JOHANNES KEPLER UNIVERSITÄT LINZ | JKU

**Unterlagen zu den Vorlesungen**

# Lineare Algebra und Analytische Geometrie 1 und 2

Wintersemester 2010/11
Sommersemester 2011

Erhard Aichinger
Institut für Algebra
Johannes Kepler Universität Linz



Version 5. Oktober 2010

Adresse:

Assoc.-Prof. Dr. Erhard Aichinger
Institut für Algebra, Johannes Kepler Universität Linz
4040 Linz, Österreich
e-mail: `erhard.aichinger@jku.at`

Version 5.10.2010

Druck: Kopierstelle, Abteilung Service, Universität Linz

# Inhaltsverzeichnis

# Teil 1

# Vektoren und Matrizen

KAPITEL 1

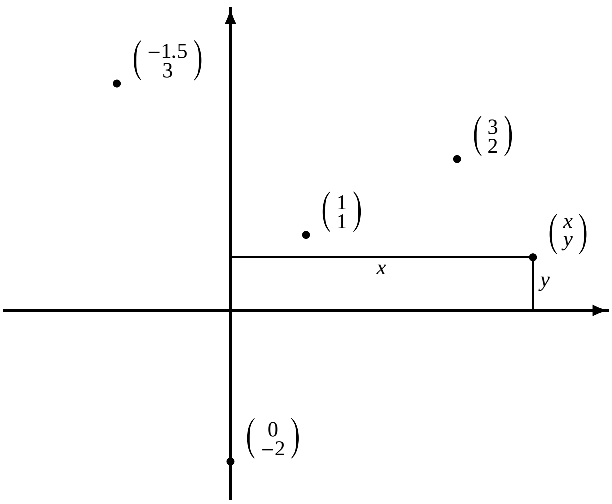
# Geometrie in der Ebene und im Raum

## 1. Koordinaten

Wir beschreiben – nach einer Idee von René Descartes (1596 – 1650) – jeden Punkt in der Ebene durch ein Paar reeller Zahlen. Die Menge der Paare reeller Zahlen kürzen wir mit $\mathbb{R} \times \mathbb{R}$ oder $\mathbb{R}^2$ ab.

$$\mathbb{R} \times \mathbb{R} := \{\binom{x}{y} \mid x \in \mathbb{R} \text{ und } y \in \mathbb{R}\}.$$

Für das Paar $\binom{x}{y}$ schreiben wir auch $(x, y)$. Aus der folgenden Skizze ist ersichtlich, wie wir jeden Punkt durch ein Zahlenpaar (seine *kartesischen Koordinaten*) beschreiben.

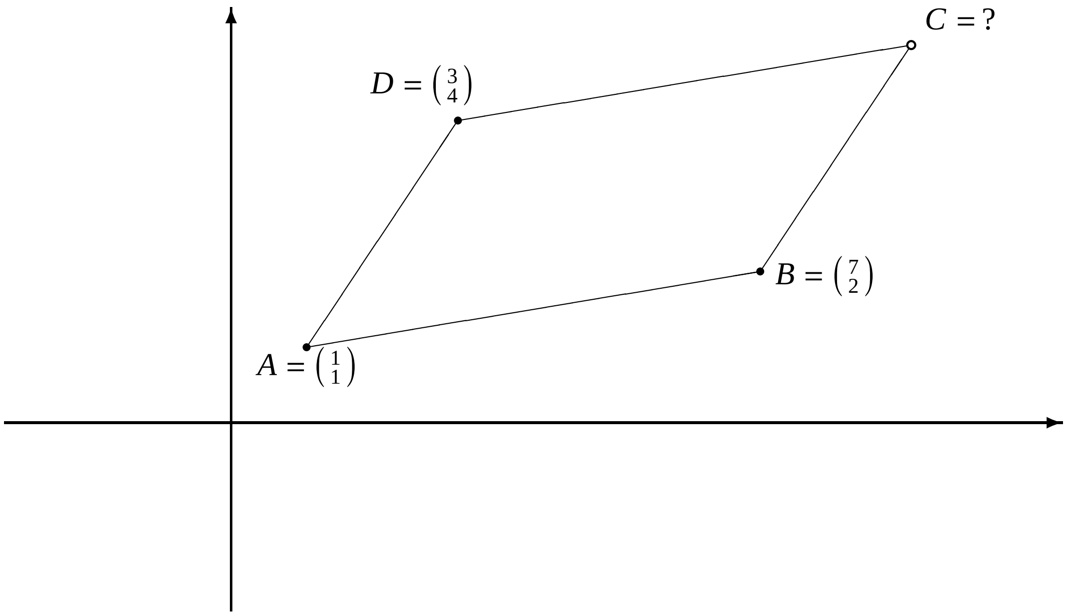



## 2. Vektoren

Wo liegt der Punkt $C$ im Parallelogramm $ABCD$, dessen Punkte $A$, $B$ und $D$ durch

$$A = \binom{1}{1},\ B = \binom{7}{2},\ D = \binom{3}{4}$$

gegeben sind?




Um von $A$ nach $B$ zu kommen, müssen wir 6 nach rechts und 1 nach oben gehen.

$$\vec{AB} = \binom{7}{2} - \binom{1}{1} = \binom{6}{1}.$$

Wenn wir von $D$ starten und um 6 nach rechts und 1 nach oben gehen, landen wir bei $C$.

$$C = D + \vec{AB} = \binom{3}{4} + \binom{6}{1} = \binom{9}{5}.$$

Wir bemerken, dass wir ein Paar reeller Zahlen, wie etwa $\binom{6}{1}$, verwenden, um zwei verschiedene Dinge zu beschreiben:

- Den Punkt, der um 6 Längeneinheiten rechts und um 1 Längeneinheit über dem Punkt $\binom{0}{0}$ liegt.
- Den Weg (*Vektor*) von $\binom{1}{1}$ nach $\binom{7}{2}$.

In Mathematica werden Vektoren als Listen dargestellt.

```
In[1]:= a = {1, 1};

In[2]:= b = {7, 2};

In[3]:= d = {3, 4};

In[4]:= ab = b - a
Out[4]= {6, 1}

In[5]:= c = d + ab
Out[5]= {9, 5}
```

## 3. Die Länge eines Vektors

Wir lösen folgendes Beispiel:

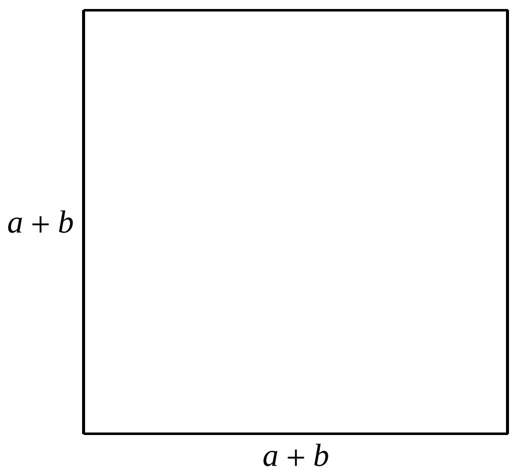
> Herr $A$ geht von $\binom{3}{2}$ aus 1 Einheit in Richtung Südosten. Wo landet er?

"Richtung Südosten" heißt "in Richtung $\binom{1}{-1}$". Allerdings hat $\binom{1}{-1}$ die Länge $\sqrt{2} \approx 1.41421$. Daher hat $v = \frac{1}{\sqrt{2}} \cdot \binom{1}{-1}$ die Länge 1 und zeigt auch in Richtung Südosten. Herr $A$ landet also im Punkt $Z$, den wir uns mit

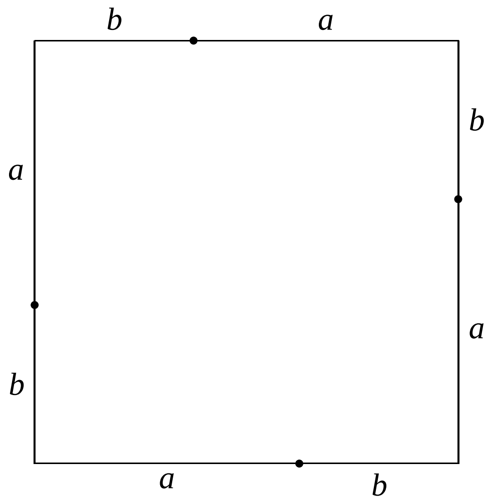
$$Z = \binom{3}{2} + \frac{1}{\sqrt{2}} \cdot \binom{1}{-1} \approx \binom{3.7}{1.3}$$

ausrechnen.

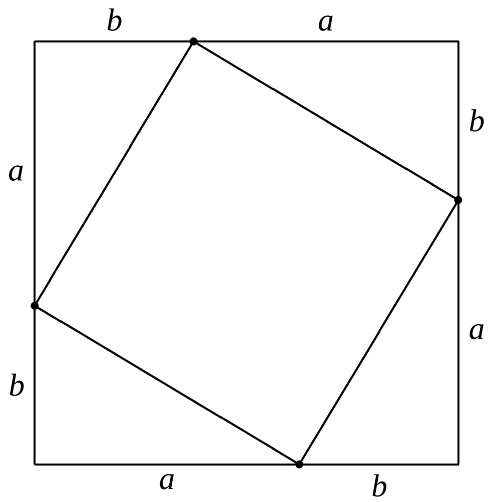
Wir überlegen uns jetzt, wie lange der Vektor $\binom{a}{b}$ ist. Das heißt, wir wollen wissen, wie lange in einem Dreieck, in dem die Seiten mit den Längen $a$ und $b$ einen rechten Winkel einschließen, die dem rechten Winkel gegenüberliegende Seite ist. Vergessen wir kurz unsere klassische Bildung, und zeichnen wir ein Quadrat mit Seitenlänge $a + b$.



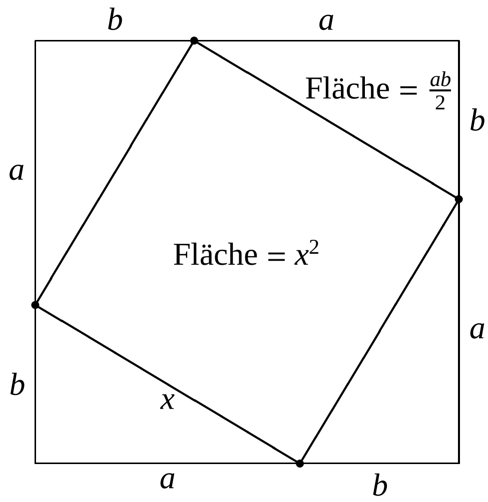

Jetzt unterteilen wir jede der vier Quadratseiten in ein Stück der Länge $a$ und ein Stück der Länge $b$.




Wir verbinden die vier Teilungspunkte.




Das innere jetzt eingezeichnete Viereck ist ein Quadrat. Das kann man so begründen: wenn man die ganze Zeichnung um $90^\circ$ gegen den Uhrzeigersinn dreht, kommt das innere Viereck auf sich selbst zu liegen: daher sind alle vier Winkel des inneren Vierecks gleich groß. In jedem Dreieck ist die Winkelsumme $180^\circ$, und daher ist in jedem Viereck die Winkelsumme $360^\circ$. Also ist jeder Winkel des inneren Vierecks gleich $\frac{360^\circ}{4} = 90^\circ$. Sei $x$ die Länge des Vektors $\binom{a}{b}$. Dann hat das innere Viereck die Fläche $x^2$. Jedes der vier rechtwinkeligen Dreiecke hat die Fläche $\frac{a\,b}{2}$.




Das innere Viereck und die vier rechtwinkeligen Dreiecke ergeben zusammen die Fläche des großen Quadrats mit der Seitenlänge $a + b$, also gilt

$$x^2 + 4\frac{a\,b}{2} = (a + b)^2.$$

Das heißt

$$x^2 + 2\,a\,b = a^2 + 2\,a\,b + b^2,$$

also

$$x^2 = a^2 + b^2.$$

Mit diesem Zusammenhang, dem Satz des Pythagoras (Pythagoras von Samos, 6. Jh. v. Chr), können wir die Länge $x$ des Vektors $\binom{a}{b}$ ausrechnen.

Wir kürzen die Länge des Vektors $\binom{a}{b}$ mit $\|\binom{a}{b}\|$ ab. Es gilt dann

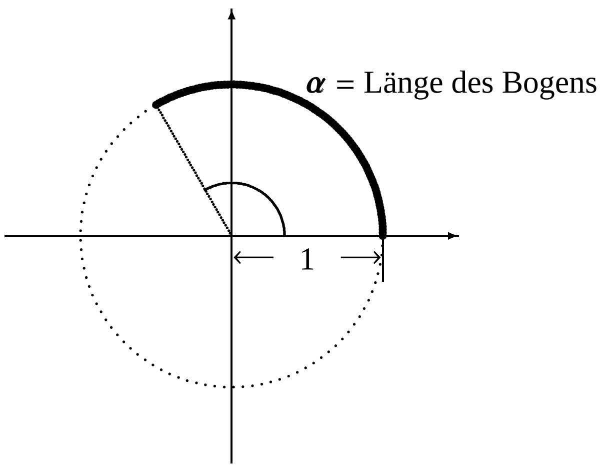
$$\left\|\binom{a}{b}\right\| = \sqrt{a^2 + b^2}.$$

## 4. Trigonometrie

In der *Trigonometrie* geht es darum, wie man – rechnerisch – aus den gegebenen Seitenlängen und Winkeln eines Dreiecks die restlichen Seitenlängen und Winkel bestimmen kann. Wenn man etwa von einem Dreieck die Längen der drei Seiten kennt, dann ist das Dreieck dadurch eindeutig bestimmt: die Winkel des Dreiecks sind also durch die Längen der drei Seiten festgelegt. (Wie konstruiert man ein Dreieck, das durch die drei Seitenlängen gegeben ist?) Ebenso ist ein Dreick dadurch bestimmt, dass man eine Seite und die beiden daran anliegenden Winkel kennt. (Wie konstruiert man dieses Dreieck?) Uns geht es jetzt darum, die fehlenden Seitenlängen und Winkel auszurechnen. Dabei geht man so vor:

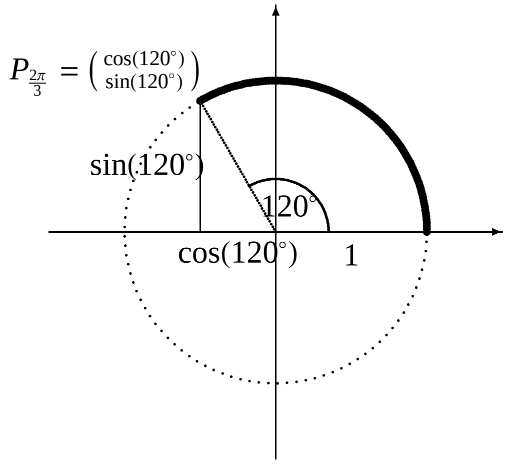
(1) Man tabelliert den Zusammenhang zwischen den Seitenlängen und den Winkeln für rechtwinkelige Dreiecke. Dazu braucht man die *Winkelfunktionen* sin (Sinus) und cos (Cosinus).
(2) Man baut sich alle anderen Dreiecke aus rechtwinkeligen Dreiecken zusammen. Da dieses Zusammenbauen aber immer gleich funktioniert, macht man es einmal für alle Dreiecke. Man gewinnt so zwei Zusammenhänge zwischen Seitenlängen und Winkeln eines Dreiecks: den *Cosinussatz* und den *Sinussatz*. Diese beiden Sätze reichen aus, um alle trigonometrischen Probleme zu lösen.

**4.1. Winkel.** Winkel misst man nicht nur in Grad ($\circ$), sondern auch in *Radiant* (rad). Dabei wird der Winkel durch die Länge des zugehörigen Bogens am Einheitskreis, dem Kreis mit Radius 1, angegeben.




Dabei entsprechen 180° dem Winkel $\pi$ rad. Demzufolge ist $1^\circ = \frac{\pi}{180}$ rad , und 1 rad $\approx$ 57.2958°.

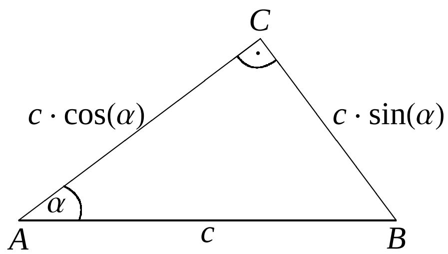
**4.2. Winkelfunktionen.**




Gegeben ist ein Winkel $x$. Der auf dem Kreis mit Mittelpunkt $\binom{0}{0}$ und Radius 1 liegende Punkt $P_x$ hat dann die Koordinaten $\binom{\cos(x)}{\sin(x)}$. Nach dem Satz des Pythagoras gilt für jeden Winkel $x$:

$$(\sin(x))^2 + (\cos(x))^2 = 1.$$

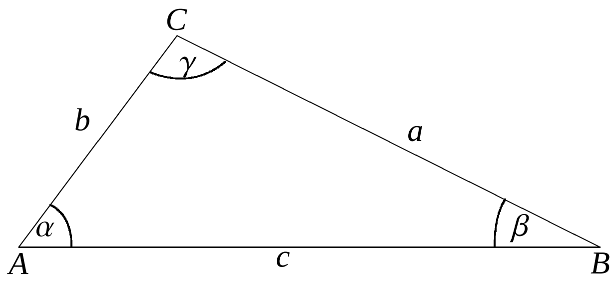
In einem rechtwinkeligen Dreieck heißt die dem rechten Winkel gegenüberliegende Seite auch *Hypotenuse*, die beiden dem rechten Winkel anliegenden Seiten heißen *Katheten*.




ÜBUNGSAUFGABEN 1.1.

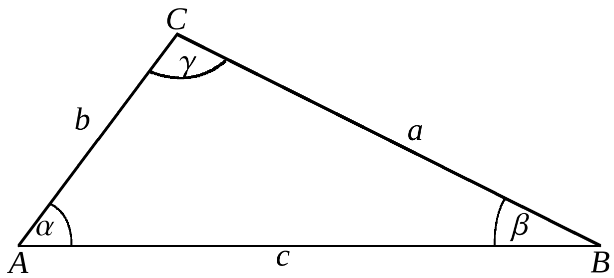
(1) Ein Kletterer kann Wände mit einer Neigung von maximal 65° besteigen. Schafft er eine Pyramide mit einer quadratischen Grundfläche von 784 $m^2$ und einer Höhe von 40 m?

**4.3. Zusammenhang zwischen Seitenlängen und Winkeln eines Dreiecks.** Wir sagen, dass drei Punkte ein *Dreieck* bilden, wenn sie nicht alle drei auf einer Geraden liegen. In einem Dreieck bezeichnet man oft die Längen der Seiten mit $a$, $b$, $c$, und den der Seite mit Länge $a$ gegenüber liegenden Winkel mit $\alpha$, den der Seite mit Länge $b$ gegenüber liegenden Winkel mit $\beta$, und den der Seite mit Länge $c$ gegenüber liegenden Winkel mit $\gamma$. Die Seiten sind üblicherweise *gegen den Uhrzeigersinn* mit $a$, $b$, $c$ beschriftet.




Der Cosinussatz löst folgendes Problem:

- Gegeben: Seitenlängen $a, b$ eines Dreiecks und der eingeschlossene Winkel $\gamma$.
- Gesucht: Die fehlende Seitenlänge $c$.




Wir betrachten zunächst den Fall $\gamma \leq 90^\circ$, $\alpha \leq 90^\circ$. Wir zeichnen in ein solches Dreieck die Höhe auf $b$ und erhalten aus dem Satz des Pythagoras:

$$c^2 = (b - a\cos(\gamma))^2 + (a\sin(\gamma))^2,$$

also

$$\begin{aligned}
c^2 &= b^2 - 2\,a\,b\cos(\gamma) + a^2(\cos(\gamma))^2 + a^2(\sin(\gamma))^2\\
c^2 &= b^2 - 2\,a\,b\cos(\gamma) + a^2\left((\cos(\gamma))^2 + (\sin(\gamma))^2\right)\\
c^2 &= b^2 - 2\,a\,b\cos(\gamma) + 1\,a^2\\
c^2 &= a^2 + b^2 - 2\,a\,b\,cos(\gamma).
\end{aligned}$$

Für den Fall $\gamma \leq 90^\circ$ und $\alpha > 90^\circ$ zeichnen wir die Höhe auf $a$ und erhalten:

$$\begin{aligned}
c^2 &= (a - b\cos(\gamma))^2 + (b\sin(\gamma))^2\\
c^2 &= a^2 - 2\,b\,a\cos(\gamma) + b^2(\cos(\gamma))^2 + b^2(\sin(\gamma))^2\\
c^2 &= a^2 - 2\,b\,a\cos(\gamma) + b^2\left((\cos(\gamma))^2 + (\sin(\gamma))^2\right)\\
c^2 &= a^2 - 2\,b\,a\;\cos(\gamma) + 1\,b^2\\
c^2 &= a^2 + b^2 - 2\,a\,b\;\cos(\gamma).
\end{aligned}$$

Zuletzt betrachten wir den Fall $\gamma > 90^\circ$. Wir zeichnen in ein solches Dreieck die Höhe auf $b$ und erhalten:

$$\begin{aligned}
c^2 &= (b + a\cos(180^\circ - \gamma))^2 + (a\sin(180^\circ - \gamma))^2 \\
c^2 &= (b - a\cos(\gamma))^2 + (a\sin(\gamma))^2 \\
c^2 &= b^2 - 2\,a\,b\cos(\gamma) + a^2(\cos(\gamma))^2 + a^2(\sin(\gamma))^2 \\
c^2 &= b^2 - 2\,a\,b\cos(\gamma) + a^2\,((\cos(\gamma))^2 + (\sin(\gamma))^2) \\
c^2 &= b^2 - 2\,a\,b\cos(\gamma) + 1\,a^2 \\
c^2 &= a^2 + b^2 - 2\,a\,b\,\cos(\gamma).
\end{aligned}$$

Insgesamt haben wir folgenden Satz bewiesen:

SATZ 1.2 (Cosinussatz). *Wir bezeichnen die Längen der Seiten eines Dreiecks mit $a$, $b$, $c$, und wir bezeichnen den der Seite mit Länge $c$ gegenüber liegenden Winkel mit $\gamma$. Dann gilt*

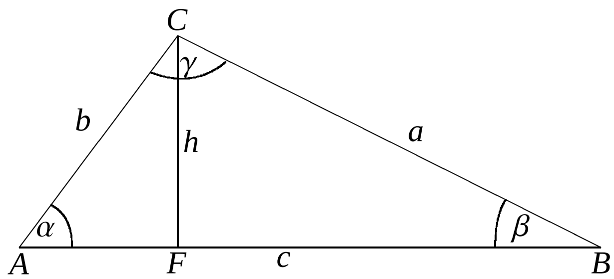
$$c^2 = a^2 + b^2 - 2\,a\,b\cos(\gamma).$$

Man findet mit dem Cosinussatz $\gamma$, wenn $a, b$ und $c$ gegeben sind. Zu jedem $y \in [-1, 1]$ gibt es genau ein $x \in [0, \pi]$, sodass $\cos(x) = y$.

### 4.4. Der Sinussatz.

Der Sinussatz löst folgendes Problem:

- Gegeben: $\alpha, \beta, a$.
- Gesucht: $b$.




Wir betrachten das rechtwinklige Dreieck $AFC$ und erhalten

$$h = b\sin(\alpha).$$

Mit dem Dreieck $FBC$ finden wir

$$h = a\sin(\beta).$$

Es gilt also $b\sin(\alpha) = a\sin(\beta)$. Sowohl $\sin(\alpha)$ also auch $\sin(\beta)$ sind ungleich 0, da in einem Dreieck kein Winkel $0^\circ$ oder $180^\circ$ sein kann. (Wir haben Dreiecke so definiert, dass die drei Punkte nicht auf einer Geraden liegen.) Es gilt also

$$\frac{a}{\sin(\alpha)} = \frac{b}{\sin(\beta)}.$$

Wenn man die gleiche Überlegung mit der Höhe auf $a$ macht, erhält man $\frac{b}{\sin(\beta)} = \frac{c}{\sin(\gamma)}$.

SATZ 1.3 (Sinussatz). *Wir bezeichnen die Längen der Seiten eines Dreiecks mit $a$, $b$, $c$, den der Seite mit Länge $a$ gegenüber liegenden Winkel mit $\alpha$, den der Seite mit Länge $b$ gegenüber liegenden Winkel mit $\beta$, und den der Seite mit Länge $c$ gegenüber liegenden Winkel mit $\gamma$. Sei $d$ der Durchmesser des Umkreises des Dreiecks. Dann gilt:*

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma} = d.$$

Wir überlegen uns jetzt, wie man Sinus- und Cosinussatz benutzen kann, um die fehlenden Winkel und Seitenlängen eines Dreiecks zu berechnen. In den Dreiecken der folgenden Beispiele bezeichnen wir die Längen der Seiten mit $a$, $b$, $c$, und wir bezeichnen den der Seite mit Länge $a$ gegenüber liegenden Winkel mit $\alpha$, den der Seite mit Länge $b$ gegenüber liegenden Winkel mit $\beta$, und den der Seite mit Länge $c$ gegenüber liegenden Winkel mit $\gamma$. Die Seiten seien dabei *gegen den Uhrzeigersinn* mit $a$, $b$, $c$ beschriftet.

(1) *Es sind die Seitenlängen $a$, $b$, $c$ gegeben:* Es gibt so ein Dreieck, wenn $a < b + c$, $b < a + c$, und $c < a + b$. Der Winkel $\alpha$ ist dann durch
$$a^2 = b^2 + c^2 - 2\,b\,c\cos(\alpha)$$
eindeutig bestimmt.

(2) *Es sind eine Seitenlänge und zwei Winkel gegeben:* Da die Winkelsumme $180^\circ$ ist, kennt man tatsächlich alle drei Winkel. Sind also $c$, $\alpha$, und $\beta$ gegeben, so gibt es so ein Dreieck, wenn $\alpha + \beta < \pi$. Man berechnet $\gamma = \pi - \alpha - \beta$, und dann $a$ und $b$ mithilfe des Sinussatzes.

(3) *Es sind zwei Seitenlängen und der eingeschlossene Winkel gegeben:* Es sind also zum Beispiel $b$, $c$, und $\alpha$ gegeben. Wenn $b$ und $c$ positive reelle Zahlen sind, und $0 < \alpha < \pi$ gilt, so gibt es sicher ein solches Dreieck. Man berechnet dann $a$ mithilfe des Cosinussatzes; dann kennt man alle drei Seitenlängen und kann mit dem Cosinussatz die verbleibenden Winkel ausrechnen.

(4) *Es sind zwei Seitenlängen gegeben, und ein Winkel, der nicht der von diesen Seiten eingeschlossene Winkel ist:* Es sind also zum Beispiel $c$, $\alpha$ und $a$ gegeben. In diesem Fall kann es sein, dass es gar kein, genau ein oder genau zwei Dreiecke mit dem vorgegebenen $c$, $\alpha$ und $a$ gibt. Es gibt mehrere Fälle:
   (a) *Es gilt $\alpha < 90^\circ$:*
      (i) *Es gilt $a \geq c$:* Es gibt genau ein Dreieck. Wir erhalten $b$ als die einzige positive Lösung von
      $$c^2 + x^2 - 2\,c\,x\cos(\alpha) = a^2.$$
      Es gilt also
      $$b = c\cos(\alpha) + \sqrt{(c\cos(\alpha))^2 + a^2 - c^2}.$$

(ii) *Es gilt* $c \sin(\alpha) < a < c$*:* Es gibt genau zwei Dreiecke. Wir erhalten die beiden Möglichkeiten für $b$ als die Lösungen der Gleichung

$$c^2 + x^2 - 2\,c\,x\cos(\alpha) = a^2,$$

also

$$b_{1,2} = c\cos(\alpha) \pm \sqrt{a^2 - (c\sin(\alpha))^2}.$$

(iii) *Es gilt* $a = c\sin(\alpha)$*:* Es gibt genau ein Dreieck. Wir erhalten $\gamma = \frac{\pi}{2}$ und $b = c\cos(\alpha)$.

(iv) *Es gilt* $a < c\sin(\alpha)$*:* Es gibt kein Dreieck.

(b) *Es gilt* $\alpha = 90^\circ$*:*

(i) *Es gilt* $a > c$*:* Es gibt genau ein Dreieck, und

$$b^2 = a^2 - c^2$$

nach dem Satz des Pythagoras.

(ii) *Es gilt* $a \leq c$*:* Es gibt kein Dreieck.

(c) *Es gilt* $\alpha > 90^\circ$*:*

(i) *Es gilt* $a > c$*:* Es gibt ein Dreieck. Die Länge $b$ ist die einzige positive Lösung von

$$c^2 + x^2 - 2\,c\,x\cos(\alpha) = a^2,$$

also

$$b = c\cos(\alpha) + \sqrt{(c\cos(\alpha))^2 + a^2 - c^2}.$$

(ii) *Es gilt* $a \leq c$*:* Es gibt kein Dreieck. Tückischerweise kann die Gleichung

$$c^2 + x^2 - 2\,c\,x\cos(\alpha) = a^2$$

trotzdem positive Lösungen haben.

## ÜBUNGSAUFGABEN 1.4.

Wir bezeichnen die Längen der Seiten eines Dreiecks mit $a$, $b$, $c$, und wir bezeichnen den der Seite mit Länge $a$ gegenüber liegenden Winkel mit $\alpha$, den der Seite mit Länge $b$ gegenüber liegenden Winkel mit $\beta$, und den der Seite mit Länge $c$ gegenüber liegenden Winkel mit $\gamma$. (Die Seiten seien dabei *gegen den Uhrzeigersinn* mit $a$, $b$, $c$ beschriftet.)

(1) Berechnen Sie $\sin(\gamma)$ für ein Dreieck mit $c = 10$, $b = \frac{10}{\sqrt{3}}$, $\beta = 30^\circ$. Das Dreieck ist mit diesen drei Bestimmungsstücken $c, b, \beta$ noch nicht eindeutig festgelegt. Warum nicht?

(2) Wie groß kann $b$ in einem Dreieck mit $\alpha = 45^\circ$, $c = 1$, $a = \frac{2}{\sqrt{6}}$ sein? (Gibt es mehr als eine Lösung?)

(3) Geben Sie eine Wahl von $a$ an, sodass es genau ein Dreieck mit den Bestimmungsstücken $\alpha = 45^\circ$, $c = 1$, und Ihrem gewählten $a$ gibt!

(4) Von einem Dreieck $ABC$ haben Sie folgende Information: $\overline{AB} = 10\,\text{cm}$, der Winkel $\alpha$ zwischen $AB$ und $AC$ ist $30^\circ$, $\overline{BC} = \frac{10}{\sqrt{2}}$ cm.
   (a) Stellen Sie diese Daten in einer Skizze dar.
   (b) Welchen Winkel schließen $CB$ und $CA$ ein? Gibt es mehr als eine Lösung?

Berechnen Sie in den folgenden Beispielen jeweils die nicht angegebenen Seitenlängen und Winkel.

(5) (a) $c = 4$, $b = 5$, $a = 3$.
   (b) $c = 4$, $b = 5$, $a = 10$.

(6) (a) $c = 5$, $b = 3$, $\alpha = \frac{\pi}{4}$.
   (b) Gibt es für jede Wahl von $c > 0$, $b > 0$, $\alpha$ mit $0 < \alpha < \pi$ ein Dreieck mit den gewählten Bestimmungsstücken?

(7) (a) $c = 5, b = 10, \beta = \frac{\pi}{6}$.
(b) $c = 5, b = 3, \beta = \frac{\pi}{6}$.
(8) (a) $c = 5, b = \frac{5}{2}, \beta = \frac{\pi}{6}$.
(b) $c = 5, b = 2, \beta = \frac{\pi}{6}$.
(9) (a) $c = 5, b = 3, \beta = \frac{5\pi}{6}$.
(b) $c = 5, b = 10, \beta = \frac{5\pi}{6}$.
(c) $c = 4, b = 5, \beta = \frac{\pi}{2}$.

(10) Fassen Sie Ihre Beobachtungen aus den letzten Beispielen zusammen: Unter welchen Voraussetzungen an $c$, $b$, $\beta$ gibt es
(a) gar kein Dreieck
(b) genau ein Dreieck
(c) genau zwei Dreiecke
(d) mehr als zwei Dreiecke
mit den Bestimmungsstücken $c > 0, b > 0, \beta \in ]0, \pi[$?

(11) (a) $c = 5, \alpha = \frac{3\pi}{4}, \beta = \frac{\pi}{8}$.
(b) $c = 5, \alpha = \frac{3\pi}{4}, \beta = \frac{\pi}{4}$.
(c) $c = 5, \alpha = \frac{3\pi}{4}, \beta = \frac{\pi}{2}$.

(12) Fassen Sie Ihre Beobachtung aus dem letzten Beispiel zusammen: Unter welchen Voraussetzungen an $c, \alpha, \beta$ gibt es
(a) gar kein Dreieck
(b) genau ein Dreieck
(c) genau zwei Dreiecke
(d) mehr als zwei Dreiecke
mit den Bestimmungsstücken $c, \alpha, \beta$?

(13) Bestimmen Sie $b$ für alle Dreiecke mit $c = 4$, $\alpha = 60^\circ$, $a = \sqrt{13}$.

(14) Bestimmen Sie ein $a$, sodass es kein Dreieck mit $c = 4$, $\alpha = 60^\circ$ und Ihrem gewählten $a$ gibt.

(15) Bestimmen Sie ein $a$, sodass es *genau ein* Dreieck mit $c = 4$, $\alpha = 60^\circ$ und Ihrem gewählten $a$ gibt.

(16) Bestimmen Sie $c$ für alle Dreiecke mit $a = 2, \beta = 45^\circ$, $b = \frac{4}{\sqrt{6}}$.

(17) Bestimmen Sie ein $b$, sodass es kein Dreieck mit $a = 2, \beta = 45^\circ$ und Ihrem gewählten $b$ gibt.

(18) Bestimmen Sie ein $b$, sodass es *genau ein* Dreieck mit $a = 2, \beta = 45^\circ$ und Ihrem gewählten $b$ gibt.

(19) Sie möchten die Entfernung eines Punktes $B$ an einem Ufer eines Flusses zu einem Punkt $C$ auf der anderen Seite des Flusses bestimmen. Dazu gehen Sie folgendermaßen vor: Sie messen an Ihrem Flussufer die Strecke von $B$ zu einem weiteren Punkt $A$ ab. Diese Strecke ist 500 m lang. Der Winkel zwischen $BC$ und $BA$ ist $60^\circ$, der Winkel zwischen $AB$ und $AC$ ist $30^\circ$.
(a) Stellen Sie diese Daten in einer Skizze dar.
(b) Wie lang ist die Strecke $BC$?
(c) Um die Breite des Flusses zu bestimmen, wollen Sie wissen, wie weit $C$ von der Strecke $AB$ entfernt liegt. Bestimmen Sie dazu den Normalabstand von $C$ auf die Gerade durch $A$ und $B$.

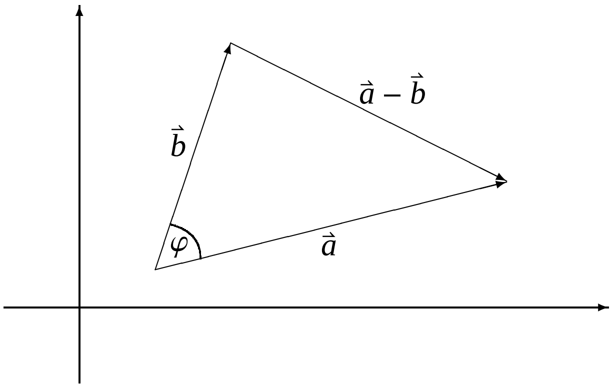
(20) Sie glauben dem italienischen Tourismusverband nicht und wollen selbst herausfinden, wie schief der Turm von Pisa ist. Dazu entfernen Sie sich in Neigerichtung des Turms 50 Meter vom Fußpunkt des Turms und blicken (vom Boden aus, damit Sie es später beim Rechnen einfacher haben) zur Turmspitze, wleche nun unter einem Winkel $\alpha$ erscheint. Sie stellen fest, dass $\alpha$ genau $47^\circ 12' 53''$ beträgt, und dass der Turm 45 m lang ist. Um wieviel Grad ist der Turm gegen die Vertikale geneigt?

(21) Sie möchten die Entfernung eines Punktes $A$ an einem Ufer eines Flusses zu einem Punkt $B$ auf der anderen Seite bestimmen. Dazu können Sie folgendermaßen vorgehen: Messen Sie an Ihrem Flussufer die Strecke $A$ zu einem Punkt $C$ ab, und bestimmen Sie dann mit Hilfe eines Kompasses den Winkel $\alpha$ zwischen der Strecke $AB$ und $AC$, sowie den Winkel $\gamma$ zwischen $AC$ und $CB$. Was ergibt sich für die Entfernung $AB$ bei $AC = 27.5$ m und $\alpha = 73^\circ$, $\gamma = 65^\circ$?

(22) Sie verlassen eine gerade Straße, die die beiden Orte A und B verbindet, 20 km bevor Sie B erreichen und gehen geradeaus, bis es Ihnen nach 10 km keinen Spaß mehr macht. Dann drehen Sie sich um $30^o$ nach rechts und erblicken nun den Ort B gerade vor sich. Wie weit müssen Sie jetzt noch wandern, um nach B zu gelangen, wenn Sie jetzt den geraden Weg nach B einschagen?

## 5. Der Winkel zwischen zwei Vektoren

Wir berechnen den Winkel, den die Vektoren $\vec{a} = \binom{a_1}{a_2}$ und $\vec{b} = \binom{b_1}{b_2}$ miteinander einschließen. Dazu nehmen wir an, dass keiner der beiden Vektoren der Nullvektor ist.




Für den eingeschlossenen Winkel $\varphi$ gilt nach dem Cosinussatz:

$$\|b - a\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\,\|a\|\,\|b\|\cos(\varphi).$$

Diese Formel können wir vereinfachen:

$$\begin{aligned}
2 \cdot \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos(\varphi) &= -(\|\vec{b} - \vec{a}\|^2) + \|\vec{a}\|^2 + \|\vec{b}\|^2 \\
2 \cdot \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos(\varphi) &= -((b_1 - a_1)^2 + (b_2 - a_2)^2) + (a_1^2 + a_2^2) + (b_1^2 + b_2^2) \\
2 \cdot \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos(\varphi) &= -(b_1^2 - 2\,a_1\,b_1 + a_1^2) - (b_2^2 - 2\,a_2\,b_2 + a_2^2) + a_1^2 + a_2^2 + b_1^2 + b_2^2 \\
2 \cdot \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos(\varphi) &= 2\,a_1\,b_1 + 2\,a_2\,b_2 \\
\|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos(\varphi) &= a_1\,b_1 + a_2\,b_2.
\end{aligned}$$

Wir erhalten

$$\cos(\varphi) = \frac{a_1\,b_1 + a_2\,b_2}{\|\vec{a}\| \cdot \|\vec{b}\|}.$$

Die Zahl $a_1\,b_1 + a_2\,b_2$ bezeichnet man als das *Skalarprodukt* von $\vec{a}$ und $\vec{b}$, und man kürzt es mit $\langle \vec{a}, \vec{b} \rangle$ ab.

$$\langle \vec{a}, \vec{b} \rangle = \langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \rangle = a_1 \cdot b_1 + a_2 \cdot b_2.$$

Die Winkelformel heißt jetzt:

$$\cos(\varphi) = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \cdot \|\vec{b}\|}.$$

Außerdem gilt für jeden Vektor $\vec{a}$

$$\langle \vec{a}, \vec{a} \rangle = (\|a\|)^2.$$

```
In[6]:= {3, 4}.{-5, 2}
Out[6]= -7

In[7]:= Laenge[v_] := Sqrt[v.v]

In[8]:= Laenge[{2, 3}]
Out[8]= √13
```

Unter Verwendung der Mathematica-Function `Norm` kann man diese Länge auch mit `Norm[{2,3}]` ausrechnen.

DEFINITION 1.5. Seien $\vec{a}, \vec{b} \in \mathbb{R}^2$. Die Vektoren $\vec{a}$ und $\vec{b}$ *sind aufeinander normal*, wenn $\langle \vec{a}, \vec{b} \rangle = 0$.

Zwei Vektoren sind also aufeinander normal, wenn einer von ihnen der Nullvektor ist, oder wenn sie einen rechten Winkel einschließen. Damit erhält man, dass (wenn $\binom{a}{b} \neq \binom{0}{0}$) der Vektor $\binom{a}{b}$ mit den Vektoren $\binom{b}{-a}$ und $\binom{-b}{a}$ einen rechten Winkel einschließt.

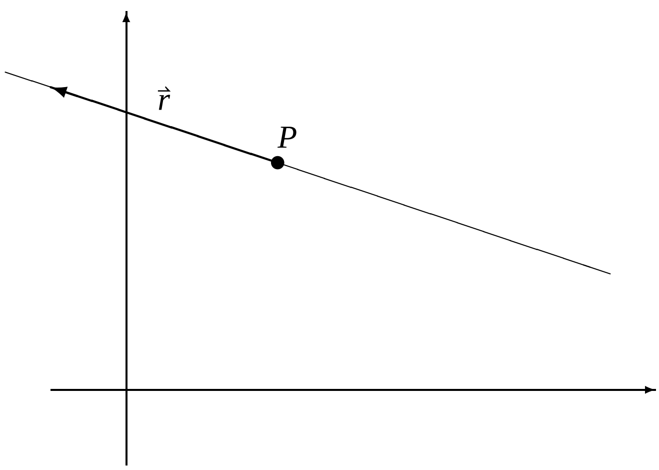
ÜBUNGSAUFGABEN 1.6.

(1) Von einem gleichschenkeligen Dreieck sind zwei Basiseckpunkte $\binom{-2}{3}$ und $\binom{4}{11}$ bekannt. Ergänzen Sie diese Punkte mit einer Spitze, sodaß das entstehende Dreieck die Höhe 5 besitzt. Wie viele verschiedene Lösungen gibt es? (Sie brauchen nur eine wirklich auszurechnen.)
(2) Berechnen Sie den Cosinus des Winkels zwischen $x$ und $y$ für $x = \binom{3}{2}$, $y = \binom{-6}{-4}$.
(3) Berechnen Sie jeweils den Winkel zwischen folgenden beiden Vektoren. Geben Sie die Ergebnisse in Grad und in Radiant an!
   (a) $a = \binom{3}{2}$, $b = \binom{-3}{2}$.
   (b) $a = \binom{3}{4}$, $b = \binom{4}{-3}$.
   (c) $a = \binom{3}{4}$, $b = \binom{6}{8}$.
   (d) $a = \binom{3}{4}$, $b = \binom{-9}{-12}$.
(4) Zeigen Sie, dass das Skalarprodukt im $\mathbb{R}^2$ folgende Eigenschaften erfüllt:
   (a) $\langle a + b, a + b \rangle = \langle a, a \rangle + 2\langle a, b \rangle + \langle b, b \rangle$.
   (b) $\langle a + b, a - b \rangle = \langle a, a \rangle - \langle b, b \rangle$.
(5) Verwenden Sie das Skalarprodukt, um folgenden geometrischen Satz zu beweisen.
   In einem Parallelogramm mit Seitenlängen $a$, $b$, und Diagonalenlängen $e$, $f$ gilt:
$$2\,(a^2 + b^2) = e^2 + f^2.$$

## 6. Geraden in der Ebene

Wir überlegen uns, wie man Geraden in der Ebene beschreiben kann.

### 6.1. Geraden, die durch einen Punkt und eine Richtung gegeben sind.



$P = \binom{2}{3}$, $\vec{r} = \binom{-3}{1}$.

Die Gerade $g$ ist die Menge aller Punkte, die man erreicht, indem man von $P$ ein Stück in Richtung $\vec{r}$ geht.

$$g = \{P + \lambda \cdot \vec{r} \mid \lambda \in \mathbb{R}\}.$$

Lies: “g ist gleich der Menge aller Punkte $P + \lambda$ mal $\vec{r}$, wobei $\lambda$ eine reelle Zahl ist.” Mit den Zahlen für $P$ und $\vec{r}$:

$$g = \{\binom{2}{3} + \lambda \cdot \binom{-3}{1} \mid \lambda \in \mathbb{R}\},$$

oder, anders geschrieben,

$$g = \{\binom{2-3\lambda}{3+\lambda} \mid \lambda \in \mathbb{R}\}.$$

Man kann $g$ auch so schreiben:

$$g = \{\binom{x}{y} \in \mathbb{R}^2 \mid \text{ es gibt } \lambda \in \mathbb{R}, \text{ sodass } \binom{x}{y} = \binom{2}{3} + \lambda \cdot \binom{-3}{1}\}.$$

Lies: “g ist gleich der Menge aller Punkte $\binom{x}{y}$ in $\mathbb{R}$ hoch 2, für die es ein $\lambda$ in den reellen Zahlen gibt, sodass $\binom{x}{y}$ gleich $\binom{2}{3} + \lambda$ mal $\binom{-3}{1}$ ist.” Diese Darstellung von $g$ durch *Punkt* und *Richtungsvektor* heißt *Parameterdarstellung der Geraden g.* Man schreibt oft kurz:

$$g : \binom{x}{y} = \binom{2}{3} + \lambda \cdot \binom{-3}{1}$$

oder

$$g : X = \binom{2}{3} + \lambda \cdot \binom{-3}{1}.$$

Wir überprüfen, ob der Punkt $\binom{-4}{5}$ auf der Geraden $g$ liegt. Er liegt auf $g$, falls es eine reelle Zahl $\lambda$ gibt, sodass $\binom{-4}{5}$ gleich $\binom{2}{3} + \lambda \cdot \binom{-3}{1}$ ist. Wir suchen also ein $\lambda \in \mathbb{R}$, das die Gleichungen

$$\begin{aligned} -4 &= 2 - 3\lambda \qquad \text{I} \\ 5 &= 3 + 1\lambda \qquad \text{II} \end{aligned}$$

erfüllt. Aus der Gleichung I erhalten wir $\lambda = 2$; da auch $5 = 3 + 1 \cdot 2$ gilt, ist $\lambda = 2$ eine Lösung des Gleichungssystems. Daher liegt der Punkt $\binom{-4}{5}$ auf $g$; wir schreiben dafür

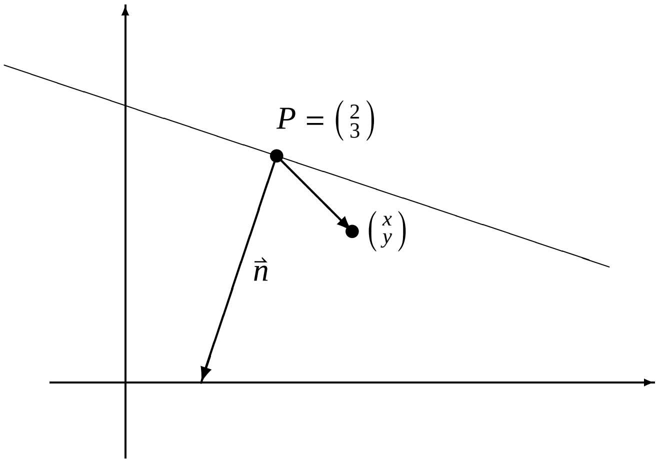
$$\binom{-4}{5} \in g.$$

**6.2. Geraden, die durch eine Gleichung gegeben sind.** Wir haben im letzten Beispiel überprüft, ob ein Punkt auf einer Geraden liegt. Dabei war die Gerade in Parameterform gegeben. Zur Überprüfung war es notwendig, festzustellen, ob es einen Wert für den Parameter $\lambda$ gibt, der uns genau den getesteten Punkt liefert. Wir mussten also für jeden Punkt ein Gleichungssystem (mit zwei Gleichungen und einer Variable) lösen.

Wir testen nun wieder, ob $\binom{x}{y}$ auf der Geraden $g$ liegt, die durch

$$g : X = \binom{2}{3} + \lambda \cdot \binom{-3}{1}$$

gegeben ist.




Anstatt zu fragen, ob $\binom{x}{y}$ auf $g$ liegt, fragen wir, ob $\binom{x}{y} - \binom{2}{3}$ normal auf $\vec{n}$ ist. Das ist nämlich genau für die Punkte $\binom{x}{y}$ auf $g$ der Fall. Zunächst finden wir den Vektor $\vec{n}$. Auf den Vektor $\binom{a}{b}$ steht immer der Vektor $\binom{-b}{a}$ normal, denn das Skalarprodukt $\langle \binom{a}{b}, \binom{-b}{a} \rangle$ ergibt $-a\,b + a\,b = 0$. Also finden wir $\vec{n}$ durch

$$\vec{n} = \binom{-1}{-3}.$$

Nun überprüfen wir, ob $\binom{x}{y} - \binom{2}{3}$ normal auf $\binom{-1}{-3}$ steht. Das gilt genau dann, wenn

$$\langle \binom{x-2}{y-3}, \binom{-1}{-3} \rangle = 0.$$

Wir rechnen das Skalarprodukt aus und erhalten

$$-x - 3\,y + 11 = 0.$$

Ein Punkt $\binom{x}{y}$ liegt also genau dann auf der Geraden, wenn $-x - 3\,y + 11 = 0$ ist. Wir können also jetzt viel einfacher überprüfen, ob ein Punkt auf der Geraden $g$ liegt. Wir berechnen $-x - 3\,y + 11$. Ist das 0, so liegt der Punkt auf der Geraden, und sonst nicht. Außerdem können wir die Gerade jetzt kürzer angeben durch

$$g = \{\binom{x}{y} \in \mathbb{R}^2 \mid -x - 3\,y = -11\}$$

(lies: “g ist gleich der Menge aller $\binom{x}{y}$ in $\mathbb{R}$ hoch 2, für die $-x - 3\,y$ gleich $-11$ ist.”) Das kürzt man auch zu

$$g : -x - 3\,y = -11$$

ab. $-x - 3\,y = -11$ heißt *Gleichung* der Geraden, diese Darstellung der Geraden *Gleichungsform* oder *implizite Darstellung* der Geraden.

### 6.3. Verwandlung zwischen Gleichungs- und Parameterform.

6.3.1. *Verwandlung von parametrisierter in implizite Darstellung.* Wir verwandeln $g : X = \binom{2}{3} + \lambda \cdot \binom{-3}{1}$ in $g : -x - 3\,y = -11$ so, wie das in obigem Beispiel erklärt worden ist.

6.3.2. *Verwandlung von impliziter in parametrisierte Form.* Wir verwandeln $g : 5x - 2y = 1$ in parametrisierte Form. Dazu setzen wir $y := t$ und rechnen uns aus diesem $y$ das $x$ aus. Wir erhalten $x = \frac{1}{5} + \frac{2}{5}t$. Somit ist für jedes $t \in \mathbb{R}$ der Punkt $\binom{\frac{1}{5}+\frac{2}{5}t}{t}$ ein Geradenpunkt. Die Gerade hat also die parametrisierte Darstellung

$$g : X = \begin{pmatrix} \frac{1}{5} \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} \frac{2}{5} \\ 1 \end{pmatrix}.$$

Eine andere Darstellung derselben Geraden ist

$$g : X = \begin{pmatrix} \frac{1}{5} \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix},$$

oder

$$g : X = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \cdot \begin{pmatrix} -22 \\ -55 \end{pmatrix}.$$

Spezialfall: Wir verwandeln $g : y = -1$ in Parameterform. Dazu setzen wir $x := t$, und rechnen uns dann das $y$ aus. Das ist aber immer $-1$. Für jedes $t \in \mathbb{R}$ ist also $\binom{t}{-1}$ ein Geradenpunkt. Die Gerade hat die parametrisierte Darstellung

$$g : X = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

## ÜBUNGSAUFGABEN 1.7.

(1) Geben Sie die Gerade durch die Punkte $\binom{3}{2}$ und $\binom{3}{-3}$ in Parameterform und in impliziter Form an!

(2) Bestimmen Sie jeweils eine Parameterform (= Punkt-Richtungs-Form) folgender Geraden.
  - (a) $3x + 4y = 17$.
  - (b) $x = 1$.
  - (c) $y = -4$.

(3) Bestimmen Sie eine Gleichung, deren Lösungsmenge die Gerade

$$X = \binom{-3}{6} + \lambda \cdot \binom{-1}{2}$$

ist.

(4) Bestimmen Sie die implizite Darstellung jener Geraden, die parallel zur Geraden $g$ mit

$$g : X = \binom{2}{-1} + t \cdot \binom{3}{-4}$$

sind und von dieser Abstand 10 haben.

(5) Ein Radfahrer startet im Punkt $\binom{2}{3}$ und fährt auf den Punkt $\binom{8}{0}$ zu. Ein Fußgänger startet im Punkt $\binom{4}{-3}$ und geht auf den Punkt $\binom{4}{10}$ zu. In welchem Punkt schneiden sich die Wege der beiden?

(6) Ein Radfahrer im Punkt $\binom{3}{3}$ und ein Fußgänger im Punkt $\binom{9}{-5}$ bewegen sich aufeinander zu, der Radfahrer mit 20 km/h, der Fußgänger mit 5 km/h. Wann und wo treffen die beiden einander?

(7) Vom Quadrat *ABCD* haben wir folgende Angaben:
  - $A = \binom{-1}{2}$.
  - $B$ liegt auf der Geraden
  $$g_B : X = \binom{-1}{2} + \lambda \cdot \binom{4}{3}.$$
  - Die Seitenlänge des Quadrats ist 10.
  - Die Eckpunkte sind gegen den Uhrzeigersinn mit $A, B, C, D$ beschriftet.

  Berechnen Sie die Koordinaten des Eckpunktes $C$!

(8) Vom Quadrat *ABCD* haben wir folgende Angaben:
  - $A = \binom{1}{2}$.
  - $B$ liegt auf der Geraden
  $$g_B : X = \binom{1}{2} + \lambda \cdot \binom{3}{-4}.$$
  - Die Seitenlänge des Quadrats ist 15.
  - Die Eckpunkte sind gegen den Uhrzeigersinn mit $A, B, C, D$ beschriftet.

  Berechnen Sie die Koordinaten des Eckpunktes $C$!

(9) Zeigen Sie, dass sich die Schwerlinien des Dreiecks *ABC* mit $A = \binom{-2}{-1}$, $B = \binom{5}{1}$ und $C = \binom{3}{5}$ in einem Punkt schneiden, und berechnen Sie diesen Schnittpunkt.

(10) Berechnen Sie den Umkreismittelpunkt $U = \binom{u_1}{u_2}$ des Dreiecks $ABC$ mit $A = \binom{-2}{-1}$, $B = \binom{5}{1}$ und $C = \binom{3}{5}$, indem Sie die Bedingung, dass $U$ gleich weit von $A$, $B$ und $C$ entfernt ist, in Gleichungen in den Variablen $u_1$ und $u_2$ umwandeln. Verwenden Sie zur Lösung der auftretenden Gleichungen den Mathematica-Befehl `Solve`.

(11) Bestimmen Sie die implizite Darstellung jener Geraden, die parallel zur Geraden $g$ mit

$$g : X = \binom{2}{0} + t \cdot \binom{3}{-4}$$

sind, und von dieser Abstand 10 haben.

(12) Berechnen Sie den Durchschnitt der Geraden $h$ und $j$, wobei

$$h : X = \binom{1}{2} + t \cdot \binom{2}{5}$$

und

$$j : 10x - 4y = 0.$$

(13) Bestimmen Sie die Schnittmenge der Geraden

$$g_1 : X = \binom{3}{4} + t \cdot \binom{-2}{1}$$

und

$$g_2 : 2x + 4y = 22.$$

(14) Bestimmen Sie den Cosinus des Schnittwinkels der Geraden

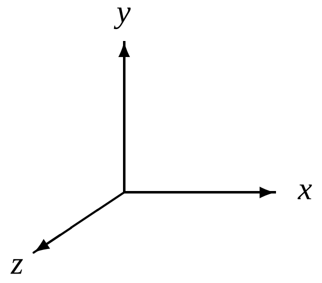
$$g_1 : X = \binom{0}{4} + t \cdot \binom{3}{4}$$

und

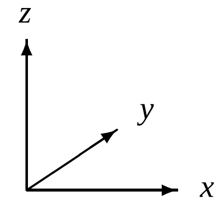
$$g_2 : 12x - 5y = 22.$$

## 7. Vektoren im $\mathbb{R}^n$

Bisher haben wir uns auf die Geometrie in der Ebene beschränkt. Man kann nun auch den Raum mit Tripeln reeller Zahlen, also mit Elementen aus $\mathbb{R}^3$, koordinatisieren. Die Konvention ist es, die Richtungen der Koordinatenachsen wie in folgenden Skizzen zu wählen:




oder



Hält man Daumen, Zeigefinger und Mittelfinger der rechten Hand so, dass sie paarweise im rechten Winkel aufeinander stehen, dann zeigen sie jeweils in die Richtung der positiven $x$-Achse, $y$-Achse und $z$-Achse.

Wir definieren die Operationen, die im $\mathbb{R}^2$ hilfreich waren, allgemein für

$$\mathbb{R}^n := \{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \ldots, x_n \in \mathbb{R}\},$$

wobei $n$ eine beliebige natürliche Zahl ist.

DEFINITION 1.8. Seien $\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ in $\mathbb{R}^n$. Wir definieren:

(1) $\vec{a} + \vec{b} := \begin{pmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{pmatrix}$.
(2) $\lambda\vec{a} := \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}$ für $\lambda \in \mathbb{R}$.
(3) $\langle \vec{a}, \vec{b} \rangle := a_1 \cdot b_1 + \cdots + a_n \cdot b_n = \sum_{i=1}^n a_i b_i$ (Skalarprodukt).
(4) Die *Länge* von $\vec{a}$ ist $\|\vec{a}\| := \sqrt{\langle \vec{a}, \vec{a} \rangle}$.

SATZ 1.9. *Sei $n \in \mathbb{N}$, seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$, und sei $\lambda \in \mathbb{R}$. Dann gilt:*

(1) $\langle \vec{a}, \vec{b} \rangle = \langle \vec{b}, \vec{a} \rangle$,
(2) $\langle \vec{a} + \vec{b}, \vec{c} \rangle = \langle \vec{a}, \vec{c} \rangle + \langle \vec{b}, \vec{c} \rangle$,
(3) $\langle \lambda\vec{a}, \vec{b} \rangle = \lambda \langle \vec{a}, \vec{b} \rangle$.
(4) $\langle \vec{a}, \vec{a} \rangle \geq 0$.
(5) *Wenn* $\langle \vec{a}, \vec{a} \rangle = 0$, *so gilt* $\vec{a} = 0$.

SATZ 1.10 (Projektionseigenschaft des Skalarprodukts). *Seien $\vec{a}, \vec{b} \in \mathbb{R}^n$ mit $\vec{b} \neq 0$, und sei $\vec{c} := \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \vec{b}$. Dann ist $\vec{a} - \vec{c}$ normal auf $\vec{b}$.*

*Beweis:* $\langle \vec{a} - \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \cdot \vec{b}, \vec{b} \rangle = \langle \vec{a}, \vec{b} \rangle - \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \langle \vec{b}, \vec{b} \rangle = 0.$ ■

Der folgende Satz ist als *Cauchy-Schwarz-Ungleichung* bekannt. Für $\vec{a}, \vec{b} \in \mathbb{R}^n$ sagen wir, dass *$\vec{a}$ ein Vielfaches von $\vec{b}$* ist, wenn es ein $\lambda \in \mathbb{R}$ gibt, sodass $\vec{a} = \lambda \vec{b}$.

SATZ 1.11 (Augustin Cauchy, 1821). *Sei $n \in \mathbb{N}$, und seien $\vec{a}, \vec{b} \in \mathbb{R}^n$. Dann gilt:*

(1) $|\langle \vec{a}, \vec{b} \rangle| \leq \|\vec{a}\| \cdot \|\vec{b}\|$.
(2) $|\langle \vec{a}, \vec{b} \rangle| = \|\vec{a}\| \cdot \|\vec{b}\|$ *gilt genau dann, wenn* $\vec{b} = 0$ *oder $\vec{a}$ ein Vielfaches von $\vec{b}$ ist.*

*Beweis:* Wenn $\vec{b} = 0$, so sind beide Seiten der Ungleichung $= 0$. Wir nehmen nun an, dass $\vec{b} \neq 0$. Wir wissen, dass

$$\langle \vec{a} - \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \cdot \vec{b}, \vec{a} - \frac{\langle \vec{a}, \vec{b} \rangle}{\langle \vec{b}, \vec{b} \rangle} \cdot \vec{b} \rangle \geq 0.$$

Nun gilt $\langle \vec{a} - \frac{\langle \vec{a},\vec{b}\rangle}{\langle \vec{b},\vec{b}\rangle} \cdot \vec{b}, \vec{a} - \frac{\langle \vec{a},\vec{b}\rangle}{\langle \vec{b},\vec{b}\rangle} \cdot \vec{b}\rangle = \langle \vec{a}, \vec{a}\rangle - 2\frac{\langle \vec{a},\vec{b}\rangle^2}{\langle \vec{b},\vec{b}\rangle} + \frac{\langle \vec{a},\vec{b}\rangle^2}{\langle \vec{b},\vec{b}\rangle} = \langle \vec{a}, \vec{a}\rangle - \frac{\langle \vec{a},\vec{b}\rangle^2}{\langle \vec{b},\vec{b}\rangle}$. Da dieser Ausdruck $\geq 0$ ist, gilt

$$\langle \vec{a}, \vec{a}\rangle \cdot \langle \vec{b}, \vec{b}\rangle \geq \langle \vec{a}, \vec{b}\rangle^2.$$

Folglich gilt auch

$$||\vec{a}|| \cdot ||\vec{b}|| \geq |\langle \vec{a}, \vec{b}\rangle|.$$

Somit gilt (1).

Um (2) zu zeigen, nehmen wir an, dass $|\langle \vec{a}, \vec{b}\rangle| = ||\vec{a}|| \cdot ||\vec{b}||$ gilt. Wir zeigen, dass dann $\vec{b} = 0$ gilt, oder dass $\vec{a}$ ein Vielfaches von $\vec{b}$ ist. Es gilt $\langle \vec{a}, \vec{b}\rangle^2 = \langle \vec{a}, \vec{a}\rangle \cdot \langle \vec{b}, \vec{b}\rangle$. Wenn also $\vec{b} \neq 0$, so gilt

$$\langle \vec{a} - \frac{\langle \vec{a}, \vec{b}\rangle}{\langle \vec{b}, \vec{b}\rangle} \cdot \vec{b}, \vec{a} - \frac{\langle \vec{a}, \vec{b}\rangle}{\langle \vec{b}, \vec{b}\rangle} \cdot \vec{b}\rangle = 0.$$

Also gilt $\vec{a} - \frac{\langle \vec{a},\vec{b}\rangle}{\langle \vec{b},\vec{b}\rangle} \cdot \vec{b} = 0$, uns somit ist $\vec{a}$ ein Vielfaches von $\vec{b}$.

Nun zeigen wir folgendes: wenn $\vec{b} = 0$ oder $\vec{b}$ ein Vielfaches von $\vec{a}$ ist, so gilt $\langle \vec{a}, \vec{b}\rangle^2 = ||\vec{a}|| \cdot ||\vec{b}||$. Wenn $\vec{b} = 0$, so sind beide Seiten der Gleichung $= 0$. Wenn $\vec{b} \neq 0$ und $\vec{a} = \lambda\vec{b}$, so gilt $\langle \vec{a}, \vec{b}\rangle^2 = \langle \lambda\vec{b}, \vec{b}\rangle^2 = \lambda^2\langle \vec{b}, \vec{b}\rangle^2 = \lambda^2\langle \vec{b}, \vec{b}\rangle\langle \vec{b}, \vec{b}\rangle = \langle \lambda\vec{b}, \lambda\vec{b}\rangle \cdot \langle \vec{b}, \vec{b}\rangle = ||\vec{a}||^2||\vec{b}||^2$. ∎

ÜBUNGSAUFGABEN 1.12.

(1) Sei $\vec{a} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Zeigen Sie:
   (a) $||\lambda\vec{a}|| = |\lambda| \cdot ||\vec{a}||$.
   (b) $||\frac{1}{||\vec{a}||}\vec{a}|| = 1$.
(2) Seien $\vec{a}, \vec{b} \in \mathbb{R}^n$. Zeigen Sie: $||\vec{a} + \vec{b}||^2 - ||\vec{a} - \vec{b}||^2 = 4\langle \vec{a}, \vec{b}\rangle$.

Nun können wir wie im $\mathbb{R}^2$ zeigen:

SATZ 1.13. *Seien $\vec{a}, \vec{b}$ im $\mathbb{R}^n$, und sei $\varphi$ der Winkel, den $\vec{a}$ und $\vec{b}$ einschließen. Dann gilt*

$$\cos(\varphi) = \frac{\langle \vec{a}, \vec{b}\rangle}{||\vec{a}|| \cdot ||\vec{b}||}.$$

*Beweisskizze:* $\langle \vec{a}, \vec{a}\rangle + \langle \vec{b}, \vec{b}\rangle - 2||\vec{a}||\,||\vec{b}|| \cos(\varphi) = \langle \vec{a} - \vec{b}, \vec{a} - \vec{b}\rangle$. ∎

Im $\mathbb{R}^3$ gibt es eine Rechenoperation, die einen Vektor liefert, der auf zwei gegebene Vektoren $\vec{a}$ und $\vec{b}$ normal steht: das Kreuzprodukt.

DEFINITION 1.14. Der Vektor

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} := \begin{pmatrix} a_2b_3 - a_3b_2 \\ -(a_1b_3 - a_3b_1) \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

ist das *Kreuzprodukt* von $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ und $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ in $\mathbb{R}^3$.

SATZ 1.15. *Seien $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ in $\mathbb{R}^3$. Dann gilt:*

(1) *$\vec{a} \times \vec{b}$ ist normal auf $\vec{a}$ und auf $\vec{b}$.*

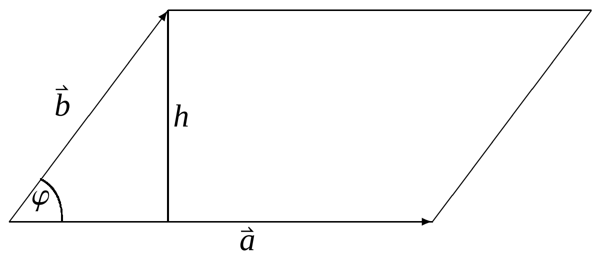
(2) $\|\vec{a} \times \vec{b}\|^2 = \|\vec{a}\|^2 \cdot \|\vec{b}\|^2 - \langle \vec{a}, \vec{b} \rangle^2$.

(3) *$\|\vec{a} \times \vec{b}\|$ ist die Fläche des Parallelogramms, das von den Vektoren $\vec{a}$ und $\vec{b}$ aufgespannt wird.*

Beweis: (1): Es gilt

$$\langle \vec{a} \times \vec{b}, \vec{a} \rangle = (a_2 b_3 - a_3 b_2) a_1 - (a_1 b_3 - a_3 b_1) a_2 + (a_1 b_2 - a_2 b_1) a_3 = 0.$$

$\langle \vec{a} \times \vec{b}, \vec{b} \rangle = 0$ wird in der Vorlesung nachgerechnet. (2): Vorlesung. (3): Wir nehmen an, dass $\vec{a} \neq 0$ und $\vec{b} \neq 0$.




Wir erhalten für die Höhe $h$ auf $\vec{a}$

$$h = \|\vec{b}\| \cdot \sin(\varphi)$$

und für den Flächeninhalt

$$F = \|\vec{a}\| \cdot h = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin(\varphi).$$

Somit gilt

$$\begin{aligned} F^2 &= \|\vec{a}\|^2 \cdot \|\vec{b}\|^2 \cdot (\sin(\varphi))^2, \\ &= \|\vec{a}\|^2 \cdot \|\vec{b}\|^2 \cdot (1 - (\cos(\varphi))^2) \\ &= \|\vec{a}\|^2 \cdot \|\vec{b}\|^2 - \langle \vec{a}, \vec{b} \rangle^2. \end{aligned}$$

Aus (2) ergibt sich jetzt

$$F^2 = \|\vec{a} \times \vec{b}\|^2.$$

■

Durch die Bedingungen an Richtung und Länge ist der Vektor $\vec{a} \times \vec{b}$ fast schon eindeutig bestimmt. Zusätzlich gilt: Zeigt der Daumen der rechten Hand in Richtung $\vec{a}$, der Zeigefinger in Richtung $\vec{b}$, und ist der Mittelfinger normal auf $\vec{a}$ und $\vec{b}$, dann zeigt er in die Richtung von $\vec{a} \times \vec{b}$.

```
In[9]:= Kreuzprodukt[{a1_, a2_, a3_}, {b1_, b2_, b3_}] :=
          {a2 * b3 - a3 * b2, -(a1 * b3 - a3 * b1), a1 * b2 - a2 * b1}

In[10]:= Kreuzprodukt[{1, -2, 3}, {0, 2, -1}]

Out[10]= {-4, 1, 2}
```

Die Mathematica-Funktion `Cross` liefert ebenfalls das Kreuzprodukt.

ÜBUNGSAUFGABEN 1.16.

(1) Zeigen Sie, dass für alle $\vec{a}, \vec{b} \in \mathbb{R}^3$ gilt: $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$.
(2) Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$. Zeigen Sie unter Verwendung der Gleichheitsaussage in der Cauchy-Ungleichung (Satz 1.11) und Satz 1.15 (2), dass $\vec{a} \times \vec{b}$ genau dann 0 ist, wenn $\vec{a} = 0$ oder $\vec{b}$ ein Vielfaches von $\vec{a}$ ist.
(3) Zeigen Sie, dass folgende Gleichheit für alle $\vec{a}, \vec{b}, \vec{c}, \vec{d} \in \mathbb{R}^3$ gilt (*Lagrange-Identität*).

$$\langle \vec{a} \times \vec{b}, \vec{c} \times \vec{d} \rangle = \langle \vec{a}, \vec{c} \rangle \cdot \langle \vec{b}, \vec{d} \rangle - \langle \vec{b}, \vec{c} \rangle \cdot \langle \vec{a}, \vec{d} \rangle.$$

(4) Zeigen Sie, dass folgende Gleichheit für alle $\vec{a}, \vec{b}, \vec{c}, \vec{d} \in \mathbb{R}^3$ gilt:

$$\langle \vec{a} \times \vec{b}, \vec{c} \rangle = \langle \vec{b} \times \vec{c}, \vec{a} \rangle.$$

(5) Zeigen Sie, dass für alle $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ gilt:

$$\vec{a} \times (\vec{b} \times \vec{c}) = \langle \vec{c}, \vec{a} \rangle \cdot \vec{b} - \langle \vec{b}, \vec{a} \rangle \cdot \vec{c}.$$

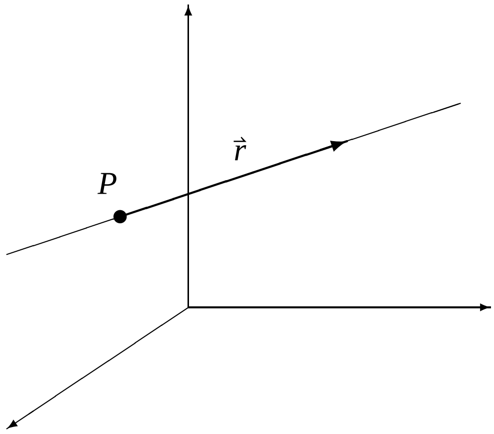
(6) Zeigen Sie, dass folgende Gleichheit für alle $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ gilt (*Jacobi-Identität*):

$$\vec{a} \times (\vec{b} \times \vec{c}) + \vec{b} \times (\vec{c} \times \vec{a}) + \vec{c} \times (\vec{a} \times \vec{b}) = 0.$$

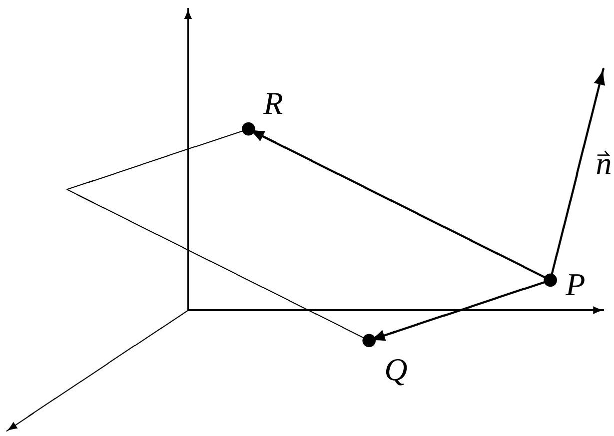
## 8. Geraden und Ebenen im Raum

**8.1. Parameterdarstellung einer Geraden.** Genau wie im $\mathbb{R}^2$ lässt sich eine Gerade im $\mathbb{R}^3$ durch eine Parameterdarstellung mit einem Punkt und einem Richtungsvektor angeben. Zum Beispiel,

$$g : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}.$$




**8.2. Parameterdarstellung einer Ebene.** Wie kann man die Ebene $e$ beschreiben, die die drei Punkte $P = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$, $Q = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}$, und $R = \begin{pmatrix} 4 \\ 2 \\ -4 \end{pmatrix}$ enthält?




Die Ebene $e$ ist die Menge aller Punkte, die man erreicht, indem man von $P$ aus ein Stück in Richtung $Q$, und dann ein Stück in die Richtung von $P$ nach $R$ geht.

$$e = \{P + \lambda \cdot \vec{PQ} + \mu \cdot \vec{PR} \mid \lambda, \mu \in \mathbb{R}\},$$

das heißt, die Punkte der Ebene sind von der Form

$$e : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}.$$

Eine Ebene lässt sich also durch einen Punkt und 2 Richtungsvektoren beschreiben.

**8.3. Implizite Darstellung einer Ebene.** Es sei $e$ die Ebene durch den Punkt $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$, die normal auf den Vektor $\vec{n} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ ist. Wir nennen $\vec{n}$ den Normalvektor von $e$.

Die Ebene $e$ ist die Menge aller Punkte $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, sodass der Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$ normal auf $\vec{n}$ ist, also

$$\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}, \vec{n} \rangle = 0.$$

Einsetzen der Werte ergibt die *implizite Darstellung der Ebene*,

$$e = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + 3y - 2z = -9 \}.$$

Ein Normalvektor von $e$ lässt sich direkt aus den Koeffizienten der Ebenengleichung ablesen.

8.3.1. *Verwandlung von Parameterdarstellung in implizite Darstellung.* Wir verwandeln

$$e : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}$$

in implizite Form Dazu suchen wir zuerst einen Vektor $\vec{n}$, der auf beide Richtungsvektoren der Ebene normal ist. Dann ist $\vec{n}$ auf die ganze Ebene normal. Wir beschreiben 2 Möglichkeiten einen solchen Normalvektor zu finden:

(1) Wir suchen $\vec{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$ so, dass

$$\begin{aligned} \langle \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \vec{n} \rangle &= 0 \text{ und} \\ \langle \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}, \vec{n} \rangle &= 0. \end{aligned}$$

Das heißt, wir müssen das lineare Gleichungssystem

$$\begin{aligned} n_1 - n_2 + 3n_3 &= 0 \\ 2n_1 - n_2 - 3n_3 &= 0 \end{aligned}$$

lösen. Klarerweise ist $n_1 = n_2 = n_3 = 0$ eine Lösung, aber wir wollen einen Vektor $\vec{n}$, der nicht der Nullvektor ist. Wie man alle Lösungen eines linearen Gleichungssystems findet, werden wir im Kapitel 3 sehen.

(2) Alternativ finden wir $\vec{n}$ auch als Kreuzprodukt der beiden Richtungsvektoren.

$$\vec{n} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ 1 \end{pmatrix}$$

Der Vektor $\vec{n}$ ist normal auf die Ebene. Ein Punkt $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ liegt also genau dann in $e$, wenn der Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ normal auf $\vec{n}$ ist. Wir berechnen

$$\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}, \vec{n} \rangle = 0$$

und erhalten

$$6x + 9y + z = 38.$$

Somit hat die Ebene $e$ die implizite Darstellung

$$e = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid 6x + 9y + z = 38 \}.$$

8.3.2. *Verwandlung von impliziter Darstellung in Parameterdarstellung.* Wir verwandeln $e : x + 3y - 2z = -9$ in parametrisierte Form. Wir beschreiben die Lösungsmenge der Gleichung, indem wir $z = \mu$ und $y = \lambda$ setzen und dann $x$ durch $\lambda$ und $\mu$ ausdrücken. Wir erhalten $x = -9 - 3\lambda + 2\mu$. Somit liegt für alle Werte $\lambda, \mu \in \mathbb{R}$ der Punkt $\begin{pmatrix} -9-3\lambda+2\mu \\ \lambda \\ \mu \end{pmatrix}$ in der Ebene. Die Ebene hat also die parametrisierte Darstellung

$$e : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -9 \\ 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

**8.4. Implizite Darstellung einer Geraden.** Offenbar kann man eine Gerade im Raum nicht durch eine einzige lineare Gleichung in $x, y, z$ beschreiben. Solche Gleichungen beschreiben nämlich Ebenen im Raum. Jede Gerade kann man aber implizit als Schnitt zweier Ebenen, das heißt als Lösungsmenge von 2 linearen Gleichungen beschreiben. Beispielsweise ist

$$g = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid 2x - y + 3z = 1, x + 4y - 2z = 3 \}$$

die Gerade, die sowohl in der Ebene mit der Gleichung $2x - y + 3z = 1$ als auch in der Ebene mit der Gleichung $x + 4y - 2z = 3$ liegt.

Zwei Ebenen im Raum, die nicht parallel sind, schneiden sich immer in einer Geraden. Parallele Ebenen erkennt man daran, dass ihre Normalvektoren in dieselbe Richtung zeigen. Also sind etwa $2x - y + 3z = 1$ und $-4x + 2y - 6z = 3$ parallel.

8.4.1. *Verwandlung von Parameterdarstellung in implizite Darstellung.* Wir verwandeln

$$g : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$$

in implizite Form. Dazu suchen wir 2 Ebenen, die die Gerade enthalten. Liegt $g$ in einer Ebene mit Normalvektor $\vec{n}$, dann ist $\vec{n}$ normal auf den Richtungsvektor $\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ der Geraden. Zusätzlich liegt der Punkt $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ in der Ebene.

Auf einen Vektor $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ sind beispielsweise $\begin{pmatrix} b \\ -a \\ 0 \end{pmatrix}$, $\begin{pmatrix} c \\ 0 \\ -a \end{pmatrix}$ und $\begin{pmatrix} 0 \\ c \\ -b \end{pmatrix}$ normal. Wir wählen $n_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ und $n_2 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$ als Vektoren, die im rechten Winkel auf $\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ stehen. Damit liegt $g$ in den Ebenen durch den Punkt $\left(\begin{smallmatrix}2\\3\\-1\end{smallmatrix}\right)$, die normal auf $n_1$ bzw. $n_2$ sind. Die Gerade ist also der Durchschnitt der beiden Ebenen

$$\begin{aligned} e_1 : &\quad x + y = 5, \text{ und} \\ e_2 : &\quad 3x - z = 7. \end{aligned}$$

Wir haben

$$g = \{\left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) \in \mathbb{R}^3 \mid x + y = 5, 3x - z = 7\}.$$

8.4.2. *Verwandlung von impliziter Darstellung in Parameterdarstellung.* Um eine Parameterdarstellung von

$$g = \{\left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) \in \mathbb{R}^3 \mid 2x - y + 3z = 1, x + 4y - 2z = 3\}$$

zu erhalten, lösen wir das lineare Gleichungssystem

$$\begin{aligned} 2x - y + 3z &= 1, \\ x + 4y - 2z &= 3. \end{aligned}$$

Eine Methode dafür werden wir im Kapitel 3 vorstellen.

KAPITEL 2

# Matrizen

## 1. Die Definition von Matrizen

Wir haben bereits *Vektoren* kennen gelernt; solche Paare reeller Zahlen haben wir etwa benutzt, um Punkte in der Ebene zu beschreiben. In der Geometrie brauchen wir auch *Matrizen*. Matrizen eignen sich besonders gut, um etwa Drehungen oder Spiegelungen zu beschreiben.

Eine *Matrix* ist ein rechteckiges Zahlenschema. Zunächst einige Beispiele:

BEISPIELE 2.1.

- $\begin{pmatrix} 1 & 3 & 5 \\ 7 & 2 & 1 \end{pmatrix}$ ist eine $2 \times 3$-Matrix.
- $\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$ ist eine $2 \times 2$-Matrix.
- $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ ist eine $2 \times 1$-Matrix.
- $\begin{pmatrix} 1 & 2 & 7 \end{pmatrix}$ ist eine $1 \times 3$-Matrix.

Eine Matrix mit $m$ Zeilen und $n$ Spalten bezeichnen wir als eine $m \times n$-Matrix. Wenn $A$ eine $m \times n$-Matrix ist, und $i \in \{1, 2, \ldots, m\}$ und $j \in \{1, 2, \ldots, n\}$, so bezeichnen wir mit $A(i, j)$, $A[i, j]$ oder $A_{i,j}$ den Eintrag, der bei $A$ in der $i$-ten Zeile und $j$-ten Spalte steht. Für

$$A = \left(\begin{smallmatrix} 1 & 3 & 5 \\ 7 & 2 & 1 \end{smallmatrix}\right)$$

gilt zum Beispiel $A_{2,1} = 7$. Die Menge aller $m \times n$-Matrizen kürzen wir mit $\mathbb{R}^{m \times n}$ ab.

Wir müssen noch den Begriff "rechteckiges Zahlenschema" klären. Man kann eine $m \times n$-Matrix $A$ mit Einträgen aus $\mathbb{R}$ als Funktion von $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ nach $\mathbb{R}$ definieren. Der Eintrag, der in der 2. Zeile und 4. Spalte steht, ist dann der Funktionswert $A(2, 4)$. Diese Sichtweise gibt auch recht gut wieder, was eine Implementation des abstrakten Datentyps "Matrix" können muss. Es muss möglich sein, eine Funktion `LiefereEintrag` zu schreiben, sodass `LiefereEintrag (A, i, j)` den Eintrag von $A$ an der $i$-ten Zeile und $j$-ten Spalte, also den Funktionswert $A(i, j)$, zurückgibt.

In Mathematica geben wir die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

wie folgt ein.

```
In[11]:= A = {{1, 2, 3}, {4, 5, 6}}
Out[11]= {{1, 2, 3}, {4, 5, 6}}

In[12]:= MatrixForm[A]
Out[12]= (1  2  3)
         (4  5  6)

In[13]:= A = {{5, 7, 8}, {-2, 3, 5}}
Out[13]= {{5, 7, 8}, {-2, 3, 5}}

In[14]:= MatrixForm[A]
Out[14]= ( 5  7  8)
         (-2  3  5)

In[15]:= A[[2]][[1]]
Out[15]= -2
```

## 2. Die Addition von Matrizen

Zwei Matrizen kann man addieren, wenn sie gleich viele Zeilen und gleich viele Spalten haben. Wie man zwei Matrizen von gleichem Format addiert, erklären wir mit folgenden Beispielen.

AUFGABEN 2.2.

- $\begin{pmatrix} 1 & 3 & 5 \\ 7 & 2 & 1 \end{pmatrix} + \begin{pmatrix} -1 & -2 & 3 \\ 6 & 5 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 8 \\ 13 & 7 & -2 \end{pmatrix}$
- $\begin{pmatrix} 1 & -3 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ 7 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 9 & 1 \end{pmatrix}$
- $\begin{pmatrix} 1 & 5 \\ 3 & 7 \end{pmatrix} + \begin{pmatrix} 2 & -3 \\ -4 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & 7 \end{pmatrix}$
- $\begin{pmatrix} 1 & 3 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -2 & -5 \end{pmatrix}$.

