Element* bzgl. m ist ein Element $e \in X$, welches $e \, m \, x = x \, m \, e = x$ für alle $x \in X$ erfüllt.

(A.67) Beispiel. Es ist 1 ein neutrales Element bzgl. der Verknüpfung $(m, n) \mapsto m \cdot n$ auf $\mathbb{N}$.

(A.68) Bemerkung. Es seien eine Menge X und eine Verknüpfung m auf X gegeben. Dann gibt es höchstens ein neutrales Element bzgl. m.

Beweis. Siehe [12, Bem. (6.11)]. □

(A.69) Definition (inverses Element)**.** Es seien eine Menge X, eine Verknüpfung m auf X, ein neutrales Element e bzgl. m und ein $x \in X$ gegeben. Ein *inverses Element* zu x bzgl. m ist ein Element $y \in X$, welches $y \, m \, x = e$ und $x \, m \, y = e$ erfüllt.

(A.70) Beispiel. Es ist $\frac{4}{3}$ ein zu $\frac{3}{4}$ inverses Element bzgl. der Verknüpfung $(m, n) \mapsto m \cdot n$ auf $\mathbb{Q}$.

(A.71) Bemerkung. Es seien eine Menge X, eine assoziative Verknüpfung m auf X, ein neutrales Element e bzgl. m und ein $x \in X$ gegeben. Dann gibt es höchstens ein inverses Element zu x bzgl. m.

Beweis. Siehe [12, Kor. (6.16)]. □

Monoide und Gruppen

(A.72) Definition (Monoid, kommutatives Monoid)**.**

(a) Ein *Monoid* besteht aus einer Menge M zusammen mit einer assoziativen Verknüpfung $\cdot$ auf M, genannt *Multiplikation* (oder *Monoidverknüpfung*) von M, derart, dass M ein neutrales Element bzgl. $\cdot$ besitzt. Das neutrale Element bzgl. der Multiplikation wird auch *Eins* (oder *Einselement*) von M genannt und als 1 notiert.

(b) Ein Monoid M heißt *kommutativ*, falls die Multiplikation von M kommutativ ist.

(A.73) Definition (abelsches Monoid)**.** Ein *abelsches Monoid* ist ein kommutatives Monoid A mit Monoidverknüpfung $+$, genannt *Addition* von A. Das neutrale Element bzgl. der Addition wird auch *Null* (oder *Nullelement*) von A genannt und als 0 notiert.

Ein abelsches Monoid ist also strukturell gesehen das Gleiche wie ein kommutatives Monoid; wir verwenden lediglich in abstrakten abelschen Monoiden eine andere Standardnotation: Abstrakte Monoide (die ggf. auch mal kommutativ sein dürfen, aber im Allgemeinen nicht müssen) werden multiplikativ geschrieben, abstrakte abelsche Monoide werden additiv geschrieben.

(A.74) Konvention. Wegen der Assoziativität der Multiplikation einer Halbgruppe bzw. der Addition einer abelschen Halbgruppe kommt es bei iterierter Bildung nicht auf die Klammerung an. Im Regelfall lassen wir daher die Klammern im Folgenden weg.

(A.75) Beispiel.

(a) (i) Die Menge $\mathbb{N}$ zusammen mit der üblichen Addition ist eine abelsche Halbgruppe, aber kein abelsches Monoid.

(ii) Die Menge $\mathbb{N}$ zusammen mit der üblichen Multiplikation ist ein kommutatives Monoid. Die Eins von $\mathbb{N}$ ist die übliche Eins.

(b) (i) Die Menge $\mathbb{N}_0$ zusammen mit der üblichen Addition ist ein abelsches Monoid. Die Null von $\mathbb{N}_0$ ist die übliche Null.

(ii) Die Menge $\mathbb{N}_0$ zusammen mit der üblichen Multiplikation ist ein kommutatives Monoid. Die Eins von $\mathbb{N}_0$ ist die übliche Eins.

(A.76) Definition (Invertierbarkeit)**.**

(a) Es sei ein Monoid M gegeben. Ein Element x in M heißt *invertierbar* in M (oder eine *Einheit* von M), falls es ein inverses Element zu x bzgl. $\cdot$ gibt. Das zu einem invertierbaren Element x in M bzgl. $\cdot$ inverse Element y wird auch das *Inverse* (oder das *inverse Element*) zu x in M genannt und als $x^{-1} = (x^{-1})^M := y$ notiert.

Die *Menge der invertierbaren Elemente* in M bezeichnen wir mit

$$M^\times = \{x \in M \mid x \text{ ist invertierbar}\}.$$

(b) Es sei ein abelsches Monoid A gegeben. Ein Element x in A heißt *negierbar* in A, falls es ein inverses Element zu x bzgl. $+^A$ gibt. Das zu einem negierbaren Element x in A bzgl. $+^A$ inverse Element y wird auch das *Negative* (oder das *negative Element*) zu x in A genannt und als $-x = (-x)^A := y$ notiert.

(A.77) Beispiel.

(a) Es ist ${\mathbb{N}_0}^\times = \{1\}$, d.h. das einzige invertierbare Element in $\mathbb{N}_0$ (bzgl. der üblichen Multiplikation) ist 1.

(b) Das einzige negierbare Element in $\mathbb{N}_0$ (bzgl. der üblichen Addition) ist 0.

(A.78) Proposition. Es sei ein Monoid M gegeben.

(a) Für $x, y \in M^\times$ ist auch $xy \in M^\times$ mit $(xy)^{-1} = y^{-1}x^{-1}$.

(b) Es ist $1 \in M^\times$ mit $1^{-1} = 1$.

(c) Für $x \in M^\times$ ist auch $x^{-1} \in M^\times$ mit $(x^{-1})^{-1} = x$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(A.79) Bemerkung. Es seien ein Monoid M und $a \in M^\times$, $b, x \in M$ gegeben.

(a) Genau dann gilt $ax = b$, wenn $x = a^{-1}b$ ist.

(b) Genau dann gilt $xa = b$, wenn $x = ba^{-1}$ ist.

Beweis. Siehe [12, Bem. (6.28)]. □

(A.80) Korollar. Es seien ein Monoid M und $a \in M^\times$, $x, y \in M$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $ax = ay$.

(b) Es ist $xa = ya$.

(c) Es ist $x = y$.

Beweis. Siehe [12, Kor. (6.29)]. □

(A.81) Korollar. Es sei ein Monoid M und $a \in M^\times$, $x \in M$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $ax = a$.

(b) Es ist $xa = a$.

(c) Es ist $x = 1$.

Beweis. Siehe [12, Kor. (6.30)]. □

(A.82) Definition ((abelsche) Gruppe)**.**

(a) Eine *Gruppe* ist ein Monoid G, in welchem jedes Element von G invertierbar ist. Die Monoidverknüpfung einer Gruppe G wird auch *Gruppenverknüpfung* von G genannt.

(b) Eine *abelsche Gruppe* ist ein abelsches Monoid A, in welchem jedes Element von A negierbar ist.

(A.83) Beispiel.

(a) (i) Die Menge $\mathbb{Z}$ zusammen mit der üblichen Addition ist eine abelsche Gruppe. Die Null von $\mathbb{Z}$ ist die übliche Null. Für $x \in \mathbb{Z}$ ist das Negative zu x in $\mathbb{Z}$ das übliche Negative.

(ii) Die Menge $\mathbb{Z}$ zusammen mit der üblichen Multiplikation ist ein kommutatives Monoid, aber keine Gruppe. Die Eins von $\mathbb{Z}$ ist die übliche Eins.

(b) (i) Die Menge $\mathbb{Q}$ zusammen mit der üblichen Addition ist eine abelsche Gruppe. Die Null von $\mathbb{Q}$ ist die übliche Null. Für $x \in \mathbb{Q}$ ist das Negative zu x in $\mathbb{Q}$ das übliche Negative.

(ii) Die Menge $\mathbb{Q}$ zusammen mit der üblichen Multiplikation ist ein kommutatives Monoid, aber keine Gruppe. Die Eins von $\mathbb{Q}$ ist die übliche Eins.

(iii) Die Menge $\mathbb{Q} \setminus \{0\}$ zusammen mit der üblichen Multiplikation ist eine kommutative Gruppe. Die Eins von $\mathbb{Q} \setminus \{0\}$ ist die übliche Eins. Für $x \in \mathbb{Q} \setminus \{0\}$ ist das Inverse zu x in $\mathbb{Q} \setminus \{0\}$ das übliche Inverse.

(A.84) Konvention. Wenn wir in Zukunft von der abelschen Gruppe $\mathbb{Z}$ sprechen, so meinen wir damit stets $\mathbb{Z}$ mit der üblichen Addition. Wenn wir vom kommutativen Monoid $\mathbb{Z}$ sprechen, so meinen wir damit stets $\mathbb{Z}$ mit der üblichen Multiplikation. Ähnlich für $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Q}$, $\mathbb{R}$.

(A.85) Definition (Subtraktion)**.** Es sei eine abelsche Gruppe A gegeben. Die Verknüpfung $(x, y) \mapsto x + (-y)$ auf A wird *Subtraktion* von A genannt und als $-$ notiert.

Die Gruppe der invertierbaren Elemente

(A.86) Bemerkung. Für jedes Monoid M wird $M^\times$ eine Gruppe, wobei die Multiplikation auf $M^\times$ durch

$$x \cdot^{M^\times} y = x \cdot^M y$$

für $x, y \in M$ gegeben ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(A.87) Definition (Gruppe der invertierbaren Elemente)**.** Es sei ein Monoid M gegeben. Die Gruppe $M^\times$ mit der Multiplikation aus Bemerkung (A.86) heißt *Gruppe der invertierbaren Elemente* (oder *Einheitengruppe*) von M.

Ringe und Körper

(A.88) Definition (Ring, kommutativer Ring, Körper)**.**

(a) Ein *Ring* (genauer *unitärer Ring* oder *Ring mit Einselement* oder *Ring mit Eins*) besteht aus einer abelschen Gruppe R zusammen mit einer Verknüpfung $\cdot$ auf R so, dass die unterliegende Menge von R ein Monoid mit Multiplikation $\cdot$ wird und so, dass folgendes Axiom gilt.

- *Distributivität.* Für alle $x, y, z \in R$ ist

$$x(y+z) = xy + xz,$$
$$(x+y)z = xz + yz.$$

(b) Ein Ring R heißt *kommutativ*, falls die Multiplikation von R kommutativ ist.

(c) Ein *Körper* ist ein kommutativer Ring K, in welchem $1 \neq 0$ gilt und in welchem jedes Element von $K \setminus \{0\}$ invertierbar (bzgl. der Multiplikation $\cdot$) ist.

(A.89) Beispiel.

(a) Die Menge $\mathbb{Z}$ zusammen mit der üblichen Addition und der üblichen Multiplikation ist ein kommutativer Ring, aber kein Körper. Die Null von $\mathbb{Z}$ ist die übliche Null und die Eins von $\mathbb{Z}$ ist die übliche Eins. Für $x \in \mathbb{Z}$ ist das Negative zu x in $\mathbb{Z}$ das übliche Negative.

(b) Die Menge $\mathbb{Q}$ zusammen mit der üblichen Addition und der üblichen Multiplikation ist ein Körper. Die Null von $\mathbb{Q}$ ist die übliche Null und die Eins von $\mathbb{Q}$ ist die übliche Eins. Für $x \in \mathbb{Q}$ ist das Negative zu x in $\mathbb{Q}$ das übliche Negative und für $x \in \mathbb{Q} \setminus \{0\}$ ist das Inverse zu x in $\mathbb{Q}$ das übliche Inverse.

(c) Die Menge $\mathbb{R}$ zusammen mit der üblichen Addition und der üblichen Multiplikation ist ein Körper. Die Null von $\mathbb{R}$ ist die übliche Null und die Eins von $\mathbb{R}$ ist die übliche Eins. Für $x \in \mathbb{R}$ ist das Negative zu x in $\mathbb{R}$ das übliche Negative und für $x \in \mathbb{R} \setminus \{0\}$ ist das Inverse zu x in $\mathbb{R}$ das übliche Inverse.

(A.90) Proposition. Es sei ein Ring R gegeben.

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

(b) Für $a, b \in R$ gilt $a(-b) = (-a)b = -ab$.

(c) Für $a, b \in R$ gilt $(-a)(-b) = ab$.

Beweis. Siehe [12, Prop. (6.45)]. □

(A.91) Proposition. Es sei ein Körper K gegeben. Für $a, b \in K$ folgt aus $ab = 0$ stets $a = 0$ oder $b = 0$.

Beweis. Siehe [12, Prop. (6.48)]. □

Das Induktionsprinzip

Zum Beweis einer Aussage der Form

$$(\forall n \in \mathbb{N} : A(n)) = (\forall n : n \in \mathbb{N} \Rightarrow A(n))$$

können wir zeigen, dass zum einen die Aussage der Form

$$A(1)$$

und zum anderen für jedes $n \in \mathbb{N}$ die Aussage der Form

$$A(n) \Rightarrow A(n+1)$$

gilt.

Produkt- und Summennotation

(A.92) Notation.

(a) Es sei ein Monoid M gegeben. Für jedes $n \in \mathbb{N}_0$ und alle $x \in M^n$ mit $x_i x_j = x_j x_i$ für $i, j \in [1, n]$ notieren wir rekursiv

$$\prod_{i \in [1,n]} x_i = \begin{cases} 1, & \text{falls } n = 0, \\ (\prod_{i \in [1,n-1]} x_i)\, x_n, & \text{falls } n > 0. \end{cases}$$

(b) Es sei ein abelsches Monoid A gegeben. Für jedes $n \in \mathbb{N}_0$ und alle $x \in A^n$ notieren wir rekursiv

$$\sum_{i \in [1,n]} x_i := \begin{cases} 0, & \text{falls } n = 0, \\ \sum_{i \in [1,n-1]} x_i + x_n, & \text{falls } n > 0. \end{cases}$$

(A.93) Notation.

