Institut für Algebra

Skriptum zur Vorlesung

Lineare Algebra und Analytische Geometrie 1

Wintersemester 2025

Erhard Aichinger
Institut für Algebra
Johannes Kepler Universität Linz

Version 1. Oktober 2025

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

Version 1.10.2025

Druck: Kopierstelle, Abteilung Service, Universität Linz

Vorbemerkungen

Dieses Skriptum basiert auf Skripten zu früheren Auflagen dieser Vorlesung, wie zum Beispiel einem Skriptum aus dem Wintersemester 2010: http://www.algebra.uni-linz.ac.at/Students/LineareAlgebra/linalg1ws10/lineareAlgebraSkriptum06102010.pdf. Viele Anregungen verdanke ich dem Skriptum von Manuel Kauers [**Kau23**], http://www.algebra.uni-linz.ac.at/people/mkauers/linalg/script.pdf zu den Vorlesungen an der JKU aus den Wintersemestern 2015–2024 und der Zusammenarbeit mit Kolleg*innen in zahlreichen Veranstaltungen mit ähnlichen Inhalten. Unter den Büchern über Lineare Algebra möchte ich [**Wei80, KS99, Jän08, FS25**] hervorheben. Manche dieser Bücher sind über die JKU-Bibliothek auch on-line auf https://lisss.jku.at/ verfügbar. Einige Rechnungen werden mit mathematischer Software (Mathematica, https://help.jku.at/im/de/software/software-fuer-studierende oder SageMath https://www.sagemath.org/) durchgeführt. Die dazugehörigen Mathematica-Programme sind auf http://www.algebra.uni-linz.ac.at/Students/MathInf/vlws05/MathematicaProgramme/ zu finden.

Linz, im Oktober 2025 E.A.

Inhaltsverzeichnis

Teil 1

Vektoren, Matrizen und Gleichungssysteme

KAPITEL 1

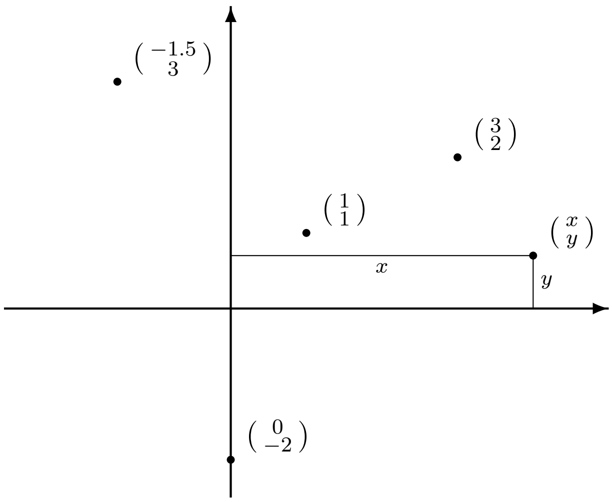
Geometrie in der Ebene und im Raum

1. Koordinaten

Wir beschreiben 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.

$$\mathbb{R} \times \mathbb{R} := \{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \mid x \in \mathbb{R} \text{ und } y \in \mathbb{R}\}.$$

Für das Paar $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ schreiben wir auch (x, y). Aus der folgenden Skizze ist ersichtlich, wie wir jeden Punkt durch ein Zahlenpaar (seine *kartesischen*[1] *Koordinaten*) beschreiben.



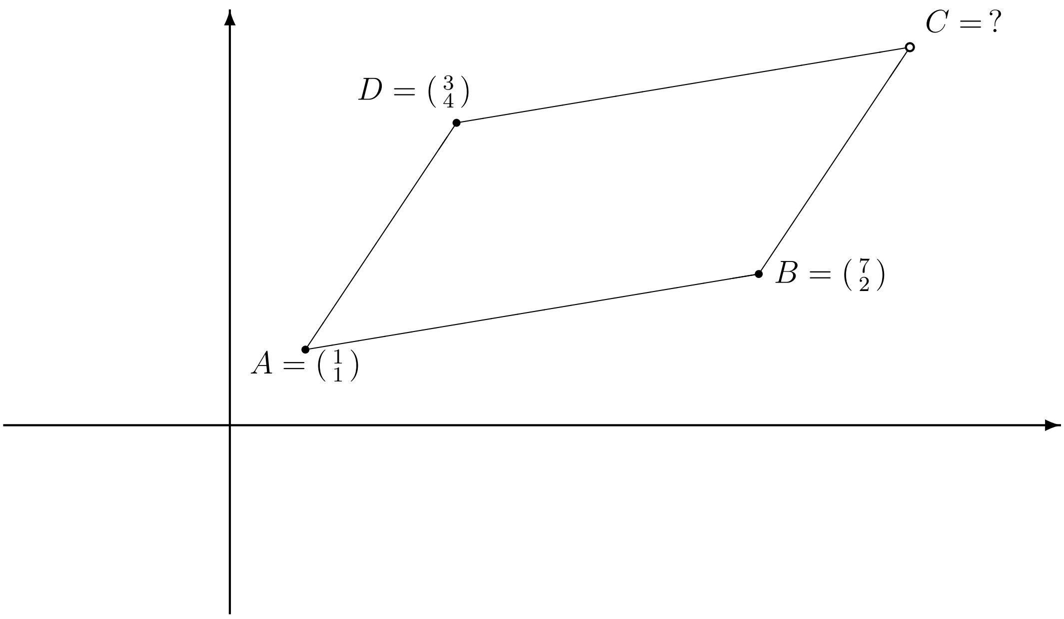
2. Vektoren

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

$$A = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right),\ B = \left(\begin{smallmatrix} 7 \\ 2 \end{smallmatrix}\right),\ D = \left(\begin{smallmatrix} 3 \\ 4 \end{smallmatrix}\right)$$

gegeben sind?

[1]nach René Descartes (1596 – 1650). Koordinaten wurden aber schon früher verwendet.



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

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

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

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

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. Länge eines Vektors

Wir lösen folgendes Beispiel:

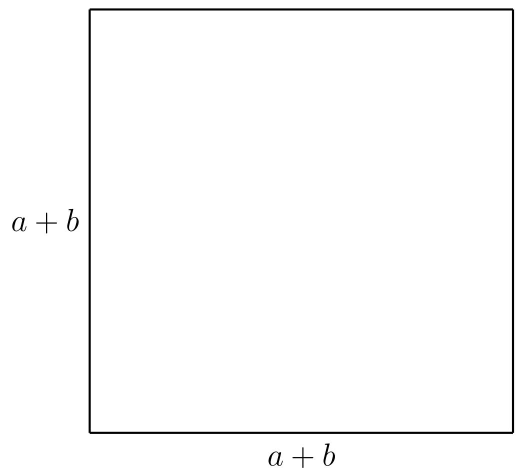
> Herr A geht von $\binom{3}{2}$ aus 1 Einheit in Richtung Südosten. Wo landet er?

"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

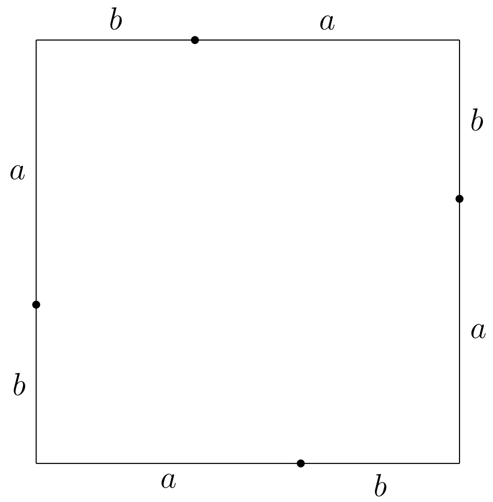
$$Z = \binom{3}{2} + \frac{1}{\sqrt{2}} \cdot \binom{1}{-1} \approx \binom{3.7}{1.3}$$

ausrechnen.

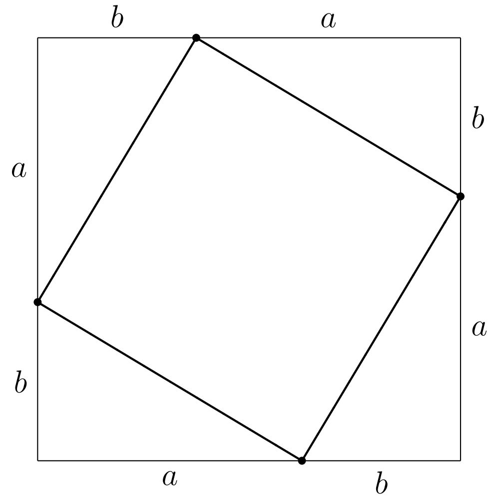
Wir überlegen uns jetzt, wie lange der Vektor $\binom{a}{b}$ ist. Das heißt, wir wollen wissen, wie lange in einem Dreieck, in dem die Seiten mit den Längen a und b einen rechten Winkel einschließen, die dem rechten Winkel gegenüberliegende Seite ist. Vergessen wir kurz unsere klassische Bildung, und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



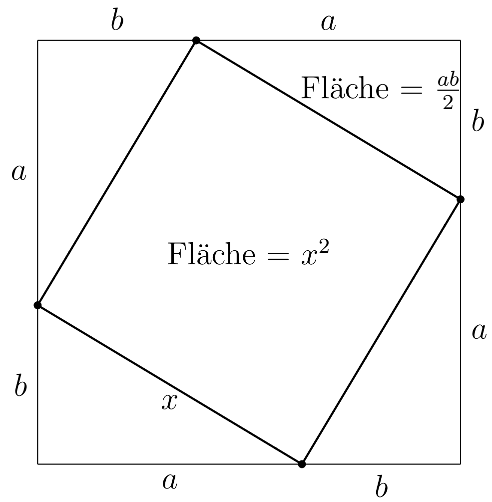
Jetzt unterteilen wir jede der vier Quadratseiten in ein Stück der Länge a und ein Stück der Länge b.



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° 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°, und daher ist in jedem Viereck die Winkelsumme 360°. Also ist jeder Winkel des inneren Vierecks gleich $\frac{360°}{4} = 90°$. 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

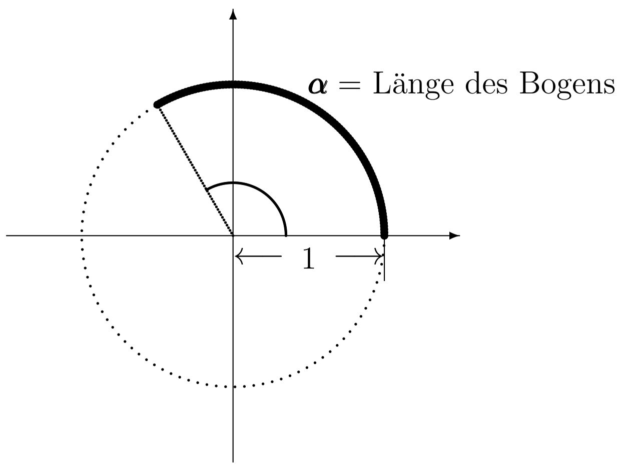
$$\left\|\binom{a}{b}\right\| = \sqrt{a^2 + b^2}.$$

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 Dreieck 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:

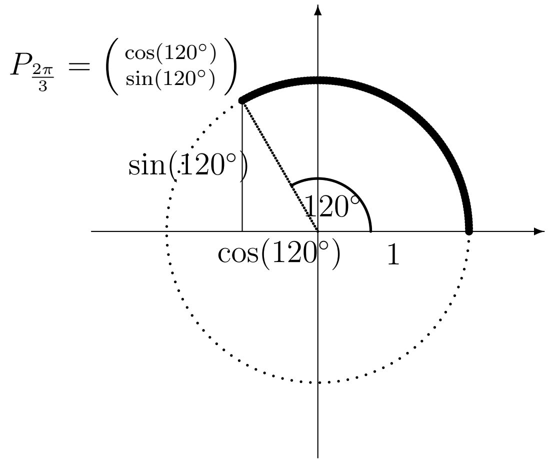
(1) Man tabelliert den Zusammenhang zwischen den Seitenlängen und den Winkeln für rechtwinkelige Dreiecke. Dazu braucht man die *Winkelfunktionen* sin (Sinus) und cos (Cosinus).
(2) Man baut sich alle anderen Dreiecke aus rechtwinkeligen Dreiecken zusammen. Da dieses Zusammenbauen aber immer gleich funktioniert, 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 (◦), 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 π rad. Demzufolge ist $1^\circ = \frac{\pi}{180}$ rad , und 1 rad $\approx 57.2958^\circ$.

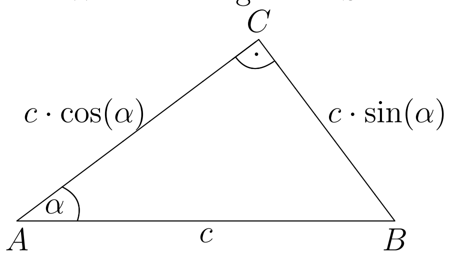
4.2. Winkelfunktionen.



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 $\begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}$. Nach dem Satz des Pythagoras gilt für jeden Winkel x:

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

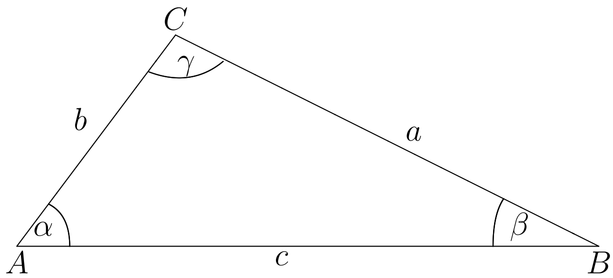
In einem rechtwinkeligen Dreieck heißt die dem rechten Winkel gegenüberliegende Seite auch *Hypotenuse*, die beiden dem rechten Winkel anliegenden Seiten heißen *Katheten*.



Übungsaufgaben 1.1.

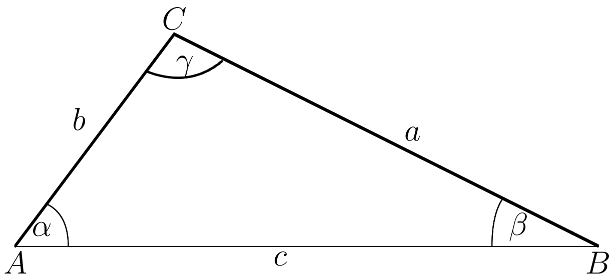
(1) Ein Kletterer kann Wände mit einer Neigung von maximal 65° besteigen. Schafft er eine Pyramide mit einer quadratischen Grundfläche von 784 m^2 und einer Höhe von 40 m?

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 α, den der Seite mit Länge b gegenüber liegenden Winkel mit β, und den der Seite mit Länge c gegenüber liegenden Winkel mit γ. 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 γ.
- ▷ Gesucht: Die fehlende Seitenlänge c.



Wir betrachten zunächst den Fall $\gamma \leq 90°$, $\alpha \leq 90°$. 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

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

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

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

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

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

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 γ. Dann gilt*

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

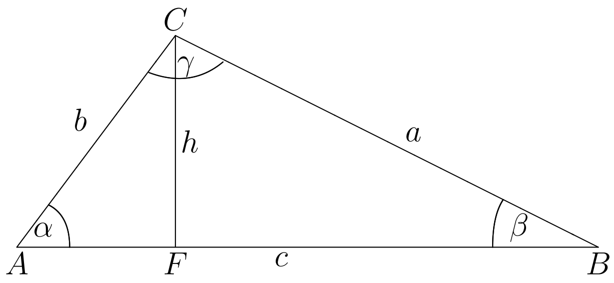
Man findet mit dem Cosinussatz γ, wenn a, b und c gegeben sind. Zu jedem $y \in [-1, 1]$ gibt es genau ein $x \in [0, \pi]$, sodass $\cos(x) = y$.

4.4. Der Sinussatz.

Der Sinussatz löst folgendes Problem:

- ▷ Gegeben: α, β, 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° oder 180° 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 α, den der Seite mit Länge b gegenüber liegenden Winkel mit β, und den der Seite mit Länge c gegenüber liegenden Winkel mit γ. Sei d der Durchmesser des Umkreises des Dreiecks. Dann gilt:*

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

Für den Beweis, dass $\frac{a}{\sin(\alpha)}$ gleich dem Durchmesser des Umkreises ist, kann man etwa den *Randwinkelsatz*[2] verwenden. 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 α, den der Seite mit Länge b gegenüber liegenden Winkel mit β, und den der Seite mit Länge c gegenüber liegenden Winkel mit γ. 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 α 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° ist, kennt man tatsächlich alle drei Winkel. Sind also c, α, und β 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 α gegeben. Wenn b und c positive reelle Zahlen sind, und $0 <$

[2]Seien A, B, C verschiedene Punkte auf einer Kreislinie mit Mittelpunkt M. Wir nehmen an, dass C und M auf der gleichen Seite der Sehne AB liegen. Dann gilt $\measuredangle BCA = 2 \cdot \measuredangle BMA$.

$\alpha < \pi$ 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, α und a gegeben. In diesem Fall kann es sein, dass es gar kein, genau ein oder genau zwei Dreiecke mit dem vorgegebenen c, α und a gibt. Es gibt mehrere Fälle:
   (a) *Es gilt $\alpha < 90°$:*
      (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 $\alpha = 90°$:*
      (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 $\alpha > 90°$:*
      (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.

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 α, den der Seite mit Länge b gegenüber liegenden Winkel mit β, und den der Seite mit Länge c gegenüber liegenden Winkel mit γ. (Die Seiten seien dabei *gegen den Uhrzeigersinn* mit a, b, c beschriftet.)

(1) Berechnen Sie $\sin(\gamma)$ für ein Dreieck mit $c = 10$, $b = \frac{10}{\sqrt{3}}$, $\beta = 30°$. Das Dreieck ist mit diesen drei Bestimmungsstücken c, b, β noch nicht eindeutig festgelegt. Warum nicht?
(2) Wie groß kann b in einem Dreieck mit $\alpha = 45°$, $c = 1$, $a = \frac{2}{\sqrt{6}}$ sein? (Gibt es mehr als eine Lösung?)
(3) Geben Sie eine Wahl von a an, sodass es genau ein Dreieck mit den Bestimmungsstücken $\alpha = 45°$, $c = 1$, und Ihrem gewählten a gibt!
(4) Von einem Dreieck ABC haben Sie folgende Information: $\overline{AB} = 10\,\text{cm}$, der Winkel α zwischen AB und AC ist $30°$, $\overline{BC} = \frac{10}{\sqrt{2}}\,\text{cm}$.
   (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 $c > 0$, $b > 0$, α 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, β 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, α, β gibt es
(a) gar kein Dreieck
(b) genau ein Dreieck
(c) genau zwei Dreiecke
(d) mehr als zwei Dreiecke

mit den Bestimmungsstücken c, α, β?

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

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

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

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

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

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

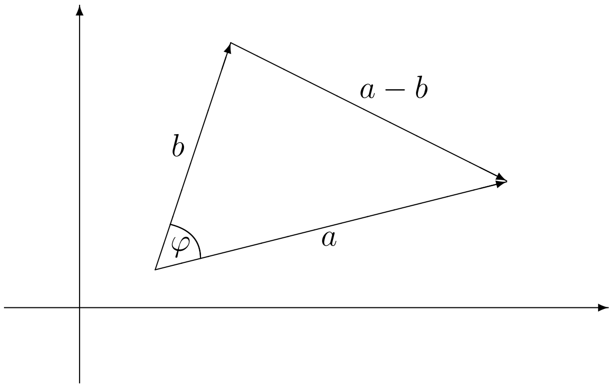
(19) Sie möchten die Entfernung eines Punktes B an einem Ufer eines Flusses zu einem Punkt C auf der anderen Seite des Flusses bestimmen. Dazu gehen Sie folgendermaßen vor: Sie messen an Ihrem Flussufer die Strecke von B zu einem weiteren Punkt A ab. Diese Strecke ist 500 m lang. Der Winkel zwischen BC und BA ist 60°, der Winkel zwischen AB und AC ist 30°.
(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 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 α zwischen der Strecke AB und AC, sowie den Winkel γ zwischen AC und CB. Was ergibt sich für die Entfernung AB bei $AC = 27.5$ m und $\alpha = 73°$, $\gamma = 65°$?

(21) 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. Winkel zwischen zwei Vektoren

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



Für den eingeschlossenen Winkel φ gilt nach dem Cosinussatz:

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

Diese Formel können wir vereinfachen:

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

Wir erhalten

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

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

$$\langle a, b\rangle = \langle \binom{a_1}{a_2}, \binom{b_1}{b_2}\rangle = a_1\cdot b_1 + a_2\cdot b_2.$$

Die Winkelformel heißt jetzt:

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

Außerdem gilt für jeden Vektor a

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

```
In[1]:= {3,4}.{-5,2}
Out[1]= -7
In[2]:= Laenge[v_]:=Sqrt[v.v]
In[3]:= Laenge[{2,3}]
Out[3]= Sqrt[13]
```

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

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

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 $\binom{c}{d} \neq \binom{0}{0}$) der Vektor $\binom{c}{d}$ mit den Vektoren $\binom{d}{-c}$ und $\binom{-d}{c}$ einen rechten Winkel einschließt.

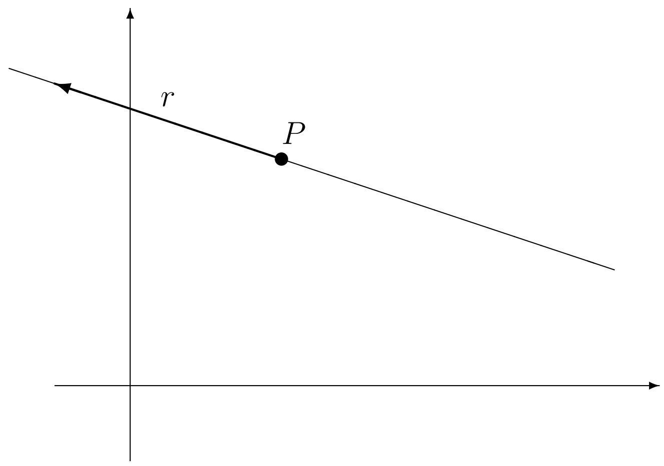
Übungsaufgaben 1.6.

(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) $\langle a+b, a-b\rangle = \langle a, a\rangle - \langle b, b\rangle$.
  (c) $\langle \lambda a, b\rangle = \lambda\langle a, b\rangle$.
(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.



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

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

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

Lies: "g ist gleich der Menge aller Punkte $P + \lambda$ mal r, wobei λ eine reelle Zahl ist." Mit den Zahlen für P und r:

$$g = \{\left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right) \mid \lambda \in \mathbb{R}\},$$

oder, anders geschrieben,

$$g = \{\left(\begin{smallmatrix} 2-3\lambda \\ 3+\lambda \end{smallmatrix}\right) \mid \lambda \in \mathbb{R}\}.$$

Man kann g auch so schreiben:

$$g = \{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 \mid \text{ es gibt } \lambda \in \mathbb{R}, \text{ sodass } \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right)\}.$$

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

$$g : \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right)$$

oder

$$g : X = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right).$$

Wir überprüfen, ob der Punkt $\left(\begin{smallmatrix} -4 \\ 5 \end{smallmatrix}\right)$ auf der Geraden g liegt. Er liegt auf g, falls es eine reelle Zahl λ gibt, sodass $\left(\begin{smallmatrix} -4 \\ 5 \end{smallmatrix}\right)$ gleich $\left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right)$ 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 $\left(\begin{smallmatrix} -4 \\ 5 \end{smallmatrix}\right)$ auf g; wir schreiben dafür

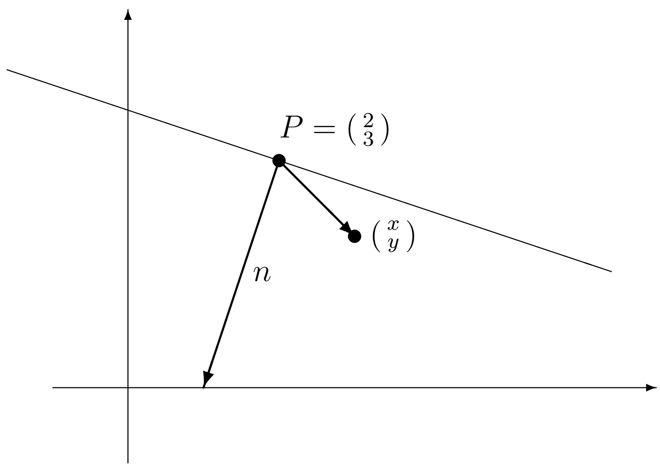
$$\left(\begin{smallmatrix} -4 \\ 5 \end{smallmatrix}\right) \in g.$$

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 λ 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 $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ auf der Geraden g liegt, die durch

$$g : X = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right)$$

gegeben ist.



Anstatt zu fragen, ob $\binom{x}{y}$ auf g liegt, fragen wir, ob $\binom{x}{y} - \binom{2}{3}$ normal auf n ist. Das ist nämlich genau für die Punkte $\binom{x}{y}$ auf g der Fall. Zunächst finden wir den Vektor 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 n durch

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

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 $g : 5\,x - 2\,y = 1$ 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 $\begin{pmatrix} \frac{1}{5} + \frac{2}{5}\,t \\ t \end{pmatrix}$ 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

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

oder

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

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

$$g : X = \binom{0}{-1} + t \cdot \binom{1}{0}.$$

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) $3\,x + 4\,y = 17$.
   (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.

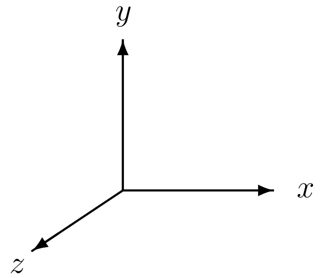
(12) Berechnen Sie den Durchschnitt der Geraden h und j, wobei
$$h : X = \binom{1}{2} + t \cdot \binom{2}{5}$$
und
$$j : 10\,x - 4\,y = 0.$$

(13) Bestimmen Sie die Schnittmenge der Geraden
$$g_1 : X = \binom{3}{4} + t \cdot \binom{-2}{1}$$
und
$$g_2 : 2\,x + 4\,y = 22.$$

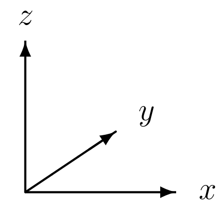
(14) Bestimmen Sie den Cosinus des Schnittwinkels der Geraden
$$g_1 : X = \binom{0}{4} + t \cdot \binom{3}{4}$$
und
$$g_2 : 12\,x - 5\,y = 22.$$

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

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

wobei n eine beliebige natürliche Zahl ist.

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

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

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

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

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

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

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

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

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

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

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

Nun gilt

$$\langle a - \frac{\langle a, b \rangle}{\langle b, b \rangle} \cdot b, a - \frac{\langle a, b \rangle}{\langle b, b \rangle} \cdot b \rangle = \langle a, a \rangle - 2\frac{\langle a, b \rangle^2}{\langle b, b \rangle} + \frac{\langle a, b \rangle^2}{\langle b, b \rangle} = \langle a, a \rangle - \frac{\langle a, b \rangle^2}{\langle b, b \rangle}. \tag{1.1}$$

Da dieser Ausdruck ≥ 0 ist, gilt

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

Folglich gilt auch

$$\|a\| \cdot \|b\| \geq |\langle a, b \rangle|.$$

Somit gilt (1).

Um (2) zu zeigen, nehmen wir an, dass $|\langle a, b \rangle| = \|a\| \cdot \|b\|$ gilt. Wir zeigen, dass dann $b = 0$ gilt, oder dass a ein Vielfaches von b ist. Es gilt $\langle a, b \rangle^2 = \langle a, a \rangle \cdot \langle b, b \rangle$. Wenn also $b \neq 0$, so gilt wegen (1.1)

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

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

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

ÜBUNGSAUFGABEN 1.12.

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

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

SATZ 1.13. *Seien $a, b \in \mathbb{R}^n$ mit $a \neq 0$ und $b \neq 0$, und sei φ der Winkel, den a und b einschließen. Dann gilt*

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

Beweisskizze: $\langle a, a \rangle + \langle b, b \rangle - 2\|a\| \, \|b\| \cos(\varphi) = \langle a - b, a - b \rangle$. □

Im $\mathbb{R}^3$ gibt es eine Rechenoperation, die einen Vektor liefert, der auf zwei gegebene Vektoren a und 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 $\mathbb{R}^3$.

Zunächst ist völlig unklar, warum man $a \times b$ genau so definiert[3]. Der folgende Satz belegt aber, dass das Kreuzprodukt nützlich ist.