Wir fassen zusammen, wie diese Addition funktioniert: Zwei Matrizen $A \in \mathbb{R}^{m\times k}$, $B \in \mathbb{R}^{n\times l}$ lassen sich genau dann addieren, wenn $m = n$ und $k = l$ gilt, d.h. wenn die Matrizen von gleichem Format sind. Wenn $C$ die Matrix $A + B$ ist, dann hat auch $C$ das Format $m \times k$, und für alle $i \in \{1, 2, \ldots, m\}$ und $j \in \{1, 2, \ldots k\}$ berechnet man den Eintrag $C_{i,j}$ durch

$$C_{i,j} = A_{i,j} + B_{i,j}.$$

```
In[16]:= A = {{1, 4, 3}, {0, 1, 0}};

In[17]:= B = {{-7, 5, 0}, {23, -7, 16}};

In[18]:= A + B
Out[18]= {{-6, 9, 3}, {23, -6, 16}}

In[19]:= MatrixForm[%]
```

$$\text{Out[19]=}\ \begin{pmatrix} -6 & 9 & 3 \\ 23 & -6 & 16 \end{pmatrix}$$

### 3. Die Multiplikation einer Matrix mit einer reellen Zahl

Eine Matrix $A$ wird mit einer reellen Zahl multipliziert, indem jeder Eintrag mit der Zahl multipliziert wird. Wir geben dazu wieder ein Beispiel:

AUFGABE 2.3.

$$2 * \begin{pmatrix} 3 & 5 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 10 \\ 14 & 16 \end{pmatrix}.$$

Wir formulieren wieder allgemein, wie man eine reelle Zahl mit einer Matrix $A$ multipliziert. Wenn $t$ eine reelle Zahl, und $A$ eine $m\times n$-Matrix ist, so ist die Matrix $C := t*A$ ebenfalls eine $m \times n$-Matrix. Die Einträge von $C$ sind dadaurch gegeben, dass für alle $i \in \{1, 2, \ldots, m\}$, $j \in \{1, 2, \ldots, n\}$ gilt:

$$C_{i,j} = t\,A_{i,j}.$$

```
In[20]:= A = {{2, 5}, {3, 4}, {10, 2}}
Out[20]= {{2, 5}, {3, 4}, {10, 2}}

In[21]:= MatrixForm[(-10) * A]
```

$$\text{Out[21]=}\ \begin{pmatrix} -20 & -50 \\ -30 & -40 \\ -100 & -20 \end{pmatrix}$$

### 4. Die Multiplikation von Matrizen

Zwei Matrizen $A$, $B$ können genau dann miteinander multipliziert werden, wenn $A$ genausoviele Spalten wie $B$ Zeilen hat. Eine $k\times l$-Matrix ist also mit einer $m\times n$-Matrix multiplizierbar, wenn $l = m$. Das Ergebnis dieser Multiplikation ist eine $k \times n$-Matrix. Wir erklären die Matrixmultiplikation zunächst anhand eines Beispiels.

AUFGABE 2.4.

$$\begin{pmatrix} 3 & 1 & 2 \\ 2 & 5 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 9 & 3 \\ 1 & 8 & 5 \\ 7 & 1 & -4 \end{pmatrix} =$$

$$\begin{pmatrix} 3\cdot 3 + 1\cdot 1 + 2\cdot 7 & 3\cdot 9 + 1\cdot 8 + 2\cdot 1 & 3\cdot 3 + 1\cdot 5 + 2\cdot(-4) \\ 2\cdot 3 + 5\cdot 1 + 4\cdot 7 & 2\cdot 9 + 5\cdot 8 + 4\cdot 1 & 2\cdot 3 + 5\cdot 5 + 4\cdot(-4) \end{pmatrix}.$$

Daher gilt

$$\begin{pmatrix} 3 & 1 & 2 \\ 2 & 5 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 9 & 3 \\ 1 & 8 & 5 \\ 7 & 1 & -4 \end{pmatrix} = \begin{pmatrix} 24 & 37 & 6 \\ 39 & 62 & 15 \end{pmatrix}.$$

Wenn man eine $k \times m$-Matrix $A$ mit einer $m \times n$-Matrix $B$ multipliziert, so ist das Produkt $C$ eine $k \times n$-Matrix. Für $i \in \{1, 2, \ldots, k\}$ und $j \in \{1, 2, \ldots, n\}$ ist der Eintrag $C_{i,j}$ das Skalarprodukt aus der $i$-ten Zeile von $A$ und der $j$-ten Spalte von $B$. Wir rechnen noch einige Beispiele:

AUFGABEN 2.5.

- $\begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 5 & 2 \\ 1 & 7 \\ 8 & -2 \end{pmatrix} = \begin{pmatrix} 24 & 19 \\ 23 & -13 \end{pmatrix}.$
- $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 5 \\ 3 & 7 \end{pmatrix}$ ist nicht definiert, da die erste Matrix 3 Spalten und die zweite Matrix 2 Zeilen hat, und 2 nicht gleich 3 ist.

Wenn $A$ eine $2 \times 3$ und $B$ eine $3 \times 1$-Matrix ist, dann ist das Produkt $A \cdot B$ eine $2 \times 1$-Matrix. Das Produkt $B \cdot A$ ist nicht definiert. Selbst dann, wenn beide Produkte $A \cdot B$ und $B \cdot A$ definiert sind, müssen die Ergebnisse nicht gleich sein. Dazu rechnen wir folgende Beispiele:

AUFGABEN 2.6.

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

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix} &= (17) \\ \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 5 \end{pmatrix} &= \begin{pmatrix} 5 & 10 & 25 \\ 1 & 2 & 5 \\ 2 & 4 & 10 \end{pmatrix} \end{aligned}$$

Das erste Beispiel noch einmal in Mathematica.

```
In[22]:= A = {{3, 1, 2}, {2, 5, 4}};
         B = {{3, 9, 3}, {1, 8, 5}, {7, 1, -4}};

In[23]:= MatrixForm[A.B]
Out[23]= (24  37  6 )
         (39  62  15)
```

Wir halten die Definition der Matrizenmultiplikation noch einmal genau fest:

DEFINITION 2.7 (Matrizenmultiplikation). Sei $A$ eine $k \times m$-Matrix und $B$ eine $m \times n$-Matrix. Dann ist das Produkt $C := A \cdot B$ eine $k \times n$-Matrix. Für $i \in \{1, 2, \dots, k\}$ und $j \in \{1, 2, \dots, n\}$ ist der Eintrag $C[i, j]$ durch

$$C[i, j] = \sum_{r=1}^{m} A[i, r] \cdot B[r, j].$$

definiert.

## 5. Rechenregeln für die Addition und Multiplikation von Matrizen

Wir haben bereits gesehen, dass nicht für alle Matrizen $A \cdot B = B \cdot A$ gelten muss. Einige Rechenregeln, die wir vom Rechnen mit Zahlen kennen, gelten aber auch für Matrizen.

SATZ 2.8 (Assoziativität der Matrizenmultiplikation). *Seien $k, l, m, n \in \mathbb{N}$, und seien $A \in \mathbb{R}^{k \times l}$, $B \in \mathbb{R}^{l \times m}$, $C \in \mathbb{R}^{m \times n}$. Dann gilt*

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

Es ist nicht schwierig, die Assoziativität der Matrizenmultiplikation zu beweisen, wenn $A, B, C$ alle $2 \times 2$-Matrizen sind. Man berechnet

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right) \cdot \begin{pmatrix} i & j \\ k & l \end{pmatrix}$$

und

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \left(\begin{pmatrix} e & f \\ g & h \end{pmatrix} \cdot \begin{pmatrix} i & j \\ k & l \end{pmatrix}\right),$$

und stellt fest, dass beide Ergebnisse gleich sind. Für Matrizen beliebigen Formats gehen wir so vor:

*Beweis von Satz 2.8:* Wir beobachten, dass $(A \cdot B) \cdot C$ und $A \cdot (B \cdot C)$ beides $k \times n$-Matrizen sind. Um zu zeigen, dass beide Matrizen gleich sind, zeigen wir, dass für alle $i \in \{1, \dots, k\}$ und $j \in \{1, \dots, n\}$ gilt, dass $(A \cdot B) \cdot C\,[i, j] = A \cdot (B \cdot C)\,[i, j]$.

Seien also $i \in \{1, \dots, k\}$ und $j \in \{1, \dots, n\}$. Wir wollen

$$(A \cdot B) \cdot C\,[i, j] = A \cdot (B \cdot C)\,[i, j] \tag{2.1}$$

zeigen. Es gilt

$$
\begin{aligned}
(A \cdot B) \cdot C\,[i,j] &= \sum_{r=1}^{m} (A \cdot B)[i,r] \cdot C[r,j] \\
&= \sum_{r=1}^{m} \Big(\sum_{s=1}^{l} A[i,s] \cdot B[s,r]\Big) \cdot C[r,j] \\
&= \sum_{r=1}^{m} \sum_{s=1}^{l} A[i,s] \cdot B[s,r] \cdot C[r,j].
\end{aligned}
$$

Wir berechnen nun die rechte Seite von (2.1), und erhalten

$$
\begin{aligned}
A \cdot (B \cdot C)[i,j] &= \sum_{r=1}^{l} A[i,r] \cdot (B \cdot C)[r,j] \\
&= \sum_{r=1}^{l} A[i,r] \cdot \Big(\sum_{s=1}^{m} B[r,s] \cdot C[s,j]\Big) \\
&= \sum_{r=1}^{l} \sum_{s=1}^{m} A[i,r] \cdot B[r,s] \cdot C[s,j]) \\
&= \sum_{s=1}^{m} \sum_{r=1}^{l} A[i,r] \cdot B[r,s] \cdot C[s,j]) \\
&= \sum_{r_1=1}^{m} \sum_{s_1=1}^{l} A[r,s_1] \cdot B[s_1,r_1] \cdot C[r_1,j] \\
&= \sum_{r=1}^{m} \sum_{s=1}^{l} A[i,s] \cdot B[s,r] \cdot C[r,j].
\end{aligned}
$$

Somit gilt die Gleichung (2.1). Folglich haben $(A \cdot B) \cdot C$ und $A \cdot (B \cdot C)$ also an jeder Stelle den gleichen Eintrag, und sind somit gleich. ■

SATZ 2.9 (Rechtsdistributivität der Matrizenmultiplikation). *Seien $k, l, m \in \mathbb{N}$, und seien $A, B \in \mathbb{R}^{k \times l}$, $C \in \mathbb{R}^{l \times m}$. Dann gilt*

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

SATZ 2.10 (Linksdistributivität der Matrizenmultiplikation). *Seien $k, l, m \in \mathbb{N}$, und seien $A \in \mathbb{R}^{k \times l}$, $B, C \in \mathbb{R}^{l \times m}$. Dann gilt*

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

SATZ 2.11 (Zusammenhang zwischen zwei Multiplikationen). *Seien $m, n \in \mathbb{N}$, sei $A \in \mathbb{R}^{m \times n}$, und sei $t \in \mathbb{R}$. Dann gilt $(tA) \cdot B = t(A \cdot B)$.*

## 6. Die Multiplikation von Vektoren und Matrizen

Sei $v = \left(\begin{smallmatrix} -2 \\ 4 \end{smallmatrix}\right)$ und $A = \begin{pmatrix} -3 & 4 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$. Dann ist der Vektor $A \cdot v$ gegeben durch

$$A \cdot v = \begin{pmatrix} -3 & 4 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 + 16 \\ -2 + 0 \\ 2 + 4 \end{pmatrix}.$$

Das Ergebnis ist ein Vektor im $\mathbb{R}^3$.

Die Multiplikation sieht also genauso aus wie die Multiplikation der $3 \times 2$-Matrix $\begin{pmatrix} -3 & 4 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$ mit der $2 \times 1$-Matrix $\left(\begin{smallmatrix} -2 \\ 4 \end{smallmatrix}\right)$. Bei der Matrizenmultiplikation ist das Ergebnis aber eine $3 \times 1$-Matrix.

Mit Mathematica wird der Unterschied deutlich:

```
In[24]:= A = {{-3, 4}, {1, 0}, {-1, 1}};

         v = {-2, 4};

         x = A.v
Out[24]= {22, -2, 6}
```

```
In[25]:= A = {{-3, 4}, {1, 0}, {-1, 1}};

         v = {{-2}, {4}};

         x = A.v
Out[25]= {{22}, {-2}, {6}}
```

```
In[26]:= A = {{-3, 4}, {1, 0}, {-1, 1}};

         v = {{-2, 4}};

         x = A.v
```

Hier liefert Mathematica eine Fehlermeldung.

```
Out[26]= {{-3, 4}, {1, 0}, {-1, 1}}.{{-2, 4}}
```

Sei $v = (-4, 3, 2)$ und $A = \begin{pmatrix} -3 & 5 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$. Dann ist der Vektor $v \cdot A$ gegeben durch

$$(-4, 3, 2) \cdot \begin{pmatrix} -3 & 5 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} = (12 + 3 - 2,\ -20 + 2) = (13, -18).$$

Das Ergebnis ist ein Vektor im $\mathbb{R}^2$.

Die Multiplikation sieht also genauso aus wie die Multiplikation der $1 \times 3$-Matrix $(-4\ 3\ 2)$ mit der $3 \times 2$-Matrix $\begin{pmatrix} -3 & 5 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$. Bei der Matrizenmultiplikation ist das Ergebnis aber eine $1 \times 2$-Matrix.

Wenn man diese Multiplikation “Matrix mal Vektor” verwendet, lassen sich lineare Gleichungssysteme kürzer anschreiben.

$$\begin{aligned} 3x + 4y - 2z &= 1 \\ 2x + 5y - 8z &= 2 \end{aligned}$$

läßt sich dann als

$$\begin{pmatrix} 3 & 4 & -2 \\ 2 & 5 & -8 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

schreiben. Im allgemeinen erhält man bei $m$ Gleichungen und $n$ Unbekannten die Form

$$A \cdot x = b,$$

wobei $A$ eine $m \times n$-Matrix ist, $x$ ein Vektor im $\mathbb{R}^n$ und $b$ ein Vektor im $\mathbb{R}^m$.

Die Funktion `LinearSolve[A,b]` liefert eine Lösung des linearen Gleichungssystems $A \cdot x = b$.

Wir lösen zum Beispiel $2x - 3y = 5$.

```
In[27]:= LinearSolve[{{2, -3}}, {5}]
```

Out[27]= $\left\{\frac{5}{2}, 0\right\}$

Später werden wir sehen, wie man alle Lösungen erhält.

## 7. Das Transponieren von Matrizen

Beim *Transponieren* einer Matrix wird die Matrix an der Hauptdiagonale gespiegelt.

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

Wenn $A$ eine $m \times n$-Matrix ist, so ist $A^T$ eine $n \times m$-Matrix, und für alle $i \in \{1, 2, \ldots, n\}$ und $j \in \{1, 2, \ldots, m\}$ gilt

$$A^T(i, j) = A(j, i).$$

```
In[28]:= A = {{1, 4, -3}, {2, -5, 3}};
```

```
In[29]:= MatrixForm[A]
```

*Out[29]=* $\begin{pmatrix} 1 & 4 & -3 \\ 2 & -5 & 3 \end{pmatrix}$

*In[30]:=* **B = Transpose[A]**

*Out[30]=* {{1, 2}, {4, -5}, {-3, 3}}

*In[31]:=* **MatrixForm[B]**

*Out[31]=* $\begin{pmatrix} 1 & 2 \\ 4 & -5 \\ -3 & 3 \end{pmatrix}$

SATZ 2.12. *Seien $l, m, n \in \mathbb{N}$, sei $A \in \mathbb{R}^{l \times m}$, und seien $B, C \in R^{m \times n}$. Dann gilt:*

(1) $(A \cdot B)^T = B^T \cdot A^T$
(2) $(B + C)^T = B^T + C^T$

ÜBUNGSAUFGABEN 2.13.

(1) Berechnen Sie für die Matrix
$$A := \begin{pmatrix} 0 & -2 & 3 \\ 8 & 7 & 6 \end{pmatrix}$$
die Matrix $B := A^T \cdot A$.

(2) Berechnen Sie $(A - B) \cdot C^T$ für die Matrizen
$$A = \begin{pmatrix} 1 & 3 \\ 4 & -2 \end{pmatrix}, B = \begin{pmatrix} -2 & -3 \\ 2 & 2 \end{pmatrix}, C = \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}.$$

(3) Finden Sie eine Matrix $X$, sodaß $A \cdot X = B$, wobei $A = \begin{pmatrix} -4 & 0 \\ 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} -4 & -8 \\ 7 & 10 \end{pmatrix}$. (Hinweis: Bestimmen Sie jede Spalte von $X$ durch Lösen eines linearen Gleichungssystems.)

## 8. Die Einheitsmatrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$$
$$\begin{pmatrix} 3 & 5 \\ 4 & 2 \\ 8 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 4 & 2 \\ 8 & 1 \end{pmatrix}$$

Die Matrix
$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$
heißt Einheitsmatrix vom Format $n \times n$. Man sieht leicht, daß für jede $m \times n$-Matrix $A$ und jede $n \times k$-Matrix $B$ gilt:
$$\begin{aligned} A \cdot E_n &= A, \\ E_n \cdot B &= B. \end{aligned}$$

Besonders einfach zu lösen sind Gleichungssysteme mit der Einheitsmatrix: Das Gleichungssystem

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

hat die Lösung $x = 4$, $y = 2$, und daher die Lösungsmenge $L = \{(4, 2)\}$.

```
In[32]:= A = -24 * IdentityMatrix[5];
         MatrixForm[A]
```

$$\mathit{Out}[32]= \begin{pmatrix} -24 & 0 & 0 & 0 & 0 \\ 0 & -24 & 0 & 0 & 0 \\ 0 & 0 & -24 & 0 & 0 \\ 0 & 0 & 0 & -24 & 0 \\ 0 & 0 & 0 & 0 & -24 \end{pmatrix}$$

## 9. Das Invertieren von Matrizen

Betrachtet man die Gleichung $5\,x = 7$, so erhält man die Lösung $x = \frac{7}{5}$ durch Multiplikation beider Seiten mit $\frac{1}{5}$ (des Inversen von 5). Wir betrachten das Gleichungssystem

$$\begin{pmatrix} 2 & 3 \\ 5 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 12 \\ -20 \end{pmatrix}$$

Seien wir nun optimistisch, und stellen wir uns vor, wir haben eine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, sodass

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 5 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Für die Lösungen des Gleichungssystems muss dann auch gelten:

$$\begin{aligned} A \cdot \begin{pmatrix} 2 & 3 \\ 5 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= A \cdot \begin{pmatrix} 12 \\ -20 \end{pmatrix} \\ \begin{pmatrix} x \\ y \end{pmatrix} &= A \cdot \begin{pmatrix} 12 \\ -20 \end{pmatrix} \end{aligned}$$

Wie bestimmen wir so eine Matrix $A$? Wir suchen eine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, die folgende Eigenschaft besitzt:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 5 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Durch Ausmultiplizieren erhalten wir folgendes Gleichungssystem:

$$\begin{aligned} 2a + 5b &= 1 \\ 3a - 5b &= 0 \\ 2c + 5d &= 0 \\ 3c - 5d &= 1 \end{aligned}$$

Lösen wir dieses, so erhalten wir

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0,2 & 0,12 \\ 0,2 & -0,08 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 5 & 3 \\ 5 & -2 \end{pmatrix}.$$

Nun können wir auch die Lösung des ursprünglichen Systems berechnen:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 5 & 3 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} 12 \\ -20 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$$

Somit ist $(0,4)$ der einzige Kandidat für eine Lösung des Systems. Da $(0,4)$ auch wirklich Lösung ist, ergibt sich als Lösungsmenge $L = \{(0,4)\}$.

DEFINITION 2.14. Sei $A$ eine $n \times n$-Matrix über $\mathbb{R}$. $A$ heißt *invertierbar*, falls es eine $n \times n$-Matrix $B$ mit $A \cdot B = B \cdot A = E_n$ gibt.

SATZ 2.15. *Seien $A_1, A_2$ invertierbare Matrizen in $\mathbb{R}^{n\times n}$. Dann ist auch $A_1 \cdot A_2$ invertierbar.*

*Beweis.* Seien $A_1, A_2 \in \mathbb{R}^{n\times n}$ invertierbar. Es gibt daher Matrizen $B_1, B_2$, sodass $A_1 \cdot B_1 = B_1 \cdot A_1 = E_n$ und $A_2 \cdot B_2 = B_2 \cdot A_2 = E_n$. Dann gilt $(A_1 \cdot A_2) \cdot (B_2 \cdot B_1) = A_1 \cdot (A_2 \cdot B_2) \cdot B_1 = A_1 \cdot E_n \cdot B_1 = A_1 \cdot B_1 = E_n$. Somit ist $A_1 \cdot A_2$ invertierbar. ■

SATZ 2.16. *Sei $A$ eine invertierbare Matrix in $\mathbb{R}^{n\times n}$, und sei $B$ so, dass $A \cdot B = B \cdot A = E_n$. Sei $C$ eine Matrix mit $A \cdot C = E_n$. Dann gilt $B = C$.*

*Beweis.* Es gilt $C = E_n \cdot C = (B \cdot A) \cdot C = B \cdot (A \cdot C) = B \cdot E_n = B$.■

Zu jeder invertierbaren Matrix $A$ gibt es also genau eine Matrix $B$ mit $A \cdot B = E_n$. Diese Matrix $B$ kürzen wir mit $A^{-1}$ ab.

DEFINITION 2.17. Sei $A$ eine $n \times n$-Matrix. $A$ ist *regulär* genau dann, wenn $A$ invertierbar ist. $A$ ist *singulär* genau dann, wenn $A$ nicht invertierbar ist.

ÜBUNGSAUFGABEN 2.18.

(1) Zeigen Sie, dass für $a, b, c, d \in \mathbb{R}$ mit $ad \neq bc$ die Matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ invertierbar ist, und dass

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

gilt.

(2) Sei $A$ eine $m \times m$- Matrix, für die es ein $n \in \mathbb{N}$ mit $A^n = 0$ gibt, und sei $E$ die $m \times m$-Einheitsmatrix. Zeigen Sie, dass $E - A$ invertierbar ist. *Hinweis:* Denken Sie beim Auffinden der inversen Matrix an $\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i$.

SATZ 2.19. *Seien $A, B \in \mathbb{R}^{n\times n}$ invertierbare Matrizen.*

(1) *$A^{-1}$ ist invertierbar, und es gilt $(A^{-1})^{-1} = A$.*
(2) *$A^T$ ist invertierbar, und es gilt $(A^T)^{-1} = (A^{-1})^T$.*
(3) *$A \cdot B$ ist invertierbar, und es gilt $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.*

*Beweis.*

(1) Es gilt $A^{-1} \cdot A = A \cdot A^{-1} = E_n$. Also ist $A^{-1}$ invertierbar, und $B := A$ ihre inverse Matrix.
(2) Es gilt $A \cdot A^{-1} = A^{-1} \cdot A = E_n$. Durch Transponieren erhält man $(A^{-1})^T \cdot A^T = A^T \cdot (A^{-1})^T = E_n$. Folglich ist $A^T$ invertierbar, und die inverse Matrix zu $A^T$ ist $(A^{-1})^T$.
(3) Es gilt $(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = E_n$ und $(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = E_n$. Folglich ist $A \cdot B$ invertierbar, und die inverse Matrix von $A \cdot B$ ist $B^{-1} \cdot A^{-1}$.

■

Den folgenden Satz werden wir erst später beweisen (Satz 12.22 und Satz 12.23):

SATZ 2.20. *Sei $A \in \mathbb{R}^{n\times n}$, und sei $B \in \mathbb{R}^{n\times n}$ so, dass $A \cdot B = E_n$. Dann ist $A$ invertierbar. Außerdem ist dann $B$ die zu $A$ inverse Matrix.*

Wir berechnen jetzt die inverse Matrix von $\begin{pmatrix} 1 & 3 \\ 2 & -4 \end{pmatrix}$ durch den Ansatz

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

```
In[33]:= A = {{1, 3}, {2, -4}};

         LinearSolve [A, {1, 0}]
```

Out[33]= $\left\{\frac{2}{5}, \frac{1}{5}\right\}$

Also $a = 0.4, c = 0.2$.

```
In[34]:= A = {{1, 3}, {2, -4}};

         LinearSolve [A, {0, 1}]
```

Out[34]= $\left\{\frac{3}{10}, -\frac{1}{10}\right\}$

Also $b = 0.3$, $d = -0.1$.

Die Funktion `Inverse` berechnet die inverse Matrix; die Funktion `^(-1)` macht leider etwas ganz anderes.

```
In[35]:= A = {{1, 3}, {2, -4}};

         B = Inverse[A];

         MatrixForm [B]
```

Out[35]= $\begin{pmatrix} \frac{2}{5} & \frac{3}{10} \\ \frac{1}{5} & -\frac{1}{10} \end{pmatrix}$

```
In[36]:= A.B
Out[36]= {{1, 0}, {0, 1}}

In[37]:= B.A
Out[37]= {{1, 0}, {0, 1}}

In[38]:= A^(-1)
```

Out[38]= $\left\{\left\{1, \frac{1}{3}\right\}, \left\{\frac{1}{2}, -\frac{1}{4}\right\}\right\}$

```
In[39]:= A.A^(-1)
```

Out[39]= $\left\{\left\{\frac{5}{2}, -\frac{5}{12}\right\}, \left\{0, \frac{5}{3}\right\}\right\}$

KAPITEL 3

# Lineare Gleichungssysteme

## 1. Beispiele

Wir betrachten zunächst vier Gleichungssysteme und bestimmen ihre Lösungsmenge. Dabei geht es uns noch nicht darum, ein Lösungsverfahren für lineare Gleichungssysteme zu entwickeln (das kommt später), sondern nur darum, ein paar typische Phänomene zu beobachten.

(1) Man bestimme alle Paare $(x, y)$ in $\mathbb{R}^2$, die sowohl die Gleichung $x + y = -1$ als auch die Gleichung $3x + 2y = -5$ erfüllen. Wir suchen also alle Lösungen des Gleichungssystems

$$\begin{aligned} x + y &= -1 \qquad (1) \\ 3x + 2y &= -5 \qquad (2). \end{aligned} \tag{3.1}$$

*Lösung:* Wenn $\binom{x}{y}$ eine Lösung des Gleichungssystems ist, dann gilt $x + y = -1$. Daher gilt auch $-3x - 3y = 3$. Somit ist jede Lösung des Systems (3.1) auch eine Lösung von

$$\begin{aligned} -3x - 3y &= 3 \qquad (1') \\ 3x + 2y &= -5 \qquad (2). \end{aligned} \tag{3.2}$$

Wenn $-3x-3y = 3$ und $3x+2y = -5$, dann gilt auch $(-3x-3y)+(3x+2y) = 3 + (-5)$, also $-y = -2$. Daher muss $y = 2$ sein. Da aber $x + y = -1$ ist, muss $x = -1 - y$ sein, und daher ist $x = -3$. Daher ist nur $\binom{-3}{2}$ als Lösung des Gleichungssystems möglich.

Wir probieren nun aus, ob $\binom{-3}{2}$ auch wirklich eine Lösung ist. Tatsächlich gilt $-3 + 2 = -1$ und $3 \cdot (-3) + 2 \cdot 2 = -5$. Daher ist die Menge

$$L = \{\tbinom{-3}{2}\}$$

die Lösungsmenge des Gleichungssystems.

Wir interpretieren dieses Beispiel jetzt geometrisch. Jene Punkte $\binom{x}{y}$ im $\mathbb{R}^2$, die die Gleichung $x + y = -1$ erfüllen, liegen auf einer Geraden (eben auf der Geraden mit Gleichung $x + y = -1$). Jene Punkte $\binom{x}{y}$, die die Gleichung $3x + 2y = -5$ erfüllen, liegen auf der Geraden mit Gleichung $3x + 2y = -5$. Die Lösungsmenge des Gleichungssystems enthält alle Punkte, die auf beiden Geraden liegen. Wenn die beiden Geraden nicht parallel sind, so gibt es genau einen Punkt, der auf beiden Geraden liegt, nämlich den Schnittpunkt der beiden Geraden. Dieser Schnittpunkt ist in diesem Beispiel der Punkt $\binom{-3}{2}$.

(2) Wir suchen alle Lösungen des Gleichungssystems

$$\begin{aligned} x + 3y &= -1 \qquad (1)\\ -3x - 9y &= 2 \qquad (2). \end{aligned} \tag{3.3}$$

*Lösung:* Wenn $\binom{x}{y}$ eine Lösung des Gleichungssystems ist, dann gilt $x+3y = -1$. Daher gilt auch $3x + 9y = -3$. Somit ist jede Lösung des Systems (3.3) auch eine Lösung von

$$\begin{aligned} 3x + 9y &= -3 \qquad (1')\\ -3x - 9y &= 2 \qquad (2). \end{aligned} \tag{3.4}$$

Wenn $3x + 9y = -3$ und $-3x - 9y = 2$, dann gilt auch

$$(3x + 9y) + (-3x - 9y) = -3 + 2. \tag{3.5}$$

Die linke Seite von (3.5) ist aber immer 0. Jede Lösung $\binom{x}{y}$ des Gleichungssystems (3.3) muss also $0 = -3 + 2$, also $0 = -1$ erfüllen. Egal welche $x, y$ man in die Gleichung (3.5) einsetzt: die Gleichung (3.5) kann nie erfüllt sein.

Somit hat das Gleichungssystem (3.3) keine Lösung. Die Lösungsmenge ist also die leere Menge, also

$$L = \{\} = \emptyset.$$

Wir interpretieren dieses Beispiel jetzt geometrisch. Die Gerade $x + 3y = -1$ hat den Normalvektor $\binom{1}{3}$. Die Gerade $-3x-9y = 2$ hat den Normalvektor $\binom{-3}{-9}$. Der Vektor $\binom{-3}{-9}$ ist ein Vielfaches des Vektors $\binom{1}{3}$. Die beiden Geraden sind also parallel. Zwei parallele Geraden sind entweder identisch, oder sie haben keinen gemeinsamen Punkt. Da das Gleichungssystem (3.3) unlösbar ist, haben die beiden Geraden keinen gemeinsamen Punkt; sie sind also zwei verschiedene parallele Geraden.

(3) Wir suchen alle Lösungen des Gleichungssystems

$$\begin{aligned} x + 5y &= -4 \qquad (1)\\ -2x - 10y &= 8 \qquad (2). \end{aligned} \tag{3.6}$$

*Lösung:* Wenn $\binom{x}{y}$ eine Lösung des Gleichungssystems ist, dann gilt $x+5y = -4$. Daher gilt auch $2x + 10y = -8$. Somit ist jede Lösung des Systems (3.6) auch eine Lösung von

$$\begin{aligned} 2x + 10y &= -8 \qquad (1')\\ -2x - 10y &= 8 \qquad (2). \end{aligned} \tag{3.7}$$

Wenn $2x + 10y = -8$ und $-2x - 10y = 8$, dann gilt auch

$$(2x + 10y) + (-2x - 10y) = -8 + 8. \tag{3.8}$$

Sowohl die linke als auch die rechte Seite der Gleichung (3.8) ist also 0. Somit ist die Gleichung (3.8) für alle $\binom{x}{y}$ in $\mathbb{R}^2$ erfüllt. Sie liefert also keine Einschränkung für die Lösungen.

Nicht jeder Punkt $\binom{x}{y}$ in $\mathbb{R}^2$ ist eine Lösung des Systems (3.6). (Der Punkt $\binom{0}{0}$ erfüllt nicht einmal die erste Gleichung.) Wir sehen aber, dass jede Lösung der ersten Gleichung von (3.6) auch eine Lösung der zweiten Gleichung von (3.6) ist: das liegt daran, dass die zweite Gleichung entsteht, wenn man beide Seiten der ersten Gleichung mit $-2$ multipliziert. Man kann also die zweite Gleichung einfach weglassen (sie liefert keine weitere Einschränkung für $x$ und $y$), und nur mehr die Lösungen von

$$x + 5y = -4$$

bestimmen. Wir sehen, dass wir für jeden Wert, den wir für $y$ vorgeben, einen Wert für $x$ erhalten. Wenn wir für $y := t$ setzen, erhalten wir für $x$ den Wert $x = -4 - 5t$. Somit können wir die Lösungsmenge $L$ so angeben:

$$L = \{\binom{-4-5t}{t} \mid t \in \mathbb{R}\}.$$

Die Lösungsmenge $L$ ist also eine Gerade durch $\binom{-4}{0}$ mit Richtungsvektor $\binom{-5}{1}$.

Wir interpretieren dieses Beispiel jetzt geometrisch. Die Gleichungen $x + 5y = -4$ und $-2x - 10y = 8$ werden von denselben $\binom{x}{y}$ erfüllt. Sie beschreiben also die gleiche Gerade. Der Schnitt dieser beiden Geraden miteinander ist also genau diese eine Gerade. Und wirklich: $X = \binom{-4}{0} + t \cdot \binom{-5}{1}$ ist genau die Parameterdarstellung der Geraden $x + 5y = -4$.

(4) Wir suchen alle Lösungen des Gleichungssystems

$$\begin{aligned} 3x + 4y &= -1 && (1) \\ 5x + 10y &= -5 && (2) \\ 2x + 8y &= -6 && (3). \end{aligned} \tag{3.9}$$

*Lösung:* Wir multiplizieren die Gleichung (1) mit 5, und die Gleichung (2) mit $-3$ und erhalten

$$\begin{aligned} 15x + 20y &= -5 && (1') \\ -15x - 30y &= 15 && (2') \\ 2x + 8y &= -6 && (3). \end{aligned} \tag{3.10}$$

Jede Lösung von (1′) und (2′) erfüllt auch

$$(15x + 20y) + (-15x - 30y) = -5 + 15,$$

also $-10y = 10$, und somit $y = -1$. Wenn $y = -1$, dann muss wegen der Gleichung (1) gelten:

$$3x = -1 - 4y,$$

also $3x = -1 + 4$, und somit $x = 1$.

Die Frage ist, ob $\binom{1}{-1}$ auch wirklich eine Lösung des Gleichungssystems ist. Wir haben bis jetzt ja nur so begründet, dass für jede Lösung $\binom{x}{y}$ des Gleichungssystems $x = 1$ und $y = -1$ gelten muss. Wir wissen aber noch nicht, ob $\binom{1}{-1}$ die Gleichungen (1), (2) und (3) erfüllt. So haben wir etwa die Gleichung (3) beim Ausrechnen von $x$ und $y$ noch gar nicht verwendet! Wir müssen also ausprobieren, ob $\binom{1}{-1}$ wirklich alle drei Gleichungen erfüllt. Es gilt $3 \cdot 1 + 4 \cdot (-1) = -1$, $5 \cdot 1 + 10 \cdot (-1) = -5$ und $2 \cdot 1 + 8 \cdot (-1) = -6$. Das Paar $\binom{1}{-1}$ erfüllt also wirklich alle drei Gleichungen, und es gilt somit

$$L = \{\tbinom{1}{-1}\}.$$

Wir interpretieren dieses Beispiel jetzt geometrisch. Von den drei Geraden, die durch die Gleichungen (1), (2) und (3) beschrieben werden, sind keine zwei parallel, da kein Normalvektor ein Vielfaches eines anderen Normalvektors ist. Alle drei Geraden gehen durch den Punkt $\binom{1}{-1}$. Die drei Geraden gehen also "sternförmig" durch den Punkt $\binom{1}{-1}$.

Das Gleichungssystem (3.6) zeigt, dass die Lösungsmenge $L$ eines linearen Gleichungssystems nicht leer oder einelementig sein muss, sondern auch eine unendliche Menge sein kann. Für die Darstellung der Lösungsmenge $L$ des Systems (3.6) gibt es zwei Möglichkeiten:

(1) *Implizite Darstellung* der Lösung: Jedes Paar $(x, y)$, das $x + 5y = -4$ erfüllt, ist auch eine Lösung für das gesamte Gleichungssystem. Die Lösung kann also in der Form
$$L = \{(x, y) \in \mathbb{R}^2 \mid x + 5y = -4\}$$
geschrieben werden.

(2) *Parametrisierte Darstellung* der Lösung: Wir können die Lösungsmenge als
$$L = \{(-4, 0) + t \cdot (-5, 1) \mid t \in \mathbb{R}\}$$
schreiben.

Wollen wir nun überprüfen, ob das Paar $(3, 4)$ in der Lösungsmenge liegt, so müssen wir bei impliziter Darstellung nur $x = 3$ und $y = 4$ in $x + 5y$ einsetzen. Da wir dabei 23 und nicht $-4$ erhalten, folgt $(3, 4) \notin L$. Bei parametrisierter Darstellung müssen wir dazu das Gleichungssystem

$$\begin{aligned} -4 - 5t &= 3 \\ t &= 4 \end{aligned}$$

lösen. Aus der Tatsache, dass dieses System keine Lösung besitzt, können wir $(3, 4) \notin L$ schließen.

Die implizite Darstellung lässt jedoch keine direkte geometrische Interpretation zu, während man aus der parametrisierten Darstellung sofort erkennt, dass es sich bei der Lösungsmenge um eine Gerade im $\mathbb{R}^2$ mit der Steigung $-\frac{1}{5}$ handelt.

Auch andere Kurven im $\mathbb{R}^2$ lassen sich sowohl implizit als auch parametrisiert darstellen. So ist zum Beispiel der Kreis mit Radius 1 um den Ursprung gegeben durch

$$\begin{aligned} K &= \{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\} \\ &= \{(\cos t, \sin t) \,|\, t \in \mathbb{R},\ 0 \le t < 2\pi\}. \end{aligned}$$

Wir werden uns überlegen, wie wir die Lösungsmenge von einer Darstellungsform in die andere umrechnen können. Die jeweiligen Übergänge nennt man *Parametrisieren* (*Lösen*) bzw. *Implizitisieren*.

## 2. Die Lösung von Gleichungssystemen in Staffelform

Wir betrachten das Gleichungssystem

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 0 & 1 & -2 & -4 & 2 \\ 0 & 0 & 0 & -16 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} -9 \\ 5 \\ 16 \\ 0 \end{pmatrix}. \tag{3.11}$$

Dieses Gleichungssystem können wir so lösen: $x_5$ können wir beliebig festlegen. Wir setzen also

$$x_5 = t.$$

Wir erhalten

$$-16\,x_4 + 8\,t = 16,$$

also

$$x_4 = -1 + \frac{1}{2}\,t.$$

Da $x_3$ frei wählbar ist, setzen wir $x_3$ auf $s$.

Dann erhalten wir

$$x_2 = 5 + 2\,x_3 + 4\,x_4 - 2\,x_5,$$

also

$$\begin{aligned} x_2 &= 5 + 2\,s - 4 + 2\,t - 2\,t \\ x_2 &= 1 + 2\,s. \end{aligned}$$

Schließlich

$$x_1 = -2 + 2\,s + t.$$

Also ergibt sich als Lösungsmenge

$$L = \{\left(\begin{smallmatrix} -2 \\ 1 \\ 0 \\ -1 \\ 0 \end{smallmatrix}\right) + s \cdot \left(\begin{smallmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{smallmatrix}\right) + t \cdot \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \\ \frac{1}{2} \\ 1 \end{smallmatrix}\right) \,|\, s, t \in \mathbb{R}\}.$$

Die Matrix

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 0 & 1 & -2 & -4 & 2 \\ 0 & 0 & 0 & -16 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ist eine besonders angenehme Koeffizientenmatrix. Sie ist nämlich in *Zeilenstaffelform*. Wir definieren:

DEFINITION 3.1. Sei $A$ eine $m \times n$-Matrix. $A$ ist in *Zeilenstaffelform*, wenn es $r \in \mathbb{N}_0$ und $j_1, j_2, \ldots, j_r \in \{1, \ldots, n\}$ gibt, sodass

(1) $j_r > j_{r-1} > \cdots > j_1$.
(2) Für alle $i \in \{1, 2, \ldots, r\}$ gilt: $A(i, j_i) \neq 0$.
(3) Für alle $i \in \{1, 2, \ldots, r\}$ und für alle $k \in \{1, 2, \ldots, n\}$ mit $k < j_i$ gilt: $A(i, k) = 0$.
(4) Für alle $i \in \{1, 2, \ldots, m\}$ mit $i > r$ und für alle $k \in \{1, 2, \ldots, n\}$ gilt: $A(i, k) = 0$.

Wenn $A$ eine $m \times n$-Matrix in Zeilenstaffelform ist, und $r$ wie in obiger Definition, dann treten in der Lösung von $A \cdot x = 0$ genau $n - r$ frei wählbare Parameter auf.

ÜBUNGSAUFGABEN 3.2.

(1) Bestimmen Sie die Lösungsmenge des linearen Gleichungssystems

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

Geben Sie die Lösungsmenge in parametrisierter Form an.

(2) Bestimmen Sie die Lösungsmenge des linearen Gleichungssystems

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

Geben Sie die Lösungsmenge in parametrisierter Form an.

(3) Bestimmen Sie alle Lösungen des folgenden Gleichungssystems:

$$\begin{aligned} 2\,x_1 - 1\,x_2 + 1\,x_3 - 1\,x_4 + 1\,x_5 &= 12 \\ 1\,x_2 + 1\,x_3 - 1\,x_4 + 1\,x_5 &= 4 \\ 1\,x_3 + 0\,x_4 - 1\,x_5 &= 14. \end{aligned}$$

(4) Lösen Sie das Gleichungssystem

$$\begin{pmatrix} 2 & 5 & 7 & 3 & 2 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix},$$

und geben Sie die Lösungsmenge parametrisiert an.

## 3. Das Gaußsche Eliminationsverfahren

Wenn die Koeffizientenmatrix eines Gleichungssystems Zeilenstaffelform hat, dann wissen wir bereits, wie wir alle Lösungen des Gleichungssystems bestimmen. Wir erklären mit einem Beispiel, wie wir sonst vorgehen. Wir betrachten das Gleichungssystem

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix} \cdot x = \begin{pmatrix} -9 \\ -4 \\ 0 \\ -18 \end{pmatrix} \tag{3.12}$$

Wir schreiben uns zunächst das System anders auf.

$$\begin{array}{rrrrrrr} I & 1 & -5 & 8 & 2 & -2 & -9 \\ II & 1 & -4 & 6 & -2 & 0 & -4 \\ III & -1 & 0 & 2 & 2 & 0 & 0 \\ IV & 5 & -8 & 6 & 0 & -5 & -18 \\ \hline \end{array}$$

Wir addieren nun passende Vielfache der Gleichung I zu jeder der anderen Gleichungen, sodass in den neuen Gleichungen die Variable $x_1$ nicht mehr vorkommt. Das führt auf

$$\begin{array}{rrrrrrrl} II' & 0 & 1 & -2 & -4 & 2 & 5 & -I + II \\ III' & 0 & -5 & 10 & 4 & -2 & -9 & I + III \\ IV' & 0 & 17 & -34 & -10 & 5 & 27 & (-5) \cdot I + IV \\ \hline \end{array} \tag{3.13}$$

Nun hat das Gleichungssystem, das aus den Gleichungen I,II,III,IV besteht, die gleiche Lösungsmenge wie das Gleichungssystem, das aus den Gleichungen I,II',III',IV' besteht.

Das kann man so begründen: wenn ein Tupel $(x_1, x_2, x_3, x_4, x_5)$ die Gleichungen I und IV erfüllt, dann erfüllt es auch die Gleichung IV', die ja eine Linearkombination (= Summer von Vielfachen) der Gleichungen I und IV ist. Sei nun $(x_1, x_2, x_3, x_4, x_5)$ ein Tupel, das die Gleichungen I und IV' erfüllt. Da IV' = −5· I + IV, gilt IV = IV' + 5· I. Daher ist die Gleichung IV eine Linearkombination der Gleichungen IV' und I. Also muss das Tupel $(x_1, x_2, x_3, x_4, x_5)$ auch die Gleichung IV erfüllen.

Wir können also anstelle des Gleichungssystems I,II,III,IV das Gleichungssystem I,II',III',IV' lösen. In den Gleichungen II', III', IV' kommt die Variable $x_1$ nicht vor. Also können wir so vorgehen: wir bestimmen die Lösungen $(x_2, x_3, x_4, x_5)$ der Gleichungen II', III', IV'. Für jede dieser Lösungen rechnen wir uns dann aus der Gleichung I den Wert von $x_1$ aus.

Um die Lösungen von II', III', IV' zu bestimmen, addieren wir wieder passende Vielfache der Gleichung II' zu III' und IV'. Wir erhalten

$$\begin{array}{rrrrrrrl} III'' & 0 & 0 & 0 & -16 & 8 & 16 & 5 \cdot II' + III' \\ IV'' & 0 & 0 & 0 & 58 & -29 & -58 & (-17) \cdot II' + IV' \\ \hline \end{array} \tag{3.14}$$

Nun können wir alle Lösungen $(x_3, x_4, x_5)$ der Gleichungen III” und IV” bestimmen. Dann können wir für jede Lösung aus II’ den Wert von $x_2$ bestimmen (und dann aus I den Wert von $x_1$).

In den Gleichungen III” und IV” kommt $x_3$ nicht vor. Wenn wir also eine Lösung $(x_4, x_5)$ für die Gleichungen III” und IV” finden, dann können wir für $x_3$ jede beliebige Zahl einsetzen. Für jede solche Setzung erhalten wir eine Lösung $(x_3, x_4, x_5)$ von III” und IV”. Wir merken uns:

$x_3$ ist frei wählbar,

sofern es Lösungen $(x_4, x_5)$ für III” und IV” gibt. Jetzt versuchen wir, alle Lösungen $(x_4, x_5)$ von III” und IV” zu finden. Dazu addieren wir ein passendes Vielfaches der Gleichung III” zur Gleichung IV”. Wir erhalten

$$\begin{array}{lcccccccl} IV''' & 0 & 0 & 0 & 0 & 0 & & 0 & 58 \cdot III'' + 16 \cdot IV'' \\ \hline \end{array} \tag{3.15}$$

Alle Lösungen $(x_5)$ dieser letzten Gleichung zu finden, ist einfach: wir können jeden Wert für $x_5$ einsetzen. Also:

$x_5$ ist frei wählbar.

Setzen wir also $x_5$ auf $t$, und schauen wir, welche Werte sich für die anderen Variablen ergeben.

Aus der Gleichung III” erhalten wir

$$-16\, x_4 + 8\, t = 16,$$

also

$$x_4 = -1 + \frac{1}{2}\, t.$$

Da $x_3$ frei wählbar ist, setzen wir $x_3$ auf $s$.

Aus der Gleichung II’ erhalten wir

$$x_2 = 5 + 2\, x_3 + 4\, x_4 - 2\, x_5,$$

also

$$\begin{aligned} x_2 &= 5 + 2\, s - 4 + 2\, t - 2\, t \\ x_2 &= 1 + 2\, s. \end{aligned}$$

Aus der Gleichung I erhalten wir schließlich

$$x_1 = -2 + 2\, s + t.$$

Also ergibt sich als Lösungsmenge

$$L = \{\begin{pmatrix} -2 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \frac{1}{2} \\ 1 \end{pmatrix} \mid s, t \in \mathbb{R}\}.$$

## ÜBUNGSAUFGABEN 3.3.

(1) Bestimmen Sie alle Lösungen des Gleichungssystems

$$\begin{aligned} x + 2y + 3z &= -10 \\ -x + 7y + 2z &= -10 \\ 5x - 8y + 5z &= -10. \end{aligned}$$

(2) Bestimmen Sie alle Lösungen des Gleichungssystems

$$\begin{aligned} -4x + 2y + 3z &= 12 \\ -6x + 3y + 0z &= -18 \\ 6x - 3y + 2z &= 34 \end{aligned}$$

(3) Bestimmen Sie alle Lösungen des Gleichungssystems

$$\begin{aligned} 1x + 0y + 2z &= 16 \\ 2x + 3y - z &= -8 \\ 0x + 2y - 3z &= -36 \end{aligned}$$

(4) Geben Sie die Lösungsmenge des folgenden linearen Gleichungssystems in parametrisierter Form an!

$$\begin{pmatrix} -2 & -1 & 0 & 8 \\ 2 & -3 & 8 & 0 \\ 10 & -7 & 24 & -16 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -32 \\ 16 \\ 112 \end{pmatrix}.$$

(5) Geben Sie die Lösungsmenge des folgenden linearen Gleichungssystems in parametrisierter Form an!

$$\begin{pmatrix} -2 & -1 & 8 & 0 \\ 2 & -3 & 0 & 8 \\ 10 & -7 & -16 & -24 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -16 \\ 8 \\ 56 \end{pmatrix}$$

(6) Ergänzen Sie die Gleichung

$$3x - 2y + 5z = 0$$

so zu einem Gleichungssystem mit drei Gleichungen, dass das System
  (a) keine Lösung
  (b) genau eine Lösung
  (c) genau zwei Lösungen
  (d) eine Gerade als Lösungsmenge
  (e) eine Ebene als Lösungsmenge

hat.

(7) Für zwei Goldbarren und acht Silbertaler erhalten Sie 69.000.– Schilling, für 7 Barren und 3 Taler 84.000.– Schilling. Wieviel ist ein Goldbarren wert? Wieviel ist ein Silbertaler wert?

(8) Bestimmen Sie alle Lösungen des Gleichungssystems

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & 10 & 8 \\ 4 & 20 & 15 \\ 1 & 6 & 5 \end{pmatrix} \cdot x = \begin{pmatrix} 16 \\ 34 \\ 67 \\ 21 \end{pmatrix}.$$

(9) Bestimmen Sie alle Lösungen des Gleichungssystems

$$\begin{pmatrix} 1 & 3 & 5 & 2 \\ 2 & 8 & 14 & 4 \\ -1 & -2 & -3 & -2 \end{pmatrix} \cdot x = \begin{pmatrix} 3 \\ 6 \\ -2 \end{pmatrix}.$$

(10) Lösen Sie folgendes Gleichungssystem, und geben Sie die Lösungsmenge paramtrisiert an!

$$\begin{pmatrix} -2 & 1 & 0 & 1 \\ -3 & 2 & 2 & 0 \\ -6 & 5 & 8 & -3 \end{pmatrix} \cdot x = \begin{pmatrix} -8 \\ -13 \\ -28 \end{pmatrix}.$$

## 4. Einige durchgerechnete Beispiele zum Gauß-Algorithmus

---

AUFGABE 3.4.

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix} \cdot x = \begin{pmatrix} -9 \\ -4 \\ 0 \\ -18 \end{pmatrix}$$

*In[40]:=* **<< GaussDemo6.m**

*In[41]:=* **Gauss [A1, b1]**

$$\begin{array}{c} \begin{array}{cccccccc} x1 & x2 & x3 & x4 & x5 & | & \\ \end{array} \\ \begin{pmatrix} 1 & -5 & 8 & 2 & -2 & | & -9 \\ 1 & -4 & 6 & -2 & 0 & | & -4 \\ -1 & 0 & 2 & 2 & 0 & | & 0 \\ 5 & -8 & 6 & 0 & -5 & | & -18 \end{pmatrix} \end{array}$$

We use equation (1) of the last system to eliminate x1

$$\begin{array}{c} \begin{array}{cccccc} x2 & x3 & x4 & x5 & | & \\ \end{array} \\ \begin{pmatrix} 1 & -2 & -4 & 2 & | & 5 \\ -5 & 10 & 4 & -2 & | & -9 \\ 17 & -34 & -10 & 5 & | & 27 \end{pmatrix} \end{array}$$

We use equation (1) of the last system to eliminate x2

$$\begin{pmatrix} x3 & x4 & x5 & | & \\ 0 & -16 & 8 & | & 16 \\ 0 & 58 & -29 & | & -58 \end{pmatrix}$$

x3 does not appear in any equation.

$$\begin{pmatrix} x4 & x5 & | & \\ -16 & 8 & | & 16 \\ 58 & -29 & | & -58 \end{pmatrix}$$

We use equation (1) of the last system to eliminate x4

$$\begin{pmatrix} x5 & | & \\ 0 & | & 0 \end{pmatrix}$$

x5 does not appear in any equation.

x5 := t1

We use

$$\begin{pmatrix} x4 & x5 & | & \\ -16 & 8 & | & 16 \end{pmatrix}$$

to compute x4

$x4 = -1 + \frac{t1}{2}$

x3 := t2

We use

$$\begin{pmatrix} x2 & x3 & x4 & x5 & | & \\ 1 & -2 & -4 & 2 & | & 5 \end{pmatrix}$$

to compute x2

x2 = 1 + 2 t2

We use

$$\begin{pmatrix} x1 & x2 & x3 & x4 & x5 & | & \\ 1 & -5 & 8 & 2 & -2 & | & -9 \end{pmatrix}$$

to compute x1

x1 = −2 + t1 + 2 t2

*Out[41]=* $\{\{-2, 1, 0, -1, 0\}, \{\{1, 0, 0, \frac{1}{2}, 1\}, \{2, 2, 1, 0, 0\}\}\}$

---

AUFGABE 3.5.

$$\begin{pmatrix} 1 & 5 & 7 & 9 \\ 2 & 10 & 14 & 18 \\ 3 & 2 & 1 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

*In[42]:=* **<< GaussDemo6.m**

*In[43]:=* **Gauss[A2, b2]**

$$\begin{pmatrix} x1 & x2 & x3 & x4 & | & \\ 1 & 5 & 7 & 9 & | & 0 \\ 2 & 10 & 14 & 18 & | & 1 \\ 3 & 2 & 1 & 0 & | & 1 \end{pmatrix}$$

We use equation (1) of the last system to eliminate x1

$$\begin{pmatrix} x2 & x3 & x4 & | & \\ 0 & 0 & 0 & | & 1 \\ -13 & -20 & -27 & | & 1 \end{pmatrix}$$

We use equation (2) of the last system to eliminate x2

$$\begin{pmatrix} x3 & x4 & | & \\ 0 & 0 & | & 1 \end{pmatrix}$$

x3 does not appear in any equation.

$$\begin{pmatrix} x4 & | & \\ 0 & | & 1 \end{pmatrix}$$

x4 does not appear in any equation.

The system has no solution

*Out[43]=* NOSOLUTION

---

AUFGABE 3.6.

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & 10 & 8 \\ 4 & 20 & 15 \\ 1 & 6 & 5 \end{pmatrix} \cdot x = \begin{pmatrix} 16 \\ 34 \\ 67 \\ 21 \end{pmatrix}$$

*In[44]:=* **<< GaussDemo6.m**

*In[45]:=* **Gauss[A3, b3]**

```
   x1  x2  x3   |
 ( 1   5   3    |  16 )
 ( 2   10  8    |  34 )
 ( 4   20  15   |  67 )
   1   6   5    |  21
```

We use equation (1) of the last system to eliminate x1

```
   x2  x3  |
 ( 0   2   |  2 )
 ( 0   3   |  3 )
   1   2   |  5
```

We use equation (3) of the last system to eliminate x2

```
 ( x3  |     )
 ( 2   |  2  )
 ( 3   |  3  )
```

We use equation (1) of the last system to eliminate x3

We use

```
 ( x3  |     )
 ( 2   |  2  )
```

to compute x3

x3 = 1

We use

```
 ( x2  x3  |     )
 ( 1   2   |  5  )
```

to compute x2

x2 = 3

We use

```
 ( x1  x2  x3  |      )
 ( 1   5   3   |  16  )
```

to compute x1

x1 = -2

*Out[45]=* {{-2, 3, 1}, {}}

---

AUFGABE 3.7.

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & 10 & 8 \\ 4 & 20 & 15 \\ 1 & 6 & 5 \end{pmatrix} \cdot x = \begin{pmatrix} 16 \\ 34 \\ 68 \\ 22 \end{pmatrix}$$

*In[46]:=* **<< GaussDemo6.m**

*In[47]:=* **Gauss [A4, b4]**

$$\begin{array}{ccccc} \text{x1} & \text{x2} & \text{x3} & | & \\ 1 & 5 & 3 & | & 16 \\ 2 & 10 & 8 & | & 34 \\ 4 & 20 & 15 & | & 68 \\ 1 & 6 & 5 & | & 22 \end{array}$$

We use equation (1) of the last system to eliminate x1

$$\begin{array}{cccc} \text{x2} & \text{x3} & | & \\ 0 & 2 & | & 2 \\ 0 & 3 & | & 4 \\ 1 & 2 & | & 6 \end{array}$$

We use equation (3) of the last system to eliminate x2

$$\begin{array}{ccc} \text{x3} & | & \\ 2 & | & 2 \\ 3 & | & 4 \end{array}$$

We use equation (1) of the last system to eliminate x3

The system has no solution

*Out[47]=* NOSOLUTION

---

AUFGABE 3.8.

$$\begin{pmatrix} 1 & 2 & -3 & 0 & 6 & 7 & 0 \\ -2 & -4 & 1 & 0 & 2 & 3 & 0 \\ 4 & 8 & 1 & 0 & 1 & -1 & 0 \\ 3 & 6 & -1 & 0 & 9 & 9 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} 5 \\ 7 \\ 3 \\ 15 \end{pmatrix}$$

*In[48]:=* **<< GaussDemo6.m**

*In[49]:=* **Gauss[A5, b5]**

$$\begin{array}{sccccccccc} \text{x1} & \text{x2} & \text{x3} & \text{x4} & \text{x5} & \text{x6} & \text{x7} & | & \\ 1 & 2 & -3 & 0 & 6 & 7 & 0 & | & 5 \\ -2 & -4 & 1 & 0 & 2 & 3 & 0 & | & 7 \\ 4 & 8 & 1 & 0 & 1 & -1 & 0 & | & 3 \\ 3 & 6 & -1 & 0 & 9 & 9 & 0 & | & 15 \end{array}$$

We use equation (1) of the last system to eliminate x1

$$\begin{array}{cccccccc} x2 & x3 & x4 & x5 & x6 & x7 & | & \\ \end{array}$$
$$\left(\begin{array}{cccccccc} 0 & -5 & 0 & 14 & 17 & 0 & | & 17 \\ 0 & 13 & 0 & -23 & -29 & 0 & | & -17 \\ 0 & 8 & 0 & -9 & -12 & 0 & | & 0 \end{array}\right)$$

x2 does not appear in any equation.

$$\begin{array}{ccccccc} x3 & x4 & x5 & x6 & x7 & | & \\ \end{array}$$
$$\left(\begin{array}{ccccccc} -5 & 0 & 14 & 17 & 0 & | & 17 \\ 13 & 0 & -23 & -29 & 0 & | & -17 \\ 8 & 0 & -9 & -12 & 0 & | & 0 \end{array}\right)$$

We use equation (1) of the last system to eliminate x3

$$\left(\begin{array}{cccccc} x4 & x5 & x6 & x7 & | & \\ 0 & 67 & 76 & 0 & | & 136 \\ 0 & 67 & 76 & 0 & | & 136 \end{array}\right)$$

x4 does not appear in any equation.

$$\left(\begin{array}{ccccc} x5 & x6 & x7 & | & \\ 67 & 76 & 0 & | & 136 \\ 67 & 76 & 0 & | & 136 \end{array}\right)$$

We use equation (1) of the last system to eliminate x5

$$\left(\begin{array}{cccc} x6 & x7 & | & \\ 0 & 0 & | & 0 \end{array}\right)$$

x6 does not appear in any equation.

$$\left(\begin{array}{ccc} x7 & | & \\ 0 & | & 0 \end{array}\right)$$

x7 does not appear in any equation.

x7 := t1

x6 := t2

We use

$$\left(\begin{array}{ccccc} x5 & x6 & x7 & | & \\ 67 & 76 & 0 & | & 136 \end{array}\right)$$

to compute x5

$$x5 = \frac{136}{67} + -\frac{76\,t2}{67}$$

x4 := t3

We use

$$\left(\begin{array}{ccccccc} x3 & x4 & x5 & x6 & x7 & | & \\ -5 & 0 & 14 & 17 & 0 & | & 17 \end{array}\right)$$

to compute x3

$$x3 = \frac{153}{67} + \frac{15\,t2}{67}$$

x2 := t4

We use

$$\begin{pmatrix} x1 & x2 & x3 & x4 & x5 & x6 & x7 & | & \\ 1 & 2 & -3 & 0 & 6 & 7 & 0 & | & 5 \end{pmatrix}$$

to compute x1

$x1 = -\frac{22}{67} + \frac{32\,t2}{67} - 2\,t4$

*Out[49]=* $\left\{\left\{-\frac{22}{67}, 0, \frac{153}{67}, 0, \frac{136}{67}, 0, 0\right\},\right.$

$\left\{\{0, 0, 0, 0, 0, 0, 1\}, \left\{\frac{32}{67}, 0, \frac{15}{67}, 0, -\frac{76}{67}, 1, 0\right\},\right.$

$\left.\{0, 0, 0, 1, 0, 0, 0\}, \{-2, 1, 0, 0, 0, 0, 0\}\right\}\right\}$

# Teil 2

# Mengen und Funktionen

KAPITEL 4

# Mengen

## 1. Eigenschaften von Mengen

Unter einer *Menge* stellen wir uns eine Zusammenfassung von Objekten zu einem Ganzen vor. Wenn das Objekt $a$ zur Menge $M$ gehört, schreiben wir

$$a \in M.$$

Zwei Mengen $A, B$ sehen wir als gleich an, wenn sie dieselben Elemente enthalten. Diese Eigenschaft nennt man das *Extensionalitätsaxiom* oder *Axiom der Umfangsbestimmtheit*.

Der Begriff *Menge* wird nicht präzise mathematisch definiert. Vielmehr gibt man einige Eigenschaften an, die Mengen unserer Vorstellung nach erfüllen sollen. Als Grundlage fast aller Teilgebiete der Mathematik haben sich jene Eigenschaften bewährt, die Ernst Zermelo, Abraham Adolf Fraenkel und Thoralf Skolem von 1907 bis 1929 von Mengen gefordert haben. Diese Eigenschaften sind die *Axiome der Zermelo-Fraenkel-Mengenlehre*. Die Zermelo-Fraenkel-Mengenlehre hat sich als ausgezeichnet geeignet herausgestellt, fast alle Resultate der Mathematik klar darzustellen und zu vermitteln. Zum anderen sind die Axiome auch theoretisch – in der Logik und der Mengenlehre – gut untersucht worden, und man hat bis heute keinen Widerspruch in ihnen gefunden. Wir werden nicht alle Axiome diskutieren, aber zumindest das Extensionalitätsaxiom explizit angeben.

AXIOM 4.1 (Extensionalitätsaxiom). *Seien $A, B$ Mengen. Dann gilt $A = B$ genau dann, wenn für alle $x \in A$ auch $x \in B$ gilt, und für alle $x \in B$ auch $x \in A$ gilt.*

Insbesondere gilt $\{1, 2, 3\} = \{3, 1, 2\} = \{3, 3, 1, 2, 2\}$.

DEFINITION 4.2 (Teilmenge). Seien $A, B$ Mengen. Dann gilt $A \subseteq B$ genau dann, wenn für alle $a \in A$ auch $a \in B$ gilt.

## 2. Operationen auf Mengen

DEFINITION 4.3. Seien $A, B$ Mengen. Dann definieren wir:

$$\begin{aligned} A \cap B &:= \{x \mid x \in A \text{ und } x \in B\}, \\ A \cup B &:= \{x \mid x \in A \text{ oder } x \in B\}. \end{aligned}$$

$A \cap B$ ist der *Durchschnitt* von $A$ und $B$. $A \cup B$ ist die *Vereinigung* von $A$ und $B$.

SATZ 4.4. *Seien $A, B, C$ Mengen. Dann gilt:*

(1) $A \cap B = B \cap A$,
(2) $A \cup B = B \cup A$,
(3) $A \cup (B \cup C) = (A \cup B) \cup C$,
(4) $A \cap (B \cap C) = (A \cap B) \cap C$,
(5) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$,
(6) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

Wir beweisen hier nur die Eigenschaft (5), geben “dafür” aber zwei Beweise an.

*Beweis I von Satz 4.4:* Es gilt

$$\begin{aligned}(A \cap B) \cup C &= \{x \,|\, x \in A \cap B \text{ oder } x \in C\} \\ &= \{x \,|\, (x \in A \wedge x \in B) \vee x \in C\}.\end{aligned}$$

Außerdem gilt

$$\begin{aligned}(A \cup C) \cap (B \cup C) &= \{x \,|\, x \in A \cup C \wedge x \in B \cup C\} \\ &= \{x \,|\, (x \in A \vee x \in C) \wedge (x \in B \vee x \in C)\}.\end{aligned}$$

Wir müssen also nachweisen, dass für alle $x$ die Eigenschaft $(x \in A \wedge x \in B) \vee x \in C$ genau dann gilt, wenn $(x \in A \vee x \in C) \wedge (x \in B \vee x \in C)$. Sei dazu $x$ fixiert. Wir beobachten, dass beide Aussagen aus den gleichen drei Aussagen $a := (x \in A)$, $b := (x \in B)$, und $c := (x \in C)$ zusammengebaut sind. Wir brauchen also nur 8 Fälle zu untersuchen, je nach dem ob $a, b, c$ jeweils wahr oder falsch sind.

| $a$ | $b$ | $c$ | $a \vee c$ | $b \vee c$ | $(a \vee c) \wedge (b \vee c)$ | $(a \wedge b)$ | $(a \wedge b) \vee c$ |
|---|---|---|---|---|---|---|---|
| $f$ | $f$ | $f$ | $f$ | $f$ | $f$ | $f$ | $f$ |
| $f$ | $f$ | $w$ | $w$ | $w$ | $w$ | $f$ | $w$ |
| $f$ | $w$ | $f$ | $f$ | $w$ | $f$ | $f$ | $f$ |
| $f$ | $w$ | $w$ | $w$ | $w$ | $w$ | $f$ | $w$ |
| $w$ | $f$ | $f$ | $w$ | $f$ | $f$ | $f$ | $f$ |
| $w$ | $f$ | $w$ | $w$ | $w$ | $w$ | $f$ | $w$ |
| $w$ | $w$ | $f$ | $w$ | $w$ | $w$ | $w$ | $w$ |
| $w$ | $w$ | $w$ | $w$ | $w$ | $w$ | $w$ | $w$ |

Wir sehen, dass $(a \vee c) \wedge (b \vee c) \equiv (a \wedge b) \vee c$. Somit gilt $(x \in A \wedge x \in B) \vee x \in C$ genau dann, wenn $(x \in A \vee x \in C) \wedge (x \in B \vee x \in C)$. Also gilt $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.

*Beweis II von Satz 4.4:* Um zu zeigen, dass

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C), \tag{4.1}$$

zeigen wir, dass $(A \cap B) \cup C \subseteq (A \cup C) \cap (B \cup C)$ und dass $(A \cap B) \cup C \supseteq (A \cup C) \cap (B \cup C)$.
“$\subseteq$”: Sei $x \in (A \cap B) \cup C$. Wir wollen zeigen, dass $x \in (A \cup C) \cap (B \cup C)$.

Dazu zeigen wir zuerst, dass $x \in A \cup C$. Da $x \in (A \cap B) \cup C$, wissen wir, dass $x \in A \cap B$ oder $x \in C$.

- *1.Fall:* $x \in A \cap B$*:* Da dann $x \in A \cap B$, gilt $x \in A$, und somit $x \in A \cup C$.
- *2.Fall:* $x \in C$*:* Dann liegt $x$ in $A \cup C$.

Somit liegt $x$ also in $A \cup C$.

Nun zeigen wir, dass $x \in B \cup C$. Da $x \in (A \cap B) \cup C$, wissen wir, dass $x$ in $A \cap B$ oder in $C$ liegt.

- *1.Fall:* $x \in A \cap B$*:* Dann gilt $x \in B$, und somit $x \in B \cup C$.
- *2.Fall:* $x \in C$*:* Dann liegt $x$ in $B \cup C$.

Somit liegt $x$ also in $B \cup C$.

Daher liegt $x$ also sowohl in $A \cup C$ als auch in $B \cup C$, und somit gilt $x \in (A \cup C) \cap (B \cup C)$.

“$\supseteq$”: Sei $x \in (A \cup C) \cap (B \cup C)$. Wir wollen zeigen, dass $x \in (A \cap B) \cup C$. Wir betrachten dazu zwei Fälle.

- *1. Fall:* $x \in C$*:* Dann liegt $x$ auch in $(A \cap B) \cup C$.
- *2. Fall:* $x \notin C$*:* Da $x$ in $(A \cup C) \cap (B \cup C)$ liegt, gilt $x \in A \cup C$. Da $x \notin C$, gilt also $x \in A$. Da $x$ in $(A \cup C) \cap (B \cup C)$ liegt, gilt auch $x \in B \cup C$. Da $x \notin C$, gilt also $x \in B$. Insgesamt gilt also $x \in A \cap B$, und somit $x \in (A \cap B) \cup C$.

■

DEFINITION 4.5. Seien $A, B$ Mengen. Dann ist $B \setminus A$ definiert durch

$$B \setminus A := \{x \mid x \in B \text{ und } x \notin A\}.$$

Dabei steht $x \notin A$ für (nicht $x \in A$).

Wenn man nur Mengen betrachtet, die Teilmengen einer Menge $U$, des *Universums*, sind, so schreibt man auch $\complement_U B$ für $U \setminus B$. Es gelten etwa folgende Zusammenhänge:

SATZ 4.6. *Seien $A, B, C, U$ Mengen, sodass $A, B, C$ Teilmengen von $U$ sind. Dann gilt:*

(1) $B \setminus A = B \cap (\complement_U A)$,
(2) $C \setminus (B \setminus A) = (A \cap C) \cup (C \setminus B)$,
(3) *(De Morgansche Regeln)* $\complement_U(A \cap B) = \complement_U(A) \cup \complement_U(B)$ *und* $\complement_U(A \cup B) = \complement_U(A) \cap \complement_U(B)$.

*Beweis von* (2)*:* Es gilt

$$\begin{aligned}
x \in (C \setminus (B \setminus A)) &\Leftrightarrow x \in C \wedge (x \notin B \setminus A) \\
&\Leftrightarrow x \in C \wedge \neg(x \in B \wedge x \notin A) \\
&\Leftrightarrow x \in C \wedge (x \notin B \vee x \in A) \\
&\Leftrightarrow (x \in C \wedge x \notin B) \vee (x \in C \wedge x \in A) \\
&\Leftrightarrow x \in C \setminus B \vee x \in C \cap A \\
&\Leftrightarrow x \in (C \setminus B) \cup (C \cap A).
\end{aligned}$$

Mit $\emptyset$ bezeichnen wir die leere Menge, die kein Element enthält. Für jede Menge $A$ gilt $\emptyset \subseteq A$ und $A \setminus A = \emptyset$.

DEFINITION 4.7. Seien $A, B$ Mengen. Mit $A \Delta B$ bezeichnen wir die Menge

$$A \Delta B := (A \setminus B) \cup (B \setminus A).$$

$A \Delta B$ heißt die *symmetrische Differenz* von $A$ und $B$.

SATZ 4.8. *Seien $A, B, C$ Mengen. Es gilt:*

(1) $A \Delta B = (A \cup B) \setminus (A \cap B)$.
(2) $(A \Delta B) \Delta C = A \Delta (B \Delta C)$.
(3) $A \Delta B = B \Delta A$.
(4) $A \Delta \emptyset = A$.
(5) $A \Delta A = \emptyset$.
(6) $(A \Delta B) \cap C = (A \cap C) \Delta (B \cap C)$.

DEFINITION 4.9. Sei $A$ eine Menge. Dann ist die Menge $\mathcal{P}(A)$, die *Potenzmenge* von $A$, definiert durch

$$\mathcal{P}(A) := \{B \mid B \subseteq A\}.$$

Es gilt etwa $\mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

SATZ 4.10. *Sei $n \in \mathbb{N}$, und sei $A = \{1, 2, \dots, n\}$. Dann hat $\mathcal{P}(A)$ genau $2^n$ Elemente.*

*Beweis:* Wir beweisen diesen Satz mit Induktion. Sei $n = 1$. Dann gilt $A = \{1\}$ und $\mathcal{P}(A) = \{\emptyset, \{1\}\}$. Die Menge $\mathcal{P}(A)$ hat also genau 2 Elemente. Wir nehmen nun an, dass $n \in \mathbb{N}$, und dass $\mathcal{P}(\{1, \dots, n\})$ genau $2^n$ Elemente hat. Wir berechnen nun $\mathcal{P}(\{1, \dots, n + 1\})$. Es gilt

$$\begin{aligned} \mathcal{P}(\{1, \dots, n, n+1\}) &= \{B \mid B \subseteq \{1, \dots, n+1\} \wedge n+1 \in B\} \\ &\quad \cup \{B \mid B \subseteq \{1, \dots, n+1\} \wedge n+1 \notin B\} \\ &= \{C \cup \{n+1\} \mid C \subseteq \{1, \dots, n\}\} \\ &\quad \cup \{C \mid C \subseteq \{1, \dots, n\}\}. \end{aligned}$$

Die Menge $\{C \mid C \subseteq \{1, \dots, n\}\}$ ist die Menge $\mathcal{P}(\{1, \dots, n\})$. Nach der Induktionsannahme hat diese Menge $2^n$ Elemente. Da verschiedene Teilmengen von $\{1, \dots, n\}$ durch Hinzufügen von $n+1$ verschieden bleiben, hat auch $\{C \cup \{n+1\} \mid C \subseteq \{1, \dots, n\}\}$ genau $2^n$ Elemente. Schließlich haben $\{C \cup \{n+1\} \mid C \subseteq \{1, \dots, n\}\}$ und $\{C \mid C \subseteq \{1, \dots, n\}\}$ keine Elemente gemeinsam, es gilt daher

$$\{C \cup \{n+1\} \mid C \subseteq \{1, \dots, n\}\} \cap \{C \mid C \subseteq \{1, \dots, n\}\} = \emptyset.$$

Also hat ihre Vereinigung $2^n + 2^n = 2^{n+1}$ Elemente. Insgesamt haben wir gezeigt, dass $\mathcal{P}(\{1, \dots, n + 1\})$ genau $2^{n+1}$ Elemente hat; damit ist auch der Induktionsschritt gelungen. ■

Wenn eine Menge $A$ genau $n$ verschiedene Elemente hat (mit $n \in \mathbb{N}_0$), so schreiben wir $|A| = n$ (oder $\#A = n$). Eine solche Menge nennen wir auch $n$-elementig, und wir sagen, $n$ ist die *Kardinalität* der Menge. Eine Menge, für die es kein solches $n \in \mathbb{N}_0$ gibt, bezeichnen wir als *unendlich*. Manchmal schreibt man $|A| = \infty$. Eine genauere Unterscheidung zwischen “verschieden großen” unendlichen Mengen werden wir in Kapitel 11 vornehmen.

Wir halten nun die Kardinalität einiger endlichen Mengen fest.

SATZ 4.11. *Seien $A, B, A_1, \ldots, A_n$ endliche Mengen. Dann gilt:*

(1) *Wenn $A \cap B = \emptyset$, so gilt $|A \cup B| = |A| + |B|$.*
(2) *Wenn für alle $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ gilt, dass $A_i \cap A_j = \emptyset$, so gilt: $|A_1 \cup A_2 \cup \ldots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n|$.*
(3) *Es gilt $|A \cup B| = |A| + |B| - |A \cap B|$.*

*Beweis:* Wir beweisen hier nur die Eigenschaft (2). Da $A \setminus B$ und $B$ disjunkt sind, und $A \cup B = (A \setminus B) \cup B$, gilt

$$|A \cup B| = |A \setminus B| + |B|. \tag{4.2}$$

Nun gilt $(A \cap B) \cup (A \setminus B) = A$ und $(A \cap B) \cap (A \setminus B) = \emptyset$. Folglich gilt $|A \cap B| + |A \setminus B| = |A|$. Wenn man nun in Gleichung (4.2) $|A \setminus B|$ durch $|A| - |A \cap B|$ ersetzt, so erhält man $|A \cup B| = |A| + |B| - |A \cap B|$. ■

SATZ 4.12. *Sei $A$ eine Menge mit $n$ Elementen, und sei $i \in \{0, \ldots, n\}$. Dann hat $A$ genau $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ Teilmengen mit $i$ Elementen.*

*Beweis:* Wir verwenden, dass für $n \in \mathbb{N}$ und $i \in \{1, \ldots, n-1\}$ gilt: $\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1}$. Nun zeigen wir den Satz mit Induktion nach $n$. Für $n = 1$ sehen wir, dass eine einelementige Menge genau eine nullelementige und eine einelementige Teilmenge hat. Sei nun $n \geq 2$. Für $i = 0$ sehen wir, dass $A$ genau eine nullelementige Teilmenge, nämlich $\emptyset$, hat. Für $i = n$ hat $A$ genau eine $n$-elementige Teilmenge, nämlich $A$. Sei nun $i \in \{1, \ldots, n-1\}$. Wir wählen ein Element $a$ aus $A$. Es gilt

$$\begin{aligned}\{B \mid B \subseteq A, |B| = i\} &= \{B \mid B \subseteq A, |B| = i, a \notin B\} \cup \{B \mid B \subseteq A, |B| = i, a \in B\} \\ &= \{B \mid B \subseteq A \setminus \{a\}, |B| = i\} \cup \{B \cup \{a\} \mid B \subseteq A \setminus \{a\}, |B| = i-1\}.\end{aligned}$$

Nun hat die erste dieser Mengen $\binom{n-1}{i}$ und die zweite $\binom{n-1}{i-1}$ Elemente. Somit gilt $|\{B \mid B \subseteq A, |B| = i\}| = \binom{n-1}{i-1} + \binom{n-1}{i} = \binom{n}{i}$. ■

Es gibt Mengen, deren Elemente alle Mengen sind. Wir geben jetzt der Vereinigung aller Elemente einer Menge und dem Durchschnitt aller Elemente einer Menge eine Abkürzung.

DEFINITION 4.13. Sei $\mathcal{A}$ eine Menge, deren Elemente alle Mengen sind. Mit $\bigcup \mathcal{A}$ oder $\bigcup\limits_{A \in \mathcal{A}} A$ bezeichnen wir die Menge, die durch

$$\bigcup \mathcal{A} = \{x \mid \exists A \in \mathcal{A} : x \in A\}$$

definiert ist.

Beispiel: $\bigcup\{\{1,2\},\{2,5\},\{1,5,\emptyset\}\} = \{\emptyset,1,2,5\}$, $\bigcup\{]n,n{+}1[\,|\,n\in\mathbb{N}\} = \bigcup\limits_{n\in\mathbb{N}} ]n,n+1[\ = \{x\in\mathbb{R}^+\,|\,x\geq 1\}\setminus\mathbb{N}$. Weiters gilt $\bigcup\emptyset = \emptyset$.

DEFINITION 4.14. Sei $\mathcal{A}$ eine nichtleere Menge, deren Elemente alle Mengen sind. Mit $\bigcap\mathcal{A}$ oder $\bigcap\limits_{A\in\mathcal{A}} A$ bezeichnen wir die Menge, die durch

$$\bigcap\mathcal{A} = \{x\,|\,\forall A\in\mathcal{A} : x\in A\}$$

definiert ist.

Beispiel: $\bigcap\{\{1,2\},\{2,5\},\{2,6\}\} = \{2\}$, $\bigcap\{]n,n+1[\,|\,n\in\mathbb{N}\} = \bigcap\limits_{n\in\mathbb{N}} ]n,n+1[\ = \emptyset$.

KAPITEL 5

# Relationen und Funktionen

## 1. Geordnete Paare

Als erstes definieren wir *geordnete Paare*.

DEFINITION 5.1. Für beliebige $a, b$ definieren wir das *geordnete Paar* $(a, b)$ durch

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

SATZ 5.2. *Für alle $a, b, c, d$ gilt $(a, b) = (c, d)$ genau dann, wenn $a = c$ und $b = d$.*

*Beweis:* Wenn $a = c$ und $b = d$, so gilt $(a, b) = \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} = (c, d)$.

Wir zeigen nun, dass aus $(a, b) = (c, d)$ folgt, dass $a = c$ und $b = d$. Seien dazu $a, b, c, d$ so, dass $(a, b) = (c, d)$. Wir wissen also, dass $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$.

- *1.Fall: $a \neq b$:* Wir wissen, dass $\{c\} \in \{\{a\}, \{a, b\}\}$. Da $\{a, b\}$ nicht einelementig ist, kann nur $\{c\} = \{a\}$ gelten. Dann gilt $a = c$. Somit gilt also $\{\{a\}, \{a, b\}\} = \{\{a\}, \{a, d\}\}$. Da $\{a, b\} \neq \{a\}$, muss $\{a, b\} = \{a, d\}$ gelten. Somit gilt $b \in \{a, d\}$, und folglich $b = d$.
- *2.Fall: $a = b$:* Dann gilt $\{\{a\}\} = \{\{c\}, \{c, d\}\}$. Also gilt $\{c, d\} = \{a\}$, und somit $c = d = a$. Also gilt $a = c$ und $d = b$.

DEFINITION 5.3. Seien $A, B$ Mengen. Wir definieren $A \times B$, das *kartesische Produkt von A und B*, durch

$$A \times B := \{(a, b) \mid a \in A \text{ und } b \in B\}.$$

SATZ 5.4. *Seien $A, B, X, Y$ Mengen. Dann gilt*

(1) $(A \cup B) \times X = (A \times X) \cup (B \times X)$,
(2) $(A \cap B) \times (X \cap Y) = (A \times X) \cap (B \times Y)$,
(3) $\emptyset \times B = \emptyset$.
(4) *Wenn $A \times B = \emptyset$, so gilt $A = \emptyset$ oder $B = \emptyset$.*

## 2. Relationen

DEFINITION 5.5. Seien $A, B$ Mengen. Jede Teilmenge von $A \times B$ heißt auch *Relation von A nach B.*

Beispiele: Sei $A := \{W,N,O,S,T,V,B,St,K\}$ die Menge der neun österreichischen Bundesländer, und sei $B := \{\text{Donau}, \text{Inn}, \text{Traun}\}$. Wir definieren eine Relation $R$ durch

$$R := \{(a,b) \in A \times B \mid a \text{ besitzt einen Teil des Ufers von } b\}.$$

Wir erhalten $R = \{(W, \text{Donau}), (N, \text{Donau}), (O, \text{Donau}), (T, \text{Inn}), (O, \text{Inn}), (St, \text{Traun}), (O, \text{Traun})\}$. Für $(a,b) \in R$ schreiben wir auch $a\,R\,b$.

Sei $A := \mathbb{R}$ und $B := \mathbb{Z}$. Wir definieren eine Relation $\rho$ durch

$$a\,\rho\,b :\Leftrightarrow a \in [b, b+1[$$

für $a \in A$, $b \in B$. Dann gilt zum Beispiel $(\pi, 3) \in \rho$, $(\sqrt{2}, 1) \in \rho$. Es gilt also $\rho = \{(r,n) \in \mathbb{R} \times \mathbb{N} \mid n \leq r < n+1\}$.

Sei nun $A := \mathbb{N}$ und $B := \mathbb{N}$. Wir definieren eine Relation $K$ durch

$$(a,b) \in K :\Leftrightarrow \exists c \in \mathbb{N}_0 : a + c = b$$

für $a \in A$, $b \in B$. Wir sehen, dass $K = \{(x,y) \in \mathbb{N} \times \mathbb{N} \mid x \leq y\}$. $K$ ist also die "kleinergleich"-Relation.

Nun definieren wir eine Relation $\equiv_5$ von $\mathbb{Z}$ nach $\mathbb{Z}$ durch

$$a \equiv_5 b :\Leftrightarrow \exists c \in \mathbb{Z} : 5 \cdot c = b - a.$$

Es gilt also

$$\equiv_5 \; = \; \{(a,b) \in \mathbb{Z} \times \mathbb{Z} \mid b - a \text{ ist Vielfaches von } 5\}.$$

## 3. Funktionen

DEFINITION 5.6. Seien $A, B$ Mengen, und sei $R$ eine Relation von $A$ nach $B$. $R$ ist eine *funktionale Relation von A nach B*, wenn es für alle $a \in A$ genau ein $b \in B$ gibt, sodass $(a,b) \in R$.

Beispiele: Seien $A := \{1,2,3\}$, $B := \{a,b,c\}$, $R := \{(1,a), (2,c), (3,c)\}$. Dann ist $R$ eine funktionale Relation von $A$ nach $B$.

Sei $A := \mathbb{R}$, $B := \mathbb{R}$, $f := \{(r, \sin(r)) \mid r \in \mathbb{R}\}$. Dann ist $f$ eine funktionale Relation von $\mathbb{R}$ nach $\mathbb{R}$.

Sei $A := \mathbb{R}$, $B := \mathbb{R}$, $g := \{(\sin(r), r) \mid r \in \mathbb{R}\}$. Dann ist $g$ keine funktionale Relation von $\mathbb{R}$ nach $\mathbb{R}$, da es kein $y \in \mathbb{R}$ gibt, sodass $(-2, y) \in \mathbb{R}$.

Sei $A := [-1, 1]$, $B := \mathbb{R}$, $h := \{(\sin(r), r) \mid r \in \mathbb{R}\}$. Dann ist $h$ keine funktionale Relation von $A$ nach $\mathbb{R}$, da $(0,0) \in h$ und $(0, \pi) \in h$. Somit gibt es für $a := 0$ mehr als ein $b \in \mathbb{R}$, sodass $(a,b) \in h$.

DEFINITION 5.7. Seien $A, B$ Mengen, und sei $f$ eine funktionale Relation von $A$ nach $B$. Für $a \in A$ bezeichnen wir mit $f(a)$ dann jenes $b \in B$, für das $(a,b) \in f$.

Funktionale Relationen von $A$ nach $B$ bezeichnen wir auch einfach als *Funktionen von A nach B*. Funktionen kann man auf verschiedene Arten angeben. Wir betrachten einige gebräuchliche Varianten für die Quadratfunktion $q$ auf den ganzen Zahlen.

(1) Direkt als Menge: $q := \{(x, x^2) \mid x \in \mathbb{Z}\}$. Die Menge kann natürlich auch anders angegeben werden, zum Beispiel durch $q := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid y = x^2\}$.

(2) Durch eine Zuordnungsvorschrift:

$$\begin{array}{cccc} q : & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & x & \longmapsto & x^2 \end{array}.$$

Man liest das als "$q$ ist eine Funktion von $\mathbb{Z}$ nach $\mathbb{Z}$, die jedes $x$ in $\mathbb{Z}$ auf $x^2$ abbildet.

(3) Durch Angabe des Funktionswerts, also etwa so: $q : \mathbb{Z} \to \mathbb{Z}$, $q(z) := z^2$ für $z \in \mathbb{Z}$. (Lies:"$q$ ist eine Funktion von $\mathbb{Z}$ nach $\mathbb{Z}$, und $q(z)$ ist gleich $z^2$ für alle $z \in \mathbb{Z}$."