(a) Es seien ein Monoid M und $x \in M$ gegeben. Für $k \in \mathbb{N}_0$ setzen wir

$$x^k := \prod_{i \in [1,k]} x.$$

Wenn x invertierbar in M ist, so setzen wir

$$x^{-k} := (x^{-1})^k$$

für $k \in \mathbb{N}$.

(b) Es seien ein abelsches Monoid A und $x \in A$ gegeben. Für $k \in \mathbb{N}_0$ setzen wir

$$kx = k \cdot x := \sum_{i \in [1,k]} x.$$

Wenn x negierbar in A ist, so setzen wir

$$(-k)x := k(-x)$$

für $k \in \mathbb{N}$.

(A.94) Proposition (Potenzgesetze)**.** Es sei ein Monoid M gegeben.

(a) Für $x \in M$, $k, l \in \mathbb{N}_0$ gilt $x^k x^l = x^{k+l}$. Für $x \in M^\times$, $k, l \in \mathbb{Z}$ gilt $x^k x^l = x^{k+l}$.

(b) Für $x \in M$, $k, l \in \mathbb{N}_0$ gilt $(x^k)^l = x^{kl}$. Für $x \in M^\times$, $k, l \in \mathbb{Z}$ gilt $(x^k)^l = x^{kl}$.

(c) Es sei M kommutativ. Für $x, y \in M$, $k \in \mathbb{N}_0$ gilt $x^k y^k = (xy)^k$. Für $x, y \in M^\times$, $k \in \mathbb{Z}$ gilt $x^k y^k = (xy)^k$.

Beweis. Siehe [12, Prop. (9.14)]. □

(A.95) Notation. Es sei ein Ring R gegeben. Für $k \in \mathbb{Z}$ schreiben wir auch

$$k = k^R := k \cdot 1^R.$$

(A.96) Notation (Kronecker-Delta)**.** Es seien ein Ring R und Objekte i und j gegeben. Das *Kronecker-Delta* ist definiert als

$$\delta_{i,j} := \begin{cases} 1^R & \text{falls } i = j, \\ 0^R & \text{falls } i \neq j. \end{cases}$$

Die symmetrische Gruppe

(A.97) Bemerkung. Es sei eine Menge X gegeben. Die Menge $\mathrm{Map}(X, X)$ wird ein Monoid mit Monoidverknüpfung $(g, f) \mapsto g \circ f$. Das Einselement von $\mathrm{Map}(X, X)$ ist id_X. Ein Element $f \in \mathrm{Map}(X, X)$ ist genau dann invertierbar in $\mathrm{Map}(X, X)$, wenn f invertierbar im Sinne von Definition (A.41)(b) ist.

Beweis. Siehe [12, Bem. (6.22)]. □

(A.98) Definition (symmetrische Gruppe)**.** Es sei $n \in \mathbb{N}_0$ gegeben. Die Gruppe

$$\mathrm{S}_n := \mathrm{Map}([1, n], [1, n])^\times$$

heißt *symmetrische Gruppe* vom Grad n. Die Elemente von S_n werden *Permutationen* von $[1, n]$ genannt. Für $\pi \in \mathrm{S}_n$ schreiben wir

$$\left(\begin{smallmatrix} 1 & 2 & 3 & \dots & n \\ \pi(1) & \pi(2) & \pi(3) & \dots & \pi(n) \end{smallmatrix}\right) := \pi.$$

Nach Satz (A.49)(c) besteht S_n für $n \in \mathbb{N}_0$ also gerade aus allen bijektiven Abbildungen von $[1, n]$ nach $[1, n]$.

(A.99) Beispiel.

(a) Es ist

$$\mathrm{S}_0 = \{\mathrm{id}_\emptyset\} = \{(\ )\}.$$

(b) Es ist

$$\mathrm{S}_1 = \{\mathrm{id}_{\{1\}}\} = \{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\}.$$

(c) Es ist

$$\mathrm{S}_2 = \{\mathrm{id}_{\{1,2\}}, (1 \mapsto 2,\, 2 \mapsto 1)\} = \{\left(\begin{smallmatrix} 1 & 2 \\ 1 & 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)\}.$$

(d) Es ist

$$\mathrm{S}_3 = \{\left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right)\}.$$

(A.100) Beispiel. In S_3 ist

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

(A.101) Definition (Transposition)**.** Eine *Transposition* (oder *Vertauschung*) in S_n ist ein $\tau \in \mathrm{S}_n$ derart, dass es $i, j \in [1, n]$ mit $i \neq j$ und

$$\begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix} = \tau$$

gibt.

(A.102) Satz. Es sei $\pi \in \mathrm{S}_n$ gegeben. Dann lässt sich π entweder als Kompositum einer geraden oder einer ungeraden Anzahl von Transpositionen schreiben.

Beweis. Dies folgt aus [12, Satz (12.29)]. □

(A.103) Definition (Signum)**.** Es sei $\pi \in \mathrm{S}_n$ gegeben. Das *Signum* (oder *Vorzeichen*) von π ist gegeben durch

$$\operatorname{sgn} \pi := \begin{cases} 1, & \pi \text{ ist Kompositum einer geraden Anzahl von Transpositionen,} \\ -1, & \pi \text{ ist Kompositum einer ungeraden Anzahl von Transpositionen.} \end{cases}$$

(A.104) Proposition (Produktsatz)**.** Es sei $n \in \mathbb{N}_0$ gegeben. Für $\pi, \sigma \in \mathrm{S}_n$ gilt

$$\operatorname{sgn}(\pi \circ \sigma) = (\operatorname{sgn} \pi)(\operatorname{sgn} \sigma).$$

Beweis. Siehe [12, Prop. (12.27)]. □

(A.105) Korollar. Es sei $n \in \mathbb{N}_0$ gegeben. Für $\pi \in \mathrm{S}_n$ ist

$$\operatorname{sgn} \pi^{-1} = \operatorname{sgn} \pi.$$

Beweis. Siehe [12, Kor. (12.30)]. □

Die komplexen Zahlen

(A.106) Arbeitsbasis (Körper der komplexen Zahlen)**.**

(a) Unter einer *komplexen Zahl* verstehen wir einen „Ausdruck“ der Form

$$z = a + b\mathrm{i}$$

für gewisse $a, b \in \mathbb{R}$.

(b) Die komplexe Zahl $\mathrm{i} = 0 + 1\mathrm{i}$ wird *imaginäre Einheit* genannt.

(c) Komplexe Zahlen $z = a + b\mathrm{i}$ und $w = c + d\mathrm{i}$, wobei $a, b, c, d \in \mathbb{R}$, sind *gleich*, geschrieben $z = w$, falls $a = c$ und $b = d$ gilt.

(d) Es seien eine komplexe Zahl z und $a, b \in \mathbb{R}$ mit $z = a + b\mathrm{i}$ gegeben. Wir nennen a den *Realteil* von z und schreiben $\operatorname{Re} z := a$. Wir nennen b den *Imaginärteil* von z und schreiben $\operatorname{Im} z := b$.

(e) Der *Körper der komplexen Zahlen* hat die unterliegende Menge

$$\mathbb{C} = \{a + b\mathrm{i} \mid a, b \in \mathbb{R}\}$$

und es gilt

$$\mathrm{i}^2 = -1.$$

(f) Wir identifizieren $\mathbb{R}$ mit der Teilmenge $\{a + 0 \cdot \mathrm{i} \mid a \in \mathbb{R}\}$ von $\mathbb{C}$.

(A.107) Definition (konjugierte komplexe Zahl)**.** Es sei $z \in \mathbb{C}$ gegeben. Wir nennen

$$\overline{z} := \operatorname{Re} z - (\operatorname{Im} z)\mathrm{i}$$

die zu z *konjugierte komplexe Zahl* (oder die zu z *komplex Konjugierte*).

(A.108) Definition (Absolutbetrag)**.** Es sei $z \in \mathbb{C}$ gegeben. Wir nennen

$$|z| := \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}$$

den *Absolutbetrag* (oder *Betrag*) von z.

Endliche Primkörper

(A.109) Arbeitsbasis (endlicher Primkörper)**.** Für jede Primzahl p ist

$$\mathbb{F}_p = \{k^{\mathbb{F}_p} \mid k \in [0, p-1]\}$$

ein Körper mit p Elementen, in welchem $p = 0$ gilt. Wir nennen $\mathbb{F}_p$ den *Primkörper* zur Primzahl p.

(A.110) Definition (Standardtransversale)**.** Für jede Primzahl p heißt $[0, p-1]$ die *Standardtransversale* von $\mathbb{F}_p$.

(A.111) Beispiel. Die Standardtransversale von $\mathbb{F}_7$ ist $[0, 6] = \{0, 1, 2, 3, 4, 5, 6\}$.

Die endlichen Körper $\mathbb{F}_4$, $\mathbb{F}_8$ und $\mathbb{F}_9$

(A.112) Arbeitsbasis ($\mathbb{F}_4$, $\mathbb{F}_8$, $\mathbb{F}_9$)**.**

(a) Es ist

$$\mathbb{F}_4 = \{a + b\alpha \mid a, b \in \mathbb{F}_2\}$$

ein Körper mit 4 Elementen, in welchem $\alpha^2 = 1 + \alpha$ gilt. Wir identifizieren $\mathbb{F}_2$ mit der Teilmenge $\{a + 0\alpha \mid a \in \mathbb{F}_2\}$ von $\mathbb{F}_4$.

(b) Es ist

$$\mathbb{F}_8 = \{a + b\beta + c\beta^2 \mid a, b, c \in \mathbb{F}_2\}$$

ein Körper mit 8 Elementen, in welchem $\beta^3 = 1 + \beta$ gilt. Wir identifizieren $\mathbb{F}_2$ mit der Teilmenge $\{a + 0\beta + 0\beta^2 \mid a \in \mathbb{F}_2\}$ von $\mathbb{F}_8$.

(c) Es ist

$$\mathbb{F}_9 = \{a + b\iota \mid a, b \in \mathbb{F}_3\}$$

ein Körper mit 9 Elementen, in welchem $\iota^2 = -1$ gilt. Wir identifizieren $\mathbb{F}_3$ mit der Teilmenge $\{a + 0\iota \mid a \in \mathbb{F}_3\}$ von $\mathbb{F}_9$.

Der Polynomring

(A.113) Notation. Es seien ein Körper K und eine Menge I gegeben. Wir setzen

$$K^{(I)} := \{x \in K^I \mid \{i \in I \mid x_i \neq 0\} \text{ ist endlich}\}.$$

(A.114) Beispiel. Es seien $x, y, z \in \mathbb{Q}^{\mathbb{N}_0}$ gegeben durch

$$\begin{aligned} x &= (0, 1, 0, 0, 0, \dots) = (\delta_{i,1})_{i \in \mathbb{N}_0}, \\ y &= (2, 0, 1, 0, 0, \dots) = 2(\delta_{i,0})_{i \in \mathbb{N}_0} + (\delta_{i,2})_{i \in \mathbb{N}_0}, \\ z &= (1, 1, 1, 1, 1, \dots) = (1)_{i \in \mathbb{N}_0}. \end{aligned}$$

Dann sind $x, y \in \mathbb{Q}^{(\mathbb{N}_0)}$ und $z \notin \mathbb{Q}^{(\mathbb{N}_0)}$.

Beweis. Es ist $\{i \in I \mid x_i \neq 0\} = \{1\}$ endlich und daher $x \in \mathbb{Q}^{(\mathbb{N}_0)}$. Es ist $\{i \in I \mid y_i \neq 0\} = \{0, 2\}$ und daher $y \in \mathbb{Q}^{(\mathbb{N}_0)}$. Es ist $\{i \in I \mid z_i \neq 0\} = \mathbb{N}_0$ unendlich und daher $z \notin \mathbb{Q}^{(\mathbb{N}_0)}$. □

(A.115) Arbeitsbasis (Polynomring)**.**

(a) Ein *Polynom* in X über K ist ein „Ausdruck“ der Form

$$f = \sum_{i \in \mathbb{N}_0} a_i X^i$$

für ein $a \in K^{(\mathbb{N}_0)}$. Die Familie a wird *Koeffizientenfolge* von f genannt. Für $i \in I$ wird a_i der *Koeffizient* von f an der Stelle i genannt.

(b) Es seien Polynome f und g in X über K und $a, b \in K^{(\mathbb{N}_0)}$ mit $f = \sum_{i \in \mathbb{N}_0} a_i X^i$ und $g = \sum_{i \in \mathbb{N}_0} b_i X^i$ gegeben. Die Polynome f und g sind *gleich*, geschrieben $f = g$, falls $a = b$ gilt.

(c) Das Polynom $X = \sum_{i \in \mathbb{N}_0} \delta_{1,i} X^i$ wird *Unbestimmte* genannt.

(d) Der kommutative Ring gegeben durch die Menge der Polynome

$$K[X] := \{\sum_{i \in \mathbb{N}_0} a_i X^i \mid a \in K^{(\mathbb{N}_0)}\}$$

in X über K mit Addition und Multiplikation gegeben durch

$$\begin{aligned} \sum_{i \in \mathbb{N}_0} a_i X^i + \sum_{i \in \mathbb{N}_0} b_i X^i &= \sum_{i \in \mathbb{N}_0} (a_i + b_i) X^i, \\ (\sum_{i \in \mathbb{N}_0} a_i X^i)(\sum_{j \in \mathbb{N}_0} b_j X^j) &= \sum_{k \in \mathbb{N}_0} (\sum_{\substack{i,j \in \mathbb{N}_0 \\ i+j=k}} a_i b_j) X^k \end{aligned}$$

für $a, b \in K^{(\mathbb{N}_0)}$ wird *Polynomring* in X über K genannt.

(e) Wir identifizieren K mit der Teilmenge $\{aX^0 \mid a \in K\}$ von $K[X]$. Das heißt, unter Missbrauch der Notation notieren wir aX^0 für $a \in K$ auch durch a.