SATZ 1.15. *Seien* $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ *und* c *in* $\mathbb{R}^3$, *und sei* φ *der von* a *und* b *eingeschlossene Winkel. Dann gilt:*

(1) $a \times b$ *ist normal auf* a *und auf* b.
(2) $\|a \times b\|^2 = \|a\|^2 \cdot \|b\|^2 - \langle a, b\rangle^2$.
(3) $\|a \times b\| = \|a\| \cdot \|b\| \cdot \sin(\varphi)$.
(4) $\|a \times b\|$ *ist die Fläche des Parallelogramms, das von den Vektoren* a *und* b *aufgespannt wird.*
(5) *Das Volumen* V *des Parallelepipeds* $P = \{\alpha a + \beta b + \gamma c \mid \alpha, \beta, \gamma \in [0, 1]\}$ *erfüllt* $V = |\langle a \times b, c\rangle|$.

Beweis: (1) Es gilt

$$\langle a \times b, a\rangle = (a_2b_3 - a_3b_2)a_1 - (a_1b_3 - a_3b_1)a_2 + (a_1b_2 - a_2b_1)a_3 = 0.$$

$\langle a \times b, b\rangle = 0$ wird in der Vorlesung nachgerechnet.

(2) Es gilt

$$\begin{aligned} \|a \times b\|^2 &= (a_1b_2 - a_2b_1)^2 + (a_3b_1 - a_1b_3)^2 + (a_2b_3 - a_3b_2)^2 \\ &= a_1^2b_2^2 + a_1^2b_3^2 - 2a_1a_2b_1b_2 - 2a_1a_3b_1b_3 + a_2^2b_1^2 + a_2^2b_3^2 - 2a_2a_3b_2b_3 + a_3^2b_1^2 + a_3^2b_2^2 \end{aligned}$$

und

$$\begin{aligned} \|a\|^2\|b\|^2 - \langle a, b\rangle^2 &= \left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\ &= a_1^2b_2^2 + a_1^2b_3^2 - 2a_1a_2b_1b_2 - 2a_1a_3b_1b_3 + a_2^2b_1^2 + a_2^2b_3^2 - 2a_2a_3b_2b_3 \\ &\quad + a_3^2b_1^2 + a_3^2b_2^2. \end{aligned}$$

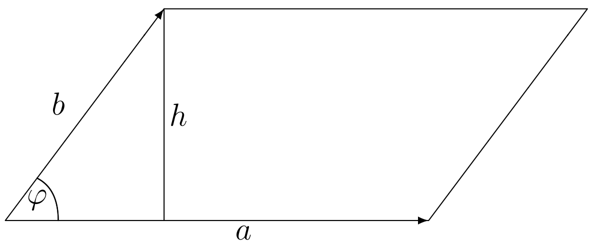
(3) Es gilt

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

Wegen (2) ist der letzte Ausdruck gleich $\|a \times b\|^2$. Da $\sin(\varphi) \geq 0$, folgt die behauptete Gleichheit (3) durch Wurzelziehen.

(4) Wir nehmen an, dass $a \neq 0$ und $b \neq 0$.

[3]Die Definition lässt sich mithilfe von *Determinanten* motivieren: $a \times b$ ist jener Vektor, sodass für jeden Vektor $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ die Determinante der Matrix $\begin{pmatrix} a_1 & b_1 & x_1 \\ a_2 & b_2 & x_2 \\ a_3 & b_3 & x_3 \end{pmatrix}$ gleich $\langle a \times b, x\rangle$ ist. Da uns aber *Determinanten* und *Matrizen* noch nicht zur Verfügung stehen, brauchen wir die etwas rechnerische Definition 1.14.



Wir erhalten für die Höhe h auf a
$$h = \|b\| \cdot \sin(\varphi)$$
und für den Flächeninhalt
$$F = \|a\| \cdot h = \|a\| \cdot \|b\| \cdot \sin(\varphi).$$
Wegen (3) ist das gleich $\|a \times b\|$.

(5) Wir wählen $A = \{\alpha a + \beta b \mid \alpha, \beta \in [0,1]\}$ als Grundfläche. Die Höhe auf A ist gleich der Projektion von c auf $a \times b$. Wenn $a \times b = 0$, dann zeigen a und b in die gleiche Richtung und das Volumen ist 0. Wir nehmen nun an, dass $a \times b \neq 0$. Wegen der Projektionseigenschaft des Skalarprodukts gilt für die Länge der Höhe $h = \|\frac{\langle c, a\times b\rangle}{\langle a\times b, a\times b\rangle} \cdot (a \times b)\| = |\frac{\langle c, a\times b\rangle}{\|a\times b\|^2}| \cdot \|a \times b\| = \frac{|\langle c, a\times b\rangle|}{\|a\times b\|}$. Das Volumen ist das Produkt aus der Grundfläche $\|a \times b\|$ und h, also $|\langle c, a \times b\rangle|$.

□

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

Die Mathematica-Funktion `Cross` liefert das Kreuzprodukt.

Übungsaufgaben 1.16.

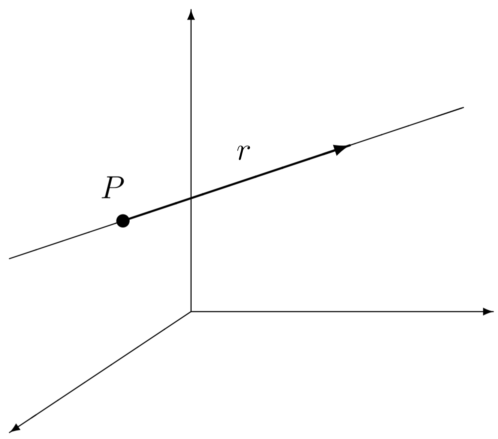
(1) Zeigen Sie, dass für alle $a, b \in \mathbb{R}^3$ gilt: $a \times b = -b \times a$.
(2) Seien $a, b \in \mathbb{R}^3$. Zeigen Sie unter Verwendung der Gleichheitsaussage in der Cauchy-Ungleichung (Satz 1.11) und Satz 1.15 (2), dass $a \times b$ genau dann 0 ist, wenn $a = 0$ oder b ein Vielfaches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

8. Geraden und Ebenen im Raum

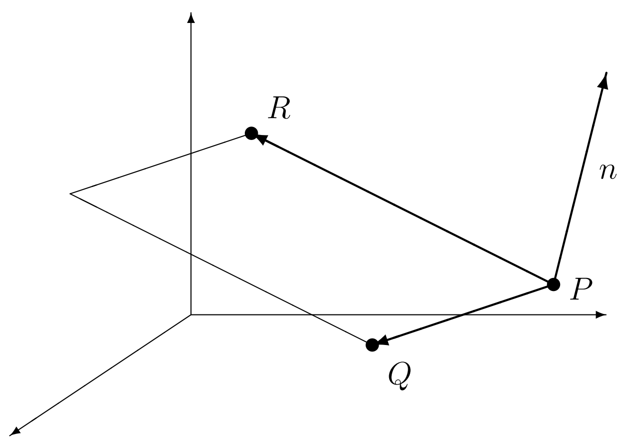
8.1. Parameterdarstellung einer Geraden. Genau wie im $\mathbb{R}^2$ lässt sich eine Gerade im $\mathbb{R}^3$ durch eine Parameterdarstellung mit einem Punkt und einem Richtungsvektor angeben.

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.

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

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 $\left(\begin{smallmatrix}2\\-1\\4\end{smallmatrix}\right)$, die normal auf den Vektor $n = \left(\begin{smallmatrix}1\\3\\-2\end{smallmatrix}\right)$ ist. Wir nennen n den Normalvektor von e.

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

$$\langle \left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) - \left(\begin{smallmatrix}2\\-1\\4\end{smallmatrix}\right), n\rangle = 0.$$

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

$$e = \{\left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) \in \mathbb{R}^3 \,|\, x + 3y - 2z = -9\}.$$

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

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

$$e : \left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) = \left(\begin{smallmatrix}2\\3\\-1\end{smallmatrix}\right) + \lambda \cdot \left(\begin{smallmatrix}1\\-1\\3\end{smallmatrix}\right) + \mu \cdot \left(\begin{smallmatrix}2\\-1\\-3\end{smallmatrix}\right)$$

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

(1) Wir suchen $n = \left(\begin{smallmatrix}n_1\\n_2\\n_3\end{smallmatrix}\right)$ so, dass

$$\begin{aligned}\langle \left(\begin{smallmatrix}1\\-1\\3\end{smallmatrix}\right), n\rangle &= 0 \text{ und}\\ \langle \left(\begin{smallmatrix}2\\-1\\-3\end{smallmatrix}\right), n\rangle &= 0.\end{aligned}$$

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 n, der nicht der Nullvektor ist. Wie man alle Lösungen eines linearen Gleichungssystems findet, werden wir im Kapitel 3 sehen.

(2) Alternativ finden wir n auch als Kreuzprodukt der beiden Richtungsvektoren.

$$n = \left(\begin{smallmatrix}1\\-1\\3\end{smallmatrix}\right) \times \left(\begin{smallmatrix}2\\-1\\-3\end{smallmatrix}\right) = \left(\begin{smallmatrix}6\\9\\1\end{smallmatrix}\right)$$

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

$$\langle \left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) - \left(\begin{smallmatrix}2\\3\\-1\end{smallmatrix}\right), n\rangle = 0$$

und erhalten

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

Somit hat die Ebene e die implizite Darstellung

$$e = \{\left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right) \in \mathbb{R}^3 \,|\, 6x + 9y + z = 38\}.$$

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 λ und μ ausdrücken. Wir erhalten $x = -9 - 3\lambda + 2\mu$. Somit liegt für alle Werte $\lambda, \mu \in \mathbb{R}$ der Punkt $\left(\begin{smallmatrix}-9-3\lambda+2\mu\\\lambda\\\mu\end{smallmatrix}\right)$ 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

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

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 n, dann ist 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 $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$, 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

$$g = \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,|\, x + y = 5, 3x - z = 7\}.$$

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

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

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. 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 2×3-Matrix.
- ▷ $\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$ ist eine 2×2-Matrix.
- ▷ $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ ist eine 2×1-Matrix.
- ▷ $\begin{pmatrix} 1 & 2 & 7 \end{pmatrix}$ ist eine 1×3-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 mit Einträgen aus $\mathbb{R}$ 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[3]:= A={{1,2,3},{4,5,6}}
Out[3]= {{1,2,3},{4,5,6}}

In[4]:= MatrixForm[A]
```

```
Out[4]= (1 2 3
         4 5 6)

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

In[6]:= MatrixForm[A]
Out[6]= (5 7 8
        -2 3 5)

In[7]:= A[[2]][[1]]
Out[7]= -2
```

---

2. 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.

Beispiele 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 $i \in \{1, 2, \dots, m\}$ und $j \in \{1, 2, \dots k\}$ berechnet man den Eintrag $C_{i,j}$ durch

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

```
In[8]:= A={{1,4,3},{0,1,0}};
In[9]:= B={{-7,5,0},{23,-7,16}};

In[10]:= A+B
Out[10]= {{-6,9,3},{23,-6,16}}

In[11]:= MatrixForm[%]
Out[11]= (-6    9   3
           23  -6  16)
```

---

3. Vervielfachen einer Matrix

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

Beispiel 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[12]:= A={{2,5},{3,4},{10,2}}
Out[12]= {{2,5},{3,4},{10,2}}

In[13]:= MatrixForm[(-10)*A]
Out[13]=
(-20 -50
 -30 -40
-100 -20)
```

---

4. 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.

Beispiel 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:

Beispiele 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.

Wenn A eine 2×3 und B eine 3×1-Matrix ist, dann ist das Produkt $A \cdot B$ eine 2×1-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:

Beispiele 2.6.

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

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

▷ $\begin{pmatrix} 1 & 2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix} = (17)$.

▷ $\begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 25 \\ 1 & 2 & 5 \\ 2 & 4 & 10 \end{pmatrix}$.

Das erste Beispiel noch einmal in Mathematica:

```
In[14]:=
A={{3,1,2},{2,5,4}};
B={{3,9,3},{1,8,5},{7,1,-4}};

In[16]:= MatrixForm[A.B]
Out[16]//MatrixForm=
(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 $i \in \{1, 2, \ldots, k\}$ und $j \in \{1, 2, \ldots, n\}$ 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 2×2-Matrizen sind. Man berechnet

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

und

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

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 $A \cdot (B \cdot C)$ beide $k \times n$-Matrizen sind. Um zu zeigen, dass beide Matrizen gleich sind, zeigen wir, dass für alle $i \in \{1, \ldots, k\}$ und $j \in \{1, \ldots, n\}$ gilt, dass $(A \cdot B) \cdot C\,[i, j] = A \cdot (B \cdot C)\,[i, j]$.

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

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

zeigen. Es gilt

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

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

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

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. 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 3×2-Matrix $\begin{pmatrix} -3 & 4 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$ mit der 2×1-Matrix $\left(\begin{smallmatrix} -2 \\ 4 \end{smallmatrix}\right)$. Bei der Matrizenmultiplikation ist das Ergebnis aber eine 3×1-Matrix.

In Mathematica wird der Unterschied deutlich:

```
In[17]:=
A={{-3,4},{1,0},{-1,1}};
v={-2,4};
x=A.v
Out[19]= {22,-2,6}

In[20]:=
A={{-3,4},{1,0},{-1,1}};
v={{-2},{4}};
x=A.v
Out[22]= {{22},{-2},{6}}

In[23]:= A={{-3,4},{1,0},{-1,1}};
v={{-2,4}};
x=A.v

Dot::dotsh:Tensors {{-3,4},{1,0},{-1,1}} and {{-2,4}} have incompatible shapes.

Out[25]= {{-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 1×3-Matrix $(-4\ 3\ 2)$ mit der 3×2-Matrix $\begin{pmatrix} -3 & 5 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$. Bei der Matrizenmultiplikation ist das Ergebnis aber eine 1×2-Matrix.

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

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

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 $2\,x - 3\,y = 5$.

```
In[26]:= LinearSolve[{{2,-3}},{5}]
Out[26]= {5/2, 0}
```

Später werden wir sehen, wie man alle Lösungen erhält.

7. 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[27]:= A={{1,4,-3},{2,-5,3}};
In[28]:= MatrixForm[A]
Out[28]//MatrixForm=
(1  4 -3
 2 -5  3)

In[29]:= B=Transpose[A]
Out[29]= {{1,2},{4,-5},{-3,3}}

In[30]:= MatrixForm[B]
Out[30]//MatrixForm=
(1  2
 4 -5
-3  3)
```

Satz 2.12. *Seien $l, m, n \in \mathbb{N}$, sei $A \in \mathbb{R}^{l\times m}$, und seien $B, C \in \mathbb{R}^{m\times n}$. 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 $A = \left(\begin{smallmatrix} -4 & 0 \\ 1 & 2 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} -4 & -8 \\ 7 & 10 \end{smallmatrix}\right)$. (Hinweis: Bestimmen Sie jede Spalte von X durch Lösen eines linearen Gleichungssystems.)

8. Einheitsmatrizen

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

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[31]:=
  A=-24*IdentityMatrix[5];
  MatrixForm[A]

Out[32]//MatrixForm=
(-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)
```

---

9. 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:

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

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 $(A_1 \cdot A_2) \cdot (B_2 \cdot B_1) = A_1 \cdot (A_2 \cdot B_2) \cdot B_1 = A_1 \cdot E_n \cdot B_1 = A_1 \cdot B_1 = E_n$ und $(B_2 \cdot B_1) \cdot (A_1 \cdot A_2) = \cdots = E_n$. 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

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^{-1} = \frac{1}{ad - bc} \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$$

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 können wir erst später beweisen (als Konsequenz von Satz 11.27), wenn wir Unterräume und deren Dimension kennen.

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[34]= {2/5,1/5}
```

Also $a = 0.4$, $c = 0.2$.

```
In[35]:=
  A={{1,3},{2,-4}};
  LinearSolve[A,{0,1}]
  Out[36]= {3/10,-(1/10)}
```

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

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

```
In[37]:=
  A={{1,3},{2,-4}};
  B=Inverse[A];
  MatrixForm[B]
Out[39]//MatrixForm=
(2/5   3/10
 1/5 -(1/10))

In[40]:= A.B
Out[40]= {{1,0},{0,1}}

In[41]:= B.A
Out[41]= {{1,0},{0,1}}
```

```
In[42]:= A^(-1)
Out[42]= {{1,1/3},{1/2,-(1/4)}}

In[43]:= A.A^(-1)
Out[43]= {{5/2,-(5/12)},{0,5/3}}
```

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.

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

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

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

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

Wenn $-3\,x - 3\,y = 3$ und $3\,x + 2\,y = -5$, dann gilt auch $(-3\,x - 3\,y) + (3\,x + 2\,y) = 3 + (-5)$, 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 $3\,x+2\,y = -5$ erfüllen, liegen auf der Geraden mit Gleichung $3\,x + 2\,y = -5$. 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 $\binom{-3}{2}$.

Wir suchen nun alle Lösungen des Gleichungssystems

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

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

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

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

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

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

$$L = \{\} = \varnothing.$$

Wir interpretieren dieses Beispiel jetzt geometrisch. Die Gerade $x + 3\,y = -1$ hat den Normalvektor $\binom{1}{3}$. Die Gerade $-3\,x - 9\,y = 2$ 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 gleich, 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.

Als nächstes suchen wir alle Lösungen des Gleichungssystems

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

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

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

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

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

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

$$x + 5\,y = -4$$

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 $x = -4 - 5\,t$. Somit können wir die Lösungsmenge L so angeben:

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

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

Wir interpretieren dieses Beispiel jetzt geometrisch. Die Gleichungen $x + 5\,y = -4$ und $-2\,x - 10\,y = 8$ 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 $x + 5\,y = -4$.

Wir suchen nun alle Lösungen des Gleichungssystems

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

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

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

Jede Lösung von $(1')$ und $(2')$ erfüllt auch

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

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

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

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

Die Frage ist, ob $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ auch wirklich eine Lösung des Gleichungssystems ist. Wir haben bis jetzt ja nur so begründet, dass für jede Lösung $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ des Gleichungssystems $x = 1$ und $y = -1$ gelten muss. Wir wissen aber noch nicht, ob $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ 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 $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ 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 $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ erfüllt also wirklich alle drei Gleichungen, und es gilt somit

$$L = \{\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\}.$$

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 $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$. Die drei Geraden gehen also "sternförmig" durch den Punkt $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$.

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

(2) *Parametrisierte Darstellung* der Lösung: Wir können die Lösungsmenge als
$$L = \{(-4, 0) + t \cdot (-5, 1) \,|\, t \in \mathbb{R}\}$$
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 $x + 5\,y$ einsetzen. Da wir dabei 23 und nicht -4 erhalten, folgt $(3, 4) \notin L$. Bei parametrisierter Darstellung müssen wir dazu das Gleichungssystem

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

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

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

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. Lösung von Gleichungssystemen in Treppenform

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

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

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 *Treppenform*. Wir definieren:

DEFINITION 3.1. Sei A eine $m \times n$-Matrix. A ist in *Treppenform*, 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 Treppenform 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 Treppenform 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
$$\begin{pmatrix} 1 & -5 & 8 & 2 & -2 \\ 1 & -4 & 6 & -2 & 0 \\ -1 & 0 & 2 & 2 & 0 \\ 5 & -8 & 6 & 0 & -5 \end{pmatrix} \cdot x = \begin{pmatrix} -9 \\ -4 \\ 0 \\ -18 \end{pmatrix} \qquad (3.12)$$
Wir schreiben uns zunächst das System anders auf.

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

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

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

Nun hat das Gleichungssystem, das aus den Gleichungen $\mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}$ besteht, die gleiche Lösungsmenge wie das Gleichungssystem, das aus den Gleichungen $\mathrm{I}, \mathrm{II}', \mathrm{III}', \mathrm{IV}'$ 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 (= Summe 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 $\mathrm{IV}' = -5 \cdot \mathrm{I} + \mathrm{IV}$, gilt $\mathrm{IV} = \mathrm{IV}' + 5 \cdot \mathrm{I}$. 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 $\mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}$ das Gleichungssystem $\mathrm{I}, \mathrm{II}', \mathrm{III}', \mathrm{IV}'$ lösen. In den Gleichungen $\mathrm{II}', \mathrm{III}', \mathrm{IV}'$ 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 $\mathrm{II}', \mathrm{III}', \mathrm{IV}'$. Für jede dieser Lösungen rechnen wir uns dann aus der Gleichung I den Wert von x_1 aus.

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

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

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

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

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

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

Übungsaufgaben 3.3.

(1) Bestimmen Sie alle Lösungen des Gleichungssystems

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

(2) Bestimmen Sie alle Lösungen des Gleichungssystems

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

(3) Bestimmen Sie alle Lösungen des Gleichungssystems

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

(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, wenn möglich, die Gleichung

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

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. Durchgerechnete Beispiele zum Gauß-Algorithmus

```
In[110]:= << GaussDemo7.m
In[111]:= (* Beispiel 1 *)

In[112]:= A1={{1,-5,8,2,-2},{1,-4,6,-2,0},{-1,0,2,2,0},{5,-8,6,0,-5}};
In[113]:= b1={-9,-4,0,-18};

In[114]:= Gauss [A1,b1]

(x1 x2 x3 x4 x5|
 1 -5  8  2 -2  |  -9
 1 -4  6 -2  0  |  -4
-1  0  2  2  0  |   0
 5 -8  6  0 -5  | -18)

We use equation (1) of the last system to eliminate x1
( x2 x3  x4 x5  |
  1  -2  -4  2  |  5
 -5  10   4 -2  | -9
 17 -34 -10  5  |  27)

We use equation (1) of the last system to eliminate x2
(x3 x4 x5  |
 0 -16   8  |  16
 0  58 -29  |  -58)

x3 does not appear in any equation.
(x4  x5  |
 -16  8  |  16
  58-29  |  -58)

We use equation (1) of the last system to eliminate x4
(x5  |
```

```
  0  |  0)

x5 does not appear in any equation.
x5:=t1

We use
(x4 x5  |
-16 8  |  16)
to compute x4
x4 = -1 + t1/2

x3:=t2

We use
(x2 x3 x4 x5  |
 1 -2 -4 2  |  5)
to compute x2
x2 = 1 + 2 t2

We use
(x1 x2 x3 x4 x5  |
 1 -5  8  2  -2  |  -9)
to compute x1
x1 = -2 + t1+2 t2

Out[114]= {{-2,1,0,-1,0},{{1,0,0,1/2,1},{2,2,1,0,0}}}

In[115]:= LinearSolve [A1,b1]
Out[115]= {-2,1,0,-1,0}

In[116]:= NullSpace [A1]
Out[116]= {{2,0,0,1,2},{2,2,1,0,0}}


In[119]:= (* Beispiel 2 *)

In[120]:= A2 = {{1,2,-3},{2,3,-11},{3,5,-14}}
Out[120]= {{1,2,-3},{2,3,-11},{3,5,-14}}

In[121]:= b2 = {3,5,-14}
Out[121]= {3,5,-14}

In[122]:= Gauss [A2,b2]

(x1 x2 x3 |
 1 2  -3   |  3
 2 3 -11   |  5
 3 5 -14   |-14)

We use equation (1) of the last system to eliminate x1
(x2 x3 |
```

```
-1 -5  |  -1
-1 -5  |  -23)

We use equation (1) of the last system to eliminate x2
(x3  |
  0  |  -22)

x3 does not appear in any equation.
The system has no solution
Out[122]= NOSOLUTION

In[123]:= LinearSolve [A2, b2]
LinearSolve::nosol: Linear equation encountered that has no solution.
Out[123]= LinearSolve[{{1,2,-3},{2,3,-11},{3,5,-14}},{3,5,-14}]

In[124]:= NullSpace [A2]
Out[124]= {{13,-5,1}}


In[125]:= (* Beispiel 3 *)
In[126]:= A3 = {{1,2,-3},{2,3,-11},{3,5,-14}}
b3 = {4,6,10}
Out[126]= {{1,2,-3},{2,3,-11},{3,5,-14}}
Out[127]= {4,6,10}

In[128]:= Gauss [A3, b3]

(x1 x2 x3  |
1 2  -3  |  4
2 3 -11  |  6
3 5 -14  |  10)

We use equation (1) of the last system to eliminate x1
(x2 x3  |
-1 -5  |  -2
-1 -5  |  -2)