Egal, welche der drei Varianten man wählt: $q$ ist dadurch jedesmal als die gleiche Teilmenge von $\mathbb{Z} \times \mathbb{Z}$ definiert. Die Schreibweise $f : A \to B$ bedeutet *$f$ ist eine Funktion von A nach B,* also einfach *$f$ ist eine funktionale Relation von A nach B.*

DEFINITION 5.8 (Einschränkung). Seien $A, B$ Mengen, sei $T$ eine Teilmenge von $A$, und sei $f$ eine Funktion von $A$ nach $B$. Mit $f|_T$ bezeichnen wir die Funktion, die durch

$$\begin{array}{cccc} f : & T & \longrightarrow & B \\ & t & \longmapsto & f(t) \end{array}$$

gegeben ist. Sie heißt *Einschränkung von $f$ auf $T$*.

Es gilt also $f|_T = \{(x, y) \in f \mid x \in T\}$.

DEFINITION 5.9. Seien $A, B$ Mengen. Mit $B^A$ bezeichnet man die Menge aller Funktion von $A$ nach $B$. Genauer:

$$B^A := \{f \in \mathcal{P}(A \times B) \mid f : A \to B\}.$$

SATZ 5.10. *Seien $n, m \in \mathbb{N}$, und sei $A = \{1, \ldots, m\}$, $B := \{1, \ldots, n\}$. Dann hat $B^A$ genau $n^m$ Elemente.*

## 4. Definitions- und Wertebereich

DEFINITION 5.11. Seien $A, B$ Mengen, und sei $f$ eine Funktion von $A$ nach $B$. Dann ist $A$ auch der *Definitionsbereich* von $f$. Der *Wertebereich* von $f$ ist die Menge $\{f(a) \mid a \in A\}$.

Der Wertebereich einer Funktion enthält also jene Elemente in $B$, die tatsächlich als Funktionswert auftreten.

DEFINITION 5.12. Seien $A, B$ Mengen, und sei $f$ eine Funktion von $A$ nach $B$. Sei $T \subseteq A$. Dann bezeichnen wir mit $f[T]$ das *Bild von $T$ unter $f$*, das wir mit

$$f[T] = \{f(t) \mid t \in T\}$$

definieren.

Wenn keine Verwechslungen möglich sind, so schreibt man auch $f(T)$ anstelle von $f[T]$. Für die Sinusfunktion sin von $\mathbb{R}$ nach $\mathbb{R}$ ist der Wertebereich also das Intervall $[-1, 1]$. Außerdem gilt $\sin[\{n \cdot \frac{\pi}{4} \mid n \in \mathbb{N}\}] = \{-1, -\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}, 1\}$.

SATZ 5.13. *Seien $A, B$ Mengen, und sei $f$ eine Funktion von $A$ nach $B$. Seien $C, D \subseteq A$. Dann gilt*

(1) $f(C \cup D) = f(C) \cup f(D)$,
(2) $f(C \cap D) \subseteq f(C) \cap f(D)$.

DEFINITION 5.14. Seien $A, B$ Mengen, und sei $f$ eine Funktion von $A$ nach $B$.

(1) Die Funktion $f$ ist *injektiv*, wenn

$$\forall x, y \in A : f(x) = f(y) \Rightarrow x = y$$

gilt.
(2) Die Funktion $f$ ist *surjektiv auf $B$*, wenn es für alle $b \in B$ ein $a \in A$ gibt, sodass $f(a) = b$.
(3) Die Funktion $f$ ist *bijektiv von $A$ nach $B$*, wenn sie injektiv und surjektiv ist.

Die Funktion $f$ ist also injektiv, wenn es kein $x, y \in A$ mit $x \neq y$ und $f(x) = f(y)$ gibt.

SATZ 5.15. *Seien $A, B$ Mengen, und sei $f$ eine Funktion von $A$ nach $B$. Sei*

$$g := \{(b, a) \in B \times A \mid (a, b) \in f\}.$$

*Dann sind äquivalent:*

(1) *$g$ ist eine Funktion von $B$ nach $A$.*
(2) *$f$ ist eine bijektive Funktion von $A$ nach $B$.*

*Beweis:* (2)⇒(1): Sei $b \in B$. Wir zeigen, dass es genau ein $a \in A$ gibt, sodass $(b, a) \in g$. Da $f$ bijektiv ist, gibt es ein $a \in A$ mit $f(a) = b$, also mit $(a, b) \in f$. Dann gilt $(b, a) \in g$, und wir haben ein geeignetes $a \in A$ gefunden. Wir zeigen nun, dass es höchstens ein $a \in A$ mit $(b, a) \in g$ gibt. Seien $a_1, a_2 \in A$ so, dass $(b, a_1) \in g$ und $(b, a_2) \in g$. Dann gilt $(a_1, b) \in f$ und $(a_2, b) \in f$, also $b = f(a_1) = f(a_2)$. Da $f$ injektiv ist, gilt $a_1 = a_2$. (1)⇒(2): Wir zeigen als erstes, dass $f$ injektiv ist. Seien $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$. Also gilt $(a_1, f(a_1)) \in f$ und $(a_2, f(a_2)) \in f$, und somit $(f(a_1), a_1) \in g$ und $(f(a_2), a_2) \in g$. Da $g$ eine Funktion ist, muss wegen $f(a_1) = f(a_2)$ also auch $a_1 = a_2$ gelten. Somit ist $f$ injektiv. Wir zeigen nun, dass $f$ surjektiv ist. Sei dazu $b \in B$. Da $g$ eine Funktion ist, gibt es ein $a \in A$ mit $(b, a) \in g$. Somit gilt $(a, b) \in f$, und folglich $f(a) = b$. Also liegt $b$ im Wertebereich von $f$. Somit ist $f$ surjektiv. ■

DEFINITION 5.16. Seien $A, B$ Mengen, und sei $f$ eine bijektive Funktion von $A$ nach $B$. Die Funktion $g : B \to A$ mit $g = \{(b, a) \in B \times A \mid f(a) = b\}$ heißt *die zu $f$ inverse Funktion*, und wird mit $f^{-1}$ abgekürzt.

Die gleiche Schreibweise, $f^{-1}$, verwendet man auch für etwas anderes:

DEFINITION 5.17. Seien $A, B$ Mengen, und sei $f$ eine Funktion von $A$ nach $B$. Sei $D \subseteq B$. Dann bezeichnet man mit $f^{-1}[D]$ (oder $f^{-1}(D)$) die Menge, die durch

$$f^{-1}[D] := \{a \in A \mid f(a) \in D\}$$

gegeben ist, und man nennt $f^{-1}[D]$ das *Urbild von $D$ unter $f$*.

## 5. Familien und Folgen

Wir können es bestimmt nicht besser formulieren als P. Halmos [**Halmos, 1976**, S. 48].

> Gelegentlich wird der Wertebereich einer Funktion für wichtiger gehalten als die Funktion selbst. In einem solchen Falle werden Terminologie und Notation stark verändert. Sei zum Beispiel $x$ eine Funktion von einer Menge $I$ in eine Menge $X$. [... ] Wir wollen jetzt ein Element des Definitionsbereiches $I$ einen *Index* und $I$ selbst die *Indexmenge* nennen; der Wertebereich der Funktion $x$ soll *indizierte Menge* und die Funktion selbst *Familie* heißen; der Wert der Funktion an einer Stelle $i$, *Term* der Familie genannt, wird (anstelle von $x(i)$) nun $x_i$ geschrieben.

DEFINITION 5.18. Seien $I, X$ Mengen, und sei $x$ eine Funktion von $I$ nach $X$. Wir schreiben $x_i$ für $x(i)$. Wir definieren nun $\langle x_i \mid i \in I\rangle$ durch

$$\langle x_i \mid i \in I\rangle := \{(i, x_i) \mid i \in I\}.$$

Es gilt also $x = \langle x_i \mid i \in I\rangle$.

Für $\langle x_i \mid i \in I\rangle$ schreibt man auch $(x_i)_{i\in I}$. Etwas allgemeiner schreibt man für eine Funktion $f$ von $A$ nach $B$ und eine Teilmenge $C$ von $A$ auch

$$\langle f(c) \mid c \in C\rangle \text{ oder } (f(c))_{c\in C}$$

für die Menge $\{(c, f(c)) \mid c \in C\}$.

DEFINITION 5.19. Sei $A$ eine Menge, sei $n \in \mathbb{N}$, und seien $a_1, \dots, a_n \in A$. Mit dem *$n$-Tupel* $\langle a_1, \dots, a_n\rangle$ meinen wir die Familie $\langle a_i \mid i \in \{1, \dots, n\}\rangle$.

Ein $n$-Tupel $\langle a_1, \dots, a_n\rangle$ sehen wir also als eine mit der Indexmenge $\{1, \dots, n\}$ indizierte Familie an. Das $n$-Tupel $\langle a_1, \dots, a_n\rangle$ schreiben wir auch als $(a_1, \dots, a_n)$; dass jetzt $(a, b)$ zwei verschiedene (aber in der Praxis sehr ähnliche) Bedeutungen haben kann, stört meist nicht.

DEFINITION 5.20. Sei $A$ eine Menge, und sei $n \in \mathbb{N}$. Mit $A^n$ bezeichnen wir die Menge aller $n$-Tupel aus $A$, also

$$A^n := \{\langle a_1, \ldots, a_n\rangle \mid a_1, \ldots, a_n \in A\} = \{f \mid f : \{1, \ldots, n\} \to A\}.$$

DEFINITION 5.21. Sei $(X_i)_{i\in I}$ eine mit $I$ indizierte Familie von Mengen. Dann ist

$$\begin{aligned}\prod_{i\in I} X_i &:= \{x : I \to \bigcup\{X_i \mid i \in I\} \ \mid\ \forall i \in I : x(i) \in X_i\} \\ &= \{(x_i)_{i\in I} \mid (x_i)_{i\in I} \text{ ist eine Familie mit } \forall i \in I : x_i \in X_i\}.\end{aligned}$$

Wenn alle $X_i$ die gleiche Menge $X$ sind, erhält man $\prod_{i\in I} X_i = X^I$. Die Menge der reellen Zahlenfolgen ist also zum Beispiel genau die Menge $\mathbb{R}^{\mathbb{N}}$.

Jetzt können wir noch ein Axiom der Mengenlehre angeben, das nicht aus den anderen Axiomen der Mengenlehre folgt. Es hat so überraschende Konsequenzen, dass man seine Verwendung, im Unterschied zur Verwendung der anderen Axiome der Mengenlehre, manchmal explizit macht, und etwa schreibt: “unter Verwendung des Auswahlaxioms gilt”.

AXIOM 5.22 (Auswahlaxiom). *Sei $I$ eine Menge, und sei $(X_i)_{i\in I}$ eine Familie von Mengen. Wir nehmen an, dass für alle $i \in I$ die Menge $X_i$ nicht leer ist. Dann ist auch $\prod_{i\in I} X_i$ nicht leer.*

In einer anderen Formulierung:

> Sei $(X_i)_{i\in I}$ eine Familie von nichtleeren Mengen. Dann gibt es eine Funktion $f$ mit Definitionsbereich $I$, sodass für alle $i \in I : f(i) \in X_i$ gilt.

Ein solches $f$ heißt auch *Auswahlfunktion*; daher der Name *Auswahlaxiom*.

Im Jahr 1937 zeigte K. Gödel: wenn die üblichen Axiome der Mengenlehre widerspruchsfrei sind, so sind auch die Axiome zusammen mit dem Auswahlaxiom widerspruchsfrei. Das Auswahlaxiom bringt also keine “neuen” Widersprüche.

Im Jahr 1963 zeigte P. Cohen, dass man auch das Gegenteil des Auswahlaxioms, also die Existenz einer Familie nichtleerer Mengen, für die es keine Auswahlfunktion gibt, annehmen kann, ohne dadurch neue Widersprüche zu erhalten. Wenn also die üblichen Axiome der Mengenlehre widerspruchsfrei sind, so sind auch die Axiome zusammen mit der Negation des Auswahlaxioms widerspruchsfrei.

Das Auswahlaxiom ist also unabhängig von den anderen Axiomen der Mengenlehre; seine Wahrheit wird von den anderen Axiomen nicht bestimmt. Legt man nur die üblichen Axiome der Mengenlehre zu Grunde, liegt das Auswahlaxiom also im “gesetzlich nicht geregelten Raum”.

Wir werden das Auswahlaxiom als gültig voraussetzen.

ÜBUNGSAUFGABEN 5.23.

(1) (Funktionen) Für eine Funktion $f : X \to Y$ und $A \subseteq X$ schreiben wir $f[A]$ für $\{f(a) \mid a \in X\}$. Für welche Funktionen gilt, dass für alle Teilmengen $A, B$ von $X$ die Menge $f[A \cap B]$ gleich $f[A] \cap f[B]$ ist?

## 6. Hintereinanderausführung von Funktionen

DEFINITION 5.24. Seien $A, B, C$ Mengen, sei $f$ eine Funktion von $A$ nach $B$, und sei $g$ eine Funktion von $B$ nach $C$. Wir definieren $g \circ f$ durch

$$\begin{array}{rcccl} g \circ f & : & A & \longrightarrow & C \\ & & a & \longmapsto & g(f(a)). \end{array}$$

Die Funktion $g \circ f$ heißt die *Hintereinanderausführung* oder *funktionale Komposition* von $f$ und $g$. Man spricht "$g$ nach $f$" für $g \circ f$.

SATZ 5.25 (Assoziativität der Hintereinanderausführung). *Seien $A, B, C, D$ Mengen, und sei $f : A \to B$, $g : B \to C$, $h : C \to D$. Dann gilt $(h \circ g) \circ f = h \circ (g \circ f)$.*

SATZ 5.26 (Hintereinanderausführung und inverse Funktion). *Seien $A, B$ Mengen, sei $f$ eine bijektive Funktion von $A$ nach $B$, und sei $f^{-1} := \{(b, a) \in B \times A \mid (a, b) \in f\}$ die zu $f$ inverse Funktion. Dann gilt $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.*

*Beweis:* Sei $a \in A$. Dann gilt $(a, f(a)) \in f$, und somit $(f(a), a) \in f^{-1}$. Also gilt $f^{-1}(f(a)) = a$. Sei nun $b \in B$, und sei $a \in A$ so, dass $f(a) = b$. Dann gilt $(a, b) \in f$ und somit $(b, a) \in f^{-1}$. Also gilt $b = f(a) = f(f^{-1}(b))$. ■

SATZ 5.27. *Seien $A, B, C$ Mengen, sei $f$ eine Funktion von $A$ nach $B$, und sei $g$ eine Funktion von $B$ nach $C$.*

(1) *Wenn $g \circ f$ surjektiv auf $C$ ist, so ist auch $g$ surjektiv auf $C$.*
(2) *Wenn $g \circ f$ injektiv ist, so ist auch $f$ injektiv.*

Mit $\mathrm{id}_A$ bezeichnen wir die Funktion von $A$ nach $A$ mit $\mathrm{id}_A(x) = x$ für alle $x \in A$.

SATZ 5.28. *Seien $A, B$ Mengen, sei $f$ eine Funktion von $A$ nach $B$, und seien $l, r$ Funktionen von $B$ nach $A$. Wenn $l \circ f = \mathrm{id}_A$ und $f \circ r = \mathrm{id}_B$, so ist $f$ bijektiv, und es gilt $l = r = f^{-1}$.*

*Beweis:* Nach Satz 5.27 ist $f$ bijektiv. Es gilt also $l = l \circ \mathrm{id}_B = l \circ (f \circ f^{-1}) = (l \circ f) \circ f^{-1} = \mathrm{id}_A \circ f^{-1} = f^{-1}$ und $r = \mathrm{id}_A \circ r = (f^{-1} \circ f) \circ r = f^{-1} \circ (f \circ r) = f^{-1} \circ \mathrm{id}_B = f^{-1}$. ■

SATZ 5.29. *Seien $A, B, C$ Mengen, sei $f$ eine bijektive Funktion von $A$ nach $B$, und sei $g$ eine bijektive Funktion von $B$ nach $C$. Dann ist $g \circ f$ eine bijektive Funktion von $A$ nach $C$, und es gilt $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

# Teil 3

# Vektorräume

KAPITEL 6

# Unterräume des $\mathbb{R}^n$

## 1. Die Definition eines Unterraums

Mengen, mit denen man so rechnen kann wie mit Vektoren, werden wir als *Vektorräume* bezeichnen. Wir behandeln zunächst den für die Geometrie wichtigsten Vektorraum, nämlich $\mathbb{R}^3$, und, in Verallgemeinerung davon, für jedes $n \in \mathbb{N}$ den Vektorraum

$$\mathbb{R}^n = \{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}.$$

Für einen Vektor $x \in \mathbb{R}^n$ und $i \in \{1, 2, \ldots, n\}$ schreiben wir $x(i)$, $x[i]$, und $x_i$ für den $i$-ten Eintrag von $x$.

Manche Teilmengen des $\mathbb{R}^n$ sind abgeschlossen bezüglich der Addition von Vektoren und der Multiplikation mit reellen Zahlen. Solche Teilmengen bezeichnen wir als *Unterräume* des $\mathbb{R}^n$.

DEFINITION 6.1. $T \subseteq \mathbb{R}^n$ ist *Unterraum* des $\mathbb{R}^n$ $:\Leftrightarrow$

(1) Die Menge $T$ enthält zumindest ein Element.
(2) Für alle $\lambda \in \mathbb{R}$ und für alle $t \in T$ gilt $\lambda t \in T$.
(3) Für alle $s, t \in T$ gilt $s + t \in T$.

Wir geben einige Beispiele von Unterräumen des $\mathbb{R}^n$ :

BEISPIELE 6.2.

(1) $T_1 = \{(x, y) \in \mathbb{R}^2 \mid 2x - 3y = 0\}$ ist Unterraum des $\mathbb{R}^2$. Begründung: Wegen $(0, 0) \in T_1$ ist die Menge $T_1$ nicht leer. Sei $(u, v) \in T_1$ und $\lambda \in \mathbb{R}$. Dann gilt $2u - 3v = 0$ und damit $2(\lambda u) - 3(\lambda v) = 0$, also gilt $\lambda \cdot (u, v) \in T_1$. Für $(u_1, v_1), (u_2, v_2) \in T_1$ gilt $2 \cdot (u_1 + u_2) - 3 \cdot (v_1 + v_2) = 2u_1 - 3v_1 + 2u_2 - 3v_2 = 0$, also ist $(u_1, v_1) + (u_2, v_2) \in T_1$.
    Damit haben wir gezeigt, dass $T_1$ ein Unterraum des $\mathbb{R}^2$ ist.
(2) $T_2 = \{(x, y) \in \mathbb{R}^2 \mid 3x + 2y = 1\}$ ist kein Unterraum des $\mathbb{R}^2$, denn $(1, -1) \in T_2$, aber $2 \cdot (1, -1) \notin T_2$.
(3) $T_3 = \{(x, y, z) \in \mathbb{R}^3 \mid$ es gibt $s, t \in \mathbb{R}$, sodass $(x, y, z) = s \cdot (1, -2, 4) + t \cdot (0, 1, 8)\}$ ist ein Unterraum des $\mathbb{R}^3$.

(4) $T_4 = \{(0,0)\}$ ist Unterraum des $\mathbb{R}^2$.
(5) $T_5 = \{(0,1)\}$ ist kein Unterraum des $\mathbb{R}^2$, da $2 \cdot (0,1) \notin T_5$.

ÜBUNGSAUFGABEN 6.3.

(1) Vervollständigen Sie die folgenden Begründungen dafür, dass die Menge

$$T = \{\binom{x}{y} \mid \text{ es gibt } \alpha \in \mathbb{R}, \text{ sodass } \binom{x}{y} = \alpha \cdot \binom{-2}{1}\}$$

die Unterraumeigenschaften (V1) und (V2) erfüllt.

(a) $T$ ist nicht die leere Menge, weil ________________________.

(b) *Für alle $\lambda \in \mathbb{R}$ und $t \in T$ liegt $\lambda \cdot t$ in $T$:* Wir fixieren $t$ aus $T$ und $\lambda \in \mathbb{R}$. Wir wollen zeigen, dass __________________ in ______ liegt. Da $t$ in $T$ liegt, gibt es ein $\alpha \in \mathbb{R}$, sodass $t = \alpha \cdot \binom{-2}{1}$. Um zu zeigen, dass $\lambda \cdot t$ in $T$ liegt, müssen wir ein $\alpha' \in \mathbb{R}$ finden, sodass

$$\lambda \cdot t = \alpha' \cdot \_\_\_\_\_\_\_\_\_\_.$$

Nun wissen wir, dass $t = \alpha \cdot \binom{-2}{1}$ ist. Daher gilt $\lambda \cdot t =$ __________________. Das heißt, dass für $\alpha' =$ ____________ gilt:

$$\lambda \cdot t = \alpha' \cdot \binom{-2}{1}.$$

Daher liegt auch __________________ in $T$.

(2) Welche der folgenden Mengen sind Unterräume des Vektorraumes $\mathbb{R}^2$? Geben Sie jeweils an, welche Unterraumeigenschaften erfüllt sind.
   (a) $\{\binom{x}{y} \in \mathbb{R}^2 \mid 3x + 2y = 0\}$.
   (b) $\{\binom{x}{y} \in \mathbb{R}^2 \mid 3x + 2y = 1\}$.

(3) Welche der folgenden Mengen sind Unterräume des Vektorraumes $\mathbb{R}^2$? Geben Sie jeweils an, welche Unterraumeigenschaften erfüllt sind.
   (a) $\{\binom{x}{y} \in \mathbb{R}^2 \mid \exists \lambda \in \mathbb{R} : \binom{x}{y} = \lambda \cdot \binom{1}{3}\}$.
   (b) $\{\binom{x}{y} \in \mathbb{R}^2 \mid \exists \lambda \in \mathbb{R} : \binom{x}{y} = \binom{-4}{-12} + \lambda \cdot \binom{1}{3}\}$.

(4) Welche der folgenden Mengen sind Unterräume des Vektorraumes $\mathbb{R}^2$? Geben Sie jeweils an, welche Unterraumeigenschaften erfüllt sind.
   (a) $\{\binom{x}{y} \in \mathbb{R}^2 \mid x + 3y \leq 0\}$.
   (b) $\{\binom{x}{y} \in \mathbb{R}^2 \mid x^4 + y^2 = 0\}$.

(5) Zeigen Sie: Wenn ein Unterraum des $\mathbb{R}^n$ ($\mathbb{R}^n$ als Vektorraum über $\mathbb{R}$) zwei Punkte enthält, so enthält er bereits die gesamte Verbindungsgerade.

Die Lösungsmenge eines linearen Gleichungssystems mit rechter Seite = 0 ist immer ein Unterraum.

SATZ 6.4. *Seien $m, n \in \mathbb{N}$, und sei $A$ eine $m \times n$-Matrix. Dann ist die Lösungsmenge des Gleichungssystems $A \cdot x = 0$ ein Unterraum des $\mathbb{R}^n$.*

*Beweis.* Sei $U = \{x \in \mathbb{R}^n \mid A \cdot x = 0\}$.

(1) Wegen $0 \in U$ ist $U$ nicht die leere Menge.
(2) Sei $x \in U, \lambda \in \mathbb{R}$. Dann gilt $0 = \lambda\,(A \cdot x) = A \cdot (\lambda\, x)$, also $\lambda\, x \in U$.
(3) Seien $u, v \in U$. Dann gilt $A \cdot (u + v) = A \cdot u + A \cdot v = 0$, d.h. $u + v \in U$.

Somit ist $U$ ein Unterraum des $\mathbb{R}^n$. ∎

AUFGABE 6.5. Wir bestimmen diesen Unterraum für die Matrix $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$. Als Lösungsmenge von $A \cdot x = 0$ erhalten wir

$$L = \{(-2\,t, -3\,t, t) \mid t \in \mathbb{R}\}.$$

## 2. Die lineare Hülle von Vektoren

DEFINITION 6.6. Sei $m \in \mathbb{N}_0$, und seien $v_1, \ldots, v_m \in \mathbb{R}^n$. Die Menge

$$L(v_1, \ldots, v_m) := \{\sum_{i=1}^{m} \lambda_i \cdot v_i \,|\, \lambda_1, \ldots, \lambda_m \in \mathbb{R}\}$$

heißt die *lineare Hülle* der Vektoren $v_1, \ldots, v_m$.

$L(\left(\begin{smallmatrix} 3 \\ 2 \end{smallmatrix}\right))$ ist also die Gerade im $\mathbb{R}^2$, die durch $\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 6 \\ 4 \end{smallmatrix}\right)$ geht. $L(\,)$ definiert man als $\{\vec{0}\}$.

Die lineare Hülle von $v_1, \ldots, v_m$ ist der kleinste Unterraum, der $v_1, \ldots, v_m$ enthält:

SATZ 6.7. *Sei $m \in \mathbb{N}_0$, $n \in \mathbb{N}$, und seien $v_1, \ldots, v_m$ Vektoren im $\mathbb{R}^n$. Dann gilt*

(1) *$L(v_1, \ldots, v_m)$ ist ein Unterraum des $\mathbb{R}^n$.*
(2) *Sei $M$ ein Unterraum der $v_1, \ldots, v_m$ enthält. Dann gilt $L(v_1, \ldots, v_m) \subseteq M$.*

Wollen wir etwa überprüfen, ob z.B. $(3, 0, 1)$ in der linearen Hülle von $(2, 1, -3)$ und $(7, 2, -5)$ liegt, müssen wir ein Gleichungssystem lösen:

$$\lambda_1 \cdot \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 7 \\ 2 \\ -5 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix},$$

$$\text{das heißt} \begin{pmatrix} 2 & 7 \\ 1 & 2 \\ -3 & -5 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Lösen wir dieses System, so erhalten wir $\lambda_1 = -2$ und $\lambda_2 = 1$. Also ist $(3, 0, 1)$ eine Linearkombination von $(2, 1, -3)$ und $(7, 2, -5)$, und liegt somit in der linearen Hülle dieser beiden Vektoren.

SATZ 6.8. *Seien $m, n \in \mathbb{N}$, seien $b_1, b_2, \ldots, b_m$ Vektoren im $\mathbb{R}^n$, und sei $\overline{B} = (b_1, b_2, \ldots, b_m)$ die Matrix mit den Vektoren $b_1, b_2, \ldots, b_m$ als Spaltenvektoren. Dann liegt $v$ genau dann in $L(b_1, \ldots, b_m)$, wenn das Gleichungssystem $\overline{B} \cdot x = v$ eine Lösung $x \in \mathbb{R}^m$ hat.*

ÜBUNGSAUFGABEN 6.9.

(1) Liegt $\left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right)$ in der linearen Hülle von $\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} -2 \\ -4 \end{smallmatrix}\right)$?
(2) Liegt $\left(\begin{smallmatrix} 1 \\ -5 \end{smallmatrix}\right)$ in der linearen Hülle von $\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 4 \end{smallmatrix}\right)$?
(3) Testen Sie, ob $\left(\begin{smallmatrix} 7 \\ -2 \\ -2 \end{smallmatrix}\right)$ in der linearen Hülle von $\left(\begin{smallmatrix} 1 \\ 0 \\ -2 \end{smallmatrix}\right), \left(\begin{smallmatrix} -2 \\ 0 \\ 4 \end{smallmatrix}\right), \left(\begin{smallmatrix} 3 \\ -1 \\ 0 \end{smallmatrix}\right))$ liegt.

Sei nun $n \in \mathbb{N}$ eine natürliche Zahl, und sei $M$ eine (möglicherweise unendliche) Teilmenge von $\mathbb{R}^n$. Wir definieren die lineare Hülle von $M$ als

$$L(M) = \{\sum_{i=1}^{k} \lambda_i \cdot m_i \,|\, k \in \mathbb{N}_0,\ \lambda_1, \ldots, \lambda_k \in \mathbb{R},\ m_1, \ldots, m_k \in M\}.$$

Die lineare Hülle von $M$ ist also die Menge aller Linearkombinationen endlich vieler Vektoren aus $M$.

DEFINITION 6.10. Für eine $m \times n$-Matrix $A$ mit Einträgen aus $\mathbb{R}$ definieren wir ihren *Zeilenraum* $Z(A)$ als die lineare Hülle der Zeilen von $A$. Der Zeilenraum ist ein Unterraum von $\mathbb{R}^n$.

Den *Spaltenraum* $S(A)$ definieren wir als die lineare Hülle der Spalten von $A$. Der Spaltenraum ist ein Unterraum von $\mathbb{R}^m$.

Den *Nullraum* $N(A)$ definieren wir als die Lösungsmenge des Gleichungssystems
$A \cdot x = 0$. Er ist ein Unterraum des $\mathbb{R}^n$.

### 3. Die lineare Unabhängigkeit von Vektoren

DEFINITION 6.11. Seien $m, n \in \mathbb{N}$, und seien $v_1, \dots, v_m$ in $R^n$. Die Folge $(v_1, \dots, v_m)$ heißt *linear unabhängig*, wenn für alle $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit

$$\sum_{i=1}^{m} \lambda_i \cdot v_i = \lambda_1 \cdot v_1 + \cdots + \lambda_m \cdot v_m = 0$$

gilt, dass alle $\lambda_i = 0$ sind.

Man sagt dann oft auch einfach, dass die Vektoren $v_1, \dots, v_m$ linear unabhängig sind. Als Spezialfall definiert man noch für $m = 0$, dass die Folge $()$ aus 0 Vektoren immer linear unabhängig ist.

Vektoren $v_1, \dots, v_m$, die nicht linear unabhängig sind, nennt man *linear abhängig*. Die Folge $(v_1, \dots, v_m)$ ist also genau dann linear abhängig, wenn es $(\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$ gibt, sodass $\sum_{i=1}^{m} \lambda_i \cdot v_i = 0$.

AUFGABE 6.12. Sind $(3, 2)$ und $(1, 3)$ linear unabhängig ?

*Lösung.* Wir betrachten

$$\lambda_1 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dieses System besitzt nur die Lösung $(0, 0)$. Daher sind die beiden Vektoren linear unabhängig.

AUFGABE 6.13. Sind $(3, 2)$, $(1, 4)$ und $(5, 3)$ linear unabhängig ?

*Lösung.* Hier erhalten wir

$$\lambda_1 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 4 & 3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eine Lösung des Gleichungssystems ist $\lambda_1 = -1.7$, $\lambda_2 = 0.1$ und $\lambda_3 = 1$. Die drei Vektoren sind also linear abhängig.

SATZ 6.14. *Seien $m, n \in \mathbb{N}$, seien $b_1, b_2, \dots, b_m$ Vektoren im $\mathbb{R}^n$, und sei $\overline{B} = (b_1, b_2, \dots, b_m)$ die Matrix mit den Vektoren $b_1, b_2, \dots, b_m$ als Spaltenvektoren. Dann sind $(b_1, b_2, \dots, b_m)$ genau dann linear abhängig, wenn das System $\overline{B} \cdot x = 0$ eine Lösung $x \neq 0$ hat.*

SATZ 6.15. *Sei $m \geq 1$ und seien $b_1, \dots, b_m \in \mathbb{R}^n$. Die folgenden zwei Aussagen sind äquivalent:*

(1) *$(b_1, \dots, b_m)$ ist linear abhängig.*
(2) *Es gibt ein $k \in \{1, \dots, m\}$, sodass $b_k$ in $L(b_1, \dots, b_{k-1})$ liegt.*

SATZ 6.16. *Sei $m \in \mathbb{N}_0$, sei $n \in \mathbb{N}$, und seien $v_1, \dots, v_m, v \in \mathbb{R}^n$. Wir nehmen an, dass $v_1, \dots, v_m$ linear unabhängig sind. Dann sind äquivalent:*

(1) *$v \in L(v_1, \dots, v_m)$.*
(2) *$(v_1, \dots, v_m, v)$ ist linear abhängig.*

ÜBUNGSAUFGABEN 6.17.

(1) Zeigen Sie, dass die Vektoren $\left(\begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 10 \\ -10 \end{pmatrix}\right)$ linear abhängig sind, indem Sie eine Linearkombination finden, bei der nicht jeder Vektor 0 mal genommen wird, und die trotzdem den Nullvektor ergibt.
(2) Testen Sie jeweils, ob folgende Mengen von Vektoren linear abhängig sind. Finden Sie, falls die Vektoren linear abhängig sind, eine Linearkombination, die den Nullvektor ergibt, und bei der nicht jeder Vektor 0 mal genommen wird.
   (a) $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix}$.
   (b) $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}$.
(3) Sind $\begin{pmatrix} 3 \\ 4 \\ -5 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 16 \\ 8 \\ -25 \end{pmatrix}$ linear abhängig?
(4) Sind die Vektoren $\left(\begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 10 \\ -8 \end{pmatrix}\right)$ linear abhängig?
(5) Geben Sie einen Vektor $v \in \mathbb{R}^2$ an, sodass $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ und $v$ linear abhängig sind.
(6) Finden Sie drei Vektoren $a, b, c \in \mathbb{R}^3$, sodass $(b, c)$ linear unabhängig und $(a, b, c)$ linear abhängig sind.
(7) Vervollständigen Sie die Begründung für folgende Aussage.

   Seien $v_1, v_2, w \in \mathbb{R}^n$ so, dass $w$ in der linearen Hülle von $v_1$ und $v_2$ liegt. Dann sind $(v_1, v_2, w)$ linear abhängig.

   *Begründung:* Da $w$ in der linearen Hülle von $v_1$ und $v_2$ liegt, gibt es $\lambda_1, \lambda_2 \in \mathbb{R}$, sodass

   $$w = \underline{\hspace{6cm}}.$$

   Daher gilt

   $$\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \underline{\hspace{3cm}} = 0.$$

   Das ist eine Linearkombination, die 0 ergibt, obwohl nicht jeder Vektor ____________ mal genommen wurde. Daher sind $(v_1, v_2, w)$ ________________.

## 4. Basen eines Vektorraums

### 4.1. Definition.

DEFINITION 6.18. Sei $T$ Unterraum von $\mathbb{R}^n$. Die Folge $B = (b_1, \dots, b_m)$ heißt Basis von $T$ :$\Leftrightarrow$

(1) $(b_1, \ldots, b_m)$ ist linear unabhängig,
(2) $L(b_1, \ldots, b_m) = T$.

Wir geben einige Beispiele:

BEISPIEL 6.19.

(1) $\{(1,0),(0,1)\}$ ist Basis des $\mathbb{R}^2$: Jeder Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ lässt sich als

$$x \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

schreiben, und liegt somit in der linearen Hülle von $(1,0)$ und $(0,1)$. Somit ist die lineare Hülle der Vektoren $\{(1,0),(0,1)\}$ der ganze $\mathbb{R}^2$. Die beiden Vektoren sind außerdem linear unabhängig.

(2) $\{(2,3)\}$ ist keine Basis des $\mathbb{R}^2$, da es kein $\lambda$ gibt, sodass $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

(3) $\{(2,3)\}$ ist Basis von $L((2,3))$.

Sei $U$ ein Unterraum des $\mathbb{R}^n$. Dann ist eine Basis eine Möglichkeit, die Menge $U$ anzugeben. Mathematica nutzt das, um die Lösungen des Gleichungssystems $A \cdot x = 0$ durch die Funktion `NullSpace` auszugeben. Dabei wird die (möglicherweise unendliche) Menge $U = \{x \in \mathbb{R}^n \,|\, A \cdot x = 0\}$ durch eine Basis dieses Raums angegeben.

```
In[50]:= A = {{1, 2, 3}}
Out[50]= {{1, 2, 3}}

In[51]:= NullSpace[A]
Out[51]= {{-3, 0, 1}, {-2, 1, 0}}
```

ÜBUNGSAUFGABEN 6.20.

(1) Finden Sie eine Basis $B$ der Ebene $L(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix})$, die weder $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ noch $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ enthält.

**4.2. Bestimmen der Basis eines Unterraums.** Wir können Unterräume des $\mathbb{R}^n$ auf zwei Arten angeben.

- explizit, das heißt, als lineare Hülle von Vektoren.
- implizit, das heißt, durch ein Gleichungssystem, dessen Lösungsmenge der anzugebende Unterraum ist.

Wir überlegen uns als erstes, wie wir eine Basis eines explizit gegebenen Unterraums berechnen können. Dazu berechnen wir die Basis des Raums

$$V = L(\begin{pmatrix} -5 \\ 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -33 \\ 17 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 85 \\ -44 \\ -3 \\ 2 \end{pmatrix}, \begin{pmatrix} -19 \\ 10 \\ 1 \\ 0 \end{pmatrix}).$$

Wir wollen also eine Basis des Zeilenraums der Matrix

$$\begin{pmatrix} -5 & 3 & 1 & 1 \\ -33 & 17 & 1 & -1 \\ 85 & -44 & -3 & 2 \\ -19 & 10 & 1 & 0 \end{pmatrix}$$

bestimmen. Der Zeilenraum ändert sich nicht, wenn wir eine Zeile mit einer von 0 verschiedenen Zahl multiplizieren, oder wenn wir zu einer Zeile ein Vielfaches einer anderen Zeile addieren. Wir können uns also jetzt systematisch Matrizen erzeugen, die alle den gleichen Zeilenraum haben wie die ursprüngliche Matrix. Das machen wir mit Mathematica.

```
In[52]:= << RowRed9.m
```

```
In[53]:= RowEchelonForm[{{-5, 3, 1, 1}, {-33, 17, 1, -1}, {85, -44, -3, 2}, {-19, 10, 1, 0}}]
```

$$\begin{pmatrix} -5 & 3 & 1 & 1 \\ -33 & 17 & 1 & -1 \\ 85 & -44 & -3 & 2 \\ -19 & 10 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} -5 & 3 & 1 & 1 \\ 0 & -14 & -28 & -38 \\ 0 & 7 & 14 & 19 \\ 0 & -7 & -14 & -19 \end{pmatrix}$$

$$\begin{pmatrix} -5 & 3 & 1 & 1 \\ 0 & -14 & -28 & -38 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

```
Out[53]= {{{1, 0, 0, 0}, {-33, 5, 0, 0}, {1, 5, 2, 0}, {-5, -5, 0, 10}},
          {{-5, 3, 1, 1}, {0, -14, -28, -38},
           {0, 0, 0, 0}, {0, 0, 0, 0}}}
```

Wir haben also eine Matrix in Zeilenstaffelform erzeugt, deren Zeilenraum der gleiche wie der Zeilenraum der gegebenen Matrix ist.

ALGORITHMUS 6.21 (Zeilenstaffelform).
Eingabe: Eine $m \times n$-Matrix $A$.
Ausgabe: Eine $m \times n$-Matrix $B$, sodass $B$ in Zeilenstaffelform ist, und $Z(A) = Z(B)$.

```
B ← A
zeile ← 1
spalte ← 1
while zeile ≤ m do
    while spalte ≤ n und B(i, spalte) = 0 für alle i mit zeile ≤ i ≤ m do
        spalte ← spalte + 1
    endwhile
    if spalte ≤ n then
        gewaehlteZeile ← ein i, sodass zeile ≤ i ≤ m und B(i, spalte) ≠ 0
        if gewaehlteZeile ≠ zeile then
            Vertausche die gewaehlteZeile-te mit der zeile-ten Zeile von B
        endif
        i ← zeile + 1
        while i ≤ m do
            Addiere passendes Vielfaches der zeile-ten Zeile zur i-ten Zeile von B, sodass B(i, spalte) = 0
            i ← i + 1
        endwhile
    endif
    zeile ← zeile + 1
    spalte ← spalte + 1
endwhile
return B
```

Dieser Algorithmus liefert also folgenden Satz.

SATZ 6.22. *Seien $m, n \in \mathbb{N}$, und sei $A$ eine $m \times n$-Matrix. Dann gibt es eine Matrix $B$ in Zeilenstaffelform, sodass $Z(A) = Z(B)$.*

Die Zeilen einer Matrix in Zeilenstaffelform, die nicht 0 sind, sind linear unabhängig. Daher haben wir auch eine Basis von

$$V = L((-5, 3, 1, 1), (-33, 17, 1, -1), (85, -44, -3, 2), (-19, 10, 1, 0))$$

gefunden, nämlich

$$B = \left(\begin{pmatrix} -5 \\ 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -14 \\ -28 \\ -38 \end{pmatrix}\right).$$

Wir fassen zusammen:

ALGORITHMUS 6.23 (Basis eines explizit gegebenen Unterraums).
Eingabe: Vektoren $v_1, \ldots, v_m \in \mathbb{R}^n$.
Ausgabe: Eine Basis $b_1, \ldots, b_k$ von $L(v_1, \ldots, v_m)$.

```
Bilde die m × n-Matrix V, in deren Zeilen die Vektoren v_1, …, v_m stehen
Berechne eine Matrix B in Zeilenstaffelform, sodass Z(V) = Z(B)
return (b_1, …, b_k) als jene Zeilen von B, die nicht 0 sind
```

Was wir noch nicht begründet haben ist, dass die Zeilen einer Matrix in Zeilenstaffelform, die nicht 0 sind, linear unabhängig sind. Betrachten wir dazu die Matrix

$$A = \begin{pmatrix} 1 & 2 & -3 & 0 & 6 & 7 & 0 \\ 0 & 0 & -5 & 0 & 14 & 17 & 0 \\ 0 & 0 & 0 & 0 & 67 & 76 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Um zu zeigen, dass die von 0 verschiedenen Zeilen von $A$ linear unabhängig sind, müssen wir zeigen, dass das Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ -3 & -5 & 0 \\ 0 & 0 & 0 \\ 6 & 14 & 67 \\ 7 & 17 & 76 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

nur die Lösung $(0, 0, 0)$ hat. Wir sehen, dass uns die erste Gleichung $\lambda_1 = 0$, dann die dritte Gleichung $\lambda_2 = 0$, und dann die fünfte Gleichung $\lambda_3 = 0$ liefert.

Das machen wir jetzt allgemein:

SATZ 6.24. *Sei $A$ eine $m \times n$-Matrix in Zeilenstaffelform. Dann sind die Zeilen von $A$, die nicht $0$ sind, linear unabhängige Vektoren im $\mathbb{R}^n$.*

*Beweis.* Seien $r \in \mathbb{N}$ und $j_1 < j_2 < \cdots < j_r$ wie in Definition 3.1. Wir betrachten eine Linearkombination der ersten $r$ Zeilen von $A$, die 0 ist. Seien also $(\lambda_1, \ldots, \lambda_r) \in \mathbb{R}^r$ so, dass für alle $k \in \{1, \ldots, n\}$ gilt:

$$\sum_{l=1}^{r} \lambda_l \cdot A[l, k] = 0. \tag{6.1}$$

Wir zeigen jetzt, dass für alle $i \in \{1, \ldots, r\}$ gilt: $\lambda_i = 0$. Wir zeigen das mit Induktion nach $i$. Betrachten wir die Gleichung 6.1 für $k := j_1$. Dann gilt

$$\sum_{l=1}^{r} \lambda_l \cdot A[l, j_1] = 0.$$

Für $l \geq 2$ gilt $A[l, j_1] = 0$, da Bedingung (3) von Definition 3.1 für $i := l$ und $k := j_1$ ergibt, dass $A(l, j_1) = 0$. Also gilt $\lambda_1 \cdot A[1, j_1] = 0$. Da $A[1, j_1] \neq 0$, gilt $\lambda_1 = 0$.

Sei nun $i \in \{1, \dots, r\}$. Wir nehmen an, dass $\lambda_1 = \lambda_2 = \cdots = \lambda_{i-1} = 0$. Wir betrachten die Gleichung 6.1 für $k := j_i$. Es gilt dann

$$\sum_{l=1}^{r} \lambda_l \cdot A[l, j_i] = 0,$$

also

$$\lambda_i \cdot A[i, j_i] + \sum_{l=i+1}^{r} \lambda_l \cdot A[l, j_i] = 0.$$

Für $l > i$ ergibt Bedingung (3) von Definition 3.1 (mit $i' := l$ und $k' := j_i$), dass $A[l, j_i] = 0$ gilt. Also gilt $\lambda_i \cdot A[i, j_i] = 0$. Da $A[i, j_i] \neq 0$, gilt auch $\lambda_i = 0$.

Es ist also $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. Also sind die ersten $r$ Zeilen von $A$ linear unabhängig. ■

ÜBUNGSAUFGABEN 6.25.

(1) Bestimmen Sie eine Basis von $L(\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix})$.
(2) Finden Sie jeweils eine Basis folgender Unterräume!
   (a) $\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + 3y + z = 0\}$.
   (b) $L(\begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix})$.
   (c) $L(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \end{pmatrix})$.

## 5. Die Zeilenstaffelnormalform

Wir studieren folgendes Problem: Gegeben sind zwei Unterräume $U, V$ von $\mathbb{R}^n$. Beide Unterräume sind explizit gegeben, das heißt, durch $u_1, \dots, u_l$ und $v_1, \dots, v_m$ so, dass $U = L(u_1, \dots, u_l)$ und $V = L(v_1, \dots, v_m)$. Wir fragen uns nun, ob $U = V$.

DEFINITION 6.26 (Zeilenstaffelnormalform). Sei $A$ eine $m \times n$-Matrix. $A$ ist in *Zeilenstaffelnormalform*, wenn es $r \in \mathbb{N}_0$ und $j_1, j_2, \dots, j_r \in \{1, \dots, n\}$ gibt, sodass

(1) $j_r > j_{r-1} > \cdots > j_1$.
(2) Für alle $i \in \{1, 2, \dots, r\}$ gilt: $A(i, j_i) = 1$.
(3) Für alle $i \in \{1, 2, \dots, r\}$ und für alle $k \in \{1, 2, \dots, r\}$ gilt: Wenn $k \neq i$, dann gilt $A(k, j_i) = 0$.
(4) Für alle $i \in \{1, 2, \dots, r\}$ und für alle $k \in \{1, 2, \dots, n\}$ mit $k < j_i$ gilt: $A(i, k) = 0$.
(5) Für alle $i \in \{1, 2, \dots, m\}$ mit $i > r$ und für alle $k \in \{1, 2, \dots, n\}$ gilt: $A(i, k) = 0$.

ALGORITHMUS 6.27 (Zeilenstaffelnormalform).
Eingabe: Eine $m \times n$-Matrix $A$.
Ausgabe: Eine $m \times n$-Matrix $B$, sodass $B$ in Zeilenstaffelnormalform ist, und $Z(A) = Z(B)$.

```
B ← A
zeile ← 1
spalte ← 1
while zeile ≤ m do
    (* Die ersten spalte − 1 Spalten sind in Zeilenstaffelnormalform *)
    while spalte ≤ n und B(i, spalte) = 0 für alle i mit zeile ≤ i ≤ m do
        spalte ← spalte + 1
    endwhile
    if spalte ≤ n then
        gewaehlteZeile ← ein i, sodass zeile ≤ i ≤ m und B(i, spalte) ≠ 0
        if gewaehlteZeile ≠ zeile then
            Vertausche die gewaehlteZeile-te mit der zeile-ten Zeile von B
        endif
        Multipliziere die zeile-te Zeile von B mit 1/B(zeile,spalte)
        i ← 1
        while i ≤ m do
            if i ≠ zeile then
                Addiere passendes Vielfaches der zeile-ten Zeile zur i-ten Zeile von B, sodass B(i, spalte) = 0
            endif
            i ← i + 1
        endwhile
    endif
    zeile ← zeile + 1
    spalte ← spalte + 1
endwhile
return B
```

Wir geben einige Beispiele für das Berechnen einer Matrix in Zeilenstaffelnormalform, die den gleichen Zeilenraum wie die Ausgangsmatrix besitzt.

AUFGABE 6.28.

```
In[54]:= << RowRed9.m

In[55]:= MatrixForm [A1]
```

Out[55]=
$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix}$$

*In[56]:=* **RowEchelonNormalForm[A1]**

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 0 & 1 & -2 & -4 & 2 \\ 0 & -5 & 10 & 4 & -2 \\ 0 & 17 & -34 & -10 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 & -18 & 8 \\ 0 & 1 & -2 & -4 & 2 \\ 0 & 0 & 0 & -16 & 8 \\ 0 & 0 & 0 & 58 & -29 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 & -18 & 8 \\ 0 & 1 & -2 & -4 & 2 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 58 & -29 \end{pmatrix}$$

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

*Out[56]=* $\left\{\left\{\left\{\frac{1}{2}, -\frac{5}{8}, -\frac{9}{8}, 0\right\}, \left\{0, -\frac{1}{4}, -\frac{1}{4}, 0\right\}, \left\{\frac{1}{4}, -\frac{5}{16}, -\frac{1}{16}, 0\right\}, \left\{-\frac{5}{2}, \frac{9}{8}, \frac{29}{8}, 1\right\}\right\}, \left\{\{1, 0, -2, 0, -1\}, \{0, 1, -2, 0, 0\}, \left\{0, 0, 0, 1, -\frac{1}{2}\right\}, \{0, 0, 0, 0, 0\}\right\}\right\}$

AUFGABE 6.29.

*In[57]:=* **<< RowRed9.m**

*In[58]:=* **MatrixForm[A2]**

*Out[58]=* $\begin{pmatrix} 1 & 5 & 7 & 9 \\ 2 & 10 & 14 & 18 \\ 3 & 2 & 1 & 0 \end{pmatrix}$

*In[59]:=* **RowEchelonNormalForm[A2]**

$$\begin{pmatrix} 1 & 5 & 7 & 9 \\ 2 & 10 & 14 & 18 \\ 3 & 2 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 7 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & -13 & -20 & -27 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 7 & 9 \\ 0 & -13 & -20 & -27 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 7 & 9 \\ 0 & 1 & \frac{20}{13} & \frac{27}{13} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -\frac{9}{13} & -\frac{18}{13} \\ 0 & 1 & \frac{20}{13} & \frac{27}{13} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

*Out[59]=* $\left\{\left\{\left\{-\frac{2}{13}, 0, \frac{5}{13}\right\}, \left\{\frac{3}{13}, 0, -\frac{1}{13}\right\}, \{-2, 1, 0\}\right\},\right.$

$\left.\left\{\left\{1, 0, -\frac{9}{13}, -\frac{18}{13}\right\}, \left\{0, 1, \frac{20}{13}, \frac{27}{13}\right\}, \{0, 0, 0, 0\}\right\}\right\}$

AUFGABE 6.30.

*In[60]:=* **<< RowRed9.m**

*In[61]:=* **MatrixForm[A3]**

*Out[61]=* $\begin{pmatrix} 1 & 5 & 3 \\ 2 & 10 & 8 \\ 4 & 20 & 15 \\ 1 & 6 & 5 \end{pmatrix}$

*In[62]:=* **RowEchelonNormalForm[A3]**

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & 10 & 8 \\ 4 & 20 & 15 \\ 1 & 6 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

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

*Out[62]=* $\left\{\left\{\left\{-\frac{10}{3}, 0, \frac{7}{3}, -5\right\}, \left\{\frac{5}{3}, 0, -\frac{2}{3}, 1\right\}, \left\{-\frac{4}{3}, 0, \frac{1}{3}, 0\right\}, \left\{\frac{2}{3}, 1, -\frac{2}{3}, 0\right\}\right\}, \{\{1, 0, 0\}, \{0, 1, 0\}, \{0, 0, 1\}, \{0, 0, 0\}\}\right\}$

AUFGABE 6.31.

*In[63]:=* **<< RowRed9.m**

*In[64]:=* **MatrixForm[A5]**

*Out[64]=* $\begin{pmatrix} 1 & 2 & -3 & 0 & 6 & 7 & 0 \\ -2 & -4 & 1 & 0 & 2 & 3 & 0 \\ 4 & 8 & 1 & 0 & 1 & -1 & 0 \\ 3 & 6 & -1 & 0 & 9 & 9 & 0 \end{pmatrix}$

*In[65]:=* **RowEchelonNormalForm[A5]**

$$\begin{pmatrix} 1 & 2 & -3 & 0 & 6 & 7 & 0 \\ -2 & -4 & 1 & 0 & 2 & 3 & 0 \\ 4 & 8 & 1 & 0 & 1 & -1 & 0 \\ 3 & 6 & -1 & 0 & 9 & 9 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -3 & 0 & 6 & 7 & 0 \\ 0 & 0 & -5 & 0 & 14 & 17 & 0 \\ 0 & 0 & 13 & 0 & -23 & -29 & 0 \\ 0 & 0 & 8 & 0 & -9 & -12 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -3 & 0 & 6 & 7 & 0 \\ 0 & 0 & 1 & 0 & -\frac{14}{5} & -\frac{17}{5} & 0 \\ 0 & 0 & 13 & 0 & -23 & -29 & 0 \\ 0 & 0 & 8 & 0 & -9 & -12 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 0 & 0 & -\frac{12}{5} & -\frac{16}{5} & 0 \\ 0 & 0 & 1 & 0 & -\frac{14}{5} & -\frac{17}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{67}{5} & \frac{76}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{67}{5} & \frac{76}{5} & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 0 & 0 & -\frac{12}{5} & -\frac{16}{5} & 0 \\ 0 & 0 & 1 & 0 & -\frac{14}{5} & -\frac{17}{5} & 0 \\ 0 & 0 & 0 & 0 & 1 & \frac{76}{67} & 0 \\ 0 & 0 & 0 & 0 & \frac{67}{5} & \frac{76}{5} & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 0 & -\frac{32}{67} & 0 \\ 0 & 0 & 1 & 0 & 0 & -\frac{15}{67} & 0 \\ 0 & 0 & 0 & 0 & 1 & \frac{76}{67} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

*Out[65]=* $\left\{\left\{\left\{\frac{1}{67}, -\frac{9}{67}, \frac{12}{67}, 0\right\}, \left\{-\frac{10}{67}, \frac{23}{67}, \frac{14}{67}, 0\right\}, \left\{\frac{6}{67}, \frac{13}{67}, \frac{5}{67}, 0\right\}, \{-1, -1, -1, 1\}\right\}, \left\{\left\{1, 2, 0, 0, 0, -\frac{32}{67}, 0\right\}, \left\{0, 0, 1, 0, 0, -\frac{15}{67}, 0\right\}, \left\{0, 0, 0, 0, 1, \frac{76}{67}, 0\right\}, \{0, 0, 0, 0, 0, 0, 0\}\right\}\right\}$

Der Algorithmus liefert den folgenden Satz.

SATZ 6.32. *Seien $m, n \in \mathbb{N}$, und sei $A$ eine $m \times n$-Matrix. Dann gibt es eine Matrix $B$ in Zeilenstaffelnormalform, sodass $Z(A) = Z(B)$.*

Mathematica hat einen eigenen Befehl zur Berechnung der Zeilenstaffelnormalform, nämlich `RowReduce`.

*In[66]:=* **MatrixForm [A1]**

*Out[66]=* $\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix}$

*In[67]:=* **MatrixForm [RowReduce [A1]]**

*Out[67]=* $\begin{pmatrix} 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

## 6. Der Nullraum einer Matrix

Wir überlegen uns jetzt, wie wir eine Basis eines implizit gegebenen Unterraums des $\mathbb{R}^n$ berechnen.

DEFINITION 6.33 (Skalarprodukt). Seien $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$. Wir definieren das Skalarprodukt $\langle v, w \rangle$ durch

$$\langle v, w \rangle = \sum_{i=1}^{n} v_i \, w_i.$$

DEFINITION 6.34. Sei $M$ eine Teilmenge des $\mathbb{R}^n$. Dann definieren wir

$$M^{\wedge} := \{v \in \mathbb{R}^n \mid \langle m, v \rangle = 0 \text{ für alle } m \in M\}.$$

SATZ 6.35. *Sei $A$ eine $m \times n$-Matrix. Dann gilt $N(A) = (Z(A))^{\wedge}$.*

KOROLLAR 6.36. *Sei $A$ eine $l \times n$, und sei $B$ eine $m \times n$-Matrix. Wenn $Z(A) = Z(B)$, dann gilt auch $N(A) = N(B)$.*

Wenn eine Matrix $B$ in Zeilenstaffelnormalform ist, dann kann man ihren Nullraum $N(B)$ besonders schnell berechnen.

SATZ 6.37. *Sei $B$ eine $m \times n$-Matrix in Zeilenstaffelnormalform, seien $j_1 < j_2 < \cdots < j_r \in \{1, 2, \ldots, n\}$ wie in Definition 6.26. Seien $i_1 < i_2 < \ldots < i_{n-r} \in \{1, 2, \ldots, n\}$ so, dass $\{i_1, \ldots, i_{n-r}\} \cup \{j_1, \ldots, j_r\} = \{1, 2, \ldots, n\}$. Sei $C$ eine $(n - r) \times n$-Matrix, die so definiert ist: Für alle $k \in \{1, 2, \ldots, n - r\}$ gilt:*

- *$C(k, i_k) = 1$,*
- *für alle $s \in \{1, \ldots, r\} : C(k, j_s) = -B(s, i_k)$, und*
- *für alle $l \in \{i_1, i_2, \ldots, i_{n-r}\} \setminus \{i_k\} : C(k, l) = 0$.*

*Dann steht in den Zeilen von $C$ eine Basis des Nullraumes von $B$.*

*Beweis.* Wir zeigen als erstes, dass jede Zeile von $C$ im Nullraum von $B$ liegt. Sei dazu $c$ die $k$-te Zeile von $C$. Wir berechnen jetzt $B \cdot c$ (dazu stellen wir uns $c$ als Spaltenvektor vor). Für $t \in \{1, \ldots, r\}$ berechnen wir den $t$-ten Eintrag von $B \cdot c$. Wir erhalten:

$$\begin{aligned}
(B \cdot c)[t] &= \sum_{l=1}^{n} B[t, l] \cdot c[l] \\
&= \sum_{l=1}^{n} B[t, l] \cdot C[k, l] \\
&= B[t, i_k] \cdot C[k, i_k] + \sum_{s=1}^{r} B[t, j_s] \cdot C[k, j_s] + \sum_{l \in \{i_1, i_2, \ldots, i_{n-r}\} \setminus \{i_k\}} B[t, l] \cdot C[k, l] \\
&= B[t, i_k] \cdot 1 + 1 \cdot C[k, j_t] + 0 \\
&= B[t, i_k] - B[t, i_k] = 0.
\end{aligned}$$

Daher gilt $Z(C) \subseteq N(B)$.

Bevor wir die Inklusion $N(B) \subseteq Z(C)$ beweisen, zeigen wir als Vorbereitung folgende Aussage:

> Wenn
> $$x = (x(1), x(2), \ldots, x(n))$$
> und
> $$y = (y(1), y(2), \ldots, y(n))$$
> Vektoren in $\mathbb{R}^n$ sind, sodass $B \cdot x = 0$, und $B \cdot y = 0$ und $x(i_l) = y(i_l)$ für alle $l \in \{1, 2, \ldots, n - r\}$ gilt, so gilt $x = y$.

Das begründen wir so: wir zeigen, dass $x$ und $y$ auch an den Stellen $j_1, \ldots, j_r$ übereinstimmen. Sei also $k \in \{1, \ldots, r\}$. Dann gilt

$$\sum_{j=1}^{n} B(k, j) \cdot x(j) = 0.$$

Das bedeutet

$$B(k, j_k) \cdot x(j_k) + \sum_{s \in \{1,\ldots,r\} \setminus \{k\}} B(k, j_s) \cdot x(j_s) = -\sum_{l=1}^{n-r} B(k, i_l) \cdot x(i_l),$$

also

$$x(j_k) = -\sum_{l=1}^{n-r} B(k, i_l) \cdot x(i_l).$$

Ebenso errechnen wir $y(j_k) = -\sum_{l=1}^{n-r} B(k, i_l) \cdot y(i_l)$. Also gilt $x = y$.

Nach dieser Vorbereitung zeigen wir $N(B) \subseteq Z(C)$. Sei dazu $x \in N(B)$. Wir setzen

$$\alpha = (\alpha(1), \ldots, \alpha(n-r)) := (x(i_1), \ldots, x(i_{n-r}))$$

und

$$y := C^T \cdot \alpha.$$

Da $B \cdot C^T = 0$, gilt $B \cdot y = 0$. Wir zeigen nun, dass für alle $l \in \{1, 2, \ldots, n-r\}$ gilt:

$$y(i_l) = x(i_l).$$

Dazu berechnen wir

$$\begin{aligned} y(i_l) &= (C^T \cdot \alpha)[i_l] \\ &= \sum_{k=1}^{n-r} C(k, i_l) \cdot \alpha(k) \\ &= C(l, i_l) \cdot \alpha(l) \\ &= 1 \cdot x(i_l). \end{aligned}$$

Aus unserer vorbereitenden Bemerkung erhalten wir

$$x = y.$$

Der Vektor $y$ liegt offensichtlich im Zeilenraum von $C$, also gilt $x \in Z(C)$.

Wir wissen also jetzt, dass die lineare Hülle der Zeilenvektoren von $C$ genau $N(B)$ ist. Jetzt zeigen wir noch, dass die Zeilen von $C$ linear unabhängig sind. Sei $\alpha =$

$(\alpha(1), \ldots, \alpha(n-r))$ so, dass $C^T \cdot \alpha = 0$. Sei $l \in \{1, 2, \ldots, n-r\}$. Dann gilt

$$\begin{aligned} 0 &= (C^T \cdot \alpha)[i_l] \\ &= \sum_{k=1}^{n-r} C(k, i_l) \cdot \alpha(k) \\ &= C(l, i_l) \cdot \alpha(l) \\ &= 1 \cdot \alpha(l). \end{aligned}$$

Daher gilt $\alpha(l) = 0$. ■

Wir haben also folgenden Satz gezeigt:

SATZ 6.38. *Sei $B$ eine $m \times n$-Matrix in Zeilenstaffelnormalform, und sei $r$ die Anzahl der Zeilen von $B$, die nicht $0$ sind. Dann hat $N(B)$ eine Basis, die genau $n - r$ Vektoren enthält.*

Wir fassen den konstruktiven Teil des Beweises des Satzes 6.37 nocheinmal zusammen.

ALGORITHMUS 6.39 (Nullraum einer Matrix in Zeilenstaffelnormalform).
Eingabe: Eine $m \times n$-Matrix $B$ in Zeilenstaffelnormalform.
Ausgabe: Eine Matrix $C$, deren Zeilen eine Basis des Nullraumes von $B$ sind.

$r \leftarrow$ Anzahl der Zeilen $\neq 0$ von $B$
**for** $s = 1$ **to** $r$ **do**
  $J[s] \leftarrow \min\{k \in \mathbb{N} \mid B[s,k] \neq 0\}$
**endfor**
$C \leftarrow$ die $(n-r) \times n$-Matrix mit allen Einträgen 0
*freieVariablen* $\leftarrow$ aufsteigend geordnete Liste der Elemente von $\{1, \ldots, n\} \setminus \{J[s] \mid s \in \{1, \ldots, r\}\}$
**for** $k = 1$ **to** $n - r$ **do**
  $i \leftarrow$ *freieVariablen*$[k]$
  (* die $i$-te Variable wird frei gewählt *)
  $C[k, i] \leftarrow 1$
  **for** $s = 1$ **to** $r$ **do**
    $C[k, J[s]] \leftarrow -B[s, i]$
  **endfor**
**endfor**

Als Beispiel berechnen wir den Nullraum von zwei Matrizen in Zeilenstaffelnormalform.

AUFGABE 6.40. *In[68]:=* **(* 1. Beispiel *)**

```
In[69]:= << RowRed9.m
```

```
In[70]:= B = RowReduce[A6];
```

*Out[70]=* MatrixForm[B]

*In[71]:=*
$$\begin{pmatrix} 1 & 2 & 0 & -17 & 0 & -22 \\ 0 & 0 & 1 & -5 & 0 & -5 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

*Out[71]=* CC = NullSpaceViaREF2[B];

MatrixForm[CC]

*In[72]:=*
$$\begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 17 & 0 & 5 & 1 & 0 & 0 \\ 22 & 0 & 5 & 0 & -2 & 1 \end{pmatrix}$$

*In[73]:=*

**(* 2. Beispiel *)**

*In[74]:=* **B = RowReduce[A5];**

*Out[74]=* MatrixForm[B]

*In[75]:=*
$$\begin{pmatrix} 1 & 2 & 0 & 0 & 0 & -\frac{32}{67} & 0 \\ 0 & 0 & 1 & 0 & 0 & -\frac{15}{67} & 0 \\ 0 & 0 & 0 & 0 & 1 & \frac{76}{67} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

*In[76]:=* **CC = NullSpaceViaREF2[B];**

*Out[76]=* MatrixForm[CC]

*In[77]:=*
$$\begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{32}{67} & 0 & \frac{15}{67} & 0 & -\frac{76}{67} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

ÜBUNGSAUFGABEN 6.41.

(1) Bestimmen Sie eine Basis für den Unterraum $U$ des $\mathbb{R}^4$, der durch
$$U = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid \begin{pmatrix} 1 & 0 & -2 & 3 \\ 2 & 1 & 0 & 4 \end{pmatrix} \cdot (x_1, x_2, x_3, x_4) = \binom{0}{0}\}$$
gegeben ist.

Jede Matrix hat den gleichen Zeilenraum wie eine Matrix in Zeilenstaffelnormalform, also folgt aus Satz 6.38:

SATZ 6.42. *Sei $m \in \mathbb{N}$, und sei $A$ eine $m \times (m+1)$-Matrix. Dann enthält der Nullraum von $A$ einen von $0$ verschiedenen Vektor.*

## 7. Die Dimension eines Unterraumes

LEMMA 6.43. *Seien $m, n \in \mathbb{N}$, sei $V$ ein Unterraum von $\mathbb{R}^n$, und seien $v_1, \dots, v_m \in \mathbb{R}^n$ so, dass $L(v_1, \dots, v_m) = V$. Seien $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ und sei $w := \sum_{i=1}^m \lambda_i v_i$. Sei $j$ so, dass $\lambda_j \neq 0$. Dann gilt*

$$L(v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_m) = V.$$

SATZ 6.44 (Austauschsatz von Steinitz). *Seien $m, n \in \mathbb{N}$, sei $V$ ein Unterraum des $\mathbb{R}^n$, und seien $x_1, \dots, x_m$ so, dass $L(x_1, \dots, x_m) = V$. Sei $r \in \mathbb{N}$, und sei $(w_1, \dots, w_r)$ eine linear unabhängige Folge von Elementen aus $V$. Dann gibt es für alle $i \in \{0, 1, \dots, \min(r, m)\}$ eine injektive Abbildung $\pi : \{i + 1, \dots, m\} \to \{1, \dots, m\}$, sodass*

$$V = L(w_1, \dots, w_i\, , x_{\pi(i+1)}, \dots, x_{\pi(m)})$$

*ist.*

*Beweis:* Induktion nach $i$. Für $i = 0$ setzen wir $\pi := \mathrm{id}_{\{1,\dots,m\}}$.

Sei nun $i \geq 1$. Wir nehmen an, dass

$$V = L(w_1, \dots, w_{i-1}, x_{\pi(i)}, \dots, x_{\pi(m)}) \tag{6.2}$$

ist. Wir wollen nun eines der $x_{\pi(j)}$ durch $w_i$ ersetzen. Da $w_i \in L(w_1, \dots, w_{i-1}, x_{\pi(i)}, \dots, x_{\pi(m)})$, gibt es $\lambda_1, \dots, \lambda_m \in \mathbb{R}$, sodass

$$w_i = \sum_{j=1}^{i-1} \lambda_j w_j + \sum_{j=i}^{m} \lambda_j x_{\pi(j)}.$$

Wenn nun für alle $j \in \{i, i + 1, \dots, m\}$ auch $\lambda_j = 0$ gilt, so sind $(w_1, \dots, w_i)$ linear abhängig. Daher gibt es ein $j \in \{i, i + 1, \dots, m\}$, sodass $\lambda_j \neq 0$. Nach Lemma 6.43 gilt also

$$V = L(w_1, \dots, w_{i-1}, x_{\pi(i)}, \dots, x_{\pi(j-1)}, w_i, x_{\pi(j+1)}, \dots, x_{\pi(m)}).$$

Wir definieren nun

$$\sigma : \{i, \dots, m\} \to \{1, \dots, m\}$$

durch $\sigma(j) := \pi(i)$, $\sigma(i) := \pi(j)$, und $\sigma(r) = \pi(r)$ für $r \in \{i, \dots, m\} \setminus \{i, j\}$. Nun gilt

$$\begin{aligned}
&L(w_1, \dots, w_i, x_{\sigma(i+1)}, \dots, x_{\sigma(m)}) \\
&= L(w_1, \dots, w_i, x_{\pi(i)}, \dots, x_{\pi(j-1)}, x_{\pi(j+1)}, \dots, x_{\pi(m)}) \\
&= V.
\end{aligned}$$

Somit leistet $\sigma|_{\{i+1,\dots,m\}}$ das Gewünschte. ■

KOROLLAR 6.45. *Seien $m, n \in \mathbb{N}$ und sei $V$ ein Unterraum des $\mathbb{R}^n$, und seien $x_1, \dots, x_m$ so, dass $L(x_1, \dots, x_m) = V$. Sei $r \in \mathbb{N}$, und sei $(w_1, \dots, w_r)$ eine linear unabhängige Folge von Elementen aus $V$. Dann gilt $r \leq m$.*

*Beweis:* Wir nehmen an $r > m$. Aus dem Austauschsatz (Satz 6.44) erhalten wir, dass $L(w_1, \dots, w_m) = V$. Also gilt $w_{m+1} \in L(w_1, \dots, w_m)$. Somit ist $(w_1, \dots, w_{m+1})$ linear abhängig, im Widerspruch zur Voraussetzung. ■

KOROLLAR 6.46. *Sei $n \in \mathbb{N}$. Dann gilt: Jede Folge von mehr als $n$ Vektoren im $\mathbb{R}^n$ ist linear abhängig.*

*Beweis.* $\mathbb{R}^n$ hat die sogenannte *kanonische* Basis $(e_1, \dots, e_n)$ mit

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

Aus Korollar 6.45 folgt, dass eine Folge aus mehr als $n$ Vektoren linear abhängig ist. ∎

SATZ 6.47. *Sei $n \in \mathbb{N}$, und sei $T$ ein Unterraum von $\mathbb{R}^n$. Dann haben alle Basen von $T$ gleich viele Elemente.*

SATZ 6.48. *Sei $n \in \mathbb{N}$, und sei $T$ ein Unterraum des $\mathbb{R}^n$. Dann hat $T$ eine Basis.*

*Beweis.*Sei $C := \{(v_1, \dots, v_i) \in \bigcup_{k=0}^{n} T^k \mid i \in \{0, \dots, n\}, (v_1, \dots, v_i) \text{ ist linear unabhängig}\}$. Wir wählen nun $m$ maximal, sodass $C \cap T^m \neq \emptyset$. Sei $(v_1, \dots, v_m) \in C \cap T^m$. Es gilt nun für alle $w \in T$, dass $(v_1, \dots, v_m, w)$ linear abhängig ist: Falls $m = n$, gilt das, da $n + 1$ Vektoren in $\mathbb{R}^n$ stets linear abhängig sind, sonst wegen der Maximalität von $m$. Also gilt $L(v_1, \dots, v_m) = T$. ∎

Jeder Unterraum des $\mathbb{R}^n$ ist also die lineare Hülle endlich vieler Vektoren. Der Satz 6.47 ermöglicht folgende Definition:

DEFINITION 6.49 (Dimension). Sei $T$ ein Unterraum des $\mathbb{R}^n$. Dann ist die *Dimension* von $T$ die Anzahl der Elemente einer Basis von $T$. Wir kürzen die Dimension von $T$ mit $\dim(T)$ ab.

SATZ 6.50. *Sei $n \in \mathbb{N}$, sei $k \in \mathbb{N}_0$, sei $T$ ein Unterraum des $\mathbb{R}^n$ mit Dimension $k$, und sei $S$ ein Unterraum von $\mathbb{R}^n$ mit $S \subseteq T$ und $\dim(S) = k$. Dann gilt $S = T$.*

*Beweis.* Sei $(s_1, \dots, s_k)$ eine Basis von $S$. Falls $L(s_1, \dots, s_k) = T$, so gilt $S = T$. Wenn es ein $t \in T \setminus L(s_1, \dots, s_k)$ gibt, so ist nach Satz 6.16 die Folge $(s_1, \dots, s_k, t)$ linear unabhängig. Dann haben wir $k + 1$ linear unabhängige Vektoren in einem Vektorraum mit einer $k$-elementigen Basis gefunden. Das widerspricht Korollar 6.45. ∎

ÜBUNGSAUFGABEN 6.51.

(1) Beweisen Sie:

Seien $k \in \mathbb{N}_0$, $n \in \mathbb{N}$, und sei $M$ eine Teilmenge von $\mathbb{R}^n$. Sei

$$B = (b_1, b_2, \dots, b_k)$$

eine linear unabhängige Folge von Vektoren aus $M$. Wir nehmen an, dass $B$ so ist, dass es kein $m \in M$ gibt, sodass $(b_1, b_2, \dots, b_k, m)$ ebenfalls linear unabhängig ist. (Wir fordern also, dass $B$ eine maximale linear unabhängige Folge aus $M$ ist.) Dann ist $B$ eine Basis für $L(M)$.

(2) Zeigen Sie:

Seien $k, n \in \mathbb{N}$, und sei $M = (m_1, m_2, \dots, m_k)$ eine Folge von Vektoren in $\mathbb{R}^n$. Dann gibt es $r \in \{0, \dots, k\}$ und $i_1, i_2, \dots, i_r \in \{1, \dots, k\}$, sodass folgendes gilt: $i_1 < i_2 < \cdots < i_r$, und $(m_{i_1}, m_{i_2}, \dots, m_{i_r})$ ist eine Basis von $L(M)$.

(3) Zeigen Sie:
Sei $k \in \mathbb{N}_0$, $n \in \mathbb{N}$, und sei $M = (m_1, m_2, \dots, m_k)$ eine Folge von Vektoren in $\mathbb{R}^n$. Dann gilt $\dim(L(M)) \leq k$.

## 8. Der Rang einer Matrix

DEFINITION 6.52. Sei $A$ eine $m \times n$-Matrix. Der *Rang* von $A$ ist die Dimension des Zeilenraums von $A$.

Man berechnet den Rang, indem man aus der Matrix $A$ eine Matrix $B$ in Zeilenstaffelform erzeugt, die den gleichen Zeilenraum wie $A$ hat. Die Zeilen von $B$, die nicht 0 sind, sind stets linear unabhängig. Sie bilden also eine Basis von $Z(B) = Z(A)$.

ALGORITHMUS 6.53 (Rang einer Matrix).
Eingabe: eine $m \times n$-Matrix $A$.
Ausgabe: der Rang von $A$.

1 Berechne mit Algorithmus 6.21 eine Matrix $B$ in Zeilenstaffelform, sodass $Z(B) = Z(A)$.
2 **return** Anzahl der Zeilen von $B$, die nicht 0 sind.

SATZ 6.54. *Seien $m, n \in \mathbb{N}$, sei $A$ eine $m \times n$-Matrix, und sei $k$ der Rang von $A$. Dann hat $N(A)$ die Dimension $n - k$.*

*Beweis.* Sei $B$ eine Matrix in Zeilenstaffelnormalform, sodass $Z(B) = Z(A)$. Wegen Korollar 6.36 hat $B \cdot x = 0$ die gleiche Lösungsmenge wie $A \cdot x = 0$. Sei $r$ die Anzahl der Zeilen von $B$, die nicht 0 sind. Diese Zeilen sind nach Satz 6.24 linear unabhängig, und bilden somit eine Basis von $Z(B)$. Da alle Basen eines Unterraums gleich viele Vektoren enthalten, gilt $r = k$. Nach Satz 6.37 hat $N(B)$ die Dimension $n - k$. Da $N(B) = N(A)$, hat auch $N(A)$ die Dimension $n - k$. ∎

## 9. Die Eindeutigkeit der Zeilenstaffelnormalform

Die Zeilenstaffelnormalform hat folgende wichtige Eigenschaft:

SATZ 6.55. *Seien $m, n \in \mathbb{N}$, und seien $B, C$ zwei $m \times n$-Matrizen in Zeilenstaffelnormalform. Wir nehmen an, dass $Z(B) = Z(C)$. Dann gilt $B = C$.*

*Beweis.* Wir fixieren $m, n \in \mathbb{N}$. Wir zeigen, dass für alle $r \in \mathbb{N}_0$ folgendes gilt:

> Für alle $a \in \mathbb{N}$ und für alle $a \times n$-Matrizen $B, C$ in Zeilenstaffelnormalform mit $\dim(Z(B)) = r$ und $Z(B) = Z(C)$ gilt $B = C$.

Diese Aussage beweisen wir durch Induktion nach $r$. Für $r = 0$ gilt $Z(B) = Z(C) = \{\vec{0}\}$, also $B = C = 0$.

Sei nun $r \geq 1$, sei $a \in \mathbb{N}$, und seien $B, C$ zwei $a \times n$-Matrizen in Zeilenstaffelnormalform mit $Z(B) = Z(C)$ und $\dim(Z(B)) = r$. Dann haben sowohl $B$ also auch $C$ genau $r$ Zeilen, die nicht 0 sind.

Für einen Unterraum $U$ von $\mathbb{R}^n$ und für $l \in \{1, \ldots, n-1\}$ definieren wir eine Teilmenge des $\mathbb{R}^l$ durch

$$U^{(l)} := \{(u_1, \ldots, u_l) \in \mathbb{R}^l \,|\, \text{es gibt } u_{l+1}, \ldots, u_n \in \mathbb{R}, \text{ sodass } (u_1, \ldots, u_n) \in U\}.$$

$U^{(n)}$ definieren wir durch $U^{(n)} := U$. Außerdem definieren wir für jede $m \times n$-Matrix $A$ und für jedes $l \in \{1, \ldots, n\}$

$$A_l := \text{ die } m \times l\text{-Matrix, die aus den ersten } l \text{ Spalten von } A \text{ besteht.}$$

Dann gilt:

$$\text{Für alle } l \in \{1, \ldots, n\} \;:\; Z(B_l) = Z(B)^{(l)}.$$

Seien $j_1, \ldots, j_r$ wie in der Definition der Zeilenstaffelnormalform. Dann gilt $\dim(Z(B)^{(j_r)}) = r$, da die ersten $r$ Zeilen von $B_{j_r}$ eine Basis von $Z(B)^{(j_r)}$ sind. Ebenso gilt $\dim(Z(B)^{(j_r-1)}) = r-1$, da die ersten $r-1$ Zeilen von $B_{j_r-1}$ eine Basis von $Z(B)^{(j_r-1)}$ sind. Daher gilt: $j_r$ ist jenes $l \in \{1, 2, \ldots, n\}$, sodass $\dim(Z(B)^{(l)}) = r$ und $\dim(Z(B)^{(l-1)}) = r-1$.

Die Matrix $C$ hat ebenfalls genau $r$ Zeilen, die nicht 0 sind. Seien $j'_1, \ldots, j'_r$ wie in der Definition der Zeilenstaffelnormalform. Wie oben gilt, dass $j'_r$ jenes $l' \in \{1, 2, \ldots, n\}$ ist , sodass $\dim(Z(C)^{(l')}) = r$ und $\dim(Z(C)^{(l'-1)}) = r-1$. Da $Z(B) = Z(C)$, gilt also $j_r = j'_r$.

Die ersten $r-1$ Zeilen der Matrix $B$ sind in Zeilenstaffelnormalform. Sie bilden eine Basis des Raums

$$U = \{(v_1, \ldots, v_n) \in Z(B) \,|\, v_{i_r} = 0\}.$$

Ebenso sind die ersten $r-1$ Zeilen der Matrix $C$ in Zeilenstaffelnormalform. Sie sind eine Basis von

$$V = \{(v_1, \ldots, v_n) \in Z(C) \,|\, v_{i_r} = 0\}.$$

Sei $B'$ die Matrix, die aus den ersten $r-1$ Zeilen von $B$ besteht, und sei $C'$ die Matrix, die aus den ersten $r-1$ Zeilen von $C$ besteht. Nach Voraussetzung gilt $Z(B) = Z(C)$. Daher gilt auch $U = V$. Das heißt aber, dass $Z(B') = Z(C')$. Nach Induktionsvoraussetzung gilt also $B' = C'$.

Es bleibt zu zeigen, dass $B$ und $C$ die gleiche $r$-te Zeile haben. Sei $b$ die $r$-te Zeile von $B$ und $c$ die $r$-te Zeile von $C$. Der Eintrag an der $j_r$-ten Stelle von $b - c$ ist $1 - 1 = 0$; auch alle Einträge von $b - c$ vor der $j_r$-ten Stelle sind gleich 0. Da $b, c \in Z(B)$, gilt $b - c \in Z(B)$. Seien $b_1, \ldots, b_r$ die ersten $r$ Zeilen von $B$. Seien $\lambda_1, \ldots, \lambda_r$ so, dass $\sum_{k=1}^{r} \lambda_k \cdot b_k = b - c$. Wenn wir für $i \in \{1, \ldots, r\}$ den $j_i$-ten Eintrag betrachten, erhalten wir

$$\sum_{k=1}^{r} \lambda_k \cdot B(k, j_i) = 0.$$

Da $B$ in Zeilenstaffelnormalform ist, ist nur der $i$-te Summand dieser Summe ungleich 0. Wir erhalten also $\lambda_i = 0$. Folglich sind alle $\lambda_i = 0$, also $b - c = 0$. Also sind auch die $r$-te Zeile von $B$ und $C$ gleich. ■

Dieser Satz hat folgende Anwendung:

ALGORITHMUS 6.56 (Gleichheit von Unterräumen).
Eingabe: $v_1, \ldots, v_k \in \mathbb{R}^n$ und $w_1, \ldots, w_l \in \mathbb{R}^n$.
Ausgabe: **true**, falls $L(v_1, \ldots, v_k) = L(w_1, \ldots, w_l)$, **false** sonst.

$m \leftarrow \max(k, l)$
$\tilde{V} \leftarrow$ die $m \times n$-Matrix, deren erste $k$ Zeilen die Vektoren $v_1, \ldots, v_k$ sind, und deren restliche Zeilen 0 sind
$\tilde{W} \leftarrow$ die $m \times n$-Matrix, deren erste $l$ Zeilen die Vektoren $w_1, \ldots, w_l$ sind, und deren restliche Zeilen 0 sind
Berechne eine Matrix $\tilde{V}'$ in Zeilenstaffelnormalform, sodass $Z(\tilde{V}') = Z(\tilde{V})$
Berechne eine Matrix $\tilde{W}'$ in Zeilenstaffelnormalform, sodass $Z(\tilde{W}') = Z(\tilde{W})$
**return true**, falls $\tilde{V}' = \tilde{W}'$, und **false** sonst.

## 10. Die Lösungsmenge inhomogener linearer Gleichungssysteme

DEFINITION 6.57. Sei $A \in \mathbb{R}^{m \times n}$, und sei $b \in \mathbb{R}^m$. Das Gleichungssystem $A \cdot x = b$ heißt *homogen*, wenn $b = 0$, und *inhomogen*, wenn $b \neq 0$.

DEFINITION 6.58. Sei $T$ ein Unterraum des $\mathbb{R}^n$, $x_0 \in \mathbb{R}^n$. Dann ist die *lineare Mannigfaltigkeit* $x_0 + T$ definiert durch:

$$x_0 + T := \{x_0 + t \mid t \in T\}.$$

Die lineare Mannigfaltigkeit $x_0 + T$ ist also der um den Vektor $x_0$ verschobene Unterraum $T$. Lösungsmengen linearer Gleichungssysteme sind stets leer oder lineare Mannigfaltigkeiten.

SATZ 6.59. *Seien $m, n \in \mathbb{N}$, sei $A$ eine $m \times n$-Matrix, und sei $x_0$ eine Lösung des Systems $A \cdot x = b$. Sei $L$ die Lösungsmenge von $A \cdot x = b$ und $N(A)$ der Nullraum von $A$. Dann gilt: $L = x_0 + N(A)$.*

AUFGABE 6.60. Sei $A = \begin{pmatrix} 3 & 1 & 2 & -5 \\ 0 & 1 & -3 & 8 \end{pmatrix}$ und $b = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$. Berechne die Lösungsmenge von $A \cdot x = b$.

Wir lösen dieses Beispiel mit Mathematica. Mathematica bietet dazu die Funktionen `NullSpace` und `LinearSolve`.

```
In[78]:= A = {{3, 1, 2, -5}, {0, 1, -3, 8}};

In[79]:= b = {4, 1};

In[80]:= NullSpace[A]
Out[80]= {{13, -24, 0, 3}, {-5, 9, 3, 0}}

In[81]:= LinearSolve[A, b]
Out[81]= {1, 1, 0, 0}
```

Die Lösungsmenge ist damit gegeben durch

$$L = \{(1, 1, 0, 0) + s \cdot (13, -24, 0, 3) + t \cdot (-5, 9, 3, 0) \mid s, t \in \mathbb{R}\}.$$

Wir überlegen uns jetzt noch, wie wir algorithmisch bestimmen können, ob ein lineares Gleichungssystem eine Lösung hat. Sei dazu $A \cdot x = b$ ein lineares Gleichungssystems. Wir bilden die erweiterte Koeffizientenmatrix $(A\ b)$, also jene Matrix, die wir erhalten, wenn wir $A$ um eine Spalte vergrößern und in diese Spalte $b$ schreiben.