(A.116) Beispiel. Es seien $f, g \in \mathbb{Q}[X]$ gegeben durch $f = X^2 - 1$ und $g = -X^3 + 2X^2 + X + 1$. Dann ist

$$\begin{aligned} f + g &= -X^3 + 3X^2 + X, \\ fg &= -X^5 + 2X^4 + 2X^3 - X^2 - X - 1, \\ -2f &= -2X^2 + 2. \end{aligned}$$

(A.117) Definition (Polynomfunktion)**.** Es seien ein Körper K, $f \in K[X]$ und $a \in K^{(\mathbb{N}_0)}$ mit $f = \sum_{i \in \mathbb{N}_0} a_i X^i$ gegeben. Die Abbildung

$$K \to K,\ x \mapsto \sum_{i \in \mathbb{N}_0} a_i x^i$$

heißt *Polynomfunktion* zu f und wird unter Missbrauch der Notation wieder als f bezeichnet.

(A.118) Beispiel. Es sei $f \in \mathbb{Q}[X]$ gegeben durch $f = X^2 - 1$. Dann ist $f(5) = 24$.

Beweis. Es gilt $f(5) = 5^2 - 1 = 24$. □

Polynome *sind* keine Funktionen, sie *liefern* lediglich zugehörige Polynomfunktionen. So gibt es etwa über dem Körper $\mathbb{F}_2$ unendlich viele Polynome, aber lediglich vier Polynomfunktionen (alle vier Funktionen von $\mathbb{F}_2$ nach $\mathbb{F}_2$ sind Polynomfunktionen).

(A.119) Definition (Grad, normiertes Polynom)**.** Es seien $f \in K[X] \setminus \{0\}$ und $a \in K^{(\mathbb{N}_0)}$ mit $f = \sum_{i \in \mathbb{N}_0} a_i X^i$ gegeben.

(a) Der *Grad* von f ist definiert als

$$\deg f := \max \{i \in \mathbb{N}_0 \mid a_i \neq 0\}.$$

(b) Wir sagen, dass f ein *normiertes* Polynom ist, falls $a_{\deg f} = 1$ ist.

(A.120) Beispiel.

(a) Das Polynom $X^2 - 1$ über $\mathbb{Q}$ hat den Grad 2 und ist normiert.

(b) Das Polynom $-X^3 + 2X^2 + X + 1$ über $\mathbb{Q}$ hat den Grad 3 und ist nicht normiert.

(A.121) Proposition. Es sei ein Körper K gegeben. Dann ist

$$K[X]^\times = K^\times.$$

Beweis. Siehe [12, Kor. (11.9)]. □

(A.122) Notation. Es sei ein Körper K gegeben. Für $n \in \mathbb{N}_0$ setzen wir

$$K[X]_{<n} := \{f \in K[X] \mid f = 0 \text{ oder } \deg f < n\}.$$

(A.123) Bemerkung. Für $n \in \mathbb{N}_0$ ist

$$K[X]_{<n} = \{\sum_{i \in [0,n-1]} a_i X^i \mid a \in K^{[0,n-1]}\}.$$

Insbesondere gilt $K[X]_{<0} = \{0\}$ und $K[X]_{<1} = K$.

(A.124) Definition (Nullstelle)**.** Es seien ein Körper K und $f \in K[X]$ gegeben. Eine *Nullstelle* von f ist ein $a \in K$ mit

$$f(a) = 0.$$

(A.125) Beispiel. Es sei $f \in \mathbb{Q}[X]$ gegeben durch $f = X^2 - 1$. Dann sind 1 und -1 Nullstellen von f.

Beweis. Es gilt $f(1) = 1^2 - 1 = 0$ und $f(-1) = (-1)^2 - 1 = 0$. Folglich sind 1 und -1 Nullstellen von f. □

(A.126) Definition (Teilbarkeit)**.** Es seien ein Körper K und $f, g \in K[X]$ gegeben. Wir sagen f *teilt* g (oder dass f ein *Teiler* von g ist oder dass f ein *Faktor* von g ist oder dass g ein *Vielfaches* von f ist), geschrieben $f \mid g$, falls es ein $q \in K[X]$ mit $g = qf$ gibt. Wenn f kein Teiler von g ist, so schreiben wir $f \nmid g$.

(A.127) Proposition. Es seien ein Körper K, $f \in K[X]$ und $a \in K$ gegeben. Genau dann ist a eine Nullstelle von f, wenn $X - a$ ein Teiler von f ist.

Beweis. Siehe [12, Prop. (12.15)]. □

(A.128) Proposition. Es seien ein Körper K und $f \in K[X] \setminus \{0\}$ gegeben. Für jedes $a \in K$ gibt es eindeutige $m \in \mathbb{N}_0$ und $g \in K[X]$ mit

$$f = (X - a)^m g$$

und $g(a) \neq 0$.

Beweis. Dies folgt aus [12, Satz (12.36)]. □

(A.129) Definition (Vielfachheit)**.** Es seien ein Körper K, $f \in K[X] \setminus \{0\}$ und $a \in K$ gegeben. Das $m \in \mathbb{N}_0$, für welches es ein $g \in K[X]$ mit $f = (X - a)^m g$ und $g(a) \neq 0$ gibt, wird *Vielfachheit* (oder *Multiplizität*) von a als Nullstelle von f genannt und als $\mathrm{m}_a(f) := m$ notiert.

(A.130) Definition (Zerfällung in Linearfaktoren)**.** Es seien ein Körper K und ein $f \in K[X] \setminus \{0\}$ gegeben.

(a) Wir sagen, dass f in *Linearfaktoren zerfällt*, falls es $b \in K^\times$ und $a_i \in K$ für $i \in [1, \deg f]$ mit

$$f = b \prod_{i \in [1, \deg f]} (X - a_i)$$

gibt.

(b) Wir sagen, dass f in *verschiedene Linearfaktoren zerfällt*, falls es $b \in K^\times$ und verschiedene $a_i \in K$ für $i \in [1, \deg f]$ mit

$$f = b \prod_{i \in [1, \deg f]} (X - a_i)$$

gibt.

(A.131) Beispiel.

(a) Es sei ein Körper K gegeben. Das Polynom $X^2 - 1$ über K zerfällt in Linearfaktoren.

(b) Das Polynom $X^2 + 1$ über $\mathbb{Q}$ zerfällt nicht in Linearfaktoren.

(A.132) Bemerkung. Es seien ein Körper K und ein $f \in K[X] \setminus \{0\}$ gegeben. Dann ist

$$\sum_{a \in K} \mathrm{m}_a(f) \leq \deg f.$$

Genau dann gilt

$$\sum_{a \in K} \mathrm{m}_a(f) = \deg f,$$

wenn f in Linearfaktoren zerfällt.

Beweis. Siehe [12, Bem. (12.47)]. □

Matrizen

Im Folgenden bezeichnen wir bis zum Ende des Abschnitts mit R stets einen beliebig gegebenen kommutativen, unitären Ring.

(A.133) Definition (Matrix)**.** Es seien $m, n \in \mathbb{N}_0$ gegeben. Die *Menge der* $(m \times n)$*-Matrizen* in R ist definiert als

$$R^{m \times n} := R^{[1,m] \times [1,n]}.$$

Ein Element von $R^{m \times n}$ wird $(m \times n)$*-Matrix* in R (oder $(m \times n)$*-Matrix* über R) genannt.
Für eine $(m \times n)$-Matrix A in R schreiben wir auch

$$\begin{pmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \dots & A_{m,n} \end{pmatrix} := (A_{i,j})_{i \in [1,m], j \in [1,n]} := A.$$

Eine präzise Definition lässt sich mit Hilfe des Begriffs einer *Familie* geben, siehe oben.

(A.134) Notation. Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Für $x \in R^{n \times 1}$ schreiben wir $x_i := x_{i,1}$ für den Eintrag an der Stelle $(i, 1)$ für ein $i \in [1, n]$ sowie

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_i)_{i \in [1,n]}.$$

(b) Für $x \in R^{1 \times n}$ schreiben wir $x_i := x_{1,i}$ für den Eintrag an der Stelle $(1, i)$ für ein $i \in [1, n]$ sowie

$$x = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} = (x_i)_{i \in [1,n]}.$$

(A.135) Definition (Zeile, Spalte)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in R^{m \times n}$ gegeben.

(a) Für $i \in [1, m]$ heißt $A_{i,-} \in R^{1 \times n}$ gegeben durch

$$A_{i,-} = \begin{pmatrix} A_{i,1} & \dots & A_{i,n} \end{pmatrix} = (A_{i,j})_{j \in [1,n]}$$

die i*-te Zeile* von A.

(b) Für $j \in [1, n]$ heißt $A_{-,j} \in R^{m \times 1}$ gegeben durch

$$A_{-,j} = \begin{pmatrix} A_{1,j} \\ \vdots \\ A_{m,j} \end{pmatrix} = (A_{i,j})_{i \in [1,m]}$$

die j*-te Spalte* von A.

(A.136) Beispiel. Es sei

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

(a) Es ist

$$A_{1,-} = \begin{pmatrix} 1 & 0 & -2 \end{pmatrix}, \; A_{2,-} = \begin{pmatrix} 2 & -1 & 3 \end{pmatrix}.$$

(b) Es ist

$$A_{-,1} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \; A_{-,2} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \; A_{-,3} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}.$$

(A.137) Definition (Matrixaddition)**.** Es seien $m, n \in \mathbb{N}_0$ gegeben.

(a) Für $A, B \in R^{m \times n}$ heißt

$$A + B := \begin{pmatrix} A_{1,1} + B_{1,1} & \dots & A_{1,n} + B_{1,n} \\ \vdots & & \vdots \\ A_{m,1} + B_{m,1} & \dots & A_{m,n} + B_{m,n} \end{pmatrix} = (A_{i,j} + B_{i,j})_{i \in [1,m], j \in [1,n]}$$

die *Summe* (oder *Matrixsumme*) von A und B.

(b) Die Matrix $0 \in R^{m \times n}$ gegeben durch

$$0 = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} = (0)_{i \in [1,m], j \in [1,n]}$$

heißt *Nullmatrix* (genauer $(m \times n)$-*Nullmatrix* über R).

(c) Für $A \in R^{m \times n}$ heißt

$$-A := \begin{pmatrix} -A_{1,1} & \dots & -A_{1,n} \\ \vdots & & \vdots \\ -A_{m,1} & \dots & -A_{m,n} \end{pmatrix} = (-A_{i,j})_{i \in [1,m], j \in [1,n]}$$

das *Negative* (oder die *negative Matrix*) von A.

(A.138) Beispiel.

(a) Es seien $A, B \in \mathbb{Z}^{2 \times 3}$ gegeben durch

$$A = \begin{pmatrix} 5 & 0 & -3 \\ 1 & -1 & 0 \end{pmatrix}, B = \begin{pmatrix} -2 & 2 & 4 \\ 1 & 1 & 1 \end{pmatrix}.$$

Dann ist

$$A + B = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 0 & 1 \end{pmatrix}.$$

(b) Es ist $0 \in \mathbb{Z}^{2 \times 3}$ gegeben durch

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

(c) Es sei $A \in \mathbb{Z}^{2 \times 3}$ gegeben durch

$$A = \begin{pmatrix} 5 & 0 & -3 \\ 1 & -1 & 0 \end{pmatrix}.$$

Dann ist

$$-A = \begin{pmatrix} -5 & 0 & 3 \\ -1 & 1 & 0 \end{pmatrix}.$$

(A.139) Definition (Skalarmultiplikation von Matrizen)**.** Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $c \in R$, $A \in R^{m \times n}$ heißt

$$cA = c \cdot A := \begin{pmatrix} cA_{1,1} & \dots & cA_{1,n} \\ \vdots & & \vdots \\ cA_{m,1} & \dots & cA_{m,n} \end{pmatrix} = (cA_{i,j})_{i \in [1,m], j \in [1,n]}$$

das *c-fache* von A.

(A.140) Beispiel. Es sei $A \in \mathbb{Z}^{2\times 3}$ gegeben durch

$$A = \begin{pmatrix} 5 & 0 & -3 \\ 1 & -1 & 0 \end{pmatrix}.$$

Dann ist

$$(-2)A = \begin{pmatrix} -10 & 0 & 6 \\ -2 & 2 & 0 \end{pmatrix}.$$

(A.141) Definition (Matrixmultiplikation)**.**

(a) Es seien $m, n, p \in \mathbb{N}_0$ gegeben. Für $A \in R^{m\times n}$, $B \in R^{n\times p}$ heißt $AB \in R^{m\times p}$ gegeben durch

$$AB = A \cdot B = (\sum_{j\in[1,n]} A_{i,j} B_{j,k})_{i\in[1,m],k\in[1,p]}$$

das *Matrixprodukt* von A und B.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Die Matrix $\mathrm{E}_n \in R^{n\times n}$ gegeben durch

$$\mathrm{E}_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

heißt *Einheitsmatrix* (genauer n*-te Einheitsmatrix* oder n*-te Identitätsmatrix*) über R.

(A.142) Beispiel.

(a) Es seien $A \in R^{3\times 4}$, $B \in R^{4\times 2}$ gegeben durch

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

Dann ist

$$AB = \begin{pmatrix} 20 & 10 \\ 3 & -3 \\ 3 & 7 \end{pmatrix}.$$

(b) Es ist $\mathrm{E}_3 \in R^{3\times 3}$ gegeben durch

$$\mathrm{E}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(A.143) Proposition.

(a) Es seien $m, n \in \mathbb{N}_0$ gegeben. Die Menge $R^{m\times n}$ zusammen mit der Matrixaddition ist eine abelsche Gruppe. Die Null von $R^{m\times n}$ ist die Nullmatrix. Für $A \in R^{m\times n}$ ist das Negative zu A in $R^{m\times n}$ die negative Matrix.

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Die Menge $R^{n\times n}$ zusammen mit der Matrixaddition und der Matrixmultiplikation ist ein Ring. Die Null von $R^{n\times n}$ ist die Nullmatrix. Die Eins von $R^{n\times n}$ ist die Einheitsmatrix. Für $A \in R^{n\times n}$ ist das Negative zu A in $R^{n\times n}$ die negative Matrix.

Beweis.

(a) Siehe [12, Prop. (15.1)].