We use equation (1) of the last system to eliminate x2
(x3  |
  0  |  0)

x3 does not appear in any equation.
x3:=t1

We use
(x2 x3  |
 -1 -5  |  -2)
to compute x2
x2 = 2 + -5 t1

We use
```

```
 (x1 x2 x3  |
  1 2 -3    | 4)
to compute x1
x1 = 0 + 13 t1

Out[128]= {{0,2,0},{{13,-5,1}}}

In[129]:= LinearSolve [A3, b3]
Out[129]= {0,2,0}

In[130]:= NullSpace [A3]
Out[130]= {{13,-5,1}}
```

Teil 2

Logik und Mengenlehre

KAPITEL 4

Aussagenlogik

1. Aussagen

Wir bezeichnen als eine *Aussage* eine „Behauptung“, von der man sinnvollerweise fragen kann, ob sie wahr oder falsch ist. Beispiele:

(1) $5 > 2$.
Das ist eine Aussage, und sie ist wahr.

(2) $5 < 2$.
Das ist eine Aussage, und sie ist falsch.

(3) $5 + 2$.
Es hat keinen Sinn, zu fragen, ob $5 + 2$ wahr oder falsch ist. $5 + 2$ ist daher keine Aussage, sondern ein Ausdruck oder *Term*.

(4) Es gibt eine natürliche gerade Zahl, deren Quadrat ungerade ist.
Das ist eine Aussage, und sie ist falsch.

(5) Es gibt ein $n \in \mathbb{N}$, sodass n gerade und n^2 ungerade ist.
Das ist eine Aussage, und sie ist falsch. (Wir bezeichnen mit $\mathbb{N} := \{1, 2, 3, \ldots\}$ die natürlichen Zahlen, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.)

(6) Für alle $n \in \mathbb{N}$ gilt: wenn n gerade ist, so ist n^2 gerade.
Das ist eine Aussage, und sie ist wahr.

(7) n ist gerade oder durch 3 teilbar.
Das ist keine Aussage, weil eine Variable, n, vorkommt. Wir bezeichnen eine solche Behauptung, die noch von Variablen abhängt, als *Aussageform*. Obige Aussageform stimmt für manche n, zum Beispiel für $n = 8$ und $n = 9$, und für andere nicht, zum Beispiel für $n = 7$.

(8) n ist gerade oder $n + 1$ ist gerade.
Das stimmt zwar für alle natürlichen Zahlen, ist aber keine Aussage, da n vorkommt.

(9) Für alle $n \in \mathbb{N}$ ist n gerade oder $n + 1$ gerade.
Das ist eine Aussage, und sie ist wahr. Hier kommt zwar n noch vor, aber es ist durch „für alle“ gebunden.

(10) Jede gerade natürliche Zahl n mit $n \geq 4$ lässt sich als Summe zweier (nicht notwendigerweise voneinander verschiedener) Primzahlen schreiben.
Das ist eine Aussage. Man weiß nicht, ob sie wahr ist. Der Mathematiker Christian Goldbach hat 1742 in einem Brief an Leonhard Euler (1707-1783) vermutet, dass diese Aussage wahr ist. Euler hielt sie für „ein ganz gewisses theorema, ungeachtet ich dasselbe nicht demonstriren kann“. Diese Aussage heißt *Goldbachsche Vermutung*.

(11) Jede natürliche gerade Zahl ist Summe von höchstens 6 (nicht notwendigerweise voneinander verschiedenen) Primzahlen.
Das ist eine Aussage. Seit 1995 weiß man, dass sie wahr ist.

Wir werden nun genauer untersuchen, wie Aussagen aufgebaut sein können.

2. Die Junktoren „und“, „oder“ und „nicht“

Atomare Aussagen sind Aussagen folgender Form: $5 > 2$, $3 = 4$, $2 + 5 > 6$.

Diese Aussagen können nun mit logischen Junktoren zu neuen Aussagen verbunden werden.

2.1. Logisches „und“.

- ▷ $5 > 2$ und $3 = 4$.
  Diese Aussage ist falsch.
- ▷ $5 > 2$ und $3 < 4$.
  Diese Aussage ist wahr.
- ▷ $5 < 2$ und $3 < 4$.
  Diese Aussage ist falsch.
- ▷ $5 < 2$ und $3 = 4$.
  Diese Aussage ist falsch,

Definition 4.1. Wenn A und B Aussagen sind, dann betrachten (definieren) wir die neue Aussage

$$A \text{ und } B$$

dann als wahr, wenn A und B beide wahr sind.

Diese Festlegung präzisiert die Bedeutung von „und“ in der Mathematik, schränkt aber gleichzeitig ein. Wenn wir im Alltag sagen: „Sie ging zum Arzt und wurde krank“, oder wenn wir sagen: „Sie wurde krank und ging zum Arzt“, so schwingt im ersten Satz, neben einer zeitlichen Abfolge, auch „der Arztbesuch war Schuld“ mit, im zweiten „weil sie krank war, ging sie zum Arzt“. Wenn wir mathematische Zusammenhänge in der Sprache der Prädikatenlogik ausdrücken, verzichten wir bewusst auf mitschwingende Nebenbedeutungen.

Für die Aussage „A und B“ schreiben wir auch $A \wedge B$ und bezeichnen sie als die *Konjunktion* von A und B.

Definition 4.2. Wir bezeichnen den *Wahrheitswert* einer Aussage A mit w oder 1, wenn die Aussage wahr ist, und mit f oder 0, wenn die Aussage falsch ist. Wir kürzen den Wahrheitswert von A auch mit $W(A)$ ab.

Wir kürzen die Menge $\{(0,0),(0,1),(1,0),(1,1)\}$ mit $\{0,1\}^2$ ab. Sei $\sqcap$ die Funktion von $\{0,1\}^2$ nach $\{0,1\}$, die durch $0 \sqcap 0 = 0 \sqcap 1 = 1 \sqcap 0 = 0$ und $1 \sqcap 1 = 1$ definiert ist. Dann ist der Wahrheitswert von $A \wedge B$ durch

$$W(A \wedge B) = W(A) \sqcap W(B)$$

gegeben. Wir halten insbesondere fest, dass der Wahrheitswert von $A \wedge B$ nur von den Wahrheitswerten von A und B abhängt.

2.2. Logisches „oder“.

- ▷ $5 > 2$ oder $3 = 4$.
  Diese Aussage ist wahr.
- ▷ $5 > 2$ oder $3 < 4$.
  Diese Aussage ist wahr.
- ▷ $5 < 2$ oder $3 < 4$.
  Diese Aussage ist wahr.
- ▷ $5 < 2$ oder $3 = 4$.
  Diese Aussage ist falsch.

Definition 4.3. Seien A und B Aussagen, und sei C die Aussage

$$A \text{ oder } B.$$

Dann ist C genau dann wahr, wenn zumindest eine der beiden Aussagen A und B wahr ist. Wir schreiben für „A oder B" auch $A \vee B$, und bezeichnen C als die *Disjunktion* von A und B.

Wir normieren hier den Gebrauch des Wortes „oder".

Wenn etwa eine Mutter ihrem Kind verspricht: „Du bekommst ein Eis oder eine Torte", so ist das Versprechen auch dann erfüllt, wenn das Kind Eis und Torte bekommt. Es ist aber vorstellbar, dass die Mutter gemeint hat: „Du bekommst ein Eis oder eine Torte, aber nicht beides". Für den mathematischen Gebrauch normieren wir, dass $A \vee B$ auch dann wahr ist, wenn beide Aussagen A und B wahr sind.

Definition 4.4. Sei $\sqcup$ die Funktion von $\{0,1\}^2$ nach $\{0,1\}$, die durch $0 \sqcup 0 = 0$ und $0 \sqcup 1 = 1 \sqcup 0 = 1 \sqcup 1 = 1$ definiert ist. Dann ist der Wahrheitswert von $A \vee B$ durch

$$W(A \vee B) = W(A) \sqcup W(B)$$

gegeben.

Der Wahrheitswert von $A \vee B$ hängt also nur von den Wahrheitswerten von A und B ab.

2.3. Verneinung.

- ▷ 5 ist nicht größer als 7.
  Da $5 > 7$ falsch ist, ist diese Aussage wahr.
- ▷ Nicht alle Primzahlen sind ungerade.
  Die Aussage „alle Primzahlen sind ungerade" ist falsch, da 2 gerade und eine Primzahl ist. Also ist die Aussage „nicht alle Primzahlen sind ungerade" wahr.
- ▷ 5 ist nicht größer als 3.
  Diese Aussage ist falsch, weil 5 größer als 3 ist.

Definition 4.5. Sei A eine Aussage, und sei C die Aussage

$$\text{nicht } A.$$

Dann ist C dann wahr, wenn A falsch ist, und C ist dann falsch, wenn A wahr ist. Wir schreiben für „nicht A" auch $\neg A$.

Definition 4.6. Sei $\sim$ die Funktion von $\{0,1\}$ nach $\{0,1\}$, die durch $\sim(0) = 1$ und $\sim(1) = 0$ definiert ist. Dann ist der Wahrheitswert von $\neg A$ durch

$$W(\neg A) = \sim(W(A))$$

gegeben.

Die Bedeutung von „nicht" in der Mathematik weicht also nicht vom üblichen Gebrauch ab. Verneinungen können aber manchmal durchaus schwierig zu durchblicken sein. Die Behauptung

„nicht alle natürlichen Zahlen sind nicht ungerade"

lässt sich etwa viel einfacher als

„nicht alle natürlichen Zahlen sind gerade"

oder

„es gibt ungerade natürliche Zahlen"

ausdrücken. Wir werden in Kürze sehen, wie wir solche Vereinfachungen fast mechanisch durchführen können.

Mehrfache Verneinungen sind logische Stolperfallen: in „Emilia Galotti“ von G.E. Lessing warnt eine Mutter ihre Tochter:

> „Wie wild er [der Vater] schon war, als er nur hörte, dass der Prinz dich jüngst nicht ohne Missfallen gesehen!“.

Die Tragödie beruht aber dann darauf, dass der Prinz die Tochter vielmehr ganz ohne Missfallen gesehen hat.

3. Rechengesetze für Junktoren

Definition 4.7. Seien A und B Aussagen. Wir bezeichnen A und B als *äquivalent*, wenn A und B beide wahr, oder beide falsch sind. Wir schreiben dafür $A \equiv B$.

Satz 4.8. *Seien A, B Aussagen. Dann gilt:*

(1) $\neg(\neg A) \equiv A$.
(2) $\neg(A \wedge B) \equiv (\neg A) \vee (\neg B)$ *(De Morgansches Gesetz*[1]*).*
(3) $\neg(A \vee B) \equiv (\neg A) \wedge (\neg B)$ *(De Morgansches Gesetz).*

Beweis: (1) Wir betrachten zuerst den Fall, dass A wahr ist. Dann ist $\neg A$ falsch, also ist $\neg(\neg A)$ wahr. Nun betrachten wir den Fall, dass A falsch ist. Dann ist $\neg A$ wahr, also ist $\neg(\neg A)$ falsch. In jedem Fall haben A und $\neg(\neg A)$ also den gleichen Wahrheitswert.

(2) Wir nehmen zuerst an, dass $\neg(A \wedge B)$ wahr ist. Dann ist $A \wedge B$ falsch. Das bedeutet, dass A und B nicht beide wahr sein können; eine der beiden Aussagen ist also falsch. Wenn A falsch ist, so ist $\neg A$ wahr, und somit erst recht $(\neg A) \vee (\neg B)$. Wenn B falsch ist, so ist $\neg B$ wahr. Auch dann ist $(\neg A) \vee (\neg B)$ wahr.

Nehmen wir nun an, dass $\neg(A \wedge B)$ falsch ist. Dann ist $A \wedge B$ wahr, also sind sind beide Aussagen A und B wahr. Folglich sind beide Aussagen $\neg A$ und $\neg B$ falsch; dann ist auch $(\neg A) \vee (\neg B)$ falsch.

Den Beweis von (3) lassen wir hier aus. □

Übungsaufgaben 4.9.

(1) Zeigen Sie $\neg(A \vee B) \equiv (\neg A) \wedge (\neg B)$. Verwenden Sie dazu ähnliche Formulierungen wie im Beweis von Satz 4.8 (2).

Wir geben nun eine zweite Variante des Beweises von Satz 4.8 (2) an.

Beweis II: Sei a der Wahrheitswert von A, und sei b der Wahrheitswert von B. Dann gilt

$$\begin{aligned} W(\neg(A \wedge B)) &= \sim(W(A \wedge B)) \\ &= \sim(W(A) \sqcap W(B)) \\ &= \sim(a \sqcap b). \end{aligned}$$

Ebenso gilt

$$\begin{aligned} W((\neg A) \vee (\neg B)) &= W(\neg A) \sqcup W(\neg B) \\ &= (\sim(W(A))) \sqcup (\sim(W(B))) \\ &= (\sim a) \sqcup (\sim b). \end{aligned}$$

[1] Augustus De Morgan, 1806-1871, englischer Mathematiker.

Nun bleibt zu zeigen, dass für alle 4 Belegungen von a und b gilt, dass $\sim(a \sqcap b) = (\sim a) \sqcup (\sim b)$. Wir machen dazu folgende Tabelle; eine solche Tabelle heißt auch *Wahrheitstafel*.

a	b	$a \sqcap b$	$\sim(a \sqcap b)$	$\sim a$	$\sim b$	$(\sim a) \sqcup (\sim b)$
0	0	0	1	1	1	1
0	1	0	1	1	0	1
1	0	0	1	0	1	1
1	1	1	0	0	0	0

Da die 4. und die 7. Spalte dieser Tabelle gleich sind, gilt für alle $(a, b) \in \{0, 1\}^2$, dass $\sim(a \sqcap b) = (\sim a) \sqcup (\sim b)$. Somit gilt insgesamt $W(\neg(A \wedge B)) = W(\neg(A) \vee \neg(B))$, und somit sind $\neg(A \wedge B))$ und $(\neg A) \vee (\neg B)$ äquivalent.

Satz 4.10. *Seien A, B, C Aussagen. Dann gilt:*

(1) $(A \wedge B) \vee C \equiv (A \vee C) \wedge (B \vee C)$ *(Distributivgesetz).*
(2) $(A \vee B) \wedge C \equiv (A \wedge C) \vee (B \wedge C)$ *(Distributivgesetz).*
(3) $A \wedge A \equiv A$ *(Idempotenz von $\wedge$).*
(4) $A \vee A \equiv A$ *(Idempotenz von $\vee$).*
(5) $A \wedge (A \vee B) \equiv A$ *(Verschmelzungs- oder Absorptionsgesetz).*
(6) $A \vee (A \wedge B) \equiv A$ *(Verschmelzungs- oder Absorptionsgesetz).*
(7) $A \wedge B \equiv B \wedge A$ *(Kommutativität von $\wedge$).*
(8) $A \vee B \equiv B \vee A$ *(Kommutativität von $\vee$).*
(9) $(A \wedge B) \wedge C \equiv A \wedge (B \wedge C)$ *(Assoziativität von $\wedge$).*
(10) $(A \vee B) \vee C \equiv A \vee (B \vee C)$ *(Assoziativität von $\vee$).*

Beweis: Wir beweisen diese Eigenschaften jeweils mit vollkommen verschiedenen Strategien.

(1) Wir nehmen zuerst an, dass $(A \wedge B) \vee C$ wahr ist und zeigen, dass dann auch $(A \vee C) \wedge (B \vee C)$ wahr ist. Da $(A \wedge B) \vee C$ wahr ist, ist entweder $A \wedge B$ wahr, oder C ist wahr. Wir betrachten zuerst den Fall, dass $A \wedge B$ wahr ist. Dann sind A und B beide wahr, also sind auch $A \vee C$ und $B \vee C$ beide wahr. Somit ist auch $(A \vee C) \wedge (B \vee C)$ wahr. Wir betrachten nun den Fall, dass C wahr ist. Dann sind $A \vee C$ und $B \vee C$ beide wahr, und somit ist $(A \vee C) \wedge (B \vee C)$ wahr.

Nun nehmen wir an, dass $(A \wedge B) \vee C$ falsch ist und zeigen, dass dann auch $(A \vee C) \wedge (B \vee C)$ falsch ist. Wenn $(A \wedge B) \vee C$ falsch ist, dann ist sowohl $(A \wedge B)$ als auch C falsch. Da $A \wedge B$ falsch ist, muss zumindest eine der beiden Aussagen A und B falsch sein. Wir betrachten zuerst den Fall, dass A falsch ist. Da dann A und C beide falsch sind, ist $A \vee C$ falsch; damit ist aber $(A \vee C) \wedge (B \vee C)$ auch falsch. Wir betrachten nun den Fall, dass B falsch ist. Da dann B und C beide falsch sind, ist $B \vee C$ falsch; damit ist aber $(A \vee C) \wedge (B \vee C)$ auch falsch.

(2)

$$\begin{aligned} (A \vee B) \wedge C &\equiv \neg\big(\neg((A \vee B) \wedge C)\big) && \text{wegen Satz 4.8 (1)} \\ &\equiv \neg\big((\neg(A \vee B)) \vee (\neg C)\big) && \text{wegen Satz 4.8 (2)} \\ &\equiv \neg\big((\neg(A) \wedge \neg(B)) \vee (\neg C)\big) && \text{wegen Satz 4.8 (3).} \end{aligned}$$

Nun verwenden wir (1) (für $A' := \neg A$, $B' := \neg B$, $C' := \neg C$) und erhalten, dass der letzte Ausdruck gleich $\neg\big(((\neg A) \vee (\neg C)) \wedge ((\neg B) \vee (\neg C))\big)$ ist. Wir verwenden nun wieder die De Morganschen Gesetze und erhalten

$$\begin{aligned} \neg\big(((\neg A) \vee (\neg C)) \wedge ((\neg B) \vee (\neg C))\big) &\equiv \neg\big((\neg A) \vee (\neg C)\big) \vee \neg\big((\neg B) \vee (\neg C)\big) \\ &\equiv \big((\neg(\neg A)) \wedge (\neg(\neg C))\big) \vee \big((\neg(\neg B)) \wedge (\neg(\neg C))\big) \\ &\equiv (A \wedge C) \vee (B \wedge C). \end{aligned}$$

(3) Wenn der Wahrheitswert von A gleich a ist, so ist der Wahrheitswert von $A \wedge A$ gleich $a \sqcap a$. Da $0 \sqcap 0 = 0$ und $1 \sqcap 1 = 1$, gilt also für alle $a \in \{0, 1\}$, dass $a \sqcap a = a$. Somit sind die Wahrheitswerte von $A \wedge A$ und A gleich.

(5) Wir zeigen folgendes:

(1) Wenn $A \wedge (A \vee B)$ wahr ist, so ist A wahr.
(2) Wenn A wahr ist, so ist $A \wedge (A \vee B)$ wahr.

Wir überlegen uns, warum das ausreicht. Wir sollten ja eigentlich zeigen, dass $A \wedge (A \vee B)$ und A den gleichen Wahrheitswert haben. Wenn sie verschiedenen Wahrheitswert haben, dann ist entweder $A \wedge (A \vee B)$ wahr und A falsch, oder es ist $A \wedge (A \vee B)$ falsch und A wahr. Die erste Alternative wird aber von der Überlegung (1) ausgeschlossen, die zweite von der Überlegung (2).

Wir nehmen also an, dass $A \wedge (A \vee B)$ wahr ist. Dann sind A und $A \vee B$ beide wahr. Insbesondere ist also A wahr.

Nun nehmen wir an, dass A wahr ist. Dann ist $A \vee B$ (erst recht) wahr, also sind A und $A \vee B$ beide wahr. Folglich ist $A \wedge (A \vee B)$ war.

Übungsaufgaben 4.11.

(1) Wir haben im Beweis von Satz 4.10 die Eigenschaft (2) aus (1) und den De Morganschen Gesetzen hergeleitet. Geben Sie einen Beweis von (2), der so aufgebaut ist, wie der Beweis von (1); starten Sie also damit, dass Sie annehmen, dass $(A \vee B) \wedge C$ wahr ist, ...
(2) Geben Sie einen Beweis von Satz 4.10 (1) durch Wahrheitstafeln.
(3) Finden Sie zwei Ausdrücke p, q der Form $A \wedge B$, $(\neg A) \vee B$, ..., sodass folgendes gilt: wenn p wahr ist, ist auch q wahr, aber wenn q wahr ist, muss deshalb p nicht notwendigerweise wahr sein.
(4) Zeigen Sie Satz 4.10 (6), indem Sie die Funktionen $f(a, b) = a \sqcap (a \sqcup b) = a$ und $g(a, b) = a$ tabellieren.
(5) Zeigen Sie Satz 4.10 (6), indem Sie so vorgehen wie im angegebenen Beweis von Satz 4.10 (5).
(6) Zeigen Sie, dass für alle Aussagen A, B gilt: $A \vee B \equiv B \vee A$ und $A \wedge B = B \wedge A$ (*Kommutativgesetz*).
(7) Zeigen Sie, dass für alle Aussagen A, B, C gilt: $(A \vee B) \vee C \equiv A \vee (B \vee C)$ und $(A \wedge B) \wedge C \equiv A \wedge (B \wedge C)$ (*Assoziativgesetz*).
(8) Seien A, B, C Aussagen. Zeigen Sie, dass $(A \vee B) \wedge C$ und $(A \wedge C) \vee (B \wedge C)$ äquivalent sind, indem Sie die Funktionen l und r tabellieren, die durch $l(a, b, c) := (a \sqcup b) \sqcap c$ und $r(a, b, c) := (a \sqcap c) \sqcup (b \sqcap c)$ für $a, b, c \in \{0, 1\}$ definiert sind.
(9) Seien A, B, C Aussagen. Zeigen Sie, dass $(A \vee B) \wedge C$ und $(A \wedge C) \vee (B \wedge C)$ äquivalent sind, und geben Sie einen Beweis in der Form des Beweises von Satz 4.10(1) oder Satz 4.10(5) aus dem Skriptum an. Nehmen Sie also an, dass $(A \vee B) \wedge C$ wahr ist, und zeigen Sie, dass dann auch $(A \wedge C) \vee (B \wedge C)$ wahr ist. Dann nehmen Sie an, dass $(A \wedge C) \vee (B \wedge C)$ wahr ist, und zeigen Sie, dass $(A \vee B) \wedge C$ wahr ist.

Wir kürzen nun die wahre Aussage $0 = 0$ mit $\mathbf{T}$, und die falsche Aussage $0 \neq 0$ mit $\mathbf{F}$ ab.

Satz 4.12. *Sei A eine Aussage. Dann gilt:*

(1) $A \wedge \mathbf{T} \equiv A$,
(2) $A \vee \mathbf{T} \equiv \mathbf{T}$,
(3) $A \wedge \mathbf{F} \equiv \mathbf{F}$,
(4) $A \vee \mathbf{F} \equiv A$.

Beweis: (1) Wir nehmen an, dass $A \wedge \mathbf{T}$ wahr ist. Dann ist A wahr und $\mathbf{T}$ wahr. Somit ist A wahr. Nun nehmen wir an, dass A wahr ist. Da dann A und $\mathbf{T}$ beide wahr sind, ist auch $A \wedge \mathbf{T}$ wahr.

(2) Sei a der Wahrheitswert von A. Dann ist der Wahrheitswert von $A \vee \mathbf{T}$ gleich $a \sqcup 1$. Nun sehen wir, dass $0 \sqcup 1 = 1 \sqcup 1 = 1$, also ist der Wahrheitswert von $a \sqcup 1$ gleich 1. Der Wahrheitswert

von $\mathbf{T}$ ist ebenfalls 1. Somit haben $A \vee \mathbf{T}$ und $\mathbf{T}$ den gleichen Wahrheitswert und sind daher äquivalent.

(3) Wir verwenden De Morgan und erhalten $A \wedge \mathbf{F} \equiv \neg(\neg(A \wedge \mathbf{F})) \equiv \neg((\neg A) \vee (\neg \mathbf{F})) \equiv \neg((\neg A) \vee \mathbf{T}) \equiv \neg \mathbf{T} = \mathbf{F}$. (Warum gilt jedes der vier $\equiv$?)

(4) wird ausgelassen. $\square$

Übungsaufgaben 4.13.

(1) Geben Sie einen Beweis für Satz 4.12 (4) in jener Form an, in der wir Satz 4.12 (1) bewiesen haben.
(2) Geben Sie einen Beweis für Satz 4.12 (4) in jener Form an, in der wir Satz 4.12 (2) bewiesen haben.
(3) Geben Sie einen Beweis für Satz 4.12 (4), indem Sie die De Morganschen Gesetze und $\neg(\neg X) \equiv X$ ausnützen und damit (4) auf eine der bereits bewiesenen Äquivalenzen zurückführen.

4. Die Implikation

In diesem Abschnitt normieren wir den Gebrauch von „wenn …, dann“. Betrachten wir dazu folgendes Beispiel (cf. [**BT09**, S. 49]). Anton sagt zu Berta:

„Wenn du mir das Buch morgen zurückbringst, zahle ich dir einen Kaffee“.

In welchen der möglichen vier Fälle hat Anton seine Zusage gehalten?

(1) Berta bringt das Buch zurück, und Anton zahlt ihr einen Kaffee: Zusage gehalten.
(2) Berta bringt das Buch zurück, und Anton zahlt ihr keinen Kaffee: Zusage nicht gehalten.
(3) Berta bringt das Buch nicht zurück, und Anton zahlt ihr keinen Kaffee: Auch hier kann sich Berta nicht beschweren: für diesen Fall hatte ihr Anton nichts versprochen. Zusage gehalten.
(4) Berta bringt das Buch nicht zurück, und Anton zahlt ihr einen Kaffee: Das ist sehr nett von Anton, und verletzt die Abmachung sicher nicht: darüber, was Anton tut, wenn Berta das Buch nicht zurückbringt, hat er nichts zugesagt. Insgesamt: Zusage gehalten.

Die einzige Möglichkeit, dass eine Aussage „wenn A, dann B“ falsch ist, ist also die, dass die Prämisse A eintritt, die Konklusion B aber nicht.

Definition 4.14. Seien A und B Aussagen, und sei C die Aussage

wenn A, dann B.

Dann ist C genau dann falsch, wenn A wahr und B falsch ist. Sonst ist C wahr. Wir kürzen die Aussage „wenn A, dann B“ auch mit $A \Rightarrow B$ ab und bezeichnen C als eine *Implikation.*

Sei $\rightarrow$ die Funktion von $\{0,1\}^2$ nach $\{0,1\}$, die durch folgende Tabelle erklärt ist:

a	b	$a \rightarrow b$
0	0	1
0	1	1
1	0	0
1	1	1

Satz 4.15. *Seien A und B Aussagen, sei a der Wahrheitswert von A, und sei b der Wahrheitswert von B. Dann ist der Wahrheitswert von $A \Rightarrow B$ gleich $a \rightarrow b$.*

Beweis: Wir betrachten als erstes den Fall, dass $A \Rightarrow B$ falsch ist. Das bedeutet, dass A wahr und B falsch ist. Dann ist $a = 1$ und $b = 0$. Wegen $1 \to 0 = 0$ gilt dann $a \to b = 0$.

Nun betrachten wir den Fall, dass $A \Rightarrow B$ wahr ist. Nehmen wir im Widerspruch zur Behauptung an, dass $a \to b = 0$. Dann gilt $a = 1$ und $b = 0$ und somit ist A wahr und B falsch. Dann ist ist $A \Rightarrow B$ falsch, im Widerspruch zur Fallannahme. Somit kann $a \to b = 0$ nicht gelten. Also gilt $a \to b = 1$. □

Ähnlich wie bei „oder" gibt die Normierung nur eine der normalsprachlichen Bedeutungen von „wenn, dann" wieder. Wir betrachten folgende Beispiele:

(1) $(5 > 2) \;\Rightarrow\; (10 > 4)$.
$(5 > 2)$ und $(10 > 4)$ sind beide wahr. Also ist $(5 > 2) \;\Rightarrow\; (10 > 4)$ wahr.
Im Satz „wenn 5 größer als 2 ist, dann ist auch 10 größer als 4" schwingt aber auch mit "klar, denn ich brauche $5 > 2$ nur zu verdoppeln. Der Wahrheitswert von $A \Rightarrow B$ sagt aber nichts über einen kausalen Zusammenhang von A und B aus, sondern nur, dass es nicht so ist, dass A wahr und B falsch ist.
(2) $(5 > 2) \;\Rightarrow$ (6 ist eine gerade Zahl).
Beide Teilaussagen sind wahr, also ist die Implikation wahr.
Dass die beiden Aussagen inhaltlich keine Verbindung haben, stört nicht. Der Wahrheitswert der Implikation hängt nur von den Wahrheitswerten der Teilaussagen ab.
(3) $(5 < 4) \;\Rightarrow\; (5 < 11)$.
Hier ist $A = (5 < 4)$ falsch und $B = (5 < 11)$ wahr. Ingesamt ist $A \Rightarrow B$ also wahr.
(4) $(3 < 2) \;\Rightarrow\; (103 < 102)$.
Da $3 < 2$ und $103 < 102$ beide falsch sind, ist die Implikation wahr.
(5) $(5 > 3) \;\Rightarrow\; (-5 > -3)$.
Da $5 > 3$ wahr ist, und $-5 > -3$ falsch, ist die Implikation falsch.
(6) Wenn Paris in Ungarn liegt, so ist Schnee schwarz.
Die Implikation ist wahr.

In der mathematischen „Umgangssprache" kommt es manchmal vor, dass „wenn, dann" auch ein heimliches „für alle" enthält. Wir betrachten die Aussage:

Wenn n gerade ist, so ist n^2 durch 4 teilbar.

Wir können das so sehen: $G(n)$ ist die *Aussageform* „n ist gerade", $V(n)$ ist die Aussageform „n ist Vielfaches von 4". Dann meint obige Behauptung vermutlich:

Für alle n gilt: wenn $G(n)$, dann $V(n^2)$,

was man auch als

für alle $n \in \mathbb{N} : (G(n) \Rightarrow V(n^2))$

schreiben kann. Diese Aussage ist wahr: wenn n gerade ist, so gibt es ein $k \in \mathbb{N}$ mit $n = 2k$. Also gilt $n^2 = 4k^2$, und das ist ein Vielfaches von 4. Wir betrachten nun die Aussage:

Wenn n gerade ist, so ist n durch 3 teilbar.

Wenn wir $t \mid n$ für „t teilt n" schreiben, so wäre

$$2 \mid n \Rightarrow 3 \mid n.$$

eine Formalisierung. Die Aussageform $I(n) := (2 \mid n \Rightarrow 3 \mid n)$ ist für manche n wahr (etwa für $n = 5$, $n = 6$, $n = 3$), für andere n falsch (etwa $n = 2$). Es könnte aber auch sein, dass