SATZ 6.61. *Seien $A \cdot x = b$ und $C \cdot x = d$ zwei lineare Gleichungssysteme, und sei $E = (A\ b)$ und $F = (C\ d)$. Wenn $E$ und $F$ den gleichen Zeilenraum haben, dann haben die beiden Gleichungssysteme die gleiche Lösungsmenge.*

SATZ 6.62. *Sei $A$ eine $m \times n$-Matrix, sei $b \in \mathbb{R}^m$, sei $E$ die Matrix $(A\ b)$, und sei $L := \{x \in \mathbb{R}^n \mid A \cdot x = b\}$. Dann gilt*

$$L = \{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ -1 \end{pmatrix} \in N(E)\}.$$

SATZ 6.63. *Seien $m, n \in \mathbb{N}$, sei $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, und sei $E := (A\ b)$. Sei $F$ eine Matrix in Zeilenstaffelform mit $Z(F) = Z(E)$, und seien $j_1 < \ldots < j_r$ wie in Definition 3.1. Wenn $j_r = n + 1$, dann hat das System $A \cdot x = b$ keine Lösung.*

*Beweis.* Wenn $j_r = n + 1$, dann gilt für alle Elemente $x \in N(F)$, dass $x_{n+1} = 0$ ist. Nach Satz 6.62 hat $A \cdot x = b$ also keine Lösung. ■

Wir bestimmen jetzt alle Lösungen von $A \cdot x = b$ folgendermaßen:

ALGORITHMUS 6.64 (Lineare Gleichungssysteme).
Eingabe: Eine $m \times n$-Matrix $A$, und $b \in \mathbb{R}^m$.
Ausgabe: “keine Lösung”, falls es kein $x \in \mathbb{R}^n$ mit $A \cdot x = b$ gibt. Sonst geben wir einen Vektor $x_0 \in \mathbb{R}^n$ und eine $t \times n$-Matrix $D$ aus, sodass die Lösungsmenge von $A \cdot x = b$ gleich $x_0 + Z(D)$ ist.

$B \leftarrow$ die erweiterte Matrix $(A\ b)$
$C \leftarrow$ die $m \times (n + 1)$-Matrix in Zeilenstaffelnormalform mit $Z(C) = Z(B)$
$r \leftarrow$ Anzahl der Zeilen $\neq 0$ von $B$.
**for** $s = 1$ **to** $r$ **do**
    $J[s] \leftarrow \min\{k \in \mathbb{N} \mid C[s, k] \neq 0\}$

**endfor**
**if** $J[r] = n + 1$ **then**
  *resultat* $\leftarrow$ *“keine Lösung”*
**else**
  $D \leftarrow$ die $(n - r) \times n$-Matrix mit allen Einträgen 0
  $x_0 \leftarrow$ der Nullvektor im $\mathbb{R}^n$
  $I \leftarrow$ aufsteigend geordnete Liste der Elemente von $\{1, \dots, n\} \setminus \{J[s] \mid s \in \{1, \dots, r\}\}$
  **for** $k = 1$ **to** $n - r$ **do**
    $i \leftarrow I[k]$
    (* die $i$-te Variable wird frei gewählt *)
    $D[k, i] \leftarrow 1$
    **for** $s = 1$ **to** $r$ **do**
      $D[k, J[s]] \leftarrow -C[s, i]$
    **endfor**
  **endfor**
  **for** $s = 1$ **to** $r$ **do**
    $x_0[J[s]] \leftarrow C[s, n + 1]$
  **endfor**
  *resultat* $\leftarrow (x_0, C)$
**endif**
**return** *resultat*

Wir berechnen jetzt die Lösung von zwei linearen Gleichungssystemen mit Mathematica, indem wir die Zeilenstaffelnormalform von $(A\ b)$ ausrechnen.

AUFGABE 6.65. *In[82]:=* **<< RowRed9.m**

*In[83]:=* **MatrixForm[A1]**

*Out[83]=* $\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix}$

*In[84]:=* **MatrixForm[b1]**

*Out[84]=* $\begin{pmatrix} -9 \\ -4 \\ 0 \\ -18 \end{pmatrix}$

*In[85]:=*

**LinSolveViaREF[A1, b1]**

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 & -9 \\ 1 & -4 & 6 & -2 & 0 & -4 \\ -1 & 0 & 2 & 2 & 0 & 0 \\ 5 & -8 & 6 & 0 & -5 & -18 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 & -9 \\ 0 & 1 & -2 & -4 & 2 & 5 \\ 0 & -5 & 10 & 4 & -2 & -9 \\ 0 & 17 & -34 & -10 & 5 & 27 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 & -18 & 8 & 16 \\ 0 & 1 & -2 & -4 & 2 & 5 \\ 0 & 0 & 0 & -16 & 8 & 16 \\ 0 & 0 & 0 & 58 & -29 & -58 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 & -18 & 8 & 16 \\ 0 & 1 & -2 & -4 & 2 & 5 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 0 & 0 & 58 & -29 & -58 \end{pmatrix}$$

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

*Out[85]=* $\left\{\{-2, 1, 0, -1, 0\}, \left\{\{2, 2, 1, 0, 0\}, \left\{1, 0, 0, \frac{1}{2}, 1\right\}\right\}\right\}$

Jetzt noch ein System, das keine Lösung hat.

AUFGABE 6.66. *In[86]:=* **<< RowRed9.m**

*In[87]:=* **MatrixForm[A2]**

*Out[87]=* $\begin{pmatrix} 1 & 5 & 7 & 9 \\ 2 & 10 & 14 & 18 \\ 3 & 2 & 1 & 0 \end{pmatrix}$

*In[88]:=* **MatrixForm[b2]**

*Out[88]=* $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

*In[89]:=*

**LinSolveViaREF[A2, b2]**

$$\begin{pmatrix} 1 & 5 & 7 & 9 & 0 \\ 2 & 10 & 14 & 18 & 1 \\ 3 & 2 & 1 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 7 & 9 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & -13 & -20 & -27 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 7 & 9 & 0 \\ 0 & -13 & -20 & -27 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 5 & 7 & 9 & 0 \\ 0 & 1 & \frac{20}{13} & \frac{27}{13} & -\frac{1}{13} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -\frac{9}{13} & -\frac{18}{13} & \frac{5}{13} \\ 0 & 1 & \frac{20}{13} & \frac{27}{13} & -\frac{1}{13} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -\frac{9}{13} & -\frac{18}{13} & 0 \\ 0 & 1 & \frac{20}{13} & \frac{27}{13} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

*Out[89]*= NOSOLUTION

ÜBUNGSAUFGABEN 6.67.

(1) Widerlegen Sie durch Angabe eines Gegenbeispiels:
Wenn das Gleichungssystem $A \cdot x = 0$ eine Lösung hat, so besitzt für jede rechte Seite $b$ das Gleichungssystem $A \cdot x = b$ zumindest eine Lösung.

## 11. Koordinaten

SATZ 6.68. *Seien $n \in \mathbb{N}$, $m \in \mathbb{N}_0$, sei $B = (b_1, \dots, b_m)$ eine Basis des Unterraums $T$ von $\mathbb{R}^n$, und sei $t \in T$. Dann gibt es genau ein $m$-Tupel $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ mit*

$$t = \sum_{i=1}^{m} \lambda_i \cdot b_i.$$

*Beweis.* Ein solches $m$-Tupel existiert, da $t \in L(b_1, \dots, b_m)$. Sei nun $\overline{B}$ die $n \times m$-Matrix, in deren Spalten die Vektoren $(b_1, \dots, b_m)$ stehen. Angenommen, $\alpha = (\lambda_1, \dots, \lambda_m)$ und $\beta = (\mu_1, \dots, \mu_m)$ seien zwei verschiedene Tupel mit $\overline{B} \cdot \alpha = \overline{B} \cdot \beta = t$. Dann gilt $\overline{B} \cdot (\alpha - \beta) = 0$ und $\alpha - \beta \neq 0$. Das ist ein Widerspruch zur linearen Unabhängigkeit von der Spalten von $\overline{B}$. ■

Dieses $m$-Tupel gibt also an, wie wir den Vektor $t$ aus den Basisvektoren zusammenbauen können. Wir nennen dieses Tupel die *Koordinaten* des Vektors $t$ bezüglich der Basis $B$.

DEFINITION 6.69. Sei $B = (b_1, \dots, b_m)$ Basis von $T \subseteq \mathbb{R}^n$, sei $t \in T$, und sei $\overline{B}$ die $n \times m$-Matrix, in deren Spalten die Vektoren $(b_1, \dots, b_m)$ stehen. Dann heißt das $\alpha = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ mit

$$\overline{B} \cdot \alpha = t$$

das *Koordinatentupel* von $t$ bezüglich der Basis $B$. Wir schreiben auch $(t)_B = \alpha$.

Das Koordinatentupel $(t)_B$ erfüllt also die Gleichung

$$\overline{B} \cdot (t)_B = t.$$

AUFGABEN 6.70. (1) Sei $B = ((1,0,3),(2,1,6))$, $T = L(B)$, und $v = (3,1,9)$. Offenbar gilt

$$\begin{pmatrix} 3 \\ 1 \\ 9 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix},$$

woraus $(v)_B = (1,1)$ folgt.

(2) Sei $T = L((3,2,1),(0,1,2))$, und $B = ((3,2,1),(0,1,2))$. Für das erste Basiselement gilt offensichtlich

$$(\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix})_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

weiters ist zum Beispiel

$$(4 \cdot \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix})_B = \begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

ÜBUNGSAUFGABEN 6.71.

(1) Der Vektor $v$ hat bezüglich der Basis $A = (\begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}, \begin{pmatrix} 2 \\ 8 \\ -4 \end{pmatrix})$ der Ebene $\varepsilon$ die Koordinaten $(v)_A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Berechnen Sie seine Koordinaten bezüglich der Basis $B = (\begin{pmatrix} 5 \\ 16 \\ -11 \end{pmatrix}, \begin{pmatrix} 12 \\ 40 \\ -26 \end{pmatrix})$.

(2) Die Ebene $\varepsilon$ hat die Basen

$$A = (\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix})$$

und

$$B = (\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}).$$

Der Vektor $v$ hat bezüglich $B$ die Koordinaten $(v)_B = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$. Berechnen Sie seine Koordinaten bezüglich $A$!

(3) Die Ebene $e$ ist durch $e = L(\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix})$ gegeben. Sie hat die Basen

$$B = (\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix})$$

und

$$C = (\begin{pmatrix} 24 \\ -2 \\ 31 \end{pmatrix}, \begin{pmatrix} 18 \\ -2 \\ 24 \end{pmatrix}).$$

Der Vektor $v$ ist gegeben durch $(v)_C = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$. Berechnen Sie $(v)_B$.

(4) Sei $a = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}$ und $b = \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix}$, und sei $B = (a, b)$. Zeigen Sie, dass ein Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ bezüglich der Basis $B$ die Koordinaten $\begin{pmatrix} \langle \begin{pmatrix} x \\ y \end{pmatrix}, a \rangle \\ \langle \begin{pmatrix} x \\ y \end{pmatrix}, b \rangle \end{pmatrix}$ hat; das heißt, zeigen Sie

$$(\begin{pmatrix} x \\ y \end{pmatrix})_B = \begin{pmatrix} \langle \begin{pmatrix} x \\ y \end{pmatrix}, a \rangle \\ \langle \begin{pmatrix} x \\ y \end{pmatrix}, b \rangle \end{pmatrix}.$$

Stimmt diese Formel für jede Basis $(a, b)$ des $\mathbb{R}^2$?

(5) Sei $B = (\begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix})$, und sei $E$ die lineare Hülle von $B$. $B$ ist dann eine Basis von $E$.
   (a) Welcher Vektor $w$ hat bezüglich $B$ die Koordinaten $(w)_B = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$?
   (b) Wie lauten die Koordinaten von $\begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$ bezüglich $B$ ?

(c) Geben Sie eine Basis $C$ von $E$ an, bezüglich der der Punkt $\left(\begin{smallmatrix}-1\\1\\5\end{smallmatrix}\right)$ die Koordinaten $\left(\begin{smallmatrix}0\\2\end{smallmatrix}\right)$ hat.

## 12. Summen und Durchschnitte von Unterräumen

SATZ 6.72. *Seien $U, V$ Unterräume des $\mathbb{R}^n$. Dann ist $U \cap V$ auch ein Unterraum des $\mathbb{R}^n$.*

DEFINITION 6.73. Seien $U, V$ Unterräume des $\mathbb{R}^n$. Wir definieren

$$U + V := \{u + v \mid u \in U, v \in V\}.$$

SATZ 6.74. *Seien $U, V$ Unterräume des $\mathbb{R}^n$. Dann ist $U + V$ auch ein Unterraum des $\mathbb{R}^n$.*

Wir überlegen uns nun, wie wir Basen von $U \cap V$ und $U + V$ berechnen können.

ALGORITHMUS 6.75 (Summe explizit gegebener Unterräume).
Eingabe: $u_1, \ldots, u_l \in \mathbb{R}^n$ und $v_1, \ldots, v_m \in \mathbb{R}^n$ so, dass $L(u_1, \ldots, u_l) = U$ und $L(v_1, \ldots, v_l) = V$.
Ausgabe: Eine Basis von $U + V$.

1 Bilde eine $(l + m) \times n$-Matrix $W$, deren erste $l$ Zeilen die Vektoren $u_1, \ldots, u_l$ sind, und deren $(l + 1)$-te bis $(l + m)$-te Zeile die Vektoren $v_1, \ldots, v_l$ sind
2 Bestimme eine Matrix $W'$ in Zeilenstaffelform, sodass $Z(W') = Z(W)$
3 **return** die Zeilen von $W'$, die nicht 0 sind

ALGORITHMUS 6.76 (Durchschnitt implizit gegebener Unterräume).
Eingabe: Eine $r \times n$-Matrix $A$ und eine $s \times n$-Matrix $B$, sodass $N(A) = U$ und $N(B) = V$.
Ausgabe: Eine Basis von $U \cap V$.

1 Bilde die $(r + s) \times n$-Matrix $C$, deren erste $r$ Zeilen die Matrix $A$ und deren letzte $s$ Zeilen die Matrix $B$ sind
2 Bestimme eine Basis $X$ für $N(C)$
3 **return** $X$

Um die Summe zweier implizit gegebener Unterräume zu berechnen, können wir uns zuerst von beiden Unterräumen mit den Algorithmen 6.27 und 6.39 eine Basis ausrechnen, und dann Algorithmus 6.75 verwenden.

AUFGABE 6.77. Bestimmen Sie eine Basis des Raums $U + V$, wobei

$$U = \{x \in \mathbb{R}^4 \mid \begin{pmatrix} 4 & -3 & 0 & 1 \\ 2 & -2 & 1 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\},$$

und

$$V = \{x \in \mathbb{R}^4 \mid \begin{pmatrix} 2 & -3 & 0 & 1 \\ -3 & -2 & 1 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\}.$$

*Lösung:* Wir berechnen Matrizen $C, D$, in deren Zeilen Basen von $U$ bzw. $V$ stehen.

```
In[90]:= A = {{4, -3, 0, 1}, {2, -2, 1, 0}}
Out[90]= {{4, -3, 0, 1}, {2, -2, 1, 0}}

In[91]:= B = {{2, -3, 0, 1}, {-3, -2, 1, 0}}
Out[91]= {{2, -3, 0, 1}, {-3, -2, 1, 0}}

In[92]:= CC = NullSpace[A]
Out[92]= {{-1, -1, 0, 1}, {3, 4, 2, 0}}

In[93]:= DD = NullSpace[B]
Out[93]= {{-2, 3, 0, 13}, {3, 2, 13, 0}}
```

Jetzt bestimmen wir eine Matrix $E$, deren Zeilenraum der Raum $U + V$ ist.

```
In[94]:= EE = Join[CC, DD];

In[95]:= MatrixForm[EE]
```

Out[95]= $\begin{pmatrix} -1 & -1 & 0 & 1 \\ 3 & 4 & 2 & 0 \\ -2 & 3 & 0 & 13 \\ 3 & 2 & 13 & 0 \end{pmatrix}$

```
In[96]:= FF = RowReduce[EE];

In[97]:= MatrixForm[FF]
```

Out[97]= $\begin{pmatrix} 1 & 0 & 0 & -\frac{16}{5} \\ 0 & 1 & 0 & \frac{11}{5} \\ 0 & 0 & 1 & \frac{2}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Also ist

$$((1, 0, 0, -16/5), (0, 1, 0, 11/5), (0, 0, 1, 2/5))$$

eine Basis von $U + V$.

Nun wollen wir den Durchschnitt zweier parametrisiert gegebener Unterräume berechnen.

SATZ 6.78. *Sei $U$ ein Unterraum des $\mathbb{R}^n$. Dann gilt* $\dim(U^\wedge) = n - \dim(U)$.

*Beweis.* Sei $(u_1, \ldots, u_l)$ eine Basis von $U$. $U^\wedge$ ist die Lösungsmenge des Gleichungssystems $\tilde{U} \cdot x = 0$, wobei $\tilde{U}$ die $l \times n$-Matrix ist, deren Zeilenvektoren $u_1, \ldots, u_l$ sind. Der Rang von $U$ ist $l$, daher ist die Dimension von $U^\wedge$ nach Satz 6.54 gleich $n - l$. ■

SATZ 6.79. *Sei $U$ ein Unterraum des $\mathbb{R}^n$. Dann gilt* $(U^\wedge)^\wedge = U$.

*Beweis.* Wir zeigen zunächst $\supseteq$. Sei $u \in U$. Um zu zeigen, dass $u \in (U^\wedge)^\wedge$ liegt, ist zu zeigen, dass $\langle v, u\rangle = 0$ für alle $v \in U^\wedge$. Sei $v \in U^\wedge$. Da $v \in U^\wedge$, gilt $\langle u, v\rangle = 0$. Also gilt $U \subseteq (U^\wedge)^\wedge$. Sei nun $l$ die Dimension von $U$. Es gilt $\dim((U^\wedge)^\wedge) = n - \dim(U^\wedge) = n - (n - l) = l$. Wegen Satz 6.50 gilt also $U = (U^\wedge)^\wedge$. ■

SATZ 6.80. *Sei $A$ eine $m \times n$-Matrix, und sei $B$ eine $n \times l$-Matrix, in deren Spalten eine Basis für den Nullraum von $A$ steht. Dann gilt*

$$N(B^T) = Z(A).$$

*Der Nullraum von $B^T$ ist also der Zeilenraum von $A$.*

*Beweis.* In den Spalten von $B$ steht eine Basis für den Nullraum $N(A)$ von $A$. Der Nullraum $N(A)$ ist gleichzeitig der Raum $(Z(A))^\wedge$. In den Spalten von $B$ steht also eine Basis für $(Z(A))^\wedge$. Wieder nach Satz 6.35 gilt $N(B^T) = (Z(B^T))^\wedge$. Da in den Spalten von $B$ eine Basis von $Z(A)^\wedge$ steht, sind auch die Zeilen von $B^T$ eine Basis von $(Z(A))^\wedge$. Es gilt also $Z(B^T) = (Z(A))^\wedge$. Somit gilt insgesamt $N(B^T) = (Z(B^T))^\wedge = ((Z(A))^\wedge)^\wedge$. Wegen Satz 6.79 gilt also $N(B^T) = Z(A)$. ■

Dieser Satz lässt sich noch umformulieren:

KOROLLAR 6.81. *Sei $A$ eine $m \times n$-Matrix, und sei $C$ eine $l \times n$-Matrix, in deren Zeilen eine Basis für den Nullraum von $A$ steht. Dann gilt*

$$N(C) = Z(A).$$

*Der Nullraum von $C$ ist also der Zeilenraum von $A$.*

Das führt zu folgendem Algorithmus:

ALGORITHMUS 6.82 (Implizitisieren eines Unterraums).
Eingabe: Eine $m \times n$-Matrix $A$.
Ausgabe: Eine $l \times n$-Matrix $C$, sodass $N(C) = Z(A)$.

1 $C \leftarrow$ eine Matrix, in deren Zeilen eine Basis von $N(A)$ steht
2 **return** $C$

ALGORITHMUS 6.83 (Durchschnitt von explizit gegebenen Unterräumen).
Eingabe: $u_1, \dots, u_l \in \mathbb{R}^n$ und $v_1, \dots, v_m \in \mathbb{R}^n$ so, dass $L(u_1, \dots, u_l) = U$ und $L(v_1, \dots, v_l) = V$.
Ausgabe: Eine Basis von $U \cap V$.

1 Bestimme mit Algorithmus 6.82 eine Matrix $C_1$, sodass $N(C_1) = U$
2 Bestimme mit Algorithmus 6.82 eine Matrix $C_2$, sodass $N(C_2) = U$
3 Verwende Algorithmus 6.76, um eine Basis von $N(C_1) \cap N(C_2)$ auszurechnen

Wir verwenden diesen Algorithmus, um folgendes Beispiel zu lösen.

AUFGABE 6.84. Sei $U = L((-1, -2, -2, -2), (1, 4, 6, 8))$, und $V = L((1, 0, 3, -2), (-1, 2, 1, 8))$. Berechnen Sie eine Basis von $U \cap V$.

*Lösung:*

```
In[98]:= CfuerU1 = NullSpace[{{-1, -2, -2, -2}, {1, 4, 6, 8}}]
Out[98]= {{4, -3, 0, 1}, {2, -2, 1, 0}}

In[99]:= CfuerU2 = NullSpace[{{1, 0, 3, -2}, {-1, 2, 1, 8}}]
Out[99]= {{2, -3, 0, 1}, {-3, -2, 1, 0}}

In[100]:= CfuerU1geschnittenU2 = Join[CfuerU1, CfuerU2]
Out[100]= {{4, -3, 0, 1}, {2, -2, 1, 0}, {2, -3, 0, 1}, {-3, -2, 1, 0}}

In[101]:= BasisFuerU1geschnittenU2 = NullSpace[CfuerU1geschnittenU2]
Out[101]= {{0, 1, 2, 3}}
```

Somit ist $((0,1,2,3))$ eine Basis von $U \cap V$.

Zum Schneiden parametrisiert gegebener linearer Mannigfaltigkeiten hilft folgender Algorithmus.

ALGORITHMUS 6.85 (Implizitisieren einer linearen Mannigfaltigkeit).
Eingabe: Eine $m \times n$-Matrix $A$, und $b \in \mathbb{R}^n$.
Ausgabe: Eine $l \times n$-Matrix $C$ und $d \in \mathbb{R}^l$, sodass $b + Z(A) = \{x \in \mathbb{R}^n \,|\, C \cdot x = d\}$.

1 $C \leftarrow$ eine Matrix, in deren Zeilen eine Basis von $N(A)$ steht
2 $d \leftarrow C \cdot b$
3 **return** $(C, d)$

KAPITEL 7

# Orthogonalität

## 1. Der Winkel zwischen zwei Vektoren

DEFINITION 7.1. Für $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ ist das *Skalarprodukt* von $x$ und $y$ definiert als

$$\langle x, y\rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

Es gilt

$$\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 = ||x||^2.$$

LEMMA 7.2. *Für $x, y, z \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt:*

(1) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,
(2) $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$.

SATZ 7.3. *Der Winkel $\varphi$ zwischen $x$ und $y$ ist gegeben durch*

$$\cos\varphi = \frac{\langle x, y\rangle}{||x|| \cdot ||y||}.$$

DEFINITION 7.4. Zwei Vektoren *stehen* genau dann *normal aufeinander*, wenn $\langle x, y\rangle = 0$. Wir schreiben dann $x \perp y$.

ÜBUNGSAUFGABEN 7.5.

(1) Bestimmen Sie einen Vektor, der auf $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ normal steht.
(2) Sei $B = (\begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix})$, und sei $E$ die lineare Hülle von $B$. $B$ ist dann eine Basis von $E$. Außerdem bezeichnen wir mit $\overline{B}$ die Matrix

$$\overline{B} := \begin{pmatrix} 3 & 2 \\ 0 & 1 \\ -1 & 4 \end{pmatrix}.$$

(a) Welches Gleichungssystem muss ein Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ erfüllen, der auf alle Vektoren in $E$ normal steht? Was hat die Koeffizientenmatrix dieses Systems mit $\overline{B}$ zu tun?
(b) Mit welchem Gleichungssystem können Sie feststellen, ob ein Vektor $v$ in $E$ liegt? Was hat die Koeffizientenmatrix dieses Systems mit $\overline{B}$ zu tun?
(c) Sei $w = \begin{pmatrix} 16 \\ 44 \\ -6 \end{pmatrix}$. Finden Sie einen Vektor $e$, der in $E$ liegt, sodass $w - e$ normal auf alle Vektoren in $E$ steht.

## 2. Der Normalraum auf eine Menge von Vektoren

Wenn $U$ eine Teilmenge des $\mathbb{R}^n$ ist, dann bezeichnen wir den Raum $U^\wedge$ auch als "$U$ orthogonal" und verwenden anstelle von $U^\wedge$ die Bezeichnung $U^\perp$. Für eine Teilmenge $A$ des $\mathbb{R}^n$ ist $A^\perp$ also definiert durch

$$A^\perp := \{x \in \mathbb{R}^n \mid x \perp a \text{ für alle } a \in A\}.$$

Für jede Teilmenge $A$ von $\mathbb{R}^n$ gilt $A^\perp = L(A)^\perp$ und $(A^\perp)^\perp = L(A)$.

ÜBUNGSAUFGABEN 7.6.

(1) Gegeben seien die Vektoren

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \;\; y = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \;\; z = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}.$$

Berechnen Sie
  (a) $\{x\}^\perp$,
  (b) $\{x, y\}^\perp$,
  (c) $\{x, y, z\}^\perp$,

und jeweils die Dimension davon.

(2) Welche Vektoren des $\mathbb{R}^4$ stehen auf alle drei Vektoren

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

normal? (Zwei Vektoren im $\mathbb{R}^4$ stehen aufeinander normal, wenn ihr Skalarprodukt gleich 0 ist.)

## 3. Orthonormalbasen

DEFINITION 7.7. Sei $T$ ein Unterraum des $\mathbb{R}^n$ und $B = (b_1, \ldots, b_k)$ eine Basis von $T$. $B$ ist eine *Orthonormalbasis* (ONB), wenn

(1) $\langle b_i, b_j\rangle = 0$ für alle $i, j \in \{1, 2, \ldots, k\}$ mit $i \neq j$.
(2) $\|b_i\| = \sqrt{\langle b_i, b_i\rangle} = 1$ für alle $i \in \{1, 2, \ldots, k\}$.

BEISPIELE 7.8.

(1) Die kanonische Basis $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ ist ONB des $\mathbb{R}^3$.
(2) $B = ((\frac{1}{2}, \frac{-\sqrt{3}}{2}, 0), (\frac{\sqrt{3}}{2}, \frac{1}{2}, 0), (0, 0, 1))$ ist ONB des $\mathbb{R}^3$.
(3) $B = (\frac{1}{2}(1, -1, 1, -1), \frac{1}{2}(1, 1, 1, 1))$ ist ONB von $L((1, -1, 1, -1), (1, 1, 1, 1))$.
(4) Sei $b_1 = (3, 4), b_2 = (-4, 3)$. Dann ist $(\frac{b_1}{5}, \frac{b_2}{5})$ ONB des $\mathbb{R}^2$.

ÜBUNGSAUFGABEN 7.9.

(1) (Orthonormalbasen) Welche der folgenden Basen sind ONB?
  (a) $(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{13}}\begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}, \frac{-1}{\sqrt{13}}\begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix})$.
  (b) $(\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$.
  (c) $(\begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix}, \begin{pmatrix} -0.8 \\ 0.6 \end{pmatrix})$.

SATZ 7.10. *Sei $B = (b_1, \ldots, b_k)$ eine ONB und $v \in L(B)$ mit den Koordinaten $(v)_B = (\lambda_1, \ldots, \lambda_k)$. Dann gilt*

$$\|v\|^2 = \lambda_1^2 + \cdots + \lambda_k^2.$$

*Beweis.*Da $B$ eine ONB ist, gilt

$$\|v\|^2 = \langle v, v\rangle = \langle \sum_{i=1}^{k} \lambda_i\, b_i, \sum_{j=1}^{k} \lambda_j\, b_j\rangle = \sum_{i=1}^{k}\sum_{j=1}^{k} \lambda_i\, \lambda_j\, \langle b_i, b_j\rangle = \sum_{i=1}^{k} \lambda_i\, \lambda_i\, \langle b_i, b_i\rangle$$
$$= \sum_{i=1}^{k} \lambda_i^2 \cdot 1.$$

■

Wir wenden diesen Satz an:

Sei $b_1 = (0.6, 0.8)$, $b_2 = (-0.8, 0.6)$. Berechnen Sie die Länge von $v = 12 \cdot b_1 + 5 \cdot b_2$!

Lösung: Da $(b_1, b_2)$ eine ONB des $\mathbb{R}^2$ ist, gilt

$$\|v\| = \sqrt{12^2 + 5^2} = 13.$$

Wir kontrollieren das Ergebnis, indem wir $v$ berechnen. Da

$$v = 12 \cdot (0.6, 0.8) + 5 \cdot (-0.8, 0.6) = (3.2, 12.6),$$

gilt

$$\|v\| = \sqrt{(3.2)^2 + (12.6)^2} = 13.$$

SATZ 7.11. *Sei $T$ ein Unterraum des $\mathbb{R}^n$ mit der ONB $B = (b_1, \dots, b_k)$. Sei $v \in L(B)$ mit den Koordinaten $(v)_B = (\lambda_1, \dots, \lambda_k)$. Dann gilt*

$$\lambda_i = \langle v, b_i\rangle \text{ für alle } i \in \{1, \dots, k\}.$$

*Beweis.* Sei $i \in \{1, \dots, k\}$. Wir wissen, dass

$$\sum_{j=1}^{k} \lambda_j\, b_j = v,$$

und somit

$$\langle \sum_{j=1}^{k} \lambda_j\, b_j, b_i\rangle = \langle v, b_i\rangle.$$

Daraus folgt, dass

$$\sum_{j=1}^{k} \lambda_j\, \langle b_j, b_i\rangle = \langle v, b_i\rangle$$

oder

$$\lambda_i\, \langle b_i, b_i\rangle = \langle v, b_i\rangle.$$

Somit gilt

$$\lambda_i = \langle v, b_i\rangle.$$

■

Bei Orthonormalbasen müssen wir also kein Gleichungssystem lösen, um die Koordinaten eines Vektors zu bestimmen. Dazu folgendes Beispiel: Gegeben sei die ONB $B = ((0.6, 0.8), (-0.8, 0.6))$. Gesucht sind die Koordinaten $(\lambda_1, \lambda_2)$ von $v = (2, 1)$ bzgl. $B$.

$$\lambda_1 = \langle v, b_1 \rangle = 2, \quad \lambda_2 = \langle v, b_2 \rangle = -1.$$

Also gilt:

$$(v)_B = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Eine Orthonormalbasis besitzt also zwei wesentliche Vorteile gegenüber normalen Basen:

(1) Mit $(v)_B = (\lambda_1, \dots, \lambda_k)$ kann die Länge des Vektors $v$ als

$$||v|| = \sqrt{\lambda_1^2 + \cdots + \lambda_k^2}$$

berechnet werden.

(2) Die Koordinaten eines Vektors $v$ erhält man als

$$(v)_B = \begin{pmatrix} \langle v, b_1 \rangle \\ \vdots \\ \langle v, b_k \rangle \end{pmatrix}.$$

## 4. Das Gram-Schmidtsche Orthonormalisierungsverfahren

Gegeben ist ein Unterraum $T$ mit der Basis $B = (b_1, \dots, b_m)$. Gesucht ist eine ONB $(d_1, \dots, d_m)$ von $T$.

Wir konstruieren $d_1, \dots, d_m$ mit folgenden Eigenschaften:

(1) $d_i \perp d_j$ für $i \neq j$,
(2) $||d_i|| = 1$ für alle $i$,
(3) für alle $k \leq m$ gilt $L(b_1, \dots, b_k) = L(d_1, \dots, d_k)$.

(1) und (2) sollen garantieren, dass wir eine ONB erhalten. (3) fordert, dass die ersten $k$ Vektoren der gesuchten ONB den gleichen Vektorraum aufspannen wie die ersten $k$ Vektoren der Basis $B$.

ALGORITHMUS 7.12.

(1) *Konstruktion von $d_1$*: Wir setzen $c_1 = b_1$ und $d_1 = \frac{c_1}{||c_1||}$.
(2) *Konstruktion von $d_2$*: Um (3) erfüllen zu können, machen wir den Ansatz

$$c_2 = b_2 + \alpha_1 \cdot d_1$$

mit $\langle c_2, d_1 \rangle = 0$. Aus

$$\langle b_2 + \alpha_1 \cdot d_1, d_1 \rangle = \langle b_2, d_1 \rangle + \alpha_1 \langle d_1, d_1 \rangle = \langle b_2, d_1 \rangle + \alpha_1 = 0$$

folgt

$$\alpha_1 = -\langle b_2, d_1\rangle.$$

Also erhalten wir mittels

$$c_2 = b_2 - \langle b_2, d_1\rangle \cdot d_1,$$
$$d_2 = \frac{c_2}{||c_2||}$$

den gewünschten Vektor $d_2$.

(3) *Konstruktion von $d_3$*: Für $c_3$ machen wir analog zu $c_2$ den Ansatz

$$c_3 = b_3 + \alpha_1 \cdot d_1 + \alpha_2 \cdot d_2.$$

Fordern wir dann Normalität zu den ersten beiden Vektoren, so erhalten wir

$$\langle c_3, d_1\rangle = \langle b_3, d_1\rangle + \alpha_1 \underbrace{\langle d_1, d_1\rangle}_{=1} + \alpha_2 \underbrace{\langle d_2, d_1\rangle}_{=0} = 0,$$

also

$$\alpha_1 = -\langle b_3, d_1\rangle,$$

und

$$\langle c_3, d_2\rangle = \langle b_3, d_2\rangle + \alpha_1 \underbrace{\langle d_1, d_2\rangle}_{=0} + \alpha_2 \underbrace{\langle d_2, d_2\rangle}_{=1} = 0,$$

also

$$\alpha_2 = -\langle b_3, d_2\rangle.$$

Somit ergibt sich $c_3$ als

$$c_3 = b_3 - \langle b_3, d_1\rangle \cdot d_1 - \langle b_3, d_2\rangle \cdot d_2$$

und der gesuchte Vektor $d_3$ der Länge 1 als

$$d_3 = \frac{c_3}{||c_3||}.$$

(4) *Konstruktion von $d_i$*: Seien $d_1, \ldots, d_{i-1}$ bereits bekannt. Dann setzen wir

$$c_i = b_i - \langle b_i, d_1\rangle \cdot d_1 - \cdots - \langle b_i, d_{i-1}\rangle \cdot d_{i-1}$$

und erhalten mit

$$d_i = \frac{c_i}{||c_i||}$$

den gewünschten Vektor.

AUFGABEN 7.13. (1) Sei

$$B = (\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 3 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}).$$

Gesucht ist eine ONB von $L(B)$.

Offenbar ist $c_1 = (1, -1, 1, -1)$ und somit

$$d_1 = \frac{c_1}{||c_1||} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}.$$

Dann ist

$$c_2 = \begin{pmatrix} -1 \\ 1 \\ 3 \\ -3 \end{pmatrix} - \langle(-1, 1, 3, -3), d_1\rangle \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 2 \\ -2 \end{pmatrix}$$

und

$$d_2 = \frac{c_2}{||c_2||} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

Schließlich ist

$$c_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \langle(1, 0, 0, 0), d_1\rangle\, d_1 - \langle(1, 0, 0, 0), d_2\rangle\, d_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

und

$$d_3 = \frac{c_3}{||c_3||} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

(2) Sei $B = ((3, 4), (5, 6))$. Gesucht ist eine ONB von $L(B)$.
Wir erhalten

$$d_1 = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix},$$

$$d_2 = \begin{pmatrix} 0.8 \\ -0.6 \end{pmatrix}.$$

Man prüft leicht nach, dass die beiden Vektoren tatsächlich normal aufeinander stehen.

SATZ 7.14. *Sei $T$ ein Unterraum des $\mathbb{R}^n$. Dann hat $T$ eine ONB.*

*Beweis.* Aus einer beliebigen Basis von $T$ kann mit dem Gram-Schmidtschen Orthonormalisierungsverfahren eine ONB berechnet werden. ■

ÜBUNGSAUFGABEN 7.15.

(1) Orthonormalisieren Sie die folgende Familie von Vektoren mit dem Verfahren von Gram-Schmidt!

$$A = (\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\3\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}).$$

(2) Bestimmen Sie jeweils eine Orthonormalbasis (ONB) für folgende Unterräume des $\mathbb{R}^3$!
   (a) $U = L(\begin{pmatrix}1\\2\\-2\end{pmatrix}, \begin{pmatrix}3\\2\\-1\end{pmatrix})$.
   (b) $V = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + 2x_3 + 2x_4 = 0 \text{ und } x_2 - 2x_3 + 2x_4 = 0\}$. (Hinweis: Berechnen Sie zuerst eine Basis von $V$ und orthonormalisieren Sie diese mit dem Verfahren von Gram-Schmidt).
(3) Geben Sie eine Orthonormalbasis der Ebene $e : x - 2y + 3z = 0$ an.
(4) Bestimmen Sie mithilfe des Verfahrens von Gram-Schmidt eine Orthonormalbasis von

$$\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 - x_2 + x_3 - x_4 = 0\}.$$

## 5. Orthogonalprojektionen

SATZ 7.16. *Sei $T$ ein Unterraum des $\mathbb{R}^n$, $v \in \mathbb{R}^n$. Dann gibt es genau ein $x \in \mathbb{R}^n$ mit den Eigenschaften*

(1) $x \in T$,
(2) $v - x \in T^\perp$.

Um diesen Satz zu zeigen, brauchen wir zwei Resultate als Vorbereitung:

SATZ 7.17. *Sei $n \in \mathbb{N}$, und sei $A \in \mathbb{R}^{n \times n}$. Wir nehmen an, dass die Spalten von $A$ linear unabhängig sind. Dann ist $A$ invertierbar.*

*Beweis:* Sei $A = \begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix}$. Da $(s_1, \ldots, s_n)$ linear unabhängig ist, gilt $\dim(L(s_1, \ldots, s_n)) = n$, und somit $L(s_1, \ldots, s_n) = \mathbb{R}^n$. Also gibt es $\lambda_{1,1}, \ldots, \lambda_{1,n} \in \mathbb{R}$, sodass

$$\begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix} \cdot \begin{pmatrix}\lambda_{1,1}\\ \vdots \\ \lambda_{1,n}\end{pmatrix} = \begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}.$$

Insgesamt erhalten wir auf diese Weise eine Matrix $R$ mit

$$\begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix} \cdot R = E_n.$$

Wir zeigen nun, dass auch die Zeilen von $A$ linear unabhängig sind. Sei dazu $A = \begin{pmatrix} — & z_1 & — \\ & \vdots & \\ — & z_n & — \end{pmatrix}$, und seien $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ so, dass $\sum_{i=1}^n \alpha_i * z_i = 0$. Dann gilt

$$(\alpha_1\, \alpha_2 \ldots \alpha_n) \cdot \begin{pmatrix} — & z_1 & — \\ & \vdots & \\ — & z_n & — \end{pmatrix} = (0\, 0 \ldots 0).$$

Sei $\overline{\alpha} := (\alpha_1\, \alpha_2 \dots \alpha_n)$. Dann gilt $\overline{\alpha} = \overline{\alpha} \cdot E_n = \overline{\alpha} \cdot (A \cdot R) = (\overline{\alpha} \cdot A) \cdot R = 0 \cdot R = 0$. Also sind alle $\alpha_i = 0$. Folglich sind die Zeilen von $A$ linear unabhängig.

Es gilt also $L(z_1, \dots, z_n) = \mathbb{R}^n$. Genauso wie die Matrix $R$ können wir nun eine Matrix $L$ finden, sodass $L \cdot A = E_n$. Es gilt dann $L = L \cdot E_n = L \cdot (A \cdot R) = (L \cdot A) \cdot R = E_n \cdot R = R$, und somit gilt $L = R$. Dann ist $A$ invertierbar mit inverser Matrix $A^{-1} := R$. ■

LEMMA 7.18. *Seien $b_1, \dots, b_k \in \mathbb{R}^n$ linear unabhängig. $B$ sei definiert als die $n \times k$-Matrix mit den Spalten $b_1, \dots, b_k$. Dann ist $B^T \cdot B$ invertierbar.*

Wir zeigen, dass die Spaltenvektoren von $B^T \cdot B$ linear unabhängig sind. Sei dazu $v \in \mathbb{R}^{k \times 1}$ so, dass $(B^T \cdot B) \cdot v = 0$. Dann gilt $v^T \cdot B^T \cdot B \cdot v = 0$, also $(B \cdot v)^T \cdot (B \cdot v) = (0)$. Dann gilt $B \cdot v = 0$. Da die Spalten von $B$ linear unabhängig sind, gilt also $v = 0$. Somit sind die Spaltenvektoren von $B^T \cdot B$ linear unabhängig. Nun folgt aus Satz 7.17, dass $B^T \cdot B$ invertierbar ist. ■

*Beweis von Satz 7.16.*

- Zunächst zeigen wir, dass so ein $x$ existiert. Sei $(b_1, \dots, b_k)$ eine Basis von $T$, und sei $B$ die $n \times k$-Matrix, in deren Spalten die Vektoren $b_1, \dots, b_k$ stehen. Wir versuchen $(\alpha_1, \dots, \alpha_k)$ zu finden, sodass
$$x := \sum_{i=1}^{k} \alpha_i \cdot b_i = B \cdot \underbrace{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}}_{\alpha}$$
die Eigenschaft $v - B \cdot \alpha \in T^\perp$ erfüllt. Dann muss gelten: $\langle b_j, v - x\rangle = \langle b_j, v - B \cdot \alpha\rangle = 0$ für alle $j$, oder, anders ausgedrückt,
$$\begin{pmatrix} \langle b_1, v - B \cdot \alpha\rangle \\ \langle b_2, v - B \cdot \alpha\rangle \\ \vdots \\ \langle b_k, v - B \cdot \alpha\rangle \end{pmatrix} = B^T \cdot (v - B \cdot \alpha) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$
und somit
$$B^T \cdot v - B^T \cdot B \cdot \alpha = 0,$$
also
$$B^T \cdot v = (B^T \cdot B) \cdot \alpha.$$
Da die Basisvektoren linear unabhängig sind, ist nach Lemma 7.18 die Matrix $B^T \cdot B$ invertierbar. Somit muss $\alpha$ die Gleichung
$$(B^T \cdot B)^{-1} \cdot B^T \cdot v = (B^T \cdot B)^{-1}(B^T \cdot B) \cdot \alpha = \alpha$$
erfüllen. Wir erhalten als Kandidaten für $x$ also
$$x = B \cdot (B^T \cdot B)^{-1} \cdot B^T \cdot v.$$

Wir überprüfen jetzt, ob $v - x$ wirklich in $T^\perp$ liegt, und erhalten

$$\begin{aligned} B^T(v-x) &= B^T(v - B\cdot(B^T\cdot B)^{-1}\cdot B^T\cdot v) \\ &= B^T\cdot v - B^T\cdot B\cdot(B^T\cdot B)^{-1}\cdot B^T\cdot v) \\ &= B^Tv - B^Tv = 0. \end{aligned}$$

- Eindeutigkeit: Seien $x, y$ zwei Vektoren, die den Bedingungen des Satzes 7.16 genügen, d.h.
  (1) $x, y \in T$,
  (2) $v - x, v - y \in T^\perp$.
  Wegen (1) gilt $x - y \in T$. Aus (2) erhalten wir $(v-x)-(v-y) = -x+y \in T^\perp$. Daher gilt auch $-(-x+y) = x - y \in T^\perp$. Da $x - y \in T$ und $x - y \in T^\perp$, gilt $\langle x-y, x-y\rangle = 0$. Also gilt $\|(x-y)\|^2 = 0$, und folglich $x = y$.

Somit existiert genau ein $x$, das die obigen Eigenschaften erfüllt. ■

DEFINITION 7.19. Sei $n \in \mathbb{N}$, sei $T$ ein Unterraum des $\mathbb{R}^n$, und sei $v \in \mathbb{R}^n$. Das $x \in T$ mit $v - x \in T^\perp$ nennen wir *orthogonale Projektion* von $v$ auf $T$ und schreiben $x = P_T(v)$.

Wenn $T$ der Spaltenraum der Matrix $B$ mit linear unabhängigen Spalten $(b_1, \ldots, b_k)$ ist, also $T = S(B)$, dann ist $(b_1, \ldots, b_k)$ eine Basis von $T$. Wie im Beweis von Satz 7.16 ermittelt, gilt dann

$$P_T(v) = B\cdot(B^T\cdot B)^{-1}\cdot B^Tv.$$

AUFGABEN 7.20.

(1) Gesucht ist die orthogonale Projektion von $(4,2)$ auf die Gerade $G = \{(x,y) \mid 3x - 4y = 0\}$.
    Lösung: Für $B = \binom{4}{3}$ gilt $G = S(B)$. Wegen $B^T\cdot B = (4^2+3^2) = (25)$ und $B^T\cdot v = (16+6) = (22)$ ist

$$P_G(v) = B\cdot(B^T\cdot B)^{-1}\cdot B^T\cdot v = \frac{22}{25}\begin{pmatrix}4\\3\end{pmatrix}.$$

(2) Wir berechnen die orthogonale Projektion des Vektors $v = (6,0,0)$ auf die Ebene $e = L((1,2,1),(-1,0,1))$. Wir bilden eine Matrix $B$, in deren Spalten

wir die Basisvektoren von $e$ schreiben. In diesem Fall gilt

$$B^T \cdot B = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix},$$

$$B^T \cdot v = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ -6 \end{pmatrix}$$

$$P_e(v) = B \cdot (B^T \cdot B)^{-1} \cdot B^T \cdot v = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 6 \\ -6 \end{pmatrix}$$

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

ÜBUNGSAUFGABEN 7.21.
Wir wollen die Pyramide mit der quadratischen Grundfläche $(A, B, C, D)$ und der Spitze $S$ als $2D$-Graphik auf dem Bildschirm darstellen. Lösen Sie dazu folgende Probleme:

(1) Sei die Ebene $e = L(B)$ durch ihre Basis
$$B = \left(\frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{19}}\begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}\right)$$
gegeben. Bestimmen Sie für $v = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$ die Projektion von $v$ auf $L(B)$, also $P_{L(B)}(v)$.

(2) Bestimmen Sie die Koordinaten von $P_{L(B)}(v)$ bezüglich $B$.

(3) Bestimmen Sie eine Orthonormalbasis der Ebene $e$, die durch
$$e : x - 2y + 2z = 0$$
gegeben ist.

(4) Welcher Punkt auf $e$ hat von $\begin{pmatrix} -2 \\ -5 \\ 5 \end{pmatrix}$ den geringsten Abstand?

SATZ 7.22. *Sei $T$ ein Unterraum des $\mathbb{R}^n$, $v \in \mathbb{R}^n$. Dann gilt*
$$P_{T^\perp}(v) = v - P_T(v).$$

*Beweisskizze.* Wir definieren $y := v - P_T(v)$. Man kann zeigen, dass $y$ die Eigenschaften der Projektion von $v$ auf $T^\perp$ erfüllt. ■

Aus diesem Satz können wir sofort eine Formel für die Projektion auf $T^\perp$ ableiten. Sei $B$ eine Matrix mit linear unabhängigen Spalten, sodass $T = S(B)$. Dann gilt:
$$P_{T^\perp}(v) = (E_n - B \cdot (B^T \cdot B)^{-1} \cdot B^T) \cdot v.$$

ÜBUNGSAUFGABEN 7.23.

(1) (a) Bestimmen Sie für jeden Vektor $v = (x, y, z) \in \mathbb{R}^3$ die Projektion auf $U := \{(x, y, z) \mid -x + 2y + 2z = 0\}$.
(b) Bestimmen Sie eine Matrix $P_U$, sodass die Projektion von $v$ auf $U$ gleich $P_U \cdot v$ ist.
(c) Bestimmen Sie $P_U \cdot P_U$!

(2) Sei $U := \{(x, y, z) \mid -x + 2y + 2z = 0\}$.
(a) Bestimmen sie für jeden Vektor $v = (x, y, z) \in \mathbb{R}^3$ die Projektion von $v$ auf den Orthogonalraum von $U$, also auf $U^\perp$.

(b) Bestimmen Sie eine Matrix $P_{U^\perp}$, sodass die Projektion von $v$ auf $U^\perp$ gleich $P_{U^\perp} \cdot v$ ist.
(c) Bestimmen Sie $P_{U^\perp} \cdot P_{U^\perp}$.

## 6. Abstandsberechnungen mit Hilfe der Orthogonalprojektion

**6.1. Abstand eines Punktes von einem Unterraum.** Gegeben sei ein Unterraum $T$ des $\mathbb{R}^n$ sowie ein Vektor $v \in \mathbb{R}^n$. Gesucht ist jenes $x \in T$, das von $v$ den kleinsten Abstand hat.

SATZ 7.24. *Sei $T$ ein Unterraum des $\mathbb{R}^n$, $v \in \mathbb{R}^n$, und sei $x \in T$. Dann sind äquivalent:*

(1) $||v - x|| = \inf\{||v - t|| \, : \, t \in T\}$.
(2) *Der Vektor $x$ ist die Projektion von $v$ auf $T$, also $x = P_T(v)$.*

*Beweis.* (2)⇒(1): Wir nehmen an, $x = P_T(v)$. Sei $t \in T$. Dann gilt

$$||v - t||^2 = ||(v - x) + (x - t)||^2.$$

Da $v - x \in T^\perp$ und $x - t \in T$, gilt

$$||(v - x) + (x - t)||^2 = ||v - x||^2 + ||x - t||^2.$$

Es gilt also

$$\text{Für alle } t \in T\text{: } ||v - t|| \geq ||v - x||.$$

Also gilt

$$||v - x|| = \inf\{||v - t|| \, : \, t \in T\}.$$

(1)⇒(2): Sei $x$ so, dass $||v - x|| = \inf\{||v - t|| \, : \, t \in T\}$, und sei $y = P_T(v)$. Es gilt dann $||v - x|| \leq ||v - y||$ und $||v - x||^2 = ||(v - y) + (y - x)||^2 = ||v - y||^2 + ||y - x||^2$. Das ergibt ingesamt $||v - y||^2 \geq ||v - y||^2 + ||y - x||^2$, also $||y - x||^2 = 0$. Somit gilt $x = y = P_T(v)$. ■

ÜBUNGSAUFGABEN 7.25.

(1) Welcher Punkt auf der Geraden

$$g = L(\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix})$$

hat von $\begin{pmatrix} 2 \\ -1 \\ 9 \end{pmatrix}$ den kleinsten Abstand?

(2) Wir betrachten den Punkt $p \in \mathbb{R}^4$ und den Unterraum $U$ des $\mathbb{R}^4$.

$$p = (8, 10, 1, 14)\,,$$

$$U = L((-1, -2, 0, 3)).$$

Welcher Punkt $p_0$ aus $U$ hat von $p$ den kleinsten Abstand?

**6.2. Abstand zwischen zwei linearen Mannigfaltigkeiten.** Wenn $A$ und $B$ nichtleere Teilmengen des $\mathbb{R}^n$ sind, dann kann man ihren *Abstand* durch

$$\operatorname{dist}(A, B) = \inf\{\|a - b\| \mid a \in A, b \in B\}$$

definieren.

Gegeben sind zwei lineare Mannigfaltigkeiten $M_1 = v_1 + T_1$ und $M_2 = v_2 + T_2$. Dabei sind $v_1, v_2 \in \mathbb{R}^n$, und $T_1, T_2$ sind Unterräume des $\mathbb{R}^n$. Wir suchen $p_1 \in M_1$ und $p_2 \in M_2$, sodass

$$\|p_1 - p_2\| = \operatorname{dist}(M_1, M_2).$$

Dazu bestimmen wir $t_1 \in T_1$ und $t_2 \in T_2$, sodass

$$\|(v_1 + t_1) - (v_2 + t_2)\| = \operatorname{dist}(M_1, M_2).$$

Nun gilt

$$\begin{aligned}\operatorname{dist}(M_1, M_2) &= \inf\{\|(v_1 + t_1) - (v_2 + t_2)\| \mid t_1 \in T_1, t_2 \in T_2\}\\ &= \inf\{\|(v_1 - v_2) - (t_2 - t_1)\| \mid t_1 \in T_1, t_2 \in T_2\}\\ &= \inf\{\|(v_1 - v_2) - s\| \mid s \in T_1 + T_2\}.\end{aligned}$$

Wir wissen, dass $(v_1 - v_2) - \overline{s}$ genau dann das Infimum von $\{(v_1 - v_2) - s \mid s \in T_1 + T_2\}$ annimmt, wenn $\overline{s} = P_{T_1+T_2}(v_1 - v_2)$. Also berechnen wir

$$\overline{s} := P_{T_1+T_2}(v_1 - v_2).$$

Folglich gilt für $\overline{t_1} \in T_1$ und $\overline{t_2} \in T_2$ die Gleichung

$$(v_1 + \overline{t_1}) - (v_2 + \overline{t_2}) = \inf\{(v_1 + t_1) - (v_2 + t_2) \mid t_1 \in T_1, t_2 \in T_2\}$$

genau dann, wenn $\overline{t_2} - \overline{t_1} = \overline{s}$. Wir müssen also alle $(\overline{t_1}, \overline{t_2}) \in T_1 \times T_2$ bestimmen, sodass $\overline{t_2} - \overline{t_1} = \overline{s}$. Wir bestimmen uns also ein $t_1' \in T_1$ und ein $t_2' \in T_2$, sodass $t_2' - t_1' = \overline{s}$. Dann gilt

$$\{(t_1, t_2) \in T_1 \times T_2 \mid t_1 - t_2 = s\} = \{t_1' + r, t_2' + r) \mid r \in T_1 \cap T_2\}.$$

Damit erhalten wir

$$\{(p_1, p_2) \in M_1 \times M_2 \mid \|p_1 - p_2\| = \operatorname{dist}(M_1, M_2)\} = \{\big(v_1 + t_1' + r, v_2 + t_2' + r\big) \mid r \in T_1 \cap T_2\}.$$

AUFGABE 7.26. Bestimmen Sie jene Punkte auf den Mannigfaltigkeiten

$$M_1 := (5, 0, 0, 0) + L((1, 2, -1, 0), (1, 1, 1, 1))$$

und

$$M_2 := (11, 27, 2, 11) + L((1, -1, -1, 1), (1, 0, 0, 1)),$$

die voneinander geringsten Abstand haben.

Sei $T_1 = L((1, 2, -1, 0), (1, 1, 1, 1))$ und $T_2 = L((1, -1, -1, 1), (1, 0, 0, 1))$. Wir berechnen eine Basis von $T_1 + T_2$.

```
In[102]:= B1 = {{1, 2, -1, 0}, {1, 1, 1, 1}};

          B2 = {{1, -1, -1, 1}, {1, 0, 0, 1}};
```

```
Out[102]= {{1, -1, -1, 1}, {1, 0, 0, 1}}
```

```
In[103]:= BasisVonT1PlusT2 = RowReduce [B1 ~Join~ B2]
```

$$\mathtt{Out[103]=}\ \left\{\{1, 0, 0, 1\}, \left\{0, 1, 0, -\frac{1}{3}\right\}, \left\{0, 0, 1, \frac{1}{3}\right\}, \{0, 0, 0, 0\}\right\}$$

Also erhalten wir

$$B = ((1, 0, 0, 1), (0, 1, 0, -1/3), (0, 0, 1, 1/3))$$

als Basis von $T_1 + T_2$. Jetzt können wir die Projektion von $v_1 - v_2$ auf $T_1 + T_2$ ausrechnen. Sei $C$ die Matrix

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

Dann gilt

$$P_{T_1+T_2}(v_1 - v_2) = C \cdot (C^T \cdot C)^{-1} \cdot C^T \cdot (v_1 - v_2) = \begin{pmatrix} -12 \\ -25 \\ -4 \\ -5 \end{pmatrix}.$$

Also gilt $\overline{s} = (-12, -25, -4, -5)$. Dieses $\overline{s}$ wollen wir als $-t_1' + t_2'$ mit $t_1' \in T_1$ und $t_2' \in T_2$ darstellen. Durch die folgenden Mathematica-Berechnungen erhalten wir

$$\overline{s} = -7 \cdot (1, 2, -1, 0) - 8 \cdot (1, 1, 1, 1) + 3 \cdot (1, -1, -1, 1) + 0 \cdot (1, 0, 0, 1).$$

```
In[104]:= s
Out[104]= {-12, -25, -4, -5}

In[105]:= B1B2 = B1 ~Join~ B2
Out[105]= {{1, 2, -1, 0}, {1, 1, 1, 1}, {1, -1, -1, 1}, {1, 0, 0, 1}}

In[106]:= M = Transpose [B1B2]
Out[106]= {{1, 1, 1, 1}, {2, 1, -1, 0}, {-1, 1, -1, 0}, {0, 1, 1, 1}}

In[107]:= MatrixForm [M]
```

$$\mathtt{Out[107]=}\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

```
In[108]:= LinearSolve [M, s]
Out[108]= {-7, -8, 3, 0}
```

Für $t_1' = 7 \cdot (1, 2, -1, 0) + 8 \cdot (1, 1, 1, 1)$ und $t_2' = 3 \cdot (1, -1, -1, 1)$ gilt also

$$\overline{s} = -t_1' + t_2'.$$

Wir erhalten

$$t_1' = (15, 22, 1, 8)\,,\ \ t_2' = (3, -3, -3, 3).$$

Also sind die Punkte

$$p_1 = (5, 0, 0, 0) + (15, 22, 1, 8) = (20, 22, 1, 8)$$

und

$$p_2 = (11, 27, 2, 11) + (3, -3, -3, 3) = (14, 24, -1, 14)$$

so, dass $p_1 \in M_1$, $p_2 \in M_2$, und $\|p_1 - p_2\| = \mathrm{dist}(M_1, M_2)$.

Um alle Punkte zu erhalten, für die der Abstand minimal wird, müssen wir noch $T_1 \cap T_2$ berechnen.

```
In[109]:= C1 = NullSpace[{{1, 2, -1, 0}, {1, 1, 1, 1}}];
          MatrixForm[C1]
```

Out[109]= $\begin{pmatrix} -2 & 1 & 0 & 1 \\ -3 & 2 & 1 & 0 \end{pmatrix}$

```
In[110]:= C2 = NullSpace[{{1, -1, -1, 1}, {1, 0, 0, 1}}];
          MatrixForm[C2]
```

Out[110]= $\begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}$

```
In[111]:= NullSpace[Join[C1, C2]]
Out[111]= {{1, 1, 1, 1}}
```

Wir erhalten $T_1 \cap T_2 = L((1, 1, 1, 1))$. Somit ist die Menge aller Paare in $M_1 \times M_2$, die minimalen Abstand voneinander haben, gegeben als

$$\{\big((20, 22, 1, 8) + t \cdot (1, 1, 1, 1), (14, 24, -1, 14) + t \cdot (1, 1, 1, 1)\big) \mid t \in \mathbb{R}\}.$$

Folgendes Mathematica-Programm liefert zwei Punkte mit minimaler Distanz. Achtung: die Basisvektoren von $T_1$ und $T_2$ stehen hier in den *Zeilen* der Argumente $B_1$ und $B_2$.

```
Proj [B_, v_] :=
   Transpose [B] . Inverse [B . Transpose[B]]. B. v;


PointsWithMinimalDistance [v1_, B1_, v2_, B2_] :=
Module[{},
   B12 = RowReduce [B1 ~Join~ B2];
   If [MatrixRank[B12] == 0,
         {v1, v2},
      Basis12 = Take [B12, MatrixRank [B12]];
      s = Proj [Basis12, v1 - v2];
      t1 = - Transpose [B1].Take [LinearSolve
              [Transpose [B1 ~Join~ B2], s], Length [B1]];
      t2 = s + t1;
      {v1 + t1, v2 + t2}
  ]
];
```

Wir testen dieses Programm:

```
In[112]:= << abstand2.m

In[113]:= v1 = {5, 0, 0, 0};

          v2 = {11, 27, 2, 11};

          T1 = {{1, 1, 1, 1}, {1, 2, -1, 0}};

          T2 = {{1, 0, 0, 1}, {1, -1, -1, 1}};

In[114]:= PointsWithMinimalDistance[v1, T1, v2, T2]
Out[114]= {{23, 25, 4, 11}, {17, 27, 2, 17}}
```

ÜBUNGSAUFGABEN 7.27.

(1) Welcher Punkt auf der Geraden

$$g : X = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

hat von $\begin{pmatrix} 6 \\ -7 \end{pmatrix}$ den geringsten Abstand?

(2) Bestimmen Sie den Punkt auf $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + L(\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix})$, der am nächsten beim Punkt $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ liegt.

(3) (a) Bestimmen Sie eine Orthonormalbasis des Unterraums $W$ von $\mathbb{R}^4$, der durch

$$W = \{(x_1, x_2, x_3, x_4) \mid -x_1 + x_3 = 0 \text{ und } -x_2 + x_4 = 0\}.$$

gegeben ist.

(b) Welcher Punkt in $W$ hat von $(4, -1, 6, -1)$ den geringsten Abstand?

(4) Bestimmen Sie den Abstand des Punktes $\begin{pmatrix} -6 \\ 4 \\ -2 \end{pmatrix}$ von der Geraden $g$, die durch den Punkt $\begin{pmatrix} 0 \\ 0 \\ 10 \end{pmatrix}$ geht, und deren Richtungsvektor gleich $\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$ ist.

(5) Bestimmen Sie den Abstand des Punktes $(8, 2, 11, 17)$ von der Geraden $g$, die durch

$$g : X = (1, 2, -4, 7) + t \cdot (12, 0, 3, 4)$$

gegeben ist. Welcher Geradenpunkt ist diesem Punkt am nächsten?

(6) (a) Welcher Punkt $p$ auf der Geraden

$$g : -2x + z = 0 \text{ und } 2x + y = 0$$

hat vom Punkt $v = \begin{pmatrix} -10 \\ -25 \\ 25 \end{pmatrix}$ den geringsten Abstand?

(b) Wie groß ist dieser Abstand?

(c) Berechnen Sie $v - p$! Welchen Winkel schließen $p$ und $v - p$ miteinander ein?

(7) Welcher Punkt der Ebene

$$e : X = \begin{pmatrix} -24 \\ -53 \\ 13 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} -1 \\ 7 \\ 2 \end{pmatrix}$$

kommt dem Punkt $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ am nächsten ?

(8) Welche Punkte der Geraden

$$g_1 : X = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

und

$$g_1 : X = \begin{pmatrix} -24 \\ -52 \\ 15 \end{pmatrix} + \mu \cdot \begin{pmatrix} -1 \\ 7 \\ 2 \end{pmatrix}$$

kommen einander am nächsten?

(9) Welcher Punkt der Ebene

$$e : X = \begin{pmatrix} -24 \\ -53 \\ 13 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} -1 \\ 7 \\ 2 \end{pmatrix}$$

kommt dem Punkt $\begin{pmatrix} -47 \\ -95 \\ 34 \end{pmatrix}$ am nächsten?

(10) Welcher Punkt der Ebene

$$e : X = (5, 2, 1, 0) + \lambda \cdot (1, 2, 0, 0) + \mu \cdot (0, 1, 2, -1)$$

kommt der Geraden

$$g : X = (5, 3, 0, 5) + \alpha \cdot (1, 0, 2, 0)$$

am nächsten?

(11) Berechnen Sie den Inkreismittelpunkt $I = \binom{i_1}{i_2}$ des Dreiecks $ABC$ mit $A = \binom{-3}{-2}$, $B = \binom{4}{0}$ und $C = \binom{2}{4}$, indem Sie die Bedingung, dass $I$ gleich weit von $AB$, $BC$ und $AC$ entfernt ist, in Gleichungen in den Variablen $i_1$ und $i_2$ umwandeln. Verwenden Sie zur Lösung der auftretenden Gleichungen den Mathematica-Befehl `Solve`.

## 7. Die bestapproximierende Lösung eines linearen Gleichungssystems

Oft trifft man auf Gleichungssysteme, die keine Lösung besitzen, wie z.B.

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

In solchen Fällen wird man ein $x$ suchen, das dem Gleichungssystem “möglichst gut” genügt. Naheliegend ist es, jenes $x$ als “Lösung” zu betrachten, das den kleinsten Fehler liefert, d.h. für das $||A \cdot x - b||$ minimal ist.

DEFINITION 7.28. Sei $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Eine *bestapproximierende Lösung* des Gleichungssystems $A \cdot x = b$ ist ein $x^* \in \mathbb{R}^n$, sodass $||A \cdot x^* - b||$ minimal ist.

Wir wissen, dass $\{A \cdot x \,|\, x \in \mathbb{R}^n\}$ genau der Spaltenraum von $A$ ist. Wir suchen also jenes Element des Spaltenraumes von $A$, das von $b$ minimalen Abstand hat. Dazu sollte $A \cdot x$ die Projektion von $b$ auf den Spaltenraum von $A$ sein. Dann steht $b - A \cdot x$ auf den Spaltenraum von $A$ normal. Es gilt also

$$A^T \cdot (b - A \cdot x) = 0.$$

Wenn $A^T \cdot A$ invertierbar ist, dann können wir $x$ durch $x = (A^T \cdot A)^{-1} \cdot A^T \cdot b$ bestimmen.

SATZ 7.29. *Sei $A$ eine Matrix, deren Spaltenvektoren linear unabhängig sind, und sei $b$ ein Vektor. Dann ist die bestapproximierende Lösung $x^*$ von $A \cdot x = b$ gegeben durch*

$$x^* = (A^T \cdot A)^{-1} \cdot A^T \cdot b.$$

AUFGABE 7.30. Wir bestimmen die bestapproximierende Lösung von

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

Hier ist

$$A^T \cdot A = \begin{pmatrix} 6 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{und} \quad (A^T \cdot A)^{-1} = \frac{1}{11} \begin{pmatrix} 2 & -1 \\ -1 & 6 \end{pmatrix}.$$

Also erhalten wir

$$x^* = \frac{1}{11} \begin{pmatrix} 8 \\ 7 \end{pmatrix}.$$

ÜBUNGSAUFGABEN 7.31.

(1) Das Gleichungssystem $A \cdot x = b$ ist für folgendes $A$ und $b$ unlösbar:

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

Bestimmen Sie die “beste Näherungslösung”. Warum (bzw. in welchem Sinn) ist diese Lösung “besser” als $(1, 1)$?

Nun versuchen wir, die “beste Gerade durch eine Punktwolke” zu legen.

Gegeben seien die Punkte $(x_1, y_1), \ldots, (x_n, y_n)$. Gesucht ist die bestapproximierende Gerade $y = k\,x + d$ durch diese Punktwolke.

Dafür berechnet man die bestapproximierende Lösung des Gleichungssystems

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ d \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

wie oben beschrieben. Wir überlegen uns noch, welcher “Abstand” zwischen Punkten und Gerade hier wirklich minimiert wird. Die bestaproximierende Lösung minimiert

$$\left\| \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ d \end{pmatrix} \right\| = \left\| \begin{pmatrix} y_1 - (k\,x_1 + d) \\ \vdots \\ y_n - (k\,x_n + d) \end{pmatrix} \right\|,$$

also die Summe der Quadrate der Vertikalabstände zwischen den Punkten und der Gerade.

ÜBUNGSAUFGABEN 7.32.

(1) Bestimmen Sie jene Gerade der Form $y = k\,x + d$, die die Punkte $(0, 3)$, $(1, 4)$ und $(2, 7)$ bestmöglich approximiert. “Bestmöglich” heißt dabei, dass $k$ und $d$ so zu bestimmen sind, dass

$$(y_1 - (k\,x_1 + d))^2 + (y_2 - (k\,x_2 + d))^2 + (y_3 - (k\,x_3 + d))^2$$

minimal wird.

(2) Bestimmen Sie jene Gerade der Form $y = k\,x + d$, die die Punkte $(2, 3)$, $(3, 0)$ und $(6, 5)$ bestmöglich approximiert. “Bestmöglich” heißt dabei, dass $k$ und $d$ so zu bestimmen sind, dass

$$(y_1 - (k\,x_1 + d))^2 + (y_2 - (k\,x_2 + d))^2 + (y_3 - (k\,x_3 + d))^2$$

minimal wird.

KAPITEL 8

# Ringe, Körper und Vektorräume

Nachrichten codiert man gerne als Listen von Bits. Deswegen wird es sich als günstig herausstellen, dass man nicht nur mit Vektoren, deren Einträge reelle Zahlen sind, rechnen kann, sondern auch mit Vektoren, deren Einträge nur 0 oder 1 sein können. Unser Wissen über das Lösen linearer Gleichungssysteme lässt sich auf solche 0/1-Vektoren verallgemeinern; das ist z.B. in der Codierungstheorie hilfreich.

## 1. Kommutative Ringe mit Eins

Unser Ziel ist, anstelle der reellen Zahlen auch andere Objekte verwenden zu können, solange man diese Objekte sinnvoll addieren und multiplizieren kann.

Sei $n \in \mathbb{N}$. Eine *$n$-stellige Operation* auf $A$ ist eine Funktion von $A^n$ nach $A$. Eine *Algebra* ist ein Paar $(A, F)$, wobei $A$ eine nichtleere Menge, und $F$ eine Familie von endlichstelligen Operationen auf $A$ ist. Wenn $F = (f_1, \ldots, f_k)$, so bezeichnet man sowohl das Paar $(A, F)$ als auch das Tupel $(A, f_1, \ldots, f_k)$ als Algebra.

DEFINITION 8.1. Eine Algebra $\langle R, +, -, \cdot, 0, 1\rangle$ ist ein *kommutativer Ring mit Eins*, wenn $+, \cdot$ binäre Operationen auf $R$ sind, $-$ eine unäre Operation auf $R$ ist, und $0, 1$ Elemente aus $R$ sind, sodass für alle $x, y, z \in R$ die folgenden Eigenschaften erfüllt sind:

(1) $x + 0 = x$
(2) $x + (-x) = 0$
(3) $(x + y) + z = x + (y + z)$
(4) $x + y = y + x$
(5) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
(6) $x \cdot y = y \cdot x$
(7) $x \cdot 1 = x$
(8) $x \cdot (y + z) = x \cdot y + x \cdot z$.

SATZ 8.2. *Sei $\langle R, +, -, \cdot, 0, 1\rangle$ ein kommutativer Ring mit* 1, *und seien $x, y \in R$. Dann gilt*

(1) $-(-(x)) = x$
(2) $x \cdot 0 = 0$.
(3) $-(x \cdot y) = (-x) \cdot y = x \cdot (-y)$.

*Beweis:* (1): $-(-x) = -(-x) + 0 = 0 + (-(-x)) = (x + (-x)) + (-(-x)) = x + ((-x) + (-(-x)) = x + 0 = x$. (2): $x \cdot 0 = x \cdot 0 + 0 = x \cdot 0 + (x \cdot 0 + (-x \cdot 0)) = (x \cdot 0 + x \cdot 0) + (-x \cdot 0) = x \cdot (0 + 0) + (-x \cdot 0) = x \cdot 0 + (-x \cdot 0) = 0$. (3): Wir verwenden jetzt außer den bei der Definition von kommutativen Ringen verwendeten Gleichungen auch die Folgerungen, dass für alle $z \in R$ auch $(-z) + z = 0$ und $0 + z = z$ gilt. $-(x \cdot y) = -(x \cdot y) + x \cdot 0 = -(x \cdot y) + x \cdot (y + (-y)) = -(x \cdot y) + (x \cdot y + x \cdot (-y)) = (-(x \cdot y) + x \cdot y) + x \cdot (-y) = 0 + x \cdot (-y) = x \cdot (-y)$. Mithilfe des Kommutativgesetzes folgt nun auch $(-x) \cdot y = -(x \cdot y)$. ∎

DEFINITION 8.3. Sei $\boldsymbol{R} = \langle R, +, -, 0, \cdot, 1 \rangle$ ein kommutativer Ring mit Eins. $\boldsymbol{R}$ ist ein *Körper*, wenn folgendes gilt:

(1) $|R| \geq 2$,
(2) für alle $x \in R$ mit $x \neq 0$ gibt es ein $y \in R$, sodass $x \cdot y = 1$.

ÜBUNGSAUFGABEN 8.4.

(1) Zeigen Sie, dass es in einem Körper für jedes $x$ höchstens ein $y$ mit $x \cdot y = 1$ geben kann.
(2) Zeigen Sie, dass das Produkt zweier Elemente in einem Körper nur dann 0 ist, wenn einer der Faktoren gleich 0 ist.

In einem Körper hat jedes Element $a \neq 0$ genau ein multiplikativ inverses Element; wir bezeichnen es mit $a^{-1}$.

Der zweielementige Ring $\mathbb{Z}_2 = \{\mathbf{0}, \mathbf{1}\}$ mit $\mathbf{0} \oplus \mathbf{0} = \mathbf{0}$, $\mathbf{0} \oplus \mathbf{1} = \mathbf{1}$, $\mathbf{1} \oplus \mathbf{0} = \mathbf{1}$, $\mathbf{1} \oplus \mathbf{1} = \mathbf{0}$, $\mathbf{0} \odot \mathbf{0} = \mathbf{0}$, $\mathbf{0} \odot \mathbf{1} = \mathbf{0}$, $\mathbf{1} \odot \mathbf{0} = \mathbf{0}$, $\mathbf{1} \odot \mathbf{1} = \mathbf{1}$ ist ein Körper.

## 2. Vektorräume

DEFINITION 8.5. Sei $K$ ein Körper. Ein Tupel $\langle V, \boldsymbol{+}, \boldsymbol{-}, \mathbf{0}, * \rangle$ heißt *Vektorraum* über $K$, wenn $V \neq \emptyset$, $\boldsymbol{+} : V \times V \to V$, $\boldsymbol{-} : V \to V$, $\mathbf{0} \in V$ und $* : K \times V \to V$, und für alle $x, y, z \in V$ und $\alpha, \beta \in K$ gilt:

(1) $(x \boldsymbol{+} y) \boldsymbol{+} z = x \boldsymbol{+} (y \boldsymbol{+} z)$,
(2) $\mathbf{0} \boldsymbol{+} x = x$,
(3) $(\boldsymbol{-}x) \boldsymbol{+} x = \mathbf{0}$,
(4) $x \boldsymbol{+} y = y \boldsymbol{+} x$,
(5) $\alpha * (\beta * x) = (\alpha \cdot \beta) * x$,
(6) $(\alpha + \beta) * x = \alpha * x \boldsymbol{+} \beta * x$,
(7) $\alpha * (x \boldsymbol{+} y) = \alpha * x \boldsymbol{+} \alpha * y$,
(8) $1 * x = x$.

BEISPIELE 8.6.

(1) Für jeden Körper $K$ und jedes $n \in \mathbb{N}$ ist $K^n$ (mit geeignet definierten Operationen) ein Vektorraum über $K$.
(2) Für jeden Unterraum $U$ des $\mathbb{R}^n$ ist $U$ (mit geeignet definierten Operationen) ein Vektorraum über $\mathbb{R}$.

ÜBUNGSAUFGABEN 8.7.
Sei in den folgenden Beispielen $K$ ein Körper, und $\langle V, \mathbf{+}, \mathbf{-}, \mathbf{0}, * \rangle$ ein Vektorraum über $K$.

(1) Sei $x \in V$. Zeigen Sie $0 * x = \mathbf{0}$.
(2) Sei $\alpha \in K$. Zeigen Sie $\alpha * \mathbf{0} = \mathbf{0}$.
(3) Sei $x \in V$. Zeigen Sie $(\mathbf{-}x) = (-1) * x$.
(4) Sei $\alpha \in K, x \in V$ so, dass $\alpha * v = \mathbf{0}$. Zeigen Sie, dass $\alpha = 0$ oder $v = \mathbf{0}$.

DEFINITION 8.8. Sei $K$ ein Körper und $V$ ein Vektorraum über $K$. Sei $U$ eine Teilmenge von $V$. $U$ ist ein *Unterraum* von $V$, wenn $U \neq \emptyset$, und für alle $u, v \in U$ und $\alpha \in K$ gilt $u + v \in U$ und $\alpha * u \in U$.

Wenn $U$ ein Unterraum von $\boldsymbol{V} = \langle V, +, -, 0, * \rangle$ ist, dann ist $\boldsymbol{U} = \langle U, +|_{U \times U}, -|_U, 0, *|_{K \times U} \rangle$ ebenfalls ein Vektorraum über $K$.

SATZ 8.9 (Unterraumkriterium). *Sei $K$ ein Körper und $V$ ein Vektorraum über $K$. Sei $U$ eine Teilmenge von $V$. Dann ist $U$ ist genau dann ein* Unterraum *von $V$, wenn $U \neq \emptyset$, und für alle $\alpha, \beta \in K$ und $u, v \in U$ gilt $\alpha * u + \beta * v \in U$.*

KAPITEL 9

# Lineare Abbildungen

## 1. Beispiele

Wir betrachten einige Funktionen, die von einem Vektorraum in einen Vektorraum gehen.

AUFGABE 9.1. Sei $s$ jene Funktion von $\mathbb{R}^2$ nach $\mathbb{R}^2$, die jeden Punkt auf den Punkt abbildet, auf dem er nach der Spiegelung an der $x$-Achse landet. Dann lässt sich $s$ so schreiben:

$$\begin{array}{rcccl} s & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}. \end{array}$$

AUFGABE 9.2. Wir überlegen uns, wo der Punkt $\binom{x}{y}$ nach einer Drehung um den Nullpunkt um $60^\circ$ gegen den Uhrzeigersinn landet. Sei $d$ diese Drehung. Dann lässt sich $d$ so schreiben:

$$\begin{array}{rcccl} d & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} \cos(\frac{\pi}{3}) & -\sin(\frac{\pi}{3}) \\ \sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3}) \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}. \end{array}$$

AUFGABE 9.3. Wir bestimmen die Projektion des Punktes $(x, y, z) \in \mathbb{R}^3$ auf $U = L((0, 1, 2), (3, 0, 4))$. Sei $p_U$ diese Projektionsabbildung. Wir berechnen die Projektionsmatrix $P_U$ in folgender Rechnung:

```
In[115]:= B = Transpose[{{0, 1, 2}, {3, 0, 4}}]
Out[115]= {{0, 3}, {1, 0}, {2, 4}}

In[116]:= PU = B . Inverse[Transpose[B].B] . Transpose[B];

          MatrixForm[PU]
```

Out[116]= $\begin{pmatrix} \frac{45}{61} & -\frac{24}{61} & \frac{12}{61} \\ -\frac{24}{61} & \frac{25}{61} & \frac{18}{61} \\ \frac{12}{61} & \frac{18}{61} & \frac{52}{61} \end{pmatrix}$

Somit haben wir eine Matrix $P_U$ gefunden, sodass $P_U \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ die Projektion von $(x, y, z)$ auf $U$ ist.

Solche Abbildungen, die man durch Matrizen beschreiben kann, werden der Inhalt dieses Kapitels sein.

## 2. Die Definition linearer Abbildungen

DEFINITION 9.4. Seien $U$ und $V$ Vektorräume über dem Körper $K$. Eine Funktion $h : U \to V$ ist eine *lineare Abbildung*, wenn:

(1) für alle $u_1, u_2 \in U : h(u_1 + u_2) = h(u_1) + h(u_2)$
(2) für alle $u \in U$ und für alle $\lambda \in K : h(\lambda\, u) = \lambda\, h(u)$.

BEISPIEL 9.5. Sei $K$ ein Körper, und sei $A \in K^{n \times m}$. Dann ist die Abbildung $h$, die durch

$$\begin{array}{rcrcl} h & : & K^m & \longrightarrow & K^n \\ & & x & \longmapsto & A \cdot x \end{array}$$

definiert ist, eine lineare Abbildung vom $K$-Vektorraum $K^m$ in den $K$-Vektorraum $K^n$.

AUFGABE 9.6. Welche der folgenden Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}$ sind linear?

(1) $h_1((x, y)) = 3\,x - 2\,y$ für $x, y \in \mathbb{R}$.
(2) $h_2((x, y)) = 3\,x + 1$ für $x, y \in \mathbb{R}$.
(3) $h_3((x, y)) = 0$ für $x, y \in \mathbb{R}$.

*Lösung.*

(1) Wir zeigen, dass die Abbildung $h_1$ linear ist. Seien dazu $u, v \in \mathbb{R}^2$. Es gibt dann $x_1, y_1, x_2, y_2 \in \mathbb{R}$, sodass $u = (x_1, y_1)$ und $v = (x_2, y_2)$. Es gilt dann $h_1(u+v) = 3\,(x_1+x_2) - 2\,(y_1+y_2) = (3\,x_1 - 2\,y_1) + (3\,x_2 - 2\,y_2) = h_1(u) + h_1(v)$ und $h_1(\lambda\, u) = 3\,(\lambda\, x_1) - 2\,(\lambda\, y_1) = \lambda\,(3\,x_1 - 2\,y_1) = \lambda\, h_1(u)$ .
(2) $h_2$ ist keine lineare Abbildung, da $h_2((1, 1) + (1, 1)) = h_2((2, 2)) = 7$ und $h_2((1, 1)) + h_2((1, 1)) = 4 + 4 = 8$.
(3) $h_3$ ist linear.

SATZ 9.7. *Sei $K$ ein Körper, seien $U, V, W$ Vektorräume über $K$, und seien $f : U \to V$ und $g : V \to W$ lineare Abbildungen. Dann ist die Hintereinanderausführung $g \circ f$ ebenfalls linear.*

SATZ 9.8. *Sei $K$ ein Körper, sei $m \in \mathbb{N}$, und sei $V$ ein Vektorraum über $K$ mit Basis $B = (b_1, \ldots, b_m)$. Dann ist die Abbildung $c$, die durch*

$$\begin{array}{rcrcl} c & : & V & \longrightarrow & K^m \\ & & v & \longmapsto & (v)_B \end{array}$$

*definiert ist, eine lineare Abbildung.*

*Beweis.* Wir geben hier nur den Beweis dafür, dass $c(u+v) = c(u)+c(v)$ für alle $u, v \in V$. Seien $u, v \in V$. Zu zeigen ist, dass $(u + v)_B = (u)_B + (v)_B$. Seien $\lambda_1, \dots, \lambda_m \in K$ und $\mu_1, \dots, \mu_m \in K$ so, dass

$$\sum_{i=1}^{m} \lambda_i \, b_i = u$$

und

$$\sum_{i=1}^{m} \mu_i \, b_i = v.$$

Dann gilt also $(u)_B = (\lambda_1, \dots, \lambda_m)$ und $(v)_B = (\mu_1, \dots, \mu_m)$. Wir berechnen nun

$$\sum_{i=1}^{m} (\lambda_i + \mu_i) \, b_i.$$

Es gilt $\sum_{i=1}^{m}(\lambda_i + \mu_i)\, b_i = \sum_{i=1}^{m} \lambda_i \, b_i + \sum_{i=1}^{m} \mu_i \, b_i = u + v$. Da also $\sum_{i=1}^{m}(\lambda_i + \mu_i)\, b_i = u + v$, gilt $(u + v)_B = (\lambda_1 + \mu_1, \dots, \lambda_m + \mu_m) = (\lambda_1, \dots, \lambda_m) + (\mu_1, \dots, \mu_m) = (u)_B + (v)_B$. ■

Auch die inverse Abbildung dieser Abbildung ist linear:

SATZ 9.9. *Sei $K$ ein Körper, sei $m \in \mathbb{N}$, und sei $V$ ein Vektorraum über $K$ mit Basis $B = (b_1, \dots, b_m)$. Dann ist die Abbildung $d$, die durch*

$$\begin{array}{rcrcl} d & : & K^m & \longrightarrow & V \\ & & (\lambda_1, \dots, \lambda_m) & \longmapsto & \sum_{i=1}^{m} \lambda_i \, b_i \end{array}$$

*definiert ist, eine lineare Abbildung.*

SATZ 9.10. *Seien $U$ und $V$ Vektorräume über $K$ mit den Basen $B = (b_1, \dots, b_m)$ und $C = (c_1, \dots, c_n)$. Sei $A$ eine $m \times n$-Matrix. Die Funktion $f : U \to V$ bilde $x$ auf jenes $y$, das durch*

$$(y)_C = A \cdot (x)_B$$

*gegeben ist, ab. Dann ist $h$ eine lineare Abbildung.*

*Beweisskizze.* Es gilt $f = g \circ h \circ i$, wobei $i : U \to K^m, u \mapsto (u)_B$, $h : K^m \to K^n, x \mapsto A \cdot x$, $g : K^n \to V, (\lambda_1, \dots, \lambda_m) \mapsto \sum_{i=1}^{m} \lambda_i \, c_i$. Alle drei Abbildungen $g, h, i$ sind linear, also auch ihre Hintereinanderausführung. ■

AUFGABE 9.11. Sei $U$ der Unterraum des $\mathbb{R}^3$ mit Basis $B = ((1, 1, 1), (1, 0, 0))$, und sei $h : \mathbb{R}^3 \to \mathbb{R}^2, v \mapsto (v)_B$. Dann ist etwa

$$\begin{array}{rcl} h((4, 1, 1)) & = & \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \\ h((-1, -1, -1)) & = & \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \end{array}$$

ÜBUNGSAUFGABEN 9.12.