(b) Siehe [12, Kor. (15.13)]. □

Wie in jedem Ring gibt es auch in $R^{n \times n}$ für $n \in \mathbb{N}_0$ den Begriff des invertierbaren Elements: Ein $A \in R^{n \times n}$ ist invertierbar, falls ein $B \in R^{n \times n}$ mit $AB = BA = \mathrm{E}_n$ existiert. In diesem Fall ist das Inverse zu A eindeutig bestimmt und wird mit $A^{-1} = B$ bezeichnet.

(A.144) Definition (allgemeine lineare Matrixgruppe)**.** Es sei $n \in \mathbb{N}_0$ gegeben. Die Gruppe

$$\mathrm{GL}_n(R) := (R^{n \times n})^\times$$

heißt *allgemeine lineare Gruppe* (oder *volle lineare Gruppe*) vom Grad n über R.

(A.145) Beispiel. Es sei $A \in \mathbb{Z}^{2 \times 2}$ gegeben durch

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

Dann ist A invertierbar mit

$$A^{-1} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}.$$

(A.146) Definition (Transponierte)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in R^{m \times n}$ gegeben. Die Matrix $A^{\mathrm{tr}} \in R^{n \times m}$ gegeben durch

$$A^{\mathrm{tr}} = \begin{pmatrix} A_{1,1} & \dots & A_{m,1} \\ \vdots & & \vdots \\ A_{1,n} & \dots & A_{m,n} \end{pmatrix} = (A_{i,j})_{j \in [1,n], i \in [1,m]}$$

heißt *Transponierte* von A (oder die *zu A transponierte Matrix*).

(A.147) Beispiel. Es sei $A \in \mathbb{Z}^{2 \times 3}$ gegeben durch

$$A = \begin{pmatrix} 5 & 0 & -3 \\ 1 & -1 & 0 \end{pmatrix}.$$

Dann ist

$$A^{\mathrm{tr}} = \begin{pmatrix} 5 & 1 \\ 0 & -1 \\ -3 & 0 \end{pmatrix}.$$

(A.148) Proposition.

(a) Es seien $m, n, p \in \mathbb{N}_0$ gegeben. Für $A \in R^{m \times n}$, $B \in R^{n \times p}$ gilt

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

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Dann gilt

$$\mathrm{E}_n^{\mathrm{tr}} = \mathrm{E}_n.$$

(c) Es sei $n \in \mathbb{N}_0$ gegeben. Für eine invertierbare Matrix $A \in R^{n \times n}$ ist auch A^{tr} invertierbar mit

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

(d) Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $A, B \in R^{m \times n}$ gilt

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

(e) Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $c \in R$, $A \in R^{m \times n}$ gilt

$$(cA)^{\mathrm{tr}} = cA^{\mathrm{tr}}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(A.149) Definition (Adjungierte)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in \mathbb{C}^{m \times n}$ gegeben. Die Matrix $A^{\mathrm{ad}} \in \mathbb{C}^{n \times m}$ gegeben durch

$$A^{\mathrm{ad}} = (\overline{A_{i,j}})_{j \in [1,n], i \in [1,m]}$$

wird *Adjungierte* von A (oder die *zu A adjungierte Matrix*) genannt.

(A.150) Beispiel. Über $\mathbb{C}$ ist

$$\begin{pmatrix} 1+\mathrm{i} & -3\mathrm{i} & 2-2\mathrm{i} \\ 0 & -1+2\mathrm{i} & \mathrm{i} \end{pmatrix}^{\mathrm{ad}} = \begin{pmatrix} 1-\mathrm{i} & 0 \\ 3\mathrm{i} & -1-2\mathrm{i} \\ 2+2\mathrm{i} & -\mathrm{i} \end{pmatrix}.$$

(A.151) Proposition.

(a) Es seien $m, n, p \in \mathbb{N}_0$ gegeben. Für $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times p}$ gilt

$$(AB)^{\mathrm{ad}} = B^{\mathrm{ad}} A^{\mathrm{ad}}.$$

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Dann gilt

$$\mathrm{E}_n^{\mathrm{ad}} = \mathrm{E}_n.$$

(c) Es sei $n \in \mathbb{N}_0$ gegeben. Für $A \in \mathrm{GL}_n(\mathbb{C})$ ist auch $A^{\mathrm{ad}} \in \mathrm{GL}_n(\mathbb{C})$ mit

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

(d) Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $A, B \in \mathbb{C}^{m \times n}$ gilt

$$(A+B)^{\mathrm{ad}} = A^{\mathrm{ad}} + B^{\mathrm{ad}}.$$

(e) Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $c \in \mathbb{C}$, $A \in \mathbb{C}^{m \times n}$ gilt

$$(cA)^{\mathrm{ad}} = \bar{c} A^{\mathrm{ad}}.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(A.152) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ gegeben. Für $A \in \mathbb{C}^{m \times n}$ gilt

$$(A^{\mathrm{ad}})^{\mathrm{ad}} = A.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

(A.153) Definition (Dreiecksmatrix)**.** Es sei $n \in \mathbb{N}_0$ gegeben.

(a) Eine *obere $(n \times n)$-Dreiecksmatrix* in R ist ein $A \in R^{n \times n}$ mit

$$A_{i,j} = 0$$

für $i, j \in [1, n]$ mit $i > j$.

(b) Eine *untere $(n \times n)$-Dreiecksmatrix* in R ist ein $A \in R^{n \times n}$ mit

$$A_{i,j} = 0$$

für $i, j \in [1, n]$ mit $i < j$.

(A.154) Beispiel. (a) Es ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

eine obere Dreiecksmatrix in $\mathbb{Z}$.

(b) Es ist

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

eine untere Dreiecksmatrix in $\mathbb{Z}$.

Lineare Gleichungssysteme

Im Folgenden, bis zum Ende des Abschnitts und mit Ausnahme einiger Beispiele, sei stets ein Körper K gegeben. Es seien $m, n \in \mathbb{N}_0$ gegeben. Ein *lineares Gleichungssystem* aus m Gleichungen und n Unbekannten x_j für $j \in [1, n]$ über K ist durch

$$\begin{array}{ccccccccc} A_{1,1}x_1 & + & A_{1,2}x_2 & + & \dots & + & A_{1,n}x_n & = & b_1, \\ A_{2,1}x_1 & + & A_{2,2}x_2 & + & \dots & + & A_{2,n}x_n & = & b_2, \\ & & & & & & & \vdots & \\ A_{m,1}x_1 & + & A_{m,2}x_2 & + & \dots & + & A_{m,n}x_n & = & b_m \end{array}$$

für $A_{i,j}, b_i \in K$, wobei $i \in [1, m]$, $j \in [1, n]$, gegeben. Kurz können wir hierfür auch

$$\sum_{j \in [1,n]} A_{i,j} x_j = b_i$$

für $i \in [1, m]$ schreiben. Wir nennen das lineare Gleichungssystem *homogen*, falls $b_i = 0$ für $i \in [1, m]$, und ansonsten *inhomogen*.
Da ein lineares Gleichungssystem aus m Gleichungen und n Unbekannten wie oben nur von den gegebenen Koeffizienten $A_{i,j}$ und b_i für $i \in [1, m]$, $j \in [1, n]$ abhängt, werden wir lineare Gleichungssysteme im Folgenden durch Matrizen kodieren. Das lineare Gleichungssystem lässt sich dann umschreiben als eine einzige Gleichung:

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ \vdots & \vdots & & \vdots \\ A_{m,1} & A_{m,2} & \dots & A_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Die gegebenen Daten des linearen Gleichungssystems lassen sich außerdem in der Matrix

$$\left(\begin{array}{cccc|c} A_{1,1} & A_{1,2} & \dots & A_{1,n} & b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ A_{m,1} & A_{m,2} & \dots & A_{m,n} & b_m \end{array}\right)$$

zusammenfassen.

(A.155) Definition (Lösung eines linearen Gleichungssystems)**.** Es seien $m, n \in \mathbb{N}_0$ gegeben.

(a) Es seien $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ gegeben. Die *Lösungsmenge des linearen Gleichungssystems* zur erweiterten Koeffizientenmatrix $\left(A \mid b\right)$ ist definiert als

$$\operatorname{Sol}(A, b) := \{x \in K^{n \times 1} \mid Ax = b\}.$$

Ein Element von $\operatorname{Sol}(A, b)$ wird *Lösung des linearen Gleichungssystems* zur erweiterten Koeffizientenmatrix $\left(A \mid b\right)$ genannt.

(b) Es sei $A \in K^{m \times n}$ gegeben. Die Lösungsmenge $\operatorname{Sol}(A, 0)$ des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\left(A \mid 0\right)$ wird auch *Lösungsmenge des homogenen linearen Gleichungssystems* (oder *Lösungsraum des homogenen linearen Gleichungssystems*) zur Koeffizientenmatrix A genannt. Ein Element von $\operatorname{Sol}(A, 0)$ wird auch *Lösung des homogenen linearen Gleichungssystems* zur Koeffizientenmatrix A genannt.

(A.156) Proposition. Es seien ein kommutativer Ring R, $m, n \in \mathbb{N}_0$, $A \in R^{m \times n}$, $b \in R^{m \times 1}$ und $x \in \operatorname{Sol}(A, b)$ gegeben. Dann ist

$$\operatorname{Sol}(A, 0) \to \operatorname{Sol}(A, b),\ x_0 \mapsto x + x_0$$

eine wohldefinierte Bijektion. Insbesondere ist

$$\operatorname{Sol}(A, b) = \{x + x_0 \mid x_0 \in \operatorname{Sol}(A, 0)\} = x + \operatorname{Sol}(A, 0).$$

Beweis. Siehe [12, Prop. (16.4)]. $\square$

(A.157) Vorstellung ((reduzierte) Zeilenstufenform)**.** Es seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$ gegeben.

(a) Wir sagen, dass A in *Zeilenstufenform* ist, falls

$$A = \left(\begin{array}{ccc|cccccccccccc}
0 & \dots & 0 & \blacksquare & * & \dots & * & * & * & \dots & * & * & * & \dots & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \blacksquare & * & \dots & * & * & * & \dots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \blacksquare & * & \dots & * \\
\hline
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0
\end{array}\right)$$

ist, wobei die mit $*$ markierten Einträge beliebig und die mit $\blacksquare$ markierten Einträge ungleich 0 sind.

(b) Wir sagen, dass A in *reduzierter Zeilenstufenform* ist, falls

$$A = \left(\begin{array}{ccc|cccccccccccc}
0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 1 & * & \dots & * \\
\hline
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0
\end{array}\right)$$

ist, wobei die mit $*$ markierten Einträge beliebig sind.

(A.158) Beispiel.

(a) Über $\mathbb{R}$ ist die Matrix

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

nicht in Zeilenstufenform.

(b) Über $\mathbb{R}$ ist die Matrix

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

in Zeilenstufenform, aber nicht in reduzierter Zeilenstufenform.

(c) Über $\mathbb{R}$ ist die Matrix

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

in reduzierter Zeilenstufenform.

(A.159) Proposition (Lösbarkeitskriterium für lineare Gleichungssysteme in Zeilenstufenform)**.** Es seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ in Zeilenstufenform und ein $b \in K^{m \times 1}$ gegeben. Ferner sei r die Zeilenstufenanzahl von A. Genau dann gibt es eine Lösung des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$, wenn die Zeilenstufenanzahl von $\begin{pmatrix} A \mid b \end{pmatrix}$ auch gleich r ist.

Beweis. Siehe [12, Prop. (16.10)(a)]. $\square$

Der konstruktive Beweis von Proposition (A.159) liefert folgenden Algorithmus zur Bestimmung der Lösungsmenge eines linearen Gleichungssystems zu einer erweiterten Koeffizientenmatrix in Zeilenstufenform.

$$\left(\begin{array}{cccccccccccccc|c}
0 & \dots & 0 & \blacksquare & * & \dots & * & * & * & \dots & * & * & * & \dots & * & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \blacksquare & * & \dots & * & * & * & \dots & * & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \blacksquare & * & \dots & * & * \\
\hline
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0
\end{array}\right)$$

Die r Unbekannten an den $\blacksquare$-Spalten nennen wir *abhängige Variablen*, die anderen $n - r$ Unbekannten nennen wir *freie Variablen*. Zunächst bringen wir die freien Variablen auf die rechte Seite des linearen Gleichungssystems und ersetzen sie der Reihe nach durch $a_1, \dots, a_{n-r} \in K$. Danach lösen wir, von unten nach oben, nach den abhängigen Variablen auf; nach jedem dieser Schritte steht in der jeweils nächsten zu lösenden Gleichung genau eine abhängige Variable.

(A.160) Beispiel. Es seien $A \in \mathbb{R}^{3\times 4}$ und $b \in \mathbb{R}^{3\times 1}$ gegeben durch

$$A := \begin{pmatrix} 2 & 2 & -2 & -6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ b := \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}.$$

Die Lösungsmenge des linearen Gleichungssystems zur erweiterten Koeffizientenmatrix $\begin{pmatrix} A \mid b \end{pmatrix}$ ist gegeben durch

$$\mathrm{Sol}(A, b) = \begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

Beweis. Es ist

$$\begin{pmatrix} A \mid b \end{pmatrix} = \left(\begin{array}{cccc|c} 2 & 2 & -2 & -6 & 4 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Für $x \in \mathbb{R}^{4\times 1}$ gilt genau dann $x \in \mathrm{Sol}(A, b)$, wenn es $a_1, a_2 \in \mathbb{R}$ gibt mit

$$\begin{aligned}
x_2 &= a_1, \\
x_4 &= a_2, \\
x_3 &= 1 - 2x_4 = 1 - 2a_2, \\
x_1 &= 2^{-1}(4 - 2x_2 + 2x_3 + 6x_4) = 2 - x_2 + x_3 + 3x_4 = 2 - a_1 + (1 - 2a_2) + 3a_2 = 3 - a_1 + a_2.
\end{aligned}$$

Folglich ist

$$\begin{aligned}
\mathrm{Sol}(A, b) &= \{ \begin{pmatrix} 3 - a_1 + a_2 \\ a_1 \\ 1 - 2a_2 \\ a_2 \end{pmatrix} \mid a_1, a_2 \in \mathbb{R} \} = \{ \begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + a_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix} \mid a_1, a_2 \in \mathbb{R} \} \\
&= \begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix}.
\end{aligned}$$

□

(A.161) Vorstellung (elementare Zeilenoperationen)**.** Es seien $m, n \in \mathbb{N}_0$ und $A, A' \in K^{m\times n}$ gegeben. Wir sagen, dass A' durch eine *elementare Zeilenoperation* aus A entsteht, falls einer der folgenden Fälle zutrifft.