$$\text{für alle } n \in \mathbb{N} : (2 \mid n \Rightarrow 3 \mid n)$$

gemeint ist. Das könnte man so ausdrücken:

> Für alle natürlichen Zahlen $n \in \mathbb{N}$ gilt: wenn n gerade ist, so ist n durch 3 teilbar.

Diese Aussage ist falsch, da die Aussageform $I(n)$ für $n = 2$ nicht gilt.

Wir werden die Formulierung

> Für alle $n \in \mathbb{N}$ gilt: wenn n gerade ist, so ist n^2 durch 4 teilbar

gegenüber

> Wenn n gerade ist, so ist n^2 durch 4 teilbar.

bevorzugen.

ÜBUNGSAUFGABEN 4.16.

(1) Schreiben Sie folgende Sätze so um, dass sie die Form „für alle $x \ldots \in \ldots$ gilt : $A(x) \Rightarrow B(x)$“ haben.
   (a) Wenn x eine reelle Zahl ist, so ist $x^2 \geq 0$.
   (b) Wenn n kein Vielfaches von 3 ist, so hat n^2 bei Division durch 3 Rest 1. *Hinweis:* $B(x)$ ist dann „x hat bei Division durch 3 Rest 1“.
(2) Schreiben Sie folgende Sätze so um, dass sie die Form „für alle $x \ldots \in \ldots$ gilt : $A(x) \Rightarrow B(x)$“ haben.
   (a) Die Wurzel einer natürlichen Zahl ist eine natürliche Zahl oder irrational. (Für eine reelle Zahl x sei $A(x)$ die Eigenschaft, dass x Wurzel einer natürlichen Zahl ist, $N(x)$ die Eigenschaft, dass x natürlich und $R(x)$ die Eigenschaft, dass x rational ist.)
   (b) Ein Quadrat ist auch ein Rechteck.

Die meisten mathematischen Zusammenhänge lassen sich erst darstellen, wenn man *Aussageformen* (also „Aussagen über Variablen“) und die *Quantoren* „für alle ... gilt:“ und „es gibt ein ..., sodass“ verwendet. Die *Prädikatenlogik* ist jenes Teilgebiet der Mathematik, dass sich mit Ausdrücken, die aus Aussagen, Aussageformen, Junktoren und Quantoren gebildet werden, beschäftigt. Die *Aussagenlogik* beschäftigt sich den Ausdrücken, die aus Aussagen und Junktoren gebildet werden.

Die Implikation lässt sich auch durch Konjunktion, Disjunktion und Negation ausdrücken.

SATZ 4.17. *Seien A, B Aussagen. Dann gilt:*

(1) $A \Rightarrow B \equiv (\neg A) \vee B$.
(2) $A \Rightarrow B \equiv \neg(A \wedge (\neg B))$.
(3) $A \Rightarrow B \equiv ((\neg B) \Rightarrow (\neg A))$ *(*Kontrapositionsregel*).*

Beweis: Wahrheitstafeln.

Wir halten nun noch einige sprachliche Möglichkeiten, die Tatsache, dass $A \Rightarrow B$ wahr ist, auszudrücken, fest:

(1) A impliziert B.
(2) Wenn A, dann B.
(3) A gilt nur dann, wenn B gilt.
(4) B gilt, wenn A gilt.
(5) Wenn B nicht gilt, so gilt auch A nicht.

ÜBUNGSAUFGABEN 4.18.

(1) Sei p die Aussage
$$\big((A \vee B) \wedge ((\neg A) \vee C)\big) \Rightarrow (B \vee C).$$

(a) Finden Sie einen zu p äquivalenten Ausdruck, der nur die Junktoren $\wedge, \vee, \neg$ verwendet.
(b) Zeigen Sie, dass p eine Tautologie ist, dass also p für alle Aussagen A, B, C wahr ist.

5. Weitere Junktoren

DEFINITION 4.19. Seien A und B Aussagen. Dann definieren wir:

(1) $A \Leftarrow B$ steht für $B \Rightarrow A$.
(2) $A \Leftrightarrow B$ ist genau dann wahr, wenn $(A \Rightarrow B) \wedge (B \Rightarrow A)$ wahr ist.

Wir bezeichnen $A \Leftrightarrow B$ als die *Äquivalenz von A und B*.

SATZ 4.20. *Seien A und B Aussagen. Dann ist $A \Leftrightarrow B$ genau dann wahr, wenn A und B beide wahr oder beide falsch sind.*

Dass die Äquivalenz $A \Leftrightarrow B$ wahr ist, kann auch durch folgende Formulierungen ausgedrückt werden.

(1) A genau dann, wenn B.
(2) A gilt dann, und nur dann, wenn B gilt.
(3) (englisch) A if and only if B.
(4) (englisch + P. Halmos[2]) A iff B.
(5) A und B sind äquivalent.

ÜBUNGSAUFGABEN 4.21. Überprüfen Sie jeweils, ob die die Aussagen p und q für alle Aussagen A und B äquivalent sind. Geben Sie dafür (im Fall der Äquivalenz) einen Beweis an, und finden Sie im Fall, dass die Aussagen nicht äquivalent sind, Belegungen für die Wahrheitswerte von A und B, sodass eine Seite wahr und die andere falsch ist.

(1) $p = \neg(A \Rightarrow B)$, $q = A \wedge (\neg B)$.
(2) $p = (A \Rightarrow B) \Rightarrow C$, $q = A \Rightarrow (B \Rightarrow C)$.
(3) $p = A \Leftrightarrow B$, $q = (A \vee (\neg B)) \wedge ((\neg A) \vee B)$.
(4) $p = A \Rightarrow (B \Rightarrow C)$, $q = (A \wedge B) \Rightarrow C$.
(5) $p = A \Rightarrow (B \Rightarrow C)$, $q = B \Rightarrow (A \Rightarrow C)$.
(6) $p = A \Rightarrow (B \Rightarrow B)$, $q = B \Rightarrow (A \Rightarrow A)$.
(7) $p = (A \Rightarrow B) \Rightarrow A$, $q = A$.

DEFINITION 4.22. Seien A und B Aussagen. Dann ist $A \dot{\vee} B$ definiert als $(A \wedge (\neg B)) \vee ((\neg A) \wedge B)$.

$A \dot{\vee} B$ ist das „ausschließende oder“: eines von beiden, und nicht beide.

SATZ 4.23. *Seien A, B Aussagen. Dann gilt $A \dot{\vee} B \equiv \neg(A \Leftrightarrow B) \equiv (A \vee B) \wedge (\neg(A \wedge B))$.*

Die dazu gehörige Boole'sche Funktion bezeichnet man mit $\oplus$. Sie ist durch $0 \oplus 0 = 1 \oplus 1 = 0$, $0 \oplus 1 = 1 \oplus 0 = 1$ definiert.

DEFINITION 4.24. Die Sheffer-Funktion[3] ist definiert durch

$$x|y := \sim(x \sqcap y).$$

Die Shefferfunktion verdankt ihre Popularität der Tatsache, dass sie imstande ist, alle anderen logischen Junktoren auszudrücken: $\sim x = x|x$, $x \sqcap y = (x|y)|(x|y)$, $x \sqcup y = (x|x)|(y|y)$, $x \oplus y = (x|(x|y))|(y|(x|y))$ für alle $x, y \in \{0, 1\}$. Das kann, anders als die Shefferfunktion, keine der Funktionen $\sqcup, \sqcap, \sim, \oplus$ allein. Allerdings reichen zum Beispiel auch $\sqcap$ und $\sim$ gemeinsam aus, um alle anderen Junktoren auszudrücken.

[2] Paul Halmos, 1916-2006
[3] Henry Sheffer, 1882-1964

Übungsaufgaben 4.25.

(1) Finden Sie für jede der 16 möglichen Funktionen von $\{0,1\}^2$ nach $\{0,1\}$ einen möglichst einfachen Ausdruck, der 0, 1, $\sqcup, \sqcap, \oplus, \sim$ verwendet und die entsprechende Funktion beschreibt.

a	b	?	$a \sqcap b$	?	?	?	?	?	?	?	?	?	?	?	?	?	?
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

(2) Finden Sie jeweils einen Ausdruck, der nur $0, x, y$ und den Junktor $\to$ verwendet, und der folgende Funktionen beschreibt:
   (a) $\sim x$.
   (b) $x \sqcup y$.
   (c) $x \sqcap y$.

6. Logisches Schließen in der Aussagenlogik

In diesem Abschnitt ermitteln wir, wann sich aus bestimmten Behauptungen zwingend eine weitere Behauptung ergibt. Betrachten wir folgendes Beispiel:

Beispiel 4.26. Wir nehmen an, dass folgende Annahmen gelten:

(1) Wenn schönes Wetter ist, dann ist Oma guter Laune.
(2) Entweder es ist schlechtes Wetter, oder Klaus ist glücklich.
(3) Wenn Oma guter Laune und Klaus glücklich ist, dann spielen sie Fußball.
(4) Es ist schönes Wetter.

Spielen Oma und Klaus sicher Fußball?

Wir kürzen die Aussagen ab:

A ... Es ist schönes Wetter.
B ... Oma ist guter Laune.
C ... Klaus ist glücklich.
D ... Oma und Klaus spielen Fußball.

Dann können wir die Voraussetzungen so abkürzen:

(1) $A \Rightarrow B$,
(2) $(\neg A) \vee C$,
(3) $(B \wedge C) \Rightarrow D$,
(4) A.

In (2) könnte auch $(\neg A)\dot{\vee}C$ gemeint sein. Wir gehen aber jetzt davon aus, dass $(\neg A) \vee C$ gemeint ist. Die Frage ist nun, ob für alle Wahrheitswertkombinationen, die A, B, C, D annehmen können, und für die die Voraussetzungen (1)-(4) wahr sind, auch D wahr sein muss.

Die Antwort ist in diesem Fall „Ja“: Wegen (4) und (1) ist B wahr. Wegen (4) und (2) ist C wahr. Also gilt wegen (3) auch D.

Wir betrachten nun folgendes Beispiel:

Beispiel 4.27. Wir nehmen an, dass folgende Annahmen gelten:

(1) Wenn schönes Wetter ist, dann ist Klaus glücklich oder er spielt mit Oma Fußball.
(2) Wenn Oma schlechter Laune ist, dann spielen Oma und Klaus nicht Fußball.
(3) Wenn schlechtes Wetter ist, ist Oma guter Laune.

Ist Oma dann sicher guter Laune?

Die Aussagen sind dann

(1) $A \Rightarrow (C \vee D)$,
(2) $(\neg B) \Rightarrow (\neg D)$,
(3) $(\neg A) \Rightarrow B$.

Die Frage ist, ob dann sicher B gilt. Wir sehen (etwa durch Probieren aller 16 Möglichkeiten), dass für $W(A) = w$, $W(B) = f$, $W(C) = w$, $W(D) = f$ alle drei Voraussetzungen erfüllt sind. Da dann B nicht gilt, ist die Antwort „Nein“.

Wir überlegen uns nun, wie man algorithmisch vorgehen kann, um solche Beispiele zu lösen. Wir bezeichnen dazu Terme wie $a \Rightarrow (c \vee d)$ als *Boolesche Ausdrücke*. In solche Ausdrücke können wir für a, b, c, d jeweils T oder F (für „true“ und „false“) einsetzen, und erhalten als Ergebnis T oder F. So ergibt zum Beispiel $a \Rightarrow (c \vee d)$ an der Stelle $(a, b, c, d) := (T, T, F, F)$ den Wert T und an der Stelle $(a, b, c, d) := (F, F, F, F)$ den Wert T. Eine Menge Φ von Ausdrücken heißt *erfüllbar*, wenn es eine Belegung der Variablen $a, b, c, \ldots$ gibt, sodass alle Ausdrücke in Φ wahr werden; so ist zum Beispiel $\{a, \neg a \vee b\}$ erfüllbar und $\{a, \neg a \vee b, \neg b\}$ unerfüllbar (mit $\neg a \vee b$ meinen wir $(\neg a) \vee b$). Ein Ausdruck φ ist *allgemeingültig*, wenn φ für alle Belegungen der Variablen wahr wird. Wir sagen, dass der Ausdruck φ eine *Konsequenz* aus der Menge Φ ist, wenn φ für jede Belegung, für die alle Ausdrücke in Φ wahr sind, ebenfalls wahr ist. Offensichtlich ist φ genau dann eine Konsequenz von Φ, wenn $\Phi \cup \{\neg\varphi\}$ unerfüllbar ist. Im Beispiel (4.26) ist also die Frage, ob

$$\{a \Rightarrow b, \neg a \vee c, (b \wedge c) \Rightarrow d, a, \neg d\} \tag{4.1}$$

erfüllbar ist. Das kann man durch Testen aller 16 Belegungen überprüfen.

Manchmal ist man mithilfe des *Resolutionsverfahrens* schneller. Wir schreiben x^1 für x und x^{-1} für $\neg x$. Eine *Klausel* ist ein Boolescher Ausdruck φ der Form $x_{i_1}^{e_1} \vee \ldots \vee x_{i_k}^{e_k}$; für $k = 0$ bezeichnen wir auch den konstanten Ausdruck F als Klausel. Die Menge $\{x_{i_1}^{e_1}, \ldots, x_{i_k}^{e_k}\}$ ist die Menge der *Literale* von φ; φ ist eine *Teilklausel* von ψ, wenn die Literalmenge von φ eine Teilmenge der Literalmenge von ψ ist. So ist zum Beispiel $a \vee \neg c$ eine Teilklausel von $a \vee \neg b \vee \neg c$, aber nicht von $a \vee c$. F ist Teilklausel jeder Klausel. Die Menge (4.1) ist zur Menge

$$\{\neg a \vee b, \neg a \vee c, \neg b \vee \neg c \vee d, a, \neg d\}$$

äquivalent. Die Erfüllbarkeit einer Menge von Klauseln lässt sich mit folgendem Satz bestimmen.

Satz 4.28. *Sei Φ eine Menge von Klauseln der Form*

$$\begin{aligned}\Phi \quad = \quad & \{x_1 \vee \varphi_1(x_2, \ldots, x_n), \ldots, x_1 \vee \varphi_k(x_2, \ldots, x_n), \\ & \neg x_1 \vee \psi_1(x_2, \ldots, x_n), \ldots, \neg x_1 \vee \psi_l(x_2, \ldots, x_n), \\ & \rho_1(x_2, \ldots, x_n), \ldots, \rho_m(x_2, \ldots, x_n)\}.\end{aligned}$$

Wir nehmen an, dass für alle $i \in \{1, \ldots, k\}$ und $j \in \{1, \ldots, l\}$ die Resolvente

$$\sigma_{i,j} := \varphi_i(x_2, \ldots, x_n) \vee \psi_j(x_2, \ldots, x_n)$$

bezüglich x_1 eine Variable x_r und ihre Negation $\neg x_r$ enthält, oder dass es ein $t \in \{1, \ldots, m\}$ gibt, sodass $\rho_t(x_2, \ldots, x_n)$ eine Teilklausel von $\sigma_{i,j}$ ist.

Dann ist Φ genau dann erfüllbar, wenn $R = \{\rho_1(x_2, \ldots, x_n), \ldots, \rho_m(x_2, \ldots, x_n)\}$ erfüllbar ist.

Die Resolvente $\sigma := \varphi_i \vee \psi_j$ von $x_1 \vee \varphi_i$ und $\neg x_1 \vee \psi_j$ ist stets eine Konsequenz von $\{x_1 \vee \varphi_i, \neg x_1 \vee \psi_j\}$. Wir betrachten noch folgende Sonderfälle beim Bilden der Resolvente: die Resolvente von x_1 und $\neg x_1$ ist F, die Resolvente von $x_1 \vee \varphi$ und $\neg x_1$ ist φ, und die Resolvente von x_1 und $\neg x_1 \vee \psi$ ist ψ. Beim Bilden der Resolvente von $a \vee b$ und $\neg a \vee b \vee \neg c$ vereinfachen wir $b \vee b \vee \neg c$

zu $b \vee \neg c$. Die Resolvente von $a \vee b$ und $\neg a \vee \neg b \vee c$ ist $b \vee \neg b \vee c$. Diese Aussage ist stets erfüllt und kann daher weggelassen werden.

Beweis. Wir nehmen an, dass R erfüllbar ist, und zeigen, dass dann auch Φ erfüllbar ist. Sei $(a_2, \dots, a_n) \in \{T, F\}^{n-1}$ eine Belegung, die R erfüllt.

Wenn $\varphi_i(a_2, \dots, a_n)$ für alle $i \in \{1, \dots, k\}$ wahr ist, so ist $(F, a_2, \dots, a_n)$ eine Belegung, die Φ erfüllt.

Wenn $\psi_j(a_2, \dots, a_n)$ für alle $j \in \{1, \dots, l\}$ wahr ist, so ist $(T, a_2, \dots, a_n)$ eine Belegung, die Φ erfüllt.

Wir nehmen nun an, dass $i \in \{1, \dots, k\}$ und $j \in \{1, \dots, l\}$ so sind, dass $\varphi_i(a_2, \dots, a_n)$ und $\psi_j(a_2, \dots, a_n)$ beide falsch sind. Dann ist auch $\varphi_i(a_2, \dots, a_n) \vee \psi_j(a_2, \dots, a_n)$ falsch. Laut Voraussetzung gibt es ein $t \in \{1, \dots, m\}$, sodass ρ_t eine Teilklausel von $\varphi_i \vee \psi_j$ ist. Dann ist auch $\rho_t(a_2, \dots, a_n)$ falsch. Das steht im Widerspruch dazu, dass $(a_2, \dots, a_n)$ die Klauselmenge R erfüllt. Daher kann dieser Fall nicht eintreten. □

Beispiel 4.29. Wir benutzen diesen Satz, um die Erfüllbarkeit von

$$\Phi = \{a \vee b \vee c, b \vee \neg c, a \vee \neg b, \neg a \vee c, \neg a \vee \neg b \vee \neg c\}$$

zu testen. Da Φ die Voraussetzungen des Satzes noch nicht erfüllt, fügen wir die Resolventen

$$b \vee c, \neg b \vee c, \neg b \vee \neg c$$

zur Menge Φ hinzu und erhalten

$$\Phi_1 = \{a \vee b \vee c, b \vee \neg c, a \vee \neg b, \neg a \vee c, \neg a \vee \neg b \vee \neg c, b \vee c, \neg b \vee c, \neg b \vee \neg c\}.$$

Das ist wegen Satz 4.28 genau dann erfüllbar, wenn

$$\Phi_2 = \{b \vee \neg c, b \vee c, \neg b \vee c, \neg b \vee \neg c\}$$

erfüllbar ist. Die Voraussetzungen von Satz 4.28 sind nicht erfüllt, also ergänzen wir die Resolventen und erhalten

$$\Phi_3 = \{b \vee \neg c, b \vee c, \neg b \vee c, \neg b \vee \neg c, c, \neg c\}.$$

Das erfüllt die Voraussetzungen von Satz 4.28 und ist daher genau dann erfüllbar, wenn

$$\Phi_4 = \{c, \neg c\}$$

erfüllbar ist. Wir ergänzen Φ_4 durch die Resolvente F und erhalten, dass Φ_4 genau erfüllbar ist, wenn $\Phi_6 = \{F\}$ erfüllbar ist. Da Φ_6 unerfüllbar ist, ist Φ nicht erfüllbar.

Beispiel 4.30. Wir untersuchen, ob

$$\Phi = \{a \vee \neg b, \neg a \vee b \vee c, a \vee \neg b, a \vee \neg b \vee c, \neg c\}$$

erfüllbar ist. Jede Resolvente bezüglich a enthält ein Paar $x_i \vee \neg x_i$. Daher ist Φ erfüllbar, wenn $\Phi_1 = \{\neg c\}$ erfüllbar ist. Das ist erfüllbar, zum Beispiel mit $b = F$, $c = F$. Nun ist die Teilklausel $b \vee c$ der Klausel $\neg a \vee b \vee c$ nicht erfüllt, also setzen wir $a := F$ und erhalten die erfüllende Belegung (F, F, F).

Wenn wir Φ_1 mit $b = W$, $c = F$ erfüllen, so ist die Teilklausel $\neg b$ der Klausel $a \vee \neg b$ nicht erfüllt, also setzen wir $a := W$ und erhalten die erfüllende Belegung (W, W, F).

Übungsaufgaben 4.31.

(1) Zeigen Sie: wenn $(A \vee B) \wedge (\neg A \vee C)$ gilt, so gilt auch $B \vee C$.
(2) Finden Sie Aussagen A, B, C, sodass $B \vee C$ wahr, aber $(A \vee B) \wedge (\neg A \vee C)$ falsch ist.
(3) Zeigen Sie, dass $B \vee C$ stets den gleichen Wahrheitswert wie $((A \vee B) \wedge (\neg A \vee C)) \vee ((\neg A \vee B) \wedge (A \vee C))$ hat.
(4) Wir nehmen an, dass folgende Annahmen gelten:

   (a) Wenn schönes Wetter ist und Oma guter Laune ist, dann ist Klaus glücklich.
   (b) Wenn schlechtes Wetter ist, so spielen Oma und Klaus Fußball.
   (c) Wenn Klaus glücklich ist, so spielen Oma und Klaus Fußball.
   (d) Oma und Klaus spielen nicht Fußball.

   Ist Oma dann sicher guter Laune?

(5) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schönes Wetter ist, dann ist Klaus glücklich oder er spielt mit Oma Fußball.
   (b) Wenn Oma schlechter Laune ist, dann spielen Oma und Klaus nicht Fußball.
   (c) Wenn schlechtes Wetter ist, ist Oma guter Laune.

   Kann es sein, dass die Annahmen gelten, Oma schlechter Laune und Klaus unglücklich ist?

(6) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schönes Wetter ist, dann ist Klaus glücklich oder er spielt mit Oma Fußball.
   (b) Wenn Oma schlechter Laune ist, dann spielen Oma und Klaus nicht Fußball.
   (c) Wenn schlechtes Wetter ist, ist Oma guter Laune.

   Kann es sein, dass die Annahmen gelten und Oma schlechter Laune ist?

(7) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schönes Wetter ist, dann ist Klaus glücklich oder er spielt mit Oma Fußball.
   (b) Wenn Oma schlechter Laune ist, dann spielen Oma und Klaus nicht Fußball.
   (c) Wenn schlechtes Wetter ist, ist Oma guter Laune.

   Kann es sein, dass die Annahmen gelten und Oma guter Laune ist?

(8) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schönes Wetter ist, dann ist Klaus glücklich.
   (b) Wenn schlechtes Wetter und Oma guter Laune ist, dann ist Klaus glücklich.
   (c) Wenn schönes Wetter und Oma guter Laune ist, dann ist Klaus unglücklich.
   (d) Wenn schönes Wetter ist, ist Oma guter Laune.
   (e) Wenn schlechtes Wetter ist, ist Oma guter Laune.
   (f) Wenn schlechtes Wetter ist, ist Klaus glücklich.

   Ist Klaus dann sicher glücklich?

(9) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schlechtes Wetter ist und Oma schlechter Laune ist, so ist Klaus glücklich.
   (b) Wenn Klaus glücklich ist, ist Oma guter Laune.
   (c) Wenn schlechtes Wetter ist, ist Oma schlechter Laune.
   (d) Wenn schönes Wetter ist, ist Klaus glücklich.
   (e) Wenn Oma guter Laune und Klaus glücklich ist, dann ist schlechtes Wetter.

   Ist Klaus dann sicher glücklich?

(10) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schlechtes Wetter ist und Oma schlechter Laune ist, so ist Klaus glücklich.
   (b) Wenn Klaus glücklich ist, ist Oma guter Laune.
   (c) Wenn schlechtes Wetter ist, ist Oma schlechter Laune.
   (d) Wenn schönes Wetter ist, ist Klaus glücklich.
   (e) Wenn Oma guter Laune und Klaus glücklich ist, dann ist schlechtes Wetter.

   Ist Klaus dann sicher unglücklich?

(11) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schlechtes Wetter ist, so ist Oma guter Laune.
   (b) Wenn schönes Wetter ist und Oma schlechter Laune ist, dann ist Klaus glücklich.
   (c) Es kommt nie vor, dass schönes Wetter und Oma guter Laune ist.
   (d) Wenn Oma guter Laune ist, dann ist schönes Wetter, oder Klaus ist glücklich.

   Ist Klaus dann sicher glücklich?

(12) Wir nehmen an, dass folgende Annahmen gelten:
   (a) Wenn schlechtes Wetter ist, so ist Oma schlechter Laune.
   (b) Wenn schönes Wetter ist und Oma schlechter Laune ist, dann ist Klaus glücklich.
   (c) Es kommt nie vor, dass schönes Wetter und Oma guter Laune ist.
   (d) Wenn Oma guter Laune ist, dann ist schönes Wetter, oder Klaus ist glücklich.

   Ist Klaus dann sicher glücklich?

(13) Zeigen Sie folgende Variante von Satz 4.28: Sei Φ eine Menge von Booleschen Ausdrücken (nicht notwendigerweise Klauseln) der Form

$$\Phi = \{x_1 \vee \varphi_1(x_2,\dots,x_n),\dots,x_1 \vee \varphi_k(x_2,\dots,x_n),\\ \neg x_1 \vee \psi_1(x_2,\dots,x_n),\dots,\neg x_1 \vee \psi_l(x_2,\dots,x_n),\\ \rho_1(x_2,\dots,x_n),\dots,\rho_m(x_2,\dots,x_n)\}.$$

Wenn

$$R = \{\rho_1(x_2,\dots,x_n),\dots,\rho_m(x_2,\dots,x_n)\} \cup \\ \{\varphi_i(x_2,\dots,x_n) \vee \psi_j(x_2,\dots,x_n) \mid i \in \{1,\dots,k\}, j \in \{1,\dots,l\}\}$$

erfüllbar ist, so ist auch Φ erfüllbar.

(14) Wir schreiben x^1 für x und x^{-1} für $\sim x$. Zeigen Sie, dass es für jede Funktion $f : \{0,1\}^n \to \{0,1\}$ einen Ausdruck p der Form

$$(x_{i_{1,1}}^{e_{1,1}} \sqcap \dots \sqcap x_{i_{1,k_1}}^{e_{1,k_1}}) \sqcup \dots \sqcup (x_{i_{l,1}}^{e_{l,1}} \sqcap \dots \sqcap x_{i_{l,k_l}}^{e_{l,k_l}})$$

gibt, sodass für alle $a_1,\dots,a_n \in \{0,1\}$ die Gleichheit $p(a_1,\dots,a_n) = f(a_1,\dots,a_n)$ gilt. Ein solcher Ausdruck ist ein Ausdruck in *disjunktiver Normalform.* Finden Sie wie in Übungsbeispiel 4.25(1) für jede der 16 möglichen Funktionen einen solchen Ausdruck, der die entsprechende Funktion beschreibt. Wie sieht eine disjunktive Normalform einer unerfüllbaren Aussageform aus? *Hinweis:* Gehen Sie mit Induktion nach n vor und bilden Sie aus einem Ausdruck p_0 für $f(0,x_2,\dots,x_n)$ und einem Ausdruck p_1 für $f(1,x_2,\dots,x_n)$ einen Ausdruck für $f(x_1,\dots,x_n)$, und verwenden Sie das Distributivgesetz.

(15) Wir schreiben x^1 für x und x^{-1} für $\sim x$. Zeigen Sie, dass es für jede Funktion $f : \{0,1\}^n \to \{0,1\}$ einen Ausdruck p der Form

$$(x_{i_{1,1}}^{e_{1,1}} \sqcup \dots \sqcup x_{i_{1,k_1}}^{e_{1,k_1}}) \sqcap \dots \sqcap (x_{i_{l,1}}^{e_{l,1}} \sqcup \dots \sqcup x_{i_{l,k_l}}^{e_{l,k_l}})$$

gibt, sodass für alle $a_1,\dots,a_n \in \{0,1\}$ die Gleichheit $p(a_1,\dots,a_n) = f(a_1,\dots,a_n)$ gilt. Ein solcher Ausdruck ist ein Ausdruck in *konjunktiver Normalform.* Finden Sie wie in Übungsbeispiel 4.25(1) für jede der 16 möglichen Funktionen einen solchen Ausdruck, der die entsprechende Funktion beschreibt. Wie sieht eine konjunktive Normalform einer allgemeingültigen Aussageform aus? *Hinweis:* Gehen Sie mit Induktion nach n vor und bilden Sie aus einem Ausdruck p_0 für $f(0,x_2,\dots,x_n)$ und einem Ausdruck p_1 für $f(1,x_2,\dots,x_n)$ einen Ausdruck für $f(x_1,\dots,x_n)$, und verwenden Sie das Distributivgesetz.

KAPITEL 5

Prädikatenlogik

1. Aussageformen

Kapitel 6 vorwegnehmend verwenden wir in der Folge einige Mengen:

$$\begin{aligned} \mathbb{N} &= \{1, 2, 3, 4, \ldots\} \\ \mathbb{N}_0 &= \{0, 1, 2, 3, 4, \ldots\} \\ \mathbb{R} &= \text{die Menge der reellen Zahlen} \end{aligned}$$

Es gilt $4 \in \mathbb{N}$, $\pi \notin \mathbb{N}$, $\mathbb{N} \notin \mathbb{N}_0$. Als *Aussageform* bezeichnen wir eine Aussage über Variablen, etwa

(1) x ist gerade,
(2) $x \geq 2 + y$,
(3) $3x - 2y = 8$,
(4) x und y haben denselben Rest bei der Division durch 5.

Wenn wir für alle Variablen Werte einsetzen, dann erhalten wir Aussagen, die wahr oder falsch sein können. Wenn etwa $A(x)$ die Aussageform „x ist gerade“ ist, so ist $A(3)$ die Aussage „3 ist gerade“ und $A(4)$ die Aussage „4 ist gerade“. Bei der Definition einer Aussageform ist es sinnvoll, anzugeben, welche Werte x annehmen darf, etwa in folgender Form:

> Für $x \in \mathbb{N}$ definieren wir
> $$A(x) :\Leftrightarrow x \text{ ist gerade.}$$

Damit ist klar, dass $A(\pi)$ nicht definiert wurde.

Wir können aus Aussageformen mithilfe der logischen Junktoren neue Aussageformen bauen. Die Aussageform $A(n)$ über den natürlichen Zahlen, die durch

$$A(n) :\Leftrightarrow n \text{ ist gerade oder } n \text{ ist durch 3 teilbar}$$

gegeben ist, gilt nicht für alle $n \in \mathbb{N}$, da 5 weder gerade noch durch 3 teilbar ist, aber doch für manche, zum Beispiel für $n = 2$ und $n = 18$.