(1) Eine lineare Abblidung von $h$ von $\mathbb{R}^2$ nach $\mathbb{R}^2$ bildet den Punkt $\binom{1}{1}$ auf $\binom{2}{0}$ und den Punkt $\binom{1}{-1}$ auf $\binom{0}{-2}$ ab. Auf welchen Punkt wird $\binom{1}{2}$ abgebildet?
(2) Zeigen oder widerlegen Sie: Wenn $h$ eine lineare Abbildung von $\mathbb{R}^3$ nach $\mathbb{R}^2$ ist und $h \neq 0$, dann gilt für alle $v_1, v_2 \in \mathbb{R}_3$:

$$(v_1, v_2) \text{ linear unabhängig } \Rightarrow (h(v_1), h(v_2)) \text{ linear unabhängig.}$$

## 3. Abbildungsmatrizen linearer Abbildungen

Sei $h : U \to V$ eine lineare Abbildung, $B = (b_1, \ldots, b_m)$ eine Basis von $U$. Sei $x \in U$ mit $x = \sum_{i=1}^m \lambda_i \cdot b_i$. Dann gilt wegen der Linearität von $h$:

$$h(x) = h(\sum_{i=1}^m \lambda_i \cdot b_i) = \sum_{i=1}^m \lambda_i \cdot h(b_i).$$

Eine lineare Abbildung ist also durch die Bilder der Basisvektoren bereits vollständig bestimmt.

Wir werden jetzt sehen, dass sich jede lineare Abbildung, deren Definitions- und Bildbereich endlichdimensionale Vektorräume sind, durch eine Matrix darstellen lässt.

DEFINITION 9.13. Seien $U, V$ Vektorräume über $K$, seien $m, n \in \mathbb{N}$, sei $B = (b_1, \ldots, b_m)$ eine Basis von $U$ und $C = (c_1, \ldots, c_n)$ eine Basis von $V$. Sei $h$ eine lineare Abbildung von $U$ nach $V$. Wir definieren nun die $n \times m$-Matrix $S_h(B, C)$. Dazu legen wir fest, dass für $i \in \{1, \ldots, m\}$ in der $i$-ten Spalte von $S_h(B, C)$ der Vektor $(h(b_i))_C$ steht. Es gilt also

$$S_h(B, C) = \begin{pmatrix} | & | & & | \\ (h(b_1))_C & (h(b_2))_C & \cdots & (h(b_m))_C \\ | & | & & | \end{pmatrix}.$$

Die Matrix $S_h(B, C) \in K^{n \times m}$ heißt *Abbildungsmatrix* oder *Darstellungsmatrix* von $h$ bezüglich der Basen $B$ und $C$.

SATZ 9.14. *Seien $U, V$ Vektorräume über $K$, seien $m, n \in \mathbb{N}$, sei $B = (b_1, \ldots, b_m)$ eine Basis von $U$ und $C = (c_1, \ldots, c_n)$ eine Basis von $V$. Sei $h$ eine lineare Abbildung von $U$ nach $V$. Dann gilt*

$$(h(u))_C = S_h(B, C) \cdot (u)_B \text{ für alle } u \in U. \tag{9.1}$$

*Beweis.* Sei $u \in U$ mit $u = \sum_{i=1}^{m} \lambda_i b_i$. Dann gilt

$$\begin{aligned}
(h(u))_C &= \Big(h(\sum_{i=1}^{m} \lambda_i b_i)\Big)_C \\
&= \Big(\sum_{i=1}^{m} \lambda_i h(b_i)\Big)_C \\
&= \sum_{i=1}^{m} \lambda_i \big(h(b_i)\big)_C \\
&= \begin{pmatrix} | & | & & | \\ (h(b_1))_C & (h(b_2))_C & \cdots & (h(b_m))_C \\ | & | & & | \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} \\
&= S_h(B,C) \cdot (u)_B.
\end{aligned}$$

■

Wir zeigen nun, dass die Abbildungsmatrix durch die Gleichung (9.1) eindeutig bestimmt ist.

SATZ 9.15. *Seien $U, V$ Vektorräume über $K$, seien $m, n \in \mathbb{N}$, sei $B = (b_1, \ldots, b_m)$ eine Basis von $U$ und $C = (c_1, \ldots, c_n)$ eine Basis von $V$. Sei $h$ eine lineare Abbildung von $U$ nach $V$, und sei $A$ eine $n \times m$-Matrix, sodass für alle $u \in U$:*

(9.2) $$(h(u))_C = A \cdot (u)_B.$$

*Dann gilt $A = S_h(B,C)$.*

*Beweis.* Sei $i \in \{1, \ldots, m\}$. Wegen (9.2) gilt $(h(b_i))_C = A \cdot (b_i)_B$, also

$$(h(b_i))_C = A \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

wobei der Einser an der $i$ten Stelle steht. Die rechte Seite dieser Gleichung ergibt die $i$-te Spalte von $A$. Also ist die $i$-te Spalte von $A$ gleich $(h(b_i))_C$. Daher $A = S_h(B,C)$. ■

AUFGABE 9.16. Sei $h : \mathbb{R}^2 \to \mathbb{R}^3$ die folgende lineare Abbildung.

$$\begin{array}{rccc} h & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & & \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} 3x-2y \\ 2x+y \\ -x+4y \end{pmatrix}. \end{array}$$

Sei $B$ die Basis $(\begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix})$ des $\mathbb{R}^2$, und sei $C$ die Basis $(\begin{pmatrix} -9 \\ -13 \\ -7 \end{pmatrix}, \begin{pmatrix} 20 \\ 39 \\ 30 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix})$ des $\mathbb{R}^3$. Bestimmen Sie $S_h(B,C)$.

Lösung: Es gilt

$$h(b_1) = \begin{pmatrix} -7 \\ 0 \\ 9 \end{pmatrix}, \; h(b_2) = \begin{pmatrix} 2 \\ 13 \\ 16 \end{pmatrix}.$$

Aus dem Gleichungssystem

$$\begin{pmatrix} -9 & 20 & 0 \\ -13 & 39 & 1 \\ -7 & 30 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} -7 \\ 0 \\ 9 \end{pmatrix}$$

erhalten wir $(h(b_1))_C = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ und ebenso $(h(b_2))_C = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$. Insgesamt erhalten wir

$$S_h(B, C) = \begin{pmatrix} 3 & 2 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

In Mathematica kann man die notwendigen Rechnungen so durchführen:

```
In[117]:= h[x_, y_] = {{3, -2}, {2, 1}, {-1, 4}} . {x, y}
Out[117]= {3 x - 2 y, 2 x + y, -x + 4 y}

In[118]:= hb1 = h[-1, 2]
Out[118]= {-7, 0, 9}

In[119]:= hb2 = h[4, 5]
Out[119]= {2, 13, 16}

In[120]:= CC = {{-9, 20, 0}, {-13, 39, 1}, {-7, 30, 0}}
Out[120]= {{-9, 20, 0}, {-13, 39, 1}, {-7, 30, 0}}

In[121]:= MatrixForm[CC]
```

$$\text{Out[121]=} \begin{pmatrix} -9 & 20 & 0 \\ -13 & 39 & 1 \\ -7 & 30 & 0 \end{pmatrix}$$

```
In[122]:= LinearSolve[CC, hb1]
Out[122]= {3, 1, 0}

In[123]:= LinearSolve[CC, hb2]
Out[123]= {2, 1, 0}
```

AUFGABE 9.17. Sei $U = \mathbb{R}^2, V = \mathbb{R}^1$ mit den kanonischen Basen $B = ((1, 0), (0, 1))$ und $C = ((1))$. Die lineare Abbildung $h : U \to V$ sei definiert durch

$$h((x, y)) = (3\,x - 2\,y)$$

für alle $x, y \in \mathbb{R}$. In diesem Fall ist $(h(b_1))_C = (3)$ und $(h(b_2))_C = (-2)$. Somit gilt

$$S_h(B, C) = (3 \quad -2).$$

AUFGABE 9.18. Sei $U = \mathbb{R}^2, V = \mathbb{R}^1$ mit den Basen $B = ((2, 3), (3, -2)), C = ((2))$. Die lineare Abbildung $h : U \to V$ sei definiert durch

$$h((x, y)) = (3\,x - 2\,y)$$

für alle $x, y \in \mathbb{R}$. Dann ist $h(b_1) = 0$, also $(h(b_1))_B = (0)$ und $h(b_2) = (13)$, also $(h(b_2))_B = 6.5$, und folglich $S_h(B, C) = (0 \;\; 6.5)$.

ÜBUNGSAUFGABEN 9.19.

(1) Eine lineare Abbildung $h : \mathbb{R}^2 \to \mathbb{R}^2$ sei gegeben durch $h(\binom{1}{-1}) = \binom{0}{2}$ und $h(\binom{-2}{1}) = \binom{1}{7}$. Bestimmen Sie die Abbildungsmatrix $S_h(E, E)$, wobei $E = (\binom{1}{0}, \binom{0}{1})$.
(2) Seien $E_2, E_3$ die kanonischen Basen von $\mathbb{R}^2$ und $\mathbb{R}^3$, und sei
$$\sigma((x, y, z)) := (3x - 2y,\ 2z).$$
Geben Sie die Abbildungsmatrix $S_\sigma(E_3, E_2)$ an.

## 4. Abbildungsmatrizen für Spiegelungen und Drehungen

Wir wollen nun die Abbildungsmatrizen für bestimmte Drehungen und Spiegelungen bestimmen. Wir betrachten folgende Beispiele:

AUFGABE 9.20. Sei $e$ die Ebene $L(\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix})$, und sei $\sigma$ jene Abbildung, die jeden Punkt im $\mathbb{R}^3$ auf den Punkt abbildet, auf dem er nach der Spiegelung an der Ebene $e$ landet.

Gesucht ist die Abbildungsmatrix $S_\sigma(E, E)$ dieser Spiegelung bezüglich der kanonischen Basis $E$ des $\mathbb{R}^3$.

AUFGABE 9.21. Sei $g$ die Gerade mit der Gleichung $5x - 2y = 0$ im $\mathbb{R}^2$. Sei $\sigma$ jene Abbildung, die jeden Punkt im $\mathbb{R}^2$ auf den Punkt abbildet, auf dem er nach der Spiegelung an $g$ landet.

Gesucht ist die Abbildungsmatrix $S_\sigma(E, E)$ dieser Spiegelung bezüglich der kanonischen Basis $E$ des $\mathbb{R}^2$.

AUFGABE 9.22. Sei $g$ die Gerade $L(\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix})$, und sei $\delta : \mathbb{R}^3 \to \mathbb{R}^3$ jene Abbildung, die jeden Punkt auf den Punkt abbildet, auf dem er nach der Drehung um $60^\circ$ um die Gerade $g$ landet. Wir müssen noch die Richtung der Drehung festlegen: Wenn wir vom Punkt $\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$ auf die Ebene $x + 2y - 2z = 0$ hinunterschauen, dann sollen sich die Punkte dieser Ebene gegen den Uhrzeigersinn drehen.

Alle drei Beispiele kann man mit dem gleichen “Programm” lösen. Sei $h$ die angegebene lineare Abbildung. Dann gehen wir so vor:

(1) Bestimme eine Basis $B$, sodass $S_h(B, B)$ leicht zu bestimmen ist.
(2) Bestimme $S_h(B, B)$.
(3) Bestimme $S_h(E, E)$ aus $S_h(B, B)$.

Die Schritte (1) und (2) können wir bereits jetzt durchführen; für den Schritt (3) müssen wir uns noch überlegen, wie man aus den Koordinaten eines Vektors bezüglich einer Basis die Koordinaten bezüglich einer anderen Basis ausrechnet.

Wir starten mit Beispiel 9.20. Wir wählen eine Basis $(b_1, b_2, b_3)$ des $\mathbb{R}^3$, sodass $b_1, b_2$ in der Ebene $e$ liegen, und $b_3$ auf $b_1$ und $b_2$ normal steht. Es gilt dann $\sigma(b_1) = b_1$,

$\sigma(b_2) = b_2$ und $\sigma(b_3) = -b_3$. Also gilt $(\sigma(b_1))_B = \left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$, $(\sigma(b_2))_B = \left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right)$ und $(\sigma(b_3))_B = (-b_3)_B = \left(\begin{smallmatrix}0\\0\\-1\end{smallmatrix}\right)$. Dann gilt

$$S_\sigma(B,B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Wir bestimmen nun eine Basis $B$ mit den gewünschten Eigenschaften. Dazu wählen wir $b_1 = \left(\begin{smallmatrix}1\\2\\-1\end{smallmatrix}\right)$ und $b_2 = \left(\begin{smallmatrix}0\\1\\-1\end{smallmatrix}\right)$. Wir bestimmen einen Vektor, der auf beiden normal steht:

```
In[124]:= NullSpace[{{1, 2, -1}, {0, 1, -1}}]
Out[124]= {{-1, 1, 1}}
```

Also wählen wir $b_3 = \left(\begin{smallmatrix}-1\\1\\1\end{smallmatrix}\right)$, und somit

$$B = (\left(\begin{smallmatrix}1\\2\\-1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\\-1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1\\1\\1\end{smallmatrix}\right).$$

Wir bestimmen nun $B$ und $S_\sigma(B,B)$ für Beispiel 9.21. Wir wählen eine Basis $(b_1, b_2)$ des $\mathbb{R}^2$, sodass $b_1$ auf der Gerade $g$ liegt, und $b_2$ auf $b_1$ normal steht. Es gilt dann $\sigma(b_1) = b_1$ und $\sigma(b_2) = -b_2$, und somit $(\sigma(b_1))_B = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ und $(\sigma(b_2))_B = (-b_2)_B = \left(\begin{smallmatrix}0\\-1\end{smallmatrix}\right)$. Dann gilt

$$S_\sigma(B,B) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Wir bestimmen nun eine Basis $B$ mit den gewünschten Eigenschaften. Dazu wählen wir $b_2 = \left(\begin{smallmatrix}5\\-2\end{smallmatrix}\right)$ und $b_1 = \left(\begin{smallmatrix}2\\5\end{smallmatrix}\right)$. Wir erhalten

$$B = (\left(\begin{smallmatrix}2\\5\end{smallmatrix}\right), \left(\begin{smallmatrix}5\\-2\end{smallmatrix}\right)).$$

Nun bestimmen wir $B$ und $S_\delta(B,B)$ für Beispiel 9.22. Dazu bestimmen wir eine Orthonormalbasis $B$ von $\mathbb{R}^3$, sodass $b_3 = \frac{1}{\left\|\left(\begin{smallmatrix}1\\2\\-2\end{smallmatrix}\right)\right\|}\left(\begin{smallmatrix}1\\2\\-2\end{smallmatrix}\right)$. Außerdem soll $(b_1, b_2, b_3)$ *positiv orientiert* sein. Das heißt: "Wenn man $b_1$ in $b_2$ hineinschraubt, dann soll sich eine Schraube mit Rechtsgewinde in Richtung $b_3$ bewegen." Eine ONB $(b_1, b_2, b_3)$ ist dann positiv orientiert, wenn $b_1 \times b_2 = b_3$. Zu dieser Basis $B$ bestimmen wir jetzt die Matrix $S_\delta(B,B)$. Es gilt $h(b_3) = b_3$, $h(b_1) = \cos(60^\circ)\, b_1 + \sin(60^\circ)\, b_2$, und $h(b_2) = -sin(60^\circ)\, b_1 + \cos(60^\circ)\, b_2$. Mit $c := \cos(60^\circ)$ und $s := \sin(60^\circ)$ lässt sich $S_\delta(B,B)$ also durch

$$S_\delta(B,B) = \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

angeben. Jetzt bestimmen eine Basis $B$ des $\mathbb{R}^3$ mit den gewünschten Eigenschaften. Dazu bestimmen wir zunächst eine ONB der Ebene $e = (L(\left(\begin{smallmatrix}1\\2\\-2\end{smallmatrix}\right)))^\perp$. Wir bestimmen eine Basis für diese Ebene:

```
In[125]:= NullSpace[{{1, 2, -2}}]
Out[125]= {{2, 0, 1}, {-2, 1, 0}}
```

Daher ist $((2,0,1),(-2,1,0))$ eine Basis von $e$. Wir bestimmen nun eine ONB von $e$.

```
In[126]:= e1 = {2, 0, 1};

          e2 = {-2, 1, 0};

          d1 = e1 / Sqrt[e1.e1]
```

Out[126]= $\left\{\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right\}$

```
In[127]:= c2 = e2 - (e2.d1) * d1
```

Out[127]= $\left\{-\frac{2}{5}, 1, \frac{4}{5}\right\}$

```
In[128]:= d2 = c2 / Sqrt[c2.c2]
```

Out[128]= $\left\{-\frac{2}{3\sqrt{5}}, \frac{\sqrt{5}}{3}, \frac{4}{3\sqrt{5}}\right\}$

Daher ist die Basis $F$, die durch

$$F = \left(\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{45}}\begin{pmatrix} -2 \\ 5 \\ 4 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}\right)$$

gegeben ist, eine ONB des $\mathbb{R}^3$. Wir bestimmen, ob $F$ positive Orientierung hat:

```
In[129]:= Cross[d1, d2]
```

Out[129]= $\left\{-\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right\}$

Die Orientierung ist also noch falsch. Wenn wir das Vorzeichen von $f_1$ umdrehen, erhalten wir eine Basis mit positiver Orientierung. (Man kann genausogut das Vorzeichen von $f_2$ umdrehen. Man darf aber nicht das Vorzeichen von $f_3$ umdrehen; die Basis, die man durch Umdrehen des Vorzeichens von $f_3$ erhält, ist zwar positiv orientiert, aber ihr dritter Basisvektor ist nicht mehr der gewünschte.) Wir schreiben die Basisvektoren von $B$ in die Spalten der Matrix $\overline{B}$.

```
In[130]:= d3 = Cross[-d1, d2]
```

Out[130]= $\left\{\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right\}$

```
In[131]:= B = Transpose[{-d1, d2, d3}]
```

Out[131]= $\left\{\left\{-\frac{2}{\sqrt{5}}, -\frac{2}{3\sqrt{5}}, \frac{1}{3}\right\}, \left\{0, \frac{\sqrt{5}}{3}, \frac{2}{3}\right\}, \left\{-\frac{1}{\sqrt{5}}, \frac{4}{3\sqrt{5}}, -\frac{2}{3}\right\}\right\}$

```
In[132]:= MatrixForm[B]
```

Out[132]= $\begin{pmatrix} -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} & \frac{1}{3} \\ 0 & \frac{\sqrt{5}}{3} & \frac{2}{3} \\ -\frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & -\frac{2}{3} \end{pmatrix}$

Somit haben wir auch für Beispiel 9.22 eine Basis $B$ und die Matrix $S_\delta(B, B)$ gefunden.

## 5. Basistransformationen

Wir lösen folgendes Problem: Gegeben sind zwei Basen $B, C$ des gleichen Unterraums $U$ des $\mathbb{R}^n$, und die Koordinaten $(v)_B$ eines Vektors $v \in U$. Gesucht sind die Koordinaten von $v$ bezüglich $C$, also $(v)_C$.

AUFGABE 9.23. Sei $B = ((1,0,1),(1,-1,0))$ und $C = ((3,-2,1),(1,1,2))$, und sei $(x)_B = (1,-1)$. Gesucht ist $(x)_C$.

*Lösung:* Aus $B$ und $(x)_B$ können wir

$$x = B \cdot (x)_B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

berechnen. Wegen $C \cdot (x)_C = x$ gilt:

$$\begin{pmatrix} 3 & 1 \\ -2 & 1 \\ 1 & 2 \end{pmatrix} \cdot (x)_C = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Als Lösung dieses Gleichungssystems erhalten wir $(x)_C = (-0.2, 0.6)$.

Allgemein gehen wir also so vor: Wenn $\overline{B}$ und $\overline{C}$ die Matrizen sind, in deren Spalten die Vektoren von $B$ beziehungsweise $C$ stehen. Dann lässt sich $(x)_C$ aus der Gleichung

$$\overline{C} \cdot (x)_C = \overline{B} \cdot (x)_B$$

berechnen. Wenn $\overline{C}$ gleich viele Spalten wie Zeilen hat, dann ist $\overline{C}$ invertierbar, und wir erhalten

$$(x)_C = \overline{C}^{-1} \cdot \overline{B} \cdot (x)_B.$$

Wenn $\overline{C}$ weniger Spalten als Zeilen hat, dann wissen wir, dass die Spaltenvektoren von $\overline{C}$ linear unabhängig sind. Da $\overline{C}$ eine Matrix über den reellen Zahlen ist, ist $\overline{C}^T \cdot \overline{C}$ invertierbar. Dann gilt: $\overline{C}^T \cdot \overline{C} \cdot (x)_C = \overline{C}^T \cdot \overline{B} \cdot (x)_B$, und somit

$$(x)_C = (\overline{C}^T \cdot \overline{C})^{-1} \cdot \overline{C}^T \cdot \overline{B} \cdot (x)_B.$$

DEFINITION 9.24. Sei $K$ ein Körper, sei $U$ ein Unterraum von $K^n$, und seien $B$ und $C$ Basen von $U$. Eine Matrix $T$ ist eine *Basistransformationsmatrix* von $B$ nach $C$, wenn für alle $u \in U$ gilt

$$(u)_C = T \cdot (u)_B.$$

Es gibt genau eine solche Matrix. Wir kürzen sie mit ${}_C T_B$ ab. Es gilt also

$$(u)_C = {}_C T_B \cdot (u)_B \text{ für alle } u \in U.$$

Die Basistransformationsmatrix ist genau die Abbildungsmatrix $S_{\mathrm{id}}(B, C)$, wobei id : $U \to U$, $u \mapsto u$ die identische Abbildung ist.

Seien $\overline{B}$ und $\overline{C}$ die Matrizen, in deren Spalten die Vektoren der Basen $B$ bzw. $C$ stehen. Dann kann die Basistransformationsmatrizen so berechnen:

(1) Falls $\overline{B}$ und $\overline{C}$ quadratische Matrizen sind, so gilt ${}_CT_B := \overline{C}^{-1} \cdot B$.
(2) Falls $\overline{C}^T \cdot \overline{C}$ invertierbar ist (das ist immer der Fall, wenn $K$ der Körper der reellen Zahlen ist), so gilt ${}_CT_B := (\overline{C}^T \cdot \overline{C})^{-1} \cdot \overline{C}^T \cdot \overline{B}$.
(3) In jedem Fall kann man ${}_CT_B$ als Abbildungsmatrix $S_{\mathrm{id}}(B, C)$ berechnen. In der $i$-ten Spalte von ${}_CT_B$ steht also $(b_i)_C$, also die Lösung des Gleichungssystems $\overline{C} \cdot x = b_i$.

ÜBUNGSAUFGABEN 9.25.

(1) (Koordinaten) Die Ebene $\varepsilon$ hat die Basen
$$A = (\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix})$$
und
$$B = (\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}).$$
(a) Der Vektor $v$ hat bezüglich $B$ die Koordinaten $(v)_B = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$. Berechnen Sie seine Koordinaten bezüglich $A$!
(b) Bestimmen Sie eine Matrix $T$, sodass für alle $v \in \varepsilon$ gilt:
$$(v)_A = T \cdot (v)_B.$$
(Diese Matrix heißt *Basistransformationsmatrix*).

## 6. Die Hintereinanderausführung und die Matrizenmultiplikation

SATZ 9.26. *Seien $U, V, W$ Vektorräume über dem Körper $K$ mit den Basen $A, B, C$, und seien $f : U \to V$ und $g : V \to W$ lineare Abbildungen. Dann gilt*
$$S_{g\circ f}(A, C) = S_g(B, C) \cdot S_f(A, B). \tag{9.3}$$

*Beweis.* Wir zeigen, dass für alle $u \in U$ gilt:
$$(S_g(B, C) \cdot S_f(A, B)) \cdot (u)_A = ((g \circ f)\ (u))_C.$$
Dann folgt aus Satz 9.15 die Gleichung 9.3. Sei $u \in U$. Dann gilt $(S_g(B, C) \cdot S_f(A, B)) \cdot (u)_A = S_g(B, C) \cdot (S_f(A, B) \cdot (u)_A) = S_g(B, C) \cdot (f(u))_B = \big(g(f(u))\big)_C = \big(g \circ f\ (u)\big)_C$. ■

AUFGABE 9.27. Man spiegle den Punkt $(2, 3)$ an der $x$-Achse, und drehe anschließend den gespiegelten Punkt um $90^\circ$ gegen den Uhrzeigersinn um den Nullpunkt:

*Lösung:* Die Spiegelung $\sigma$ an der $x$-Achse hat bzgl. der kanonischen Basis $E$ des $\mathbb{R}^2$ die Darstellungsmatrix $S_\sigma(E, E) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Die Drehung $\delta$ um $90^\circ$ hat die Darstellungsmatrix $S_\delta(E, E) = \begin{pmatrix} \cos(90^\circ) & -\sin(90^\circ) \\ \sin(90^\circ) & \cos(90^\circ) \end{pmatrix}$. Die Darstellungsmatrix der Spiegelung mit anschließender Drehung ist also

$$S_{\delta\circ\sigma}(E,E) = S_\delta(E,E)\cdot S_\sigma(E,E) = \begin{pmatrix} \cos(90^\circ) & \sin(90^\circ) \\ \sin(90^\circ) & -\cos(90^\circ) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Wir erhalten $\delta(\sigma((2,3))) = (3,2)$.

SATZ 9.28. *Seien $U,V$ Vektorräume über dem Körper $K$, und seien $A,B$ Basen von $U$ und $C,D$ Basen von $V$. Sei $h$ eine lineare Abbildung von $U$ nach $V$. Dann gilt*

$$S_h(A,D) = {}_DT_C \cdot S_h(B,C)\cdot {}_BT_A.$$

*Beweisskizze.* Für alle $u\in U$ gilt ${}_DT_C\cdot S_h(B,C)\cdot {}_BT_A\cdot (u)_A = (h(u))_D$. Also gilt nach Satz 9.15 ${}_DT_C\cdot S_h(B,C)\cdot {}_BT_A = S_h(A,D)$. ■

Das folgende Korollar wird uns helfen, die Beispiele 9.20, 9.21 und 9.22 zu lösen.

KOROLLAR 9.29. *Sei $K$ ein Körper, sei $B$ eine Basis des $K$-Vektorraums $K^n$, und sei $E$ die kanonische Basis von $K^n$. Sei $h$ eine lineare Abbildung von $K^n$ nach $K^n$. Dann gilt*

$$S_h(E,E) = {}_ET_B\cdot S_h(B,B)\cdot {}_BT_E.$$

*Wenn $\overline{B}$ die Matrix ist, in deren Spalten die Vektoren von $B$ stehen, so gilt also*

$$S_h(E,E) = \overline{B}\cdot S_h(B,B)\cdot \overline{B}^{-1}.$$

## 7. Abbildungsmatrizen für Spiegelungen und Drehungen bezüglich der kanonischen Basis

Wir lösen das Beispiel 9.20. Dazu müssen wir die Abbildungsmatrix $S_h(E,E)$ aus der Abbildungsmatrix $S_h(B,B)$ ausrechnen.

```
In[133]:= B = Transpose[{{1, 2, -1}, {0, 1, -1}, {-1, 1, 1}}]
Out[133]= {{1, 0, -1}, {2, 1, 1}, {-1, -1, 1}}

In[134]:= ShBB = {{1, 0, 0}, {0, 1, 0}, {0, 0, -1}}
Out[134]= {{1, 0, 0}, {0, 1, 0}, {0, 0, -1}}

In[135]:= ETB = B
Out[135]= {{1, 0, -1}, {2, 1, 1}, {-1, -1, 1}}

In[136]:= BTE = Inverse[B]
```

Out[136]= $\left\{\left\{\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right\}, \{-1, 0, -1\}, \left\{-\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}\right\}$

```
In[137]:= ShEE = ETB . ShBB . BTE
```

Out[137]= $\left\{\left\{\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right\}, \left\{\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right\}, \left\{\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right\}\right\}$

```
In[138]:= MatrixForm[ShEE]
```

*Out[138]=* $\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$

Nun lösen wir das Beispiel 9.21.

```
In[139]:= B = Transpose[{{2, 5}, {5, -2}}]
Out[139]= {{2, 5}, {5, -2}}

In[140]:= ShBB = {{1, 0}, {0, -1}}
Out[140]= {{1, 0}, {0, -1}}

In[141]:= ETB = B
Out[141]= {{2, 5}, {5, -2}}

In[142]:= BTE = Inverse[B]
```

*Out[142]=* $\left\{\left\{\frac{2}{29}, \frac{5}{29}\right\}, \left\{\frac{5}{29}, -\frac{2}{29}\right\}\right\}$

```
In[143]:= ShEE = ETB . ShBB . BTE
```

*Out[143]=* $\left\{\left\{-\frac{21}{29}, \frac{20}{29}\right\}, \left\{\frac{20}{29}, \frac{21}{29}\right\}\right\}$

```
In[144]:= MatrixForm[ShEE]
```

*Out[144]=* $\begin{pmatrix} -\frac{21}{29} & \frac{20}{29} \\ \frac{20}{29} & \frac{21}{29} \end{pmatrix}$

Schließlich lösen wir noch Beispiel 9.22

```
In[145]:= B = Transpose[{-1/Sqrt[5] * {2, 0, 1},
              1/Sqrt[45] * {-2, 5, 4}, 1/3 * {1, 2, -2}}]
```

*Out[145]=* $\left\{\left\{-\frac{2}{\sqrt{5}}, -\frac{2}{3\sqrt{5}}, \frac{1}{3}\right\}, \left\{0, \frac{\sqrt{5}}{3}, \frac{2}{3}\right\}, \left\{-\frac{1}{\sqrt{5}}, \frac{4}{3\sqrt{5}}, -\frac{2}{3}\right\}\right\}$

```
In[146]:= c = Cos[π/3];

          s = Sin[π/3];

          ShBB = {{c, -s, 0}, {s, c, 0}, {0, 0, 1}}
```

*Out[146]=* $\left\{\left\{\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0\right\}, \left\{\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right\}, \{0, 0, 1\}\right\}$

```
In[147]:= ETB = B
```

*Out[147]=* $\left\{\left\{-\frac{2}{\sqrt{5}}, -\frac{2}{3\sqrt{5}}, \frac{1}{3}\right\}, \left\{0, \frac{\sqrt{5}}{3}, \frac{2}{3}\right\}, \left\{-\frac{1}{\sqrt{5}}, \frac{4}{3\sqrt{5}}, -\frac{2}{3}\right\}\right\}$

```
In[148]:= BTE = Inverse[B]
```

*Out[148]=* $\left\{\left\{-\frac{2}{\sqrt{5}}, 0, -\frac{1}{\sqrt{5}}\right\}, \left\{-\frac{2}{3\sqrt{5}}, \frac{\sqrt{5}}{3}, \frac{4}{3\sqrt{5}}\right\}, \left\{\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right\}\right\}$

```
In[149]:= ShEE = Simplify[ETB.ShBB.BTE]
```

$$\text{Out[149]=}\ \left\{\left\{\frac{5}{9}, \frac{1}{9}+\frac{1}{\sqrt{3}}, -\frac{1}{9}+\frac{1}{\sqrt{3}}\right\}, \left\{\frac{1}{9}-\frac{1}{\sqrt{3}}, \frac{13}{18}, \frac{1}{18}\left(-4-3\sqrt{3}\right)\right\}, \left\{-\frac{1}{9}-\frac{1}{\sqrt{3}}, \frac{1}{18}\left(-4+3\sqrt{3}\right), \frac{13}{18}\right\}\right\}$$

```
In[150]:= MatrixForm[N[ShEE, 6]]
```

$$\text{Out[150]=}\ \begin{pmatrix} 0.555556 & 0.688461 & 0.466239 \\ -0.466239 & 0.722222 & -0.510897 \\ -0.688461 & 0.0664529 & 0.722222 \end{pmatrix}$$

## ÜBUNGSAUFGABEN 9.30.

(1) Finden Sie eine Basis $B$, bezüglich der die Spiegelung $\delta$ an der Ebene $x + y + z = 0$ die Abbildungsmatrix

$$S_\delta(B, B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

hat.

(2) Bestimmen Sie eine geeignete Basis $B$ des $\mathbb{R}^2$, für die die Abbildungsmatrix der Spiegelung $s$ an der Geraden $g : 7x - 4y$ folgende Form besitzt:

$$S_s(B, B) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(3) (Spiegelung an einer Geraden in der Ebene) Wir bezeichnen mit $\sigma$ jene Spiegelung, die jeden Punkt der Ebene $\mathbb{R}^2$ an der Geraden $-3x + 4y = 0$ spiegelt.

  (a) Bestimmen Sie eine Basis $B$ des $\mathbb{R}^2$, sodass für die Abbildungsmatrix $S_\sigma(B, B)$ der Spiegelung $\sigma$ bezüglich der Basis $B$ folgende Gleichung gilt.

$$S_\sigma(B, B) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

  (b) Berechnen Sie die Abbildungsmatrix $S_\sigma(E, E)$ der Spiegelung $\sigma$ bezüglich der kanonischen Basis $E$.

  (c) Wo landet der Punkt $(a, b)$ nach dieser Spiegelung $\sigma$?

(4) Sei $h$ die lineare Abbildung, die jeden Punkt an der Ebene $e : x + 2y + 2z = 0$ spiegelt.

  (a) Berechnen Sie $h(v)$ für $v \in \{\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}\}$.

  (b) Bestimmen Sie die Abbildungsmatrix $S_h(B, B)$ für die Basis

$$B = (\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}).$$

(5) Wir betrachten die Abbildung $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$, die jeden Punkt an der $y$-Achse spiegelt. Bestimmen Sie die Abbildungsmatrix $S_\sigma(E, E)$ dieser Abbildung bezüglich der kanonischen Basis $E$. Testen Sie Ihre Abbildungsmatrix, indem Sie damit das Spiegelbild von $\begin{pmatrix} 4 \\ -3 \end{pmatrix}$ ausrechnen.

(6) Geben Sie die Darstellungsmatrix für die Spiegelung an der Ebene $2x - y + z = 0$ bezüglich der kanonischen Basis $E$ an. (Sie brauchen die auftretenden inversen Matrizen nicht zu berechnen.)

(7) (Spiegelung an einer Geraden in der Ebene)

  (a) Wir bezeichnen mit $\sigma$ jene Spiegelung, die jeden Punkt an der Geraden

$$x + y = 0$$

  spiegelt. Wo landet der Punkt $(a, b)$ nach dieser Spiegelung $\sigma$? Bestimmen Sie die Abbildungsmatrix dieser Spiegelung bezüglich der kanonischen Basis.

  (b) Finden Sie mithilfe einer Skizze, wo der Punkt $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ landet, und überprüfen Sie das Ergebnis Ihrer Rechnung anhand der Skizze.

(8) Sei $h$ die lineare Abbildung, die jeden Punkt an der Ebene $e : x+2y+2z = 0$ spiegelt. Wir suchen in diesem Beispiel $S_h(E, E)$, wobei $E$ die kanonische Basis des $\mathbb{R}^3$ ist. Gehen Sie dazu so vor:

  (a) Bestimmen Sie $S_{\mathrm{id}}(E, B)$. Dabei ist id die identische Abbildung. $S_{\mathrm{id}}(E, B)$ ist also eine Basistransformationsmatrix, und erfüllt die Eigenschaft $(v)_B = S_{\mathrm{id}}(E, B) \cdot (v)_E$ für alle $v \in \mathbb{R}^3$.)

  (b) Bestimmen Sie $S_{\mathrm{id}}(B, E)$.

  (c) Bauen Sie aus diesen beiden Matrizen und $S_h(B, B)$ die Matrix $S_h(E, E)$ zusammen.

(9) (Spiegelung an einer Ebene im Raum) Wo landet der Punkt $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ nach der Spiegelung $\sigma$ an der Ebene

$$-2x - y + z = 0 \ ?$$

Bestimmen Sie die Abbildungsmatrix dieser Spiegelung!

(10) (Spiegelung an einer Geraden in der Ebene) Wir bezeichnen mit $\sigma$ jene Spiegelung, die jeden Punkt an der Geraden $12x + 5y = 0$ spiegelt. Wo landet der Punkt $(a, b)$ nach dieser Spiegelung $\sigma$? Bestimmen Sie die Abbildungsmatrix dieser Spiegelung bezüglich der kanonischen Basis.

(11) Wir bezeichnen mit $\sigma$ jene Spiegelung, die jeden Punkt an der Geraden $15x - 8y = 0$ spiegelt. Wo landet der Punkt $(a, b)$ nach dieser Spiegelung $\sigma$? Bestimmen Sie die Abbildungsmatrix dieser Spiegelung bezüglich der kanonischen Basis.

(12) Geben Sie die Darstellungsmatrix für die Spiegelung an der Geraden $x - 2y = 0$ bezüglich der kanonischen Basis $E$ an.

(13) Wo landet der Punkt $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ nach der Drehung $\delta$ um $90^\circ$ um die Gerade

$$X = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 4 \\ 0 \\ -3 \end{pmatrix}.$$

Bestimmen Sie die Abbildungsmatrix dieser Drehung bezüglich der kanonischen Basis! Wir drehen dabei *gegen den Uhrzeigersinn*, wenn man von $\begin{pmatrix} 4 \\ 0 \\ -3 \end{pmatrix}$ nach $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ schaut.

(14) (Drehung um eine Gerade im Raum) Wo landet der Punkt $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ nach der Drehung $\delta$ um $90^\circ$ um die Gerade $g$, die durch

$$g : X = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 12 \\ 0 \\ 5 \end{pmatrix}$$

gegeben ist? Dabei drehen wir *gegen den Uhrzeigersinn*, wenn man von $\begin{pmatrix} 12 \\ 0 \\ 5 \end{pmatrix}$ in Richtung $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ schaut. Bestimmen Sie die Abbildungsmatrix dieser Drehung bezüglich der kanonischen Basis!

# Teil 4

# Mengen und Relationen

KAPITEL 10

# Äquivalenzrelationen und Faktormengen

## 1. Äquivalenzrelationen

Wir nennen eine Relation von $A$ nach $A$ auch eine *Relation auf $A$*.

DEFINITION 10.1. Sei $\rho$ eine Relation auf $A$.

(1) $\rho$ ist *reflexiv*, wenn für alle $a \in A$ gilt: $(a, a) \in \rho$.
(2) $\rho$ ist *transitiv*, wenn für alle $a, b, c \in A$ gilt: wenn $(a, b) \in \rho$ und $(b, c) \in \rho$, so gilt auch $(a, c) \in \rho$.
(3) $\rho$ ist *symmetrisch*, wenn für alle $a, b \in A$ gilt: wenn $(a, b) \in \rho$, so gilt auch $(b, a) \in \rho$.

DEFINITION 10.2. Sei $\rho$ eine Relation auf $A$. Die Relation $\rho$ ist eine *Äquivalenzrelation auf $A$*, wenn sie reflexiv, transitiv und symmetrisch ist.

DEFINITION 10.3. Sei $\rho$ eine Äquivalenzrelation auf $A$, und sei $a \in A$. Die *Äquivalenzklasse von $a$ bezüglich $\rho$* wird mit $[a]_\rho$ oder $a/\rho$ abgekürzt, und ist definiert durch

$$a/\rho := \{b \in A \mid (a, b) \in \rho\}.$$

Eine Teilmenge $C$ von $A$ ist eine *Äquivalenzklasse von $\rho$*, wenn es ein $a \in A$ gibt, sodass $C = a/\rho$.

LEMMA 10.4. *Sei $\rho$ eine Äquivalenzrelation auf $A$, und seien $a, b \in A$. Wenn $(a, b) \in \rho$, so gilt $[a]_\rho = [b]_\rho$.*

*Beweis:* Sei $c \in [a]_\rho$. Dann gilt $(a, c) \in \rho$. Wegen der Symmetrie von $\rho$ gilt auch $(b, a) \in \rho$, und somit wegen der Transitivität von $\rho$ auch $(b, c) \in \rho$. Somit gilt $c \in [b]_\rho$. Sei nun $c \in [b]_\rho$. Dann gilt $(b, c) \in \rho$ und somit wegen $(a, b) \in \rho$ und der Transitivität von $\rho$ auch $(a, c) \in \rho$, und somit $c \in [a]_\rho$. ■

ÜBUNGSAUFGABEN 10.5.

(1) Sei $\rho$ eine Äquivalenzrelation auf $A$, und seien $a, b \in A$. Zeigen Sie, dass folgende Aussagen äquivalent sind:
    (a) $(a, b) \in \rho$.
    (b) $[a]_\rho = [b]_\rho$.
    (c) $a \in [b]_\rho$.
    (d) $[a]_\rho \cap [b]_\rho \neq \emptyset$.

## 2. Partitionen

DEFINITION 10.6. Sei $A$ eine Menge. Eine Teilmenge $\mathcal{P}$ von $\mathcal{P}(A)$ ist eine *Partition von A*, wenn

(1) für alle $P \in \mathcal{P} : P \neq \emptyset$,
(2) $\bigcup \{P \mid P \in \mathcal{P}\} = A$,
(3) für alle $P_1, P_2 \in \mathcal{P}$ mit $P_1 \neq P_2$ gilt $P_1 \cap P_2 = \emptyset$.

Wenn $\mathcal{P}$ eine Partition von $A$ ist, so gibt es für jedes $a \in A$ genau ein $P \in \mathcal{P}$, sodass $a \in P$.

DEFINITION 10.7. Sei $\rho$ eine Äquivalenzrelation auf $A$. Die *Faktormenge von A modulo* $\rho$ ist die Menge

$$A/\rho := \{[a]_\rho \mid a \in A\}.$$

SATZ 10.8. *Sei $\rho$ eine Äquivalenzrelation auf A. Dann ist die Faktormenge von A bezüglich $\rho$ eine Partition von A.*

*Beweis:* Sei $P \in A/\rho$. Dann gibt es ein $a \in A$, sodass $P = [a]_\rho = \{b \in A \mid (a,b) \in \rho\}$. Wegen der Reflexivität von $\rho$ gilt $(a,a) \in \rho$, und folglich $a \in [a]_\rho$, also $a \in P$. Somit gilt $P \neq \emptyset$.

Wir zeigen nun, dass jedes $a \in A$ Element eines Elementes von $A/\rho$ ist. Sei dazu $a \in A$. Dann gilt wegen der Reflexivität von $\rho$, dass $a \in [a]_\rho$. Somit ist $a$ Element eines Elementes von $A/\rho$, nämlich von $[a]_\rho$.

Seien nun $P, Q \in A/\rho$. Seien $a, b \in A$ so, dass $P = [a]_\rho$ und $Q = [b]_\rho$. Wir nehmen nun an, dass $P \cap Q \neq \emptyset$. Es gibt dann also ein $c \in A$ mit $c \in P$ und $c \in Q$. Also gilt wegen $c \in [a]_\rho$ auch $(a,c) \in \rho$, und wegen $c \in [b]_\rho$ auch $(b,c) \in \rho$. Wegen der Symmetrie von $\rho$ gilt daher auch $(c,b) \in \rho$, und daher, wegen der Transitivität von $\rho$, auch $(a,b) \in \rho$. Somit gilt nach Lemma 10.4 auch $P = Q$. ■

SATZ 10.9. *Sei A eine Menge, und sei $\mathcal{P}$ eine Partition von A. Dann ist*

$$\rho := \{(a,b) \in A \times A \mid \exists P \in \mathcal{P} : a \in P \text{ und } b \in P\}$$

*eine Äquivalenzrelation auf A.*

SATZ 10.10. *Sei A eine Menge, sei $\mathbf{P}$ die Menge aller Partitionen auf A, und sei $\mathbf{E}$ die Menge aller Äquivalenzrelationen auf A. Dann sind die Abbildungen e und p, die durch*

$$\begin{array}{rcccc} e & : & \mathbf{P} & \longrightarrow & \mathbf{E} \\ & & \mathcal{P} & \longmapsto & \rho(\mathcal{P}) \end{array}$$

*mit* $\rho(\mathcal{P}) = \{(a,b) \in A \times A \mid \exists P \in \mathcal{P} : a \in P \text{ und } b \in P\}$ *und*

$$\begin{array}{rcccc} p & : & \mathbf{E} & \longrightarrow & \mathbf{P} \\ & & \rho & \longmapsto & \{[a]_\rho \mid a \in A\} \end{array}$$

*definiert sind, zueinander inverse Bijektionen.*

*Beweis:* Wir zeigen, dass für jede Äquivalenzrelation $\rho$ auf $A$ gilt:

$$e(p(\rho)) = \rho. \tag{10.1}$$

Sei dazu $\rho \in \mathbf{E}$. Wir zeigen nun als erstes $\subseteq$ von (10.1), und nehmen dazu, dass $a, b \in A$ so sind, dass $(a, b) \in e(p(\rho))$. Dann gibt es ein $P \in p(\rho)$, sodass $a \in P$ und $b \in P$. Es gibt also ein $P \in \{[c]_\rho \mid c \in A\}$, sodass $a \in P$ und $b \in P$. Somit gibt es ein $c \in A$, sodass $a \in [c]_\rho$ und $b \in [c]_\rho$. Somit gilt $(c, a) \in \rho$ und $(c, b) \in \rho$, also $(a, b) \in \rho$. Das beweist $\subseteq$ von (10.1).

Sei nun $(a, b) \in \rho$. Dann gilt $b \in [a]_\rho$, und, wegen der Reflexivität von $\rho$, auch $a \in [a]_\rho$. Nun gilt $[a]_\rho \in p(\rho)$. Da $a, b$ beide Elemente von $[a]_\rho$ sind, gilt $(a, b) \in e(p(\rho))$. Das beweist $\supseteq$ von (10.1).

Nun zeigen wir, dass für jede Partition $\mathcal{P}$ von $A$ gilt:

$$p(e(\mathcal{P})) = \mathcal{P}. \tag{10.2}$$

Wir zeigen als erstes die Inklusion $\supseteq$ von (10.2). Sei dazu $P \in \mathcal{P}$, und sei $\sigma := e(\mathcal{P})$. Da $P$ nicht leer ist, gibt es $a \in P$. Wir zeigen als erstes

$$P = [a]_\sigma. \tag{10.3}$$

Um $\subseteq$ von (10.3) zu zeigen, wählen wir $p \in P$. Dann gilt $a \in P$ und $p \in P$, also $(a, p) \in e(\mathcal{P}) = \sigma$. Somit gilt $p \in [a]_\sigma$. Um $\supseteq$ von (10.3) zu zeigen, wählen wir $q \in [a]_\sigma$. Es gilt dann also $(a, q) \in \sigma$. Folglich gibt es $Q \in \mathcal{P}$, sodass $a \in Q$ und $q \in Q$. Nun gilt $a \in P \cap Q$, also $P = Q$. Folglich gilt $q \in P$. Das beweist (10.3). Da $[a]_\sigma$ offensichtlich in $p(\sigma)$ liegt, gilt auch $P \in p(\sigma)$. Somit ist $\supseteq$ von (10.2) bewiesen.

Um $\subseteq$ von (10.2) zu zeigen, wählen wir $Q \in p(e(\mathcal{P}))$. Sei $\sigma := e(\mathcal{P})$. Dann gibt es ein $a \in A$, sodass $Q = [a]_\sigma$. Sei nun $P \in \mathcal{P}$ so, dass $a \in P$. Wir zeigen nun

$$P = Q. \tag{10.4}$$

Sei dazu $p \in P$. Dann gilt $(a, p) \in \sigma$. Folglich gilt $p \in [a]_\sigma$, also $p \in Q$. Sei nun $q \in Q$. Dann gilt $(a, q) \in \sigma$. Es gibt also ein $R \in \mathcal{P}$, sodass $a \in R$ und $q \in R$. Wegen $a \in R \cap P$ gilt $R = P$. Somit gilt $q \in P$. Das beweist (10.4). Somit gilt also auch $Q \in \mathcal{P}$; also gilt auch $\subseteq$ in (10.2).

Aus (10.1), (10.2) und Satz 5.28 erhalten wir nun, dass $e$ und $p$ zueinander inverse bijektive Abbildungen sind. ■

DEFINITION 10.11. Sei $A$ eine Menge, und sei $\rho$ eine Äquivalenzrelation auf $A$. Eine Teilmenge $R$ von $A$ ist ein *Repräsentantensystem* von $A$ modulo $\rho$, wenn für alle $a \in A$ die Menge $[a]_\rho \cap R$ genau ein Element enthält.

*Beispiel:* Sei $A := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, und sei $(\binom{a}{b}, \binom{c}{d}) \in \rho$ genau dann, wenn $ad = bc$. Dann ist $\rho$ eine Äquivalenzrelation, und $R := \{\binom{a}{b} \in A \mid b > 0, \operatorname{ggT}(a, b) = 1\}$ ist ein Repräsentantensystem. Die Faktormenge $A/\rho$ bezeichnet man als die Menge der *rationalen Zahlen*. Für $[\binom{a}{b}]_\rho$ schreibt man $\frac{a}{b}$.

KAPITEL 11

# Die Mächtigkeit von Mengen

## 1. Ordnungsrelationen

DEFINITION 11.1. Sei $M$ eine Menge, und sei $\rho$ eine Relation auf $M$. Die Relation $\rho$ ist *antisymmetrisch*, wenn für alle $x, y \in M$ mit $(x, y) \in \rho$ und $(y, x) \in \rho$ gilt: $x = y$.

DEFINITION 11.2. Sei $M$ eine Menge, und sei $\rho$ eine Relation auf $M$. Die Relation $\rho$ ist eine *Ordnungsrelation*, wenn sie *reflexiv*, *transitiv* und *antisymmetrisch* ist.

DEFINITION 11.3. Sei $M$ eine Menge, und sei $\leq$ eine Ordnungsrelation auf $M$. Die Relation $\leq$ ist *linear*, wenn für alle $x, y \in M$ gilt, dass $x \leq y$ oder $y \leq x$.

Ein Paar $(M, \leq)$ aus einer Menge und einer Ordnungsrelation bezeichnen wir als *geordnete Menge*. Wir schreiben auch $a < b$, wenn $a \leq b$ und $a \neq b$.

DEFINITION 11.4. Sei $(M, \leq)$ eine geordnete Menge, und sei $a \in M$.

(1) $a$ ist ein *kleinstes Element* von $M$, wenn für alle $b \in M$ gilt: $a \leq b$.
(2) $a$ ist ein *minimales Element* von $M$, wenn es kein $b \in M$ mit $b < a$ gibt.
(3) Sei $T$ eine Teilmenge von $M$, und sei $m \in M$. Das Element $m$ ist eine *untere Schranke für $T$*, wenn für alle $t \in T$ gilt: $m \leq t$. (Eine untere Schranke kann, aber muss nicht, in $T$ liegen.)
(4) $a$ is ein *größtes Element* von $M$, wenn für alle $b \in M$ gilt: $b \leq a$.
(5) $a$ ist ein *maximales Element* von $M$, wenn es kein $b$ in $M$ mit $a < b$ gibt.
(6) Sei $T$ eine Teilmenge von $M$, und sei $m \in M$. Das Element $m$ ist eine *obere Schranke für $T$*, wenn für alle $t \in T$ gilt: $t \leq m$.

Eine geordnete Menge $(M, \leq)$ hat höchstens ein kleinstes Element. Jedes kleinste Element ist minimal.

SATZ 11.5 (Lemma von Zorn). *Sei $(M, \leq)$ eine geordnete Menge mit folgender Eigenschaft:*

> *Für alle Teilmengen $T$ von $M$ mit der Eigenschaft, dass $(T, \leq)$ linear geordnet ist, gibt es ein $m \in M$, sodass für alle $t \in T$: $t \leq m$.*

*Dann hat $M$ ein maximales Element.*

Zur Formulierung: exakterweise muss es statt $(T, \leq)$ natürlich $(T, \leq \cap (T \times T))$ heißen. Die Forderung an $M$ ist, dass jede linear geordnete Teilmenge $T$ von $M$ eine obere

Schranke besitzt, die zwar nicht in $T$, aber sehr wohl in $M$ liegen muss. Der Beweis des Lemmas von Zorn ist aufwändig und benötigt das Auswahlaxiom.

Aus dem Lemma von Zorn werden wir nun folgern, dass jeder Vektorraum eine Basis besitzt. Manchmal ist es nützlich, als Basen für Vektorräume nicht nur *Folgen* von Vektoren, sondern auch einfach *Teilmengen* des Vektorraums zuzulassen; die Definitionen lassen sich rasch anpassen:

DEFINITION 11.6. Sei $K$ ein Körper, und sei $V$ ein Vektorraum über $K$. Eine Teilmenge $B$ von $V$ ist *linear unabhängig*, wenn für jede endliche Teilmenge $W$ von $B$ und für alle $\lambda : W \to K$ mit

$$\sum_{w \in W} \lambda(w)\, w = 0$$

gilt, dass für alle $w \in W$ gilt: $\lambda(w) = 0$. Die *lineare Hülle* der Menge $B$ ist definiert als $\{\sum_{w \in W} \lambda(w)\, w \mid W \text{ ist endlich}, \lambda : W \to K\}$. Die Menge $B$ ist eine *Basis für $V$*, wenn sie linear unbhängig ist, und ihre lineare Hülle ganz $V$.

LEMMA 11.7. *Sei $V$ ein Vektorraum über dem Körper $F$, und sei*

$$\mathcal{U} := \{B \mid B \text{ ist linear unabhängige Teilmenge von } V\}.$$

*Jedes maximale Element aus $(\mathcal{U}, \subseteq)$ ist eine Basis von $V$.*

*Beweis:* Sei $B$ ein maximales Element von $(\mathcal{U}, \subseteq)$. Da $B \in \mathcal{U}$, ist $B$ linear unabhängig. Wir zeigen nun, dass $L(B) = V$. Sei dazu $v \in V$; wir wollen zeigen, dass $v \in L(B)$. Wenn $v \in B$, so gilt klarerweise $v \in L(B)$. Wir betrachten nun den Fall $v \notin B$. Wir bilden $B' := B \cup \{v\}$. Wegen der Maximalität von $B$ ist $B'$ linear abhängig. Es gibt also eine endliche Teilmenge $W$ von $B \cup \{v\}$ und $\lambda : W \to F$, sodass es $w \in W$ mit $\lambda(w) \neq 0$ gibt, und

$$\sum_{w \in W} \lambda(w)\, w = 0. \tag{11.1}$$

Wenn $v \notin W$ oder $\lambda(v) = 0$, so erhalten wir aus (11.1), dass $B$ linear abhängig ist, im Widerspruch zu $B \in \mathcal{U}$. Somit gilt $v \in W$ und $\lambda(v) \neq 0$. Es gilt also

$$v = -\frac{1}{\lambda(v)} \sum_{w \in W \setminus \{v\}} \lambda(w)\, w,$$

und somit $v \in L(B)$.

Somit gilt $V = L(B)$. ■

SATZ 11.8. *Sei $F$ ein Körper, und sei $V$ ein Vektorraum über $F$. Dann besitzt $V$ eine Basis.*

*Beweis:* Sei $\mathcal{U} := \{B \mid B \subseteq V, B \text{ ist linear unabhängig}\}$. Nach Lemma 11.7 genügt es zu zeigen, dass $(\mathcal{U}, \subseteq)$ ein maximales Element hat. Dazu verwenden wir das Zornsche Lemma. Sei $\mathcal{K}$ eine Teilmenge von $\mathcal{U}$, sodass $(\mathcal{K}, \subseteq)$ linear geordnet ist. Es gilt also für alle $K, L \in \mathcal{K}$: $K \subseteq L$ oder $L \subseteq K$. Wir zeigen nun, dass $\mathcal{K}$ eine obere Schranke in $\mathcal{U}$ besitzt. Sei dazu $M := \bigcup\{K \mid K \in \mathcal{K}\}$, also die Vereinigung aller Elemente aus $\mathcal{K}$. Klarerweise ist $M$ eine obere Schranke für $\mathcal{K}$. Es bleibt zu zeigen, dass $M \in \mathcal{U}$. Dazu ist zu zeigen, dass auch $M$ linear unabhängig ist. Sei dazu $W$ eine endliche nichtleere Teilmenge von $M$, und sei $\lambda : W \to \mathbb{R}$ so, dass

$$\sum_{w \in W} \lambda(w)\, w = 0.$$

Da jedes $w \in W$ in $M$ liegt, gibt es für jedes $w$ ein $K_w \in \mathcal{K}$, sodass $w \in K_w$.

Nun gilt für jede endliche nichtleere Teilmenge $\mathcal{E}$ von $\mathcal{K}$, dass $\bigcup\{E \mid E \in \mathcal{E}\} \in \mathcal{E}$. (Das kann man mit Induktion nach der Anzahl der Elemente von $\mathcal{E}$ zeigen. Die Beobachtung ist aber tatsächlich einfach: in einer endlichen linear geordneten Menge gibt es immer ein größtes Element.)

Wenn wir diese Beobachtung für $\mathcal{E} := \{K_w \mid w \in W\}$ verwenden, so erhalten wir, dass es ein $w_0 \in W$ gibt, sodass $W \subseteq K_{w_0}$. Da $K_{w_0} \in \mathcal{U}$, ist $K_{w_0}$ linear unabhängig. Daher gilt für alle $w \in W$: $\lambda(w) = 0$.

Somit ist $M$ linear unabhängig, und es gilt $M \in \mathcal{U}$.

Das Lemma von Zorn liefert nun, dass $(\mathcal{U}, \subseteq)$ zumindest ein maximales Element besitzt. ■

## 2. Mächtigkeit

DEFINITION 11.9. Seien $A, B$ Mengen. Wir sagen, dass $A$ und $B$ *gleichmächtig* sind ($A \sim B$), wenn es eine bijektive Funktion von $A$ nach $B$ gibt.

SATZ 11.10. *Sei $C$ eine Menge. Dann ist $\sim$ auf $\mathcal{P}(C)$ eine Äquivalenzrelation.*

PROPOSITION 11.11. $\mathbb{Z} \sim \mathbb{N} \sim \mathbb{N} \times \mathbb{N}$.

*Beweis:* Die Funktion

$$\begin{array}{rccl} f & : & \mathbb{Z} & \longrightarrow & \mathbb{N} \\ & & x & \longmapsto & \begin{cases} -2x + 2 & \text{wenn } x \leq 0 \\ 2x - 1 & \text{wenn } x \geq 1 \end{cases} \end{array}$$

ist bijektiv. Ebenso ist

$$\begin{array}{rccl} f & : & \mathbb{N} \times \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & (x, y) & \longmapsto & 2^{x-1} \cdot (2y - 1) \end{array}$$

bijektiv. ■

DEFINITION 11.12. Eine Menge $A$ ist *endlich*, wenn es ein $n \in \mathbb{N}_0$ gibt, sodass $A \sim \{x \in \mathbb{N} \mid 1 \leq x \leq n\}$. Eine Menge $B$ ist *abzählbar unendlich*, wenn sie gleichmächtig zu $\mathbb{N}$ ist. Eine Menge ist *abzählbar*, wenn sie endlich oder abzählbar unendlich ist.

DEFINITION 11.13. Seien $A$, $B$ Mengen. Wir sagen, dass $B$ *mindestens so mächtig wie* $A$ ($A \precsim B$) ist, wenn es eine injektive Funktion von $A$ nach $B$ gibt. $B$ ist *mächtiger als* $A$, wenn $B$ mindestens so mächtig wie $A$ ist und $A$ und $B$ nicht gleichmächtig sind.

ÜBUNGSAUFGABEN 11.14.

(1) Seien $A, B$ Mengen mit $A \precsim B$. Zeigen Sie, dass es eine surjektive Funktion von $B$ auf $A$ gibt.
(2) Seien $A, B$ Mengen, sodass es eine surjektive Funktion $s$ von $B$ auf $A$ gibt. Zeigen Sie, dass dann $A \preceq B$. *Hinweis:* Verwenden Sie das Auswahlaxiom für $\prod_{a \in A} s^{-1}[\{a\}]$.
(3) Wir nehmen an, dass $A_1 \sim A_2$ und $B_1 \sim B_2$. Zeigen Sie, dass dann auch $A_1 \times B_1 \sim A_2 \times B_2$ und $\mathcal{P}(A_1) \sim \mathcal{P}(A_2)$.
(4) Sei $A$ eine Menge. Finden Sie eine bijektive Abbildung von $\mathcal{P}(A)$ nach $\{0, 1\}^A$.

Für jede Menge $A$ gilt $A \precsim A$.

SATZ 11.15. *Seien $A, B, C$ Mengen mit $A \precsim B$ und $B \precsim C$. Dann gilt $A \precsim C$.*

*Beweis:* Die Hintereinanderausführung injektiver Funktionen ist injektiv. ∎

Nun überlegen wir uns, was passiert, wenn $A \precsim B$ und $B \precsim A$. Dazu beweisen wir zuerst folgendes Lemma.

LEMMA 11.16. *Sei $Y$ eine Menge, und sei $U$ eine Teilmenge von $Y$. Wir nehmen an, dass es eine injektive Funktion $g : Y \to U$ gibt. Dann sind $Y$ und $U$ gleichmächtig.*

*Beweis:* Sei $V := Y \setminus U$, und

$$U_1 := \bigcap \{B \subseteq U \mid g[V \cup B] \subseteq B\}.$$

Wir zeigen nun

$$g[V \cup U_1] \subseteq U_1. \tag{11.2}$$

Sei dazu $w \in V \cup U_1$. Wir wollen zeigen, dass $g(w) \in \bigcap \{B \mid B \subseteq U$ und $g[V \cup B] \subseteq B\}$. Dazu zeigen wir, dass $g(w)$ in jeder Teilmenge $B$ von $U$ mit $g[V \cup B] \subseteq B$ liegt. Sei also $B \subseteq U$ so, dass $g[V \cup B] \subseteq B$. Wegen $U_1 \subseteq B$ gilt auch $w \in V \cup B$. Daher gilt $g(w) \in g[V \cup B]$, und somit $g(w) \in B$. Somit gilt (11.2).

Nun zeigen wir

$$g[V \cup U_1] = U_1. \tag{11.3}$$

Nehmen wir nun an, dass $g[V \cup U_1] \neq U_1$. Dann gibt es ein $u_1 \in U_1$, sodass $u_1 \notin g[V \cup U_1]$. Dann gilt $g[V \cup (U_1 \setminus \{u_1\})] \subseteq U_1 \setminus \{u_1\}$. Somit ist $B := U_1 \setminus \{u_1\}$ eine der Mengen, die bei der Bildung von $U_1$ geschnitten wurden. Also gilt $u_1 \notin U_1$, im Widerspruch zur Wahl von $u_1$. Somit gilt (11.3).

Somit ist $g|_{V \cup U_1}$ eine bijektive Funktion von $V \cup U_1$ nach $U_1$. Da $Y = V \cup U = (V \cup U_1) \cup (U \setminus U_1)$ und $U = U_1 \cup (U \setminus U_1)$, ist $h := g|_{V \cup U_1} \cup \mathrm{id}_{U \setminus U_1}$ eine bijektive Funktion von $Y$ nach $U$. ∎

SATZ 11.17 (Satz von Schröder-Bernstein). *Seien $A, B$ Mengen mit $A \precsim B$ und $B \precsim A$. Dann gilt $A \sim B$.*

*Beweis:* Sei $f : A \to B$ injektiv und $g : B \to A$ injektiv. Dann ist $f \circ g$ eine injektive Funktion, und es gilt $f \circ g[B] \subseteq f[A]$.

Sei nun $Y := B$ und $U := f[A]$. Nun ist $f \circ g$ eine injektive Funktion von $Y$ nach $U$. Nach Lemma 11.16 gibt es eine bijektive Funktion $h : Y \to U$. Nun ist $f$ bijektiv von $A$ nach $f[A]$, und $h^{-1}$ bijektiv von $f[A]$ nach $B$, also ist $h^{-1} \circ f$ bijektiv von $A$ nach $B$. Somit gilt $A \sim B$. ■

## 3. Einige abzählbar unendliche Mengen

SATZ 11.18. *Es gilt* $\mathbb{Q} \sim \mathbb{N}$.

*Beweis:* Nach dem Satz von Schröder-Bernstein genügt es $\mathbb{N} \precsim \mathbb{Q}$ und $\mathbb{Q} \precsim \mathbb{N}$ zu zeigen. Klarerweise ist $f : \mathbb{N} \to \mathbb{Q}, x \mapsto \frac{x}{1}$ injektiv, also gilt $\mathbb{N} \precsim \mathbb{Q}$.

Wir bilden nun eine injektive Abbildung $g : \mathbb{Q} \to \mathbb{Z} \times \mathbb{N}$ durch

$$g(\frac{a}{b}) := (a/\mathrm{ggT}(a, b), b/\mathrm{ggT}(a, b)),$$

wenn $b > 0$. Diese Abbildung ist wohldefiniert und injektiv. Da $\mathbb{Z} \times \mathbb{N} \sim \mathbb{N} \times \mathbb{N} \sim \mathbb{N}$, gibt es eine injektive Abbildung von $\mathbb{Z} \times \mathbb{N}$ nach $\mathbb{N}$, und somit gilt $\mathbb{Q} \precsim \mathbb{N}$. (Ebenso ist $h : \mathbb{Z} \times (\mathbb{N} \setminus \{0\}) \to \mathbb{Q}, (a, b) \mapsto \frac{a}{b}$ surjektiv auf $\mathbb{Q}$. Unter Verwendung des Auswahlaxioms gilt also deshalb $\mathbb{Q} \precsim \mathbb{Z} \times (\mathbb{N} \setminus \{0\})$, und folglich $\mathbb{Q} \precsim \mathbb{Z} \times \mathbb{N} \precsim \mathbb{N} \times \mathbb{N} \precsim \mathbb{N}$.) ■

SATZ 11.19. *Sei* $\langle A_i \mid i \in \mathbb{N} \rangle$ *eine Familie von Mengen. Wir nehmen an, dass für alle* $i \in \mathbb{N}$ *gilt :* $A_i \precsim \mathbb{N}$. *Dann gilt auch* $\bigcup_{i \in \mathbb{N}} A_i \precsim \mathbb{N}$.

*Beweis:* Sei $f_i : A_i \to \mathbb{N}$ injektiv. Wir bilden nun $f : \bigcup \{A_i \mid i \in \mathbb{N}\} \to \mathbb{N} \times \mathbb{N}$ durch $f(a) := (k_1, k_2)$, wobei $k_1 := \min\{j \in \mathbb{N} \mid a \in A_j\}$ und $k_2 := f_{k_1}(a)$. Diese Abbildung ist injektiv und beweist $\bigcup \{A_i \mid i \in \mathbb{N}\} \precsim \mathbb{N} \times \mathbb{N}$. Wegen $\mathbb{N} \times \mathbb{N} \precsim \mathbb{N}$ folgt die Behauptung. ■

ÜBUNGSAUFGABEN 11.20.

(1) Zeigen Sie, dass für jedes $a$ mit $a \notin \mathbb{N}$ die Menge $\{a\} \cup \mathbb{N}$ gleichmächtig zu $\mathbb{N}$ ist.
(2) Zeigen Sie, dass die Vereinigung einer abzählbar unendlichen mit einer endlichen Menge abzählbar unendlich ist.
(3) Zeigen Sie, dass eine Vereinigung von abzählbar vielen abzählbaren Mengen abzählbar ist, indem Sie eine surjektive Abbildung von $\mathbb{N} \times \mathbb{N}$ auf diese Menge definieren.
(4) Zeigen Sie, dass für jedes $n \in \mathbb{N}$ die Menge $\mathbb{N}^n$ gleichmächtig zu $\mathbb{N}$ ist.
(5) Zeigen Sie, dass für nichtleere abzählbare Menge $A$ die Menge $A^* := \bigcup \{A^n \mid n \in \mathbb{N}\}$ abzählbar unendlich ist.
(6) Zeigen Sie, dass die Menge der endlichen Teilmengen von $\mathbb{N}$ gleichmächtig zu $\mathbb{N}$ ist.

## 4. Einige überabzählbar unendliche Mengen

Eine Menge $C$ ist *überabzählbar unendlich*, wenn $C$ unendlich und nicht gleichmächtig zu $\mathbb{N}$ ist. Zunächst überlegen wir uns, warum es solche Mengen gibt.

LEMMA 11.21. *Sei $A$ eine Menge. Dann gibt es keine surjektive Funktion von $A$ auf $\mathcal{P}(A)$.*

*Beweis:* Sei $f : A \to \mathcal{P}(A)$. Wir zeigen, dass $f$ nicht surjektiv auf $\mathcal{P}(A)$ sein kann.

Wir betrachten dazu

$$B := \{x \in A \mid x \notin f(x)\}.$$

Wir zeigen nun, dass $B$ nicht im Wertebereich von $f$ liegt. Dazu zeigen wir, dass für alle $a \in A$ gilt: $f(a) \neq B$. Sei also $a \in A$.

- *1.Fall: $a \in f(a)$.* Wenn $a \in f(a)$, so gilt $a \notin B$. Das Element $a$ liegt also in $f(a)$, aber nicht in $B$. Somit gilt $f(a) \neq B$.
- *2. Fall: $a \notin f(a)$.* Wenn $a \notin f(a)$, so gilt $a \in B$. Das Element $a$ liegt also in $B$, aber nicht in $f(a)$. Somit gilt $f(a) \neq B$.

$B$ liegt also nicht im Wertebereich von $f$; somit ist $f$ nicht surjektiv auf $\mathcal{P}(A)$. ■

Somit gilt:

SATZ 11.22 (Satz von Cantor). *Sei $A$ eine Menge. Dann gilt $A \precsim \mathcal{P}(A)$, und $A \nsim \mathcal{P}(A)$.*

*Beweis:* Die Abbildung $f : A \to \mathcal{P}(A)$, $a \mapsto \{a\}$ ist injektiv, also gilt $A \precsim \mathcal{P}(A)$.

Wenn $A \sim \mathcal{P}(A)$, so gibt es eine bijektive Abbildung von $A$ nach $\mathcal{P}(A)$. Diese Abbildung ist surjektiv auf $\mathcal{P}(A)$, im Widerspruch zu Lemma 11.21. Also gilt $A \nsim \mathcal{P}(A)$. ■

Nach dem Satz von Cantor ist also $\mathcal{P}(\mathbb{N})$ überabzählbar.

SATZ 11.23. *Die Menge $\mathbb{R}$ der reellen Zahlen ist gleichmächtig zu $\mathcal{P}(\mathbb{N})$, also überabzählbar.*

*Beweis:* Die Funktion $f : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$, $I \mapsto \sum_{i \in I} 10^{-i}$ ist injektiv und belegt $\mathcal{P}(\mathbb{N}) \precsim \mathbb{R}$.

Sei nun $q$ eine bijektive Funktion von $\mathbb{N}$ nach $\mathbb{Q}$. Wir schreiben für $q(i)$ kurz $q_i$. Dann ist $g : \mathbb{R} \to \mathcal{P}(\mathbb{N})$, $r \mapsto \{i \in \mathbb{N} \mid q_i < r\}$ injektiv, da zwischen zwei verschiedenen reellen Zahlen stets eine rationale Zahl liegt.

Somit gilt nach dem Satz von Schröder-Bernstein $\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$. ■

## 5. Unendliche Mengen

In dieser Sektion stellen wir noch einige Sätze über unendliche Mengen zusammen. Diese Sätze haben gemeinsam, dass man für die Beweise das Auswahlaxiom benötigt.

SATZ 11.24. *Jede unendliche Menge $M$ enthält eine abzählbar unendliche Teilmenge.*

*Beweis:* Sei $f : \mathcal{P}(M) \setminus \{\emptyset\} \to M$ so, dass $f(A) \in A$ für alle $A \subseteq M$. So ein $f$ existiert, weil nach dem Auswahlaxiom die Menge $\prod_{A \in \mathcal{P}(M) \setminus \{\emptyset\}} A$ nicht leer ist.

Sei $\mathcal{F}$ die Menge aller endlichen Teilmengen von $M$. Wir definieren eine Funktion $E : \mathbb{N} \to \mathcal{F}$ rekursiv. Da $M$ nicht leer ist, können wir können wir $E(1) := \{f(M)\}$ definieren, und für alle $n \in \mathbb{N}$: $E(n+1) = E(n) \cup \{f(M \setminus E(n))\}$.

Sei nun $g : \mathbb{N} \to M$ definiert durch $g(n) := f(M \setminus E(n))$. Wir zeigen nun, dass $g$ injektiv ist. Sei $n_1 < n_2$. Es gilt $g(n_2) = f(M \setminus E(n_2)) \notin E(n_2)$. Da $g(n_1) = f(M \setminus E(n_1))$, gilt $g(n_1) \in E(n_1) \cup \{f(M \setminus E(n_1))\}$, also $g(n_1) \in E(n_1 + 1)$, und somit $g(n_1) \in E(n_2)$. Also gilt $g(n_1) \neq g(n_2)$. Folglich ist $g[\mathbb{N}]$ eine abzählbare Teilmenge von $M$. ■

ÜBUNGSAUFGABEN 11.25.

(1) Sei $A$ unendlich und $E$ endlich. Zeigen Sie $A \cup E \sim A$. *Hinweis:* Benutzen Sie eine abzählbare Teilmenge $B$ von $A$ und verwenden Sie $B \cup E \sim B$.
(2) Sei $B$ unendlich. Zeigen Sie, dass es eine Funktion $f : B \to B$ gibt, die injektiv, aber nicht surjektiv ist. *Hinweis:* Lösen Sie das Beispiel zuerst für $B := \mathbb{N}$.
(3) Zeigen Sie $[0, 1] \sim \,]0, 1[\, \sim \mathbb{R}$.

SATZ 11.26 (Vergleichbarkeitssatz). *Seien $A, B$ Mengen. Dann gilt $A \precsim B$ oder $B \precsim A$.*

*Beweisskizze:* Sei

$$\mathcal{F} := \{f \subseteq A \times B \mid \text{es gibt } C \subseteq A, \text{ sodass } f \text{ eine injektive Funktion von } C \text{ nach } B \text{ ist.}\}$$

Mit dem Lemma von Zorn kann man zeigen, dass $(\mathcal{F}, \subseteq)$ ein maximales Element $f_0$ besitzt. Wenn der Definitionsbereich $C$ von $f_0$ gleich $A$ ist, so gilt $A \precsim B$.

Wenn $f_0$ surjektiv auf $B$ ist, so ist $f_0$ eine bijektive Funktion von $C$ nach $B$; also ist $f_0^{-1}$ injektiv von $B$ nach $C$, und es gilt $B \precsim A$.

Wenn $C \neq A$ und $f_0[C] \neq B$, so wählen wir $a \in A \setminus C$ und $b \in B \setminus f_0[C]$. Dann ist $f_0 \cup \{(a, b)\}$ ebenfalls injektiv, im Widerspruch zur Maximalität von $f_0$. ■

SATZ 11.27. *Seien $A, B$ Mengen mit $A \precsim B$. Wir nehmen an, dass $B$ unendlich ist. Dann gilt*

(1) *$A \cup B \sim B$;*
(2) *Wenn $A$ nicht leer ist, so gilt $A \times B \sim B$;*
(3) *Wenn $A$ zumindest zwei Elemente enthält, so gilt $A^B \sim \mathcal{P}(B)$.*

*Beweis:* [**Halmos, 1976**, Kapitel 24]. ■

ÜBUNGSAUFGABEN 11.28.

(1) Zeigen Sie $\mathbb{N}^{\mathbb{R}} \sim \mathcal{P}(\mathbb{R})$.
(2) Zeigen Sie $\mathbb{R}^{\mathbb{N}} \sim \mathbb{R}$. *Hinweis:* Finden Sie eine injektive Abbildung von $\mathcal{P}(N)^{\mathbb{N}}$ nach $\mathcal{P}(\mathbb{N} \times \mathbb{N})$.

Auf Satz 11.27 aufbauend kann man nun beweisen, dass alle Basen eines Vektorraums gleichmächtig sind.

SATZ 11.29 (Dimensionssatz für Vektorräume). *Sei $V$ ein Vektorraum über dem Körper $K$, und seien $B$ und $C$ Basen von $V$. Dann gilt $B \sim C$.*

## 6. Erstaunliches über Mengen

SATZ 11.30. *Sei $A$ eine Menge, und sei $B := \{a \in A \mid a \notin a\}$. Dann gilt $B \notin A$.*

*Beweis:* Nehmen wir an, dass $B \in A$.

- *1.Fall: $B \notin B$*: Dann gilt $B \in A$ und $B \notin B$. Also gilt $B \in B$, im Widerspruch zur Fallannahme. Dieser Fall kann also nicht auftreten.
- *2. Fall: $B \in B$*: Dann erfüllt $B$ die Eigenschaft, die unter den Elementen von $A$ jene in $B$ auswählt; es gilt also $B \notin B$, im Widerspruch zur Fallannahme.

Somit ist die Annahme $B \in A$ falsch; es gilt also $B \notin A$. ∎

Damit gibt es auch keine Menge, die alle Mengen als Elemente enthalten würde: jede Menge $M$ enthält zumindest die Menge $\{m \in M \mid m \notin m\}$ nicht als Element. Der Begriff "die Menge aller Mengen" ist also widersprüchlich, weil er von einem Objekt, das es nicht gibt, nämlich einer Menge aller Mengen, so spricht, als ob es dieses Objekt gäbe. Dass es eine "Menge aller Mengen" nicht gibt, ist das *Russell'sche Paradoxon*.

Eine berühmte Vermutung (die Kontinuumshypothese) sagt, dass folgende Frage die Antwort "ja" hat.

PROBLEM 11.31. *Gilt für jede unendliche Teilmenge $A$ von $\mathbb{R}$ : $A \sim \mathbb{R}$ oder $A \sim \mathbb{N}$?*

K. Gödel zeigte, dass die Axiome der Zermelo-Fraenkelschen Mengenlehre, wenn widerspruchsfrei, auch unter Zuhilfenahme des Auswahlaxioms nicht erlauben, die Antwort "nein" herzuleiten. P. Cohen zeigte, dass die Axiome der Zermelo-Fraenkelschen Mengenlehre und das Auswahlaxiom, wenn widerspruchsfrei, nicht erlauben, die Antwort "ja" herzuleiten.

# Teil 5

# Analyse linearer Abbildungen

KAPITEL 12

# Determinanten

## 1. Volumen eines Parallelepipeds

Sei $A = \begin{pmatrix} -z_1- \\ -z_2- \\ \vdots \\ -z_m- \end{pmatrix}$ eine reelle $m \times m$-Matrix. Das *von den Zeilen von A aufgespannte Parallelepiped* ist die Teilmenge

$$\{\sum_{i=1}^{m} \lambda_i \cdot z_i \mid \lambda_1, \dots, \lambda_m \in [0, 1]\}$$

von $\mathbb{R}^m$. Wir versuchen jetzt, das Volumen dieses Parallelepipeds zu messen, und schreiben $D(z_1, z_2, \dots, z_m)$ für dieses Volumen.

Ohne den Begriff Volumen im $\mathbb{R}^m$ definiert zu haben, könnte man vermuten, dass eine Funktion $D : \mathbb{R}^m \times \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}$, die dieses Volumen misst, folgende Eigenschaften hat. Da wir die Eigenschaften nicht nur für $\mathbb{R}$ benötigen, formulieren wir diese Eigenschaften für einen kommutativen Ring mit Eins $K$. $D$ ist dann eine Funktion von $K^m \times \cdots \times K^m$ nach $K$.

(D1) Für alle $z_1, \dots, z_m, y \in K^m$, $\alpha, \beta \in K$ und $i \in \{1, \dots, m\}$ gilt

$$\begin{aligned} D(z_1, \dots, z_{i-1},\ \alpha * z_i + \beta * y,\ z_{i+1}, \dots, z_m) \\ = \alpha\, D(z_1, \dots, z_{i-1},\ z_i,\ z_{i+1}, \dots, z_m) \\ + \beta\, D(z_1, \dots, z_{i-1},\ y,\ z_{i+1}, \dots, z_m). \end{aligned}$$

Eine Funktion $D$, die diese Eigenschaft besitzt, nennen wir *multilinear*.

(D2) Für alle $z_1, \dots, z_m$ gilt: wenn es $i, j \in \{1, 2, \dots, m\}$ gibt, sodass $i \neq j$ und $z_i = z_j$, so gilt

$$D(z_1, \dots, z_m) = 0.$$

Eine Funktion $D$, die diese Eigenschaft besitzt, nennen wir *alternierend*.

(D3) Für die Einheitsvektoren $e_1, \dots, e_m$ im $K^m$ gilt

$$D(e_1, \dots, e_m) = 1.$$

SATZ 12.1. *Sei $K$ ein kommutativer Ring mit $1$, sei $m \in \mathbb{N}$, sei $D : (K^m)^m \to K$ eine Abbildung, die (D1), (D2) und (D3) erfüllt, seien $z_1, \dots, z_m \in K^m$, $\alpha \in K$, und seien $i, j \in \{1, \dots, m\}$ mit $i \neq j$. Dann gilt*

(1) $D(z_1, \dots, z_{j-1},\ \alpha * z_i + z_j,\ z_{j+1}, \dots, z_m) = D(z_1, \dots, z_m)$.

(2) *Wenn $i < j$, so gilt*

$$D(z_1, \ldots, z_{i-1},\ z_j,\ z_{i+1}, \ldots, z_{j-1},\ z_i, z_{j+1}, \ldots, z_m) = -D(z_1, \ldots, z_m).$$

*Beweis:*

(1) Wegen der Multilinearität von $D$ gilt $D(z_1, \ldots, z_{j-1}, \alpha * z_i + z_j, z_{j+1}, \ldots, z_m) = \alpha\, D(z_1, \ldots, z_{j-1},\ z_i,\ z_{j+1}, \ldots, z_m) + D(z_1, \ldots, z_m)$. Da $D$ alternierend ist, ist der erste Summand gleich 0.

(2) Da $D$ alternierend ist, gilt

$$D(z_1, \ldots, z_{i-1},\ z_i + z_j,\ z_{i+1}, \ldots, z_{j-1},\ z_i + z_j,\ z_{j+1}, \ldots, z_m) = 0.$$

Wegen der Multlinearität gilt also

$$\begin{aligned}
0 &= D(z_1, \ldots, z_{i-1},\ z_i,\ z_{i+1}, \ldots, z_{j-1},\ z_i, z_{j+1}, \ldots, z_m) \\
&\quad + D(z_1, \ldots, z_{i-1},\ z_i,\ z_{i+1}, \ldots, z_{j-1},\ z_j, z_{j+1}, \ldots, z_m) \\
&\quad + D(z_1, \ldots, z_{i-1},\ z_j,\ z_{i+1}, \ldots, z_{j-1},\ z_i,\ z_{j+1}, \ldots, z_m) \\
&\quad + D(z_1, \ldots, z_{i-1},\ z_j,\ z_{i+1}, \ldots, z_{j-1},\ z_j, z_{j+1}, \ldots, z_m) \\
&= 0 + D(z_1, \ldots, z_{i-1},\ z_i,\ z_{i+1}, \ldots, z_{j-1},\ z_j, z_{j+1}, \ldots, z_m) \\
&\quad + D(z_1, \ldots, z_{i-1},\ z_j,\ z_{i+1}, \ldots, z_{j-1},\ z_i,\ z_{j+1}, \ldots, z_m) + 0.
\end{aligned}$$

■

## 2. Permutationen

DEFINITION 12.2. Sei $X$ eine Menge, $m \in \mathbb{N}$. Eine *Permutation von $X$* ist eine bijektive Abbildung von $X$ nach $X$. Die Menge aller Bijektionen auf $X$ kürzt man auch mit $S_X$, die Menge aller Bijektionen auf $\{1, \ldots, m\}$ mit $S_m$ ab.

Permutationen von $\{1, \ldots, m\}$ kann man durch die Wertetabelle angeben. So beschreibt

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 4 & 6 & 7 & 1 & 5 \end{pmatrix}$$

die Abbildung $f = \{(1, 3), (2, 2), \ldots, (7, 5)\}$; es gilt also $f(1) = 3, \ldots, f(7) = 5$.

DEFINITION 12.3 (Signatur). Sei $f : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$ eine bijektive Abbildung. Wir definieren die *Signatur* von $f$ folgendermaßen: Wenn $m = 1$, so gilt $\operatorname{sgn}(f) = 1$. Wenn $m \geq 2$, so definieren wir

$$\operatorname{sgn}(f) := \prod_{(i,j)\in\{(k,l)\in\{1,\ldots,m\}\times\{1,\ldots,m\}\ \mid\ k<l\}} \frac{f(i) - f(j)}{i - j}.$$

SATZ 12.4. *Sei $m \in \mathbb{N}$, und seien $f, g \in S_m$. Sei $k$ die Anzahl der Elemente der Menge*

$$\{(i, j) \in \{1, 2, \ldots, m\}^2 \mid i < j \text{ und } f(i) > f(j)\}.$$

*Dann gilt:*

(1) $\operatorname{sgn}(f) = (-1)^k$.
(2) $\operatorname{sgn}(f \circ g) = \operatorname{sgn}(f) \cdot \operatorname{sgn}(g)$.
(3) $\operatorname{sgn}(\mathrm{id}) = 1$.
(4) $\operatorname{sgn}(f^{-1}) = \operatorname{sgn}(f)$.