- Es entsteht A' durch Vertauschen der k-ten und l-ten Zeile aus A für gewisse $k, l \in [1, m]$. In diesem Fall schreiben wir

  $$A \xmapsto{\mathrm{sw}_{k,l}} A'.$$

- Es entsteht A' durch Addition des c-fachen der l-ten zur k-ten Zeile aus A für gewisse $k, l \in [1, m]$ mit $k \neq l$ und $c \in K$. In diesem Fall schreiben wir

  $$A \xmapsto{\mathrm{add}_{k,l,c}} A'.$$

- Es entsteht A' durch Multiplikation der k-ten Zeile um das c-fache aus A für gewisse $k \in [1, m]$, $c \in K^\times$. In diesem Fall schreiben wir

  $$A \xmapsto{\mathrm{mul}_{k,c}} A'.$$

(A.162) Beispiel. Über $\mathbb{Q}$ gilt

$$\begin{pmatrix} 0 & 2 & 1 & -3 \\ 2 & 0 & -2 & 6 \\ 0 & 4 & 2 & -7 \end{pmatrix} \xmapsto{\mathrm{sw}_{1,2}} \begin{pmatrix} 2 & 0 & -2 & 6 \\ 0 & 2 & 1 & -3 \\ 0 & 4 & 2 & -7 \end{pmatrix} \xmapsto{\mathrm{add}_{3,2,-2}} \begin{pmatrix} 2 & 0 & -2 & 6 \\ 0 & 2 & 1 & -3 \\ 0 & 0 & 0 & -1 \end{pmatrix} \xmapsto{\mathrm{mul}_{1,\frac{1}{2}}} \begin{pmatrix} 1 & 0 & -1 & 3 \\ 0 & 2 & 1 & -3 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

(A.163) Proposition. Es seien $m, n \in \mathbb{N}_0$, $A, A' \in K^{m \times n}$, $b, b' \in K^{m \times 1}$ so gegeben, dass $\big(A' \mid b'\big)$ durch eine endliche Folge elementarer Zeilenoperationen aus $\big(A \mid b\big)$ entsteht. Dann ist

$$\mathrm{Sol}(A, b) = \mathrm{Sol}(A', b').$$

Beweis. Siehe [12, Prop. (16.17)]. □

(A.164) Satz. Es seien $m, n \in \mathbb{N}_0$ gegeben.

(a) Jedes $A \in K^{m \times n}$ lässt sich durch eine Folge von Vertauschungs- und Additionsoperationen in eine Matrix in Zeilenstufenform überführen.

(b) Jedes $A \in K^{m \times n}$ in Zeilenstufenform lässt sich durch eine Folge von Multiplikations- und Additionsoperationen in eine Matrix in reduzierter Zeilenstufenform überführen.

(c) Jedes $A \in K^{m \times n}$ lässt sich durch eine Folge von elementaren Zeilenoperationen in eine Matrix in reduzierter Zeilenstufenform überführen.

Beweisskizze. Da eine Nullmatrix bereits in reduzierter Zeilenstufenform ist, dürfen wir ohne Beschränkung der Allgemeinheit annehmen, dass $A \neq 0$ ist.

(a) Wir betrachten die vorderste Spalte von A, in welcher ein von 0 verschiedener Eintrag steht, und vertauschen diese Zeile mit der ersten Zeile von A. Danach räumen wir die Einträge in dieser Spalte unterhalb der ersten Zeile mittels geeigneten Additionsoperationen aus.

$$\left(\begin{array}{cc|c|cc} & \vdots & \vdots & \vdots & \\ \cdots & 0 & * & * & \cdots \\ \hline \cdots & 0 & \blacksquare & * & \cdots \\ \hline \cdots & 0 & * & * & \cdots \\ & \vdots & \vdots & \vdots & \end{array}\right) \mapsto \left(\begin{array}{cc|c|cc} \cdots & 0 & \blacksquare & * & \cdots \\ \hline \cdots & 0 & * & * & \cdots \\ \cdots & 0 & * & * & \cdots \\ & \vdots & \vdots & \vdots & \\ \cdots & 0 & * & * & \cdots \end{array}\right) \mapsto \cdots \mapsto \left(\begin{array}{cc|c|cc} \cdots & 0 & \blacksquare & * & \cdots \\ \hline \cdots & 0 & 0 & * & \cdots \\ \cdots & 0 & 0 & * & \cdots \\ & \vdots & \vdots & \vdots & \\ \cdots & 0 & 0 & * & \cdots \end{array}\right)$$

Nun iterieren wir das Verfahren mit der Matrix unterhalb der ersten Zeile.

(b) Wir betrachten die letzte Zeile von A, in welcher ein von 0 verschiedener Eintrag steht. Durch Multiplikation dieser Zeile mit einem geeigneten Element aus K erhalten wir eine Matrix, deren vorderster von 0 verschiedener Eintrag in dieser Zeile eine 1 ist. Danach räumen wir die Einträge oberhalb dieser 1 mittels geeigneten Additionsoperationen aus.

$$\begin{pmatrix} \begin{array}{cc|c|cc} & \vdots & \vdots & \vdots & \\ \cdots & * & * & * & \cdots \\ \hline \cdots & 0 & \blacksquare & * & \cdots \\ \hline \cdots & 0 & 0 & 0 & \cdots \\ & \vdots & \vdots & \vdots & \end{array} \end{pmatrix} \mapsto \begin{pmatrix} \begin{array}{cc|c|cc} & \vdots & \vdots & \vdots & \\ \cdots & * & * & * & \cdots \\ \hline \cdots & 0 & 1 & * & \cdots \\ \hline \cdots & 0 & 0 & 0 & \cdots \\ & \vdots & \vdots & \vdots & \end{array} \end{pmatrix} \mapsto \cdots \mapsto \begin{pmatrix} \begin{array}{cc|c|cc} & \vdots & \vdots & \vdots & \\ \cdots & * & 0 & * & \cdots \\ \hline \cdots & 0 & 1 & * & \cdots \\ \hline \cdots & 0 & 0 & 0 & \cdots \\ & \vdots & \vdots & \vdots & \end{array} \end{pmatrix}$$

Nun iterieren wir das Verfahren mit der Matrix oberhalb der letzten Zeile, in welcher ein von 0 verschiedener Eintrag steht.

(c) Dies folgt aus (a) und (b). □

(A.165) Beispiel. Es seien $B, B', B'' \in \mathbb{R}^{3\times 5}$ gegeben durch

$$B = \begin{pmatrix} -2 & 2 & -6 & -10 & -24 \\ 2 & 3 & -9 & -7 & -15 \\ 1 & 1 & -3 & -2 & -4 \end{pmatrix}, \; B' = \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ 0 & 1 & -3 & -3 & -7 \\ 0 & 0 & 0 & -2 & -4 \end{pmatrix}, \; B'' = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Dann lässt sich B durch eine Folge elementarer Zeilenoperationen in die Matrix B' und in die Matrix B'' überführen.

Beweis. Wir benutzen das Gaußsche Eliminationsverfahren:

$$\begin{pmatrix} -2 & 2 & -6 & -10 & -24 \\ 2 & 3 & -9 & -7 & -15 \\ 1 & 1 & -3 & -2 & -4 \end{pmatrix} \xmapsto{\mathrm{sw}_{1,3}} \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ -2 & 2 & -6 & -10 & -24 \\ 2 & 3 & -9 & -7 & -15 \end{pmatrix} \xmapsto{\mathrm{add}_{2,1,2}} \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ 0 & 4 & -12 & -14 & -32 \\ 2 & 3 & -9 & -7 & -15 \end{pmatrix}$$

$$\xmapsto{\mathrm{add}_{3,1,-2}} \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ 0 & 4 & -12 & -14 & -32 \\ 0 & 1 & -3 & -3 & -7 \end{pmatrix} \xmapsto{\mathrm{sw}_{2,3}} \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ 0 & 1 & -3 & -3 & -7 \\ 0 & 4 & -12 & -14 & -32 \end{pmatrix}$$

$$\xmapsto{\mathrm{add}_{3,2,-4}} \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ 0 & 1 & -3 & -3 & -7 \\ 0 & 0 & 0 & -2 & -4 \end{pmatrix} \xmapsto{\mathrm{mul}_{3,-\frac{1}{2}}} \begin{pmatrix} 1 & 1 & -3 & -2 & -4 \\ 0 & 1 & -3 & -3 & -7 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \xmapsto{\mathrm{add}_{1,3,2}} \begin{pmatrix} 1 & 1 & -3 & 0 & 0 \\ 0 & 1 & -3 & -3 & -7 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\xmapsto{\mathrm{add}_{2,3,3}} \begin{pmatrix} 1 & 1 & -3 & 0 & 0 \\ 0 & 1 & -3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \xmapsto{\mathrm{add}_{1,2,-1}} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$ □

Die Lösungsmenge des linearen Gleichungssystems über $\mathbb{R}$ zur erweiterten Koeffizientenmatrix

$$\left(\begin{array}{cccc|c} -2 & 2 & -6 & -10 & -24 \\ 2 & 3 & -9 & -7 & -15 \\ 1 & 1 & -3 & -2 & -4 \end{array}\right)$$

ist also durch

$$\begin{aligned} \mathrm{Sol}(\begin{pmatrix} -2 & 2 & -6 & -10 \\ 2 & 3 & -9 & -7 \\ 1 & 1 & -3 & -2 \end{pmatrix}, \begin{pmatrix} -24 \\ -15 \\ -4 \end{pmatrix}) = \mathrm{Sol}(\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}) &= \{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix} + a \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix} \mid a \in \mathbb{R} \} \\ &= \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

gegeben.

(A.166) Definition (Elementarmatrizen)**.** Es sei $n \in \mathbb{N}_0$ gegeben. Eine Matrix $R \in K^{n \times n}$ heißt *Elementarmatrix*, falls R durch eine elementare Zeilenoperation aus E_n entsteht.

- Gilt

  $$\mathrm{E}_n \xmapsto{\mathrm{sw}_{k,l}} R$$

  für gewisse $k, l \in [1, n]$, so schreiben wir

  $$\mathrm{Sw}_{k,l} := R.$$

- Gilt

  $$\mathrm{E}_m \xmapsto{\mathrm{add}_{k,l,c}} R$$

  für gewisse $k, l \in [1, n]$ mit $k \neq l$ und $c \in K$, so schreiben wir

  $$\mathrm{Add}_{k,l,c} := R.$$

- Gilt

  $$\mathrm{E}_m \xmapsto{\mathrm{mul}_{k,c}} R$$

  für gewisse $k \in [1, n]$, $c \in K$, $c \neq 0$, so schreiben wir

  $$\mathrm{Mul}_{k,c} := R.$$

(A.167) Beispiel.

(a) In $\mathbb{Q}^{3\times 3}$ ist

$$\mathrm{Sw}_{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

(b) In $\mathbb{Q}^{3\times 3}$ ist

$$\mathrm{Add}_{3,1,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

(c) In $\mathbb{Q}^{3\times 3}$ ist

$$\mathrm{Mul}_{2,-3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(A.168) Bemerkung. Es seien $m, n \in \mathbb{N}_0$ und $A, A' \in K^{m\times n}$ so gegeben, dass A' durch eine elementare Zeilenoperation aus A entsteht. Bezeichnet R die Elementarmatrix, welche durch die gleiche elementare Zeilenoperation auf E_m entsteht, so gilt

$$A' = RA.$$

Beweis. Siehe [12, Kor. (16.27)]. □

(A.169) Satz. Es seien $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$ gegeben. Genau dann ist $A \in \mathrm{GL}_n(K)$, wenn sich $\begin{pmatrix} A \mid \mathrm{E}_n \end{pmatrix}$ durch eine Folge elementarer Zeilenoperationen in eine Matrix $\begin{pmatrix} \mathrm{E}_n \mid B \end{pmatrix}$ für ein $B \in K^{n\times n}$ überführen lässt. In diesem Fall ist $A^{-1} = B$.

Beweis. Siehe [12, Satz (16.28)]. □

(A.170) Beispiel. Es sei $A \in \mathbb{Q}^{3\times 3}$ gegeben durch

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

Dann ist

$$A^{-1} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Beweis. Wir werden elementare Zeilenoperationen auf die Matrix $\left(A \mid \mathrm{E}_3\right)$ an:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \xmapsto{\mathrm{add}_{1,2,-1}} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & -1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \xmapsto{\mathrm{add}_{3,1,-1}} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & -1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & 1 & 1 \end{array}\right)$$

$$\xmapsto{\mathrm{add}_{2,3,-1}} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & -1 & 1 & 1 \end{array}\right) \xmapsto{\mathrm{sw}_{1,3}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right)$$

Folglich ist

$$A^{-1} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

□

B Algebren und Körpererweiterungen

Ziel dieses Abschnitts ist das Kennenlernen einer Methode zur Konstruktion von Körpererweiterungen. Ausgehend von einem Körper K wollen wir einen Körper L so konstruieren, dass Addition bzw. Multiplikation von K durch Einschränkung von Addition bzw. Multiplikation von L entstehen.
Dabei gehen wir in zwei Schritten vor: Zuerst betrachten wir einen Ring A, in welchem K liegt (zumindest im Wesentlichen) und welcher die geforderten Vererbungseigenschaften hinsichtlich der Verknüpfungen erfüllt. Dabei wird A so gewählt, dass dieser Ring gleichzeitig ein Vektorraum über K ist, etwa $A = K^{n\times n}$ für ein $n \in \mathbb{N}$. Dies führt uns auf den Begriff der K-Algebra.
In einem zweiten Schritt suchen wir in A nach einer Unteralgebra L, welche von den Potenzen eines Elements $x \in A$ erzeugt wird, und von welcher wir dann im jeweiligen Fall zeigen können, dass alle Elemente ungleich der Null invertierbar sind, so dass es sich sogar um einen Körper handelt.

$$K \subseteq L = \langle x^i \mid i \in \mathbb{N}_0 \rangle \subseteq A$$

Als Hilfsmittel entwickeln wir zunächst die elementare Theorie der K-Algebren, welche in weiten Teilen ganz analog zu der für K-Vektorräume verläuft.
Im Folgenden bezeichnen wir bis zum Ende des Abschnitts mit K stets einen beliebig gegebenen Körper.