Definition 5.1. Seien $A(x_1, \ldots, x_n)$ und $B(x_1, \ldots, x_n)$ Aussageformen, die für alle Belegungen von $x_1, \ldots, x_n$ mit Werten $a_1, \ldots, a_n \in X$ definiert sind.

(1) $A(x_1, \ldots, x_n)$ und $B(x_1, \ldots, x_n)$ sind *äquivalent*, wenn für alle $a_1, \ldots, a_n \in X$ die Aussagen $A(a_1, \ldots, a_n)$ und $B(a_1, \ldots, a_n)$ den gleichen Wahrheitswert haben. Wir schreiben dafür $A(x_1, \ldots, x_n) \equiv B(x_1, \ldots, x_n)$.
(2) $A(x_1, \ldots, x_n)$ ist *hinreichend für* $B(x_1, \ldots, x_n)$, wenn für alle $a_1, \ldots, a_n \in X$, für die $A(a_1, \ldots, a_n)$ wahr ist, auch $B(a_1, \ldots, a_n)$ wahr ist.
(3) $B(x_1, \ldots, x_n)$ ist *notwendig für* $A(x_1, \ldots, x_n)$, wenn $A(x_1, \ldots, x_n)$ hinreichend für $B(x_1, \ldots, x_n)$ ist.

So ist sind etwa über den reellen Zahlen die Aussageformen $x + y = 3$ und $x = 3 - y$ äquivalent. Daher kommt der Begriff „Äquivalenzumformung“. Man formt eine Aussageform in eine äquivalente Aussageform um.

2. Quantoren

Wir betrachten die Aussageform $A(x, y)$ über den natürlichen Zahlen, die durch

$$A(x, y) :\Leftrightarrow x = 2y - 1$$

geben ist. Aus dieser Aussageform bilden wir eine Aussageform $B(x)$, die durch

$$B(x) :\Leftrightarrow \text{es gibt ein } y \in \mathbb{N}, \text{sodass } A(x, y) \text{ gilt.}$$

definiert ist. $B(x)$ gilt also genau dann, wenn es ein y in den natürlichen Zahlen gibt, sodass $x = 2y - 1$. Die Aussageform $A(x, y)$ ist für folgende Paare (x, y) wahr: $(1, 1)$, $(3, 2)$, $(5, 3)$, $(7, 4)$, $(9, 5)$, $\ldots$. Die Aussageform $B(x)$ ist genau dann wahr, wenn x eine ungerade natürliche Zahl ist.

Wir betrachten als nächstes folgende Aussageform $C(x, y)$ über den reellen Zahlen:

$$C(x, y) :\Leftrightarrow x^2 = y.$$

Wir bilden eine neue Aussageform $D(y)$ durch

$$D(y) :\Leftrightarrow \text{es gibt ein } x \in \mathbb{R}, \text{sodass } C(x, y) \text{ gilt.}$$

Wir fragen uns nun, ob $D(2)$ gilt. Es gibt zwei x, die die Eigenschaft $C(x, 2)$ erfüllen, da $(\sqrt{2})^2 = 2$ und $(-\sqrt{2})^2 = 2$. Wir normieren die Bedeutung von „es gibt ein“ dahingehend, dass wir damit immer „es gibt mindestens ein“ meinen, und nicht „es gibt genau ein“. Somit ist $D(y)$ genau für die $y \in \mathbb{R}$ erfüllt, die $y \geq 0$ erfüllen. Die Aussageform $D(y)$ ist also äquivalent zu $y \geq 0$.

Definition 5.2. Sei $A(x, y_1, \ldots, y_n)$ eine Aussageform, die für alle $x, y_1, \ldots, y_n \in X$ definiert ist. Wir definieren eine neue Aussageform

$$B(y_1, \ldots, y_n) :\Leftrightarrow \text{es gibt ein } x \in X, \text{sodass } A(x, y_1, \ldots, y_n).$$

Seien $b_1, b_2, \ldots, b_n \in X$. Dann ist $B(b_1, \ldots, b_n)$ genau dann wahr, wenn es mindestens ein $a \in X$ gibt, sodass $A(a, b_1, \ldots, b_n)$ wahr ist.

Für die neue Aussage $B(y_1, \ldots, y_n)$ schreiben wir auch

$$\exists x \in X : A(x, y_1, \ldots, y_n),$$

$(\exists x \in X)\,(A(x, y_1, \ldots, y_n))$, oder $(\exists x \in X)\,A(x, y_1, \ldots, y_n)$. Lies:

- ▷ Es gibt ein $x \in X$, sodass $A(x, y_1, \ldots, y_n)$.
- ▷ Es existiert ein $x \in X$, sodass $A(x, y_1, \ldots, y_n)$.
- ▷ Es gibt ein $x \in X$ mit $A(x, y_1, \ldots, y_n)$.
- ▷ Es existiert ein $x \in X$ mit $A(x, y_1, \ldots, y_n)$.

Wir nennen $\exists$ den *Existenzquantor*. Wenn wir aus einer Aussageform $A(x, y)$ die Aussageform $B(y) = (\exists x \in X)(A(x, y))$ machen, so wird die Variable x „verschluckt“; der Fachausdruck dafür ist, dass x durch den Quantor *gebunden* wird. Wenn eine Variable nicht durch einen Quantor gebunden wird, so heißt sie *frei*.

Wir betrachten nun folgende Aussageform über den reellen Zahlen:

$$E(x, y) :\Leftrightarrow x^2 \geq y.$$

Wir bilden nun eine neue Aussageform $F(y)$ durch

$$F(y) :\Leftrightarrow \text{für alle } x \in \mathbb{R} \text{ gilt } E(x, y).$$

$F(y)$ ist also genau dann wahr, wenn für alle $x \in \mathbb{R}$ gilt, dass $x^2 \geq y$. $F(y)$ ist äquivalent zu $y \leq 0$.

Definition 5.3. Sei $A(x, y_1, \ldots, y_n)$ eine Aussageform, die für alle $x, y_1, \ldots, y_n \in X$ definiert ist. Wir definieren eine neue Aussageform

$$B(y_1, \ldots, y_n) :\Leftrightarrow \text{für alle } x \in X \text{ gilt } A(x, y_1, \ldots, y_n).$$

Seien $b_1, b_2, \ldots, b_n \in X$. Dann ist $B(b_1, \ldots, b_n)$ genau dann wahr, wenn $A(a, b_1, \ldots, b_n)$ für alle $a \in X$ wahr ist.

Für die neue Aussage $B(y_1, \ldots, y_n)$ schreiben wir auch

$$\forall x \in X : A(x, y_1, \ldots, y_n),$$

$(\forall x \in X)\,(A(x, y_1, \ldots, y_n))$ oder $(\forall x \in X)\,A(x, y_1, \ldots, y_n)$. Lies:

$$\text{Für alle } x \in X \text{ gilt } A(x, y_1, \ldots, y_n).$$

Wir nennen $\forall$ den *Allquantor.*

Wir formulieren nun ein paar Zusammenhänge mithilfe dieser Quantoren. Wir betrachten folgende Aussage:

Es gibt eine positive gerade Zahl, die durch 3 und durch 5 teilbar ist.

Wir könnten das so schreiben: sei $G := \{2, 4, 6, \ldots\} = \{x \in \mathbb{N} \mid \exists k \in \mathbb{N} : x = 2k\}$ die Menge der positiven geraden Zahlen. Dann können wir die Aussage so schreiben:

$$\exists x \in G : (3|x) \wedge (5|x)$$

oder so:

$$\exists x \in \mathbb{N} : (x \in G) \wedge (3|x) \wedge (5|x).$$

Diese Aussage ist wahr. Das wird zum Beispiel von $x = 60$ belegt. Die Aussage „das Quadrat einer geraden Zahl ist ein Vielfaches von 4“ können wir so ausdrücken:

$$\forall x \in G : x^2 \text{ ist ein Vielfaches von } 4.$$

Oder, anders:

$$\forall x \in \mathbb{N} : (x \in G) \Rightarrow (4 \mid x^2).$$
$$\forall x \in \mathbb{N} : ((x \in G) \Rightarrow (\exists y \in \mathbb{N} : x^2 = 4y)).$$

Die letzte Zeile liest man etwa so

> Für alle x aus den natürlichen Zahlen gilt: wenn x ein Element von G ist, so gibt es ein y aus den natürlichen Zahlen, sodass x Quadrat gleich 4 mal y ist.

oder so:

> Für alle x aus $\mathbb{N}$ mit $x \in G$ gibt es ein y aus $\mathbb{N}$, sodass $x^2 = 4y$.

Interessant ist, dass die Formalisierung von

$$\text{Es gibt ein } x \in \mathbb{N} \text{ mit } A(x), \text{ sodass } B(x)$$

durch die logische Formel

$$\exists x \in \mathbb{N} : (A(x) \wedge B(x))$$

gegeben ist, die Formalisierung von

$$\text{Für alle } x \in \mathbb{N} \text{ mit } A(x) \text{ gilt } B(x)$$

aber durch

$$\forall x \in \mathbb{N} : A(x) \Rightarrow B(x).$$

Mithilfe mehrerer Quantoren kann man kompliziertere Aussageformen zusammenbauen. Betrachten wir etwa folgende Aussageformen über den reellen Zahlen:

$$\text{Für alle } x \in \mathbb{R} \text{ gibt es ein } y \in \mathbb{R}, \text{ sodass } ay = x.$$

Die Aussageform $ay = x$ ist eine Aussage über a, x, y. Somit ist

$$\forall x \in \mathbb{R} : (\exists y \in \mathbb{R} : ay = x)$$

eine Aussage über a. Die Variablen x und y sind durch Quantoren gebunden. Sie ist äquivalent zur Aussageform $a \neq 0$. Um das zu beweisen, nehmen wir zunächst an, dass $a \neq 0$ ist. Sei nun $x \in \mathbb{R}$. Es ist zu zeigen, dass es $y \in \mathbb{R}$ gibt, sodass $ay = x$. Wir wählen $y := \frac{x}{a}$. Dann gilt $ay = \frac{ax}{a} = x$. Somit belegt $y := \frac{x}{a}$ die Gültigkeit von $\exists y \in \mathbb{R} : ay = x$.

Nehmen wir nun an, dass $a \neq 0$ falsch ist, dass also $a = 0$. Dann gibt es für $x = 2$ kein y mit $0y = 2$. Somit widerlegt $x := 2$ die Gültigkeit von $\forall x \in \mathbb{R} : (\exists y \in \mathbb{R} : ay = x)$. Also ist dann $\forall x \in \mathbb{R} : (\exists y \in \mathbb{R} : ay = x)$ falsch.

Somit sind die beiden Aussagen äquivalent. Andere mögliche Schreibweisen sind:

$$\forall x \in \mathbb{R}\, \exists y \in \mathbb{R} \,:\, ay = x$$

oder

$$(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})\,(ay = x).$$

Die Reihenfolge der Quantoren darf man nicht einfach umdrehen. Wir betrachten dazu

$$\exists y \in \mathbb{R}\ \forall x \in \mathbb{R} \,:\, ay = x. \tag{5.1}$$

Diese Aussageform ist für alle $a \in \mathbb{R}$ falsch: Nehmen wir an, $a \in \mathbb{R}$ ist so, dass $\exists y \in \mathbb{R}\, \forall x \in \mathbb{R} : ay = x$ gilt. Dann gibt es ein $y \in \mathbb{R}$, sodass für alle $x \in \mathbb{R}$ gilt, dass $ay = x$. Damit gilt aber auch für $x = ay + 1$, dass $ay = ay + 1$. Das ist falsch. Dieser Widerspruch zeigt, dass die Annahme, dass es ein a gibt, für das (5.1) gilt, falsch ist. Somit ist (5.1) für jedes $a \in \mathbb{R}$ falsch.

Wir betrachten nun

$$\forall x \in \mathbb{R}\ \forall y \in \mathbb{R} \,:\, ay = x.$$

Das ist für alle $a \in \mathbb{R}$ falsch, da $x = 1$ und $y = 0$ die Gleichung $ay = x$ nicht erfüllen.

Die Aussageform

$$\exists x \in \mathbb{R}\ \forall y \in \mathbb{R} \,:\, ay = x$$

ist für $a = 0$ wahr, und sonst falsch. Wenn $a = 0$, so belegt $x = 0$, dass $\exists x \in \mathbb{R}\, \forall y \in \mathbb{R} : ay = x$ wahr ist. Dann gilt nämlich für jedes $y \in \mathbb{R}$, dass $ay = 0y = 0 = x$. Wenn $a \neq 0$, so ist die Aussage falsch. Nehmen wir an, $x \in \mathbb{R}$ ist so, dass $\forall y \in \mathbb{R} : ay = x$ wahr ist. Dann gilt $a = a \cdot 1 = x = a \cdot 0 = 0$, im Widerspruch zur Annahme, dass $a \neq 0$. Also gibt es kein $x \in \mathbb{R}$ mit $\forall y \in \mathbb{R} : ay = x$.

Die Aussageform

$$\forall y \in \mathbb{R}\ \exists x \in \mathbb{R} \,:\, ay = x$$

ist für alle $a \in \mathbb{R}$ wahr. Sei dazu $a \in \mathbb{R}$ und $y \in \mathbb{R}$. Wir finden das gesuchte x als $x := ay$. Die Aussage

$$\forall a \in \mathbb{R}\ \forall y \in \mathbb{R}\ \exists x \in \mathbb{R} \,:\, ay = x$$

ist also wahr.

Die Aussageform

$$\exists y \in \mathbb{R}\ \exists x \in \mathbb{R} \,:\, ay = x$$

ist für alle $a \in \mathbb{R}$ wahr. Die Werte $y = a$ und $x = a^2$ belegen, dass $\exists y \in \mathbb{R}\ \exists x \in \mathbb{R} : ay = x$ wahr ist.

Mithilfe des Allquantors können wir auch die Begriffe *äquivalent*, *hinreichend* und *notwendig* besser beschreiben.

Satz 5.4. *Seien $A(x_1, \dots, x_n)$ und $B(x_1, \dots, x_n)$ Aussageformen, die für alle $x_1, \dots, x_n \in X$ definiert sind.*

(1) *$A(x_1, \dots, x_n)$ und $B(x_1, \dots, x_n)$ sind genau dann* äquivalent*, wenn die Aussage*
$$(\forall x_1 \in X)(\forall x_2 \in X) \dots (\forall x_n \in X)\ (A(x_1, \dots, x_n) \Leftrightarrow B(x_1, \dots, x_n))$$
wahr ist.

(2) *$A(x_1, \dots, x_n)$ ist genau dann* hinreichend *für $B(x_1, \dots, x_n)$, wenn die Aussage*
$$(\forall x_1 \in X)(\forall x_2 \in X) \dots (\forall x_n \in X)\ (A(x_1, \dots, x_n) \Rightarrow B(x_1, \dots, x_n))$$
wahr ist. Genau dann ist auch $B(x_1, \dots, x_n)$ notwendig *für $A(x_1, \dots, x_n)$.*

Übungsaufgaben 5.5. Die folgenden Beispiele sind Übertragungen aus [**Smi20**].

(1) Mindestens ein blauer Baustein ist eckig. Eckige Bausteine sind nie aus Holz. Kann man daraus folgern: "Mindestens ein blauer Baustein ist nicht aus Holz."?
(2) Mindestens ein blauer Baustein ist eckig. Mindestens ein hölzerner Baustein ist blau. Kann man daraus folgern: "Mindestens ein eckiger Baustein ist aus Holz." ?
(3) Alle blauen Bausteine sind eckig. Mindestens ein blauer Baustein ist aus Holz. Kann man daraus schließen: "Mindestens ein hölzerner Baustein ist eckig."?
(4) Blaue Bausteine sind niemals eckig. Mindestens ein eckiger Baustein ist aus Holz. Kann man daraus schließen: "Mindestens ein hölzerner Baustein ist nicht blau."?
(5) Blaue Bausteine sind niemals eckig. Hölzerne Bausteine sind niemals eckig. Kann man daraus schließen: "Mindestens ein blauer Baustein ist nicht aus Holz."?
(6) Alle blauen Bausteine sind eckig. Kein eckiger Baustein ist aus Holz. Kann man daraus schließen: "Kein hölzerner Baustein ist blau."?
(7) Die Anzahl der blauen Bausteine ist ungerade. Die Anzahl der eckigen Bausteine ist ungerade. Kann man daraus schließen: "Die Anzahl der Bausteine, die blau oder eckig sind, ist gerade."?
(8) Die Anzahl der blauen Bausteine ist ungerade. Die Anzahl der eckigen Bausteine ist ungerade. Kann man daraus schließen: "Die Anzahl der Bausteine, die blau oder eckig, aber nicht beides sind, ist gerade."?

3. Rechenregeln für Quantoren

Die Rechenregeln für Quantoren sind schwieriger als die Rechenregeln für die Aussagenlogik. Wir halten nur einige Äquivalenzen fest.

Satz 5.6 (Umbenennung von Variablen). *Seien $x, y_1, \dots, y_n, z$ voneinander paarweise verschiedene Variablen, und sei $A(x, y_1, \dots, y_n)$ eine Aussageform, die für alle Belegungen von $x, y_1, \dots, y_n$ mit Werten aus X definiert ist. Dann gilt*
$$(\exists x \in X)\, A(x, y_1, \dots, y_n) \equiv (\exists z \in X)\, A(z, y_1, \dots, y_n)$$
und
$$(\forall x \in X)\, A(x, y_1, \dots, y_n) \equiv (\forall z \in X)\, A(z, y_1, \dots, y_n).$$

Beweisskizze: Seien $b_1, \dots, b_n \in X$ so, dass $(\exists x \in X)\, A(x, b_1, \dots, b_n)$ gilt. Dann gibt es ein $a \in X$, sodass $A(a, b_1, \dots, b_n)$ gilt. Also belegt $z := a$, dass auch
$$(\exists z \in X)\, A(z, b_1, \dots, b_n)$$
gilt. Die umgekehrte Implikation und die Aussage über den Allquantor werden ganz ähnlich bewiesen. □

Satz 5.7 (Vertauschung gleicher Quantoren). *Sei $A(x,y)$ eine Aussageform, die für alle $x \in X$ und $y \in Y$ definiert ist. Dann gilt:*

(1) $(\exists x \in X)(\exists y \in Y)\, A(x,y) \equiv (\exists y \in Y)(\exists x \in X)\, A(x,y)$.
(2) $(\forall x \in X)(\forall y \in Y)\, A(x,y) \equiv (\forall y \in Y)(\forall x \in X)\, A(x,y)$.

Der nächste Satz sagt, dass „nicht alle Schafe sind weiß" äquivalent zu „es gibt ein Schaf, das nicht weiß ist" ist.

Satz 5.8 (Regel von De Morgan für Quantoren). *Sei $A(x)$ eine Aussageform, die für alle $x \in X$ definiert ist. Dann gilt:*

(1) $\neg((\exists x \in X)\, A(x)) \equiv (\forall x \in X)(\neg A(x))$.
(2) $\neg((\forall x \in X)\, A(x)) \equiv (\exists x \in X)(\neg A(x))$.

Beweis: (1): Wir nehmen zunächst an, dass $\neg((\exists x \in X)\, A(x))$ wahr ist. Wir wollen zeigen, dass $(\forall x \in X)(\neg A(x))$ gilt. Sei dazu $y \in X$. Wir wollen zeigen, dass $\neg A(y)$ wahr ist. Wenn $A(y)$ wahr ist, so gilt belegt dieses y, dass $(\exists x \in X) A(x)$ wahr ist, im Widerspruch zur Annahme. Also ist $A(y)$ falsch, und somit $\neg A(y)$ wahr. Insgesamt gilt also $(\forall x \in X)\,(\neg A(x))$.

Nehmen wir nun an, dass $(\forall x \in X)(\neg A(x))$ wahr ist. Wir wollen zeigen, dass $\neg((\exists x \in X)\, A(x))$ wahr ist. Dazu zeigen wir, dass $(\exists x \in X)\, A(x)$ falsch ist. Nehmen wir, im Widerspruch zur Behauptung, an, dass $(\exists x \in X)\, A(x)$ wahr ist. Sei $y \in X$ ein Element aus X, das das belegt, also so, dass $A(y)$ gilt. Dann ist $\neg A(y)$ falsch, im Widerspruch zur Annahme, dass $(\forall x \in X)\neg A(x)$ wahr ist. Somit ist die $(\exists x \in X)\, A(x)$ falsch.

(2): wird hier ausgelassen. □

Satz 5.9 (Vorziehen des Quantors). *Seien $A(x)$ und $B(x,y)$ Aussageformen, die für alle $x \in X$ und $y \in Y$ definiert sind. Dann gilt:*

(1) $A(x) \wedge (\exists y \in Y)\, B(x,y) \equiv (\exists y \in Y)\,(A(x) \wedge B(x,y))$
(2) $A(x) \vee (\forall y \in Y)\, B(x,y) \equiv (\forall y \in Y)\,(A(x) \vee B(x,y))$.

Beweis: (1) Nehmen wir an, dass $a \in X$ so ist, dass $A(a) \wedge (\exists y \in Y) B(a,y)$ gilt. Dann gibt es ein $b \in Y$, sodass $B(a,b)$ gilt. Nun belegt dieses b, dass $(\exists y \in Y)\,(A(a) \wedge B(a,y))$ gilt.

Sei umgekehrt $a \in X$ so, dass $(\exists y \in Y)\,(A(a) \wedge B(a,y))$ gilt. Sei $b \in Y$ so, dass $A(a) \wedge B(a,b)$ gilt. Dann gilt $A(a)$, und b belegt, dass $(\exists y \in Y) B(a,y)$ gilt. Somit gilt $A(a) \wedge (\exists y \in Y) B(a,y)$.

Die beiden Aussageformen gelten also für die gleichen $a \in X$, und sind somit äquivalent.

(2) Vorlesung. □

Entscheidend ist hier, dass im Ausdruck $A(x)$ die Variable y nicht vorkommen darf.

Satz 5.10 (Vorziehen des Quantors). *Seien $A(x)$ und $B(x,y)$ Aussageformen, die für alle $x \in X$ und $y \in Y$ definiert sind. Wenn die Menge Y nicht leer ist, so gilt:*

(1) $A(x) \vee (\exists y \in Y) B(x,y) \equiv (\exists y \in Y)\,(A(x) \vee B(x,y))$.
(2) $A(x) \wedge (\forall y \in Y) B(x,y) \equiv (\forall y \in Y)\,(A(x) \wedge B(x,y))$.

Satz 5.11. *Sei $A(x,y)$ eine Aussageform, die für alle $x \in X$ und $y \in Y$ definiert ist. Wenn $(\exists x \in X)(\forall y \in Y)\, A(x,y)$ gilt, so auch $(\forall y \in Y)(\exists x \in X)\, A(x,y)$.*

Wir schließen diesen Abschnitt mit einer auf den ersten Blick überraschenden Tautologie. Als *Tautologie* bezeichnet man eine Eigenschaft von Aussageformen, die für alle Aussageformen gilt, wie etwa $(\forall x \in X)\,(A(x) \vee (\neg A(x)))$.

Satz 5.12. *Sei X eine nichtleere Menge, und sei $A(x)$ eine Aussageform, die für alle $x \in X$ definiert ist. Dann sind die folgenden Aussagen wahr:*

(1) $(\exists x \in X)\Big(A(x) \Rightarrow \big((\forall y \in X)\, A(y)\big)\Big)$.
(2) $(\exists x \in X)(\forall y \in X)\,\big((\neg A(x)) \vee A(y)\big)$.

Beweis: (1) Wir müssen ein $x \in X$ finden, sodass $A(x) \Rightarrow ((\forall y \in X)\, A(y))$ gilt. Wir betrachten dazu zwei Fälle.

- ▷ 1.Fall: $(\forall y \in X)\, A(y)$ *gilt*: Wir benutzen, dass X nicht leer ist und somit ein Element a besitzt. Wir setzen $x := a$. Es gilt dann $A(a) \Rightarrow ((\forall y \in X)\, A(y))$, da wegen der Fallannahme die Konklusion dieser Implikation wahr ist.
- ▷ 2.Fall: $(\forall y \in X)\, A(y)$ *gilt nicht*: Dann gilt $\neg((\forall y \in X)\, A(y))$, also $(\exists y \in X)(\neg A(y))$. Sei $a \in X$ so, dass $\neg A(a)$ gilt. Dann gilt $A(a) \Rightarrow ((\forall y \in X)\, A(y))$, da die Prämisse der Implikation falsch ist.

(2) Übung.

Die Aussage von Satz 5.12 wird manchmal so formuliert:

> Es gibt immer einen, sodass, wenn der einen Hut trägt, alle einen Hut tragen.

Die Aussage $A(x)$ ist dann „x trägt einen Hut“. Abgesehen davon, dass diese Formulierung die Bedingung übergeht, dass die Menge X nicht leer sein soll, könnte sie in folgender Weise missverstanden werden:

> Es gibt ein $x \in X$, sodass für alle Aussageformen A gilt: wenn $A(x)$ gilt, so gilt für alle $y \in X$, dass $A(y)$.

In dieser Weise missverstanden heißt der obige Satz:

> Es gibt einen in unserer Gruppe, wir nennen ihn Franz, sodass folgendes gilt: wann immer Franz einen Hut auf hat, haben alle einen Hut auf.

Es ist aber offensichtlich, dass es so einen Franz nicht geben kann. Es ist nämlich möglich, dass Franz einen Hut trägt, aber sonst niemand. Der Satz mit dem Hut sollte also so verstanden werden:

> Für alle Aussageformen A gibt es ein $x \in X$ mit folgender Eigenschaft: wenn $A(x)$ gilt, so gilt für alle $y \in X$, dass $A(y)$.

Das ist jetzt genau die Aussage von Satz 5.12(3), und wir können dieses x schnell finden. Übertragen auf die Formulierung mit dem Hut ist dieser eine jemand, der keinen Hut trägt, falls es so jemanden gibt, und irgendjemand sonst.

Übungsaufgaben 5.13.

(1) Zeigen Sie: Wenn Y nicht leer ist und $(\forall y \in Y)\,(C(y))$ gilt, so gilt auch $(\exists y \in Y)\,(C(y))$.

Zeigen Sie in den folgenden Beispielen jeweils, dass die Aussagen p und q nicht äquivalent sein müssen, indem Sie konkrete Aussageformen finden, sodass p und q nicht äquivalent sind.

(2) $p = (\forall x \in \mathbb{R})(\exists y \in \mathbb{R})\, A(x,y)$, $q = (\exists y \in \mathbb{R})(\forall x \in \mathbb{R})\, A(x,y)$. *Hinweis*: Versuchen Sie für $A(x,y)$ eine Gleichung in x und y.
(3) $p = (\forall x \in \mathbb{R})(\exists y \in \mathbb{R})\, A(x,y)$, $q = (\exists x \in \mathbb{R})(\forall y \in \mathbb{R})\, A(x,y)$.
(4) $p = (\exists x \in \mathbb{R})\,(A(x) \Rightarrow B)$, $q = ((\exists x \in \mathbb{R})\, A(x)) \Rightarrow B$.
Bestimmen Sie in den folgenden Beispielen, ob p und q für alle Aussageformen äquivalent sind.

(5) $p = \exists x \in \mathbb{R} : A(x) \wedge B(x)$, $q = (\exists x \in \mathbb{R} : A(x)) \wedge (\exists x \in \mathbb{R} : B(x))$.
(6) $p = \exists x \in \mathbb{R} : A(x) \vee B(x)$, $q = (\exists x \in \mathbb{R} : A(x)) \vee (\exists x \in \mathbb{R} : B(x))$.
(7) $p = \forall x \in \mathbb{R} : A(x) \vee B(x)$, $q = (\forall x \in \mathbb{R} : A(x)) \vee (\forall x \in \mathbb{R} : B(x))$.
(8) Seien A, B Aussageformen, sodass $(\forall x \in \mathbb{R})\,(A(x) \Rightarrow B(x))$ und $(\exists x \in \mathbb{R})\,A(x)$ beide gelten. Zeigen Sie, dass $(\exists x \in \mathbb{R})\,(A(x) \wedge B(x))$ gilt.
(9) Ist folgende Aussage für alle Aussageformen wahr? $(\exists x \in \mathbb{R})\,(\forall y \in \mathbb{R})\ (A(y) \Rightarrow A(x))$.
(10) Zeigen Sie Satz 5.12 (2).
(11) Bestimmen Sie jeweils den Wahrheitswert folgender Aussagen.
   (a) $\forall x \in \mathbb{N} : 2 \mid x^2 \Rightarrow 2 \mid x$.
   (b) $\forall x \in \mathbb{N} : 4 \mid x^2 \Rightarrow 4 \mid x$.
   (c) $\exists x \in \mathbb{N} : 4 \mid x^2 \wedge 4 \mid x$.
   (d) $\exists x \in \mathbb{N} : 4 \mid x^2 \wedge (\neg(4 \mid x))$.
(12) Beweisen Sie folgende Aussagen:
   (a) $\forall x \in \mathbb{N}\,\exists z \in \mathbb{N} : x \leq z^2 \wedge z^2 \leq x + 3\sqrt{x}$.
   (b) $\forall x, y \in \mathbb{N} : x \cdot y = 7 \Rightarrow (x = 1 \vee y = 1)$.
   (c) $\forall x, y \in \mathbb{N} : 2 \mid x \cdot y \Rightarrow (2 \mid x \vee 2 \mid y)$.
   Was bedeutet jede dieser Aussagen?
(13) Wir betrachten Aussageformen über den reellen Zahlen.
   (a) Finden Sie eine Aussageform $B(y)$, die zur Aussageform
   $$\exists x \in \mathbb{R} : y + x^2 = 3$$
   äquivalent ist, aber keinen Quantor enthält.
   (b) Finden Sie eine Aussageform $B(y, z)$, die zur Aussageform
   $$\exists x \in \mathbb{R} : (x + 2z = 0 \text{ und } x - 3y = 0)$$
   äquivalent ist, aber keine Quantoren enthält.
(14) Wir betrachten Aussageformen über den reellen Zahlen.
   (a) Finden Sie eine Aussageform $B(y)$, die zur Aussageform
   $$\exists x \in \mathbb{R} : (x + y = 5 \text{ und } x - 3y = -7)$$
   äquivalent ist, aber keine Quantoren enthält.
   (b) Finden Sie eine Aussageform $B(y)$, die zur Aussageform
   $$(\exists z \in \mathbb{R})(\forall x \in \mathbb{R})(x \cdot (y - 3) = z)$$
   äquivalent ist, aber keine Quantoren enthält.

4. Weitere Quantoren

Manchmal verwendet man auch andere Quantoren, etwa in

$$\text{es gibt genau ein } x \in \mathbb{R}, \text{ sodass } x > 0 \text{ und } x^2 = 4.$$

Eine Schreibweise dafür ist

$$\exists!\, x \in \mathbb{R} : (x > 0 \wedge x^2 = 4).$$

Das kann man auch ohne den neuen Quantor $\exists!$ ausdrücken, etwa durch

$$\exists x \in \mathbb{R} : \Big((x > 0 \wedge x^2 = 4) \wedge \big(\forall y \in \mathbb{R} : ((y > 0 \wedge y^2 = 4) \Rightarrow y = x)\big)\Big).$$