*Beweis:* (1) $\operatorname{sgn}(f) = \dfrac{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} f(i) - f(j)}{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} i - j}$. Für den Zähler dieses Produkts gilt:

$$\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} f(i) - f(j)$$
$$= \Big(\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j \text{ und } f(i)<f(j)}} f(i) - f(j)\Big) \cdot \Big(\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j \text{ und } f(i)>f(j)}} f(j) - f(i)\Big) \cdot (-1)^k.$$

Wir behaupten nun, dass dieses Produkt gleich $(-1)^k \prod\limits_{\substack{(s,t)\in\{1,\ldots,m\}^2\\ s<t}} (s - t)$ ist. Seien $a, b$ so, dass $f(a) = s$ und $f(b) = t$. Wenn $a < b$, so tritt der Faktor $s - t$ im ersten Produkt auf, wenn $b < a$, so tritt der Faktor $s - t$ im zweiten Produkt auf. (2) Es gilt

$$\operatorname{sgn}(f \circ g) = \frac{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} f(g(i)) - f(g(j))}{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} i - j} \tag{12.1}$$
$$= \frac{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} f(g(i)) - f(g(j))}{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} g(i) - g(j)} \cdot \frac{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} g(i) - g(j)}{\prod\limits_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} i - j}.$$

Für den ersten Bruch in der rechten Seite von (12.1) gilt

$$\begin{aligned}
\prod_{\substack{(i,j)\in\{1,\dots,m\}^2\\ i<j}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)}
&= \prod_{\substack{(i,j)\in\{1,\dots,m\}^2\\ i<j \text{ und } g(i)<g(j)}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)} \cdot \prod_{\substack{(i,j)\in\{1,\dots,m\}^2\\ i<j \text{ und } g(i)>g(j)}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)}\\
&= \prod_{\substack{(i,j)\in\{1,\dots,m\}^2\\ i<j \text{ und } g(i)<g(j)}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)} \cdot \prod_{\substack{(i,j)\in\{1,\dots,m\}^2\\ i<j \text{ und } g(i)>g(j)}} \frac{f(g(j))-f(g(i))}{g(j)-g(i)}\\
&= \prod_{\substack{(s,t)\in\{1,\dots,m\}^2\\ s<t}} \frac{f(s)-f(t)}{s-t} = \operatorname{sgn}(f).
\end{aligned}$$

Der zweite Bruch in der rechten Seite von (12.1) ist $\operatorname{sgn}(g)$. (3) folgt unmittelbar aus der Definition der Signatur. (4) $\operatorname{sgn}(f^{-1}) \cdot \operatorname{sgn}(f) = \operatorname{sgn}(f^{-1} \circ f) = \operatorname{sgn}(\mathrm{id}) = 1$. ■

DEFINITION 12.5 (Zyklen). Seien $i_1, i_2, \dots, i_n$ paarweise verschiedene Zahlen in $\{1, \dots, m\}$. Dann bezeichnen wir mit $(i_1\ i_2\ i_3\ \dots\ i_n)$ die Permutation $f$ mit $f(i_1) = i_2, \dots, f(i_{n-1}) = i_n$, $f(i_n) = i_1$, $f(k) = k$ für alle $k \in \{1, \dots, m\} \setminus \{i_1, \dots, i_n\}$.

Für $i \neq j$ bezeichnen wir den Zweierzyklus $(i\ \ j)$ auch als *Transposition*.

SATZ 12.6. *Sei $m \in \mathbb{N}$, $a, b \in \mathbb{N}_0$, $i, j \in \{1, \dots, m\}$, und sei $\sigma \in S_m$.*

(1) *$\sigma$ ist ein Produkt von endlich vielen Transpositionen. (Das Produkt von* 0 *Transpositionen definieren wir dabei als* id.*)*
(2) *Wenn $i \neq j$, so gilt* $\operatorname{sgn}((i\ \ j)) = -1$.
(3) *Für alle Transpositionen $\rho_1, \dots, \rho_a$ und $\tau_1, \dots, \tau_b$ mit $\sigma = \rho_1 \circ \dots \circ \rho_a = \tau_1 \circ \dots \circ \tau_b$ teilt* 2 *die Differenz $a - b$.*
(4) *Für $n \in \mathbb{N}$ und paarweise verschiedene* $i_1, \dots, i_n \in \{1, \dots, m\}$ *gilt*

$$\operatorname{sgn}((i_1\ i_2\ \dots\ i_n)) = (-1)^{n+1}.$$

*Beweis:* (1) Sei

$$M(\sigma) := \max(\{0\} \cup \{k \in \{1, \dots, m\} \mid \sigma(k) \neq k\}).$$

Wir zeigen nun mit Induktion nach $n$, dass alle Permutationen mit $M(\sigma) = n$ Produkt von Transpositionen sind. Für $n = 0$ gilt $\sigma = \mathrm{id}$; $\sigma$ ist dann also das Produkt von 0 Transpositionen. Sei nun $n \geq 1$, und sei $\sigma$ so, dass $M(\sigma) = n$. Sei $k := \sigma(n)$. Es gilt $k < n$. Sei $\rho := (k\ n) \circ \sigma$. Es gilt $\rho(n) = n$ und $\rho(r) = r$ für alle $r > n$. Also gilt $M(\rho) < n$. Somit gibt es nach Induktionsvoraussetzung Transpositionen $\tau_1, \dots, \tau_l$ mit

$\rho = \tau_1 \circ \cdots \circ \tau_l$. Also gilt $\sigma = (k\ n)^{-1} \circ \tau_1 \circ \cdots \circ \tau_l = (k\ n) \circ \tau_1 \circ \cdots \circ \tau_l$. Somit ist $\sigma$ ebenfalls ein Produkt von Transpositionen.

(2) Sei $\sigma := (i\ j)$. Wir nehmen an, dass $i < j$. Wir bestimmen die Anzahl der Elemente in $\{(k,l) \in \{1,\ldots,m\}^2 \mid k < l \text{ und } \sigma(k) > \sigma(l)\}$. Diese Menge ist gleich $\{(i,i+1),(i,i+2),\ldots,(i,j),(i+1,j),(i+2,j),\ldots,(i-1,j)\}$. Sie hat also $(i-j)+(i-j-1) = 2(i-j)-1$ viele Elemente. Da diese Zahl ungerade ist, gilt wegen Satz 12.4 (1), dass die Signatur von $(i\ j)$ gleich $-1$ ist.

Im Fall $i > j$ beobachten wir, dass $(i\ j) = (j\ i)$, und erhalten dann aus dem ersten Fall, dass $\mathrm{sgn}((j\ i)) = -1$.

(3) Die Signatur von $\sigma$ ist $(-1)^a = (-1)^b$.

(4) Es gilt $(i_1\ i_2\ \ldots\ i_n) = (i_1\ i_n) \circ (i_1\ i_2\ \ldots,\ i_{n-1})$, somit folgt die behauptete Gleichheit aus (2) durch Induktion nach $n$. ■

SATZ 12.7. *Sei $m \geq 2$, und seien $i, j \in \{1,\ldots,m\}$ mit $i < j$. Sei*

$$A_m := \{f \in S_m \mid \mathrm{sgn}(f) = 1\},$$

*und sei $(i\ j) \circ A_m := \{(i\ j) \circ f \mid f \in A_m\}$. Dann gilt $A_m \cap ((i\ j) \circ A_m) = \emptyset$ und $A_m \cup ((i\ j) \circ A_m) = S_m$; außerdem ist $\varphi : A_m \to (i\ j) \circ A_m,\ f \mapsto (i\ j) \circ f$ bijektiv.*

*Beweis:* Alle Elemente in $A_m$ haben Signatur 1, alle Elemente in $(i\ j) \circ A_m$ haben Signatur $-1$, folglich ist ihr Schnitt leer.

Sei nun $f \in S_m$. Wenn $\mathrm{sgn}(f) = 1$, so liegt $f$ in $A_m$. Wenn $\mathrm{sgn}(f) = -1$, so gilt $f = (i\ j) \circ (i\ j) \circ f$, und da $(i\ j) \circ f$ in $A_m$ liegt, gilt $f \in (i\ j) \circ A_m$.

Um die Bijektivität von $\varphi$ zu zeigen, definieren wir $\psi : (i\ j) \circ A_m \to A_m,\ f \mapsto (i\ j) \circ f$. Dann gilt $\psi \circ \varphi = \mathrm{id}_{A_m}$ und $\varphi \circ \psi = \mathrm{id}_{(i\ j)\circ A_m}$, folglich ist $\varphi$ bijektiv. ■

ÜBUNGSAUFGABEN 12.8.

(1) Der Beweis von Satz 12.6 (1) liefert eine Zerlegung von $\begin{pmatrix} 1&2&3&4&5&6&7 \\ 3&2&4&6&7&1&5 \end{pmatrix}$ in ein Produkt von Transpositionen. Geben Sie diese Transpositionen an!
(2) Seien $f, g \in S_m$. Sei $F : S_m \to S_m,\ h \mapsto f \circ h \circ g$. Zeigen Sie, dass $F$ bijektiv ist.
(3) Sei $F : S_m \to S_m,\ F(\sigma) := \sigma^{-1}$ für $\sigma \in S_m$. Zeigen Sie, dass $F$ bijektiv ist.

SATZ 12.9. *Sei $K$ ein kommutativer Ring mit 1, sei $m \in \mathbb{N}$, sei $D : (K^m)^m \to K$ eine Abbildung, die (D1), (D2) und (D3) erfüllt, seien $z_1,\ldots,z_m \in K^m$, und sei $f \in S_m$. Dann gilt*

$$D(z_{f(1)},\ldots,z_{f(m)}) = \mathrm{sgn}(f)\,D(z_1,\ldots,z_m). \tag{12.2}$$

*Beweis:* Wir zeigen durch Induktion nach $n$, dass die Gleichung (12.2) für alle Permutationen gilt, die Produkt von genau $n$ Transpositionen sind. Klarerweise gilt (12.2) für $f = \mathrm{id}$. Sei nun $f = \tau_n \circ \cdots \circ \tau_1$, wobei alle $\tau_i$ Transpositionen sind, und sei $g := \tau_n \circ \cdots \circ \tau_2$. Dann gilt wegen Satz 12.1

$$D(z_{g(\tau_1(1))},\ldots,z_{g(\tau_1(m))}) = -D(z_{g(1)},\ldots,z_{g(m)}).$$

Nach Induktionsvoraussetzung ist dieser Ausdruck gleich $-\mathrm{sgn}(g)\, D(z_1, \ldots, z_m)$. Da $f = g \circ \tau_1$, gilt $\mathrm{sgn}(f) = -\mathrm{sgn}(g)$. Also gilt $D(z_{f(1)}, \ldots, z_{f(m)}) = \mathrm{sgn}(f)\, D(z_1, \ldots, z_m)$. Das beschließt den Induktionsbeweis.

Da sich jede Permutation als Produkt von endlich vielen Transpositionen schreiben lässt, gilt (12.2) also für alle $f \in S_m$. ■

## 3. Determinante einer quadratischen Matrix

Eine Funktion $D$ mit den in Sektion 1 geforderten Eigenschaften erhält man mithilfe der *Determinante*.

DEFINITION 12.10. Sei $K$ ein kommutativer Ring mit Eins, $m \in \mathbb{N}$, und sei $A \in K^{m\times m}$. Dann definieren wir die *Determinante* von $A$ durch

$$\det(A) := \sum_{f \in S_m} \mathrm{sgn}(f) \prod_{i=1}^{m} A(i, f(i)).$$

SATZ 12.11. *Sei $K$ ein kommutativer Ring mit Eins, sei $m \in \mathbb{N}$, sei $A$ eine $m\times m$-Matrix über $K$, und sei $\mathbf{D} : K^m \times K^m \times \cdots \times K^m \to K$ gegeben durch*

$$\begin{array}{rccc} \mathbf{D} : & (K^m)^m & \longrightarrow & K \\ & (z_1, z_2, \ldots, z_m) & \longmapsto & \det\left(\begin{pmatrix} -z_1- \\ -z_2- \\ \vdots \\ -z_m- \end{pmatrix}\right) \end{array}.$$

*Dann erfüllt $\mathbf{D}$ die Eigenschaften (D1), (D2), und (D3).*

*Beweis:* Wir zeigen als erstes (D1). Seien dazu $z_1, \ldots, z_m, y \in K^m$, $\alpha, \beta \in K$ und $i \in \{1, \ldots, m\}$. Seien $A, B, C$ die $m \times m$-Matrizen, die durch

$$A = \begin{pmatrix} - & z_1 & - \\ & \vdots & \\ - & z_{i-1} & - \\ - & \alpha * z_i + \beta * y & - \\ - & z_{i+1} & - \\ & \vdots & \\ - & z_m & - \end{pmatrix}, \quad B = \begin{pmatrix} - & z_1 & - \\ & \vdots & \\ - & z_{i-1} & - \\ - & z_i & - \\ - & z_{i+1} & - \\ & \vdots & \\ - & z_m & - \end{pmatrix}, \quad C = \begin{pmatrix} - & z_1 & - \\ & \vdots & \\ - & z_{i-1} & - \\ - & y & - \\ - & z_{i+1} & - \\ & \vdots & \\ - & z_m & - \end{pmatrix}$$

gegeben sind. Nun gilt

$$
\begin{aligned}
&\mathbf{D}(z_1, \ldots, z_{i-1}, \alpha * z_i + \beta * y, z_{i+1}, \ldots, z_m) \\
&\quad = \det(A) \\
&\quad = \sum_{f \in S_m} \operatorname{sgn}(f) \prod_{j=1}^{m} A(j, f(j)) \\
&\quad = \sum_{f \in S_m} \operatorname{sgn}(f) \Big( \prod_{\substack{j=1 \\ j \neq i}}^{m} A(j, f(j)) \Big) A(i, f(i)) \\
&\quad = \sum_{f \in S_m} \operatorname{sgn}(f) \Big( \prod_{\substack{j=1 \\ j \neq i}}^{m} A(j, f(j)) \Big) \big( \alpha B(i, f(i)) + \beta C(i, f(i)) \big) \\
&\quad = \sum_{f \in S_m} \operatorname{sgn}(f) \Big( \prod_{\substack{j=1 \\ j \neq i}}^{m} A(j, f(j)) \Big) \alpha\, B(i, f(i)) \\
&\qquad + \sum_{f \in S_m} \operatorname{sgn}(f) \Big( \prod_{\substack{j=1 \\ j \neq i}}^{m} A(j, f(j)) \Big) \beta\, C(i, f(i)) \\
&\quad = \alpha \sum_{f \in S_m} \operatorname{sgn}(f) \Big( \prod_{j=1}^{m} B(j, f(j)) \Big) + \beta \sum_{f \in S_m} \operatorname{sgn}(f) \Big( \prod_{j=1}^{m} C(j, f(j)) \Big) \\
&\quad = \alpha \det(B) + \beta \det(C) \\
&\quad = \alpha\, \mathbf{D}(z_1, \ldots, z_m) + \beta\, \mathbf{D}(z_1, \ldots, z_{i-1}, \; y, \; z_{i+1}, \ldots, z_m).
\end{aligned}
$$

Wir zeigen als nächstes die Eigenschaft (D2). Sei dazu $A$ eine Matrix, deren $i$-te und $j$-Zeile gleich sind. Dann gilt

$$\begin{aligned}
\det(A) &= \sum_{f\in S_m} \operatorname{sgn}(f) \prod_{k=1}^{m} A(k, f(k)) \\
&= \sum_{f\in A_m} \operatorname{sgn}(f) \prod_{k=1}^{m} A(k, f(k)) + \sum_{f\in A_m} \operatorname{sgn}(f \circ (i\ \ j)) \prod_{k=1}^{m} A(k, (f \circ (i, j))\ (k)) \\
&= \sum_{f\in A_m} \prod_{k=1}^{m} A(k, f(k)) - \sum_{f\in A_m} \Big( \prod_{\substack{k=1\\ k\notin\{i,j\}}}^{m} A(k, f(k))\Big) \cdot A(i, f(j)) \cdot A(j, f(i)) \\
&= \sum_{f\in A_m} \prod_{k=1}^{m} A(k, f(k)) - \sum_{f\in A_m} \Big( \prod_{\substack{k=1\\ k\notin\{i,j\}}}^{m} A(k, f(k))\Big) \cdot A(j, f(j)) \cdot A(i, f(i)) \\
&= \sum_{f\in A_m} \prod_{k=1}^{m} A(k, f(k)) - \sum_{f\in A_m} \prod_{k=1}^{m} A(k, f(k)) \\
&= 0.
\end{aligned}$$

Um (D3) zu zeigen, beobachten wir, dass $\det(E_m) = 1$. ■

SATZ 12.12. *Sei $K$ ein kommutativer Ring mit Eins, sei $m \in \mathbb{N}$, und seien $A$ und $B$ $m \times m$-Matrizen über $K$. Dann gilt* $\det(A) = \det(A^T)$ *und* $\det(A \cdot B) = \det(A) \cdot \det(B)$.

*Beweis.* Zuerst zeigen wir $\det(A) = \det(A^T)$. Es gilt

$$\begin{aligned}
\det(A^T) &= \sum_{f\in S_m} \operatorname{sgn}(f) \prod_{i=1}^{m} A^T(i, f(i)) \\
&= \sum_{f\in S_m} \operatorname{sgn}(f) \prod_{i=1}^{m} A(f(i), i) \\
&= \sum_{f\in S_m} \operatorname{sgn}(f) \prod_{j=1}^{m} A(j, f^{-1}(j)) \\
&= \sum_{f\in S_m} \operatorname{sgn}(f^{-1}) \prod_{j=1}^{m} A(j, f^{-1}(j)) \\
&= \sum_{g\in S_m} \operatorname{sgn}(g) \prod_{j=1}^{m} A(j, g(j)).
\end{aligned}$$

Nun beweisen wir die zweite Eigenschaft: Seien $b_1, \ldots, b_m$ die Zeilen der Matrix $B$. Dann steht in der ersten Zeile des Produkts $A \cdot B$ der Vektor

$$A(1,1)b_1 + A(1,2)b_2 + \cdots + A(1,m)b_m.$$

Wenn wir alle $m$ Zeilen des Produkts ausrechnen, dann sehen wir

$$\begin{aligned}\det(A \cdot B) = \mathbf{D}\big(\; & A(1,1)b_1 + A(1,2)b_2 + \cdots + A(1,m)b_m, \\ & A(2,1)b_1 + A(2,2)b_2 + \cdots + A(2,m)b_m, \\ & \ldots \\ & A(m,1)b_1 + A(m,2)b_2 + \cdots + A(m,m)b_m\big).\end{aligned}$$

Wir nutzen nun die Multilinearität der Funktion $\mathbf{D}$ und erhalten

$$\det(A \cdot B) = \sum_{f:\{1,\ldots,m\}\to\{1,\ldots,m\}} \Big(\prod_{i=1}^{m} A(i, f(i))\Big) \cdot \mathbf{D}(b_{f(1)}, b_{f(2)}, \ldots, b_{f(m)}).$$

Die Determinante ist 0, wenn zwei Zeilen gleich sind. Daher brauchen wir nur über die bijektiven Funktionen zu summieren und erhalten:

$$\det(A \cdot B) = \sum_{f \in S_m} \Big(\prod_{i=1}^{m} A(i, f(i))\Big) \cdot \mathbf{D}(b_{f(1)}, b_{f(2)}, \ldots, b_{f(m)}).$$

Wegen Satz 12.9 gilt für eine Bijektion $f$: $\mathbf{D}(b_{f(1)}, b_{f(2)}, \ldots, b_{f(m)}) = \operatorname{sgn}(f)\, \mathbf{D}(b_1, b_2, \ldots, b_m)$. Also erhalten wir

$$\det(A \cdot B) = \sum_{f \in S_m} \Big(\Big(\prod_{i=1}^{m} A(i, f(i))\Big) \operatorname{sgn}(f)\ \det(B)\Big) = \det(B) \cdot \det(A).$$

■

SATZ 12.13. *Sei $K$ ein Körper, sei $m \in \mathbb{N}$, und sei $A$ eine $m \times m$-Matrix über $K$. Dann sind äquivalent:*

(1) *Die Zeilen von $A$ sind linear unabhängig.*
(2) $\det(A) \neq 0$.

*Beweis.* (1)⇒(2): Da die Zeilenvektoren linear unabhängig sind, ist der Zeilenraum von $A$ ganz $K^n$. Somit gibt es eine Matrix $L$ mit $L \cdot A = E_n$. Dann gilt $1 = \det(L)\det(A)$, also $\det(A) \neq 0$. (2)⇒(1): Wir gehen so vor: Wenn $z_i$ in der linearen Hülle von $z_1, \ldots, z_{i-1}$ liegt, so kann man Satz 12.1 (1) verwenden, um eine Matrix mit gleicher Determinante und $i$-ter Zeile $= 0$ zu erzeugen. Wegen der Eigenschaft (D1) ist diese Determinante $= 0$. ■

## 4. Berechnen der Determinante in Körpern

Es ist besonders leicht, die Determinante einer Matrix in Zeilenstaffelform zu berechnen:

SATZ 12.14. *Sei $K$ ein Körper, sei $m \in \mathbb{N}$, und sei $A$ eine $m \times m$-Matrix über $K$, sodass für alle $i$, $j$ mit $i > j$ gilt: $A(i, j) = 0$. Dann gilt*

$$\det(A) = \prod_{i=1}^{m} A(i, i).$$

*Beweisskizze:* Sei $f \in S_m$, $f \neq \mathrm{id}$. Wenn für alle $i$ die Ungleichung $i \leq f(i)$ gilt, so gilt $f(m) = m, \ldots, f(1) = 1$, also $f = \mathrm{id}$. Es gibt also $i$ mit $i > f(i)$, und somit ist $A(i, f(i)) = 0$. ■

Wir wissen, dass sich jede Matrix durch Zeilenumformungen in Zeilenstaffelnormalform bringen lässt. Dabei ändert sich die Determinante folgendermaßen:

(1) Wenn wir ein Vielfaches einer Zeile zu einer anderen dazu addieren, bleibt die Determinante unverändert.
(2) Wenn wir eine Zeile mit einem Körperelement vervielfachen, dann vervielfacht sich die Determinante um eben dieses Körperelement.
(3) Beim Vertauschen von zwei Zeilen ändert sich das Vorzeichen der Determinante.

Aus diesen Überlegungen erhält man einen Algorithmus zum Berechnen der Determinante. Wir rechnen einige Beispiele.

```
In[151]:= << RowRed10.m
```

```
In[152]:= DeterminantenDemo [B]
```

$$\mathrm{Det}\left[\begin{pmatrix} 1 & 2 & -3 \\ 2 & 29 & -16 \\ -3 & -16 & 22 \end{pmatrix}\right]$$

Wir addieren das − 2 fache
der 1. Zeile zum 1 fachen der 2. Zeile.

$$= 1 * 1 * det\left[\begin{pmatrix} 1 & 2 & -3 \\ 0 & 25 & -10 \\ -3 & -16 & 22 \end{pmatrix}\right]$$

Wir addieren das 3 fache
der 1. Zeile zum 1 fachen der 3. Zeile.

$$= 1 * 1 * det\left[\begin{pmatrix} 1 & 2 & -3 \\ 0 & 25 & -10 \\ 0 & -10 & 13 \end{pmatrix}\right]$$

Wir addieren das 2 fache
der 2. Zeile zum 5 fachen der 3. Zeile.

$$= 1 * \frac{1}{5} * det\left[\begin{pmatrix} 1 & 2 & -3 \\ 0 & 25 & -10 \\ 0 & 0 & 45 \end{pmatrix}\right]$$

$= 225$

*Out[152]=* 225

*In[153]:=* **DeterminantenDemo** [**A**]

$\text{Det}\left[\begin{pmatrix} 5 & 0 & 0 \\ 3 & 0 & -2 \\ 4 & 1 & 1 \end{pmatrix}\right]$

Wir addieren das − 3 fache
der 1. Zeile zum 5 fachen der 2. Zeile.

$$= 1 * \frac{1}{5} * det\left[\begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & -10 \\ 4 & 1 & 1 \end{pmatrix}\right]$$

Wir addieren das − 4 fache
der 1. Zeile zum 5 fachen der 3. Zeile.

$$= \frac{1}{5} * \frac{1}{5} * det\left[\begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & -10 \\ 0 & 5 & 5 \end{pmatrix}\right]$$

$$= -\frac{1}{25} det\left[\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 5 \\ 0 & 0 & -10 \end{pmatrix}\right]$$

$= 10$

*Out[153]=* 10

*In[154]:=* **DeterminantenDemo**[**B2**]

$\text{Det}\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 3 & 1 & 3 & 5 \\ 4 & 3 & 0 & 9 \\ 8 & 0 & 0 & 1 \end{pmatrix}\right]$

Wir addieren das − 3 fache
der 1. Zeile zum 1 fachen der 2. Zeile.

$$= 1 * 1 * det\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 12 & -7 \\ 4 & 3 & 0 & 9 \\ 8 & 0 & 0 & 1 \end{pmatrix}\right]$$

Wir addieren das − 4 fache
der 1. Zeile zum 1 fachen der 3. Zeile.

$$= 1 * 1 * det\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 12 & -7 \\ 0 & -5 & 12 & -7 \\ 8 & 0 & 0 & 1 \end{pmatrix}\right]$$

Wir addieren das − 8 fache
der 1. Zeile zum 1 fachen der 4. Zeile.

$$= 1 * 1 * det\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 12 & -7 \\ 0 & -5 & 12 & -7 \\ 0 & -16 & 24 & -31 \end{pmatrix}\right]$$

Wir addieren das − 1 fache
der 2. Zeile zum 1 fachen der 3. Zeile.

$$= 1 * 1 * det\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 12 & -7 \\ 0 & 0 & 0 & 0 \\ 0 & -16 & 24 & -31 \end{pmatrix}\right]$$

Wir addieren das − 16 fache
der 2. Zeile zum 5 fachen der 4. Zeile.

$$= 1 * \frac{1}{5} * det\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 12 & -7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -72 & -43 \end{pmatrix}\right]$$

$$= -\frac{1}{5}\, det\left[\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 12 & -7 \\ 0 & 0 & -72 & -43 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right]$$

$= 0$

*Out[154]=* 0

Schließlich berechnen wir noch die Determinante von $\begin{pmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}$

```
In[155]:= << RowRed10.m
In[156]:= M = {{0, 2, 1}, {0, 0, 1}, {1, 2, 3}}

Out[156]= {{0, 2, 1}, {0, 0, 1}, {1, 2, 3}}

In[157]:= DeterminantenDemo [M]
```

$$\text{Det}\left[\begin{pmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}\right]$$

$$= -1\, det\left[\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\right]$$

$$= 1\, det\left[\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\right]$$

$= 2$

*Out[157]=* 2

```
In[158]:= Det [M]
Out[158]= 2
```

## 5. Die adjungierte Matrix

DEFINITION 12.15. Sei $n \in \mathbb{N}$, sei $A$ eine $n \times n$-Matrix, und seien $i, j \in \{1, \ldots, n\}$. Dann bezeichnen wir mit $A^{[i,j]}$ die $n \times n$-Matrix, die durch

$$A^{[i,j]}(k,l) := \begin{cases} A(k,l) & \text{wenn } k \neq i \text{ und } l \neq j \\ 0 & \text{wenn } (k = i \text{ und } l \neq j) \text{ oder } (k \neq i \text{ und } l = j) \\ 1 & \text{wenn } k = i \text{ und } l = j \end{cases}$$

für $k, l \in \{1, \ldots, n\}$ definiert ist. $A^{[i,j]}$ ist also die Matrix, die man aus $A$ durch Ersetzen der $i$-ten Zeile durch den Einheitsvektor $e_j$ und durch Ersetzen der $j$-ten Spalte durch den Einheitsvektor $e_i$ erhält.

Für

$$A = \begin{pmatrix} -5 & 7 & 8 \\ 6 & 3 & -2 \\ 1 & -5 & 4 \end{pmatrix}$$

gilt also

$$A^{[3,2]} = \begin{pmatrix} -5 & 0 & 8 \\ 6 & 0 & -2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Für $n \geq 2$ beteichnet man mit $A^{(i,j)}$ die $(n-1) \times (n-1)$-Matrix, die man aus $A$ durch Streichen der $i$-ten Zeile und $j$-ten Spalte erhält. Es gilt also

$$A^{(i,j)}(k,l) := \begin{cases} A(k,l) & \text{wenn } k < i \text{ und } l < j \\ A(k+1,l) & \text{wenn } k \geq i \text{ und } l < j \\ A(k,l+1) & \text{wenn } k < i \text{ und } l \geq j \\ A(k+1,l+1) & \text{wenn } k \geq i \text{ und } l \geq j \end{cases}$$

für alle $k, l \in \{1, \ldots, n-1\}$.

Für $A = \begin{pmatrix} -5 & 7 & 8 \\ 6 & 3 & -2 \\ 1 & -5 & 4 \end{pmatrix}$ gilt also $A^{(3,2)} = \begin{pmatrix} -5 & 8 \\ 6 & -2 \end{pmatrix}$.

LEMMA 12.16. *Seien $B$ eine $n \times n$-Matrix, und seien $\sigma, \tau \in S_n$. Sei $C$ eine $n \times n$-Matrix, die durch $C(i,j) := B(\sigma(i), \tau(j))$ definiert ist. Dann gilt* $\det(C) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\,\det(B)$.

*Beweis:*

$$\begin{aligned}
\det(C) &= \sum_{f\in S_n} \operatorname{sgn}(f) \prod_{i=1}^{n} C(i, f(i)) \\
&= \sum_{f\in S_n} \operatorname{sgn}(f) \prod_{i=1}^{n} B(\sigma(i), \tau(f(i))) \\
&= \sum_{f\in S_n} \operatorname{sgn}(f) \prod_{k=1}^{n} B(k, \tau\circ f\circ\sigma^{-1}(k)) \\
&= \sum_{f\in S_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\operatorname{sgn}(\tau\circ f\circ\sigma^{-1}) \prod_{k=1}^{n} B(k, \tau\circ f\circ\sigma^{-1}(k)) \\
&= \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) \sum_{g\in S_n} \operatorname{sgn}(g) \prod_{k=1}^{n} B(k, g(k)) \\
&= \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)\ \det(B).
\end{aligned}$$

LEMMA 12.17. *Sei $n \geq 2$, sei $A$ eine $n\times n$-Matrix, und seien $i, j \in \{1, \dots, n\}$. Dann gilt* $\det(A^{[i,j]}) = (-1)^{i+j}\det(A^{(i,j)})$.

*Beweis:* Wir beweisen den Satz zunächst für $i = j = n$. Dann gilt

$$\begin{aligned}
\det(A^{[n,n]}) &= \sum_{f\in S_n} \operatorname{sgn}(f) \prod_{k=1}^{n} A^{[n,n]}(k, f(k)) \\
&= \sum_{\substack{f\in S_n \\ f(n)=n}} \operatorname{sgn}(f) \prod_{k=1}^{n-1} A(k, f(k)).
\end{aligned}$$

Sei $f$ eine Permutation von $\{1, \dots, n\}$ mit $f(n) = n$. Dann ist $g := f|_{\{1,\dots,n-1\}}$ eine Permutation von $\{1, \dots, n-1\}$. Da sich $f$ als Produkt gleich vieler Transpositionen wie $g$ schreiben lässt, ist die Signatur von $g$ gleich der Signatur von $f$. Es gilt also

$$\begin{aligned}
&\sum_{\substack{f\in S_n \\ f(n)=n}} \operatorname{sgn}(f) \prod_{k=1}^{n-1} A(k, f(k)) \\
&= \sum_{g\in S_{n-1}} \operatorname{sgn}(g) \prod_{k=1}^{n-1} A(k, f(k)) \\
&= \det(A^{(i,j)}).
\end{aligned}$$

Wir beweisen nun den Fall $(i, j) \neq (n, n)$. Sei $\sigma$ der Zyklus $(i \quad i+1 \quad \dots \quad n)$, und $\tau := (j \quad j+1 \quad \dots \quad n)$

$$B(k, l) := A(\sigma(k), \tau(l).)$$

Dann gilt $B^{(n,n)} = A^{(i,j)}$. Außderdem gilt für alle $k, l \in \{1, \dots, n\}$ auch $B^{[n,n]}(k, l) = A^{[i,j]}(\sigma(k), \tau(l))$. Wir berechnen nun $\det(A^{[i,j]})$. Nach Lemma 12.16 gilt $\det(A^{[i,j]}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\det(B^{[n,n]})$. Nach dem bereits betrachteten Fall gilt $\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\det(B^{[n,n]}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\det(B^{(n,n)})$. Wegen $B^{(n,n)} = A^{(i,j)}$ gilt nun insgesamt

$$\det(A^{[i,j]}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\det(A^{(i,j)}).$$

Nach Satz 12.6 (4) gilt $\operatorname{sgn}(\sigma) = (-1)^{n-i+1+1} = (-1)^{n-i}$ und $\operatorname{sgn}(\tau) = (-1)^{n-j}$. Also gilt $\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = (-1)^{i+j}$. ■

DEFINITION 12.18. Sei $n \in \mathbb{N}$, und sei $A$ eine $n \times n$-Matrix. Die zu $A$ adjungierte (oder adjunkte[1] ) Matrix $A^{\mathrm{ad}}$ ist wie folgt definiert: wenn $n \geq 2$, so gilt

$$A^{\mathrm{ad}}(i, j) := (-1)^{i+j}\det(A^{(j,i)}) \text{ für } i, j \in \{1, \dots, n\}.$$

Für $n = 1$ definiert man die adjungierte Matrix durch $A^{\mathrm{ad}} := (1)$.

Es gilt also $A^{\mathrm{ad}}(i, j) = \det(A^{[j,i]})$. Wir beobachten zunächst folgenden einfachen Satz:

SATZ 12.19. *Sei $A$ eine $n \times n$-Matrix. Dann gilt* $(A^{\mathrm{ad}})^T = (A^T)^{\mathrm{ad}}$.

*Beweis:* Für $i, j \in \{1, \dots, n\}$ gilt $(A^{\mathrm{ad}})^T(i, j) = A^{\mathrm{ad}}(j, i) = \det(A^{[i,j]}) = \det((A^{[i,j]})^T) = \det((A^T)^{[j,i]}) = (A^T)^{\mathrm{ad}}(i, j)$. ■

SATZ 12.20. *Sei $R$ ein kommutativer Ring mit Eins, und sei $A$ eine $n \times n$ Matrix mit Einträgen aus $R$. Dann gilt*

$$A \cdot A^{\mathrm{ad}} = A^{\mathrm{ad}} \cdot A = \begin{pmatrix} \det(A) & 0 & 0 & \dots & 0 \\ 0 & \det(A) & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \det(A) \end{pmatrix}.$$

*Beweis:* Wir werden als Abkürzung für die $n \times n$-Matrix

$$\begin{pmatrix} r & 0 & 0 & \dots & 0 \\ 0 & r & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & r \end{pmatrix}$$

mit $r \in R$ auch oft kürzer $r * E_n$ schreiben.

[1]Der Begriff *adjunkte Matrix* hat folgenden Vorteil: es gibt in der linearen Algebra auch den Begriff des *adjungierten Operators*, der aber nicht mit $A^{\mathrm{ad}}$ zu tun hat, sondern mit $A^T$. Der Begriff der zu $A$ *adjunkten Matrix* vermeidet diese Inkongruenz.

Wir nehmen an, dass $n \geq 2$. Für $j \in \{1, \ldots, n\}$ sei

$$e_j := (0, 0, \ldots, 0, 1, 0, \ldots, 0)$$

der $j$-te Einheitsvektor in $R^n$; der Einser steht dabei an der $j$-ten Stelle. Wir zeigen zunächst $A \cdot A^{\mathrm{ad}} = \det(A) * E_n$.

Sei nun $i \in \{1, \ldots, n\}$. Wir bilden nun Matrizen $A_1, \ldots, A_n$ wie folgt: $A_1$ ist die Matrix, die wir aus $A$ erhalten, wenn wir die $i$-te Zeile durch $e_1$ ersetzen; allgemein sei $A_j$ die Matrix, die wir erhalten, wenn wir die $i$-te Zeile durch $e_j$ ersetzen. Es gilt also

$$A_j = \begin{pmatrix} A(1,1) & \ldots & A(1,j-1) & A(1,j) & A(1,j+1) & \ldots & A(1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(i-1,1) & \ldots & A(i-1,j-1) & A(i-1,j) & A(i-1,j+1) & \ldots & A(i-1,n) \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ A(i+1,1) & \ldots & A(i+1,j-1) & A(i+1,j) & A(i+1,j+1) & \ldots & A(i+1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(n,1) & \ldots & A(n,j-1) & A(n,j) & A(n,j+1) & \ldots & A(n,n) \end{pmatrix}.$$

Aus Satz 12.1 (1) und Satz 12.11 erhalten wir, dass für alle $j \in \{1, \ldots, n\}$ gilt:

$$\det(A_j) = \det(A^{[i,j]}).$$

Sei nun $k \in \{1, \ldots, n\}$ Wir berechnen nun den $(k, i)$-ten Eintrag von $A \cdot A^{\mathrm{ad}}$. Wir erhalten

$$\begin{aligned} A \cdot A^{\mathrm{ad}}(k,i) &= \sum_{j=1}^{n} A(k,j) A^{\mathrm{ad}}(j,i) \\ &= \sum_{j=1}^{n} A(k,j) \det(A^{[i,j]}) \\ &= \sum_{j=1}^{n} A(k,j) \det(A_j) \end{aligned}$$

Wir nützen nun die Linearität der Determinante in der $i$-ten Zeile aus und erhalten, dass die letzte Summe gleich der Determinante der Matrix $B$ ist, die wir aus $A$ erhalten, indem wir die $i$-te Zeile von $A$ durch

$$(A(k,1), A(k,2), \ldots, A(k,n)),$$

also durch die $k$-te Zeile von $A$ ersetzen. Es gilt also

$$B = \begin{pmatrix} A(1,1) & \ldots & A(1,j-1) & A(1,j) & A(1,j+1) & \ldots & A(1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(i-1,1) & \ldots & A(i-1,j-1) & A(i-1,j) & A(i-1,j+1) & \ldots & A(i-1,n) \\ A(k,1) & \ldots & A(k,j-1) & A(k,j) & A(k,j+1) & \ldots & A(k,n) \\ A(i+1,1) & \ldots & A(i+1,j-1) & A(i+1,j) & A(i+1,j+1) & \ldots & A(i+1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(n,1) & \ldots & A(n,j-1) & A(n,j) & A(n,j+1) & \ldots & A(n,n) \end{pmatrix}.$$

Die Determinante dieser Matrix $B$ ist wegen Satz 12.11 gleich 0, wenn $k \neq i$, da dann in $B$ die $k$-te und die $i$-te Zeile gleich sind. Ist $k = i$, so gilt $\det(B) = \det(A)$.

Insgesamt haben wir also bewiesen, dass für jede $n \times n$-Matrix $A$ die Gleichheit $A \cdot A^{\mathrm{ad}} = \det(A) * E_n$ gilt. Dann gilt auch $A^{\mathrm{ad}} \cdot A = (A^T \cdot (A^{\mathrm{ad}})^T)^T = (A^T \cdot (A^T)^{\mathrm{ad}})^T = (\det(A^T) * E_n)^T = \det(A) * E_n$. ■

KOROLLAR 12.21 (Entwicklungssatz von Laplace). *Sei $n \geq 2$, sei $A$ eine $n \times n$-Matrix über einem kommutativen Ring mit Eins $K$, und seien $i, j \in \{1, \ldots, n\}$. Dann gilt*

(1) $\det(A) = \sum_{k=1}^{n} (-1)^{i+k} A(i,k) \cdot \det(A^{(i,k)})$. *(Entwicklung nach der $i$-ten Zeile.)*
(2) $\det(A) = \sum_{k=1}^{n} (-1)^{k+j} A(k,j) \cdot \det(A^{(k,j)})$. *(Entwicklung nach der $j$-ten Spalte.)*

*Beweis:* (1) Nach Satz 12.20 gilt

$$\begin{aligned} \det(A) &= (A \cdot A^{\mathrm{ad}})\ (i,i) \\ &= \sum_{k=1}^{n} A(i,k) A^{\mathrm{ad}}(k,i) \\ &= \sum_{k=1}^{n} A(i,k)(-1)^{i+k} \det(A^{(i,k)}). \end{aligned}$$

(2) Nach Satz 12.20 gilt

$$\begin{aligned} \det(A) &= (A^{\mathrm{ad}} \cdot A)\ (j,j) \\ &= \sum_{k=1}^{n} A^{\mathrm{ad}}(j,k) A(k,j) \\ &= \sum_{k=1}^{n} (-1)^{j+k} \det(A^{(k,j)}) A(k,j). \end{aligned}$$

## 6. Determinanten und Invertierbarkeit

SATZ 12.22. *Sei $R$ ein kommutativer Ring mit Eins, und sei $A$ eine $(n \times n)$-Matrix mit Einträgen aus $R$. Dann sind äquivalent:*

(1) *Es gibt eine Matrix $B$ mit $A \cdot B = E_n$.*
(2) *Es gibt eine Matrix $C$ mit $C \cdot A = E_n$.*
(3) *Es gibt ein $y \in R$, sodass* $\det(A)\ y = 1$.

*Beweis:* (1)⇒(3): wegen Satz 12.12 gilt $1 = \det(E_n) = \det(A) \det(B)$, also leistet $y := \det(B)$ das Gewünschte. (3)⇒(1): Sei $B := y * A^{\mathrm{ad}}$. Dann gilt $A \cdot B = y * (A \cdot A^{\mathrm{ad}})$, und das ist wegen Satz 12.20 gleich

$$y * \begin{pmatrix} \det(A) & 0 & 0 & \dots & 0 \\ 0 & \det(A) & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \det(A) \end{pmatrix},$$

also gleich $E_n$.

Die Äquivalenz von (1) und (2) zeigt man genauso. ■

Somit wissen wir, dss eine ganzzahlige Matrix genau dann eine ganzzahlige inverse Matrix hat, wenn ihre Determinante $+1$ oder $-1$ ist.

SATZ 12.23. *Sei $K$ ein Körper, und sei $A$ eine $(n \times n)$-Matrix mit Einträgen aus $K$. Dann sind äquivalent:*

(1) $\det(A) \neq 0$.
(2) *$A$ ist invertierbar.*
(3) *Die Zeilenvektoren von $A$ sind linear unabhängig.*
(4) *Die Spaltenvektoren von $A$ sind linear unabhängig.*

*Beweis:* (1)⇒(2): $B := \frac{1}{\det(A)} A^{\text{ad}}$ ist nach Satz 12.20 zu $A$ invers. (2)⇒(3): Sei $(y_1, \dots, y_n) \in K^n$ mit $(y_1, \dots, y_n) \cdot A = 0$. Dann gilt $0 = 0 \cdot A^{-1} = ((y_1, \dots, y_n) \cdot A) \cdot A^{-1} = (y_1, \dots, y_n)$, also sind die Zeilen linear unabhängig. (3)⇒(1): Da die Zeilenvektoren linear unabhängig sind, ist der Zeilenraum von $A$ ganz $K^n$. Somit gibt es eine Matrix $L$ mit $L \cdot A = E_n$. Dann gilt $1 = \det(L) \det(A)$, also $\det(A) \neq 0$.

Die Bedingungen (1), (2) und (3) sind also für jede Matrix $A$ äquivalent.

(1)⇒(4): Sei $A$ eine Matrix mit $\det(A) \neq 0$. Dann gilt $\det(A^T) \neq 0$, also sind die Zeilen von $A^T$ nach einer der bereits bewiesenen Implikationen linear unabhängig. Somit gilt (4). (4)⇒(3): Wenn die Spalten von $A$ linear unabhängig sind, so sind es auch die Zeilen von $A^T$. Also gilt $\det(A^T) \neq 0$, und somit $\det(A) \neq 0$. ■

## 7. Determinanten und Volumina

Wir betrachten in dieser Sektion Matrizen mit Einträgen aus $\mathbb{R}$. Wir wissen (wegen $\det(A) = \det(A^T)$), dass $|\det(A)|$ das Volumen des von den Spalten von $A$ aufgespannten Parallelepipeds ist.

SATZ 12.24. *Seien $A, B \in \mathbb{R}^{n \times n}$, und seien $b_1, \dots, b_n$ die Spaltenvektoren von $B$. Sei $h_A : \mathbb{R}^n \to R^n$ gegeben durch $h_A(y) = A \cdot y$ für $y \in R^n$. Das Volumen des von $h_A(b_1), \dots, h_A(b_n)$ aufgespannten Parallelepipeds ist $|\det(A) \det(B)|$.*

*Beweis:* Das Parallelepiped wird von den Spaltenvektoren von $A \cdot B$ aufgespannt und hat also das Volumen $|\det(A \cdot B)|$. Aus der Multiplikativität der Determinante folgt die Aussage.

SATZ 12.25. *Seien $m, n \in \mathbb{R}$ mit $m \geq n$, seien $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times n}$, und seien $b_1, \ldots, b_n$ die Spaltenvektoren von $B$. Sei $h_A : \mathbb{R}^n \to R^m$ gegeben durch $h_A(y) = A \cdot y$ für $y \in R^n$. Das $n$-dimensionale Volumen des von $h_A(b_1), \ldots, h_A(b_n)$ aufgespannten Parallelepipeds ist* $\sqrt{\det(A^T \cdot A)} \cdot |\det(B)|$.

*Beweis:* Sei $Q$ eine $m \times n$-Matrix, in deren Spalten eine Orthonormalbasis des Spaltenraums von $A$ steht. Es gilt dann $Q^T \cdot Q = E_n$. Dann stehen in den Spalten von $Q^T \cdot (A \cdot B)$ die Koordinaten des von $h_A(b_1), \ldots, h_A(b_n)$ aufgespannten Parallelepipeds bezüglich dieser Orthonormalbasis. Sei nun $R \in \mathbb{R}^{n\times n}$ mit $A = QR$. Dann gilt für das gesuchte Volumen $V$:

$$\begin{aligned} V^2 &= \det(Q^T AB)^2 \\ &= \det(B^T A^T Q Q^T AB) \\ &= \det(B^T R^T Q^T Q Q^T QRB) \\ &= \det(B^T) \det(B) \det(R^T Q^T QR) \\ &= \det(B)^2 \det(A^T \cdot A). \end{aligned}$$

■

Für $B := E_n$ erhalten wir als Folgerung, dass das $n$-dimensionale Volumen des von den Spalten von $A$ aufgespannten Parallelepipeds gleich $\sqrt{\det(A^T \cdot A)}$ ist.

ÜBUNGSAUFGABEN 12.26.

(1) Sei $r > s$. Zeigen Sie, dass für jede $r \times s$-Matrix $A$ gilt, dass $\det(A^T \cdot A) \geq 0$. *Hinweis:* Schreiben Sie $A$ als $Q \cdot R$, wobei in den Spalten von $Q$ eine Orthonormalbasis für den Spaltenraum von $A$ steht, und berechnen Sie $\det(Q^T \cdot A)^2$.

## 8. Gleichungssysteme

SATZ 12.27. *Sei $K$ ein kommutativer Ring mit Eins, sei $n \in \mathbb{N}$, sei $A \in K^{n\times n}$ und sei $b \in K^n$. Wir nehmen an, dass $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ eine Lösung des Gleichungssystems $A \cdot x = b$ ist.*

*Sei $i \in \{1, \ldots, n\}$, und sei $A_i^b$ die Matrix, die man aus $A$ erhält, indem man die $i$-te Spalte von $A$ durch $b$ ersetzt. Dann gilt* $\det(A)\, y_i = \det(A_i^b)$.

*Beweis:* Sei $X$ die Matrix, die man aus der Einheitsmatrix $E_n$ erhält, indem man die $i$-te Spalte durch $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ ersetzt. Dann gilt

$$A \cdot X = A_i^b.$$

Also gilt $\det(A)\ \det(X) = \det(A_i^b)$. Durch Entwicklung nach der $i$-ten Zeile erhält man $\det(X) = y_i$. ■

SATZ 12.28. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, $A \in K^{n\times n}$, $b \in K^n$. Äquivalent sind:*

(1) *Das Gleichungssystem $A \cdot x = b$ hat genau eine Lösung in $K^n$.*
(2) $\det(A) \neq 0$.

*Beweis:* Wenn $\det(A) \neq 0$, so ist $A$ nach Satz 12.23 invertierbar, und somit ist $x := A^{-1} \cdot b$ die eindeutige Lösung von $A \cdot x = b$.

Wenn $\det(A) = 0$, so sind die Spaltenvektoren von $A$ linear abhängig; der Nullraum von $A$ ist also nicht nulldimensional. Die Lösungsmenge von $A \cdot x = b$ ist also leer oder von der Form $x_0 + N(A)$, und somit in beiden Fällen nicht einelementig. ■

SATZ 12.29 (Cramersche Regel). *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, $A \in K^{n\times n}$, $b \in K^n$. Wir nehmen an, dass* $\det(A) \neq 0$. *Für jedes $i \in \mathbb{N}$ sei $A_i^b$ die Matrix, die man aus $A$ dadurch erhält, dass man die $i$-te Spalte durch $b$ ersetzt, und sei $y_i := \frac{\det(A_i^b)}{\det(A)}$.*

*Dann ist $(y_1, \ldots, y_n)$ die eindeutige Lösung von $A \cdot x = b$.*

*Beweis:* Wegen Satz 12.28 hat $A \cdot x = b$ genau eine Lösung, und wegen Satz 12.27 berechnet sich diese Lösung durch $y_i := \frac{\det(A_i^b)}{\det(A)}$. ■

KAPITEL 13

# Polynome

## 1. Primfaktorzerlegung in den ganzen Zahlen

DEFINITION 13.1 (Primzahl). Eine Zahl $p \in \mathbb{N}$ ist genau dann eine *Primzahl*, wenn folgende beiden Bedingungen gelten:

(1) Es gilt $p > 1$.
(2) Für alle $a, b \in \mathbb{N}$ mit $p = a \cdot b$ gilt $a = 1$ oder $b = 1$.

DEFINITION 13.2 (Teilbarkeit). Seien $x, y \in \mathbb{Z}$. Die Zahl $x$ *teilt* $y$ genau dann, wenn es ein $z \in \mathbb{Z}$ gibt, sodass $y = z \cdot x$ ist.

Wir schreiben dann auch $x \mid y$; die Zahl $y$ heißt ein *Vielfaches* von $x$.

SATZ 13.3. *Seien $a \in \mathbb{Z}$ und $n \in \mathbb{N}$. Dann gibt es genau ein Paar von Zahlen $(q, r) \in \mathbb{Z} \times \mathbb{Z}$, sodass $a = q \cdot n + r$ und $r \in \{0, \ldots, n-1\}$.*

Wir bezeichnen den Rest $r$ mit $a \bmod n$.

SATZ 13.4 (Euklid, 360-280 v.Chr.). *Es gibt unendlich viele Primzahlen.*

*Beweis:* Wir nehmen an, dass $p_1, \ldots, p_r$ bereits alle Primzahlen sind. Da 2 prim ist, gilt $r \geq 1$. Dann ist der kleinste Teiler $q$ von $1 + \prod_{i=1}^{r} p_i$ mit $q > 1$ eine Primzahl mit $q \notin \{p_1, \ldots, p_r\}$. ■

DEFINITION 13.5 (Größter gemeinsamer Teiler). Für zwei Zahlen $a, b \in \mathbb{Z}$ (nicht beide 0) ist $\operatorname{ggT}(a, b)$ die größte Zahl $z \in \mathbb{N}$ mit $z \mid a$ und $z \mid b$.

SATZ 13.6. *Seien $a, b \in \mathbb{Z}$ nicht beide 0, und sei $z \in \mathbb{Z}$. Dann gilt:* $\operatorname{ggT}(a, b) = \operatorname{ggT}(a + z \cdot b, b)$.

So gilt zum Beispiel $\operatorname{ggT}(25, 15) = \operatorname{ggT}(40, 15)$.

*Beweis:* Wir zeigen, dass nicht nur der ggT, sondern sogar die Mengen der gemeinsamen Teiler der beiden Zahlenpaare gleich sind. Wir zeigen also

$$\{t \in \mathbb{Z} \;:\; t \mid a \text{ und } t \mid b\} = \{t \in \mathbb{Z} \;:\; t \mid a + zb \text{ und } t \mid b\}.$$

"$\subseteq$": Falls $t$ sowohl $a$ als auch $b$ teilt, dann auch $a + zb$ und $b$. "$\supseteq$": Falls $t$ sowohl $a + zb$, als auch $b$ teilt, dann auch $a + zb - zb$ und $b$, also auch $a$ und $b$. ■

Das nützen wir jetzt möglichst geschickt aus, um ggT(147, 33) zu berechnen:

$$\begin{aligned}
\text{ggT}(147, 33) &= \text{ggT}(147 - 4 \cdot 33, 33)\\
&= \text{ggT}(15, 33)\\
&= \text{ggT}(15, 33 - 2 \cdot 15)\\
&= \text{ggT}(15, 3)\\
&= \text{ggT}(0, 3)\\
&= 3.
\end{aligned}$$

Günstig ist es also, $z$ so zu wählen, dass $a + zb$ der Rest von $a$ bei der Division durch $b$ wird.

Mit Hilfe des *erweiterten Euklidischen Algorithmus* findet man nicht nur den ggT von $a$ und $b$, sondern auch $u, v \in \mathbb{Z}$, sodass gilt:

$$ggT\ (a, b) = u \cdot a + v \cdot b.$$

**Beispiel:** Wir berechnen ggT(147, 33), und schreiben das so:

| | 147 | 33 | |
|---|---|---|---|
| 147 | 1 | 0 | $(147 = 1 \cdot 147 + 0 \cdot 33)$ |
| 33 | 0 | 1 | $(33 = 0 \cdot 147 + 1 \cdot 33)$ |
| 15 | 1 | -4 | $(15 = 1 \cdot 147 - 4 \cdot 33)$ |
| 3 | -2 | 9 | $(3 = -2 \cdot 147 + 9 \cdot 33)$ |
| 0 | | | |

Berechnet man ggT$(a, b)$ mithilfe dieses Algorithmus, sieht man, dass sich die Zahlen in der linken Spalte immer als Linearkombination von $a$ und $b$ schreiben lassen. Als Konsequenz davon erhalten wir folgenden Satz:

SATZ 13.7. *Seien $a, b \in \mathbb{Z}$ (nicht beide $0$). Dann gibt es $u, v \in \mathbb{Z}$, sodass*

$$\text{ggT}\,(a, b) = u \cdot a + v \cdot b.$$

*Beweis:* Wir betrachten zuerst den Fall $a \geq 0$, $b \geq 0$, und zeigen den Satz durch Induktion nach $\min(a, b)$. Wenn $b = 0$, so gilt $a > 0$, und somit nach der Definition des ggT auch ggT$(a, b) = a$. Wenn $a = 0$, so gilt $b \neq 0$ und ggT$(a, b) = b$.

Seien nun $a > 0$, $b > 0$, $b \leq a$. Durch Division mit Rest erhalten wir $q \in \mathbb{N}_0$, $r \in \{0, \ldots, b - 1\}$ sodass $a = qb + r$. Wegen Satz 13.6 gilt ggT$(a, b)$ = ggT$(r, b)$. Da $r < b$, gibt es nach Induktionsvoraussetzung $u', v' \in \mathbb{Z}$, sodass ggT$(r, b) = u'r + v'b$. Dann gilt ggT$(a, b)$ = ggT$(r, b) = u'r + v'b = u'(a - qb) + v'b = u'a + (v' - u'q)b$, also ist auch ggT$(a, b)$ als Kombination von $a$ und $b$ darstellbar. Der Fall $a > 0$, $b > 0$, $a \leq b$ funktioniert genauso. ■

Eine Folgerung davon ist:

SATZ 13.8. *Seien $a, b \in \mathbb{Z}$, nicht beide* $0$, *und sei $t \in \mathbb{Z}$ so, dass $t \mid a$ und $t \mid b$. Dann gilt auch $t \mid$* $\mathrm{ggT}(a, b)$.

*Beweis*: Seien $u, v \in \mathbb{Z}$ so, dass $\mathrm{ggT}(a, b) = ua + vb$. Da $t$ die Zahl $a$ teilt, ist auch $ua$ ein Vielfaches von $t$. Ebenso ist $vb$ ein Vielfaches von $t$. Somit ist auch die Summe $ua + vb$ ein Vielfaches von $t$. Die Zahl $t$ ist also ein Teiler von $\mathrm{ggT}(a, b)$. ■

Wenn $a$ und $b$ größten gemeinsamen Teiler 1 haben, so heißen sie *teilerfremd* oder *relativ prim*.

ÜBUNGSAUFGABEN 13.9.

(1) [**Remmert and Ullrich, 1987**, p. 28] Sei $p_n$ die $n$-te Primzahl, d. h. $p_1 = 2$, $p_2 = 3$, usw. Zeigen Sie
$$p_n \leq 2^{(2^{n-1})}.$$
(2) Seien $a, b, x \in \mathbb{N}$ und $u, v \in \mathbb{Z}$ so, dass
$$x = ua + vb.$$
Zeigen Sie: Wenn $x$ sowohl $a$ als auch $b$ teilt, so gilt $x = \mathrm{ggT}(a, b)$.
(3) Seien $a, b \in \mathbb{N}$, $y \in \mathbb{Z}$ so, dass $a \mid y$, $b \mid y$, $\mathrm{ggT}(a, b) = 1$. Zeigen Sie (ohne Vorgriff auf die Primfaktorzerlegung): $a \cdot b \mid y$.
(4) Seien $a, b \in \mathbb{Z}$ (nicht beide 0), und sei $k \in \mathbb{N}$. Zeigen Sie: $\mathrm{ggT}(ka, kb) = k\,\mathrm{ggT}(a, b)$. Gelingt es Ihnen, $\mathrm{ggT}(ka, kb) \mid k\,\mathrm{ggT}(a, b)$ auch ohne Verwendung der Primfaktorzerlegung zu zeigen?
(5) Seien $a, c \in \mathbb{Z}$, $b, d \in \mathbb{N}$. Zeigen Sie: Wenn die Brüche $\frac{a}{b}$ und $\frac{c}{d}$ gekürzt, und die Nenner $b$ und $d$ teilerfremd sind, so ist auch der Bruch $\frac{ad+bc}{bd}$ gekürzt.

SATZ 13.10. *Seien $a, b, c \in \mathbb{Z}$, und sei zumindest eine der Zahlen $a$ und $b$ nicht* $0$. *Wir nehmen an, dass $a$ die Zahl $b \cdot c$ teilt und dass* $\mathrm{ggT}(a, b) = 1$ *gilt. Dann gilt: $a$ teilt $c$.*

*Beweis:* Es gibt $u, v \in \mathbb{Z}$, sodass $1 = u \cdot a + v \cdot b$. Es gilt $a \mid uac$. Da nach Voraussetzung $a \mid bc$ gilt, gilt auch $a \mid vbc$. Daraus erhalten wir
$$a \mid (ua + vb)c,$$
und somit $a \mid c$. ■

KOROLLAR 13.11. *Seien $b, c \in \mathbb{Z}$, und sei $p$ eine Primzahl. Wenn $p$ das Produkt $b\,c$ teilt, so teilt $p$ einen der beiden Faktoren $b$ und $c$.*

SATZ 13.12. *Sei* $\langle p_i \mid i \in \mathbb{N}\rangle = (2, 3, 5, 7, 11, \ldots)$ *die Folge aller Primzahlen, und sei $n \in \mathbb{N}$. Dann gibt es genau eine Funktion $\alpha : \mathbb{N} \to \mathbb{N}_0$ mit folgenden Eigenschaften:*

(1) $\{i \in \mathbb{N} \mid \alpha(i) > 0\}$ *ist endlich.*
(2) $n = \prod_{i \in \mathbb{N}} p_i^{\alpha(i)}$.

*Beweis:* Wir zeigen zunächst durch Induktion nach $n$, dass es ein solches $\alpha$ gibt. Für $n = 1$ setzen wir $\alpha(i) := 0$ für alle $i \in \mathbb{N}$. Für $n > 1$ sei $q$ der kleinste Teiler von $n$ mit $q > 1$. Die Zahl $q$ ist eine Primzahl; es gibt also $j \in \mathbb{N}$ mit $q = p_j$. Nach

Induktionsvoraussetzung gibt es $\beta : \mathbb{N} \to \mathbb{N}_0$ mit
$$\frac{n}{q} = \prod_{i\in\mathbb{N}} p_i^{\beta(i)},$$
also gilt $n = p_j^{\beta(j)+1} \cdot \prod_{i\in\mathbb{N}\setminus\{j\}} p_i^{\beta(i)}$.

Nun zeigen wir die Eindeutigkeit. Seien $\alpha, \beta : \mathbb{N} \to \mathbb{N}_0$ so, dass $\{i \in \mathbb{N} \mid \alpha(i) > 0\}$ und $\{i \in \mathbb{N} \mid \beta(i) > 0\}$ beide endlich sind und
$$\prod_{i\in\mathbb{N}} p_i^{\alpha(i)} = \prod_{i\in\mathbb{N}} p_i^{\beta(i)}.$$
Wir zeigen, dass für alle $j \in \mathbb{N}$ gilt: $\alpha(j) = \beta(j)$. Sei dazu $j \in \mathbb{N}$. Wir nehmen an $\alpha(j) > \beta(j)$. Dann gilt
$$p_j^{\alpha(j)-\beta(j)} \prod_{i\in\mathbb{N}\setminus\{j\}} p_i^{\alpha(i)} = \prod_{i\in\mathbb{N}\setminus\{j\}} p_i^{\beta(i)}.$$
Nach Korollar 13.11 teilt $p_j$ also ein $p_i^{\beta(i)}$ mit $i \neq j$. Im Fall $\beta(i) = 0$ widerspricht das $p_j > 1$, im Fall $\beta(i) > 0$ gilt $p_j \mid p_i$. Da $p_i$ eine Primzahl ist, gilt dann $p_i = p_j$, im Widerspruch zu $i \neq j$. ■

ÜBUNGSAUFGABEN 13.13.

(1) Sei $p_n$ die $n$-te Primzahl, d. h. $p_1 = 2$, $p_2 = 3$, usw. Zeigen Sie, auch, ohne die Eindeutigkeit der Primfaktorzerlegung zu verwenden, dass Folgendes gilt: Wenn
$$\begin{aligned} a &= \prod p_i^{\alpha_i} \\ b &= \prod p_i^{\beta_i}, \end{aligned}$$
wobei $\alpha_i, \beta_i \in \mathbb{N}_0$, und fast alle $\alpha_i, \beta_i = 0$ sind, dann gilt $a \mid b$ genau dann, wenn für alle $i$ gilt: $\alpha_i \leq \beta_i$. (Zeigen Sie, dass diese Aussage für alle Primfaktorzerlegungen von $a$ und $b$ gilt. Folgt daraus die Eindeutigkeit der Primfaktorzerlegung?)

(2) Sei $p_n$ die $n$-te Primzahl, d. h. $p_1 = 2$, $p_2 = 3$, usw. Zeigen Sie: Wenn
$$\begin{aligned} a &= \prod p_i^{\alpha_i} \\ b &= \prod p_i^{\beta_i}, \end{aligned}$$
wobei $\alpha_i, \beta_i \in \mathbb{N}_0$, und fast alle $\alpha_i, \beta_i = 0$ sind, dann gilt
$$\mathrm{ggT}(a, b) = \prod p_i^{\min(\alpha_i, \beta_i)}.$$

(3) Welche Zahlen $q \in \mathbb{N}$ erfüllen folgende Eigenschaft?
Für alle $a, b \in \mathbb{Z}$ mit $q \mid a \cdot b$ gilt $q \mid a$ oder es gibt ein $n \in \mathbb{N}$, sodass $q \mid b^n$.

## 2. Polynome

DEFINITIONSVERSUCH 13.14. Sei $K$ kommutativer Ring mit Eins. Dann ist $K[x]$ die Menge aller Ausdrücke
$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n$$
mit $n \in \mathbb{N}_0$ und $a_0, a_1, \ldots, a_n \in K$. Die Elemente von $K[x]$ nennen wir *Polynome*.

Was heißt aber *Ausdruck*? Und welche Rolle spielt $x$? Mit folgender Definition stehen wir auf dem sicheren Boden der Mengenlehre.

DEFINITION 13.15. Sei $\boldsymbol{K}$ kommutativer Ring. Dann ist ein *Polynom über $K$* eine Folge $(a_0, a_1, a_2, a_3, \dots)$, sodass es ein $i \in \mathbb{N}_0$ gibt, sodass für alle $j \in \mathbb{N}_0$ mit $j \geq i$ gilt: $a_j = 0$.

Wir haben also Polynome als unendliche Liste ihrer Koeffizienten definiert.

DEFINITION 13.16. Sei $\boldsymbol{K}$ ein kommutativer Ring, und seien $(a_0, a_1, a_2, \dots)$, $(b_0, b_1, b_2, \dots)$ Polynome über $K$. Wir definieren

(1) $(a_0, a_1, a_2, \dots) + (b_0, b_1, b_2, \dots) := (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots)$
(2) $(a_0, a_1, a_2, \dots) \cdot (b_0, b_1, b_2, \dots) := (c_0, c_1, c_2, \dots)$ mit

$$c_k := \sum_{\substack{(i,j)\in\{0,\dots,k\}\times\{0,\dots,k\} \\ i+j=k}} a_i \cdot b_j$$

für alle $k \in \mathbb{N}_0$.

SATZ 13.17. *Sei $\boldsymbol{K}$ ein kommutativer Ring mit Eins, seien $p, q, r$ Polynome über $K$, und sei $e$ das Polynom* $(1, 0, \dots, 0)$. *Dann gilt:*

(1) $(p \cdot q) \cdot r = p \cdot (q \cdot r)$.
(2) $p \cdot q = q \cdot p$.
(3) $p \cdot (q + r) = p \cdot q + p \cdot r$.
(4) $p \cdot e = p$.

*Beweis von* (1)*:* Sei $p = (p_0, p_1, \dots)$, $q = (q_0, q_1, \dots)$, $r = (r_0, r_1, \dots)$, und sei $a = (a_0, a_1, \dots) := (p \cdot q) \cdot r$. Dann gilt für $k \in \mathbb{N}$

$$\begin{aligned}
a_k &= \sum_{\substack{(i,j)\in\{0,\dots,k\}^2 \\ i+j=k}} (p \cdot q)_i \, r_j \\
&= \sum_{\substack{(i,j)\in\{0,\dots,k\}^2 \\ i+j=k}} \; \sum_{\substack{(l,m)\in\{0,\dots,i\}^2 \\ l+m=i}} p_l q_m r_j \qquad (13.1) \\
&= \sum_{\substack{(l,m,j)\in\{0,\dots,k\}^3 \\ l+m+j=k}} p_l q_m r_j.
\end{aligned}$$

Die letzte Gleichheit erklären wir so: seien $(l, m, j) \in \{0, \dots, k\}^3$ so, dass $l + m + j = k$. Dann kommt der Summand $p_l q_m r_j$ der letzten Summe in der vorletzten Summe genau mit der Wahl $i' := l + m$, $j' := j$, $l' := l$, $m' := m$ vor. Seien nun $i, j, l, m$ so, dass $i + j = k$ und $l + m = i$. Dann gilt $l + m + j = i + j = k$, und somit kommt der Summand

$p_l q_m r_j$ in der letzten Summe mit den Indizes $l' := l$, $m' := m$, $j' := j$ vor. Genauso sieht man, dass $(p \cdot (q \cdot r))_k = \sum_{\substack{(l,m,j) \in \{0,\dots,k\}^3 \\ l+m+j=k}} p_l q_m r_j$. Somit gilt insgesamt $(p \cdot q) \cdot r = p \cdot (q \cdot r)$.

Die Argumentation über Gleichheit der Summen in (13.1) ist etwas unbefriedigend. Daher kommt jetzt eine leichter verifizierbare Argumentation. Wir verwenden das "Kronecker-Symbol" $\delta(a, b)$ mit $\delta(a, b) = 1$, wenn $a = b$, und $\delta(a, b) = 0$, wenn $a \neq b$. Dann rechnen wir

$$\begin{aligned}
((p \cdot q) \cdot r)_k &= \sum_{i=0}^{k} \sum_{j=0}^{k} \delta(i + j, k)(p \cdot q)_i r_j \\
&= \sum_{i=0}^{k} \sum_{j=0}^{k} \delta(i + j, k) \sum_{l=0}^{i} \sum_{m=0}^{i} \delta(l + m, i) p_l q_m r_j \\
&= \sum_{i=0}^{k} \sum_{j=0}^{k} \delta(i + j, k) \sum_{l=0}^{k} \sum_{m=0}^{k} \delta(l + m, i) p_l q_m r_j \\
&= \sum_{j=0}^{k} \sum_{l=0}^{k} \sum_{m=0}^{k} p_l q_m r_j \cdot (\sum_{i=0}^{k} \delta(i + j, k)\delta(l + m, i)).
\end{aligned}$$

Wir betrachten nun die Summe $\sum_{i=0}^{k} \delta(i + j, k)\delta(l + m, i)$. Damit der $i$-te Summand ungleich 0 ist, muss $i = l + m$ und $i + j = k$ gelten. Wenn $l + m + j \neq k$, so sind diese beiden Gleichungen für kein $i$ erfüllt. Wenn $l + m + j = k$, so erfüllt genau $i := l + m$ diese Gleichung. Die Summe ist also 0, wenn $l + m + j \neq k$, und 1, wenn $l + m + j = k$. Somit gilt

$$\sum_{j=0}^{k} \sum_{l=0}^{k} \sum_{m=0}^{k} p_l q_m r_j \cdot (\sum_{i=0}^{k} \delta(i + j, k)\delta(l + m, i)) = \sum_{j=0}^{k} \sum_{l=0}^{k} \sum_{m=0}^{k} p_l q_m r_j \delta(l + m + j, k).$$

Wieder formen wir $(p \cdot (q \cdot r))_k$ auf den gleichen Ausdruck um, und sehen so, dass $(p \cdot q) \cdot r = p \cdot (q \cdot r)$. ■

SATZ 13.18. *Sei* $\boldsymbol{K}$ *ein kommutativer Ring mit Eins, und sei P die Menge aller Polynome über K. Dann ist* $\langle P, +, -, \cdot, (0, 0, 0, \dots), (1, 0, 0, \dots)\rangle$ *ein kommutativer Ring mit Eins.*

SATZ 13.19. *Sei* $\boldsymbol{K}$ *ein kommutativer Ring mit Eins, sei* $p = (a_0, a_1, a_2, \dots)$ *ein Polynom über K, und sei* $x := (0, 1, 0, 0, \dots, 0)$. *Dann gilt:*

(1) $x^i = (\underbrace{0, 0, \dots, 0}_{i \text{ Nuller}}, 1, 0, 0, \dots)$.

(2) $p = \sum_{\substack{i \in \mathbb{N}_0 \\ a_i \neq 0}} a_i * x^i$.

Wir werden die Menge der Polynome über $K$ nun oft mit $K[x]$ oder $K[t]$ bezeichnen. Sobald wir $K[t]$ verwenden, haben wir die Bedeutung von 2 Variablen erklärt:

(1) $K[t] = \{(a_0, a_1, \dots) \in K^{\mathbb{N}_0} \mid \{i \in \mathbb{N}_0 \mid a_i \neq 0\} \text{ ist endlich}\}$.
(2) $t = (0, 1, 0, 0, \dots)$.

## 3. Polynomfunktionen

DEFINITION 13.20. Sei $\boldsymbol{K}$ ein kommutativer Ring mit Eins, und sei $p = \sum_{i=0}^{n} a_i * x^i$ ein Element von $K[x]$. Dann bezeichen wir *die von $p$ induzierte Funktion* mit $p^{\boldsymbol{K}}$ und definieren sie durch

$$\begin{array}{cccc} p^{\boldsymbol{K}} & : & K & \longrightarrow & K \\ & & y & \longmapsto & \sum_{i=0}^{n} a_i y^i. \end{array}$$

SATZ 13.21. *Sei $\boldsymbol{K}$ ein kommutativer Ring mit Eins, und seien $p, q \in K[x]$. Dann gilt für alle $y \in K$: $(p \cdot q)^{\boldsymbol{K}}(y) = p^{\boldsymbol{K}}(y) q^{\boldsymbol{K}}(y)$.*

*Beweis:* Sei $y \in K$. Im Fall $p = 0$ oder $q = 0$ ergeben beide Seiten 0, also nehmen wir an, dass $p \neq 0$ und $q \neq 0$. Sei $l := \deg(p)$ und $m := \deg(q)$. Dann gilt

$$\begin{aligned}
(p \cdot q)^{\boldsymbol{K}}(y) &= \sum_{k=0}^{l+m} (\sum_{i=0}^{l} \sum_{j=0}^{m} \delta(i+j,k) p_i q_j) y^k \\
&= \sum_{k=0}^{l+m} \sum_{i=0}^{l} \sum_{j=0}^{m} \delta(i+j,k) p_i q_j y^i y^j \\
&= \sum_{i=0}^{l} \sum_{j=0}^{m} \sum_{k=0}^{l+m} \delta(i+j,k) p_i q_j y^i y^j \\
&= \sum_{i=0}^{l} \sum_{j=0}^{m} p_i q_j y^i y^j \cdot (\sum_{k=0}^{l+m} \delta(i+j,k)) \\
&= \sum_{i=0}^{l} \sum_{j=0}^{m} p_i q_j y^i y^j \cdot 1 \\
&= (\sum_{i=0}^{l} p_i y^i) \cdot (\sum_{j=0}^{m} q_j y^j) \\
&= p^{\boldsymbol{K}}(y) \cdot q^{\boldsymbol{K}}(y).
\end{aligned}$$

■

## 4. Teilbarkeit von Polynomen

DEFINITION 13.22 (Grad eines Polynoms). Für $f := (a_0, a_1, a_2, \dots) \in K[x] \setminus \{0\}$ ist der *Grad* von $f$, $\deg f$, jenes $n \in \mathbb{N}_0$, sodass $a_n \neq 0$ und $a_i = 0$ für alle $i > n$. Dann nennen wir $a_n$ den *führenden Koeffizienten* von $f$. Wir definieren $\deg 0 := -1$.

DEFINITION 13.23. Sei $\boldsymbol{K}$ ein Körper, und seien $f, g \in K[x]$.

(1) $f$ *teilt* $g$, wenn es ein $q \in K[x]$ gibt, sodass $g = q \cdot f$.
(2) $f$ ist *irreduzibel über* $\boldsymbol{K}$ (ein irreduzibles Polynom in $K[x]$), wenn $\deg f \geq 1$ und für alle $a, b \in K[x]$ mit $a \cdot b = f$ entweder $a$ oder $b$ Grad 0 hat.
(3) $f$ ist *normiert*, wenn es führenden Koeffizienten 1 hat.

SATZ 13.24 (Division). *Sei* $\boldsymbol{K}$ *ein Körper, und seien* $f, g \in K[x]$. *Wenn* $f \neq 0$, *so gibt es genau ein Paar* $(q, r) \in K[x] \times K[x]$ *mit* $g = q \cdot f + r$ *und* $\deg r < \deg f$.

DEFINITION 13.25 (ggT in $\boldsymbol{K}[x]$). Sei $\boldsymbol{K}$ ein Körper, und seien $f, g \in K[x]$, nicht beide 0. Dann ist $d \in K[x]$ ein *größter gemeinsamer Teiler* von $f$ und $g$, wenn folgende Bedingungen gelten:

(1) $d \mid f$ und $d \mid g$,
(2) Für alle $h \in K[x]$ mit $h \mid f$ und $h \mid g$ gilt $\deg(h) \leq \deg(d)$,
(3) $d$ ist normiert.

Wir bezeichnen den Rest von $g$ bei der Division durch $f$ mit $(g \bmod f)$. Da das Paar $(g, f)$ die gleichen gemeinsamen Teiler wie das Paar $(f, g \bmod f)$ hat, können wir einen größten gemeinsamen Teiler mithilfe des Euklidschen Algorithmus berechnen.

Wir rechnen dazu zwei Beispiele:

AUFGABE 13.26. Wir berechnen einen größten gemeinsamen Teiler von $f, g \in \mathbb{R}[x]$ für

$$f = -8\,x + 4\,x^2 + 6\,x^3 - 5\,x^4 + x^5$$

und

$$g = 4 - 4\,x - x^2 + x^3.$$

Wir bilden die gleiche Tabelle wie beim Euklidschen Algorithmus für ganze Zahlen und erhalten:

$$\begin{array}{ccc}
-8\,x + 4\,x^2 + 6\,x^3 - 5\,x^4 + x^5 & 1 & 0 \\
4 - 4\,x - x^2 + x^3 & 0 & 1 \\
-24 + 32\,x - 10\,x^2 & 1 & -6 + 4\,x - x^2 \\
-\left(\frac{32}{25}\right) + \frac{16\,x}{25} & \frac{11}{50} + \frac{x}{10} & -\left(\frac{8}{25}\right) + \frac{7\,x}{25} + \frac{9\,x^2}{50} - \frac{x^3}{10} \\
0 & &
\end{array}$$

Um einen normierten gemeinsamen Teiler zu erhalten, multiplizieren wir die vorletzte Zeile dieser Tabelle mit $\frac{25}{16}$ und erhalten $-2 + x$ als einen größten gemeinsamen Teiler.

Außerdem gilt

$$-2 + x = (\frac{11}{32} + \frac{5x}{32}) \cdot f + (-\frac{1}{2} + \frac{7x}{16} + \frac{9x^2}{32} - \frac{5x^3}{32}) \cdot g.$$

AUFGABE 13.27. Wir berechnen den größten gemeinsamen Teiler der Polynome

$$f = 1 + x^3 + x^5$$

und

$$g = 1 + x + x^3$$

in $\mathbb{Z}_2[x]$. Wir erhalten

$$\begin{array}{ccc} 1 + x^3 + x^5 & 1 & 0 \\ 1 + x + x^3 & 0 & 1 \\ 1 + x^2 & 1 & x^2 \\ 1 & x & 1 + x^3 \\ 0 & & \end{array}$$

Daher ist 1 ein größter gemeinsamer Teiler, und es gilt

$$1 = x \cdot f + (1 + x^3) \cdot g.$$

Wir können also einen größten gemeinsamen Teiler mithilfe des Euklidschen Algorithmus bestimmen. Daraus ergibt sich:

SATZ 13.28. *Sei $\boldsymbol{K}$ ein Körper, und seien $f, g \in K[x]$, nicht beide* 0*. Dann gibt es einen größten gemeinsamen Teiler $d$ von $f$ und $g$, für den es $u, v \in K[x]$ gibt, sodass $u \cdot f + v \cdot g = d$.*

SATZ 13.29. *Sei $\boldsymbol{K}$ ein Körper, seien $f, g \in K[x]$, nicht beide* 0*, und sei $d \in K[x]$. Wir nehmen an, dass es $u, v \in K[x]$ gibt, sodass $d = u \cdot f + v \cdot g$. Dann teilt jeder gemeinsame Teiler von $f$ und $g$ auch das Polynom $d$.*

*Beweis:* Sei $h$ ein gemeinsamer Teiler von $f$ und $g$. Dann gilt $h \mid uf + vg$, also $h \mid d$. ■

KOROLLAR 13.30. *Sei $\boldsymbol{K}$ ein Körper, und seien $f, g \in K[x]$, nicht beide* 0*. Seien $d_1, d_2 \in K[x]$ beide* ggT *von $f$ und $g$. Dann gilt $d_1 = d_2$.*

*Beweis*: Nach Satz 13.28 gibt es einen größten gemeinsamen Teiler $d$ von $f$ und $g$, der sich als $uf + vg$ mit $u, v \in K[x]$ schreiben lässt. Wegen Satz 13.29 gilt $d_1 \mid d$. Sowohl $d_1$ als auch $d$ haben den maximal möglichen Grad unter allen gemeinsamen Teilern von $f$ und $g$. Also gilt $\deg(d_1) = \deg(d)$. Somit gibt es ein $\alpha \in K$, sodass $d = \alpha d_1$. Da $d$ und $d_1$ normiert sind, gilt $\alpha = 1$ und somit $d = d_1$. Ebenso gilt $d = d_2$, also $d_1 = d_2$. ■

Sei $\boldsymbol{K}$ ein Körper. Ein Polynom $f \in K[x]$ ist irreduzibel über $\boldsymbol{K}$, wenn $\deg(f) \geq 1$, und wenn für alle $h, g \in K[t]$ mit $hg = f$ gilt, dass $\deg(h) = 0$ oder $\deg(g) = 0$. Jedes Polynom vom Grad 1 ist offensichtlich irreduzibel.

SATZ 13.31. *Sei $\boldsymbol{K}$ ein Körper, und seien $f, g, h \in K[x]$ so, dass $f$ irreduzibel über $\boldsymbol{K}$ ist. Wenn $f \mid gh$, so gilt $f \mid g$ oder $f \mid h$.*

*Beweis:* Wenn $f$ das Polynom $g$ nicht teilt, so gilt $\mathrm{ggT}(f, g) = 1$. Also gibt es $u, v \in K[x]$ mit $1 = uf + vg$, und somit $h = ufh + vgh$. Da $f \mid ufh$ und $f \mid vgh$, gilt auch $f \mid h$. ■

SATZ 13.32 (Zerlegung in irreduzible Polynome). *Sei $\boldsymbol{K}$ ein Körper, sei $\mathrm{Irr}(\boldsymbol{K})$ die Menge aller normierten, über $\boldsymbol{K}$ irreduziblen Polynome in $K[x]$, sei $f \in K[x] \setminus \{0\}$, sei $n := \deg(f)$, und sei $f_n$ der führende Koeffizient von $f$. Dann gibt es genau eine Funktion $\alpha : \mathrm{Irr}(\boldsymbol{K}) \to \mathbb{N}_0$, sodass $\{g \in \mathrm{Irr}(\boldsymbol{K}) \mid \alpha(g) \neq 0\}$ endlich ist, und*

$$f = f_n * \prod_{g \in \mathrm{Irr}(\boldsymbol{K})} g^{\alpha(g)}.$$

## 5. Polynomfunktionen und Nullstellen

Wir erinnern uns, dass für

$$f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in K[x]$$

die von $f$ auf $\boldsymbol{K}$ induzierte Funktion $f^{\boldsymbol{K}}$ durch

$$\begin{array}{rccl} f^{\boldsymbol{K}} : & K & \longrightarrow & K \\ & k & \longmapsto & a_0 + a_1 k + a_2 k^2 + \cdots + a_n k^n \end{array}$$

definiert ist.

DEFINITION 13.33. Sei $\boldsymbol{K}$ ein Körper, sei $f \in K[x]$, und sei $\alpha \in K$. Die Zahl $\alpha$ ist eine *Nullstelle* von $f$, wenn $f^{\boldsymbol{K}}(\alpha) = 0$.

SATZ 13.34. *Sei $\boldsymbol{K}$ ein Körper, sei $f \in K[x]$, und sei $\alpha \in K$. Dann ist $\alpha$ genau dann eine Nullstelle von $f$, wenn $x - \alpha \mid f$ gilt.*

ÜBUNGSAUFGABEN 13.35.

(1) Zeigen Sie:
Sei $\boldsymbol{K}$ ein Körper, und sei $f \in K[x]$ ein Polynom mit $\deg(f) \geq 2$ und einer Nullstelle $\alpha \in K$. Dann ist $f$ nicht irreduzibel.
(2) Sei $\boldsymbol{K}$ ein Körper. Zeigen Sie, dass jedes Polynom vom Grad 2 oder 3 über $\boldsymbol{K}$, das keine Nullstelle in $K$ hat, irreduzibel über $\boldsymbol{K}$ ist.
(3) Finden Sie ein nicht irreduzibles Polynom vom Grad 4 über $\mathbb{Q}$, das keine Nullstelle in $\mathbb{Q}$ hat und nicht irreduzibel ist.
(4) Zeigen Sie: jedes irreduzible Polynom über $\mathbb{R}$ hat Grad 1 oder geraden Grad. (Tatsächlich gilt sogar: hat Grad 1 oder 2, aber das ist viel schwieriger zu zeigen.)

SATZ 13.36. *Sei $\boldsymbol{K}$ ein Körper, sei $n \in \mathbb{N}$, und sei $f \in K[x]$ ein Polynom mit $\deg(f) = n$. Dann hat $f$ höchstens $n$ Nullstellen.*

*Beweis:* Wir beweisen diese Aussage durch Induktion nach $n$. Die Aussage stimmt für $n = 1$: ein Polynom der Form $\alpha_1 x + \alpha_2$ hat, wenn $\alpha_1 \neq 0$, nur die Nullstelle $-\alpha_2 \cdot (\alpha_1)^{-1}$.