Algebren

Der Begriff der K-Algebra entsteht als Synthese der Begriffe K-Vektorraum und Ring:

(B.1) Definition (Algebra)**.** Eine *Algebra* über K (oder *K-Algebra*) besteht aus einem K-Vektorraum A zusammen mit einer Abbildung $A \times A \to A$, $(x, y) \mapsto xy = x \cdot y$, genannt *Multiplikation*, derart, dass folgende Axiome gelten.

- *Assoziativität der Multiplikation.* Für $x, y, z \in A$ ist

$$x(yz) = (xy)z.$$

- *Existenz der Eins.* Es existiert ein $e \in A$ mit

  $$ex = xe = x$$

  für alle $x \in A$.

- *Distributivität von Addition und Multiplikation.* Für $x, y, z \in A$ ist

  $$(x + y)z = (xz) + (yz),$$
  $$x(y + z) = (xy) + (xz).$$

- *Distributivität von Skalarmultiplikation und Multiplikation.* Für $a \in K$, $x, y \in A$ ist

  $$(ax)y = a(xy),$$
  $$x(ay) = a(xy).$$

Es sei eine K-Algebra A gegeben. Es lässt sich zeigen, dass das neutrale Element bzgl. der Multiplikation in A eindeutig durch diese bestimmt ist. Wir bezeichnen es mit 1 und nennen es die *Eins* von A. Wir haben also $1x = x1 = x$ für alle $x \in A$.

Es sei eine K-Algebra A gegeben. Die ersten beiden hier gelisteten Axiome besagen gerade, dass die unterliegende Menge von A zusammen mit der Multiplikation ein Monoid ist. Die ersten drei hier gelisteten Axiome haben zur Folge, dass die unterliegende abelsche Gruppe von A (also A zusammen mit der Addition, welche durch die Vektorraumstruktur gegeben ist) zusammen mit der Multiplikation ein unitärer Ring ist. Lax gesprochen ist eine K-Algebra also eine Struktur, bei der Vektorraum- und Ringstruktur zusammentreffen, und zwar auf eine verträgliche Art und Weise (nämlich so, dass die Additionen gleich sind, und so, dass Skalarmultiplikation und Multiplikation verträglich sind, siehe das vierte hier gelistete Axiom).

(B.2) Definition (kommutative Algebra)**.** Eine K-Algebra A heißt *kommutativ*, falls die Multiplikation von A kommutativ ist.

(B.3) Konvention. Wir verwenden für Skalarmultiplikation und Multiplikation die gleiche Notation und verwenden die üblichen Konventionen wie das Weglassen von Klammern bei iterierten Produkten oder „Punkt- vor Strichrechnung“.

(B.4) Beispiel.

(a) Der Polynomring $K[X]$ wird eine kommutative K-Algebra bzgl. der üblichen Addition, Multiplikation und Skalarmultiplikation von Polynomen. ([10])

(b) Es sei ein K-Vektorraum V gegeben. Dann wird $\mathrm{End}_K(V)$ eine K-Algebra bzgl. der üblichen Addition, Komposition als Multiplikation, sowie der üblichen Skalarmultiplikation. Die Eins von $\mathrm{End}_K(V)$ ist id_V.

(c) Für $n \in \mathbb{N}_0$ wird $K^{n \times n}$ eine K-Algebra bzgl. der üblichen Addition, Multiplikation und Skalarmultiplikation von Matrizen. Die Eins von $K^{n \times n}$ ist E_n.

(d) Es sei ein Körper L so gegeben, dass die unterliegende Menge von K eine Teilmenge von L ist, und so, dass Addition bzw. Multiplikation von K gerade durch Einschränkung der Addition bzw. der Multiplikation von L gegeben ist. ([11]) Dann wird L eine kommutative K-Algebra mit Addition gegeben durch die Addition des Körpers L, mit Multiplikation gegeben durch die Multiplikation des Körpers L, und mit Skalarmultiplikation gegeben durch die Einschränkung $K \times L \to L$, $(a, x) \mapsto ax$.

(e) Es wird K eine kommutative K-Algebra mit Addition gegeben durch die Addition des Körpers K und mit Multiplikation und Skalarmultiplikation gegeben durch die Multiplikation des Körpers K.

(f) Jede einelementige Menge wird eine K-Algebra (mit der einzig möglichen Addition, der einzig möglichen Skalarmultiplikation und der einzig möglichen Multiplikation).

[10] Aus diesem Grund wird $K[X]$ manchmal auch die *Polynomalgebra* über K in der Unbestimmten X genannt.
[11] Man nennt K dann einen *Unterkörper* von L.

Beweis.

(d) Dies ist ein Spezialfall von (d). □

(B.5) Beispiel. Es wird $\mathbb{R}$ eine $\mathbb{Q}$-Algebra mit Addition gegeben durch die übliche Addition der reellen Zahlen, mit Multiplikation gegeben durch die übliche Multiplikation der reellen Zahlen, und mit Skalarmultiplikation gegeben durch $\mathbb{Q} \times \mathbb{R} \to \mathbb{R}$, $(a, x) \mapsto ax$.

Beweis. Dies ist ein Spezialfall von Beispiel (B.4)(d). □

Im Folgenden wollen wir, ausgehend von K, einen Körper L so konstruieren, dass die Situation in Beispiel (B.4)(d) eintritt. Hierzu bedienen wir uns der Polynomalgebra aus Beispiel (B.4)(a) und der Matrixalgebra aus Beispiel (B.4)(c).

Homomorphismen

Analog zu den Vektorraumhomomorphismen gibt es auch für Algebren die strukturverträglichen Abbildungen, welche wir als nächstes einführen wollen. Unsere Kandidaten für Körpererweiterungen werden Bilder von speziellen solchen Homomorphismen sein.

(B.6) Definition (K-Algebrenhomomorphismus)**.** Es seien K-Algebren A und B gegeben. Ein *K-Algebrenhomomorphismus* von A nach B ist ein K-Vektorraumhomomorphismus $\varphi\colon A \to B$ so, dass folgende Axiome gelten.

- *Verträglichkeit mit den Multiplikationen.* Für alle $x, x' \in A$ gilt
$$\varphi(xx') = \varphi(x)\,\varphi(x').$$
- *Verträglichkeit der Einselemente.* Es ist
$$\varphi(1) = 1.$$

Die beiden hier gelisteten Axiome besagen gerade, dass ein K-Algebrenhomomorphismus bzgl. der unterliegenden Monoidstrukturen von Start und Ziel ein Monoidhomomorphismus ist. Sie haben außerdem zur Folge, dass ein K-Algebrenhomomorphismus bzgl. der unterliegenden Ringstrukturen von Start und Ziel ein Ringhomomorphismus ist.

(B.7) Beispiel. Die Abbildung
$$\varphi\colon K^{2\times 2} \to K^{2\times 2},\ A \mapsto \begin{pmatrix} A_{2,2} & A_{2,1} \\ A_{1,2} & A_{1,1} \end{pmatrix}$$
ist ein K-Algebrenhomomorphismus.

Beweis. Für $A, A' \in K^{2\times 2}$ gilt
$$\begin{aligned}\varphi(A + A') &= \varphi(\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} + \begin{pmatrix} A'_{1,1} & A'_{1,2} \\ A'_{2,1} & A'_{2,2} \end{pmatrix}) = \varphi(\begin{pmatrix} A_{1,1} + A'_{1,1} & A_{1,2} + A'_{1,2} \\ A_{2,1} + A'_{2,1} & A_{2,2} + A'_{2,2} \end{pmatrix}) \\ &= \begin{pmatrix} A_{2,2} + A'_{2,2} & A_{2,1} + A'_{2,1} \\ A_{1,2} + A'_{1,2} & A_{1,1} + A'_{1,1} \end{pmatrix} = \begin{pmatrix} A_{2,2} & A_{2,1} \\ A_{1,2} & A_{1,1} \end{pmatrix} + \begin{pmatrix} A'_{2,2} & A'_{2,1} \\ A'_{1,2} & A'_{1,1} \end{pmatrix} = \varphi(A) + \varphi(A'),\end{aligned}$$
und für $c \in K$, $A \in K^{2\times 2}$ gilt
$$\varphi(cA) = \varphi(c\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}) = \varphi(\begin{pmatrix} cA_{1,1} & cA_{1,2} \\ cA_{2,1} & cA_{2,2} \end{pmatrix}) = \begin{pmatrix} cA_{2,2} & cA_{2,1} \\ cA_{1,2} & cA_{1,1} \end{pmatrix} = c\begin{pmatrix} A_{2,2} & A_{2,1} \\ A_{1,2} & A_{1,1} \end{pmatrix} = c\,\varphi(A).$$
Nach dem Kriterium für Vektorraumhomomorphismen (2.2) ist φ somit ein Vektorraumhomomorphismus. Für $A, A' \in K^{2\times 2}$ gilt ferner
$$\varphi(AA') = \varphi(\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}\begin{pmatrix} A'_{1,1} & A'_{1,2} \\ A'_{2,1} & A'_{2,2} \end{pmatrix}) = \varphi(\begin{pmatrix} A_{1,1}A'_{1,1} + A_{1,2}A'_{2,1} & A_{1,1}A'_{1,2} + A_{1,2}A'_{2,2} \\ A_{2,1}A'_{1,1} + A_{2,2}A'_{2,1} & A_{2,1}A'_{1,2} + A_{2,2}A'_{2,2} \end{pmatrix})$$

$$= \begin{pmatrix} A_{2,1}A'_{1,2} + A_{2,2}A'_{2,2} & A_{2,1}A'_{1,1} + A_{2,2}A'_{2,1} \\ A_{1,1}A'_{1,2} + A_{1,2}A'_{2,2} & A_{1,1}A'_{1,1} + A_{1,2}A'_{2,1} \end{pmatrix} = \begin{pmatrix} A_{2,2}A'_{2,2} + A_{2,1}A'_{1,2} & A_{2,2}A'_{2,1} + A_{2,1}A'_{1,1} \\ A_{1,2}A'_{2,2} + A_{1,1}A'_{1,2} & A_{1,2}A'_{2,1} + A_{1,1}A'_{1,1} \end{pmatrix}$$
$$= \begin{pmatrix} A_{2,2} & A_{2,1} \\ A_{1,2} & A_{1,1} \end{pmatrix} \begin{pmatrix} A'_{2,2} & A'_{2,1} \\ A'_{1,2} & A'_{1,1} \end{pmatrix} = \varphi(A)\,\varphi(A'),$$

und es ist

$$\varphi(1) = \varphi(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

Insgesamt ist φ ein Algebrenhomomorphismus. □

Wir betonen, dass beim Nachweis eines Algebrenhomomorphismus auf die Verträglichkeit der Einselemente *nicht* verzichtet werden kann (im Gegensatz zur Verträglichkeit der Nullvektoren bei Vektorraumhomomorphismen, vgl. Definition (2.1) und das Kriterium für Vektorraumhomomorphismen (2.2)). Beispielsweise ist

$$\varphi\colon K \to K^{2\times 2},\ x \mapsto \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}$$

ein K-Vektorraumhomomorphismus, welcher verträglich mit den Multiplikationen von K und $K^{2\times 2}$ ist, aber nicht verträglich mit den Einselementen von K und $K^{2\times 2}$ ist: Es ist

$$\varphi(1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{E}_2 = 1^{K^{2\times 2}}.$$

Insbesondere ist φ daher *kein* K-Algebrenhomomorphismus.
Nichtsdestotrotz sind die Axiome eines K-Algebrenhomomorphismus etwas redundant, wie wir nun sehen werden.

(B.8) Bemerkung (Kriterium für Algebrenhomomorphismen)**.** Es seien K-Algebren A und B und eine Abbildung $\varphi\colon A \to B$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist φ ein K-Algebrenhomomorphismus.

(b) Es gilt:

- *Verträglichkeit mit den Additionen.* Für $x, x' \in A$ ist $\varphi(x + x') = \varphi(x) + \varphi(x')$.
- *Verträglichkeit mit den Multiplikationen.* Für $x, x' \in A$ ist $\varphi(xx') = \varphi(x)\,\varphi(x')$.
- *Verträglichkeit mit den Skalaren.* Für $a \in K$ ist $\varphi(a \cdot 1) = a \cdot 1$.

Beweisskizze. Dies lässt sich ähnlich zum Kriterium für Vektorraumhomomorphismen (2.2) zeigen. □

Auch der Begriff des Vektorraumisomorphismus hat eine Entsprechung für Algebren:

(B.9) Definition (K-Algebrenisomorphismus)**.** Es seien K-Algebren A und B gegeben. Ein *K-Algebrenisomorphismus* von A nach B ist ein K-Algebrenhomomorphismus $\varphi\colon A \to B$ so, dass φ eine invertierbare Abbildung und $\varphi^{-1}\colon B \to A$ ein K-Algebrenhomomorphismus ist.

(B.10) Bemerkung. Es sei ein K-Algebrenhomomorphismus $\varphi\colon A \to B$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist φ ein K-Algebrenisomorphismus.

(b) Es ist φ eine invertierbare Abbildung.

(c) Es ist φ bijektiv.

Beweisskizze. Dies lässt sich ähnlich zu Bemerkung (2.9) zeigen. □

Schließlich betrachten wir Homomorphismen, welche uns erlauben, den Körper K in fast jede K-Algebra einzubetten.