Meistens ist es nützlich, die Aussage in eine Konjunktion von „es gibt mindestens ein“ und „es gibt höchstens ein“ zu verwandeln. Die Formalisierung ist dann

$$\big(\exists x \in \mathbb{R} : (x > 0 \wedge x^2 = 4)\big) \wedge$$
$$\big(\forall y \in \mathbb{R}\ \forall z \in \mathbb{R} : (y > 0 \wedge y^2 = 4 \wedge z > 0 \wedge z^2 = 4) \Rightarrow y = z\big).$$

Satz 5.14. *Sei $A(x)$ eine Aussageform, die für alle $x \in X$ definiert ist. Dann sind folgende beiden Aussagen äquivalent.*

(1) $(\exists!\, x \in X)\, A(x)$,
(2) $\Big((\exists x \in X)\, A(x)\Big) \wedge \Big((\forall y \in X)(\forall z \in X)\, \big((A(y) \wedge (A(z))) \Rightarrow y = z\big)\Big)$.

Auch andere Formulierungen, wie „es gibt kein“, „es gibt mindestens zwei“, „es gibt höchstens zwei“, ..., lassen sich mit $\exists, \forall, \neg$ und $=$ gut ausdrücken. Der Sinn davon ist, dass die gebräuchlichen Regeln unseres logischen Schließens genau zu den Quantoren $\exists$ und $\forall$ passen. Einen Beweis von „es gibt genau ein“ führt man etwa oft dadurch, dass man zeigt, dass es mindestens ein Element gibt, und dass zwei Elemente, die beide die geforderte Bedingung erfüllen, gleich sein müssen. Man benützt also Satz 5.14.

Übungsaufgaben 5.15.

(1) Geben Sie Formulierungen von „es gibt kein $x \in X$ mit $A(x)$“, von „es gibt mindestens zwei $x \in X$ mit $A(x)$“, von „es gibt höchstens zwei $x \in X$ mit $A(x)$“ und von „es gibt genau zwei $x \in X$ mit $A(x)$“ an, in denen Sie nur die Quantoren $\exists$ und $\forall$ benützen.

KAPITEL 6

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

$$a \in M,$$

und sagen, dass *a ein Element von M ist.* Zwei Mengen A, B sehen wir als gleich an, wenn sie dieselben Elemente enthalten. Diese Eigenschaft nennt man das *Extensionalitätsaxiom* oder *Axiom der Umfangsbestimmtheit*.

Wir geben zunächst in Form von Beispielen drei wichtige Arten an, in denen man eine Menge angeben kann. Die Menge $A = \{2, 3, 4\}$ ist die Menge, die genau die drei Zahlen 2,3 und 4 als Elemente besitzt. Als nächstes geben wir eine Menge B durch

$$B := \{x \in \mathbb{Z} \mid \exists y \in \mathbb{Z} : y^2 = x\}$$

an. Wir lesen das als „B ist die Menge der x in $\mathbb{Z}$, für die es ein y in $\mathbb{Z}$ gibt, sodass x gleich y zum Quadrat ist“. Die Menge B enthält also die Zahlen 1, 4, 9, 16, 25,.... Schließlich geben wir eine Menge C durch

$$C := \{x^2 \mid x \in \mathbb{Z}, x \leq -5\}$$

an. Wir lesen das als „C ist die Menge aller x^2, wobei x Element von $\mathbb{Z}$ ist, und $x \leq -5$ gilt“. C enthält also alle Quadratzahlen ab 25. Man kann diese Art, C anzugeben, auch als Kurzschreibweise für $C := \{y \in \mathbb{Z} \mid \exists x : (y = x^2 \text{ und } x \in \mathbb{Z} \text{ und } x \leq -5)\}$ sehen.

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 (1871-1953), Abraham Fraenkel (1891-1965) und Thoralf Skolem (1887-1963) 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, ohne dass man bis heute einen Widerspruch in ihnen gefunden hätte. Wir werden nicht alle Axiome diskutieren, aber zumindest das Extensionalitätsaxiom explizit angeben.

Axiom 6.1 (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 $\{1, 2\} = \{2, 1\} = \{2, 1, 1, 2, 2\} = \{x \in \mathbb{N} \mid x \leq 2\} = \{n \in \mathbb{N} \mid \exists a, b, c \in \mathbb{N} : a^n + b^n = c^n\}$, da diese Mengen alle genau dieselben Elemente enthalten. Dass auch die letzte Menge genau $\{1, 2\}$ ist, war allerdings 358 Jahre lang ein offenes Problem (Fermats letzter Satz), bevor es Andrew Wiles[1] im Jahr 1995 gelöst hat.

[1]Andrew Wiles, *1953

Das Extensionalitätsaxiom liefert insbesondere: wenn $A = B$ und $x \in A$ ist, so gilt auch $x \in B$. Zusätzlich brauchen wir auch folgende Eigenschaft von Mengen:

$$\text{Wenn } x = y \text{ und } x \in A, \text{ so gilt auch } y \in A. \tag{6.1}$$

Eine Menge soll also nicht zwischen gleichen Elementen unterscheiden können. Um diese Eigenschaft zu garantieren, fordert man, dass für alle Aussageformen $P(x)$ und alle a, b gilt:

$$(a = b \text{ und } P(a)) \Longrightarrow P(b).$$

Wenn man für das Prädikat $P(x)$ die Aussageform $x \in A$ verwendet, erhält man (6.1).

Definition 6.2 (Teilmenge). Seien A, B Mengen. Dann gilt $A \subseteq B$ genau dann, wenn für alle $a \in A$ auch $a \in B$ gilt.

Wir können nun das Extensionalitätsaxiom so umformulieren:

Seien A, B Mengen. Dann gilt $A = B$ genau dann, wenn $A \subseteq B$ und $B \subseteq A$.

Anstelle von $A \subseteq B$ schreiben wir manchmal $B \supseteq A$.

Übungsaufgaben 6.3.

(1) Geben Sie folgende Mengen in einer der Formen $\{x \in A \mid B(x)\}$ oder $\{f(x) \mid x \in A \text{ und } B(x)\}$ an.
 (a) Die Menge aller natürlicher Zahlen, die durch 7 teilbar sind.
 (b) Die Menge aller reeller Zahlen, die größer oder gleich -1 sind, aber kleiner als 15.
 (c) Die Menge aller ganzzahligen Zweierpotenzen.
 (d) Die Menge aller ganzen Zahlen, die ein Quadrat einer natürlichen Zahl sind.
(2) Geben Sie die Mengen $\{x \in \mathbb{Z} \mid x^2 \leq 7\}$ und $\{x^2 \mid x \in \mathbb{N} \text{ und } x \leq 3\}$ jeweils in der Form $\{a_1, a_2, \ldots, a_n\}$, also durch Aufzählen aller ihrer Elemente an. Wie würden Sie "$\{x \in \mathbb{Z} \mid x^2 \leq 7\}$" und "$\{x^2 \mid x \in \mathbb{N} \text{ und } x \leq 3\}$" jeweils vorlesen?

2. Operationen auf Mengen

Wir werden nun einige Operationen definieren, mithilfe derer wir aus gegebenen neue Mengen bilden können.

Definition 6.4 (Durchschnitt und Vereinigung). Seien A, B Mengen. Dann definieren wir:

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

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

Satz 6.5. *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 zeigen zwei mögliche Arten, diese Gleichungen zu beweisen, und illustrieren diese Arten am Beispiel (5).

Beweis I von Satz 6.5 (5): 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 $(a \vee c) \wedge (b \vee c)$ und $(a \wedge b) \vee c$ für alle Belegungen von a, b, c mit w und f den gleichen Wert annehmen. 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 $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$. □

In diesem Beweis haben wir die Mengengleichheit $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ also auf die Äquivalenz der Aussagen $(a \wedge b) \vee c$ und $(a \vee c) \wedge (b \vee c)$ zurückgeführt. Kürzer könnte man diesen Beweis so formulieren: Es gilt

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

Im zweiten Beweis benutzen wir jetzt eine Technik, die man oft zum Beweisen der Gleichheit zweier Mengen X und Y verwendet. Um $X = Y$ zu zeigen, beweisen wir $X \subseteq Y$ und $Y \subseteq X$. Das Extensionalitätsaxiom sagt dann, dass X und Y gleich sind.

Beweis II von Satz 6.5: Um zu zeigen, dass

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C), \tag{6.2}$$

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 $x \in (A \cap B) \cup C$.

□

Übungsaufgaben 6.6.

(1) Zeigen Sie die übrigen Aussagen von Satz 6.5. Versuchen Sie, Beweise in beiden der illustrierten Arten (also einmal durch Verwendung von Gleichheiten der Aussagenlogik, und einmal durch Beweisen von zwei Inklusionen) anzugeben.

Definition 6.7. 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 6.8. *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 $\varnothing$ bezeichnen wir die leere Menge, die kein Element enthält. Für jede Menge A gilt $\varnothing \subseteq A$ und $A \setminus A = \varnothing$.

Übungsaufgaben 6.9.

(1) Zeigen Sie die übrigen Aussagen von Satz 6.8.
(2) Seien A, B, C Mengen mit $A \cap C = \varnothing$ und $A \cup C = B$. Zeigen Sie, dass $C = B \setminus A$.
(3) Seien A, B, C Mengen, sodass $C \subseteq A \cup B$ und $C \cap A = \varnothing$. Zeigen Sie, dass dann $C \subseteq B$ gilt.
(4) Seien A, B, C Mengen mit $A \cap B \subseteq C$ und $B \subseteq A \cup C$. Zeigen Sie $B \subseteq C$.
(5) Gilt für alle Mengen A, B, C, dass $A \setminus (B \setminus C) = (A \cap C) \cup (A \setminus B)$?
(6) Gilt für alle Mengen A, B, C, dass $(A \setminus C) \setminus (C \setminus B) = A \cup B$?

Definition 6.10. 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. Ein Element x liegt also in $A\Delta B$, wenn es in genau einer der Mengen A und B enthalten ist, wenn also $(x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)$ gilt. Wir werden im folgenden auch brauchen, wann ein Element *nicht* in $A\Delta B$ liegt. Es gilt

$$x \notin A\Delta B \Leftrightarrow (x \in A \wedge x \in B) \vee (x \notin A \wedge x \notin B).$$

Satz 6.11. *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) $A\Delta\varnothing = A$.
(5) $A\Delta A = \varnothing$.
(6) $(A\Delta B) \cap C = (A \cap C)\Delta(B \cap C)$.

Beweis: Wir zeigen (2). Dazu beobachten wir, dass für jedes x gilt:

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

Ähnliche Berechnung zeigen, dass $x \in A\Delta(B\Delta C)$ ebenfalls genau dann gilt, wenn $(x \in A \wedge x \notin B \wedge x \notin C) \vee (x \notin A \wedge x \in B \wedge x \notin C) \vee (x \in A \wedge x \in B \wedge x \in C) \vee (x \notin A \wedge x \notin B \wedge x \in C)$. Das beweist (2)

Wir beweisen nun (6), indem wir beide Inklusionen zeigen. Sei dazu $x \in (A\Delta B) \cap C$. Da $x \in A\Delta B$ liegt, gilt $x \in A \wedge x \notin B$ oder $x \notin A$ und $x \in B$.

- ▷ 1.Fall: $x \in A \wedge x \notin B$: Da $x \in C$, gilt $x \in (A \cap C)$. Es gilt $x \notin (B \cap C)$. Insgesamt gilt also $x \in ((A \cap C)\Delta(B \cap C))$.
- ▷ 2.Fall: $x \notin A \wedge x \in B$: Da $x \in C$, gilt $x \in (B \cap C)$. Es gilt $x \notin (A \cap C)$. Insgesamt gilt also $x \in ((A \cap C)\Delta(B \cap C))$.

Somit ist die Inklusion $\subseteq$ gezeigt.

Für $\supseteq$ wählen wir $x \in (A \cap C)\Delta(B \cap C)$. Da $x \in (A \cap C)\Delta(B \cap C)$ liegt, gilt $x \in (A \cap C) \wedge x \notin (B \cap C)$ oder $x \notin (A \cap C) \wedge x \in (B \cap C)$.

- ▷ 1. Fall: $x \in (A \cap C) \wedge x \notin (B \cap C)$: Da $x \in A \cap C$ gilt $x \in C$ und $x \in A$. Da $x \in C$ und $x \notin B \cap C$, gilt $x \notin B$. Somit gilt $x \in (A\Delta B)$ und $x \in C$, also $x \in (A\Delta B) \cap C$.
- ▷ 2. Fall: $x \notin (A \cap C) \wedge x \in (B \cap C)$: Da $x \in B \cap C$ gilt $x \in C$ und $x \in B$. Da $x \in C$ und $x \notin A \cap C$, gilt $x \notin A$. Somit gilt $x \in (A\Delta B)$ und $x \in C$, also $x \in (A\Delta B) \cap C$.

□

Übungsaufgaben 6.12.

(1) Sei $n \in \mathbb{N}$, und seien $A_1, \dots, A_n$ Mengen. Zeigen Sie, dass

$$x \in ((\dots((A_1\Delta A_2)\Delta A_3)\Delta \dots)\Delta A_{n-1})\Delta A_n$$

genau dann gilt, wenn x in einer ungeraden Anzahl der Mengen $A_1, \dots, A_n$ enthalten ist. *Hinweis:* Induktion nach n.

DEFINITION 6.13. 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 $\mathcal{P}(\{1,2\}) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$.

SATZ 6.14. *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[2]. Sei $n = 1$. Dann gilt $A = \{1\}$ und $\mathcal{P}(A) = \{\varnothing, \{1\}\}$. 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

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

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

ÜBUNGSAUFGABEN 6.15.

(1) Wie lauten die Potenzmengen folgender Mengen:
  (a) $A_1 = \{1, 2, 3, 4\}$.
  (b) $A_2 = \{1, \{1, 3\}, 3\}$.
  (c) $A_3 = \{\{1, 2, \{3\}\}\}$.
(2) Gilt für alle Mengen A, B, dass $\mathcal{P}(A) \cap \mathcal{P}(B) = \mathcal{P}(A \cap B)$?
(3) Gilt für alle Mengen A, B, dass $\mathcal{P}(A) \cup \mathcal{P}(B) = \mathcal{P}(A \cup B)$?

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* oder *Mächtigkeit* 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 8 vornehmen.

Wir halten nun die Kardinalität einiger endlichen Mengen fest.

SATZ 6.16. *Seien $A, B, A_1, \dots, A_n$ endliche Mengen. Dann gilt:*

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

[2]Ein *Beweis mit Induktion* ist der Beweis einer Aussage $\forall n \in \mathbb{N} : A(n)$ dadurch, dass man $A(1)$ und $\forall n \in \mathbb{N} : A(n) \Rightarrow A(n+1)$ zeigt. $A(1)$ heißt dabei der *Induktionsanfang*, und $\forall n \in \mathbb{N} : A(n) \Rightarrow A(n+1)$ der *Induktionsschritt*. Im Beweis von $A(n) \Rightarrow A(n+1)$ heißt $A(n)$ auch die *Induktionsvoraussetzung* und $A(n+1)$ die *Induktionsbehauptung*.

(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

$$|A \cup B| = |A \setminus B| + |B|. \tag{6.3}$$

Nun gilt $(A \cap B) \cup (A \setminus B) = A$ und $(A \cap B) \cap (A \setminus B) = \varnothing$. Folglich gilt $|A \cap B| + |A \setminus B| = |A|$. Wenn man nun in Gleichung (6.3) $|A \setminus B|$ durch $|A| - |A \cap B|$ ersetzt, so erhält man $|A \cup B| = |A| + |B| - |A \cap B|$. □

Satz 6.17. *Sei A eine Menge mit n Elementen, und sei $i \in \{0, \dots, n\}$. Dann hat A genau $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ Teilmengen mit i Elementen.*

Beweis: Zunächst zeigen wir, dass für $n \in \mathbb{N}$ und $i \in \{1, \dots, n-1\}$ gilt:

$$\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1}. \tag{6.4}$$

Es gilt $\binom{n-1}{i} + \binom{n-1}{i-1} = \frac{(n-1)!}{(n-1-i)!\,i!} + \frac{(n-1)!}{(n-i)!\,(i-1)!} = \frac{(n-1)!\cdot(n-i)+(n-1)!\cdot i}{(n-i)!\,i!} = \frac{(n-1)!\cdot n}{(n-i)!\,i!} = \frac{n!}{(n-i)!\,i!} = \binom{n}{i}$. Das beweist (6.4). 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 $\varnothing$, hat. Für $i = n$ hat A genau eine n-elementige Teilmenge, nämlich A. Sei nun $i \in \{1, \dots, n-1\}$. Wir wählen ein Element a aus A. Es gilt

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

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

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

Definition 6.18. 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.

Beispiele:

- ▷ $\bigcup \{\{1, 2\}, \{2, 5\}, \{1, 5, 0\}\} = \{0, 1, 2, 5\}$.
- ▷ $\bigcup\limits_{n \in \mathbb{N}}]n, n+1[= \{x \in \mathbb{R}^+ \mid x \geq 1\} \setminus \mathbb{N}$.
- ▷ $\bigcup \varnothing = \varnothing$.

Definition 6.19. 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

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

definiert ist.

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

Übungsaufgaben 6.20.

(1) Geben Sie die Menge
$$A = \bigcup_{i\in\mathbb{N}} \{(a,a+i)\,|\,a\in\mathbb{N}\}$$
in der Form $A = \{(c,d)\in\mathbb{N}\times\mathbb{N}\,|\,\ldots\ldots\}$ an.

(2) Bestimmen Sie die Menge
$$B = \bigcap_{i\in\mathbb{N}} \{(a,a+i)\,|\,a\in\mathbb{N}\}.$$

3. Geordnete Paare

Definition 6.21. Für beliebige a, b definieren wir das *geordnete Paar* (a, b) durch
$$(a,b) := \{\{a\},\{a,b\}\}.$$

Satz 6.22. *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 6.23. Seien A, B Mengen. Wir definieren $A \times B$, das *kartesische Produkt von A und B*, durch
$$A \times B := \{(a,b)\,|\,a\in A \text{ und } b\in B\}.$$

Satz 6.24. *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) $\varnothing\times B = \varnothing$.
(4) *Wenn $A\times B = \varnothing$, so gilt $A = \varnothing$ oder $B = \varnothing$.*

Beweis: Wir geben nur die Beweise von (1) und (4) an. (1): Es gilt $(x,y)\in(A\cup B)\times X \Leftrightarrow (x\in A\vee x\in B)\wedge y\in X \Leftrightarrow (x\in A\wedge y\in X)\vee(x\in B\wedge y\in X) \Leftrightarrow (x,y)\in(A\times X)\cup(B\times X)$. (4): Nehmen wir an $A\neq\varnothing$ und $B\neq\varnothing$. Dann gibt es $a\in A$ und $b\in B$. Somit gilt $(a,b)\in A\times B$, und folglich gilt $A\times B\neq\varnothing$. □

4. Russellsches Paradoxon

Satz 6.25. *Sei A eine Menge, und sei $B := \{a\in A\,|\,a\notin a\}$. 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 *Russellsche Paradoxon*.

KAPITEL 7

Funktionen

1. Relationen

Definition 7.1. Seien A, B Mengen. Jede Teilmenge von $A \times B$ heißt auch *Relation von A nach B*.

Beispiele:

(1) Sei $A := \{\text{Wien}, \text{Niederösterreich}, \text{Oberösterreich}, \text{Salzburg}, \text{Tirol}, \text{Vorarlberg}, \text{Burgenland}, \text{Steiermark}, \text{Kärnten}\}$ die Menge der neun österreichischen Bundesländer, und sei $B := \{\text{Donau}, \text{Inn}, \text{Traun}\}$. Wir definieren eine Relation R durch
$$R := \{(a, b) \in A \times B \mid \text{ in } a \text{ liegt ein Teil des Ufers von } b\}.$$
Wir erhalten $R = \{(\text{Wien}, \text{Donau}), (\text{Niederösterreich}, \text{Donau}), (\text{Oberösterreich}, \text{Donau}), (\text{Tirol}, \text{Inn}), (\text{Oberösterreich}, \text{Inn}), (\text{Steiermark}, \text{Traun}), (\text{Oberösterreich}, \text{Traun})\}$.

(2) Für $(a, b) \in R$ schreiben wir auch $a\, R\, b$. Sei nun $A := \mathbb{R}$ und $B := \mathbb{Z}$. Wir definieren eine Relation ρ 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 $\rho = \{(r, n) \in \mathbb{R} \times \mathbb{N} \mid n \leq r < n + 1\}$.

(3) 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 „kleiner-gleich“-Relation.

(4) 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
$$\equiv_5 \; = \; \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid b - a \text{ ist Vielfaches von } 5\}.$$

2. Funktionen

Definition 7.2. 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 7.3. 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:
$$\begin{array}{rccl} q : & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & x & \longmapsto & x^2. \end{array}$$
Man liest das als „q ist eine Funktion von $\mathbb{Z}$ nach $\mathbb{Z}$, die jedes x aus $\mathbb{Z}$ 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.*

Übungsaufgaben 7.4.

(1) Welche der folgenden Relationen sind Funktionen von $\mathbb{N}$ nach $\mathbb{R}$? Begründen Sie Ihre Antwort, und geben Sie jene Relationen, die Funktionen sind, auch in der Form
$$\begin{array}{rccl} \ldots : & \mathbb{N} & \to & \mathbb{R} \\ & x & \mapsto & \ldots \end{array}$$
an.
(a) $f = \{(x^3, x) \mid x \in \mathbb{R}\} \cap (\mathbb{N} \times \mathbb{R})$.
(b) $g = \{((x-1)(x-2), x) \mid x \in \mathbb{R}\} \cap (\mathbb{N} \times \mathbb{R})$.
(c) $h = \{(a,b) \in \mathbb{N} \times \mathbb{R} \mid b = \frac{a}{3}\}$.

Definition 7.5 (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
$$\begin{array}{rccl} f : & T & \longrightarrow & B \\ & t & \longmapsto & f(t) \end{array}$$
gegeben ist. Sie heißt *Einschränkung von f auf T.*

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

Definition 7.6. Seien A, B Mengen. Mit B^A bezeichnet man die Menge aller Funktionen von A nach B. Genauer:
$$B^A := \{f \in \mathcal{P}(A \times B) \mid f : A \to B\}.$$

Satz 7.7. *Seien $n, m \in \mathbb{N}$, und sei $A = \{1, \ldots, m\}$, $B := \{1, \ldots, n\}$. Dann hat B^A genau n^m Elemente.*

Beweisskizze: Um zu zählen, wieviele Funktionen f von $\{1, \ldots, m\}$ nach $\{1, \ldots, n\}$ es gibt, beobachten wir, dass wir n Möglichkeiten für $f(1)$, n Möglichkeiten für $f(2)$, $\ldots$, und n Möglichkeiten für $f(m)$ haben. Insgesamt gibt es also n^m Funktionen. $\square$

3. Definitions- und Wertebereich

Definition 7.8. Seien A, B Mengen, und sei f eine Funktion von A nach B. Dann heißt A auch der *Definitionsbereich* von f. Der *Wertebereich* von f ist die Menge $\{f(a) \mid a \in A\}$.

Den Wertebereich von f bezeichnet man auch als *Bildbereich* von f. Der Wertebereich einer Funktion von A nach B enthält also jene Elemente in B, die tatsächlich als Funktionswert auftreten. Er muss nicht gleich der ganzen Menge B sein. Wenn f eine Funktion von A nach B ist, so bezeichnen wir B auch als einen *Wertevorrat* oder eine *Zielmenge* von f.

Definition 7.9. Seien A, B Mengen, und sei f eine Funktion von A nach B. Sei $T \subseteq A$. Dann bezeichnen wir mit $f[T]$ die *Bildmenge von T unter f*, die 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 7.10. *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) $f(C \cap D) \subseteq f(C) \cap f(D)$.

Übungsaufgaben 7.11.

(1) Beweisen Sie Satz 7.10.
(2) Finden Sie ein Beispiel, für das $f(C \cap D) \neq f(C) \cap f(D)$ ist.

Definition 7.12. 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 auf B ist.

Die Funktion f ist also injektiv, wenn es kein $x, y \in A$ mit $x \neq y$ und $f(x) = f(y)$ gibt. Wenn für eine Funktion $f : A \to B$ klar ist, welches B gemeint ist, sagt man oft einfach „f ist surjektiv" anstelle von „f ist surjektiv auf B". Im folgenden arbeiten wir darauf hin, die Wirkung einer Funktion wieder rückgängig zu machen, also, wenn möglich, aus dem Bild $f(x)$ einer Funktion das Argument x zu rekonstruieren.

Satz 7.13. *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 7.14. 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* oder *Umkehrfunktion von f*, und wird mit f^{-1} abgekürzt.

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

Definition 7.15. 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*.

4. Familien und Folgen

Wir können es bestimmt nicht besser formulieren als P. Halmos [**Hal76**, 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 7.16. 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 $(x_i)_{i\in I}$. Wenn f eine Funktion von A nach B, und C eine Teilmenge von A ist, schreibt man

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

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

Definition 7.17. Sei A eine Menge, sei $n \in \mathbb{N}$, und seien $a_1, \ldots, a_n \in A$. Mit dem *n-Tupel* $\langle a_1, \ldots, a_n\rangle$ meinen wir die Familie $\langle a_i \mid i \in \{1, \ldots, n\}\rangle$.

Eine mit der Menge I indizierte Familie $(a_i)_{i\in I}$ aus der Menge X ist also eine Funktion von I nach X. Familien mit Indexmenge $\mathbb{N}$ oder $\mathbb{N}_0$ heißen auch *Folgen*, Familien mit Indexmenge $\{1, \ldots, n\}$ auch *Tupel* oder *(endliche) Folgen*. Ein n-Tupel $\langle a_1, \ldots, a_n\rangle$ sehen wir also als eine mit der Indexmenge $\{1, \ldots, n\}$ indizierte Familie an. Das n-Tupel $\langle a_1, \ldots, a_n\rangle$ schreiben wir auch als $(a_1, \ldots, a_n)$; dass jetzt (a, b) zwei verschiedene (aber in der Praxis sehr ähnliche) Bedeutungen haben kann, stört meist nicht.

Definition 7.18. 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 7.19. 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 7.20 (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[1]: 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[2], 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 7.21.

(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?