Wir nehmen nun an, dass $n \geq 1$ ist, und dass jedes Polynom vom Grad $n$ höchstens $n$ Nullstellen hat. Wir zeigen, dass dann jedes Polynom vom Grad $n + 1$ höchstens

$n + 1$ Nullstellen haben kann. Sei dazu $f$ ein Polynom vom Grad $n + 1$. Wenn $f$ keine Nullstellen hat, dann sind wir fertig, denn“keine Nullstellen” heißt natürlich auch “weniger als $n + 2$ Nullstellen”. Wenn $f$ zumindest eine Nullstelle hat, dann wählen wir eine Nullstelle $\alpha$. Wir können dann ein Polynom $g$ vom Grad $n$ finden, sodass

$$f = (x - \alpha) \cdot g.$$

Sei nun $\beta$ eine Nullstelle von $f$ mit $\beta \neq \alpha$. Dann gilt $f^{\boldsymbol{K}}(\beta) = (\beta - \alpha) \cdot g^{\boldsymbol{K}}(\beta)$. Also gilt $0 = (\beta - \alpha) \cdot g^{\boldsymbol{K}}(\beta)$. Wegen $\beta - \alpha \neq 0$ gilt $g^{\boldsymbol{K}}(\beta) = 0$. Das Element $\beta$ ist daher eine Nullstelle von $g$.

Da wir angenommen haben, dass jedes Polynom vom Grad $n$ höchstens $n$ Nullstellen hat, hat $g$ höchstens $n$ Nullstellen. Jede Nullstelle von $f$ ist entweder gleich $\alpha$ oder unter diesen $n$ Nullstellen von $g$. Somit hat $f$ höchstens $n + 1$ Nullstellen. ■

DEFINITION 13.37. Sei $\boldsymbol{K}$ ein Körper, sei $f \in K[x]\backslash\{0\}$, und sei $\alpha \in K$ eine Nullstelle von $f$. Wir definieren die *Vielfachheit der Nullstelle $\alpha$ von $f$* als

$$\max\{n \in \mathbb{N} : (x - \alpha)^n \mid f\}.$$

## 6. Polynome über den reellen und den komplexen Zahlen

DEFINITION 13.38. Wir definieren $\mathbb{C} := \{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R}\}$ als die *Menge der komplexen Zahlen*.

LEMMA 13.39. *Die Menge* $\mathbb{C} := \{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R}\}$ *hat folgende Eigenschaften:*

(1) $\forall c_1, c_2 \in \mathbb{C} : c_1 + c_2 \in \mathbb{C}, -c_1 \in \mathbb{C}, c_1 \cdot c_2 \in \mathbb{C}$.
(2) $\forall c \in \mathbb{C} \setminus \{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\} : c^{-1} \in \mathbb{C}$.
(3) $\forall c_1, c_2 \in \mathbb{C} : c_1 \cdot c_2 = c_2 \cdot c_1$.

$(\mathbb{C}, +, -, \cdot, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$ ist also ein Körper.

Mit den Abkürzungen $e := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $i := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ lässt sich die komplexe Zahl $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ auch als $a * e + b * i$, oder kürzer als $a + b\,i$ schreiben. Es gilt $i^2 = -1$, das Polynom $x^2 + 1$ hat also in $\mathbb{C}$ die Nullstellen $i$ und $-i$. Komplexe Zahlen der Form $a + 0i$ bezeichnen wir auch als *reell*.

Für die Zahl $z = a + b\,i$ bezeichnen wir $\overline{z} := a - b\,i$ als die zu $z$ *konjugiert komplexe* Zahl. Schreiben wir die komplexe Zahl $z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ als Matrix, so gilt $\overline{z} = z^T$.

SATZ 13.40. *Seien $z_1, z_2, z \in \mathbb{C}$. Dann gilt $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ und $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$. Die Zahlen $\overline{z} \cdot z$ und $\overline{z} + z$ sind stets reell.*

SATZ 13.41 (Hauptsatz der Algebra, Gauß(1799), Argand (1806)). *Sei $f$ ein Polynom in $\mathbb{C}[x]$ mit $\deg(f) > 0$. Dann besitzt $f$ eine Nullstelle $\alpha \in \mathbb{C}$.*

KOROLLAR 13.42.

(1) *Jedes über* $\mathbb{C}$ *irreduzible Polynom in* $\mathbb{C}[x]$ *hat Grad* $1$.
(2) *Jedes über* $\mathbb{R}$ *irreduzible Polynom in* $\mathbb{R}[x]$ *hat Grad* $1$ *oder* $2$.

*Beweis:* Wir nehmen an, dass $f$ normiert ist. (1): Sei $f \in \mathbb{C}[x]$ irreduzibel über $\mathbb{C}$. Nach dem Hauptsatz der Algebra hat $f$ eine Nullstelle $\alpha \in \mathbb{C}$, also gilt $x - \alpha \mid f$. Somit gilt $x - \alpha = f$. (2): Sei $f \in \mathbb{R}[x]$ irreduzibel über $\mathbb{R}$. Nach dem Hauptsatz der Algebra hat $f$ eine Nullstelle $\alpha \in \mathbb{C}$. Wenn $\alpha \in \mathbb{R}$, so gilt $x - \alpha \mid f$, also $f = x - \alpha$. Wenn $\alpha \notin \mathbb{R}$, so verwenden wir, dass alle Koeffizienten von $f$ reell sind und folglich gilt: $f^{\mathbb{C}}(\overline{\alpha}) = \sum_{i=0}^{n} f_i \cdot \overline{\alpha}^i = \sum_{i=0}^{n} \overline{f_i} \cdot \overline{\alpha}^i = \sum_{i=0}^{n} \overline{f_i \alpha^i} = \overline{\sum_{i=0}^{n} f_i \alpha^i} = \overline{f^{\mathbb{C}}(\alpha)} = \overline{0} = 0$. Also gilt in $\mathbb{C}[x]$, dass $x - \alpha \mid f$ und $x - \overline{\alpha} \mid f$. Somit gilt in $\mathbb{C}[x]$, dass $(x - \alpha)(x - \overline{\alpha}) \mid f$. Das Polynom $g := (x - \alpha)(x - \overline{\alpha})$ hat nur reelle Koeffizienten. Es gilt $g \mid f$ in $\mathbb{C}[x]$. Somit gilt auch $g \mid f$ in $\mathbb{R}[x]$ (denn gäbe es bei der Division in $\mathbb{R}[x]$ einen Rest $\neq 0$, wäre der Rest bei der Division in $\mathbb{C}[x]$ nicht eindeutig). Es gilt also $g = f$. ∎

KOROLLAR 13.43. *Sei* $n \in \mathbb{N}$, *und sei* $f \in \mathbb{C}[x]$ *ein normiertes Polynom vom Grad* $n$. *Dann gibt es* $m \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ *und* $\nu_1, \ldots, \nu_m \in \mathbb{N}$ *sodass*

$$f = \prod_{i=1}^{m} (x - \lambda_i)^{\nu_i}$$

*und* $\nu_1 + \cdots + \nu_m = n$.

KAPITEL 14

# Eigenwerte und Eigenvektoren

## 1. Diagonalisieren von Matrizen

DEFINITION 14.1. Sei $K$ ein Körper, $n \in \mathbb{N}$, und sei $A \in K^{n\times n}$. Mit $h_A$ bezeichnen wir die Abbildung, die durch $h_A : K^n \to K^n$, $h_A(x) = A \cdot x$ für $x \in K^n$ definiert ist.

Für die kanonische Basis $E$ von $K^n$ gilt also $S_{h_A}(E,E) = A$. Sei nun $B$ eine Basis für $K^n$. Dann gilt

$$S_{h_A}(B,B) = {}_BT_E \cdot A \cdot {}_ET_B.$$

Wenn $P$ die Matrix ist, in deren Spalten die Vektoren der Basis $B$ stehen, so gilt also

$$S_{h_A}(B,B) = P^{-1} \cdot A \cdot P.$$

DEFINITION 14.2. Sei $K$ ein Körper und sei $n \in \mathbb{N}$. Mit $\mathrm{GL}(n,K)$ bezeichnen wir die Menge aller invertierbaren $n \times n$-Matrizen in $K^{n\times n}$. Es gilt also

$$\mathrm{GL}(n,K) = \{A \in K^{n\times n} \mid \det(A) \neq 0\}.$$

DEFINITION 14.3. Sei $K$ ein Körper und sei $n \in \mathbb{N}$. Die Matrizen $A, B \in K^{n\times n}$ sind *ähnlich über $K$*, wenn es ein $P \in \mathrm{GL}(n,K)$ gibt, sodass $B = P^{-1} \cdot A \cdot P$. Wir schreiben dann $A \sim B$.

LEMMA 14.4. *Sei $K$ ein Körper und sei $n \in \mathbb{N}$. Die Relation $\sim$ ist eine Äquivalenzrelation auf $K^{n\times n}$.*

SATZ 14.5. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, und seien $A, B \in K^{n\times n}$. Dann sind äquivalent:*

(1) *$A \sim B$.*
(2) *Es gibt eine Basis $C$ von $K^n$, sodass $S_{h_A}(C,C) = B$.*
(3) *Es gibt eine lineare Abbildung $h : K^n \to K^n$ und Basen $D, F$ von $K^n$, sodass $S_h(D,D) = A$ und $S_h(F,F) = B$.*

DEFINITION 14.6. Sei $K$ ein Körper, und sei $n \in \mathbb{N}$. Die Matrix $D \in K^{n\times n}$ ist eine *Diagonalmatrix*, wenn für alle $i, j \in \{1, \dots, n\}$ mit $i \neq j$ gilt: $D(i, j) = 0$.

Für die Matrix

$$\begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & a_n \end{pmatrix}$$

schreiben wir oft auch einfach $\operatorname{diag}(a_1, a_2, \dots, a_n)$.

DEFINITION 14.7. Sei $K$ ein Körper und sei $n \in \mathbb{N}$. Die Matrix $A$ ist *diagonalisierbar über $K$*, wenn es eine Matrix $P \in \mathrm{GL}(n, K)$ gibt, sodass $P^{-1} \cdot A \cdot P$ eine Diagonalmatrix ist.

Eine Matrix ist also diagonalisierbar über $K$, wenn sie über $K$ ähnlich zu einer Diagonalmatrix ist.

SATZ 14.8. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, und sei $A \in K^{n\times n}$. Dann ist $A$ genau dann diagonalisierbar über $K$, wenn es eine Basis $B = (b_1, \dots, b_n)$ von $K^n$ und $\lambda_1, \dots, \lambda_n \in K$ gibt, sodass für alle $i \in \{1, \dots, n\}$: $A \cdot b_i = \lambda_i * b_i$.*

*Beweisskizze:* Wenn $P^{-1} \cdot A \cdot P$ eine Diagonalmatrix $D$ ist, so gilt

$$A \cdot P = P \cdot D.$$

Seien $\mu_1, \dots, \mu_n \in K$ so, dass $D = \operatorname{diag}(\mu_1, \dots, \mu_n)$, und seien $b_1, \dots, b_n$ die Spaltenvektoren von $P$. Dann gilt

$$A \cdot P = \begin{pmatrix} | & | & & | \\ \mu_1 * b_1 & \mu_2 * b_2 & \cdots & \mu_n * b_n \\ | & | & & | \end{pmatrix}.$$

Also gilt $A \cdot b_i = \mu_i * b_i$ für $i \in \{1, \dots, n\}$.

Wenn $(b_1, \dots, b_n)$ eine Basis ist, sodass $A \cdot b_i = \lambda_i * b_i$ für alle $i \in \{1, \dots, n\}$, so gilt wegen $S_{h_A}(B, B) = {}_BT_E \cdot S_{h_A}(E, E) \cdot {}_ET_B$ die Gleichung $\operatorname{diag}(\lambda_1, \dots, \lambda_n) = P^{-1} \cdot A \cdot P$, wobei $P$ die Matrix ist, in deren Spalten die Vektoren $b_1, \dots, b_n$ stehen. ■

## 2. Eigenwerte

DEFINITION 14.9. Sei $K$ ein Körper, sei $n \in \mathbb{N}$, sei $A \in K^{n\times n}$, und sei $\lambda \in K$. Dann ist $\lambda$ ein *Eigenwert* von $A$ :⇔ es gibt ein $v \in K^n$ mit $v \neq 0$ und $A \cdot v = \lambda * v$.

SATZ 14.10. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, $A \in K^{n\times n}$, $\lambda \in K$. Dann sind äquivalent:*

(1) *$\lambda$ ist ein Eigenwert von $A$.*
(2) $\det(\lambda * E - A) = 0$.

DEFINITION 14.11. Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$. Wir bezeichen die Matrix $C := t * E - A$ in $K[t]^{n\times n}$ als die *charakteristische Matrix von $A$* und

$$c_A := \det(t * E - A)$$

als das *charakteristische Polynom von $A$*.

Wir berechnen als Beispiel das charakteristische Polynom der Matrix

$$A := \begin{pmatrix} 1 & -1 & 1 \\ 12 & -7 & 12 \\ 10 & -5 & 10 \end{pmatrix}.$$

**$A$ = {{1, −1, 1}, {12, −7, 12}, {10, −5, 10}}**

{{1, −1, 1}, {12, −7, 12}, {10, −5, 10}}

**MatrixForm[$A$]**

$$\begin{pmatrix} 1 & -1 & 1 \\ 12 & -7 & 12 \\ 10 & -5 & 10 \end{pmatrix}$$

**CC = $t$ ∗ IdentityMatrix[3] − $A$**

{{−1 + $t$, 1, −1}, {−12, 7 + $t$, −12}, {−10, 5, −10 + $t$}}

**MatrixForm[CC]**

$$\begin{pmatrix} -1+t & 1 & -1 \\ -12 & 7+t & -12 \\ -10 & 5 & -10+t \end{pmatrix}$$

**$c$ = Det[CC]**

$-5t - 4t^2 + t^3$

Somit erhalten wir $c_A = t^3 - 4t^2 - 5t$ als das charakteristische Polynom von $A$.

SATZ 14.12. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$, und sei $c_A$ das charakteristische Polynom von $A$. Dann gilt für alle $x \in K$: ${c_A}^K(x) = \det(x * E_n - A)$.*

SATZ 14.13. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, $A \in K^{n\times n}$, und sei $c_A = c_0 + c_1 t^1 + \cdots + c_{n-1}t^{n-1} + c_n t^n$ das charakteristische Polynom von $A$. Dann gilt:*

(1) *Das charakteristische Polynom von $A$ ist ein normiertes Polynom vom Grad $n$, es gilt also $c_n = 1$.*
(2) *Für alle $\lambda \in K$ gilt: $\lambda$ ist ein Eigenwert von $A \Leftrightarrow \lambda$ ist eine Nullstelle von $c_A$.*
(3) ${c_A}^K(0) = c_0 = (-1)^n \det(A)$.
(4) $c_{n-1} = -\sum_{i=1}^{n} A(i,i)$.

KOROLLAR 14.14. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, $A \in K^{n\times n}$. Dann hat $A$ höchstens $n$ Eigenwerte.*

SATZ 14.15. *Sei $K$ ein Körper, sei $n \in \mathbb{N}$, und seien $A, B \in K^{n\times n}$ ähnlich. Dann haben $A$ und $B$ dasselbe charakteristische Polynom.*

## 3. Eigenvektoren

DEFINITION 14.16. Sei $K$ ein Körper, sei $n \in \mathbb{N}$, $A \in K^{n\times n}$, und sei $\lambda$ ein Eigenwert von $A$. Dann ist $v$ ein *Eigenvektor von $A$ zum Eigenwert $\lambda$* $:\Leftrightarrow v \neq 0$ und $A \cdot v = \lambda * v$.

Der *Eigenraum* von $A$ zum Eigenwert $\lambda$ ist definiert durch

$$E(A, \lambda) := \{x \in K^n \,|\, A \cdot x = \lambda * x\}.$$

Der Eigenraum von $A$ zum Eigenwert $\lambda$ ist also genau der Nullraum von $\lambda * E_n - A$, wobei $E_n$ die $n \times n$-Einheitsmatrix ist.

Wir bestimmen nun die Eigenvektoren von

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 12 & -7 & 12 \\ 10 & -5 & 10 \end{pmatrix}.$$

**$A$ = {{1, −1, 1}, {12, −7, 12}, {10, −5, 10}}**

{{1, −1, 1}, {12, −7, 12}, {10, −5, 10}}

**MatrixForm[$A$]**

$$\begin{pmatrix} 1 & -1 & 1 \\ 12 & -7 & 12 \\ 10 & -5 & 10 \end{pmatrix}$$

**Eigenvalues[$A$]**

{5, −1, 0}

**B1 = 5 ∗ IdentityMatrix[3] − $A$**

{{4, 1, −1}, {−12, 12, −12}, {−10, 5, −5}}

**E1 = NullSpace[B1]**

{{0, 1, 1}}

**B2 = (−1) ∗ IdentityMatrix[3] − $A$**

{{−2, 1, −1}, {−12, 6, −12}, {−10, 5, −11}}

**E2 = NullSpace[B2]**

{{1, 2, 0}}

**B3 = 0 ∗ IdentityMatrix[3] − $A$**

{{−1, 1, −1}, {−12, 7, −12}, {−10, 5, −10}}

**NullSpace[B3]**

{{−1, 0, 1}}

Wir erhalten also, dass in den Spalten der Matrix

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

eine Basis $B$ von $\mathbb{R}^3$ steht, die aus lauter Eigenvektoren von $A$ besteht. Es gilt

$$P^{-1} \cdot A \cdot P = \operatorname{diag}(5, -1, 0).$$

Da $\operatorname{diag}(5, -1, 0) = S_{h_A}(B, B)$, lässt sich diese Gleichung auch als

$${}_BT_E \cdot S_{h_A}(E, E) \cdot {}_ET_B = S_{h_A}(B, B)$$

schreiben. $B$ ist also eine Basis, bezüglich der $h_A$ durch eine Diagonalmatrix dargestellt werden kann.

DEFINITION 14.17 (Vielfachheit von Eigenwerten). Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$, und sei $\lambda$ ein Eigenwert von $A$.

(1) Die *algebraische Vielfachheit von $\lambda$ in $A$*, $\mu_{\text{alg}}(A, \lambda)$, ist die Vielfachheit der Nullstelle $\lambda$ im charakteristschen Polynom $c_A$ von $A$.
(2) Die *geometrische Vielfachheit von $\lambda$ in $A$*, $\mu_{\text{geo}}(A, \lambda)$, ist die Dimension des Eigenraums $E(A, \lambda)$. Es gilt also $\mu_{\text{geo}}(A, \lambda) = \dim(E(A, \lambda))$.

SATZ 14.18. *Sei $K$ ein Körper, $n \in \mathbb{N}$, und seien $A, B \in K^{n\times n}$ ähnlich über $K$. Dann haben $A$ und $B$ die gleichen Eigenwerte, und für alle Eigenwerte $\lambda$ von $A$ gilt $\mu_{alg}(A, \lambda) = \mu_{alg}(B, \lambda)$ und $\mu_{geo}(A, \lambda) = \mu_{geo}(B, \lambda)$.*

*Beweis:* Wegen Satz 14.15 haben $A$ und $B$ das gleiche charakteristische Polynom und folglich die gleichen Eigenwerte mit jeweils den gleichen algebraischen Vielfachheiten.

Sei nun $\lambda$ ein Eigenwert von $A$, und sei $P \in \mathrm{GL}(n, K)$ so, dass $P^{-1} \cdot A \cdot P = B$. Wir betrachten die Abbildung $\varphi : E(B, \lambda) \to K^n$, $x \mapsto P \cdot x$. Wenn $x \in E(B, \lambda)$, so gilt $A \cdot \varphi(x) = A \cdot P \cdot x = P \cdot B \cdot x = P \cdot (\lambda * x) = \lambda * (P \cdot x) = \lambda * \varphi(x)$. Also gilt $\varphi(x) \in E(A, \lambda)$. Die Abbildung $\varphi$ ist wegen der Regularität von $P$ injektiv. Also wird eine Basis von $E(B, \lambda)$ auf eine Folge linear unabhängiger Vektoren in $E(A, \lambda)$ abgebildet. Es gilt also $\dim(E(B, \lambda)) \leq \dim(E(A, \lambda))$, und somit $\mu_{\text{geo}}(B, \lambda) \leq \mu_{\text{geo}}(A, \lambda)$.

Genauso kann man $\mu_{\text{geo}}(A, \lambda) \leq \mu_{\text{geo}}(B, \lambda)$ begründen, also ist die geometrische Vielfachheit von $\lambda$ in $A$ und in $B$ gleich. ■

Wir berechnen die Eigenwerte und ihre Vielfachheiten für die Matrix $F = \left(\begin{smallmatrix} 0 & -1 \\ 4 & 4 \end{smallmatrix}\right)$.

**$F$ = {{0, −1}, {4, 4}}**

$$\begin{pmatrix} 0 & -1 \\ 4 & 4 \end{pmatrix}$$

**$p$ = Det[$t$ ∗ IdentityMatrix[2] − $F$]**

$4 - 4t + t^2$

**Factor[*p*]**

$(-2 + t)^2$

**(* Bestimmen der geometrischen Vielfachheit *)**
**NullSpace[2 ∗ IdentityMatrix[2] − *F*]**

$\{\{-1, 2\}\}$

Somit hat $F$ den Eigenwert 2, und es gilt $\mu_{\text{alg}}(2, F) = 2$ und $\mu_{\text{geo}}(2, F) = 1$.

SATZ 14.19. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$, und sei $\lambda$ ein Eigenwert von $A$. Dann gilt $\mu_{geo}(\lambda, A) \leq \mu_{alg}(\lambda, A)$.*

Die geometrische Vielfachheit ist also höchstens so groß wie die algebraische Vielfachheit.

*Beweis:* Sei $(b_1, \dots, b_k)$ eine Basis von $E(\lambda, A)$, und seien $b_{k+1}, \dots, b_n \in K^n$ so, dass $B = (b_1, \dots, b_n)$ eine Basis von $K^n$ ist. Dann gilt

$$M := S_{h_A}(B, B) = \left(\begin{array}{cccc|c} \lambda & 0 & \dots & 0 & \\ 0 & \lambda & & 0 & X \\ \vdots & & \ddots & & \\ 0 & 0 & & \lambda & \\ \hline & & 0 & & Y \end{array}\right)$$

für passende Matrizen $X, Y$, wobei der linke obere Block eine $k \times k$-Matrix ist. Wir berechnen nun das charakteristische Polynom von $M$ in $K[t]$. Es gilt $c_M = \det(t * E - M)$. Durch Entwickeln nach der ersten Spalte (und “Induktion nach $k$”) sehen wir, dass es ein Polynom $r$ gibt, sodass $c_M = (t - \lambda)^k \cdot r$. Also ist die algebraische Vielfachheit von $\lambda$ in $M$ mindestens $k$. Wegen Satz 14.5 ist $M$ ähnlich zu $A$, und wegen Satz 14.18 gilt dann $\mu_{\text{alg}}(M) = \mu_{\text{alg}}(A)$. Also gilt $\mu_{\text{alg}}(A) \geq k$. ∎

LEMMA 14.20. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$, sei $m \in \mathbb{N}$, und seien $\lambda_1, \dots, \lambda_m$ paarweise verschiedene Eigenwerte von $A$. Sei $u_1 \in E(A, \lambda_1)$, …, $u_m \in E(A, \lambda_m)$ so, dass $u_1 + u_2 + \cdots + u_m = 0$. Dann gilt $u_1 = u_2 = \cdots = u_m = 0$.*

DEFINITION 14.21. Sei $K$ ein Körper, und sei $p \in K[x]$. Das Polynom $p$ *zerfällt über $K$ in Linearfaktoren* :⇔ es gibt $m \in \mathbb{N}_0$, $\alpha \in K$, $\lambda_1, \dots, \lambda_m \in K$, $e_1, \dots, e_m \in \mathbb{N}$, sodass

$$p = \alpha \prod_{i=1}^{m} (x - \lambda_i)^{e_i}.$$

SATZ 14.22. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$. Dann sind äquivalent:*

(1) *$A$ ist über $K$ diagonalisierbar.*

(2) *$c_A$ zerfällt über $K$ in Linearfaktoren, und für jeden Eigenwert $\lambda$ von $A$ ist die geometrische Vielfachheit gleich der algebraischen Vielfachheit.*

## 4. Abschätzungen für die Eigenwerte

DEFINITION 14.23. Sei $z = a + bi \in \mathbb{C}$. Dann definieren wir den *Betrag* von $z$ durch

$$|z| := \sqrt{a^2 + b^2}.$$

Es gilt also $|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}} = \sqrt{\det(\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right))}$.

LEMMA 14.24. *Seien $z_1, z_2 \in \mathbb{C}$. Dann gilt $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$ und $|z_1 + z_2| \leq |z_1| + |z_2|$.*

SATZ 14.25 (Gerschgorin, 1931). *Sei $A \in \mathbb{C}^{n\times n}$, und sei $\lambda \in \mathbb{C}$ ein Eigenwert von $A$. Dann gibt es ein $i \in \{1, \dots, n\}$, sodass*

$$|\lambda - A(i, i)| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |A(i, j)|$$

Dieser Satz liefert auch eine Abschätzung für den Betrag der Nullstellen eines Polynoms. Dazu brauchen wir folgenden Begriff.

DEFINITION 14.26. Sei $K$ ein Körper, und sei $f = f_0 + f_1 t + \cdots + f_{n-1}t^{n-1} + t^n$ ein normiertes Polynom vom Grad $n$ in $K[t]$. Dann ist die Matrix

$$B(f) := \begin{pmatrix} 0 & 0 & \dots & 0 & -f_0 \\ 1 & 0 & \dots & 0 & -f_1 \\ 0 & 1 & \ddots & \vdots & -f_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \dots & 1 & -f_{n-1} \end{pmatrix}$$

die *Begleitmatrix von $f$*.

SATZ 14.27. *Sei $K$ ein Körper, und sei $f \in K[t]$ ein normiertes Polynom vom Grad $n$. Dann ist das charakteristische Polynom der $n \times n$-Matrix $B(f)$ gleich $f$; es gilt also $c_{B(f)} = f$.*

KOROLLAR 14.28. *Sei $f \in \mathbb{C}[t]$ ein normiertes Polynom vom Grad $n$, und sei $x \in \mathbb{C}$ eine Nullstelle von $f$. Dann gilt:*

(1) $|x| \leq \max(|f_0|, 1 + |f_1|, \dots, 1 + |f_{n-1}|)$.
(2) $|x| \leq \max(1, \sum_{j=0}^{n-1} |f_j|)$.

## 5. Beweis des Hauptsatzes der Algebra

In diesem Kapitel geben wir einen Beweis dafür an, dass jedes Polynom $f \in \mathbb{C}[t]$ mit $\deg(f) > 1$ eine Nullstelle $x \in \mathbb{C}$ besitzt. Dazu brauchen wir drei Lemmata aus der Analysis.

LEMMA 14.29. *Sei $f = f_0 + \cdots + f_{n-1}t^{n-1} + t^n \in \mathbb{C}[t]$ ein normiertes Polynom vom Grad $n \geq 1$, sei $M \in \mathbb{R}$ mit $M > 0$, und sei $x \in \mathbb{C}$. Wenn $|x| > 1$ und*

$$|x| > \sum_{i=0}^{n-1} |f_i| + M,$$

*so gilt $|f^{\mathbb{C}}(x)| > M$.*

LEMMA 14.30 (Formel von de Moivre (1667-1754)). *Sei $z = a + bi = \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right) \in \mathbb{C}$. Dann gibt es $r \in \mathbb{R}$ mit $r \geq 0$ und $\varphi \in [0, 2\pi)$, sodass*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r * \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

*Für alle $n \in \mathbb{N}$ gilt dann*

$$z^n = r^n * \begin{pmatrix} \cos(n\varphi) & -\sin(n\varphi) \\ \sin(n\varphi) & \cos(n\varphi) \end{pmatrix}.$$

KOROLLAR 14.31. *Sei $z \in \mathbb{C}$, $n \in \mathbb{N}$. Es gibt ein $x \in \mathbb{C}$ mit $x^n = z$.*

LEMMA 14.32 (Satz vom Minimum). *Sei $f$ eine stetige Funktion von $\mathbb{R}^2$ nach $\mathbb{R}$, sei $R \in \mathbb{R}$, und sei $K := \{(x, y) \in \mathbb{R}^2 : \|(x, y)\| \leq R\}$. Dann gibt es ein $(x_0, y_0) \in K$, sodass für alle $(x, y) \in K$ gilt: $f(x_0, y_0) \leq f(x, y)$.*

*Beweis des Hauptsatzes der Algebra (Satz 13.41):* Sei $f = f_0 + f_1 t + \cdots + f_{n-1}t^{n-1} + t^n$ ein normiertes Polynom in $\mathbb{C}[t]$ mit $\deg(f) > 0$, und sei $n := \deg(f)$. Wir werden zeigen, dass $f$ eine Nullstelle besitzt.

Seien $R \in \mathbb{R}$, $Q : \mathbb{R}^2 \to R$ und $K \subseteq \mathbb{R}^2$ gegeben durch

$$\begin{array}{rcl} R & := & \left(\sum_{i=0}^{n-1} |f_i|\right) + 3\,|f_0| + 2, \\ Q & : & \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto |f^{\mathbb{C}}(x + i\,y)|, \\ K & := & \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq R^2\}. \end{array}$$

Wegen Lemma 14.32 nimmt $Q$ auf der Menge $K$ ein Minimum an; sei $(x_0, y_0)$ eine Stelle, an der $Q$ auf $K$ minimal ist. Es gilt nun

$$Q(0, 0) = |f^{\mathbb{C}}(0)| = |f_0|.$$

Wenn $x^2 + y^2 \geq R^2$, so gilt $|x + y\,i| \geq R > \max(1, (\sum_{i=0}^{n-1} |f_i|) + 2\,|f_0| + 1)$, und folglich wegen Lemma 14.29

$$Q(x, y) = |f^{\mathbb{C}}(x + y\,i)| > 2\,|f_0| + 1.$$

Also gilt $|x_0 + y_0\,i| < R$; die Stelle, an der $Q$ sein Minimum auf $K$ annimmt, liegt also im Inneren des Kreises $K$, und das Minimum von $Q$ auf $K$ ist sogar ein globales Minimum von $Q$ auf $\mathbb{R}^2$.

Sei nun $z_0 := x_0 + y_0\, i$, und sei $g \in \mathbb{C}[t]$ definiert durch

$$g(t) := f(t + z_0).$$

Wir wissen nun, dass $(x, y) \mapsto |g^{\mathbb{C}}(x + y\, i)|$ an der Stelle $(0, 0)$ ein lokales Minimum besitzt. Wir schreiben $g$ als

$$g = b_0 + b_k t^k + b_{k+1} t^{k+1} \cdots + t^n,$$

wobei $k := \min\{j \in \{1, \ldots, n\} : b_j \neq 0\}$.

Wenn $b_0 = 0$, so gilt $g^{\mathbb{C}}(0) = 0$. Dann gilt $f^{\mathbb{C}}(z_0) = 0$, und somit hat $f$ eine Nullstelle.

Wir betrachten nun den Fall, dass $b_0 \neq 0$. Sei nun $c \in \mathbb{C}$ so, dass

$$c^k = -b_0 \cdot \overline{b_k}.$$

Wir bilden nun $h \in \mathbb{C}[t]$ durch

$$h(t) := \frac{1}{b_0} g(c \cdot t).$$

Wir wissen, dass auch $(x, y) \mapsto |h^{\mathbb{C}}(x + y\, i)|$ an der Stelle $(0, 0)$ ein lokales Minimum besitzt, und dass $h^{\mathbb{C}}(0) = 1$. Wir berechnen nun das Polynom $h$. Es gilt

$$\begin{aligned} h(t) &= \frac{1}{b_0} \cdot \Big(b_0 + b_k c^k\, t^k + \sum_{j=k+1}^{n} b_j c^j\, t^j\Big) \\ &= 1 + \frac{b_k}{b_0}(-b_0 \overline{b_k})\, t^k + \sum_{j=k+1}^{n} \frac{b_j c^j}{b_0} t^j. \end{aligned}$$

Seien $h_0, h_1, \ldots, h_n$ die Koeffizienten von $h$. Dann gilt $h_0 = 1$ und $h_1 = -b_k \overline{b_k}$. Also ist $h_1$ eine negative reelle Zahl. Sei $a := -h_1$, und sei $z \in \mathbb{C}$. Dann gilt $a \in \mathbb{R}$, $a > 0$, und

$$\begin{aligned} |h^{\mathbb{C}}(z)| &= |1 - a\, z^k + \sum_{j=k+1}^{n} h_j z^j| \\ &\leq |1 - a\, z^k| + \sum_{j=k+1}^{n} |h_j||z|^j \\ &= |1 - a\, z^k| + |z|^k \cdot |z| \cdot \sum_{j=k+1}^{n} |h_j||z|^{j-(k+1)}. \end{aligned}$$

Wenn $|z| < 1$, so gilt

$$|h^{\mathbb{C}}(z)| \leq |1 - a\, z^k| + |z|^k \cdot |z| \cdot \sum_{j=k+1}^{n} |h_j|.$$

Wir nehmen nun an, dass $z$ eine reelle Zahl mit $0 < z < 1$, und, falls $\sum_{j=k+1}^{n} |h_j| \neq 0$, auch $z < \dfrac{a}{2 \cdot \sum\limits_{j=k+1}^{n} |h_j|}$ ist. Weiters nehmen wir an, dass $z < \frac{1}{\sqrt[k]{a}}$ ist. Dann ist $1 - a\,z^k$ eine positive reelle Zahl, und es gilt:

$$\begin{aligned}
|h^{\mathbb{C}}(z)| &\leq |1 - a\,z^k| + |z|^k \cdot \frac{a}{2} \\
&= 1 - a\,z^k + z^k \cdot \frac{a}{2} \\
&= 1 - \frac{a}{2} z^k \\
&< 1.
\end{aligned}$$

Daher nimmt $F(x, y) := |h(x + y\,i)|$ an der Stelle $(0, 0)$ kein lokales Minimum an, im Widerspruch zur Wahl von $(x_0, y_0)$ als Minimum von $Q$. Der Fall $b_0 \neq 0$ kann also nicht eintreten. ■

KAPITEL 15

# Homomorphismen zwischen Vektorräumen

## 1. Der Rang einer Matrix

Wir erinnern uns, dass der Rang einer Matrix $A$, $\mathrm{rk}(A)$, als die Dimension des Zeilenraums $Z(A)$ von $A$ definiert ist.

SATZ 15.1. *Seien $m, n \in \mathbb{N}$, und sei $A \in K^{m\times n}$. Dann gilt* $\dim(Z(A)) = \dim(S(A))$.

*Beweis:* Sei $r$ die Dimension des Spaltenraums von $A$, und seien $a_1, \ldots, a_n \in K^m$ die Spalten von $A$. Dann gibt es wegen Übungsbeispiel 6.51(2) $i_1 < i_2 < \cdots < i_r \in \{1, \ldots, n\}$, sodass $(a_{i_1}, \ldots, a_{i_r})$ eine Basis von $S(A)$ ist. Sei

$$B = \begin{pmatrix} | & | & & | \\ a_{i_1} & a_{i_2} & \cdots & a_{i_r} \\ | & | & & | \end{pmatrix}$$

eine Matrix, in deren Spalten die Vektoren dieser Basis stehen. $B$ ist eine $m\times r$-Matrix. Da sich jede Spalte von $A$ als Linearkombination von $a_{i_1}, \ldots, a_{i_r}$ ausdrücken lässt, gibt es eine Matrix $C \in K^{r\times n}$, sodass

$$A = B \cdot C. \tag{15.1}$$

Aus der Gleichung (15.1) sieht man, dass jede Zeile von $A$ eine Linearkombination der Zeilen von $C$ ist. Also liegt jede Zeile von $A$ im Zeilenraum $Z(C)$ von $C$. Da $C$ genau $r$ Zeilen hat, gilt $\dim(Z(C)) \le r$, und somit $\dim(Z(A)) \le r$. Insgesamt gilt also für die Matrix $A$, dass $\dim(Z(A)) \le \dim(S(A))$.

Wir haben also bewiesen, dass für jede Matrix die Dimension des Zeilenraums höchstens so groß wie die Dimension des Spaltenraums ist. Es gilt also auch $\dim(Z(A^T)) \le \dim(S(A^T))$, und somit $\dim(S(A)) \le \dim(Z(A))$. ■

KOROLLAR 15.2. *Sei $A$ eine $m \times n$-Matrix, und sei $r \in \mathbb{N}$. Dann sind äquivalent:*

(1) $r \le \mathrm{rk}(A)$.
(2) *Es gibt $i_1 < \cdots < i_r \in \{1, \ldots, m\}$ und $j_1 < \cdots < j_r \in \{1, \ldots, n\}$, sodass die $r \times r$-Matrix $B$, die durch*

$$B(l, k) := A(i_l, j_k)$$

*definiert ist, regulär ist.*

*Beweis:* (1)⇒(2): Seien die $i_1, \ldots, i_r$-te Zeile von $A$ linear unabhängig, und sei $C$ die $r \times n$-Matrix, die nur aus diesen Zeilen besteht. Die Matrix $C$ hat den Rang $r$, daher besitzt sie $r$ linear unabhängige Spalten; die Indizes dieser Spalten liefern $j_1, \ldots, j_r$. (2)⇒(1): Wenn $B$ regulär ist, so sind die $i_1, \ldots, i_r$-te Zeile von $A$ linear unabhängig, also gilt $\operatorname{rk}(A) \geq r$. ■

## 2. Homomorphismen zwischen Vektorräumen

DEFINITION 15.3. Seien $U, V$ Vektorräume über dem Körper $K$, und sei $h : U \to V$.

(1) $h$ ist genau dann ein *Homomorphismus* (genauer: *K-Vektorraum-Homomorphismus*) von $U$ nach $V$, wenn $h$ eine lineare Abbildung von $U$ nach $V$ ist, also wenn $h(u_1 + u_2) = h(u_1) + h(u_2)$ und $h(\alpha * u) = \alpha * h(u)$ für alle $u, u_1, u_2 \in U$, $\alpha \in K$ gilt.
(2) $h$ ist genau dann ein *Monomorphismus*, wenn $h$ ein injektiver Homomorphismus ist.
(3) $h$ ist genau dann ein *Epimorphismus*, wenn $h$ ein surjektiver Homomorphismus ist.
(4) $h$ ist genau dann ein *Isomorphismus*, wenn $h$ ein bijektiver Homomorphismus ist.
(5) $h$ ist genau dann ein *Endomorphismus*, wenn $h$ ein Homomorphismus und $U = V$ ist.
(6) $h$ ist genau dann ein *Automorphismus*, wenn $h$ ein Isomorphismus und $U = V$ ist.

DEFINITION 15.4. Seien $U, V$ Vektorräume über dem Körper $K$, und sei $h : U \to V$ ein Homomorphismus. Dann definieren wir den *Kern von h* durch

$$\ker(h) := \{u \in U \mid h(u) = 0\}.$$

Wir definieren das *Bild von h* oder *Image von h* durch

$$\operatorname{im}(h) := \{h(u) \mid u \in U\}.$$

LEMMA 15.5. *Seien $U, V$ Vektorräume über dem Körper $K$, und sei $h : U \to V$ ein Homomorphismus. Dann ist der Kern von $h$ ein Unterraum von $U$, und das Image von $h$ ein Unterraum von $V$.*

SATZ 15.6. *Seien $U, V$ Vektorräume über dem Körper $K$, und sei $h : U \to V$ ein Homomorphismus. Dann gilt:*

(1) *$h$ ist genau dann injektiv, wenn* $\ker(h) = \{0\}$.
(2) *$h$ ist genau dann surjektiv, wenn* $\operatorname{im}(h) = V$.

SATZ 15.7. *Seien $m, n \in \mathbb{N}$, sei $A$ eine $m \times n$-Matrix über dem Körper $K$, sei $h_A : K^n \to K^m$, $x \mapsto A \cdot x$, und sei $N(A)$ der Nullraum von $A$. Dann gilt*

(1) $\ker(h_A) = N(A)$.

(2) $\mathrm{im}(h_A) = S(A)$.
(3) *Der Rang von A ist die Dimension von* $\mathrm{im}(h_A)$.

SATZ 15.8. *Seien $m, n \in \mathbb{N}$, sei A eine $m \times n$-Matrix über dem Körper K, sei $h_A : K^n \to K^m$, $x \mapsto A \cdot x$ Dann sind äquivalent:*

(1) *$h_A$ ist injektiv.*
(2) $\dim(N(A)) = 0$
(3) *Die Spalten von A sind linear unabhängig.*

SATZ 15.9. *Seien $m, n \in \mathbb{N}$, sei A eine $m \times n$-Matrix über dem Körper K, sei $h_A : K^n \to K^m$, $x \mapsto A \cdot x$ Dann sind äquivalent:*

(1) *$h_A$ ist surjektiv.*
(2) $\mathrm{rk}(A) = m$.
(3) *Die Zeilen von A sind linear unabhängig.*

ÜBUNGSAUFGABEN 15.10.
Beachten Sie in den folgenden Übungsbeispiel, ob wir die Basis eines Vektorraums als Menge (wie etwa in Kapitel 11 und in Lemma 11.7) oder als Folge (wie etwa in Definition 6.18) ansehen.

(1) Seien $U, V$ Vektorräume, $h$ ein Homomorphismus von $U$ nach $V$, und sei $B$ eine Teilmenge von $U$. Zeigen Sie, dass die lineare Hülle des Bildes von $B$ gleich dem Bild der linearen Hülle von $B$ ist, also dass $L(h[B]) = h[L(B)]$.
(2) Seien $U, V$ Vektorräume, $h$ ein Monomorphismus von $U$ nach $V$, und sei $B$ eine linear unabhängige Teilmenge von $U$. Zeigen Sie, dass auch das Bild $h[B]$ linear unabhängig ist.
(3) Seien $U, V$ Vektorräume, $h$ ein Isomorphismus von $U$ nach $V$, und sei $B$ eine Basis von $U$. Zeigen Sie, dass auch das Bild $h[B]$ eine Basis von $V$ ist.
(4) Seien $U, V$ Vektorräume, $h$ ein Homomorphismus von $U$ nach $V$, und seien $u_1, \dots, u_n \in U$ so, dass $(u_1, \dots, u_n)$ linear abhängig ist. Zeigen Sie, dass $(h(u_1), \dots, h(u_n))$ linear abhängig ist.

SATZ 15.11. *Sei $(b_1, \dots, b_l)$ eine linear unabhängige Folge von Vektoren im $K^n$, und sei $(y_1, \dots, y_l)$ eine Folge von Vektoren im $K^m$. Dann gibt es einen Homomorphismus $h : K^n \to K^m$, sodass $h(b_i) = y_i$ für alle $i \in \{1, \dots, n\}$.*

*Beweis:* Wir bilden eine Basis $B = (b_1, \dots, b_l, b_{l+1}, \dots, b_n)$ von $K^n$. Dann definieren wir eine lineare Abbildung $h$ durch

$$h(v) := A \cdot (v)_B,$$

wobei $A$ die $m \times n$-Matrix ist, in deren Spalten die Vektoren $(y_1, \dots, y_l, 0, \dots, 0)$ stehen. Wegen Satz 9.10 ist $h$ eine lineare Abbildung, und sie bildet die die Vektoren $(b_1, \dots, b_l)$ auf die gewünschten Werte ab.

Es gilt dann $S_h(B, E) = A$, und wir erhalten $S_h(E, E)$ durch

$$S_h(E, E) = S_h(B, E) \cdot {}_BT_E = A \cdot P^{-1},$$

wobei $P$ die Matrix mit den Vektoren $(b_1, \dots, b_n)$ in den Spalten ist. ■

Wie in Kapitel 11 sehen wir nun eine Basis von $U$ als eine Teilmenge von $U$ (und nicht als Folge von Vektoren aus $U$) an.

SATZ 15.12. *Seien $U, V$ Vektorräume über dem Körper $K$, sei $B$ eine Basis von $U$, und sei $f$ eine Funktion von $B$ nach $V$. Dann gibt es genau einen Homomorphismus $h : U \to V$, sodass $h(b) = f(b)$ für alle $b \in B$.*

*Beweis:* Wir definieren die Relation $h$ durch

$$h := \Big\{\big(\sum_{b\in T} \lambda(b) * b, \sum_{b\in T} \lambda(b) * f(b)\big) \mid T \text{ ist eine endliche Teilmenge von } B, \lambda : T \to K\Big\}.$$

Dann ist $h$ funktional, ein Homomorphismus, und $h|_B = f$. Wir zeigen als erstes, dass $h$ eine Funktion ist. Seien dazu $T_1, T_2$ endliche Teilmengen von $B$, und sei $\lambda : T_1 \to K$, $\mu : T_2 \to K$ so, dass

$$\sum_{b\in T_1} \lambda(b) * b = \sum_{b\in T_2} \mu(b) * b.$$

Wir bilden nun eine Abbildung $\lambda_1 : T_1 \cup T_2 \to K$ durch $\lambda_1(t) := \lambda(t)$ für $t \in T_1$ und $\lambda_1(t) := 0$ für $t \in T_2 \setminus T_1$. Genauso erweitern wir $\mu$ und bilden eine Abbildung $\mu_1 : T_1 \cup T_2 \to K$ durch $\mu_1(t) := \mu(t)$ für $t \in T_2$ und $\mu_1(t) := 0$ für $t \in T_1 \setminus T_2$. Dann gilt

$$\sum_{b\in T_1\cup T_2} \lambda_1(b) * b = \sum_{b\in T_1\cup T_2} \mu_1(b) * b.$$

Also gilt $\sum_{b\in T_1\cup T_2} (\lambda_1(b) - \mu_1(b)) * b = 0$, und folglich wegen der linearen Unabhängigkeit von $B$ auch $\lambda_1 = \mu_1$. Also gilt

$$\sum_{b\in T_1} \lambda(b) * f(b) = \sum_{b\in T_2} \mu(b) * f(b).$$

Die Relation $h$ ist also funktional.

Wir zeigen nun die Homomorphismuseigenschaften. Es gilt

$$\begin{aligned} h(\sum_{b\in T_1} \lambda(b) * b + \sum_{b\in T_2} \mu(b) * b) &= h(\sum_{b\in T_1\cup T_2} (\lambda_1(b) + \mu_1(b)) * b) \\ &= \sum_{b\in T_1\cup T_2} (\lambda_1(b) + \mu_1(b)) * f(b) \\ &= \sum_{b\in T_1} \lambda(b) * f(b) + \sum_{b\in T_2} \mu(b) * f(b) \\ &= h(\sum_{b\in T_1} \lambda(b) * b) + h(\sum_{b\in T_2} \mu(b) * b). \end{aligned}$$

Dabei sind $\lambda_1, \mu_1$ wieder die Erweiterungen von $\lambda$ und $\mu$ wie oben. Die Eigenschaft $h(\alpha * v) = \alpha * h(v)$ zeigt man genauso.

Sei nun $b \in B$. Dann gilt $h(b) = h(1 * b) = 1 * f(b) = f(b)$, also gilt $h|_B = f$. ■

KOROLLAR 15.13. *Es gibt eine Abbildung $h : \mathbb{R} \to \mathbb{R}$, die $h(x + y) = h(x) + h(y)$ für alle $x, y \in \mathbb{R}$ und $h(1) = 1$ erfüllt, und die nicht gleich der identischen Abbildung auf $\mathbb{R}$ ist.*

*Beweis:* Wir sehen $\mathbb{R}$ als Vektorraum über $\mathbb{Q}$ mit der Vektorraummultiplikation $*$ : $\mathbb{Q} \times \mathbb{R} \to \mathbb{R}$, $q * r := qr$ an. Wir erweitern $\{1\}$ zu einer Basis von $\mathbb{R}$ über $\mathbb{Q}$. (Diese Basis erhält man, indem man mithilfe des Zornschen Lemmas, wie im Beweis von Satz 11.8, eine maximale linear unabhängige Teilmenge $B$ von $\mathbb{R}$ mit der Eigenschaft $1 \in B$ auswählt.) Nun definieren wir $f : B \to \mathbb{R}$ durch $f(1) := 1$ und $f(b) := 0$ für $b \in B \setminus \{1\}$. Wir können nun $f$ zu einem Homomorphismus $h$ von $\mathbb{R}$ nach $\mathbb{R}$ erweitern. Diese Abbildung $h$ ist nicht die identische Abbildung. ■

ÜBUNGSAUFGABEN 15.14.

(1) Sei $h : \mathbb{R} \to \mathbb{R}$ eine stetige Abbildung, die $h(x + y) = h(x) + h(y)$ für alle $x, y \in \mathbb{R}$ erfüllt. Zeigen Sie, dass für alle $x \in \mathbb{R}$ gilt: $h(x) = h(1) \cdot x$.
(2) $^*$ Sei $h : \mathbb{R} \to \mathbb{R}$ eine Abbildung, die $h(x + y) = h(x) + h(y)$ und $h(x \cdot y) = h(x) \cdot h(y)$ für alle $x, y \in \mathbb{R}$ erfüllt. Zeigen Sie, dass $h$ entweder die Nullabbildung oder die identische Abbildung ist.

SATZ 15.15. *Seien $U, V$ Vektorräume über dem Körper $K$, und sei $h : U \to V$ ein Isomorphismus. Dann ist die Abbildung $h^{-1}$ ebenfalls eine lineare Abbildung, und somit ein Isomorphismus von $V$ nach $U$.*

*Beweisskizze:* Wir zeigen nur, dass $h^{-1}$ additiv ist: Seien $a, b \in V$, und sei $x := h^{-1}(a)$, $y := h^{-1}(b)$. Dann gilt $h(x + y) = h(x) + h(y) = a + b$, also $h^{-1}(a + b) = x + y = h^{-1}(a) + h^{-1}(b)$.

DEFINITION 15.16. Seien $U, V$ Vektorräume über dem Körper $K$. Die Vektorräume $U, V$ sind genau dann *isomorph*, wenn es einen Isomorphismus $h : U \to V$ gibt.

SATZ 15.17. *Seien $U, V$ Vektorräume über dem Körper $K$, sei $B$ eine Basis von $U$, und sei $C$ eine Basis von von $V$. Dann sind $U$ und $V$ genau dann isomorph, wenn $B$ und $C$ gleichmächtig sind.*

*Beweisskizze:* Übung 15.10(3) und Satz 11.29. ■

## 3. Faktorräume

DEFINITION 15.18. Sei $V$ ein Vektorraum über dem Körper $K$, und sei $U$ ein Unterraum von $V$, und sei $w \in V$. Dann definieren wir

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

und

$$V/U := \{v + U \mid v \in V\}.$$

DEFINITION 15.19. Sei $V$ ein Vektorraum über dem Körper $K$, und sei $U$ ein Unterraum von $V$. Dann definieren wir $\boldsymbol{+}$ und $\boldsymbol{*}$ durch

$$(v + U) \boldsymbol{+} (w + U) := (v + w) + U$$

und

$$\alpha \boldsymbol{*} (v + U) := (\alpha v) + U$$

für alle $v, w \in V$ und $\alpha \in K$.

Wir überlegen uns, dass $\mathbf{+}$ und $\mathbf{*}$ wohldefiniert sind, das heißt, dass die Relation $\{\big((v + U, w + U), (v + w) + U\big) \mid v, w \in V\}$, die ja eine Teilmenge von $(V/U \times V/U) \times V/U$ ist, eine Funktion von $V/U \times V/U$ nach $V/U$ ist.

DEFINITION 15.20. Sei $V$ ein Vektorraum über dem Körper $K$, und sei $U$ ein Unterraum von $V$. Dann ist $\langle V/U, \mathbf{+}, \mathbf{-}, 0 + U, \mathbf{*}\rangle$ ein Vektorraum über $K$.

ÜBUNGSAUFGABEN 15.21.
In den folgenden Beispielen ist $V$ stets ein Vektorraum, und $U$ ein Unterraum von $V$.

(1) Zeigen Sie, dass die Relation $\sim_U$, die durch $v \sim_U w :\Leftrightarrow v - w \in U$ definiert ist, eine Äquivalenzrelation auf $V$ ist.
(2) Zeigen Sie, dass für die im Beispiel 1 definierte Relation $\sim_U$ und für jedes $v \in V$ gilt: $v/{\sim_U} = v + U$.
(3) Zeigen Sie, dass für die im Beispiel 1 definierte Relation $\sim_U$ gilt: $V/U$ ist gleich der Faktormenge $V/{\sim_U}$.

SATZ 15.22. *Sei $V$ ein Vektorraum über dem Körper $K$. Seien $v_1, \ldots, v_m \in V$ und $u_1, \ldots, u_n$ in $U$ so, dass $(v_1 + U, v_2 + U, \ldots, v_m + U)$ eine Basis von $V/U$ ist und $(u_1, \ldots, u_n)$ eine Basis von $U$ ist. Dann ist $B = (v_1, \ldots, v_m, u_1, \ldots, u_n)$ eine Basis von $V$.*

*Beweis.* Wir zeigen zunächst, dass $V \subseteq L(B)$. Sei dazu $w \in V$. Wegen $w + U \in V/U$ gibt es $\alpha_1, \ldots, \alpha_m \in K$, sodass

$$\sum_{i=1}^{m} \alpha_i * (v_i + U) = w + U.$$

Also gilt

$$(\sum_{i=1}^{m} \alpha_i v_i) + U = w + U.$$

Es gibt also ein $u \in U$, sodass

$$(\sum_{i=1}^{m} \alpha_i v_i) + u = w + 0.$$

Da $u \in U$, gibt es $\beta_1, \ldots, \beta_n \in K$, sodass $u = \sum_{i=1}^{n} \beta_i u_i$. Also gilt

$$w = \sum_{i=1}^{m} \alpha_i v_i + \sum_{i=1}^{n} \beta_i u_i,$$

und daher $w \in L(B)$.

Um zu zeigen, dass $B$ linear unabhängig ist, wählen wir $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \in K$ so, dass

$$\sum_{i=1}^{m} \alpha_i v_i + \sum_{i=1}^{n} \beta_i u_i = 0.$$

Dann gilt

$$\sum_{i=1}^{m} \alpha_i v_i \in U,$$

also

$$\sum_{i=1}^{m} \alpha_i * (v_i + U) = 0 + U.$$

Da $(v_1 + U, \ldots, v_m + U)$ eine Basis von $V/U$ sind, gilt $\alpha_1 = \cdots = \alpha_m = 0$. Also gilt $\sum_{i=1}^{n} \beta_i u_i = 0$, und daher sind wegen der linearen Unabhängigkeit von $(u_1, \ldots, u_n)$ auch alle $\beta_i$ gleich 0. ■

KOROLLAR 15.23. *Sei $V$ ein Vektorraum über $K$, und sei $U$ ein endlichdimensionaler Unterraum von $V$, sodass $V/U$ ebenfalls wieder endlichdimensional ist. Dann ist $V$ endlichdimensional, und es gilt*

$$\dim(V) = \dim(V/U) + \dim U.$$

Dieser Satz gilt auch für unendlichdimensionale Vektorräume. Allerdings müssen wir dazu erst erklären, wie für unendlichdimensionale Vektorräume Gleichungen der Form $\dim(U) + \dim(V) = \dim(W)$ zu lesen sind. Zunächst müssen wir $\dim(U)$ definieren, wenn $U$ keine endliche Basis besitzt. Wir definieren $\dim(U)$ dann als die "Kardinalität" (Mächtigkeit) einer Basis von $U$. Wegen Satz 11.29 sind alle Basen von $U$ gleichmächtig. Allerdings fehlen uns noch die Objekte, diese Mächtigkeiten, also etwa $|\mathbb{N}| = \aleph_0$, $|\mathbb{R}| = 2^{\aleph_0}$, … zu beschreiben. Dazu kann man *Kardinalzahlen* einführen, die man dann auch wirklich addieren kann (siehe [**Halmos, 1976**]).

Eine formal weniger aufwändige Sichtweise, der Gleichung $\dim(U)+\dim(V) = \dim(W)$ eine definierte Bedeutung zu geben, ist, diese Gleichung als eine Abkürzung für folgende Behauptung zu sehen:

> Es gibt Basen $B$ von $U$, $C$ von $V$ und $D$ von $W$, sodass $D$ gleichmächtig zu $(B \times \{0\}) \cup (C \times \{1\})$ ist.

SATZ 15.24. *Sei $V$ ein $K$-Vektorraum, und sei $U$ ein Unterraum von $V$. Wenn $U$ eine Basis $B$ und $V/U$ eine Basis $C$ hat, so hat $V$ eine Basis, die gleichmächtig zu* $(B \times \{0\}) \cup (C \times \{1\})$ *ist; es gilt also* $\dim(U) + \dim(V/U) = \dim(V)$.

## 4. Der Homomorphiesatz

SATZ 15.25 (Homomorphiesatz). *Seien $V$ und $W$ Vektorräume, und sei $h$ eine lineare Abbildung von $V$ nach $W$. Sei $U := \ker(h)$. Dann ist die Abbildung*

$$\begin{array}{rccc} H : & V/U & \longrightarrow & W \\ & v + U & \longmapsto & h(v) \end{array}$$

*wohldefiniert. Die Abbildung $H$ ist außerdem ein Isomorphismus von $V/U$ nach* $\mathrm{im}(h)$.

Anstelle "Dann ist die Abbildung [...] wohldefiniert" kann man sagen: "Dann ist die Relation $H = \{(v + U, h(v)) \mid v \in V\}$ eine Funktion von $V/U$ nach $W$."

KOROLLAR 15.26. *Seien $V$, $W$ Vektorräume über $K$ (nicht notwendigerweise endlichdimensional), und sei $h$ eine lineare Abbildung von $V$ nach $W$. Dann gilt*

$$\dim(\mathrm{im}(h)) + \dim(\ker(h)) = \dim(V).$$

*Beweis:* Sei $U := \ker(h)$. Wegen Satz 15.24 gilt

$$\dim(V/U) + \dim(U) = \dim(V).$$

Da $V/\ker(h)$ isomorph zu $\mathrm{im}(h)$ ist, gilt also

$$\dim(\mathrm{im}(h)) + \dim(\ker(h)) = \dim(V).$$

SATZ 15.27 (Rangsatz). *Sei $A$ eine $m \times n$-Matrix über $K$. Dann gilt* $\dim(N(A)) = n - \mathrm{rk}(A)$.

*Beweis:* Wir betrachten $h_A : K^n \to K^m$, $x \mapsto A \cdot x$. Dann gilt $\ker(h_A) = N(A)$, und $\dim(\mathrm{im}(h_A)) = \dim(S(A)) = \mathrm{rk}(A)$. Da $K^n/\ker(h_A)$ isomorph zu $\mathrm{im}(h_A)$ ist, sind die Dimensionen gleich. Es gilt also $n - \dim(N(A)) = \mathrm{rk}(A)$. ■

SATZ 15.28 (Erster Isomorphiesatz). *Seien $U, V$ beide Unterräume des Vektorraums $W$. Dann sind die beiden Faktorräume $U/(U \cap V)$ und $(U + V)/V$ isomorph.*

*Beweis:* Wir betrachten die Abbildung

$$\begin{array}{rcll} h & : & U & \longrightarrow & (U+V)/V \\ & & u & \longmapsto & u+V. \end{array}$$

Die Abbildung $h$ ist surjektiv auf $(U + V)/V$, und es gilt $\ker(h) = U \cap V$. Somit gilt nach dem Homomorphiesatz (Satz 15.25), dass $U/\ker(h)$ isomorph zu $(U + V)/V$ ist. ■

KOROLLAR 15.29. *Seien $U, V$ beide endlichdimensionale Unterräume des Vektorraums $W$. Dann gilt* $\dim(U + V) = \dim(U) + dim(V) - \dim(U \cap V)$.

Für unendlichdimensionale Unterräume kann man diesen Satz als $\dim(U+V)+\dim(U\cap V) = \dim(U) + \dim(V)$ formulieren, und beweisen, indem man eine Basis $B$ von $U \cap V$ zu einer Basis $C$ von $U$ und einer Basis $D$ von $V$ erweitert, und beobachtet, dass $C \cup D$ eine Basis von $U + V$ ist, und $|C \cup D| + |C \cap D| = |C| + |D|$ gilt.

LEMMA 15.30. *Seien $V, W$ Vektorräume, sei $h$ ein Epimorphismus von $V$ nach $W$, und sei $U$ ein Unterraum von $W$. Dann gilt*

$$\dim(\{x \in V \mid h(x) \in U\}) = \dim(U) + \dim(\ker(h)).$$

*Beweis:* Wir wählen eine Basis $B$ von $U$. Dann wählen wir aus jedem $h^{-1}[\{b\}]$ genau ein Element aus, und wir fassen diese Elemente in einer Menge $A$ zusammen. Die Einschränkung $h|_A$ ist also injektiv, und $h[A]$ ist eine Basis von $U$. Dann ist $\ker(h) \cup A$ linear unabhängig und eine Basis von $\{x \in V \mid h(x) \in U\}$.

Wenn $V$ endlichdimensional ist, kann man Lemma 15.30 auch mit dem Homomorphiesatz zeigen: Sei $\varphi : V \to W/U$, $\varphi(v) = h(v) + U$. Dann gilt $\ker(\varphi) = \{x \in V \mid h(x) \in U\}$. Also gilt

$$\dim(\ker(\varphi)) + \dim(\mathrm{im}(\varphi)) = \dim(V) = \dim(\ker(h)) + \dim(\mathrm{im}(h)),$$

und somit

$$\dim(\ker(\varphi)) + \dim(W) - \dim(U) = \dim(\ker(h)) + \dim(W),$$

woraus $\dim(\ker(\varphi)) = \dim(\ker(h)) + \dim(U)$ folgt. ■

SATZ 15.31. *Seien $k, l, m \in \mathbb{N}$, sei $K$ ein Körper, seien $A, A_1, A_2 \in K^{k\times l}$, und sei $B \in K^{l\times m}$. Dann gilt:*

(1) $\mathrm{rk}(A \cdot B) \leq \min(\mathrm{rk}(A), \mathrm{rk}(B))$.
(2) *(Sylvesters Rangungleichung)* $\mathrm{rk}(A \cdot B) \geq \mathrm{rk}(A) + \mathrm{rk}(B) - l$.
(3) $\mathrm{rk}(A_1 + A_2) \leq \mathrm{rk}(A_1) + \mathrm{rk}(A_2)$.
(4) $\mathrm{rk}(A_1 + A_2) \geq \mathrm{rk}(A_1) - \mathrm{rk}(A_2)$.

*Beweis:* (1) Es gilt $\mathrm{rk}(A\cdot B) = \dim(\mathrm{im}(h_{A\cdot B})) = \dim(\mathrm{im}(h_A \circ h_B))$. Nun gilt $\mathrm{im}(h_A \circ h_B) \subseteq \mathrm{im}(h_A)$, also $\dim(\mathrm{im}(h_A \circ h_B)) \leq \mathrm{rk}(A)$.

Weiters gilt für jeden Homomorphismus $\varphi$ und für jeden Vektorraum $X$, dass $\dim(\varphi[X]) \leq \dim(X)$. (Das gilt etwa, weil $\dim(\varphi[X]) + \dim(\ker(\varphi)) = \dim(X)$.) Also gilt $\dim(\mathrm{im}(h_A \circ h_B)) = \dim(h_A[\mathrm{im}(h_B)]) \leq \dim(\mathrm{im}(h_B)) = \mathrm{rk}(B)$.

(2)

$$\begin{aligned}
\dim(\mathrm{im}(h_A \circ h_B)) &= m - \dim(\ker(h_A \circ h_B)) \\
&= m - \dim(\{x \in K^m \mid h_B(x) \in \ker(h_A)\}) \\
&= m - \dim(\{x \in K^m \mid h_B(x) \in \ker(h_A) \cap S(B)\}) \\
&= m - \big(\dim(\ker(h_B)) + \dim(\ker(h_A) \cap S(B))\big) \\
&\geq m - \dim(\ker(h_B)) - \dim(\ker(h_A)) \\
&= m - (m - \dim(\mathrm{im}(h_B))) - (l - \dim(\mathrm{im}(h_A))) \\
&= \mathrm{rk}(B) + \mathrm{rk}(A) - l.
\end{aligned}$$

(3) Wegen $\mathrm{im}(h_{A+B}) \subseteq \mathrm{im}(h_A) + \mathrm{im}(h_B)$ gilt $\dim(\mathrm{im}(h_{A+B})) \leq \dim(\mathrm{im}(h_A) + \mathrm{im}(h_B)) \leq \dim(\mathrm{im}(h_A)) + \dim(\mathrm{im}(h_B)) = \mathrm{rk}(A) + \mathrm{rk}(B)$. (4) $\mathrm{rk}(A) = \mathrm{rk}((A + B) + (-B)) \leq \mathrm{rk}(A + B) + \mathrm{rk}(-B) = \mathrm{rk}(A + B) + \mathrm{rk}(B)$, also $\mathrm{rk}(A + B) \geq \mathrm{rk}(A) - \mathrm{rk}(B)$. ■

## 5. Zerlegung von Vektorräumen in direkte Summen

DEFINITION 15.32. Wir definieren die Summe der Unterräume $U_1, \dots, U_n$ des Vektorraums $V$ durch

$$U_1 + U_2 + \cdots + U_n := \Big\{\sum_{i=1}^{n} u_i \,\Big|\, u_i \in U_i \text{ für alle } i \in \{1, \dots, n\}\Big\}.$$

Besonders interessant ist eine solche Zerlegung, wenn sich jedes Element der Summe eindeutig in Komponenten, die in den einzelnen $U_i$ liegen, zerlegen lässt.

DEFINITION 15.33. Seien $U_1, \dots, U_n$ Unterräume des Vektorraums $V$. Der Unterraum $U$ ist genau dann die *direkte Summe* von $U_1, \dots, U_n$, wenn

(1) $U = U_1 + \cdots + U_n$,
(2) für jedes $u \in U$ gibt es genau ein $(u_1, \dots, u_n) \in U_1 \times \cdots \times U_n$, sodass $u = u_1 + \cdots + u_n$.

Wenn $U$ die direkte Summe von $U_1, \dots, U_n$ ist, so schreiben wir $U = U_1 \dotplus \cdots \dotplus U_n$.

SATZ 15.34. *Seien $U_1, \dots, U_n$ Unterräume des Vektorraums $V$, und sei $U := U_1 + U_2 + \cdots + U_n$. Dann sind äquivalent:*

(1) *Für jedes $u \in U$ gibt es genau ein $(u_1, \dots, u_n) \in U_1 \times \cdots \times U_n$, sodass $u = u_1 + \cdots + u_n$.*
(2) *Für jedes $i \in \{1, \dots, n\}$ gilt*

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_n) = \{0\}.$$

SATZ 15.35. *Seien $U, U_1, \dots, U_n$ Unterräume des Vektorraums $V$. Wir nehmen an, dass $U = U_1 \dotplus \cdots \dotplus U_n$, und dass für jedes $i \in \{1, \dots, n\}$ die Menge $B_i$ eine Basis für $U_i$ ist. Dann ist $B_1 \cup \cdots \cup B_n$ eine Basis für $U$.*

Folglich ist die Dimension einer direkten Summe gleich der Summe der Dimensionen der Summanden.

KAPITEL 16

# Die Jordansche Normalform

## 1. Einsetzen von Matrizen in Polynome

DEFINITION 16.1. Sei $K$ ein Körper, $n \in \mathbb{N}$, und sei $A \in K^{n\times n}$. Sei $p = \sum_{i=0}^{n} p_i t^i \in K[t]$. Dann definieren wir $\hat{p}(A)$ durch

$$\hat{p}(A) := \sum_{i=0}^{n} p_i * A^i.$$

Dabei ist $A^0 := E_n$, $A^n := A^{n-1} \cdot A$ für alle $n \in \mathbb{N}$.

SATZ 16.2. *Sei $K$ ein Körper, $n \in \mathbb{N}$, sei $A \in K^{n\times n}$, und seien $p, q \in K[t]$. Dann gilt $\hat{p}(A) \cdot \hat{q}(A) = (\widehat{pq})(A)$.*

*Beweis:* Der Beweis von Satz 13.21 kann mit $y := A$ übernommen werden. ■

ÜBUNGSAUFGABEN 16.3.

(1) Sei $A$ eine $n \times n$-Matrix über $K$, und seien $p, q \in K[t]$. Zeigen Sie, dass für die Matrizen $B := \hat{p}(A)$ und $C := \hat{q}(A)$ gilt, dass $B \cdot C = C \cdot B$.

SATZ 16.4 (Satz von Cayley-Hamilton). *Sei $K$ ein Körper, $n \in \mathbb{N}$, sei $A \in K^{n\times n}$, und sei $c = \sum_{i=0}^{n} \gamma_i t^i \in K[t]$ das charakteristische Polynom von $A$. Dann gilt $\hat{c}(A) = 0$.*

*Beweis:* Sei $C := t * E_n - A$, und sei $B := C^{\mathrm{ad}}$. Die Matrix $B$ hat als Einträge Polynome in $K[t]$ vom Grad $\leq n-1$. Es gibt also Matrizen $B_0, B_1, \ldots, B_{n-1} \in K^{n\times n}$, sodass

$$B = t^0 * B_0 + t^1 * B_1 + \cdots + t^{n-1} * B_{n-1}.$$

Da

$$B \cdot C = \begin{pmatrix} c & 0 & 0 & \ldots & 0 \\ 0 & c & 0 & \ldots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \ldots & \ldots & c \end{pmatrix},$$

gilt

$$(16.1) \qquad B \cdot C = \sum_{i=0}^{n} \left( (\gamma_i t^i) * \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 1 \end{pmatrix} \right).$$

Sei nun für $i \in \{0, \dots, n\}$ die Matrix $D_i$ definiert durch $D_i := \gamma_i * E_n$. Aus Gleichung (16.1) ergibt sich

$$B \cdot C = \sum_{i=0}^{n} t^i * D_i,$$

also

$$(\sum_{i=0}^{n-1} t^i * B_i) \cdot (t * E_n - A) = \sum_{i=0}^{n} t^i * D_i.$$

Es gilt

$$\begin{aligned} \sum_{i=0}^{n-1} t^{i+1} * B_i - \sum_{i=0}^{n-1} t^i * (B_i \cdot A) &= \Big( \sum_{j=1}^{n} t^j * B_{j-1} \Big) - \Big( \sum_{i=1}^{n-1} t^i * (B_i \cdot A) \Big) - t^0 * (B_0 \cdot A) \\ &= t^0 * (-B_0 \cdot A) + \Big( \sum_{i=1}^{n-1} t^i * (B_{i-1} - B_i \cdot A) \Big) + t^n * B_{n-1}. \end{aligned}$$

Folglich gilt $D_0 = -B_0 \cdot A$, für alle $i \in \{1, \dots, n-1\}$ gilt $D_i = B_{i-1} - B_i \cdot A$, und $D_n = B_{n-1}$. Wir berechnen nun $\hat{c}(A)$. Es gilt

$$\begin{aligned} \hat{c}(A) &= \sum_{i=0}^{n} \gamma_i * A^i \\ &= \sum_{i=0}^{n} D_i \cdot A^i \\ &= -B_0 \cdot A + \sum_{i=1}^{n-1} B_{i-1} \cdot A^i - \sum_{i=1}^{n-1} B_i \cdot A^{i+1} + B_{n-1} \cdot A^n \\ &= -B_0 \cdot A + B_0 \cdot A - B_{n-1} \cdot A^n + B_{n-1} \cdot A^n \\ &= 0. \end{aligned}$$

■

ÜBUNGSAUFGABEN 16.5.

(1) Berechnen Sie die Matrizen $B_0, B_1, B_2$ aus dem Beweis von Satz 16.4 für die Matrix $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ -3 & 4 & 9 \end{pmatrix}$, und weisen Sie nach, dass $-B_0 \cdot A$, $B_0 - B_1 \cdot A$, $B_1 - B_2 \cdot A$ und $B_2$ Diagonalmatrizen sind.

## 2. Zerlegung in invariante Unterräume

In diesem Kapitel bezeichnen wir für eine Matrix $A \in K^{n \times n}$ den Endomorphismus, der $x$ auf $A \cdot x$ abbildet, stets mit $h_A$.

DEFINITION 16.6. Sei $V$ ein Vektorraum, und sei $h : V \to V$ ein Endomorphismus. Ein Unterraum $U$ von $V$ ist *$h$-invariant* $:\Leftrightarrow h(U) \subseteq U$.

LEMMA 16.7. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n \times n}$, und $p \in K[t]$. Sei $B := \hat{p}(A)$. Dann ist $N(B)$ ein $h_A$-invarianter Unterraum von $V$.*

*Beweis:* Sei $u \in N(B)$. Dann gilt $\hat{p}(A) \cdot u = 0$. Nun gilt

$$\begin{aligned}
\hat{p}(A) \cdot (A \cdot u) &= \Big(\sum_{i=0}^{\deg(p)} p_i * A^i\Big) \cdot (A \cdot u) \\
&= \Big(\sum_{i=0}^{\deg(p)} p_i * A^{i+1}\Big) \cdot u \\
&= A \cdot \Big(\sum_{i=0}^{\deg(p)} p_i * A^i\Big) \cdot u \\
&= A \cdot B \cdot u \\
&= A \cdot 0 \\
&= 0.
\end{aligned}$$

DEFINITION 16.8. Sei $K$ ein kommutativer Ring mit Eins, $m, n \in \mathbb{N}$, $A \in K^{m \times m}$, $B \in K^{n \times n}$. Dann definieren wir die $(m+n) \times (m+n)$-Matrix $A \oplus B$ durch

$$A \oplus B = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right).$$

SATZ 16.9. *Sei $K$ ein kommutativer Ring mit Eins, $m, n \in \mathbb{N}$, $A \in K^{m \times m}$, $B \in K^{n \times n}$. Dann gilt* $\det(A \oplus B) = \det(A) \cdot \det(B)$.

*Beweis*: Induktion nach $m$. Wenn $m = 1$, so entwickeln wir nach der ersten Zeile und erhalten $\det(\left(\begin{array}{c|c} A(1,1) & 0 \\ \hline 0 & B \end{array}\right)) = A(1,1) \cdot \det(B) = \det(A)\ \det(B)$.

Wenn $m \geq 2$, so erhalten wir durch Entwickeln nach der ersten Zeile

$$\begin{aligned}\det(A \oplus B) &= \sum_{i=1}^{m} A(1,i) \cdot \det((A \oplus B)^{[1,i]}) \\ &= \sum_{i=1}^{m} A(1,i) \cdot \det(A^{[1,i]} \oplus B).\end{aligned}$$

Nach Induktionsvoraussetzung ist der letzte Ausdruck gleich

$$\begin{aligned}\sum_{i=1}^{m} A(1,i) \cdot \det(A^{[1,i]}) \cdot \det(B) &= \det(B) \cdot \sum_{i=1}^{m} A(1,i) \det(A^{[1,i]}) \\ &= \det(B)\ \det(A).\end{aligned}$$

■

SATZ 16.10 (Zerlegung einer Abbildung auf invariante Unterräume). *Seien $m, n \in \mathbb{N}$, sei $K$ ein Körper, und sei $A \in K^{(m+n)\times(m+n)}$. Seien $U$ und $V$ Unterräume von $K^{m+n}$ mit folgenden Eigenschaften:*

(1) *$U$ ist $m$-dimensional mit Basis $B_U = (u_1, \dots, u_m)$.*
(2) *$V$ ist $n$-dimensional mit Basis $B_V = (v_1, \dots, v_n)$.*
(3) *$U$ und $V$ sind beide $h_A$-invariant.*
(4) *$K^{m+n}$ ist die direkte Summe von $U$ und $V$, also $U + V = K^{m+n}$ und $U \cap V = \{0\}$.*

*Sei $\varphi_1 := (h_A)|_U$ die Einschränkung von $h_A$ auf $U$, und sei $A_1 := S_{\varphi_1}(B_U, B_U)$.*
*Sei $\varphi_2 := (h_A)|_V$ die Einschränkung von $h_A$ auf $V$, und sei $A_2 := S_{\varphi_2}(B_V, B_V)$.*
*Sei $B = (u_1, \dots, u_m, v_1, \dots, v_n)$.*

*Dann gilt:*

(1) *$B$ ist eine Basis von $K^{m+n}$, und es gilt*

$$S_{h_A}(B,B) = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right).$$

(2) *Für die charakteristischen Polynome gilt $c_A = c_{A_1} \cdot c_{A_2}$.*

*Beweis:* (1) Sei $v \in K^{m+n}$, und sei $(v)_B =: (\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n)$. Dann gilt

$$\begin{aligned}
h_A(v) &= h_A\Big(\sum_{i=1}^m \alpha_i u_i + \sum_{i=1}^n \beta_i v_i\Big) \\
&= \sum_{i=1}^m \alpha_i h_A(u_i) + \sum_{i=1}^n \beta_i h_A(v_i) \\
&= \sum_{i=1}^m \alpha_i \Big(\sum_{j=1}^m A_1(j,i)\, u_j\Big) \;+\; \sum_{i=1}^n \beta_i \Big(\sum_{j=1}^n A_2(j,i)\, v_j\Big) \\
&= \sum_{j=1}^m \Big(\sum_{i=1}^m A_1(j,i)\alpha_i\Big) u_j \;+\; \sum_{j=1}^n \Big(\sum_{i=1}^n A_2(j,i)\beta_i\Big) v_j.
\end{aligned}$$

Also gilt

$$(h_A(v))_B = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

(2) Da $A$ ähnlich zu $S_{h_A}(B,B)$ ist, haben $A$ und $S_{h_A}(B,B)$ das gleiche charakteristische Polynom. Es gilt also

$$\begin{aligned}
c_A &= \det(t * E_{m+n} - \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)) \\
&= \det(\left(\begin{array}{c|c} t * E_m - A_1 & 0 \\ \hline 0 & t * E_n - A_2 \end{array}\right)) \\
&= \det(t * E_m - A_1)\ \det(t * E_n - A_2) \\
&= c_{A_1}\, c_{A_2}.
\end{aligned}$$

■

Eine Zerlegung des charakteristischen Polynoms von $A$ in teilerfremde Faktoren liefert eine Zerlegung in $h_A$-invariante Unterräume:

SATZ 16.11. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n \times n}$, und seien $p, q \in K[t]$ so dass $pq = c_A$ und* $\mathrm{ggT}(p,q) = 1$. *Sei $B_1 := \hat{p}(A)$ und $B_2 := \hat{q}(A)$, und seien $U_1 := N(B_1)$ und $U_2 := N(B_2)$ die Nullräume von $B_1$ und $B_2$. Dann sind $U_1$ und $U_2$ zwei $h_A$-invariante Unterräume von $K^n$, und $K^n$ ist die direkte Summe von $U_1$ und $U_2$.*

*Beweis:* Wegen Lemma 16.7 sind $U_1$ und $U_2$ beide $h_A$-invariant. Wir zeigen nun $K^n \subseteq U_1 + U_2$. Sei dazu $x \in K^n$. Seien $a, b \in K[t]$ mit $ap + pq = 1$. Dann gilt

$$(\widehat{ap})(A) + (\widehat{bq})(A) = \hat{1}(A),$$

also

$$(\widehat{ap})(A) + (\widehat{bq})(A) = E_n,$$

und somit

$$(\widehat{ap})(A) \cdot x + (\widehat{bq})(A) \cdot x = x.$$

Nun gilt $\hat{q}(A) \cdot (\widehat{ap})(A) \cdot x = \hat{a}(A) \cdot \hat{c_A}(A) \cdot x$. Wegen des Satzes von Cayley-Hamilton (Satz 16.4) ist $\hat{c_A}(A) = 0$, also gilt $(\widehat{ap})(A) \cdot x \in U_2$. Genauso gilt $(\widehat{bq})(A) \cdot x \in U_1$.

Wir zeigen nun, dass die Summe direkt ist: Sei $x \in U_1 \cap U_2$. Dann gilt $x = (\widehat{ap})(A) \cdot x + (\widehat{bq})(A) \cdot x = \hat{a}(A) \cdot \hat{p}(A) \cdot x + \hat{b}(A) \cdot \hat{q}(A) \cdot x = \hat{a}(A) \cdot 0 + \hat{b}(A) \cdot 0 = 0$. ∎

ÜBUNGSAUFGABEN 16.12.

(1) Sei

$$A := \begin{pmatrix} -30 & -24 & -29 & -1 \\ -5 & -2 & -5 & 0 \\ 37 & 28 & 36 & 1 \\ -241 & -200 & -250 & -7 \end{pmatrix}.$$

Das charakteristische Polynom von $A$ ist $c_A = (t^2 - 5t + 6)(t + 4)^2$.

(a) Benutzen Sie diese Zerlegung von $c_A$, um $h_A$-invariante Unterräume $U_1$ und $U_2$ mit $\mathbb{R}^4 = U_1 \dotplus U_2$, $U_1 \neq \{0\}$, $U_2 \neq \{0\}$ zu finden.

(b) Finden Sie Basen $B_1, B_2$ von $U_1$ und $U_2$.

(c) Bilden Sie aus den Vektoren von $B_1$ und $B_2$ eine Basis $B$ von $\mathbb{R}^4$, und berechnen Sie $S_{h_A}(B, B)$.

(2) Sei $A \in K^{n\times n}$, und seien $p, q \in K[t]$ so, dass $(\widehat{pq})(A) = 0$ und $\mathrm{ggT}(p, q) = 1$. Zeigen Sie, dass der Spaltenraum von $\hat{p}(A)$ gleich dem Nullraum von $\hat{q}(A)$ ist.

(3) Sei $A$ eine $n \times n$-Matrix mit $A^2 = A$. Zeigen Sie, dass $K^n = N(A) \dotplus S(A)$.

(4) Sei $A \in \mathbb{R}^3$ gegeben durch $A = \begin{pmatrix} -4 & -2 & 0 \\ -2 & 4 & -4 \\ -1 & 5 & -6 \end{pmatrix}$. Es gilt $c_A = (t - 2)(t + 4)^2$.

(a) Bestimmen Sie $U_1 = N(A - 2 * E)$ und $U_2 = N((A + 4 * E)^2)$.

(b) Finden Sie aus der Darstellung $a(t) \cdot (t - 2) + b(t) \cdot (t + 4)^2 = 1$ Matrizen $A_1, A_2 \in \mathbb{R}^{3\times 3}$, sodass für alle $x \in \mathbb{R}^3$ gilt: $A_1 \cdot x \in U_1$, $A_2 \cdot x \in U_2$, $A_1 \cdot x + A_2 \cdot x = x$. (*Hinweis:* Verwenden Sie Mathematica und `PolynomialExtendedGCD[p, q, t]`.

(c) Berechnen Sie Basen für den Spaltenraum von $A_1$ und von $A_2$.

(5) Sei $A \in K^{n\times n}$, und seien $p, q \in K[t]$ mit $p \mid q$. Zeigen Sie, dass dann $N(\hat{q}(A)) \subseteq N(\hat{p}(A))$. Gilt auch die Umkehrung?

## 3. Die Haupträume einer Matrix

DEFINITION 16.13. Sei $A \in K^{n\times n}$, sei $\lambda \in K$ ein Eigenwert von $A$, und sei $e$ die algebraische Vielfachheit von $\lambda$ in $A$. Dann ist der *Hauptraum von $A$ bezüglich $\lambda$*, $U(A, \lambda)$, definiert als der Nullraum von $(A - \lambda * E_n)^e$.

Wir untersuchen jetzt diese Haupträume, und dazu als erstes nilpotente Matrizen.

DEFINITION 16.14. Sei $K$ ein Körper, $n \in \mathbb{N}$, und sei $A \in K^{n\times n}$. Die Matrix $A$ ist *nilpotent*, wenn es ein $k \in \mathbb{N}$ gibt, sodass $A^k = 0$.

ÜBUNGSAUFGABEN 16.15.
In den folgenden Beispielen ist $K$ ein Körper, $n \in \mathbb{N}$, und $A, B \in K^{n\times n}$.

(1) Zeigen Sie, dass für alle $m \in \mathbb{N}_0$ gilt: $N(A^m) \subseteq N(A^{m+1})$.
(2) Zeigen Sie: Wenn $m \in \mathbb{N}_0$ so ist, dass $N(A^m) = N(A^{m+1})$, so gilt $N(A^m) = N(A^j)$ für alle $j \geq m$.
(3) Zeigen Sie: wenn $A$ nilpotent ist, so ist $A^n = 0$. (Die Zahl $n$ ist das Format der Matrix).
(4) Sei $x \in K^n$, und sei $k \in \mathbb{N}$ so, dass $A^k \cdot x = 0$. Sei $c_A(t) = t^e \cdot r(t)$, sodass $t$ das Polynom $r(t)$ nicht teilt. Zeigen Sie, dass $A^e \cdot x = 0$. *Hinweis:* Für $p(t) := t^e \cdot r(t)$ und für $q(t) := t^k$ gilt $\hat{p}(A) \cdot x = \hat{q}(A) \cdot x = 0$. Was bedeutet das für $h = \mathrm{ggT}(p, q)$?
(5) Seien $A, B$ nilpotent und so, dass $A \cdot B = B \cdot A$. Zeigen Sie, dass $A + B$ nilpotent ist.
(6) Sei $A$ nilpotent. Zeigen Sie, dass $E_n - A$ invertierbar ist.

Wir berechnen nun das charakteristische Polynom einer nilpotenten Matrix.