(B.11) Proposition. Es sei eine K-Algebra A gegeben. Die Abbildung

$$\iota\colon K \to A,\ a \mapsto a \cdot 1^A$$

ist ein K-Algebrenhomomorphismus. Wenn $A \neq \{0\}$ ist, dann ist ι injektiv.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Wie im Beispiel der Polynomalgebra, siehe Arbeitsbasis (A.115)(e), fassen wir im Folgenden K als eine Teilmenge einer K-Algebra ungleich $\{0\}$ auf:

(B.12) Konvention. Es sei eine K-Algebra $A \neq \{0\}$ gegeben. Von jetzt an identifizieren wir K mit dem Bild des injektiven K-Algebrenhomomorphismus $\iota\colon K \to A$, $a \mapsto a \cdot 1^A$ aus Proposition (B.11). Das heißt, unter Missbrauch der Notationen schreiben wir K anstatt $\operatorname{Im}\iota$, und für $a \in K$ notieren wir das Bild $\iota(a)$ von a auch durch a.

So fassen wir beispielsweise $\mathbb{R}$ als Teilmenge von $\mathbb{R}^{2\times 2}$ auf, indem wir $a \in \mathbb{R}$ mit $a \cdot 1 = a\mathrm{E}_2$ identifizieren. Für $i \in \mathbb{R}^{2\times 2}$ gegeben durch

$$i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

machen dann Ausdrücke wie $a + bi$ für $a, b \in \mathbb{R}$ einen Sinn, es ist

$$a + bi = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Es seien K-Algebren A und B und eine Abbildung $\varphi\colon A \to B$ gegeben. Falls $A \neq \{0\}$ und $B \neq \{0\}$ ist, so lässt sich die Verträglichkeit mit den Skalaren im Kriterium für Algebrenhomomorphismen (B.8) dank Konvention (B.12) wie folgt umformulieren:

- *Verträglichkeit mit den Skalaren.* Für $a \in K$ ist $\varphi(a) = a$.

Mit anderen Worten: Es ist $\varphi(K) \subseteq K$ und $\varphi|_K^K = \mathrm{id}_K$.

Unteralgebren

Als nächstes betrachten wir die Unterstrukturen von Algebren, welche in Analogie zu den Untervektorräumen definiert sind, vgl. Definition (1.12).

(B.13) Definition (Unteralgebra)**.** Es sei eine K-Algebra A gegeben. Eine K-*Unteralgebra* von A ist eine K-Algebra U derart, dass der unterliegende K-Vektorraum von U ein K-Untervektorraum von A ist, und so, dass für $u, u' \in U$ stets

$$u \cdot^U u' = u \cdot^A u'$$

gilt.

Eine Unteralgebra einer K-Algebra A ist also eine K-Algebra U derart, dass U eine Teilmenge von A ist und die Operationen von U gerade als Einschränkungen der jeweiligen Operationen von A entstehen.
Wie bei den Untervektorräumen treffen wir folgende Vereinbarung.

(B.14) Konvention. Es seien eine K-Algebra A und eine Teilmenge U von A gegeben. Da die Addition bzw. die Skalarmultiplikation bzw. die Multiplikation jeder K-Unteralgebra von A vollständig durch die Addition bzw. die Skalarmultiplikation bzw. die Multiplikation von A bestimmt ist, gibt es höchstens eine Algebrenstruktur auf U so, dass U mit dieser Algebrenstruktur eine K-Unteralgebra von A wird. Wir sagen daher auch, dass U eine K-Unteralgebra von A *ist*, falls so eine Algebrenstruktur auf U existiert.

Unter Verwendung von Konvention (B.14) geben wir nun ein Kriterium zur Erkennung von Unteralgebren an.

(B.15) Lemma (Unteralgebrakriterium)**.** Es seien eine K-Algebra A und eine Teilmenge U von A gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist U eine K-Unteralgebra von A.

(b) Es gilt:

- *Abgeschlossenheit unter der Addition.* Für $u, u' \in U$ ist
$$u + u' \in U.$$
- *Abgeschlossenheit unter der Multiplikation.* Für $u, u' \in U$ ist
$$uu' \in U.$$
- *Abgeschlossenheit unter den Skalaren.* Für $a \in K$ ist
$$a \cdot 1^A \in U.$$

Beweisskizze. Dies lässt sich ähnlich zum Untervektorraumkriterium (1.15) zeigen. □

Falls $A \neq \{0\}$ ist, so lässt sich die Abgeschlossenheit unter den Skalaren im Unteralgebrakriterium (B.15) dank Konvention (B.12) wie folgt umformulieren:

- *Abgeschlossenheit unter den Skalaren.* Es ist
$$K \subseteq U.$$

Unsere Beispiele für Unteralgebren werden Bilder von Algebrenhomomorphismen sein. Wir halten fest, dass solche Bilder stets Unteralgebren der Zielalgebra sind:

(B.16) Bemerkung. Es sei ein K-Algebrenhomomorphismus $\varphi\colon A \to B$ gegeben. Dann ist $\operatorname{Im} \varphi$ eine K-Unteralgebra von B. ([12])

Beweis. Dies lässt sich ähnlich zu Bemerkung (2.12)(a) zeigen. □

Der Einsetzungshomomorphismus

Als nächstes wollen wir Unteralgebren einer K-Algebra A betrachten, welche jeweils durch Einsetzen eines Elements x von A in alle Polynome über K (in einer Unbestimmten) entstehen. Hierzu betrachten wir den sogenannten *Einsetzungshomomorphismus*.
Zunächst sei daher an den Begriff des Polynoms erinnert: Ein Polynom über K in der Unbestimmten X hat die Form

$$f = \sum_{i \in \mathbb{N}_0} a_i X^i$$

für ein $a = (a_i)_{i \in \mathbb{N}_0} \in K^{(\mathbb{N}_0)}$, d.h. für ein $a \in K^{\mathbb{N}_0}$ so, dass $\{i \in \mathbb{N}_0 \mid a_i \neq 0\}$ endlich ist. Es ist $(X^i)_{i \in \mathbb{N}_0}$ eine Basis von $K[X]$. Für $a, b \in K^{(\mathbb{N}_0)}$ gilt auf Grund der linearen Unabhängigkeit somit genau dann

$$\sum_{i \in \mathbb{N}_0} a_i X^i = \sum_{i \in \mathbb{N}_0} b_i X^i,$$

wenn $a = b$ ist.

(B.17) Proposition. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Die Abbildung $\operatorname{ev}_x\colon K[X] \to A$ gegeben durch

$$\operatorname{ev}_x(\sum_{i \in \mathbb{N}_0} a_i X^i) = \sum_{i \in \mathbb{N}_0} a_i x^i$$

für $a \in K^{(\mathbb{N}_0)}$ ist ein K-Algebrenhomomorphismus.

[12]Die analoge Aussage für den Kern ist im Allgemeinen falsch.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(B.18) Definition (Einsetzungshomomorphismus)**.** Es seien eine K-Algebra A und ein $x \in A$ gegeben. Der K-Algebrenhomomorphismus $\mathrm{ev}_x \colon K[X] \to A$ gegeben durch

$$\mathrm{ev}_x(\sum_{i \in \mathbb{N}_0} a_i X^i) = \sum_{i \in \mathbb{N}_0} a_i x^i$$

für $a \in K^{(\mathbb{N}_0)}$ heißt *Einsetzungshomomorphismus*. Wir schreiben

$$f(x) := \mathrm{ev}_x(f)$$

für $f \in K[X]$.

Mit der in Definition (B.18) vereinbarten Notation lässt sich der Einsetzungshomomorphismus für ein Element x einer K-Algebra A kurz wie folgt notieren:

$$\mathrm{ev}_x \colon K[X] \to A,\ f \mapsto f(x)$$

Wir veranschaulichen das Einsetzen von Algebrenelementen in Polynome noch einmal an einigen Beispielen:

(B.19) Beispiel. Es seien ein Körper K sowie $f \in K[X]$, $a \in K^{(\mathbb{N}_0)}$ mit $f = \sum_{i \in \mathbb{N}_0} a_i X^i$ gegeben.

(a) Für $b \in K$ ist $f(b) \in K$ gegeben durch

$$f(b) = \sum_{i \in \mathbb{N}_0} a_i b^i.$$

(b) Es sei $n \in \mathbb{N}_0$ gegeben. Für $A \in K^{n \times n}$ ist

$$f(A) = \sum_{i \in \mathbb{N}_0} a_i A^i.$$

(c) Es ist

$$f(X) = \sum_{i \in \mathbb{N}_0} a_i X^i = f.$$

Wir sind an den Bildern von Einsetzungshomomorphismen interessiert:

(B.20) Notation. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Wir setzen

$$K[x] := \operatorname{Im} \mathrm{ev}_x.$$

Notation (B.20) ist verträglich mit der Notation des Polynomrings $K[X]$: Der Einsetzungshomomorphismus $\mathrm{ev}_X \colon K[X] \to K[X]$ ist gegeben durch $\mathrm{ev}_X = \mathrm{id}_{K[X]}$. Vgl. Beispiel (B.19)(c).

(B.21) Bemerkung. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Dann ist $K[x]$ eine kommutative K-Unteralgebra von A.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(B.22) Bemerkung. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Dann ist

$$(x^i)_{i \in \mathbb{N}_0}$$

ein Erzeugendensystem von $K[x]$ als K-Vektorraum.

Beweis. Es ist

$$K[x] = \operatorname{Im} \mathrm{ev}_x = \{\mathrm{ev}_x(f) \mid f \in K[X]\} = \{f(x) \mid f \in K[X]\} = \{\sum_{i \in \mathbb{N}_0} a_i x^i \mid a \in K^{(\mathbb{N}_0)}\} = \langle x^i \mid i \in \mathbb{N}_0 \rangle$$

und damit $(x^i)_{i \in \mathbb{N}_0}$ ein Erzeugendensystem von $K[x]$. □

Es seien eine K-Algebra A und ein $x \in A$ gegeben. Lax gesprochen sind die Elemente von $K[x]$ also „polynomielle Ausdrücke“ in x, es ist

$$K[x] = \{f(x) \mid f \in K[X]\} = \{\sum_{i \in \mathbb{N}_0} a_i x^i \mid a \in K^{(\mathbb{N}_0)}\}.$$

Der Unterschied zu „echten Polynomen“ liegt darin, dass die Koeffizienten der Elemente von $K[x]$ nicht durch diese bestimmt sein müssen, es kann unter Umständen $a, b \in K^{(\mathbb{N}_0)}$ mit $\sum_{i \in \mathbb{N}_0} a_i x^i = \sum_{i \in \mathbb{N}_0} b_i x^i$ und $a \neq b$ geben. Dies ist genau dann der Fall, wenn der Einsetzungshomomorphismus $\mathrm{ev}_x \colon K[X] \to A$, $f \mapsto f(x)$ nicht injektiv ist, bzw. falls $(x^i)_{i \in \mathbb{N}_0}$ ein linear abhängiges Erzeugendensystem von $K[x]$ ist. In diesem Fall werden wir, analog zum endlichen Fall im Basisauswahlsatz (1.49), das Erzeugendensystem $(x^i)_{i \in \mathbb{N}_0}$ zu einer endlichen Basis von $K[x]$ verkürzen, und zwar sogar auf eine ganz spezielle Art und Weise.

(B.23) Bemerkung. Es seien eine K-Algebra A, ein $x \in A$ und ein $n \in \mathbb{N}_0$ gegeben. Genau dann ist $(1, x, \ldots, x^{n-1})$ ein Erzeugendensystem von $K[x]$, wenn x^n eine Linearkombination von $(1, x, \ldots, x^{n-1})$ ist.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(B.24) Korollar. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Wenn es ein $n \in \mathbb{N}_0$ derart gibt, dass $(1, x, \ldots, x^{n-1})$ linear unabhängig und x^n eine Linearkombination von $(1, x, \ldots, x^{n-1})$ ist, dann ist $(1, x, \ldots, x^{n-1})$ eine Basis von $K[x]$.

Beweis. Wenn es ein $n \in \mathbb{N}_0$ so gibt, dass $(1, x, \ldots, x^{n-1})$ linear unabhängig und x^n eine Linearkombination von $(1, x, \ldots, x^{n-1})$ ist, dann ist $(1, x, \ldots, x^{n-1})$ nach Bemerkung (B.23) ein Erzeugendensystem und damit eine Basis von $K[x]$. □

Nachdem wir nun eine Methode zur Bestimmung einer Basis von endlich erzeugten Unteralgebren der Form $K[x]$ für ein Element x einer K-Algebra A kennengelernt haben, können wir nun einige Beispiele solcher Unteralgebren explizit beschreiben:

(B.25) Beispiel.

(a) Für jede K-Algebra $A \neq \{0\}$ ist (1) eine Basis von $K[1] = K$.

(b) Es ist $(1, \sqrt{2})$ eine Basis von $\mathbb{Q}[\sqrt{2}]$.

(c) Es sei $i \in \mathbb{R}^{2 \times 2}$ gegeben durch

$$i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

und es sei $\mathbb{C}$ die $\mathbb{R}$-Unteralgebra von $\mathbb{R}^{2 \times 2}$ gegeben durch

$$\mathbb{C} = \mathbb{R}[i].$$

Dann ist $(1, i)$ eine Basis von $\mathbb{C}$.

Beweis.

(a) Wegen $A \neq \{0\}$ ist $1 \neq 0$ und damit (1) linear unabhängig in A. Ferner ist $1^1 = 1$ eine Linearkombination von (1) und damit (1) nach Korollar (B.24) eine Basis von $K[1]$. Es folgt $K[1] = \langle 1 \rangle = K$.