5. Hintereinanderausführung von Funktionen

Definition 7.22. 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

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

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

[1]Kurt Gödel, 1906-1978

[2]Paul Cohen, 1934-2007

Satz 7.23 (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 $(h \circ g) \circ f = h \circ (g \circ f)$.*

Beweis: Zwei Funktionen α und β sind genau dann gleich, wenn sie den gleichen Definitionsbereich haben, und für alle x aus dem Definitionsbereich gilt, dass $\alpha(x) = \beta(x)$. Sei also $x \in A$. Dann gilt $((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x)) = h((g \circ f)(x)) = (h \circ (f \circ g))(x)$. ☐

Satz 7.24 (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 7.25. *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.*

Beweis: (1) Sei $c \in C$. Da $g \circ f$ surjektiv ist, gibt es $a \in A$, sodass $g(f(a)) = c$. Dann belegt $b := f(a)$, dass es ein $b \in B$ gibt, sodass $g(b) = c$. Somit ist g surjektiv. (2) Seien $a_1, a_2 \in A$ so, dass $f(a_1) = f(a_2)$. Dann gilt auch $g(f(a_1)) = g(f(a_2))$. Da $g \circ f$ injektiv ist, erhalten wir $a_1 = a_2$. Somit ist f injektiv. ☐

Übungsaufgaben 7.26.

(1) Finden Sie Mengen A, B, C, eine Funktion $f : A \to B$ und eine Funktion $g : B \to C$, sodass g surjektiv und $g \circ f$ nicht surjektiv ist.
(2) Finden Sie Mengen A, B, C und $f : A \to B$, $g : B \to C$, sodass f injektiv und $g \circ f$ nicht injektiv ist.
(3) Finden Sie Mengen A, B, C und $f : A \to B$, $g : B \to C$, sodass $g \circ f$ surjektiv und f nicht surjektiv ist.
(4) Finden Sie Mengen A, B, C und $f : A \to B$, $g : B \to C$, sodass $g \circ f$ injektiv und g nicht injektiv ist.

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

Satz 7.27. *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 7.25 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 7.28. *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 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Beweis Es gilt $(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ ((f \circ f^{-1}) \circ g^{-1}) = g \circ (\mathrm{id}_B \circ g^{-1}) = g \circ g^{-1} = \mathrm{id}_C$ und $(f^{-1} \circ g^{-1}) \circ (g \circ f) = (f^{-1} \circ (g^{-1} \circ g)) \circ f = (f^{-1} \circ \mathrm{id}_B) \circ f = f^{-1} \circ f = \mathrm{id}_A$. Somit gilt wegen Satz 7.27, dass $f^{-1} \circ g^{-1} = (g \circ f)^{-1}$. ☐

6. Äquivalenzrelationen und Partitionen

Wir nennen eine Relation von A nach A auch eine *Relation auf* A.

DEFINITION 7.29. Sei ρ eine Relation auf A.

(1) ρ ist *reflexiv*, wenn für alle $a \in A$ gilt: $(a, a) \in \rho$.
(2) ρ 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) ρ ist *symmetrisch*, wenn für alle $a, b \in A$ gilt: wenn $(a, b) \in \rho$, so gilt auch $(b, a) \in \rho$.

DEFINITION 7.30. Sei ρ eine Relation auf A. Die Relation ρ ist eine *Äquivalenzrelation auf* A, wenn sie reflexiv, transitiv und symmetrisch ist.

DEFINITION 7.31. Sei ρ eine Äquivalenzrelation auf A, und sei $a \in A$. Die *Äquivalenzklasse von* a *bezüglich* ρ wird mit $[a]_\rho$ oder a/ρ abgekürzt, und ist definiert durch

$$a/\rho := \{b \in A \mid (a, b) \in \rho\}.$$

Eine Teilmenge C von A ist eine *Äquivalenzklasse von* ρ, wenn es ein $a \in A$ gibt, sodass $C = a/\rho$.

LEMMA 7.32. *Sei ρ 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 ρ gilt auch $(b, a) \in \rho$, und somit wegen der Transitivität von ρ 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 ρ auch $(a, c) \in \rho$, und somit $c \in [a]_\rho$. ☐

ÜBUNGSAUFGABEN 7.33.

(1) Sei ρ 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) $[a]_\rho \cap [b]_\rho \neq \varnothing$.
(2) Geben Sie ein Beispiel für eine Äquivalenzrelation auf $A := \{2, 3, 4, 5\}$ an. Geben Sie die Relation in der Form $\rho = \{\ldots\}$ an!

DEFINITION 7.34. Sei A eine Menge. Eine Teilmenge $\mathcal{P}$ von $\mathcal{P}(A)$ ist eine *Partition von* A, wenn

(1) für alle $P \in \mathcal{P} : P \neq \varnothing$,
(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 $P_1 \cap P_2 = \varnothing$.

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 7.35. Sei ρ eine Äquivalenzrelation auf A. Die *Faktormenge von* A *modulo* ρ ist die Menge $A/\rho := \{[a]_\rho \mid a \in A\}$.

SATZ 7.36. *Sei ρ eine Äquivalenzrelation auf A. Dann ist die Faktormenge von A bezüglich ρ 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 ρ gilt $(a, a) \in \rho$, und folglich $a \in [a]_\rho$, also $a \in P$. Somit gilt $P \neq \varnothing$.

Wir zeigen nun, dass jedes $a \in A$ Element eines Elementes von A/ρ ist. Sei dazu $a \in A$. Dann gilt wegen der Reflexivität von ρ, dass $a \in [a]_\rho$. Somit ist a Element eines Elementes von A/ρ, 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 $P \cap Q \neq \varnothing$. 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 ρ gilt daher auch $(c, b) \in \rho$, und daher, wegen der Transitivität von ρ, auch $(a, b) \in \rho$. Somit gilt nach Lemma 7.32 auch $P = Q$. □

Satz 7.37. *Sei A eine Menge, und sei $\mathcal{P}$ eine Partition von A. Dann ist*

$$\rho := \{(a, b) \in A \times A \,|\, \exists P \in \mathcal{P} : a \in P \text{ und } b \in P\}$$

eine Äquivalenzrelation auf A.

Definition 7.38. Sei A eine Menge, und sei ρ eine Äquivalenzrelation auf A. Eine Teilmenge R von A ist ein *Repräsentantensystem* von A modulo ρ, wenn für alle $a \in A$ die Menge $[a]_\rho \cap R$ genau ein Element enthält.

7. Zahlen als Äquivalenzklassen

Mithilfe von Äquivalenzrelationen können wir aus den natürlichen Zahlen die ganzen Zahlen konstruieren. Sei

$$M := \mathbb{N}_0 \times \mathbb{N}_0.$$

Das Paar (a, b) soll $a - b$ beschreiben. Dann sollen (a, b) und (c, d) die gleiche Zahl beschreiben, wenn $a - b = c - d$, also wenn $a + d = c + b$. Daher definieren wir eine Äquivalenzrelation δ durch

$$\big((a, b), (c, d)\big) \in \delta :\Leftrightarrow a + d = c + b.$$

Diese Relation ist reflexiv: Sei $(x, y) \in M$. Dann gilt $x + y = x + y$, also $\big((x, y), (x, y)\big) \in \delta$. Sie ist symmetrisch: Sei $\big((a, b), (c, d)\big) \in \delta$. Dann gilt $a + d = c + b$, also $c + b = a + d$, und somit $\big((c, d), (a, b)\big) \in \delta$. Sie ist transitiv: Seien $\big((a, b), (c, d)\big) \in \delta$ und $\big((c, d), (e, f)\big) \in \delta$. Dann gilt $a + d = c + b$ und $c + f = e + d$. Also gilt $a + d + c + f = c + b + e + d$. Somit gilt $a + f = e + b$, also $\big((a, b), (e, f)\big) \in \delta$. Wir definieren nun Z als die Faktormenge M/δ. Nun ist $\{(n, 0) \mid n \in \mathbb{N}_0\} \cup \{(0, n) \mid n \in \mathbb{N}\}$ ein Repräsentantensystem von M modulo δ. Für $n \in \mathbb{N}$ kürzen wir die Klasse $(0, n)/\delta$ mit $-n$ ab. Für die Klasse $(n, 0)/\delta$ schreiben wir einfach $+n$. Dann gilt $Z = \{-3, -2, -1, +0, +1, +2, +3, \ldots\}$.

Auch für die Einführung der rationalen Zahlen verwenden wir eine Äquivalenzrelation. Dabei klären wir zum Beispiel auch, ob $\frac{3}{4} = \frac{6}{8}$ gilt. 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 ρ eine Äquivalenzrelation, und $R := \{\binom{a}{b} \in A \,|\, b > 0, \operatorname{ggT}(a, b) = 1\}$ ist ein Repräsentantensystem. Die Faktormenge A/ρ bezeichnet man als die Menge der *rationalen Zahlen*. Für $[\binom{a}{b}]_\rho$ schreibt man $\frac{a}{b}$. Da $\frac{3}{4} = [\binom{3}{4}]_\rho = [\binom{6}{8}]_\rho = \frac{6}{8}$, gilt also wirklich $\frac{3}{4} = \frac{6}{8}$. Den Repräsentanten aus R eines Bruchs bezeichnet man als seine *gekürzte Darstellung*.

Übungsaufgaben 7.39.

(1) Geben Sie die Partition $\mathcal{P}$ der Menge $M = \{1, 2, 3\}$ an, die von der Äquivalenzrelation $\alpha = \{(1, 1), (2, 2), (3, 3), (1, 3), (3, 1)\}$ induziert wird.
(2) Geben Sie die Äquivalenzrelation β auf $M = \{1, 2, 3, 4\}$ an, die die Partition $\mathcal{P} = \{\{1\}, \{2, 4\}, \{3\}\}$ induziert.

8. Ordnungsrelationen

Definition 7.40. Sei M eine Menge, und sei ρ eine Relation auf M. Die Relation ρ ist *antisymmetrisch*, wenn für alle $x, y \in M$ mit $(x, y) \in \rho$ und $(y, x) \in \rho$ gilt: $x = y$.

Definition 7.41. Sei M eine Menge, und sei ρ eine Relation auf M. Die Relation ρ ist eine *Ordnungsrelation*, wenn sie *reflexiv*, *transitiv* und *antisymmetrisch* ist.

Definition 7.42. Sei M eine Menge, und sei $\leq$ eine Ordnungsrelation auf M. Die Relation $\leq$ ist *linear* (oder *total*) 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 7.43. 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) Sei T eine Teilmenge von M, und sei $x \in M$. Dann ist x ein *Infimum* von T, wenn x ein größtes Element der Menge U aller unteren Schranken von T ist. (Es kann auch Teilmengen T geben, die kein Infimum besitzen.)
(6) a ist ein *maximales Element* von M, wenn es kein b in M mit $a < b$ gibt.
(7) 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: $t \leq m$.
(8) Sei T eine Teilmenge von M, und sei $x \in M$. Dann ist x ein *Supremum* von T, wenn x ein kleinstes Element der Menge O aller oberen Schranken von T ist. (Es kann auch Teilmengen T geben, die kein Supremum besitzen.)

Eine geordnete Menge $(M, \leq)$ hat höchstens ein kleinstes Element. Jedes kleinste Element ist minimal.

KAPITEL 8

Die Mächtigkeit von Mengen

1. Gleichmächtige Mengen

Wir rufen uns die Definition davon, dass zwei Mengen gleichmächtig sind, in Erinnerung:

Definition 8.1. 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.

Diese Definition lässt sich auch für unendliche Mengen verwenden.

Satz 8.2. *Sei C eine Menge. Dann ist $\sim$ auf $\mathcal{P}(C)$ eine Äquivalenzrelation.*

Unendliche Mengen bieten das erstaunliche Phänomen, dass eine Menge gleichmächtig zu einer echten Teilmenge sein kann.

Proposition 8.3. $\mathbb{Z} \sim \mathbb{N} \sim \mathbb{N} \times \mathbb{N}$.

Beweis: Die Funktion

$$\begin{array}{rcccl} f & : & \mathbb{Z} & \longrightarrow & \mathbb{N} \\ & & x & \longmapsto & \begin{cases} -2x+2 \text{ wenn } x \leq 0 \\ 2x-1 \text{ wenn } x \geq 1 \end{cases} \end{array}$$

ist bijektiv. Ebenso ist

$$\begin{array}{rcccl} f & : & \mathbb{N} \times \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & (x,y) & \longmapsto & 2^{x-1} \cdot (2y-1) \end{array}$$

bijektiv. ☐

Definition 8.4. 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 8.5. Seien A, B Mengen. Wir sagen, dass A *höchstens so mächtig wie* B (oder B *mindestens so mächtig wie* A) ist, wenn es eine injektive Funktion von A nach B gibt. Das kürzen wir mit $A \precsim B$ ab. B ist *mächtiger als* A, wenn B mindestens so mächtig wie A ist, und A und B nicht gleichmächtig sind.

Übungsaufgaben 8.6.

(1) Zeigen Sie, dass das reelle Intervall $[0, 4]$ gleichmächtig zu $[0, 2]$ ist.
(2) Seien A, B Mengen mit $A \precsim B$. Zeigen Sie, dass es eine surjektive Funktion von B auf A gibt.
(3) 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 $\prod_{a \in A} s^{-1}[\{a\}]$.
(4) 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 $\mathcal{P}(A_1) \sim \mathcal{P}(A_2)$.
(5) 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 8.7. *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 8.8. *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

$$g[V \cup U_1] \subseteq U_1. \tag{8.1}$$

Sei dazu $w \in V \cup U_1$. Wir wollen zeigen, dass $g(w) \in \bigcap \{B \mid B \subseteq U \text{ und } g[V \cup B] \subseteq B\}$. 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 (8.1).

Nun zeigen wir

$$g[V \cup U_1] = U_1. \tag{8.2}$$

Nehmen wir 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 (8.2).

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. □

Dieses Lemma ist der entscheidende Schritt, um den folgenden Satz zu beweisen.

SATZ 8.9 (Satz von Cantor-Schröder-Bernstein[1]). *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 8.8 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 $A \sim B$. □

Zuletzt überlegen wir uns noch, ob für zwei Mengen stets $A \precsim B$ oder $B \precsim A$ gilt, oder ob es Mengen „unvergleichbarer Mächtigkeit" geben kann. Die Antwort wird sein, dass es unter Annahme des Auswahlaxioms keine solchen Mengen unvergleichbarer Mächtigkeit geben kann. Wir werden im Beweis aber nicht das Auswahlaxiom verwenden, sondern einen Satz, das Lemma von Zorn[2]. Das Lemma von Zorn ist äquivalent zum Auswahlaxiom; wir könnten anstelle des Auswahlaxioms für Mengen also auch das Lemma von Zorn als Axiom fordern und würden dann das Auswahlaxiom als Satz erhalten.

SATZ 8.10 (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.

---

[1] Georg Cantor (1845-1918), Ernst Schröder (1841-1902), Felix Bernstein (1878-1956)

[2] Max August Zorn (1906-1993)

Dabei ist mit $(T, \leq)$ genau genommen $(T, \leq \cap (T \times T))$ gemeint. 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. In der Praxis ist das Lemma von Zorn ein hilfreiches Instrument zum Beweis für die Existenz von Dingen, die in irgendeinem Sinn „maximal" sind. Eine Anwendung geben wir im folgenden Satz.

Satz 8.11 (Vergleichbarkeitssatz). *Seien A, B Mengen. Dann gilt $A \precsim B$ oder $B \precsim A$.*

Beweis: Sei

$$\mathcal{F} := \{f \subseteq A \times B \mid \text{es gibt } C \subseteq A, \text{ sodass } f \text{ eine injektive Funktion von } C \text{ nach } B \text{ ist}\}.$$

Wir verwenden nun das Lemma von Zorn, um zu zeigen, dass $(\mathcal{F}, \subseteq)$ ein maximales Element f_0 besitzt. Sei dazu $\mathcal{T}$ eine mit $\subseteq$ linear geordnete Teilmenge von $\mathcal{F}$. Wir bilden die Menge

$$g := \bigcup \mathcal{T} = \bigcup \{f \mid f \in \mathcal{T}\}.$$

Die Menge g ist also die Vereinigung aller Funktionen in $\mathcal{T}$. Wir zeigen nun als erstes, dass g wieder eine funktionale Relation von einer Teilmenge von A nach B ist. Sei dazu $C := \{a \in A \mid \exists b \in B : (a, b) \in g\}$. Wir zeigen, dass g eine Funktion von C nach B ist. Sei dazu $a \in C$. Aus der Definition von C geht unmittelbar hervor, dass es ein b gibt, sodass $(a, b) \in g$. Um die Funktionalität zu zeigen, müssen wir noch nachweisen, dass dieses b eindeutig ist. Seien also $b_1, b_2 \in B$ so, dass $(a, b_1) \in g$ und $(a, b_2) \in g$. Dann gibt es f_1 und f_2 in $\mathcal{T}$, sodass $(a, b_1) \in f_1$ und $(a, b_2) \in f_2$. Die Menge $\mathcal{T}$ ist linear geordnet, also gilt $f_1 \subseteq f_2$ oder $f_2 \subseteq f_1$. Wenn $f_1 \subseteq f_2$, so gilt $(a, b_1) \in f_2$ und $(a, b_2) \in f_2$. Da f_2 eine Funktion ist, gilt also $b_1 = b_2$. Im Fall $f_2 \subseteq f_1$ erhalten wir $(a, b_2) \in f_1$, und somit $b_1 = b_2$, weil f_1 funktional ist. Somit ist g eine Funktion.

Wir zeigen als nächstes, dass g injektiv ist. Seien dazu $a_1, a_2 \in g$ so, dass $g(a_1) = g(a_2)$. Nach der Konstruktion von g gibt es $h_1, h_2 \in \mathcal{T}$, sodass $(a_1, g(a_1)) \in h_1$ und $(a_2, h(a_2)) \in h_2$, also $h_1(a_1) = g(a_1)$ und $h_2(a_2) = g(a_2)$. Die Menge $\mathcal{T}$ ist linear geordnet, und folglich gilt $h_1 \subseteq h_2$ oder $h_2 \subseteq h_1$. Wenn $h_1 \subseteq h_2$, so gilt $(a_1, g(a_1)) \in h_2$, und somit $h_2(a_1) = g(a_1)$. Dann gilt $h_2(a_1) = g(a_1) = g(a_2) = h_2(a_2)$. Nun verwenden wir, dass h_2 injektiv ist, und erhalten $a_1 = a_2$. Ebenso erhalten wir im Fall $h_2 \subseteq h_1$, dass $a_1 = a_2$. Die Funktion g ist also injektiv.

Somit liegt g in $\mathcal{F}$, und g ist eine obere Schranke für die Menge $\mathcal{T}$. Nun verwenden wir das Lemma von Zorn. Es liefert uns ein maximales Element f_0 von $\mathcal{F}$.

Wenn der Definitionsbereich C_0 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_0 nach B; also ist f_0^{-1} injektiv von B nach C_0, und es gilt $B \precsim A$.

Der verbleibende Fall ist, dass $C_0 \neq A$ und $f_0[C_0] \neq B$. In diesem Fall wählen wir $a_0 \in A \setminus C_0$ und $b_0 \in B \setminus f_0[C_0]$. Dann ist $f_0 \cup \{(a_0, b_0)\}$ ebenfalls eine injektive Funktion, also $f_0 \in \mathcal{F}$, im Widerspruch zur Maximalität von f_0. □

2. Abzählbar und überabzählbar unendliche Mengen

Satz 8.12. *Es gilt $\mathbb{Q} \sim \mathbb{N}$.*

Beweis: Nach dem Satz von Cantor-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

$$g(\tfrac{a}{b}) := (a/\mathrm{ggT}(a, b), b/\mathrm{ggT}(a, b)),$$

für $a, b \in \mathbb{Z}$ mit $b > 0$. 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 $h : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}$, $(a, b) \mapsto \frac{a}{b}$ 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 8.13. *Sei* $\langle A_i \,|\, i \in \mathbb{N}\rangle$ *eine Familie von Mengen. Wir nehmen an, dass für alle* $i \in \mathbb{N}$ *gilt, dass* $A_i \precsim \mathbb{N}$. *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 $f : \bigcup\{A_i \,|\, i \in \mathbb{N}\} \to \mathbb{N} \times \mathbb{N}$ durch $f(a) := (k_1, k_2)$, wobei $k_1 := \min\{j \in \mathbb{N} \,|\, a \in A_j\}$ und $k_2 := f_{k_1}(a)$. Diese Abbildung ist injektiv und beweist $\bigcup\{A_i \,|\, i \in \mathbb{N}\} \precsim \mathbb{N} \times \mathbb{N}$. Wegen $\mathbb{N} \times \mathbb{N} \precsim \mathbb{N}$ folgt die Behauptung. □

ÜBUNGSAUFGABEN 8.14.

(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 $A^* := \bigcup\{A^n \,|\, n \in \mathbb{N}\}$ abzählbar unendlich ist.
(6) Zeigen Sie, dass die Menge der endlichen Teilmengen von $\mathbb{N}$ gleichmächtig zu $\mathbb{N}$ ist.

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 8.15. *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
$$B := \{x \in A \,|\, x \notin f(x)\}.$$
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 $\mathcal{P}(A)$. □

Damit haben wir die entscheidende Information, um folgenden Satz zu beweisen:

SATZ 8.16 (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 8.15. Also gilt $A \nsim \mathcal{P}(A)$. □

Damit haben wir also unendliche Mengen gefunden, die nicht abzählbar sind, etwa $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathbb{Z})$, $\{0,1\}^{\mathbb{N}}$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$.

SATZ 8.17. *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 $g : \mathbb{R} \to \mathcal{P}(\mathbb{N})$, $r \mapsto \{i \in \mathbb{N} \,|\, q_i < r\}$ injektiv, da zwischen zwei verschiedenen reellen Zahlen stets eine rationale Zahl liegt. Somit gilt nach dem Satz von Cantor-Schröder-Bernstein $\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$. □

Übungsaufgaben 8.18.

(1) Zeigen Sie, dass das reelle Intervall $[0,2]$ gleichmächtig zu $\mathbb{R}$ ist.

Für den Beweis des nächsten Satzes brauchen wir das Auswahlaxiom.

Satz 8.19. *Jede unendliche Menge M enthält eine abzählbar unendliche Teilmenge.*

Beweis: Sei $f : \mathcal{P}(M) \setminus \{\varnothing\} \to M$ so, dass $f(A) \in A$ für alle $A \subseteq M$. So ein f existiert, weil nach dem Auswahlaxiom die Menge $\prod_{A \in \mathcal{P}(M)\setminus\{\varnothing\}} A$ 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 8.20.

(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 8.21. *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: [**Hal76**, Kapitel 24]. □

Übungsaufgaben 8.22.

(1) Zeigen Sie ohne Verwendung des Teils (3) von Satz 8.21, dass $\mathbb{N}^{\mathbb{R}} \sim \mathcal{P}(\mathbb{R})$.
(2) Zeigen Sie $\mathbb{R}^{\mathbb{N}} \sim \mathbb{R}$. *Hinweis:* Finden Sie eine injektive Abbildung von $\mathcal{P}(\mathbb{N})^{\mathbb{N}}$ nach $\mathcal{P}(\mathbb{N} \times \mathbb{N})$.

Eine berühmte Vermutung (die Kontinuumshypothese) sagt, dass folgende Frage die Antwort „ja“ hat.

Problem 8.23. *Gilt für jede unendliche Teilmenge A von $\mathbb{R}$: $A \sim \mathbb{R}$ oder $A \sim \mathbb{N}$?*

K. Gödel[3] zeigte, dass die Axiome der Zermelo-Fraenkelschen Mengenlehre, wenn widerspruchsfrei, auch unter Zuhilfenahme des Auswahlaxioms nicht erlauben, die Antwort „nein“ herzuleiten. P. Cohen[4] zeigte, dass die Axiome der Zermelo-Fraenkelschen Mengenlehre und das Auswahlaxiom, wenn widerspruchsfrei, nicht erlauben, die Antwort „ja“ herzuleiten. Die Gültigkeit von „$\forall A \subseteq \mathbb{R} : A \sim \mathbb{R}$ oder $A \sim \mathbb{N}$ oder A ist endlich“ wird also durch die Axiome der Zermelo-Fränkelschen Mengenlehre nicht geregelt.

[3]Kurt Gödel (1906-1978)
[4]Paul Cohen (1934-2007)

Teil 3

Vektorräume

KAPITEL 9

Algebraische Strukturen

1. Motivation

Die Ideen der Vektorrechnung im $\mathbb{R}^n$ lassen sich auch in anderen Bereichen anwenden, zum Beispiel in:

(1) $\mathbb{R}^{\mathbb{N}} = \{(a_1, a_2, \ldots) \mid \text{für alle } i \in \mathbb{N} \text{ gilt } a_i \in \mathbb{R}\}$, dem Raum aller *reellen Zahlenfolgen*.
(2) $\{0,1\}^n$, dem Raum aller Bitfolgen der Länge n,
(3) $\mathbb{Q}^n$, dem Raum aller Vektoren der Länge n mit rationalen Einträgen.

Daher ist es sinnvoll, die Theorie der Vektoren so allgemein zu entwicklen, dass sie diese Beispiele enthält.

In allen diesen Beispielen haben wir einen Bereich wie $\mathbb{R}$ oder $\mathbb{Q}$, in dem wir drei Grundrechnungsarten $(+, -, \cdot)$ und eine Division durch Elemente ungleich 0 ausführen können, den Bereich der *Skalare* und einen zweiten Bereich, in dem wir zwei Operationen $\boldsymbol{+}, \boldsymbol{-}$ ausführen können, die *Vektoren*. Außerdem können wir Vektoren *vervielfachen*, also mit einem Skalar multiplizieren.

2. Körper

Definition 9.1. Eine algebraische Struktur $\boldsymbol{R} = \langle R, +, -, \cdot, 0\rangle$ ist ein *Ring*, wenn $+, \cdot$ binäre Operationen auf R sind, $-$ eine unäre Operation auf R ist, und 0 ein Element aus R ist, sodass für alle $x, y, z \in R$ die folgenden Eigenschaften erfüllt sind:

(1) $x + 0 = x$ (0 ist rechtsneutral für $+$).
(2) $x + (-x) = 0$ ($-x$ ist additiv rechtsinvers zu x).
(3) $(x + y) + z = x + (y + z)$ ($+$ ist assoziativ).
(4) $x + y = y + x$ ($+$ ist kommutativ).
(5) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ ($\cdot$ ist assoziativ).
(6) $x \cdot (y + z) = x \cdot y + x \cdot z$ (Linksdistributivgesetz).
(7) $(x + y) \cdot z = x \cdot z + y \cdot z$ (Rechtsdistributivgesetz).

Satz 9.2. *Sei $\langle R, +, -, \cdot, 0\rangle$ ein Ring, und seien $x, y \in R$. Dann gilt*

(1) $-(-(x)) = x$
(2) $x \cdot 0 = 0 \cdot x = 0$.