LEMMA 16.16. *Sei $m \in \mathbb{N}$ und sei $A \in K^{m\times m}$. Wir nehmen an, dass es $k \in \mathbb{N}$ gibt, sodass $A^k = 0$. Dann gilt $c_A = t^m$.*

*Beweis:* Sei $E := E_m$, und sei $k$ so, dass $A^k = 0$. Es gilt

$$\begin{aligned}(t * E - A) \cdot (\sum_{i=0}^{k-1} t^{(k-1)-i} * A^i) &= \sum_{i=0}^{k-1} t^{k-i} * A^i - \sum_{i=0}^{k-1} t^{k-(i+1)} * A^{i+1} \\ &= t^k * A^0 - t^0 * A^k \\ &= t^k * E.\end{aligned}$$

Also gilt $\det(t * E - A) \cdot \det(\sum_{i=0}^{k-1} t^{(k-1)-i} * A^i) = t^{km}$. Es gilt also $c_A \mid t^{km}$. Wegen der Eindeutigkeit der Zerlegung von Polynomen in irreduzible Faktoren (Satz 13.32) ist jeder Teiler von $t^{km}$ von der Form $\alpha t^r$ mit $\alpha \in K \setminus \{0\}$ und $r \in \{0, \ldots, km\}$. Da $c_A$ normiert ist und Grad $m$ hat, gilt also $c_A = t^m$. ■

KOROLLAR 16.17. *Sei $K$ ein Körper, $m \in \mathbb{N}$, $A \in K^{m\times m}$, und sei $\lambda \in K$ so, dass $A - \lambda * E$ nilpotent ist. Dann gilt $c_A = (t - \lambda)^m$.*

*Beweis:* $c_A = \det(t * E - A) = \det((t - \lambda) * E - (A - \lambda * E)) = c_{A-\lambda * E}(t - \lambda) = (t - \lambda)^m$. ■

Nun können wir die Dimension der Haupträume bestimmen.

SATZ 16.18. *Sei $K$ ein Körper, $n \in \mathbb{N}$, und sei $A \in K^{n\times n}$. Sei $\lambda \in K$ ein Eigenwert von $A$ mit algebraischer Vielfachheit $e$, und sei $U := N((A - \lambda * E)^e)$ der Hauptraum von $A$ zum Eigenwert $\lambda$. Sei $h := h_A$, also $h(x) = A \cdot x$ für alle $x \in K^n$. Dann gilt:*

(1) *$U$ ist $h$-invariant.*
(2) *Es gibt einen $h$-invarianten Unterraum $V$ von $K^n$, sodass $K^n = U \dotplus V$.*
(3) $\dim(U) = e$.
(4) *Sei $B_U = (u_1, \ldots, u_e)$ eine Basis von $U$, sei $\varphi := h|_U$, und sei $A_1 := S_\varphi(B_U, B_U)$. Dann gilt $c_{A_1} = (t - \lambda)^e$.*

*Beweis:* (1) Für $p = (t-\lambda)^e$ gilt $\hat{p}(A) = (A-\lambda*E_n)^e$. Wegen Lemma 16.7 ist $U = N(\hat{p}(A))$ $h$-invariant.

(2) Wir schreiben $c_A$ als $(t-\lambda)^e \cdot q$, sodass $t-\lambda$ das Polynom $q$ nicht teilt. Wegen Satz 16.11 gilt $K^n = U \dotplus N(\hat{q}(A))$, also leistet $V := N(\hat{q}(A))$ das Gewünschte.

(3) Sei $d := \dim(U)$, und sei $B_U = (u_1, \dots, u_d)$ eine Basis von $U$. Sei $B_V = (v_1, \dots, v_{n-d})$ eine Basis von $V$. Sei $A_1 := S_{h|_U}(B_U, B_U)$ und $A_2 := S_{h|_V}(B_V, B_V)$. Wegen Satz 16.10 gilt

$$c_A = c_{A_1} \cdot c_{A_2}. \tag{16.2}$$

Wir zeigen als nächstes:

$$(A_1 - \lambda * E_d)^e = 0. \tag{16.3}$$

Dazu beobachten wir, dass für $\varphi := h|_U - \lambda * \mathrm{id}_U$ gilt: $A_1 - \lambda * E_d = S_\varphi(B_U, B_U)$. Wir wissen, dass für alle $x \in U$ gilt:

$$\begin{aligned}
(A_1 - \lambda * E_d)^e \cdot (x)_{B_U} &= (S_\varphi(B_U, B_U))^e \cdot (x)_{B_U} \\
&= \Big(\underbrace{\varphi \circ \varphi \circ \cdots \circ \varphi}_{e}\,(x)\Big)_{B_U} \\
&= \Big(\underbrace{(h - \lambda * \mathrm{id}_{K^n}) \circ \cdots \circ (h - \lambda * \mathrm{id}_{K^n})}_{e}\,(x)\Big)_{B_U} \\
&= \Big((A - \lambda * E_n)^e \cdot x\Big)_{B_U}.
\end{aligned}$$

Da $x \in N((A - \lambda * E_n)^e)$, gilt insgesamt $(A_1 - \lambda * E_d)^e \cdot (x)_{B_U} = (0)_{B_U} = 0$. Also gilt für alle $y \in K^d$, dass $(A_1 - \lambda * E_d)^e \cdot y = 0$, und somit gilt $(A_1 - \lambda * E_d)^e = 0$. Das beweist (16.3). Wegen Korollar 16.17 gilt also

$$c_{A_1} = (t-\lambda)^d. \tag{16.4}$$

Wegen (16.2) gilt $(t-\lambda)^d | c_A$. Da die Vielfachheit der Nullstelle $\lambda$ in $c_A$ gleich $e$ ist, gilt $d \le e$.

Wir zeigen nun $d = e$. Nehmen wir dazu an, dass $d < e$. Dann gilt wegen (16.2), dass $t-\lambda$ auch $c_{A_2}$ teilt. Folglich ist $\lambda$ ein Eigenwert von $A_2$. Sei $w = (w_1, \dots, w_{n-d}) \in K^{n-d}$ ein Eigenvektor von $A_2$, und sei $x \in V$ so, dass $(x)_{B_V} = w$. Dann gilt $(A \cdot x)_{B_V} = (h(x))_{B_V} = S_{h|_V}(B_V, B_V) \cdot w = A_2 \cdot w = \lambda * w = \lambda * (x)_{B_V} = (\lambda * x)_{B_V}$. Also gilt $A \cdot x = \lambda * x$, $x$ ist folglich ein Eigenvektor von $A$ zum Eigenwert $\lambda$. Daraus ergibt sich $x \in U$. Insgesamt liegt $x$ also in $U \cap V$. Wegen (2) gilt dann $x = 0$, im Widerspruch dazu, dass $w$ als Eigenvektor $\neq 0$, und somit auch $x \neq 0$ ist. Es gilt also $d = e$. Somit haben wir (3) bewiesen.

(4) folgt nun aus (3) und Gleichung (16.4). ■

SATZ 16.19. *Sei $K$ ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$. Wir nehmen an, dass $c_A$ über $K$ in Linearfaktoren zerfällt. Seien $r \in \mathbb{N}$, $\lambda_1, \dots, \lambda_r \in K$ und $e_1, \dots, e_r \in \mathbb{N}$ so, dass*

$$c_A = \prod_{i=1}^{r} (t - \lambda_i)^{e_i},$$

*und die* $\lambda_i$ *alle verschieden voneinander sind. Sei* $U_i := N((A - \lambda_i * E)^{e_i})$ *der zu* $\lambda_i$ *gehörende Hauptraum von A.*

*Dann ist* $K^n$ *die direkte Summe der* $h_A$*-invarianten Unterräume* $U_1, \ldots, U_r$*.*

*Beweis:* Wegen Satz 16.18 ist jedes $U_i$ $h_A$-invariant. Sei nun $U := U_1 + \cdots + U_r$. Wir zeigen, dass diese Summe sogar direkt ist. Sei dazu $i \in \{1, \ldots, r\}$, und sei $u \in U_i \cap \Big(\sum_{\substack{j=1\\j\neq i}}^{r} U_j\Big)$.

Wenn $j \neq i$, so gilt $U_j = N((A - \lambda_j)^{e_j})$. Sei $q \in K[t]$ definiert durch $q := c_A/(t - \lambda_i)^{e_i}$. Wegen $(t - \lambda_j)^{e_j} \mid q$ gilt $U_j \subseteq N(\hat{q}(A))$. Insgesamt gilt daher $\sum_{\substack{j=1\\j\neq i}}^{r} U_j \subseteq N(\hat{q}(A))$.

Wegen Satz 16.11 gilt $U_i \cap N(\hat{q}(A)) = \{0\}$. Also gilt $u = 0$.

$U$ ist also die direkte Summe $U_1 \dotplus \cdots \dotplus U_r$. Nun gilt $\dim(U) = \sum_{i=1}^{r} \dim(U_i)$. Nach Satz 16.18 gilt also $\dim(U) = \sum_{i=1}^{r} e_i = n$. Folglich gilt also $U = K^n$. ■

## 4. Nilpotente Endomorphismen

DEFINITION 16.20. Sei $V$ ein Vektorraum, und sei $\varphi : V \to V$ ein Endomorphismus. Der Endomorphismus $\varphi$ ist *nilpotent*, wenn es ein $k \in \mathbb{N}$ gibt, sodass für alle $v \in V$ $\varphi^k(v) = \underbrace{\varphi \circ \cdots \circ \varphi}_{k}(v) = 0$.

LEMMA 16.21. *Sei K ein Körper,* $n \in \mathbb{N}$*, sei* $A \in K^{n\times n}$*, und sei* $\lambda$ *ein Eigenwert von A mit algebraischer Vielfachheit e. Sei* $U := N((A - \lambda * E_n)^e)$*. Sei* $\varphi : U \to U$ *definiert durch* $\varphi(u) := A \cdot u - \lambda * u$*, also* $\varphi = h_A|_U - \lambda * \mathrm{id}_U$*. Dann ist* $\varphi$ *nilpotent auf U.*

*Beweis:* Sei $u \in U$. Es gilt $\varphi^e(u) = (A - \lambda * E_n)^e \cdot u = 0$, weil ja $u \in N((A - \lambda * E_n)^e)$. ■

Es folgt nun ein typisches Beispiel für einen nilpotenten Endomorphismus. Sei $e \in \mathbb{N}$ und $B = (b_1, \ldots, b_e)$ eine Basis für den Vektorraum $U$, und sei $\varphi$ eine Abbildung mit $\varphi(b_e) = b_{e-1}$, $\varphi(b_{e-1}) = b_{e-2}$, …, $\varphi(b_2) = b_1$, $\varphi(b_1) = 0$. Dann gilt $\varphi^e = 0$ und $\varphi^{e-1} \neq 0$. Die Matrixdarstellung von $\varphi$ hat folgende Form:

$$S_\varphi(B,B) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

DEFINITION 16.22. Sei $K$ ein Körper, $m \in \mathbb{N}$, und sei $\lambda \in K$. Die $m \times m$-Matrix

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

mit $J[i, i] = \lambda$ für $i \in \{1, \dots, m\}$ und $J[i, i+1] = 1$ für $i \in \{1, \dots, m-1\}$ heißt *Jordankästchen zu $\lambda$ vom Format $m$.* Sie wird mit $J(\lambda, m)$ abgekürzt.

DEFINITION 16.23. Sei $V$ ein Vektorraum über dem Körper $K$, und sei $U$ ein Unterraum von $V$. Die Folge $(v_1, \dots, v_m)$ von Elementen aus $V$ ist *eine Basis von $V$ modulo $U$*, wenn $(v_1 + U, \dots, v_m + U)$ eine Basis des Faktorraums $V/U$ ist.

LEMMA 16.24. *Sei $W$ ein Vektorraum über dem Körper $K$, und seien $U, V$ Unterräume von $W$ mit $U \subseteq V$. Wir nehmen an, dass $(w_1, \dots, w_m)$ eine Basis von $W$ modulo $V$ ist, und $(v_1, \dots, v_n)$ eine Basis von $V$ modulo $U$. Dann ist $(w_1, \dots, w_m, v_1, \dots, v_n)$ eine Basis von $W$ modulo $U$.*

*Beweisskizze:* ähnlich wie Satz 15.22. ■

ÜBUNGSAUFGABEN 16.25.

(1) Seien $U_0 \subseteq U_1 \subseteq \cdots \subseteq U_n = U$ Unterräume von $U$, und sei für jedes $i \in \{1, \dots, n\}$ die Folge $(b_{i,1}, \dots, b_{i,j(i)})$ eine Basis von $U_i$ modulo $U_{i-1}$. Zeigen Sie, dass $(b_{1,1}, \dots, b_{1,j(1)}, b_{2,1}, \dots, b_{2,j(2)}, \dots, b_{n,1}, \dots, b_{n,j(n)})$ eine Basis von $U_n$ modulo $U_0$ ist. *Hinweis:* Lemma 16.24 und Induktion nach $n$.

SATZ 16.26 (Normalform für nilpotente Endomorphismen). *Sei $n \in \mathbb{N}$, sei $U$ ein $n$-dimensionaler Vektorraum über $K$, und sei $\varphi : U \to U$ ein nilpotenter Endomorphismus. Dann gibt es eine Basis $B$ von $U$, ein $d \in \mathbb{N}$, und $m_1, \dots, m_d \in \mathbb{N}$, sodass $m_1 + \cdots + m_d = n$ und*

$$S_\varphi(B, B) = J(0, m_1) \oplus J(0, m_2) \oplus \cdots \oplus J(0, m_d).$$

*Beweisskizze:* Sei $U_0 := \{0\}$, und sei für $i \in \{1, \dots, n\}$ der Unterraum $U_i$ definiert durch

$$U_i := \ker(\varphi^i).$$

Dann gilt wegen Übungsbeispiel 16.15 (3), dass $U_n = U$, und $U_0 \subseteq U_1 \subseteq U_2 \subseteq \cdots \subseteq U_n = U$.

Wir bilden nun für jedes $i \in \{n, n-1, \dots, 0\}$ eine Basis von $U_i$ auf folgende Weise: Zuerst wählen wir $a(n) \in \mathbb{N}_0$ und

$$(x_{n,1}, \dots, x_{n,a(n)})$$

als Basis von $U_n$ modulo $U_{n-1}$.

Als nächstes konstruieren wir eine Basis von $U_{n-1}$ modulo $U_{n-2}$ folgendermaßen: Die Vektoren $\varphi(x_{n,1}), \ldots, \varphi(x_{n,a(n)})$ liegen in $U_{n-1}$, weil für alle $k \in \{1, \ldots, a(n)\}$ die Gleichung $\varphi^{n-1}(\varphi(x_{n,k})) = \varphi^n(x_{n,k}) = 0$ gilt. Wir zeigen als nächstes, dass

$$\big(\varphi(x_{n,1}) + U_{n-2}, \ldots, \varphi(x_{n,a(n)}) + U_{n-2}\big)$$

eine linear unabhängige Folge aus $U_{n-1}/U_{n-2}$ ist. Sei dazu

$$\sum_{j=1}^{a(n)} \lambda_j * \big(\varphi(x_{n,j}) + U_{n-2}\big) = 0 + U_{n-2}.$$

Dann gilt

$$\sum_{j=1}^{a(n)} \lambda_j * \varphi(x_{n,j}) \in U_{n-2},$$

also wegen $\varphi^{n-2}[U_{n-2}] = \{0\}$

$$\varphi^{n-1}\Big(\sum_{j=1}^{a(n)} \lambda_j * x_{n,j}\Big) = 0.$$

Folglich gilt

$$\sum_{j=1}^{a(n)} \lambda_j * x_{n,j} \in U_{n-1}.$$

Da $(x_{n,1}, \ldots, x_{n,a(n)})$ eine Basis von $U_n$ modulo $U_{n-1}$ ist, gilt also $\lambda_1 = \cdots = \lambda_{a(n)} = 0$. Da also $\big(\varphi(x_{n,1}) + U_{n-2}, \ldots, \varphi(x_{n,a(n)}) + U_{n-2}\big)$ eine linear unabhängige Folge aus $U_{n-1}/U_{n-2}$ ist, gibt es $a(n-1) \in \mathbb{N}_0$ und $(x_{n-1,1}, \ldots, x_{n-1,a(n-1)})$ aus $U_{n-1}$, sodass

$$\big(\varphi(x_{n,1}), \ldots, \varphi(x_{n,a(n)}),\ x_{n-1,1}, \ldots, x_{n-1,a(n-1)}\big)$$

eine Basis von $U_{n-1}$ modulo $U_{n-2}$ ist.

Nun vervollständigen wir

$$\big(\varphi^2(x_{n,1}), \ldots, \varphi^2(x_{n,a(n)}),\ \varphi(x_{n-1,1}), \ldots, \varphi(x_{n-1,a(n-1)})\big)$$

mit $x_{n-2,1}, \ldots, x_{n-2,a(n-2)}$ zu einer Basis von $U_{n-2}$ modulo $U_{n-3}$. Durch Fortführung dieser Konstruktion erhalten wir Basen für jedes $U_i$ modulo $U_{i-1}$ mit $i \in \{1, \ldots, n\}$.

Mit Lemma 16.24 und Induktion kann man nun zeigen, dass man durch Vereinigen aller dieser Basen eine Basis $B$ von $U$ erhält, die so aussieht:

$$\begin{array}{rclrclc}
x_{n,1}, & \ldots, & x_{n,a(n)}, & & & & \\
\varphi(x_{n,1}), & \ldots, & \varphi(x_{n,a(n)}), & x_{n-1,1}, & \ldots, & x_{n-1,a(n-1)}, & \\
\varphi^2(x_{n,1}), & \ldots, & \varphi^2(x_{n,a(n)}), & \varphi(x_{n-1,1}), & \ldots, & \varphi(x_{n-1,a(n-1)}), & x_{n-2,1}, \ldots, x_{n-2,a(n-2)} \\
 & & & & & & \vdots \\
\varphi^{n-1}(x_{n,1}), & \ldots, & \varphi^{n-1}(x_{n,a(n)}), & \varphi^{n-2}(x_{n-1,1}), & \ldots, & \varphi^{n-2}(x_{n-1,a(n-1)}), & \ldots\ldots \quad x_{1,1}, \ldots, x_{1,a(1)}.
\end{array}$$

Wenn wir diese Basis nun so ordnen, dass $B =$

$$\begin{aligned}
B = \big(&\varphi^{n-1}(x_{n,1}), \ldots, \varphi(x_{n,1}), x_{n,1},\\
&\ldots,\\
&\varphi^{n-1}(x_{n,a(n)}), \ldots, \varphi(x_{n,a(n)}), x_{n,a(n)},\\
&\varphi^{n-2}(x_{n-1,1}), \ldots, \varphi(x_{n-1,1}), x_{n-1,1},\\
&\ldots,\\
&\varphi^{n-2}(x_{n-1,a(n-1)}), \ldots, \varphi(x_{n-1,a(n-1)}), x_{n-1,a(n-1)},\\
&\ldots,\\
&x_{1,1},\\
&\ldots,\\
&x_{1,a(1)}\big),
\end{aligned}$$

so gilt

$$S_\varphi(B,B) = \underbrace{J(0,n) \oplus \cdots \oplus J(0,n)}_{a(n)} \oplus \underbrace{J(0,n-1) \oplus \cdots \oplus J(0,n-1)}_{a(n-1)} \oplus \cdots \oplus \underbrace{J(0,1) \oplus \cdots \oplus J(0,1)}_{a(1)}.$$

Somit hat $S_\varphi(B,B)$ die gewünschte Form. ■

ÜBUNGSAUFGABEN 16.27.

(1) Sei $\varphi$ ein Endomorphismus von $V \to V$, und sei $B$ eine Basis von $V$ mit $S_\varphi(B,B) = J(0,n)$. Berechnen Sie für jedes $i \in \mathbb{N}_0$ eine Basis von $\ker(\varphi^i)$, und die Dimension von $\ker(\varphi^i)$.

(2) Seien $m_1, \ldots, m_d \in \mathbb{N}$, sei $n := \sum_{i=1}^d m_i$, und sei $\varphi : \mathbb{R}^n \to \mathbb{R}^n$, $\varphi(x) := (J(0,m_1) \oplus \cdots \oplus J(0,m_d)) \cdot x$. Berechnen Sie $\dim(\ker(\varphi))$, und für alle $i \in \mathbb{N}$ die Dimension von $\ker(\varphi^i)$. Verwenden Sie zum Ausdrücken des Ergebnisses die Zahl $a(k) := |\{j \in \{1,\ldots,d\} \mid m_d = k\}|$. Die Zahl $a(k)$ gibt also die Anzahl der Jordankästchen der Dimension $k$ an.

(3) (Mathematica) Sei $A = \begin{pmatrix} 18 & -22 & 1 & -11 \\ 18 & -22 & 1 & -11 \\ 8 & -9 & 2 & -8 \\ -4 & 5 & 0 & 2 \end{pmatrix}$.
   (a) Bestimmen Sie $N(A^3)$, $N(A^2)$, $N(A)$.
   (b) Bestimmen Sie eine Basis $B_3$ von $N(A^3)$ modulo $N(A^2)$.
   (c) Vervollständigen Sie $\{A \cdot v \mid v \in B_3\}$ zu einer Basis $B_2$ von $N(A^2)$ modulo $N(A)$.
   (d) Vervollständigen Sie $\{A \cdot w \mid w \in B_2\}$ zu einer Basis $B_1$ von $N(A)$.
   (e) Ordnen Sie die Vektoren in $B_1 \cup B_2 \cup B_3$ passend zu einer Basis $B$ an, sodass $S_{h_A}(B,B)$ eine $\oplus$-Summe von Jordankästchen ist.

(4) (Mathematica) Sei $A = \begin{pmatrix} 73 & -1 & 0 & 0 & 0 & 0 & 3 & 10 \\ -34 & -2 & -1 & 0 & 0 & 2 & -2 & -6 \\ 1221 & -9 & 3 & 0 & 0 & -6 & 53 & 172 \\ -1252 & 10 & -1 & 0 & 1 & 5 & -54 & -176 \\ 363 & -3 & 1 & 0 & 0 & -2 & 14 & 50 \\ 1355 & -11 & 3 & 0 & 0 & -6 & 58 & 190 \\ 292 & -4 & 0 & 0 & 0 & 0 & 12 & 40 \\ -584 & 8 & 0 & 0 & 0 & 0 & -24 & -80 \end{pmatrix}$. Finden Sie eine Basis $B$ des $\mathbb{R}^8$, sodass $S_{h_A}(B,B)$ die Summe von Jordankästchen ist. *Hinweis:* $A^3 = 0$.

## 5. Die Jordansche Normalform

DEFINITION 16.28. Sei $J$ eine $n \times n$-Matrix über dem Körper $K$. $J$ ist in *Jordanscher Normalform*, wenn es $k \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_k \in K$ (nicht notwendigerweise verschieden) und $m_1, \ldots, m_k \in \mathbb{N}$ gibt, sodass

$$J = J(\alpha_1, m_1) \oplus \cdots \oplus J(\alpha_k, m_k).$$

SATZ 16.29 (Existenz einer ähnlichen Matrix in Jordanscher Normalform). *Sei $A$ eine $n \times n$-Matrix über dem Körper $K$. Wir nehmen an, dass das charakteristische Polynom von $A$ über $K$ in lauter Linearfaktoren zerfällt. Dann gibt es eine Matrix $J \in K^{n\times n}$ in Jordanscher Normalform und eine Matrix $P \in \mathrm{GL}(n, K)$, sodass $A = P \cdot J \cdot P^{-1}$.*

*Beweis:* Sei $c_A = \prod_{i=1}^{r}(t - \lambda_i)^{e_i}$ mit $\lambda_1, \ldots, \lambda_r$ paarweise verschieden. Wegen Satz 16.19 ist $K^n$ die direkte Summe der Haupträume $U_1, \ldots, U_r$ von $A$.

Wir finden nun für jeden Hauptraum $U_i$ eine Basis $B_i$, sodass für $\varphi := h_A|_{U_i}$ die Matrix $S_\varphi(B_i, B_i)$ eine $e_i \times e_i$-Matrix in Jordanscher Normalform ist. Sei dazu $\psi := \varphi - \lambda * \mathrm{id}_{U_i}$. Wegen Lemma 16.21 ist $\psi$ nilpotent auf $U_i$. Also gibt es wegen Satz 16.26 eine Basis $B_i$ von $U_i$, $d \in \mathbb{N}$ und $m_1, \ldots, m_d \in \mathbb{N}$, sodass

$$S_\psi(B_i, B_i) = J(0, m_1) \oplus \cdots \oplus J(0, m_d).$$

Also gilt

$$S_\varphi(B_i, B_i) = J(\lambda_i, m_1) \oplus \cdots \oplus J(\lambda_i, m_d).$$

Sei $J_i := S_\varphi(B_i, B_i)$.

Satz 15.35 erlaubt nun, die Basen $B_1, B_2, \ldots, B_r$ zu einer Basis $B$ von $K^n$ zusammenzuhängen. Dann gilt $S_{h_A}(B, B) = J_1 \oplus \cdots \oplus J_r$. Wenn wir in die Spalten der Matrix $P$ die Vektoren aus $B$ schreiben, so gilt $A = P \cdot S_{h_A}(B, B) \cdot P^{-1}$. ∎

ÜBUNGSAUFGABEN 16.30.

(1) Finden Sie eine zu $A = \begin{pmatrix} -4 & -2 & 0 \\ -2 & 4 & -4 \\ -1 & 5 & -6 \end{pmatrix}$ ähnliche Matrix in Jordanscher Normalform. Ist $A$ diagonalisierbar?

(2) Die Folge $(x_n)_{n\in\mathbb{N}}$ ist gegeben durch $x_1 = x_2 = x_3 = 1$, $x_{n+3} = 6x_{n+2} - 12x_{n+1} + 8x_n$ für $n \geq 1$. Für $A := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \end{pmatrix}$ gilt also also $\left(\begin{smallmatrix} x_{n+1} \\ x_{n+2} \\ x_{n+3} \end{smallmatrix}\right) = A^n \cdot \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right)$.
   (a) Bestimmen Sie eine Matrix $J$ in Jordanscher Normalform und eine Matrix $B$, sodass $B^{-1} \cdot A \cdot B = J$.
   (b) Bestimmen Sie einen Ausdruck für $J^n$.
   (c) (Mathematica) Bestimmen Sie daraus einen Ausdruck für $A^n$.
   (d) (Mathematica) Bestimmen Sie einen Ausdruck für den ersten Eintrag von $A^{m-1} \cdot \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$, also für $x_m$. (*Lösung:* $2^{m-4}\left(m^2 - 7m + 14\right)$.)

(3) Finden Sie eine zu $A = \begin{pmatrix} -30 & -24 & -29 & -1 \\ -5 & -2 & -5 & 0 \\ 37 & 28 & 36 & 1 \\ -241 & -200 & -250 & -7 \end{pmatrix}$ ähnliche Matrix in Jordanscher Normalform. ($c_A = (t - 3)(t - 2)(t + 4)^2$).

SATZ 16.31 (Eindeutigkeit der Jordanschen Normalform). *Seien $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_l \in K$ (nicht notwendigerweise verschieden), und seien $m_1, \dots, m_k, n_1, \dots, n_l \in \mathbb{N}$ so, dass die Matrizen $A$ und $B$, die durch*

$$\begin{aligned} A &= J(\alpha_1, m_1) \oplus \cdots \oplus J(\alpha_k, m_k), \\ B &= J(\beta_1, n_1) \oplus \cdots \oplus J(\beta_l, n_l) \end{aligned}$$

*definiert sind, ähnlich sind. Dann gilt $k = l$, und es gibt eine bijektive Abbildung $\pi : \{1, \dots, k\} \to \{1, \dots, l\}$, sodass für alle $i \in \{1, \dots, k\}$ gilt: $\alpha_i = \beta_{\pi(i)}$ und $m_i = n_{\pi(i)}$.*

*Beweis:* Da $A$ und $B$ ähnlich sind, gibt es wegen Satz 14.5 einen Endomorphismus $\varphi : K^n \to K^n$ und Basen $A'$, $B'$ von $K^n$, sodass $A = S_\varphi(A', A')$ und $B = S_\varphi(B', B')$. Sei $F := S_\varphi(E, E)$.

Sei $c_F$ das charakteristische Polynom von $F$. Es gilt $c_F = c_A = c_B$, also

$$\prod_{i=1}^{k} (t - \alpha_i)^{m_i} = \prod_{j=1}^{l} (t - \beta_j)^{n_j}.$$

Seien $\lambda_1, \dots, \lambda_r$ die verschiedenen Eigenwerte der Matrix $F$. Es gilt dann wegen der Eindeutigkeit der Faktorisierung des Polynoms $c_F$ (Satz 13.32) für jedes $\lambda \in \{\lambda_1, \dots, \lambda_r\}$:

$$\sum_{\substack{i=1 \\ \alpha_i = \lambda}}^{k} m_i = \sum_{\substack{j=1 \\ \beta_j = \lambda}}^{k} n_j = e,$$

wobei $e$ die algebraische Vielfachheit von $\lambda$ in $F$ ist. Wir werden nun eine Bijektion $\pi_\lambda : \{i \in \{1, \dots, k\} \mid \alpha_i = \lambda\} \to \{j \in \{1, \dots, l\} \mid \beta_j = \lambda\}$ finden. Sei dazu $\{i_1, \dots, i_s\} := \{i \in \{1, \dots, k\} \mid \alpha_i = \lambda\}$, und $\{j_1, \dots, j_t\} := \{j \in \{1, \dots, l\} \mid \beta_j = \lambda\}$.

Sei $U$ der Hauptraum von $F$ zu $\lambda$. Durch Auswahl einer passenden Teilfolge der Basis $A'$ von $K^n$ erhalten wir eine Basis $A''$ von $U$, sodass

$$S_{\varphi|_U}(A'', A'') = J(\lambda, m_{i_1}) \oplus \cdots \oplus J(\lambda, m_{i_s}).$$

Ebenso erhalten wir durch Auswahl einer passenden Teilfolge von $B'$ eine Basis $B''$ von $U$, sodass

$$S_{\varphi|_U}(B'', B'') = J(\lambda, n_{j_1}) \oplus \cdots \oplus J(\lambda, m_{j_t}).$$

Für $u \in \mathbb{N}$ sei $a(u)$ die Anzahl der $v \in \{1, \dots, s\}$ mit $m_{i_v} = u$. Genauso sei $b(u)$ die Anzahl der $v \in \{1, \dots, t\}$ mit $n_{j_v} = u$.

Wir berechnen nun für $i \in \{1, \dots, n\}$ die Dimension von $\ker((\varphi|_U - \lambda * \mathrm{id}_U)^i)$. Sei $\psi := \varphi|_U - \lambda * \mathrm{id}_U$. Es gilt

$$\dim(\ker(\psi^i)) = \sum_{j=1}^{i} j \cdot a(j) + \sum_{j=i+1}^{n} i \cdot a(j).$$

Es gilt also

$$\begin{aligned}\dim(\ker(\psi^i)) - \dim(\ker(\psi^{i-1})) &= \sum_{j=1}^{i} j \cdot a(j) + \sum_{j=i+1}^{n} i \cdot a(j) - \sum_{j=1}^{i-1} j \cdot a(j) - \sum_{j=i}^{n} (i-1) \cdot a(j) \\ &= \sum_{j=i}^{n} a(j).\end{aligned}$$

Insgesamt gilt also

$$\Big(\dim(\ker(\psi^i)) - \dim(\ker(\psi^{i-1}))\Big) - \Big(\dim(\ker(\psi^{i+1})) - \dim(\ker(\psi^i))\Big) = a(i).$$

Genauso gilt aber auch $b(i) = \Big(\dim(\ker(\psi^i)) - \dim(\ker(\psi^{i-1}))\Big) - \Big(\dim(\ker(\psi^{i+1})) - \dim(\ker(\psi^i))\Big)$. Insgesamt gilt also $a(i) = b(i)$ für alle $i \in \{1, \ldots, n\}$. Somit sind die Folgen $(m_{i_1}, \ldots, m_{i_s})$ und $(n_{j_1}, \ldots, n_{j_t})$ bis auf die Reihenfolge gleich. Daraus erhalten wir die Bijektion $\pi_\lambda$.

Durch Vereinigen der $\pi_\lambda$ erhalten wir schließlich $\pi$. ■

## 6. Das Minimalpolynom einer Matrix

DEFINITION 16.32. Sei $n \in \mathbb{N}$, $K$ ein Körper, und sei $A \in K^{n \times n}$. Jedes Polynom von minimalem Grad in

$$M := \{p \in K[t] \mid \hat{p}(A) = 0, \deg(p) \geq 0, p \text{ ist normiert}\}$$

ist ein *Minimalpolynom von $A$ über $K$*.

SATZ 16.33. *Sei $n \in \mathbb{N}$, $K$ ein Körper, und sei $A \in K^{n \times n}$. Sei $q \in K[t]$ so, dass $\hat{q}(A) = 0$, und sei $m$ ein Minimalpolynom von $A$ über $K$. Dann gilt $m | q$.*

KOROLLAR 16.34. *Sei $n \in \mathbb{N}$, $K$ ein Körper, und sei $A \in K^{n \times n}$. Dann besitzt $A$ genau ein Minimalpolynom über $K$. Sei $m_A$ das Minimalpolynom von $A$ über $K$, und sei $c_A$ das charakteristische Polynom. Dann gilt $m_A | c_A$.*

SATZ 16.35. *Sei $n \in \mathbb{N}$, $K$ ein Körper, und sei $A \in K^{n \times n}$. Dann sind äquivalent:*

(1) *$A$ ist über $K$ diagonalisierbar.*
(2) *Es gibt $r \in \mathbb{N}$ und paarweise verschiedene $\lambda_1, \ldots, \lambda_r \in K$, sodass $m_A = \prod_{i=1}^{r} (t - \lambda_i)$.*

*Beweisskizze:* (1)⇒(2): Sei $A = P \cdot D \cdot P^{-1}$ mit $P \in \mathrm{GL}(n, K)$. Sei $r$ die Anzahl der verschiedenen Diagonalelemente von $D$, und seien $\alpha_1, \ldots, \alpha_r$ diese Diagonalelemente. Dann gilt $0 = m_D(D) = \operatorname{diag}({m_D}^K(\alpha_1), \ldots, {m_D}^K(\alpha_r))$. Also gilt ${m_D}^K(\alpha_i) = 0$ für alle $i \in \{1, \ldots, r\}$, und folglich $\prod_{i=1}^{r}(t - \alpha_i) | m_D$. Da $p := \prod_{i=1}^{r}(t - \alpha_i)$ die Gleichung $\hat{p}(D) = 0$ erfüllt, gilt $m_D | p$. Also gilt $m_D = p$. Da für jedes Polynom $q \in K[t]$ die Gleichung $\hat{q}(A) = 0$ genau dann gilt, wenn $\hat{q}(D) = 0$, erhalten wir $m_D = m_A$.

(2)⇒(1): Sei $p_i := \prod_{\substack{j=1 \\ j\neq i}} (t-\lambda_j) = \frac{m_A}{t-\lambda_i}$. Wegen $\gcd(p_1, \ldots, p_r) = 1$ gibt es $u_1, \ldots, u_r \in K[t]$ mit $\sum_{i=1}^r p_i u_i = 1$. Sei $U_i$ der Spaltenraum von $\hat{p}_i(A)$, und sei $l_i := t - \lambda_i$. Dann gilt für alle $x \in K^n$, dass $0 = 0 \cdot x = \hat{m}_A(A) \cdot x = \hat{l}_i(A) \cdot \hat{p}_i(A) \cdot x$, und folglich $U_i \subseteq E(A, \lambda_i)$. Wegen $x = \sum_{i=1}^r \hat{p}_i(A) \cdot (\hat{u}_i(A) \cdot x) = x$, gilt $K^n \subseteq U_1 + \cdots + U_r \subseteq E(A, \lambda_1) + \cdots + E(A, \lambda_r)$. Somit erhält man eine Basis von $K^n$, die nur aus Eigenvektoren von $A$ besteht. Wegen Satz 14.8 ist $A$ also diagonalisierbar. ■

KAPITEL 17

# Der Dualraum eines Vektorraums

## 1. Vektorräume von Homomorphismen

DEFINITION 17.1. Seien $V, W$ Vektorräume über dem Körper $K$. Mit $\mathrm{Hom}_K(V, W)$ bezeichnen wir die Menge aller Homomorphismen von $V$ nach $W$.

Für $g, h \in \mathrm{Hom}_K(V, W)$ sind $g + h : V \to W, v \mapsto g(v) + h(v)$ und $\alpha * g : V \to W$, $v \mapsto \alpha * g(v)$ wieder Homomorphismen. Mit diesen Operationen ist $\mathrm{Hom}_K(V, W)$ ein Vektorraum über $K$.

SATZ 17.2. *Seien $V, W$ endlichdimensionale Vektorräume über dem Körper $K$ mit* $\dim(V) > 0$ *und* $\dim(W) > 0$, *sei $B = (v_1, \ldots, v_m)$ eine Basis von $V$ und $C = (w_1, \ldots, w_n)$ eine Basis von $W$, und sei $K^{n\times m}$ der Vektorraum aller $n \times m$-Matrizen über $K$. Die Abbildung $\Phi$ ist definiert durch*

$$\begin{array}{rrcl} \Phi : & \mathrm{Hom}_K(V, W) & \longrightarrow & K^{n\times m} \\ & g & \longmapsto & S_g(B, C). \end{array}$$

*Dann ist $\Phi$ ein Isomorphismus.*

KOROLLAR 17.3. *Seien $V, W$ endlichdimensionale Vektorräume über $K$. Dann gilt* $\dim(\mathrm{Hom}_K(V, W)) = \dim(V) \cdot \dim(W)$.

SATZ 17.4. *Seien $V, W$ Vektorräume über $K$, sei $B$ eine Basis von $V$, und sei $C$ eine Basis von $W$. Für $b \in B$ und $c \in C$ sei $\varphi[b, c] \in \mathrm{Hom}_K(V, W)$ jene lineare Abbildung, die $\varphi[b, c](b) = c$ und $\varphi[b, c](b') = 0$ für alle $b' \in B \setminus \{b\}$ erfüllt. Dann gilt:*

(1) *Die Menge $D = \{\varphi[b, c] \mid b \in B, c \in C\}$ ist linear unabhängig.*
(2) *$D$ ist genau dann eine Basis von $\mathrm{Hom}_K(V, W)$, wenn $V$ endlichdimensional oder $W = \{0\}$ ist.*

## 2. Der Dualraum

DEFINITION 17.5. Sei $V$ ein Vektorraum über $K$. Der *Dualraum* von $V$ ist definiert durch $V^\star := \mathrm{Hom}(V, K)$.

$V^\star$ ist also die Menge aller linearen Abbildungen von $V$ in den eindimensionalen Vektorraum $K$.

DEFINITION 17.6. Sei $V$ ein endlichdimensionaler Vektorraum mit Basis $B = (b_1, \dots, b_n)$. Wir definieren nun $B^\star = (b^\star{}_1, \dots, b^\star{}_n)$ als Folge von Vektoren in $V^\star$. Dazu sei für $i \in \{1, \dots, n\}$ die Abbildung $b^\star{}_i \in V^\star$ jener Homomorphismus von $V$ nach $K$, der

$$b^\star{}_i(b_i) = 1,\ \ b^\star{}_i(b_j) = 0 \text{ für } j \neq i.$$

erfüllt.

LEMMA 17.7. *Sei $V$ ein endlichdimensionaler Vektorraum mit Basis $B = (b_1, \dots, b_n)$, sei $B^\star = (b^\star{}_1, \dots, b^\star{}_n)$ wie in Definition 17.6, und sei $x \in V$. Dann gilt:*

(1) $b^\star{}_i(x) = (x)_B[i]$.

(2) $\varphi(x) = \sum_{i=1}^{n} \varphi(b_i) * b^\star{}_i(x)$.

*Beweis:* (1) Sei $(y_1, \dots, y_n) := (x)_B$. Dann gilt

$$\begin{aligned}\sum_{i=1}^{n} b^\star{}_i(x) \ * \ b_i &= \sum_{i=1}^{n} b^\star{}_i\Big(\sum_{j=1}^{n} y_j * b_j\Big) \ * \ b_i \\ &= \sum_{i=1}^{n}\Big(\sum_{j=1}^{n} y_j * b^\star{}_i(b_j)\Big) \ * \ b_i \\ &= \sum_{i=1}^{n} y_i \ * \ b_i \\ &= x.\end{aligned}$$

Also gilt $(b^\star{}_1(x), \dots, b^\star{}_n(x)) = (x)_B$.

(2): Es gilt

$$\begin{aligned}\sum_{i=1}^{n} \varphi(b_i) \cdot b^\star{}_i(x) &= \sum_{i=1}^{n} \varphi(b_i) \cdot (x)_B[i] \\ &= \varphi\Big(\sum_{i=1}^{n} (x)_B[i] * b_i\Big) \\ &= \varphi(x).\end{aligned}$$

■

SATZ 17.8. *Sei $V$ ein endlichdimensionaler Vektorraum, $B$ eine Basis von $V$, und $B^\star$ wie in Definition 17.6. Dann ist $B^\star$ eine Basis von $V^\star$.*

*Beweis:* Wir zeigen als erstes, dass $B^\star$ linear unabhängig ist. Seien $\lambda_1, \dots, \lambda_n \in K$ so, dass $\sum_{i=1}^{n} \lambda_i b^\star{}_i = 0$. Sei $j \in \{1, \dots, n\}$. Es gilt $0 = \sum_{i=1}^{n} \lambda_i b^\star{}_i(b_j) = \lambda_j$. Also ist $B^\star$ linear unabhängig. Wegen Lemma 17.7 (2) gilt für $\varphi \in V^\star$, dass $\varphi = \sum_{i=1}^{n} \varphi(b_i) * b^\star{}_i$, also ist $\varphi \in L(B^\star)$. ■

Für einen endlichdimensionalen Vektorraum bezeichnen wir $B^\star$ als die *zu $B$ duale Basis von $V^\star$*. Die Koordinaten eines Elements von $V^\star$ bezüglich $B^\star$ lassen sich mit Lemma 17.7 berechnen: Für jedes $\varphi \in V^\star$ gilt $(\varphi)_{B^\star} = (\varphi(b_1), \ldots, \varphi(b_n))$, oder, anders geschrieben,

$$(\varphi)_{B^\star}[i] = \varphi(b_i). \tag{17.1}$$

Wir halten noch eine Folgerung von Satz 17.8 fest:

KOROLLAR 17.9. *Sei $V$ ein endlichdimensionaler Vektorraum. Dann gilt* $\dim(V) = \dim(V^\star)$.

DEFINITION 17.10. Seien $V, W$ Vektorräume über dem Körper $K$, und sei $f : V \to W$ ein Homomorphismus. Wir definieren die *zu $f$ duale Abbildung* $f^\star$ durch

$$\begin{array}{rcrcl} f^\star & : & W^\star & \longrightarrow & V^\star \\ & & \varphi & \longmapsto & f^\star(\varphi), \end{array}$$

$$\begin{array}{rcrcl} f^\star(\varphi) & : & V & \longrightarrow & K \\ & & v & \longmapsto & \varphi(f(v)). \end{array}$$

Es gilt also $f^\star(\varphi) = \varphi \circ f$. Somit ist $f^\star(\varphi)$ wieder eine lineare Abbildung, und liegt daher in $V^\star$.

Wir werden nun sehen, dass $f^\star$ nicht nur eine Funktion, sondern sogar ein Homomorphismus von $W^\star$ nach $V^\star$ ist.

LEMMA 17.11. *Seien $V, W$ Vektorräume über dem Körper $K$, und sei $f : V \to W$ ein Homomorphismus. Dann gilt* $f^\star \in \mathrm{Hom}(W^\star, V^\star)$.

*Beweisskizze:* Für $\varphi_1, \varphi_2, \varphi \in W^\star$, $v \in V$, $\alpha \in K$ rechnet man nach: $f^\star(\varphi_1 + \varphi_2)(v) = f^\star(\varphi_1)(v) + f^\star(\varphi_2)(v)$ und $f^\star(\alpha * \varphi)(v) = (\alpha * f^\star(\varphi))(v)$. ∎

SATZ 17.12. *Seien $V, W$ endlichdimensionale Vektorräume über dem Körper $K$ mit* $\dim(V) > 0$, $\dim(W) > 0$, *und sei $f : V \to W$ ein Homomorphismus, $B$ eine Basis von $V$, und $C$ eine Basis von $W$. Dann gilt*

$$S_{f^\star}(C^\star, B^\star) = (S_f(B, C))^T.$$

*Beweis:* Sei $B = (b_1, \ldots, b_m)$, $C = (c_1, \ldots, c_n)$, und sei $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. Dann gilt $S_{f^\star}(C^\star, B^\star)[i, j] = (f^\star(c^\star_j))_{B^\star}[i]$. Wegen Gleichung (17.1) ist das gleich $f^\star(c^\star_j)(b_i) = (c^\star_j \circ f)(b_i) = c^\star_j(f(b_i))$. Wegen Lemma 17.7 (1) ist das gleich $(f(b_i))_C[j] = S_f(B, C)[j, i]$. ∎

## 3. Der Bidualraum

DEFINITION 17.13. Sei $V$ ein Vektorraum über $K$. Der *Bidualraum* von $V$ ist definiert als $(V^\star)^\star$, und wird auch einfacher mit $V^{\star\star}$ bezeichnet.

SATZ 17.14. *Sei $V$ ein Vektorraum über $K$, und sei $\Phi : V \to V^{\star\star}$ gegeben durch*

$$\begin{array}{rcccl} \Phi & : & V & \longrightarrow & V^{\star\star} \\ & & v & \longmapsto & \Phi(v), \end{array}$$

*wobei*

$$\begin{array}{rcccl} \Phi(v) & : & V^{\star} & \longrightarrow & K \\ & & \varphi & \longmapsto & \varphi(v). \end{array}$$

*Dann ist $\Phi$ ein Monomorphismus von $V$ nach $V^{\star\star}$. Wenn $V$ endlichdimensional ist, ist $\Phi$ sogar ein Isomorphismus.*

*Beweisskizze:* Wir zeigen als erstes, dasss für jedes $v \in V$ die Abbildung $\Phi(v)$ eine lineare Abbildung von $V^{\star}$ nach $V$ ist. Dazu zeigen wir, dass für alle Homomorphismen $\varphi_1, \varphi_2 : V \to K$ gilt, dass $\Phi(v)\,(\varphi_1 + \varphi_2) = \Phi(v)\,(\varphi_1) + \Phi(v)\,(\varphi_1)$.

Um zu zeigen, dass $\Phi$ linear ist, zeigen wir, dass für alle $v_1, v_2 \in V$ und $\varphi \in V^{\star}$ gilt, dass $\Phi(v_1 + v_2)\,(\varphi) = \Phi(v_1)\,(\varphi) + \Phi(v_2)\,(\varphi)$. ■

Für $v \in V$ und $\varphi \in V^{\star}$ gilt also $\Phi(v)\,(\varphi) = \varphi(v)$. Vektorräume, für die $\Phi$ ein Isomorphismus von $V$ nach $V^{\star\star}$ ist, heißen *reflexiv*. Jeder endlichdimensionale Vektorraum ist also reflexiv.

# Teil 6

# Vektorräume mit Skalarprodukt

KAPITEL 18

# Euklidische Vektorräume

## 1. Orthogonale Matrizen

DEFINITION 18.1. Sei $n \in \mathbb{N}$, $A \in \mathbb{R}^{n \times n}$. Die Matrix $A$ ist *orthogonal* $:\Leftrightarrow A^T \cdot A = E_n$.

LEMMA 18.2. *Sei $n \in \mathbb{N}$, $A \in \mathbb{R}^{n \times n}$. Dann sind äquivalent:*

(1) *$A$ ist orthogonal.*
(2) *Die Spalten von $A$ sind eine Orthonormalbasis von $\mathbb{R}^n$.*
(3) *Die Zeilen von $A$ sind eine Orthonormalbasis von $\mathbb{R}^n$.*

DEFINITION 18.3. Sei $\varphi : \mathbb{R}^n \to \mathbb{R}^n$. Die Funktion $\varphi$ ist eine *Isometrie*, wenn für alle $x, y \in \mathbb{R}^n$ gilt: $\|\varphi(x) - \varphi(y)\| = \|x - y\|$.

SATZ 18.4. *Sei $n \in \mathbb{N}$, und sei $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung. Dann ist $\varphi$ genau dann eine Isometrie, wenn $S_\varphi(E, E)$ eine orthogonale Matrix ist.*

Wir werden im folgenden das Skalarprodukt $\langle x, y\rangle$ von zwei Vektoren im $\mathbb{R}^n$ auch als $x^T \cdot y$ schreiben.

*Beweis von Satz 18.4:* Sei $A := S_\varphi(E, E)$. Seien nun $\varphi$ eine Isometrie, und $x, y \in \mathbb{R}^n$. Dann gilt $\|A \cdot (x + y) - A \cdot 0\| = \|(x + y) - 0\|$, da $\varphi$ eine Isometrie ist. Also gilt $\langle A \cdot x + A \cdot y, A \cdot x + A \cdot y\rangle = \langle x + y, x + y\rangle$, und somit

$$\langle A \cdot x, A \cdot x\rangle + 2\langle A \cdot x, A \cdot y\rangle + \langle A \cdot y, A \cdot y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle.$$

Da $\varphi$ eine Isometrie ist und $\varphi(0) = 0$, gilt $\|A \cdot x\| = \|x\|$ und $\|A \cdot y\| = \|y\|$. Somit gilt

$$\langle A \cdot x, A \cdot y\rangle = \langle x, y\rangle.$$

Also gilt $x^T \cdot A^T \cdot A \cdot y = x^T \cdot y$. Wenn wir nun $x := e_i$, $y := e_j$ setzen, so erhalten wir $(e_i \cdot A^T) \cdot (A \cdot e_j) = \delta(i, j)$. Somit steht in den Spalten von $A$ eine Orthonormalbasis von $\mathbb{R}^n$, und $A$ ist daher orthogonal.

Sei nun umgekehrt $A$ eine orthogonale Matrix. Dann gilt $\|A \cdot x - A \cdot y\|^2 = \|A \cdot (x - y)\|^2 = (x - y)^T \cdot A^T \cdot A \cdot (x - y) = (x - y)^T \cdot E_n \cdot (x - y) = (x - y)^T \cdot (x - y) = \|x - y\|^2$. ∎

## 2. Euklidische Räume

Wir werden nun auch Skalarprodukte auf anderen Räumen als $\mathbb{R}^n$ brauchen:

DEFINITION 18.5. Sei $V$ ein Vektorraum über $\mathbb{R}$, und sei $\langle . \,|\, . \rangle : V \times V \to \mathbb{R}$. Diese Abbildung ist genau dann ein *Skalarprodukt auf* $V$, wenn für alle $x, x', y \in V$ und $\alpha \in \mathbb{R}$ gilt:

(1) $\langle x + x' \,|\, y \rangle = \langle x \,|\, y \rangle + \langle x' \,|\, y \rangle$,
(2) $\langle \alpha * x \,|\, y \rangle = \alpha \langle x \,|\, y \rangle$,
(3) $\langle x \,|\, y \rangle = \langle y \,|\, x \rangle$,
(4) wenn $x \neq 0$, so gilt $\langle x \,|\, x \rangle > 0$.

Ein *Euklidischer Raum* ist ein Paar $(V, \langle . \,|\, . \rangle)$ aus einem Vektorraum $V$ über $\mathbb{R}$ und einem Skalarprodukt auf $V$.

Mit $\langle x, y \rangle$ bezeichnen wir, wie bisher, nur das Skalarprodukt auf $\mathbb{R}^n$, das durch $\langle x, y \rangle := x^T \cdot y = \sum_{i=1}^n x_i y_i$ definiert ist.

Für jedes Element $x$ eines Euklidischen Raums definieren wir $||x|| := \sqrt{\langle x \,|\, x \rangle}$. Eine *Orthonormalbasis* eines Euklidischen Raums ist eine Basis $B$, sodass $\langle b \,|\, b \rangle = 1$ für alle $b \in B$, und $\langle b \,|\, c \rangle = 0$ für $b, c \in B$ mit $b \neq c$.

SATZ 18.6. *Sei $(V, \langle . \,|\, . \rangle)$ ein endlichdimensionaler Euklidischer Raum. Dann besitzt $V$ eine Orthonormalbasis.*

*Beweisskizze:* Gram-Schmidtsches Orthonormalisierungsverfahren.

LEMMA 18.7. *Sei $(V, \langle . \,|\, . \rangle)$ ein endlichdimensionaler Euklidischer Raum, und sei $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis von $(V, \langle . \,|\, . \rangle)$. Dann gilt für alle $x, y \in V$: $\langle x \,|\, y \rangle = \langle (x)_B, (y)_B \rangle$.*

*Beweis:* Sei $n := \dim(V)$. Es gilt $\langle x \,|\, y \rangle = \langle \sum_{i=1}^n (x)_B[i] * b_i \,|\, \sum_{j=1}^n (x)_B[j] * b_j \rangle = \sum_{i=1}^n \sum_{j=1}^n (x)_B[i]\,(x)_B[j]\,\langle b_i \,|\, b_j \rangle = \sum_{i=1}^n (x)_B[i]\,(x)_B[i] = \langle (x)_B, (y)_B \rangle$. ∎

## 3. Selbstadjungierte Operatoren

DEFINITION 18.8. Sei $(V, \langle . \,|\, . \rangle)$ ein Euklidischer Raum, und sei $\varphi : V \to V$ ein Endomorphismus. Die Abbildung $\varphi$ heißt *selbstadjungiert*, wenn für alle $x, y \in V$ gilt: $\langle \varphi(x) \,|\, y \rangle = \langle x \,|\, \varphi(y) \rangle$.

SATZ 18.9. *Sei $(V, \langle . \,|\, . \rangle)$ ein endlichdimensionaler Euklidischer Raum, sei $\varphi : V \to V$ ein Endomorphismus, und sei $B$ eine Orthonormalbasis von $(V, \langle . \,|\, . \rangle)$. Dann sind äquivalent:*

(1) *$\varphi$ ist selbstadjungiert.*
(2) *$A := S_\varphi(B, B)$ ist symmetrisch (das heißt, es gilt $A^T = A$).*

*Beweis:* (1)⇒(2): Für alle $x, y \in V$ gilt $\langle \varphi(x) \,|\, y \rangle = \langle x \,|\, \varphi(y) \rangle$, also $\langle (\varphi(x))_B, (y)_B \rangle = \langle (x)_B, (\varphi(y))_B \rangle$, und somit $\langle A \cdot (x)_B, (y)_B \rangle = \langle (x)_B, A \cdot (y)_B \rangle$. Für $x := b_i$ und $y := b_j$ gilt also $\langle A \cdot e_i, e_j \rangle = \langle e_i, A \cdot e_j \rangle$, und somit $e_i^T \cdot A^T \cdot e_j = e_i^T \cdot A \cdot e_j$, also $A^T(i,j) = A(i,j)$. Die Matrix $A$ ist also symmetrisch.

(2)⇒(1): $\langle \varphi(x) \,|\, y \rangle = \langle (\varphi(x))_B, (y)_B \rangle = (A \cdot (x)_B)^T \cdot (y)_B = (x)_B^T \cdot A^T \cdot (y)_B = (x)_B^T \cdot A \cdot (y)_B = \langle (x)_B, A \cdot (y)_B \rangle = \langle (x)_B, (\varphi(y))_B \rangle = \langle x \,|\, \varphi(y) \rangle$. ■

LEMMA 18.10. *Sei $A \in \mathbb{R}^{n \times n}$ so, dass $A = A^T$. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von $A$, das heißt, sei $\lambda \in \mathbb{C}$ so, dass es ein $x \in \mathbb{C}^n \setminus \{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\}$ mit $A \cdot x = \lambda * x$ gibt. Dann gilt $\lambda \in \mathbb{R}$. Außerdem gibt es $y \in \mathbb{R}^n$ mit $y \neq 0$ und $A \cdot y = \lambda * y$*

*Beweis:* Für $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$ und $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ bezeichen wir mit $\overline{z} := (\overline{z_1}, \ldots, \overline{z_n})$ den Vektor mit den konjugiert komplexen Einträgen, und mit $y^T \cdot \overline{z}$ die Zahl $\sum_{i=1}^n y_i \overline{z_i}$. Da $x \neq 0$, gilt $x^T \cdot \overline{x} \in \mathbb{R}$ und $x^T \cdot \overline{x} > 0$. Wegen $A = A^T$ gilt $\lambda(x^T \cdot \overline{x}) = (\lambda * x)^T \cdot \overline{x} = (A \cdot x)^T \cdot \overline{x} = (x^T \cdot A^T) \cdot \overline{x} = x^T \cdot (A^T \cdot \overline{x}) = x^T \cdot (A \cdot \overline{x}) = x^T \cdot (\overline{A \cdot x}) = x^T \cdot (\overline{\lambda * x}) = x^T \cdot (\overline{\lambda} * \overline{x}) = \overline{\lambda}(x^T \cdot \overline{x})$. Somit gilt $\lambda = \overline{\lambda}$, und damit $\lambda \in \mathbb{R}$. Die Matrix $A - \lambda * E$ ist also in $\mathbb{C}^{n \times n}$ nicht invertierbar, da ihre Spalten linear abhängige Vektoren in $\mathbb{C}^n$ sind. Damit ist die Matrix $A - \lambda * E$ auch in $\mathbb{R}^{n \times n}$ nicht invertierbar, und somit gibt es $y \in \mathbb{R}^n$ mit $y \neq 0$ und $A \cdot y = \lambda * y$. ■

DEFINITION 18.11. Sei $V$ ein Vektorraum über $K$, und $\varphi : V \to V$ ein Endomorphismus. Ein $\lambda \in K$ ist ein *Eigenwert* von $\varphi$, wenn es ein $v \in V$ mit $v \neq 0$ gibt, sodass $\varphi(v) = \lambda * v$. Ein $v \in V$ ist ein *Eigenvektor* von $\varphi$, wenn es $\lambda \in K$ gibt, sodass $\varphi(v) = \lambda * v$.

LEMMA 18.12. *Sei $(V, \langle . \,|\, . \rangle)$ ein endlichdimensionaler Euklidischer Raum mit* $\dim(V) > 0$*, und sei $\varphi : V \to V$ ein selbstadjungierter Endomorphismus. Dann besitzt $\varphi$ einen Eigenvektor in $V$.*

*Beweis:* Sei $C$ eine Orthonormalbasis von $V$, $n := \dim(V)$, und sei $A := S_\varphi(C, C)$. Die Matrix $A$ liegt in $\mathbb{R}^{n \times n}$ und ist wegen Satz 18.9 symmetrisch. Da in $\mathbb{C}$ das characteristische Polynom $c_A$ zumindest eine Nullstelle hat, gibt es $\lambda \in \mathbb{C}$ und $x \in \mathbb{C}^n \setminus \{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\}$, sodass $A \cdot x = \lambda * x$. Wegen Lemma 18.10 gilt $\lambda \in \mathbb{R}$, und es gibt $y \in \mathbb{R}^n$ mit $y \neq 0$, sodass $A \cdot y = \lambda * y$. Sei nun $v \in V$ so, dass $(v)_C = y$. Dann ist $v$ ein Eigenvektor von $\varphi$ zum Eigenwert $\lambda$. ■

SATZ 18.13 (Hauptachsentransformation für selbstadjungierte Operatoren). *Sei $(V, \langle . \,|\, . \rangle)$ ein endlichdimensionaler Euklidischer Raum mit* $\dim(V) > 0$*, und sei $\varphi : V \to V$ ein selbstadjungierter Endomorphismus. Dann besitzt $V$ eine Orthonormalbasis, die nur aus Eigenvektoren von $\varphi$ besteht.*

*Beweis:* Wir beweisen den Satz mit Induktion nach der Dimension von $V$. Wenn $\dim(V) = 1$, so wählen wir $v \in V$ mit $v \neq 0$. Wir setzen $b := \frac{1}{\|v\|} * v$, und $B := (b)$. Dann ist $B$ eine Orthonormalbasis. Wegen $\dim(V) = 1$ ist $b$ ein Eigenvektor von $V$.

Sei nun $n := \dim(V) \geq 2$. Wegen Lemma 18.12 besitzt $\varphi$ einen Eigenvektor $v$. Wir bilden nun $b_1 := \frac{1}{\|v\|} * v$. Wir setzen $U_1 := L(b_1)$ und

$$U_2 := \{b_1\}^\perp = \{w \in \mathbb{R}^n \,:\, \langle w \,|\, b_1 \rangle = 0\}.$$

Wegen $b_1 \neq 0$ gilt $\dim(U_1) = 1$. Für die Abbildung $h \in V^\star$, $h(w) := \langle w \,|\, b_1 \rangle$ gilt $h(b_1) \neq 0$. Also gilt $\dim(\mathrm{im}(h)) = 1$, und somit wegen des Homomorphiesatzes $\dim(U_2) = \dim(\ker(h)) = n - 1$. Wir zeigen nun, dass $U_2$ $\varphi$-invariant ist. Sei dazu $w \in U_2$. Wir zeigen, dass auch $\varphi(w)$ in $U_2$ liegt. Es gilt $\langle \varphi(w) \,|\, b_1 \rangle = \langle w \,|\, \varphi(b_1) \rangle = \langle w \,|\, \lambda * b_1 \rangle = \lambda \langle w \,|\, b_1 \rangle = 0$. Also ist $U_2$ $\varphi$-invariant.

Wir zeigen nun $U_1 \cap U_2 = \{0\}$. Sei $w \in U_1 \cap U_2$. Wegen $w \in U_1$ gibt es $\mu \in \mathbb{R}$, sodass $w = \mu * b_1$. Dann gilt $0 = \langle w \,|\, b_1 \rangle = \langle \mu * b_1 \,|\, b_1 \rangle = \mu \langle b_1 \,|\, b_1 \rangle = \mu$, also $w = 0$. Also ist die Summe $U = U_1 \dotplus U_2$ direkt. Da $\dim(U) = \dim(U_1) + \dim(U_2) = 1 + (n - 1)$, gilt $U = V$. Der Vektorraum $V$ ist also die direkte Summe von $U_1$ und $U_2$.

Nun ist $(U_2, \langle . \,|\, . \rangle|_{U_2 \times U_2})$ ein Euklidischer Raum. Der Endomorphismus $\varphi_2 := \varphi|_{U_2}$ ist selbstadjungiert. Nach Induktionsvoraussetzung besitzt also $(U_2, \langle . \,|\, . \rangle|_{U_2 \times U_2})$ eine Orthonormalbasis $B_2$, die aus lauter Eigenvektoren von $\varphi_2$, und damit von $\varphi$, besteht.

Insgesamt bildet $\{b_1\} \cup B_2$ also eine Orthonormalbasis von $V$, die nur aus Eigenvektoren von $\varphi$ besteht. ■

KOROLLAR 18.14. *Sei $A \in \mathbb{R}^{n \times n}$ so, dass $A = A^T$. Dann gibt es eine orthogonale Matrix $Q \in \mathbb{R}^{n \times n}$, sodass $Q^T \cdot A \cdot Q$ eine Diagonalmatrix in $\mathbb{R}^{n \times n}$ ist.*

*Beweis:* Sei $\varphi := h_A$. Sei $B$ eine Orthonormalbasis des $\mathbb{R}^n$, die aus Eigenvektoren von $h_A$ besteht. Wir schreiben diese Vektoren in die Spalten von $Q$. Sei $D := S_{h_A}(B, B)$. Es gilt dann $D = Q^{-1} \cdot A \cdot Q$. Da $Q$ eine orthogonale Matrix ist, gilt $Q^{-1} = Q^T$. ■

BEISPIEL 18.15. Wir zeigen in folgendem Beispiel, wie man diese Diagonalmatrix tatsächlich finden kann.

Wir berechnen im folgenden Beispiel eine Orthonormalbasis aus Eigenvektoren der Matrix A.

**A = {{326/125, 672/625, −504/625},**
**{672/625, −3216/3125, 2412/3125},**
**{−504/625, 2412/3125, −1809/3125}}**

$$\left\{\left\{\frac{326}{125}, \frac{672}{625}, -\frac{504}{625}\right\}, \left\{\frac{672}{625}, -\frac{3216}{3125}, \frac{2412}{3125}\right\}, \left\{-\frac{504}{625}, \frac{2412}{3125}, -\frac{1809}{3125}\right\}\right\}$$

**MatrixForm[A]**

$$\begin{pmatrix} \frac{326}{125} & \frac{672}{625} & -\frac{504}{625} \\ \frac{672}{625} & -\frac{3216}{3125} & \frac{2412}{3125} \\ -\frac{504}{625} & \frac{2412}{3125} & -\frac{1809}{3125} \end{pmatrix}$$

Wir berechnen als erstes einen Eigenwert und einen Eigenvektor von $A$.

**cA = CharacteristicPolynomial[A, t]**

$6t + t^2 - t^3$

Wir berechnen die Eigenvektoren zum Eigenwert 0.

**U1 = NullSpace[*A* − 0 ∗ IdentityMatrix[3]]**

$\left\{\left\{0, \frac{3}{4}, 1\right\}\right\}$

Wir wählen einen Vektor daraus als Vektor b1

**b1 = U1[[1]]**

$\left\{0, \frac{3}{4}, 1\right\}$

Der Orthogonalraum von b1 ist U2. Wir wissen, dass auch dieser Raum wieder h_A-invariant ist.

**U2 = NullSpace[{b1}]**

$\left\{\left\{0, -\frac{4}{3}, 1\right\}, \{1, 0, 0\}\right\}$

Wir berechnen nun irgendeine Basis von U2 mit dem Ziel, einen Eigenvektor von A zu erhalten, der in U2 liegt

**C2 = Transpose[U2]**

$\left\{\{0, 1\}, \left\{-\frac{4}{3}, 0\right\}, \{1, 0\}\right\}$

Wir berechnen nun die Darstellungsmatrix von h_A bezueglich der Basis C2

**col1 = LinearSolve[C2, *A*.U2[[1]]]**

$\left\{-\frac{201}{125}, -\frac{56}{25}\right\}$

**col2 = LinearSolve[C2, *A*.U2[[2]]]**

$\left\{-\frac{504}{625}, \frac{326}{125}\right\}$

**ShC2C2 = Transpose[{col1, col2}]**

$\left\{\left\{-\frac{201}{125}, -\frac{504}{625}\right\}, \left\{-\frac{56}{25}, \frac{326}{125}\right\}\right\}$

**MatrixForm[ShC2C2]**

$$\begin{pmatrix} -\frac{201}{125} & -\frac{504}{625} \\ -\frac{56}{25} & \frac{326}{125} \end{pmatrix}$$

Wir berechnen nun einen Eigenvektor von S_h(C2, C2)

***c* = CharacteristicPolynomial[ShC2C2, *t*]**

$-6 - t + t^2$

**Factor[*c*]**

$(-3 + t)(2 + t)$

**U21bzglC2 = NullSpace[ShC2C2 + 2 ∗ IdentityMatrix[2]]**

$\left\{\left\{\frac{72}{35}, 1\right\}\right\}$

Aus diesem Eigenvektor von S_h (C2,C2) erhalten wir einen Eigenvektor von A, der in U2 liegt

**b2 = C2.U21bzglC2[[1]]**

$\left\{1, -\frac{96}{35}, \frac{72}{35}\right\}$

Nun betrachten wir den h_A-invarianten Unterraum U3 von U2 aller der Vektoren, die in U2 liegen und normal auf b2 stehen. Das ist genau das orthogonale Komplement von {b1,b2}

**U3 = NullSpace[{b1, b2}]**

$\left\{\left\{-\frac{40}{7}, -\frac{4}{3}, 1\right\}\right\}$

**b3 = U3[[1]]**

$\left\{-\frac{40}{7}, -\frac{4}{3}, 1\right\}$

Wir berechnen nun noch eine Basis C3 von U3.

**C3 = Transpose[U3]**

$\left\{\left\{-\frac{40}{7}\right\}, \left\{-\frac{4}{3}\right\}, \{1\}\right\}$

**ShC3C3 = LinearSolve[C3, *A*.U3[[1]]]**

{3}

Die Vektoren b1,b2,b3 normieren wir nun; so erhalten wir die gewuenschte Orthonormalbasis. Diese Basis schreiben wir in die Spalten von Q.

***Q* = Transpose[{1/Sqrt[(b1.b1)] ∗ b1, 1/Sqrt[(b2.b2)] ∗ b2,**
**1/Sqrt[(b3.b3)] ∗ b3}]**

$\left\{\left\{0, \frac{7}{25}, -\frac{24}{25}\right\}, \left\{\frac{3}{5}, -\frac{96}{125}, -\frac{28}{125}\right\}, \left\{\frac{4}{5}, \frac{72}{125}, \frac{21}{125}\right\}\right\}$

**DD = Transpose[*Q*].*A*.*Q***

$\{\{0, 0, 0\}, \{0, -2, 0\}, \{0, 0, 3\}\}$

SATZ 18.16. *Sei $(V, \langle . \,|\, . \rangle)$ ein Euklidischer Vektorraum, und sei $\varphi$ ein selbstadjungierter Operator. Dann sind Eigenvektoren von $\varphi$, die zu verschiedenen Eigenwerten von A gehören, aufeinander orthogonal.*

*Beweis:* Sei $\varphi(v_1) = \lambda_1 * v_1$, $\varphi(v_2) = \lambda_2 * v_2$ mit $\lambda_1 \neq \lambda_2$. Dann gilt $\lambda_1 \langle v_1 \,|\, v_2 \rangle = \langle \lambda_1 * v_1 \,|\, v_2 \rangle = \langle \varphi(v_1) \,|\, v_2 \rangle = \langle v_1 \,|\, \varphi(v_2) \rangle = \langle v_1 \,|\, \lambda_2 * v_2 \rangle = \lambda_2 \langle v_1 \,|\, v_2 \rangle$. Wegen $\lambda_1 \neq \lambda_2$ gilt $\langle v_1 \,|\, v_2 \rangle = 0$. ■

## 4. Positiv definite Matrizen

DEFINITION 18.17. Sei $A \in \mathbb{R}^{n\times n}$. Die Matrix $A$ ist *positiv definit*, wenn für alle $x \in \mathbb{R}^n$ mit $x \neq 0$ gilt $x^T \cdot A \cdot x > 0$.

Positiv definite Matrizen sind regulär: wenn nämlich die Spalten linear abhängig sind, so gibt es $x \neq 0$ mit $A \cdot x = 0$ und somit $x^T \cdot A \cdot x = 0$. Wegen Lemma 18.10 sind alle komplexen Nullstellen des charakteristischen Polynoms einer symmetrischen Matrix reell. Aus ihren Vorzeichen kann man auch die positive Definitheit einer symmetrischen Matrix $A$ ablesen.

SATZ 18.18. *Sei $n \in N$, und sei $A \in \mathbb{R}^{n\times n}$ mit $A^T = A$. Dann sind äquivalent:*

(1) *$A$ ist positiv definit.*
(2) *Alle Eigenwerte von $A$ sind positiv.*
(3) *Es existiert $C \in \mathrm{GL}(n, \mathbb{R})$, sodass $C^T = C$ und $A = C^T \cdot C = C^2$.*
(4) *Es existiert $F \in \mathrm{GL}(n, \mathbb{R})$, sodass $A = F^T \cdot F$.*

*Beweis:* (1)⇒(2): Sei $\lambda$ ein Eigenwert von $A$, und sei $x \in \mathbb{R}^n \setminus \{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\}$ so, dass $A \cdot x = \lambda * x$. Dann gilt $0 < x^T \cdot A \cdot x = x^T \cdot (\lambda * x) = \lambda(x^T \cdot x)$. Wegen $x^T \cdot x \geq 0$ bedeutet das, dass $\lambda > 0$ ist.

(2)⇒(3): Da $A$ symmetrisch ist, gibt es eine nach Koollar 18.14 eine orthogonale Matrix $Q \in \mathbb{R}^{n\times n}$ und eine Diagonalmatrix $D \in \mathbb{R}^{n\times n}$, sodass $A = Q \cdot D \cdot Q^T$. Sei $D = \mathrm{diag}(d_1, \ldots, d_n)$. Wegen $A \cdot (Q \cdot e_i) = Q \cdot D \cdot e_i = Q \cdot (d_i * e_i) = d_i * (Q \cdot e_i)$ sind alle $d_i$ Eigenwerte von $A$. Also sind alle $d_i > 0$. Sei

$$W := \mathrm{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n}).$$

Dann gilt $W^T \cdot W = D$. Sei nun

$$C := Q \cdot W \cdot Q^T.$$

Wegen $\det(W) > 0$ ist $W$ invertierbar. Somit ist auch $C$ invertierbar. Es gilt

$$\begin{aligned} C^T &= Q \cdot W^T \cdot Q^T \\ &= (Q^T)^T \cdot W \cdot Q^T \\ &= C. \end{aligned}$$

Folglich ist $C$ symmetrisch. Es gilt $C^T \cdot C = Q \cdot W^T \cdot Q^T \cdot Q \cdot W \cdot Q^T = Q \cdot W^T \cdot W \cdot Q^T = Q \cdot D \cdot Q^T = A$. Somit gilt (3)

(3)⇒(4): Die Matrix $F := C$ leistet das Gewünschte.

(4)⇒(1): Sei $x \in \mathbb{R}^n$ mit $x \neq 0$. Es gilt $x^T \cdot A \cdot x = x^T \cdot F^T \cdot F \cdot x = (F \cdot x)^T \cdot (F \cdot x)$. Da $F$ invertierbar ist, ist $F \cdot x \neq 0$, also ist $(F \cdot x)^T \cdot (F \cdot x) > 0$. ■

DEFINITION 18.19. Sei $A \in \mathbb{R}^{n\times n}$, und sei $k \in \{1, \ldots, n\}$. Der *$k$-te Hauptminor von $A$* ist die Determinante der $k \times k$-Matrix $B$, die durch $B[i, j] := A[i, j]$ für $i, j \in \{1, \ldots, k\}$ definiert ist.

Der $k$-te Hauptminor ist also die Determinante des nordwestlichen $k \times k$-Ecks von $A$.

LEMMA 18.20. *Sei $A \in \mathbb{R}^{n\times n}$, $B \in \mathrm{GL}(n-1, \mathbb{R})$, $c \in \mathbb{R}^{n-1}$ und $d \in \mathbb{R}$ so, dass*

$$A = \left(\begin{array}{c|c} & | \\ B & c \\ & | \\ \hline — \quad c^T \quad — & d \end{array}\right).$$

*Dann gilt* $\det(A) = \det(B) \cdot (d - c^T \cdot B^{-1} \cdot c)$.

*Beweis:* Sei $a := -B^{-1} \cdot c$, und sei $X$ die Matrix, die durch

$$X := \left(\begin{array}{c|c} & | \\ B^{-1} & a \\ & | \\ \hline — \quad 0 \quad — & 1 \end{array}\right)$$

definiert ist. Durch Entwickeln nach der letzten Zeile erhält man $\det(X) = \det(B^{-1})$. Es gilt

$$\begin{aligned} A \cdot X &= \left(\begin{array}{c|c} & | \\ B & c \\ & | \\ \hline — \quad c^T \quad — & d \end{array}\right) \cdot \left(\begin{array}{c|c} & | \\ B^{-1} & a \\ & | \\ \hline — \quad 0 \quad — & 1 \end{array}\right) \\ &= \left(\begin{array}{c|c} & | \\ E_{n-1} & B \cdot a + c \\ & | \\ \hline — \quad c^T \cdot B^{-1} \quad — & c^T \cdot a + d \end{array}\right) \\ &= \left(\begin{array}{c|c} & | \\ E_{n-1} & 0 \\ & | \\ \hline — \quad c^T \cdot B^{-1} \quad — & -c^T \cdot B^{-1} \cdot c + d \end{array}\right). \end{aligned}$$

Die Determinate der letzten Matrix kann man durch Entwickeln nach der letzten Spalte ausrechnen. Es gilt also

$$\det(A) \cdot \det(X) = -c^T \cdot B^{-1} \cdot c + d.$$

Somit gilt wegen $\det(X) = \frac{1}{\det(B)}$ die Gleichheit $\det(A) = \det(B) \cdot (-c^T \cdot B^{-1} \cdot c + d)$.

SATZ 18.21. *Sei $n \in \mathbb{N}$, $A \in \mathbb{R}^{n\times n}$, $A = A^T$. Dann sind äquivalent:*

(1) *$A$ ist positiv definit.*
(2) *Alle $n$ Hauptminoren von $A$ sind positiv.*

*Beweis:* (1)⇒(2): Sei $k \in \{1, \ldots, n\}$, und sei $B \in \mathbb{R}^{k\times k}$ die Matrix, die aus den ersten $k$ Zeilen der ersten $k$ Spalten von $A$ besteht. Wir zeigen nun, dass $B$ positive Determinante hat. Da $B$ eine symmetrische Matrix ist, gibt es eine orthogonale Matrix $Q \in \mathbb{R}^{k\times k}$, sodass $D = Q^T \cdot B \cdot Q$ eine Diagonalmatrix $\operatorname{diag}(d_1, \ldots, d_k)$ ist, Wir zeigen als nächstes, dass alle $d_i > 0$ sind. Da $Q$ eine orthogonale Matrix ist, ist ihre $i$-te Spalte, also $Q \cdot e_i$, nicht der Nullvektor. Sei $y \in \mathbb{R}^n$ so, dass $\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} = Q \cdot e_i$ und $y_{k+1} = \cdots = y_n = 0$. Da $A$ positiv definit ist, gilt

$$\begin{aligned} 0 &< y^T \cdot A \cdot y \\ &= (Q \cdot e_i)^T \cdot B \cdot (Q \cdot e_i) \\ &= e_i^T \cdot D \cdot e_i \\ &= d_i. \end{aligned}$$

Wegen $\det(B) = \det(Q)^2 \prod_{i=1}^{n} d_i$ gilt also auch $\det(B) > 0$.

(2)⇒(1): Wir beweisen diese Implikation mit Induktion nach $n$. Für $n = 1$ erhalten wir, dass $A = (A[1, 1])$ ist. Der erste Hauptminor von $A$ ist $A[1, 1]$. Wenn diese Zahl positiv ist, so gilt für alle $x \in \mathbb{R} \setminus \{0\}$, dass $A[1, 1]\, x^2 > 0$; somit ist $A$ positiv definit.

Sei nun $n \geq 2$. Wir finden $c \in \mathbb{R}^{n-1}$, $d \in \mathbb{R}$, sodass

$$A = \left(\begin{array}{ccc|c} & & & | \\ & B & & c \\ & & & | \\ \hline — & c^T & — & d \end{array}\right).$$

Da $B$ symmetrisch ist und alle Hauptminoren von $B$ positiv sind, ist $B$ nach Induktionsvoraussetzung positiv definit. Sei nun $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ mit $x \neq 0$. Wir berechnen $x^T \cdot A \cdot x$.

Wenn $x_n = 0$, so gilt

$$x^T \cdot A \cdot x = (x_1, \ldots, x_{n-1}) \cdot B \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}.$$

Da $B$ positiv definit ist, ist das $> 0$.

Wenn $x_n \neq 0$, so setzen wir $y = \frac{1}{x_n} * x$, und $z := (y_1, \ldots, y_{n-1})^T$. Wir zeigen nun

$$y^T \cdot A \cdot y > 0. \tag{18.1}$$

Dazu berechnen wir als erstes $y^T \cdot A \cdot y$. Es gilt

$$
\begin{aligned}
y^T \cdot A \cdot y &= y^T \cdot \left(\begin{array}{c|c} B & \begin{matrix} | \\ c \\ | \end{matrix} \\ \hline — \; c^T \; — & d \end{array}\right) \cdot y \\
&= (—z^T—,\ 1) \cdot \left(\begin{array}{c|c} B & \begin{matrix} | \\ c \\ | \end{matrix} \\ \hline — \; c^T \; — & d \end{array}\right) \cdot \begin{pmatrix} | \\ z \\ | \\ 1 \end{pmatrix} \qquad (18.2) \\
&= (—z^T \cdot B + c^T—,\ z^T \cdot c + d) \cdot \begin{pmatrix} | \\ z \\ | \\ 1 \end{pmatrix} \\
&= z^T \cdot B \cdot z + c^T \cdot z + z^T \cdot c + d \\
&= z^T \cdot B \cdot z + 2\, c^T \cdot z + d.
\end{aligned}
$$

Da $B$ symmetrisch und positiv definit ist, gibt es ein $C \in \mathrm{GL}(n-1, \mathbb{R})$, sodass $B = C^T \cdot C$ und $C^T = C$. Es gilt

$$\langle C \cdot z + C^{-1} \cdot c\,,\, C \cdot z + C^{-1} \cdot c \rangle \geq 0.$$

Das bedeutet

$$z^T \cdot C^T \cdot C \cdot z + 2\, c^T \cdot (C^{-1})^T \cdot C \cdot z + c^T \cdot (C^{-1})^T \cdot C^{-1} \cdot c \geq 0.$$

Wir verwenden nun

$$
\begin{aligned}
C^T \cdot C &= B, \\
(C^{-1})^T \cdot C = (C^{-1})^T \cdot C^T &= (C \cdot C^{-1})^T = E_{n-1},
\end{aligned}
$$

und

$$(C^{-1})^T \cdot C^{-1} = (C^T)^{-1} \cdot C^{-1} = (C \cdot C^T)^{-1} = (C^T \cdot C)^{-1} = B^{-1},$$

und erhalten somit

$$z^T \cdot B \cdot z + 2\, c^T \cdot z + c^T \cdot B^{-1} \cdot c \geq 0. \qquad (18.3)$$

Da alle Hauptminoren von $A$ positiv sind, haben $A$ und $B$ positive Determinante. Also folgt aus Lemma 18.20, dass $d > c^T \cdot B^{-1} \cdot c$. Somit erhalten wir aus Gleichung (18.3), dass der letzte Ausdruck in (18.2) positiv ist, dass also $z^T \cdot B \cdot z + 2\, c^T \cdot z + d > 0$. Es gilt also $y^T \cdot A \cdot y > 0$, und somit $x^T \cdot A \cdot x = x_n^2\, (y^T \cdot A \cdot y) > 0$. Also ist $A$ positiv definit.

■

ANHANG A

# Programme und Unterlagen

## 1. Mathematica-Programme

Die Mathematica-Files `GaussDemo6.m` (Mathematica 5) und `RowRed12.m` (Mathematica 7) enthalten Mathematica-Funktionen, die folgende Probleme mit Zwischenschritten vorrechnen:

- Lösen eines linearen Gleichungssystems (`Gauss[A,b]`).
- Bestimmen einer Matrix in Zeilenstaffelform, die den gleichen Zeilenraum wie die eingegebene Matrix hat (`RowEchelonForm[A]`).
- Bestimmen einer Matrix in Zeilenstaffelnormalform, die den gleichen Zeilenraum wie die eingegebene Matrix hat (`RowEchelonNormalForm[A]`).
- Bestimmen der Determinante einer Matrix (`DeterminantenDemo[A]`).

Die Programme können von Mathematica aus mit `<< GaussDemo6.m` und `<< RowRed12.m` geladen werden.
Sie sind auf `http://www.algebra.uni-linz.ac.at/Students/MathInf/vlws05/MathematicaProgramme/` erhältlich und werden den Studierenden ausschließlich für die Nutzung im Rahmen des Kurses "Lineare Algebra" zur Verfügung gestellt.

## 2. Literatur

Es kann hilfreich sein, auch noch andere Unterlagen als das Vorlesungsskriptum zu kennen. Solche sind etwa (Kommentare dazu in der Vorlesung):

- Kiyek, Karl-Heinz und Schwarz, Friedrich, Lineare Algebra. Teubner Studienbücher Mathematik. B. G. Teubner, Stuttgart, 1999. [**Kiyek and Schwarz, 1999**]
- Peter Weiss, Lineare Algebra und analytische Geometrie, Rudolf Trauner Verlag, Linz, 1980. [**Weiß, 1980**]
- P. Halmos, Finite Dimensional Vector Spaces, D. Van Nostrand Co., Inc., Princeton-Toronto-New York-London, 1958. [**Halmos, 1958**]
- P. Halmos, Naive Mengenlehre, Vandenhoeck & Ruprecht, Göttingen, 1976. [**Halmos, 1976**]

# Literaturverzeichnis


[Halmos, 1958] Halmos, P. R. (1958). *Finite-dimensional vector spaces*. The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Inc., Princeton-Toronto-New York-London. 2nd ed.

[Halmos, 1976] Halmos, P. R. (1976). *Naive Mengenlehre*. Vandenhoeck & Ruprecht, Göttingen. Vierte Auflage, Aus dem Englischen übersetzt von Manfred Armbrust und Fritz Ostermann, Moderne Mathematik in elementarer Darstellung, No. 6.

[Kiyek and Schwarz, 1999] Kiyek, K.-H. and Schwarz, F. (1999). *Lineare Algebra*. Teubner Studienbücher Mathematik. [Teubner Mathematical Textbooks]. B. G. Teubner, Stuttgart.

[Remmert and Ullrich, 1987] Remmert, R. and Ullrich, P. (1987). *Elementare Zahlentheorie*. Birkhäuser Verlag, Basel.

[Weiß, 1980] Weiß, P. (1980). *Lineare Algebra und analytische Geometrie*. Rudolf Trauner, Linz. Eine anwendungsbezogene Einführung. [An application oriented introduction].