(b) Wegen $1 \neq 0$ in $\mathbb{R}$ ist (1) linear unabhängig in $\mathbb{R}$ über $\mathbb{Q}$. Auf Grund der Irrationalität von $\sqrt{2}$ ist ferner $\sqrt{2} \notin \mathbb{Q} = \langle 1 \rangle_{\mathbb{Q}}$ und damit $(1, \sqrt{2})$ linear unabhängig in $\mathbb{R}$ über $\mathbb{Q}$ nach Proposition (1.37). Schließlich ist $(\sqrt{2})^2 = 2 = 2 \cdot 1 + 0 \cdot \sqrt{2}$ eine Linearkombination von $(1, \sqrt{2})$ über $\mathbb{Q}$ und damit $(1, \sqrt{2})$ nach Korollar (B.24) eine Basis von $\mathbb{Q}[\sqrt{2}]$.

(c) Dies sei dem Leser zur Übung überlassen. □

Körpererweiterungen

Schließlich können wir unser vorgegebenes Ziel präzise formulieren und charakterisieren.

(B.26) Definition (Körpererweiterung)**.** Eine *Körpererweiterung* über K (oder *K-Körpererweiterung*) ist eine K-Algebra L so, dass der unterliegende Ring von L ein Körper ist.

(B.27) Bemerkung. Es sei eine K-Algebra L gegeben. Genau dann ist L eine K-Körpererweiterung, wenn L kommutativ und $L^\times = L \setminus \{0\}$ ist.

Das nachfolgende Korollar gibt nun eine Antwort auf die Frage, wann das Bild eines Einsetzungshomomorphismus eine Körpererweiterung ist.

(B.28) Korollar. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Genau dann ist $K[x]$ eine K-Körpererweiterung, wenn $K[x]^\times = K[x] \setminus \{0\}$ ist.

Beweis. Nach Bemerkung (B.21) ist $K[x]$ eine kommutative K-Algebra und somit genau dann eine K-Körpererweiterung, wenn $K[x]^\times = K[x] \setminus \{0\}$ ist. □

(B.29) Beispiel.

(a) Für jede K-Algebra $A \neq \{0\}$ ist $K[1] = K$ eine K-Körpererweiterung.

(b) Es ist $\mathbb{Q}[\sqrt{2}]$ eine $\mathbb{Q}$-Körpererweiterung.

(c) Es sei $i \in \mathbb{R}^{2\times 2}$ gegeben durch

$$i = \mathrm{C}(X^2 + 1) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Dann ist $\mathbb{R}[i]$ eine $\mathbb{R}$-Körpererweiterung mit

$$\mathbb{R}[i] \cong \mathbb{C}.$$

(d) Es sei $\alpha \in \mathbb{F}_2^{2\times 2}$ gegeben durch

$$\alpha = \mathrm{C}(X^2 + X + 1) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Dann ist $\mathbb{F}_2[\alpha]$ eine $\mathbb{F}_2$-Körpererweiterung mit

$$\mathbb{F}_2[\alpha] \cong \mathbb{F}_4.$$

(e) Es sei $\beta \in \mathbb{F}_2^{3\times 3}$ gegeben durch

$$\beta = \mathrm{C}(X^3 + X + 1) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Dann ist $\mathbb{F}_2[\beta]$ eine $\mathbb{F}_2$-Körpererweiterung mit

$$\mathbb{F}_2[\beta] \cong \mathbb{F}_8.$$

(f) Es sei $\iota \in \mathbb{F}_3^{2\times 2}$ gegeben durch

$$\iota = \mathrm{C}(X^2 + 1) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Dann ist $\mathbb{F}_3[\iota]$ eine $\mathbb{F}_3$-Körpererweiterung mit

$$\mathbb{F}_3[\iota] \cong \mathbb{F}_9.$$

Beweis.

(b) Es sei $x \in \mathbb{Q}[\sqrt{2}] \setminus \{0\}$ gegeben. Nach Beispiel (B.25)(b) ist $(1, \sqrt{2})$ eine Basis von $\mathbb{Q}[\sqrt{2}]$ über $\mathbb{Q}$. Folglich gibt es (genau) ein $a \in \mathbb{Q}^{\{0,1\}} \setminus \{0\}$ mit $x = a_0 + a_1\sqrt{2}$. Wegen $\sqrt{2} \notin \mathbb{Q}$ und $(a_0, a_1) \neq (0,0)$ ist $a_0^2 - 2a_1^2 \neq 0$. Es gilt

$$x\,(a_0 - a_1\sqrt{2}) = (a_0 + a_1\sqrt{2})(a_0 - a_1\sqrt{2}) = (a_0^2 - 2a_1^2) + (a_1a_0 - a_0a_1)\sqrt{2} = a_0^2 - 2a_1^2$$

und damit

$$(a_0 + a_1\sqrt{2}) \cdot \frac{1}{a_0^2 - 2a_1^2}(a_0 - a_1\sqrt{2}) = 1.$$

Wegen der Kommutativität von $\mathbb{Q}[\sqrt{2}]$ nach Bemerkung (B.21) ist somit $x \in \mathbb{Q}[\sqrt{2}]^\times$ mit

$$x^{-1} = \frac{1}{a_0^2 - 2a_1^2}(a_0 - a_1\sqrt{2}).$$

Folglich ist $\mathbb{Q}[\sqrt{2}]^\times = \mathbb{Q}[\sqrt{2}] \setminus \{0\}$ und somit $\mathbb{Q}[\sqrt{2}]$ eine $\mathbb{Q}$-Körpererweiterung nach Korollar (B.28).

(c) Dies sei dem Leser zur Übung überlassen.

(d) Dies sei dem Leser zur Übung überlassen.

(e) Dies sei dem Leser zur Übung überlassen.

(f) Dies sei dem Leser zur Übung überlassen. □

Zusätzliche Konzepte

Nachdem wir weiter oben bereits das Bild des Einsetzungshomomorphismus näher beleuchtet haben, betrachten wir im Folgenden mit dem Minimalpolynom noch ein zusätzliches Konzept, welches in engem Zusammenhang zum Kern des Einsetzungshomomorphimus steht. Das Studium der Details sei dem Leser zur Übung überlassen. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Wir studieren Polynome μ über K derart, dass $\mu(x) = 0$ ist und dass für $f \in K[X]$ mit $f(x) = 0$ stets $\mu \mid f$ folgt, also kleinste Elemente in der prägeordneten Menge $\operatorname{Ker} \operatorname{ev}_x = \{f \in K[X] \mid f(x) = 0\}$ zusammen mit der Teilbarkeitsrelation, vgl. [12, Def. (8.15)(a)].

(B.30) Bemerkung. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Ferner sei $\mu \in K[X]$ derart gegeben, dass $\mu(x) = 0$ ist und dass für $f \in K[X]$ mit $f(x) = 0$ stets $\mu \mid f$ folgt. Schließlich sei $\mu' \in K[X]$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Es ist $\mu'(x) = 0$ und für $f \in K[X]$ mit $f(x) = 0$ gilt $\mu' \mid f$.

(b) Es ist $\mu'(x) = 0$ und für $f \in K[X]$ mit $f(x) = 0$ und $f \mid \mu'$ gilt auch $\mu' \mid f$.

(c) Es gilt $\mu \mid \mu'$ und $\mu' \mid \mu$.

(d) Es ist $\mu'(x) = 0$ und es gilt $\mu' \mid \mu$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(B.31) Proposition. Es seien eine K-Algebra A, $x \in A$ und $\mu \in K[X] \setminus \{0\}$ mit $\mu(x) = 0$ gegeben. Die folgenden Bedingungen sind äquivalent.

(a) Für $f \in K[X]$ mit $f(x) = 0$ gilt $\mu \mid f$.

(b) Es ist $\deg \mu = \min \{\deg f \mid f \in \operatorname{Ker} \operatorname{ev}_x \setminus \{0\}\}$, d.h. für $f \in K[X] \setminus \{0\}$ mit $f(x) = 0$ gilt $\deg \mu \leq \deg f$.

Beweis. Dies sei dem Leser zur Übung überlassen. □

Es sei eine K-Algebra A gegeben. Nach Proposition (B.31) gibt es für jedes $x \in A$ mit $\operatorname{Ker} \operatorname{ev}_x \neq \{0\}$ ein $\mu \in K[X]$ derart, dass $\mu(x) = 0$ ist und dass für $f \in K[X]$ mit $f(x) = 0$ stets $\mu \mid f$ folgt, und nach Bemerkung (B.30) sind diese im Wesentlichen eindeutig. Genauer:

(B.32) Korollar. Es seien eine K-Algebra A und ein $x \in A$ mit $\operatorname{Ker} \operatorname{ev}_x \neq \{0\}$ gegeben. Dann gibt es genau ein normiertes $\mu \in K[X] \setminus \{0\}$ derart, dass $\mu(x) = 0$ ist und dass für $f \in K[X]$ mit $f(x) = 0$ stets $\mu \mid f$ folgt.

Beweis. Dies sei dem Leser zur Übung überlassen. □

(B.33) Definition (Minimalpolynom)**.** Es seien eine K-Algebra A und ein $x \in A$ mit $\operatorname{Ker} \operatorname{ev}_x \neq \{0\}$ gegeben und es sei $\mu \in K[X] \setminus \{0\}$ das eindeutige normierte Polynom derart, dass $\mu(x) = 0$ ist und dass für $f \in K[X]$ mit $f(x) = 0$ stets $\mu \mid f$ folgt. Wir nennen $\mathrm{\mu}_x := \mu$ das *Minimalpolynom* von x (über K).

(B.34) Beispiel. Das Minimalpolynom von $\sqrt{2}$ in der $\mathbb{Q}$-Algebra $\mathbb{R}$ ist

$$\mathrm{\mu}_{\sqrt{2}} = X^2 - 2.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Die in Korollar (B.24) beschriebene Situation führt zu einer naiven Berechnungsmethode für das Minimalpolynom:

(B.35) Bemerkung. Es seien eine K-Algebra A, $x \in A$, $n \in \mathbb{N}_0$ und $a \in K^{[0,n-1]}$ mit

$$x^n = \sum_{i \in [0,n-1]} a_i x^i$$

gegeben. Ferner sei $f := X^n - \sum_{i \in [0,n-1]} a_i X^i$. Dann ist $f(x) = 0$.

(B.36) Korollar. Es seien eine K-Algebra A und ein $x \in A$ gegeben. Ferner sei $n \in \mathbb{N}_0$ so gegeben, dass $(1, x, \ldots, x^{n-1})$ linear unabhängig ist, und es sei $a \in K^{[0,n-1]}$ mit

$$x^n = \sum_{i \in [0,n-1]} a_i x^i$$

gegeben. Dann ist $\operatorname{Ker} \operatorname{ev}_x \neq \{0\}$ und

$$\mathrm{\mu}_x = X^n - \sum_{i \in [0,n-1]} a_i X^i.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Insbesondere ist die Voraussetzung aus Korollar (B.32) immer dann erfüllt, wenn die gegebene K-Algebra endlichdimensional ist:

(B.37) Korollar. Es seien eine endlichdimensionale K-Algebra A und ein $x \in A$ gegeben. Dann ist $\operatorname{Ker} \operatorname{ev}_x \neq \{0\}$ und

$$\deg \mathrm{\mu}_x \leq \dim_K A.$$

Beweis. Dies sei dem Leser zur Übung überlassen. □

Literatur

[1] Brin, Sergey; Page, Lawrence. *The anatomy of a large-scale hypertextual Web search engine.* Computer Networks and ISDN Systems **30** (1998), S. 107–117.

[2] Bryan, Kurt; Leise, Tanya. *The* $25,000,000,000 *eigenvector: the linear algebra behind Google.* SIAM Review **48**(3) (2006), S. 569–581.

[3] Hanke, Timo; Hiss, Gerhard. *Diskrete Strukturen und Lineare Algebra für Informatiker.* Vorlesungsmanuskript, RWTH Aachen, 2012.

[4] Hess, Florian. *Skript zur Linearen Algebra I + II.* Vorlesungsmanuskript, Otto-von-Guericke-Universität Magdeburg, 2010.

[5] Hiss, Gerhard. *Lineare Algebra I.* Vorlesungsnotizen, RWTH Aachen, 2006.

[6] Hiss, Gerhard. *Lineare Algebra II.* Vorlesungsnotizen, RWTH Aachen, 2006.

[7] Huppert, Bertram; Willems, Wolfgang. *Lineare Algebra.* Vieweg+Teubner, 2. Auflage, 2010. DOI: 10.1007/978-3-8348-9710-7.

[8] Kreussler, Bernd; Pfister, Gerhard. *Mathematik für Informatiker. Algebra, Analysis, Diskrete Strukturen.* eXamen.press. Springer-Verlag, Berlin/Heidelberg, 2009. DOI: 10.1007/978-3-540-89107-9.

[9] Künzer, Matthias. *Lineare Algebra für Informatiker.* Vorlesungsmanuskript, Universität Ulm, 2004.

[10] Schoenwaelder, Ulrich. *Lineare Algebra II.* Vorlesungsnotizen, RWTH Aachen, 2003.

[11] Stirner, Matthias. *JPEG - Das Bildformat Teil 1: Theorie und Grundlagen.* Internetveröffentlichung, 2011. Webseite: `http://www.burosch.de/technische-informationen/339-jpeg-das-bildformat-teil-1-theorie-und-grundlagen.html` (abgerufen am 28. März 2018)

[12] Thomas, Sebastian. *Diskrete Strukturen.* Vorlesungsmanuskript, RWTH Aachen, 2018 (Version 1.3, 16. Februar 2018).

[13] Thomas, Sebastian. *Lineare Algebra.* Vorlesungsmanuskript, Carl von Ossietzky Universität Oldenburg, 2013.

[14] Thomas, Sebastian. *Lineare Algebra für Informatiker.* Vorlesungsmanuskript, RWTH Aachen, 2017 (Version 2.4.1, 1. August 2017).

[15] Willems, Wolfgang. *Codierungstheorie und Kryptographie.* Mathematik Kompakt. Birkhäuser Verlag, Basel, 2008. DOI: 10.1007/978-3-7643-8612-2.

Sebastian Thomas
Lehrstuhl D für Mathematik
RWTH Aachen University
Pontdriesch 14/16
52062 Aachen
Germany
sebastian.thomas@math.rwth-aachen.de
`http://www.math.rwth-aachen.de/~Sebastian.Thomas/`