(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): Es gilt $x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0$, also $0 = x \cdot 0 + (-(x \cdot 0)) = (x \cdot 0 + x \cdot 0) + (-(x \cdot 0)) = x \cdot 0 + (x \cdot 0 + (-(x \cdot 0))) = x \cdot 0 + 0 = x \cdot 0$. Die Identität $0 \cdot x = 0$ beweist man genauso. (3): Es gilt $(-x) \cdot y + x \cdot y = ((-x) + x) \cdot y = (x + (-x)) \cdot y = 0 \cdot y = 0$, also $(-x) \cdot y = -(x \cdot y)$. Die Identität $x \cdot (-y) = -(x \cdot y)$ beweist man genauso. $\square$

Beispiele für Ringe: $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}[i] = \{a + b\,i \mid a, b \in \mathbb{Z}\}, \mathbb{R}^{2\times 2}$.

DEFINITION 9.3. Sei $\boldsymbol{R} = \langle R, +, -, \cdot, 0 \rangle$ ein Ring.

(1) $e \in R$ ist ein *Einselement* von $\boldsymbol{R}$, wenn für alle $r \in R$ gilt, dass $e \cdot r = r \cdot e = r$. Wir bezeichnen dann die Struktur $\langle R, +, -, \cdot, 0, 1 \rangle$ mit $1 := e$ als *Ring mit Eins*.
(2) Ein Ring mit Eins $\boldsymbol{R}$ ist ein *Schiefkörper*, wenn $|R| \geq 2$ gilt und es für alle $x \in R$ mit $x \neq 0$ ein $y \in R$ mit $x \cdot y = y \cdot x = 1$ gibt.
(3) $\boldsymbol{R}$ ist *kommutativ*, wenn für alle $r, s \in R$ gilt: $r \cdot s = s \cdot r$.
(4) Ein *Körper* ist ein kommutativer Schiefkörper.
(5) Ein kommutativer Ring mit Eins $\boldsymbol{R}$ ist ein *Integritätsbereich*, wenn $|R| \geq 2$ und für alle $r, s \in R$ gilt: $r \cdot s = 0 \Rightarrow (r = 0 \vee s = 0)$.

Beispiele für Körper: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Auch 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. Die Operation $\oplus$ ist also das ausschließende Oder, und die Multiplikation $\odot$ die Operation $\sqcap$.

ÜBUNGSAUFGABEN 9.4.

(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.
(3) Ein Ring mit Eins R ist *reduziert*, wenn für alle $x \in R : x^2 = 0 \Rightarrow x = 0$ gilt. Zeigen Sie, dass in einem reduzierten Ring R für alle $n \in \mathbb{N}$ und $x \in R$ gilt : $x^n = 0 \Rightarrow x = 0$.

In einem Körper hat jedes Element $a \neq 0$ genau ein multiplikativ inverses Element; wir bezeichnen es mit a^{-1} oder $\frac{1}{a}$.

3. Vektorräume

DEFINITION 9.5. Sei K ein Körper. Ein Tupel $\langle V, \boldsymbol{+}, \boldsymbol{-}, \mathbf{0}, * \rangle$ heißt *Vektorraum* über K, wenn $V \neq \varnothing$, $\boldsymbol{+} : V \times V \to V$, $\boldsymbol{-} : V \to V$, $\mathbf{0} \in V$ 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) $1 * x = x$.

ÜBUNGSAUFGABEN 9.6. Sei in den folgenden Beispielen K ein Körper, und $\langle V, \boldsymbol{+}, \boldsymbol{-}, \mathbf{0}, * \rangle$ 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 $(\boldsymbol{-}x) = (-1) * x$.
(4) Sei $\alpha \in K, x \in V$ so, dass $\alpha * v = \mathbf{0}$. Zeigen Sie, dass $\alpha = 0$ oder $v = \mathbf{0}$.

Für einen Körper K und $n \in \mathbb{N}$ wird K^n durch $(x_1, \ldots, x_n) \boldsymbol{+} (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$, $\boldsymbol{-}(x_1, \ldots, x_n) := (-x_1, \ldots, -x_n)$, $\mathbf{0} = (0, \ldots, 0)$ und $\alpha * (x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n)$ für $\alpha, x_1, \ldots, x_n \in K$ ein Vektorraum. Wir schreiben für $\alpha * v$ meistens αv oder $\alpha \cdot v$, und verwenden das Symbol $*$ besonders dann, wenn $\cdot$ mit der Matrixmultiplikation oder der Multiplikation von zwei Körperelementen verwechselt werden könnte.

KAPITEL 10

Die Struktur von Vektorräumen

In diesem Kapitel bezeichnet K stets einen Körper.

1. Unterräume

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 10.1. Sei V ein Vektorraum über K, und sei U eine Teilmenge von V. U ist ein *Unterraum* von V, wenn

(1) $U \neq \varnothing$,
(2) für alle $u \in U$ und $\alpha \in K$ gilt $\alpha * u \in U$, und
(3) für alle $u, v \in U$ gilt $u + v \in U$.

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.

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

Beispiele 10.2.

(1) $T_1 = \{(x, y) \in \mathbb{R}^2 \mid 2\,x - 3\,y = 0\}$ 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 $2\,u - 3\,v = 0$ 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 $2 \cdot (u_1 + u_2) - 3 \cdot (v_1 + v_2) = 2\,u_1 - 3\,v_1 + 2\,u_2 - 3\,v_2 = 0$, 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) $T_2 = \{(x, y) \in \mathbb{R}^2 \mid 3\,x + 2\,y = 1\}$ 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 10.3.

(1) Vervollständigen Sie die folgenden Begründungen dafür, dass die Menge
$$T = \{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \mid \text{ es gibt } \alpha \in \mathbb{R}, \text{ sodass } \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \alpha \cdot \left(\begin{smallmatrix} -2 \\ 1 \end{smallmatrix}\right)\}$$
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 $t = \alpha \cdot \left(\begin{smallmatrix} -2 \\ 1 \end{smallmatrix}\right)$. 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) $\{\binom{x}{y} \in \mathbb{R}^2 \mid 3\,x + 2\,y = 0\}$.
 (b) $\{\binom{x}{y} \in \mathbb{R}^2 \mid 3\,x + 2\,y = 1\}$.
(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) $\{\binom{x}{y} \in \mathbb{R}^2 \mid x + 3\,y \leq 0\}$.
 (b) $\{\binom{x}{y} \in \mathbb{R}^2 \mid x^4 + y^2 = 0\}$.
(5) Zeigen Sie: Wenn ein Unterraum des $\mathbb{R}^n$ zwei Punkte enthält, so enthält er bereits die gesamte Verbindungsgerade.

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

SATZ 10.4. *Seien $m, n \in \mathbb{N}$, und sei $A \in K^{m \times n}$. Dann ist die Lösungsmenge des Gleichungssystems $A \cdot x = 0$ ein Unterraum des K^n.*

BEWEIS. Sei $U = \{x \in K^n \mid A \cdot x = 0\}$.

(1) Wegen $0 \in U$ ist U nicht die leere Menge.
(2) Sei $x \in U, \lambda \in K$. 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 K^n. □

BEISPIEL 10.5. 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

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

SATZ 10.6 (Unterraumkriterium). *Sei V ein Vektorraum über K und 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 $\alpha * u + \beta * v \in U$.*

2. Lineare Hülle von Vektoren

DEFINITION 10.7. Sei V ein Vektorraum über K, sei $m \in \mathbb{N}_0$, und seien $v_1, \ldots, v_m \in V$. Die Menge

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

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

$L(\binom{3}{2})$ ist also die Gerade im $\mathbb{R}^2$, die durch $\binom{0}{0}$ und $\binom{6}{4}$ geht. $L(\emptyset)$ definiert man als $\{0\}$. 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 10.8. *Seien $m, n \in \mathbb{N}$, seien $b_1, b_2, \ldots, b_m$ Vektoren im K^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 $x \in K^m$ hat.*

Übungsaufgaben 10.9.

(1) Liegt $\binom{2}{-1}$ in der linearen Hülle von $\binom{0}{0}$, $\binom{1}{2}$, $\binom{-2}{-4}$?
(2) Liegt $\binom{1}{-5}$ in der linearen Hülle von $\binom{1}{2}, \binom{1}{3}, \binom{1}{4}$?
(3) Testen Sie, ob $\begin{pmatrix} 7 \\ -2 \\ -2 \end{pmatrix}$ in der linearen Hülle von $\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$) liegt.

Wir definieren nun die lineare Hülle von *Teilmengen* eines K-Vektorraums V. Die lineare Hülle $L(\{v_1, v_2, v_3\})$ wird dann gleich $L(v_1, v_2, v_3)$ sein. Die neue Definition hat den Sinn, dass wir auch von der lineare Hülle von unendlichen Mengen von Vektoren sprechen können. Sei dazu M eine (möglicherweise unendliche) Teilmenge von V. Wir definieren die lineare Hülle von M als

$$L(M) = \{\sum_{i=1}^{k} \lambda_i \cdot m_i \,|\, k \in \mathbb{N}_0,\, \lambda_1, \ldots, \lambda_k \in K,\, m_1, \ldots, m_k \in M\}. \tag{10.1}$$

Die lineare Hülle von M ist also die Menge aller Linearkombinationen endlich vieler Vektoren aus M. Die lineare Hülle von M ist der kleinste Unterraum, der alle Vektoren aus M enthält. enthält:

Satz 10.10. *Sei V ein Vektorraum über K und sei M eine Teilmenge von V. Dann gilt:*

(1) *$L(M)$ ist ein Unterraum von V.*
(2) *Für jeden Unterraum U von V mit $M \subseteq U$ gilt $L(M) \subseteq U$.*

Beweis. (1) Wegen $0 \in L(M)$ (für $M = \varnothing$ muss man in (10.1) $k = 0$ setzen, die Summe über eine leere Menge von Summanden als 0 definieren und (10.1) als

$$L(M) = \{\sum_{i=1}^{k} \lambda(i) \cdot m(i) \,|\, k \in \mathbb{N}_0,\, \lambda \in K^{\{1,\ldots,k\}},\, m \in K^{\{1,\ldots,k\}}\}$$

lesen) gilt $L(M) \neq \varnothing$. Wenn $u = \sum_{i=1}^{k} \lambda_i m_i$ und $v = \sum_{i=1}^{l} \mu_i m_i'$ Elemente aus $L(M)$ sind, so ist $u + v = \sum_{i=1}^{k} \lambda_i m_i + \sum_{i=1}^{l} \mu_i m_i'$ wieder in $L(M)$, und für $\alpha \in K$ ist auch $\alpha * u = \sum_{i=1}^{k} (\alpha\lambda_i) m_i$ ein Element von $L(M)$.

(2) Sei $\sum_{i=1}^{k} \lambda_i m_i \in L(M)$. Wegen $M \subseteq U$ gilt für jedes $i \in \{1, \ldots, k\}$ auch $m_i \in U$ und, da U ein Unterraum ist, auch $\lambda_i m_i \in U$ und schließlich $\sum_{i=1}^{k} \lambda_i m_i \in U$. □

Korollar 10.11. *Sei M eine Teilmenge des K-Vektorraums V. Dann gilt $L(L(M)) = L(M)$.*

Beweis. Es gilt für jede Teilmenge A von V, dass $A \subseteq L(A)$. Für $A := L(M)$ erhalten wir $L(M) \subseteq L(L(M))$.

Aus Satz 10.10(1) wissen wir, dass $L(M)$ ein Unterraum von V mit $L(M) \subseteq L(M)$ ist. Wegen Satz 10.10(2) (für $\tilde{M} := L(M)$ und $\tilde{U} := L(M)$) gilt dann auch $L(L(M)) \subseteq L(M)$. □

Die lineare Hülle einer (möglicherweise unendlichen) Familie von Vektoren $(v_i)_{i\in I}$ definieren wir als $L(\{v_i \mid i \in I\})$. Es gilt also $w \in L((v_i)_{i\in I})$, wenn es eine Familie $(\lambda_i)_{i\in I}$ aus K gibt, sodass $\{i \in I \mid \lambda_i \neq 0\}$ endlich ist und $w = \sum_{\substack{i\in I\\ \lambda_i\neq 0}} \lambda_i v_i$.

Übungsaufgaben 10.12.

(1) Sei A eine Teilmenge des K-Vektorraums V. Zeigen Sie:
$$L(A) = \bigcap\{U \subseteq V \mid U \text{ ist ein Unterraum von } V,\ A \subseteq U\}.$$
(2) Welche Teilmengen von V werden durch folgende Gleichungen definiert?
$$\begin{array}{rcl} L_1(A) & = & \bigcap\{U \subseteq V \mid A \subseteq U\}. \\ L_2(A) & = & \bigcap\{U \subseteq V \mid U \text{ ist ein Unterraum von } V\}. \end{array}$$

Für jede Matrix A definieren wir drei Unterräume: den Zeilenraum, den Spaltenraum, und den Nullraum.

Definition 10.13. Für eine $m \times n$-Matrix A mit Einträgen aus K definieren wir ihren *Zeilenraum* $Z(A)$ als die lineare Hülle der Zeilen von A. Der Zeilenraum ist ein Unterraum von K^n.

Den *Spaltenraum* $S(A)$ definieren wir als die lineare Hülle der Spalten von A. Der Spaltenraum ist ein Unterraum von K^m.

Den *Nullraum* $N(A)$ definieren wir als die Lösungsmenge des Gleichungssystems $A \cdot x = 0$. Er ist ein Unterraum des K^n.

Für diese Räume sind auch andere Bezeichnungen gebräuchlich: $\mathrm{im}(A)$ für den Spaltenraum, $\ker(A)$ für den Nullraum und $\mathrm{coim}(A) = \mathrm{im}(A^T)$ für den Zeilenraum.

3. Lineare Unabhängigkeit von Vektoren

Definition 10.14. Sei V ein Vektorraum über K, seien $m, n \in \mathbb{N}$, und seien $v_1, \ldots, v_m$ in V. Die Folge $(v_1, \ldots, v_m)$ heißt *linear unabhängig*, wenn für alle $\lambda_1, \ldots, \lambda_m \in K$ 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 $v_1, \ldots, v_m$ linear unabhängig sind. Als Spezialfall definiert man noch für $m = 0$, dass die Folge () aus 0 Vektoren immer linear unabhängig ist.

Vektoren $v_1, \ldots, v_m$, die nicht linear unabhängig sind, nennt man *linear abhängig*. Die Folge $(v_1, \ldots, v_m)$ ist also genau dann linear abhängig, wenn es $(\lambda_1, \ldots, \lambda_m) \neq (0, \ldots, 0)$ gibt, sodass $\sum_{i=1}^{m} \lambda_i \cdot v_i = 0$.

Beispiel 10.15. Sind $(3, 2)$ und $(1, 3)$ aus dem $\mathbb{R}^2$ 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.

BEISPIEL 10.16. Sind $(3,2)$, $(1,4)$ und $(5,3)$ aus dem $\mathbb{R}^2$ 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 10.17. *Seien $m, n \in \mathbb{N}$, seien $b_1, b_2, \ldots, b_m$ Vektoren in K^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 sind $(b_1, b_2, \ldots, b_m)$ genau dann linear abhängig, wenn das System $\overline{B} \cdot x = 0$ eine Lösung $x \neq 0$ hat.*

Der folgende Satz liefert einen Zusammenhang zwischen der linearen Abhängigkeit und der Zugehörigkeit zu einer linearen Hülle von Vektoren.

SATZ 10.18. *Sei V ein Vektorraum, sei $m \geq 1$, und seien $b_1, \ldots, b_m \in V$. Die folgenden zwei Aussagen sind äquivalent:*

(1) *$(b_1, \ldots, b_m)$ ist linear abhängig.*
(2) *Es gibt ein $k \in \{1, \ldots, m\}$, sodass b_k in $L(b_1, \ldots, b_{k-1})$ liegt.*

BEWEIS. (2)⇒(1): Seien $\alpha_1, \ldots, \alpha_k \in K$ so, dass $b_k = \sum_{i=1}^{k-1} \alpha_i b_i$. Dann belegt die Gleichheit $(\sum_{i=1}^{k-1} \alpha_i b_i) + (-1) b_k = 0$, dass $(b_1, \ldots, b_m)$ linear abhängig ist.

(1)⇒(2): Sei $(\alpha_1, \ldots, \alpha_m) \in K^m \setminus \{(0, \ldots, 0)\}$ so, dass $\sum_{i=1}^m \alpha_i b_i = 0$. Sei $k := \max\{i \in \{1, \ldots, m\} \mid \alpha_i \neq 0\}$. Dann gilt $\sum_{i=1}^k \alpha_i b_i = 0$, und daher $b_k = -\sum_{i=1}^{k-1} \frac{\alpha_i}{\alpha_k} b_i$. Somit gilt $b_k \in L(b_1, \ldots, b_{k-1})$. □

In der linearen Hülle von n Vektoren hat jede linear unabhängige Folge höchstens Länge n:

SATZ 10.19. *Sei V ein Vektorraum, seien $m, n \in \mathbb{N}_0$ mit $m > n$, seien $b_1, \ldots, b_n \in V$, und seien $c_1, \ldots, c_m \in L(b_1, \ldots, b_n)$. Dann ist $(c_1, \ldots, c_m)$ linear abhängig.*

BEWEIS. Induktion nach n. Wenn $n = 0$, dann gilt wegen $m > n$, dass $m \geq 1$. Es gilt $c_1 \in L(\varnothing)$, also $c_1 = 0$. Wegen $1 c_1 = 0$ ist (c_1), und damit auch $(c_1, \ldots, c_m)$ linear abhängig.

Sei nun $n \geq 1$. Für jedes $i \in \{1, \ldots, m\}$ schreiben wir c_i als Linearkombination von $b_1, \ldots, b_n$; wir finden also $a_{i,1}, \ldots, a_{i,n} \in K$, sodass

$$c_i = a_{i,1} b_1 + \cdots + a_{i,n} b_n.$$

Falls alle $a_{i,1} = 0$, so gilt $\{c_1, \ldots, c_m\} \subseteq L(b_2, \ldots, b_n)$. Wegen $m > n - 1$ ist $(c_1, \ldots, c_m)$ nach Induktionsvoraussetzung linear abhängig.

Im Fall, dass es ein $k \in \{1, \ldots, m\}$ mit $a_{k,1} \neq 0$ gibt, bilden wir für $i \in \{1, \ldots, m\} \setminus \{k\}$ den Vektor

$$c_i' := -a_{k,1} c_i + a_{i,1} c_k.$$

Es gilt dann

$$c_i' = \sum_{j=1}^{n} (-a_{k,1}a_{i,j} + a_{i,1}a_{k,j})\, b_j.$$

Für $j = 1$ gilt $-a_{k,1}a_{i,j} + a_{i,1}a_{k,j} = -a_{k,1}a_{i,1} + a_{i,1}a_{k,1} = 0$, also gilt $c_i' \in L(b_2, \ldots, b_n)$. Nach Induktionsvoraussetzung ist $(c_1', \ldots, c_{k-1}', c_{k+1}', \ldots, c_m')$ linear abhängig. Also gibt es $(\alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_m) \in K^{m-1} \setminus \{(0, \ldots 0)\}$, sodass

$$\sum_{\substack{i=1\\ i\neq k}}^{m} \alpha_i c_i' = 0.$$

Sei $l \in \{1, \ldots, m\} \setminus \{k\}$ so, dass $\alpha_l \neq 0$. Dann gilt

$$0 = \sum_{\substack{i=1\\ i\neq k}}^{m} \alpha_i(-a_{k,1}c_i + a_{i,1}c_k) = -\sum_{\substack{i=1\\ i\neq k}}^{m} \alpha_i a_{k,1} c_i + (\sum_{\substack{i=1\\ i\neq k}}^{m} \alpha_i a_{i,1})c_k.$$

Da $\alpha_l a_{k,1} \neq 0$, belegt diese Linearkombination, dass $(c_1, \ldots, c_m)$ linear abhängig ist. □

Korollar 10.20. *Seien $m, n \in \mathbb{N}$ mit $m < n$, und sei $A \in K^{m\times n}$. Dann enthält der Nullraum von A einen von 0 verschiedenen Vektor.*

Beweis. Die Spaltenvektoren von A sind n Vektoren im $K^m = L(e_1, \ldots, e_m)$ und daher linear abhängig. Also gibt es $x \in K^n$ mit $x \neq 0$ und $A \cdot x = 0$. □

Wir definieren nun auch die lineare Abhängigkeit einer unendlichen *Familie* von Vektoren:

Definition 10.21. Sei V ein Vektorraum über K, und sei $(v_i)_{i\in I}$ eine Familie aus V. Die Familie $(v_i)_{i\in I}$ ist *linear abhängig*, wenn es eine Familie $(\lambda_i)_{i\in I}$ gibt, sodass

(1) $\{i \in I \mid \lambda_i \neq 0\}$ ist endlich,
(2) es gibt ein $i \in I$ mit $\lambda_i \neq 0$,
(3) $\displaystyle\sum_{\substack{i\in I\\ \lambda_i\neq 0}} v_i = 0.$

Eine *Teilmenge* $M \subseteq V$ ist linear abhängig, wenn es $\lambda : M \to K$ gibt, sodass $\lambda^{-1}[K\setminus\{0\}]$ endlich ist, es ein $m' \in V$ mit $\lambda(m') \neq 0$ gibt, und $\displaystyle\sum_{\substack{m\in M\\ \lambda(m)\neq 0}} \lambda(m)*m = 0$ gilt. Die Folge $((-1, 3), (-1, 3))$ ist linear abhängig, die Menge $\{(-1, 3), (-1, 3)\}$ ist gleich der Menge $\{(-1, 3)\}$ und somit linear unabhängig. Die leere Menge ist linear unabhängig.

Übungsaufgaben 10.22. In den folgenden Beispielen verwenden wir stets $\mathbb{R}$ als Körper.

(1) Zeigen Sie, dass die Vektoren $(\begin{pmatrix} 1\\4\\-2\end{pmatrix}, \begin{pmatrix} 0\\1\\2\end{pmatrix}, \begin{pmatrix} 3\\10\\-10\end{pmatrix})$ 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) $\binom{1}{2}, \binom{1}{3}, \binom{1}{4}$.
   (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 $(\begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 10 \\ -8 \end{pmatrix})$ 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

$$\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \underline{\hspace{1.5cm}} = 0.$$

Das ist eine Linearkombination, die 0 ergibt, obwohl nicht jeder Vektor ______________ mal genommen wurde. Daher sind (v_1, v_2, w) ____________________.

(8) Folgern Sie aus Satz 10.18 folgende Konsequenz:

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:

(a) $v \in L(v_1, \dots, v_m)$.

(b) $(v_1, \dots, v_m, v)$ ist linear abhängig.

4. Basen eines Vektorraums

Definition 10.23. Sei V ein Vektorraum über K. Eine Familie $B = (b_i)_{i \in I}$ ist eine *Basis* von V, wenn

(1) $(b_i)_{i \in I}$ linear unabhängig ist, und

(2) $L((b_i)_{i \in I}) = V$.

Wir geben einige Beispiele:

Beispiel 10.24.

(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 λ gibt, sodass $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

(3) $((2, 3))$ ist Basis von $L((2, 3))$.

Satz 10.25. *Sei $n \in \mathbb{N}$ und sei V ein Unterraum von K^n. Dann hat V eine Basis.*

Beweis. Wenn $V = \{0\}$, so fassen wir die leere Folge $() = \varnothing$ als Basis auf. Wenn $V \neq \{0\}$, so konstruieren wir eine Folge $(v_1, v_2, \dots)$ aus V rekursiv folgendermaßen: v_1 ist ein beliebiger Vektor aus $V \setminus \{0\}$, und wenn $v_1, \dots, v_i$ bereits definiert sind, so bilden wir die Menge $A_{i+1} := V \setminus L(v_1, \dots, v_i)$. Wenn A_{i+1} leer ist, definieren wir $v_{i+1} = 0$, ansonsten wählen wir als v_{i+1} irgendein Element in A_{i+1}. Für jedes $m \in \mathbb{N}_0$ gilt: wenn $v_{m+1} \neq 0$, so gilt $v_{m+1} \notin L(v_1, \dots, v_m)$. Wegen Satz 10.18 ist dann $(v_1, \dots, v_{m+1})$ linear unabhängig, also gilt $m + 1 \leq n$. Somit gilt

$v_{n+1} = 0$. Sei $k \in \mathbb{N}$ maximal mit $v_k \neq 0$. Dann gilt $v_{k+1} = 0$, also $L(v_1, \ldots, v_k) = V$. Da $(v_1, \ldots, v_k)$ linear unabhängig ist, ist $B = (v_1, \ldots, v_k)$ eine Basis von V. □

Jeder Unterraum des K^n ist also die lineare Hülle endlich vieler Vektoren.

Beispiel 10.26.

- ▷ Die Folge $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ ist eine Basis von $\mathbb{R}^3$.
- ▷ Für $i \in \mathbb{N}$ sei $e_i := (0, \ldots, 1, 0, \ldots)$ mit $e_i : \mathbb{N} \to \mathbb{R}$, $e_i(i) = 1$, $e_i(j) = 0$ für $j \neq i$. Dann ist $(e_1, e_2, \ldots)$ eine Basis für $\mathbb{R}^{(\mathbb{N})} := \{(a_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \exists j \in \mathbb{N}\, \forall i \in \mathbb{N} : i \geq j \Rightarrow a_i = 0\}$.
- ▷ Es gilt aber nicht $L((e_i)_{i\in\mathbb{N}}) = \mathbb{R}^{\mathbb{N}}$. Wie könnte eine Basis von $\mathbb{R}^{\mathbb{N}}$ aussehen?

Es zeigt sich, dass (unter Annahme des Auswahlaxioms) *jeder* Vektorraum eine Basis hat. Um die Existenz einer solchen Basis zu beweisen, braucht man einen Satz aus der Mengenlehre, das *Lemma von Zorn.*

Satz 10.27 (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* oder *Familien* von Vektoren, sondern auch einfach *Teilmengen* des Vektorraums zuzulassen; wir wiederholen die Definitionen für *lineare Unabhängigkeit*, *lineare Hülle* und *Basis*:

Definition 10.28. 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 10.29. *Sei V ein Vektorraum über K, 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 $\lambda : W \to K$, sodass es $w \in W$ mit $\lambda(w) \neq 0$ gibt, und

$$\sum_{w\in W} \lambda(w)\, w = 0. \tag{10.2}$$

Wenn $v \notin W$ oder $\lambda(v) = 0$, so erhalten wir aus (10.2), 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 10.30. *Sei V ein Vektorraum über K. Dann besitzt V eine Basis.*

Beweis: Sei $\mathcal{U} := \{B \mid B \subseteq V, B$ ist linear unabhängig$\}$. Nach Lemma 10.29 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 $L_1, L_2 \in \mathcal{K}$: $L_1 \subseteq L_2$ oder $L_2 \subseteq L_1$. Wir zeigen nun, dass $\mathcal{K}$ eine obere Schranke in $\mathcal{U}$ besitzt. Sei dazu $M := \bigcup \{L \mid L \in \mathcal{K}\}$, 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 $\lambda : W \to K$ 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. □

ÜBUNGSAUFGABEN 10.31.

(1) Sei V ein Vektorraum, und seien $B_1 \supseteq B_2 \supseteq B_3 \cdots$ Teilmengen von V mit $L(B_i) = V$ für alle $i \in \mathbb{N}$. Gilt dann sicher $L(\bigcap_{i \in \mathbb{N}} B_i) = V$?
(2) Sei V ein Vektorraum, und sei B eine Menge mit $L(B) = V$, sodass für jede Teilmenge C von B mit $C \neq B$ gilt, dass $L(C) \neq V$. Zeigen Sie, dass B eine Basis von V ist.

5. Dimension eines Vektorraums

Sei $B = (b_1, \ldots, b_n)$ eine Folge von Vektoren. Dann ist $|B| = |\{(i, b_i) : i \in \{1, \ldots, n\}\}| = n$.

SATZ 10.32. *Sei V ein Vektorraum, und seien B, C Basen von V. Wenn B endlich ist, so ist auch C endlich, und es gilt $|C| = |B|$.*

BEWEIS. Sei $B = (b_1, \ldots, b_n)$. Da die Vektoren in C liner unabhängig sind, gilt wegen Satz 10.19, dass C endlich ist und $|C| \leq n$. Sei $m := |C|$. Da $B \subseteq L(c_1, \ldots, c_m)$ gilt wegen Satz 10.19 auch $n \leq m$ und somit $|B| = |C|$ □

Dieser Satz gilt auch für unendliche Basen.

SATZ 10.33. *Sei V ein Vektorraum, und seien B, C Teilmengen von V, die beide Basen von V sind. Dann gibt es eine bijektive Abbildung von B nach C.*

Für unendliche B, C braucht man für den Beweis ein paar Sätze über die Mächtigkeit unendlicher Mengen. Diese Sätze beruhen darauf, dass man jeder Menge A eine Größe $|A|$ zuordnen kann. Diese Größe ist eine *Kardinalzahl*. Kardinalzahlen sind z.B. $0, 1, 2, \ldots$, dann $\omega = \aleph_0$ für die Anzahl der Elemente von $\mathbb{N}$ und c für die Anzahl der Elemente von $\mathbb{R}$. Für alle Mengen A_1, A_2 gilt $|A_1| = |A_2|$ genau dann, wenn es eine bijektive Abbildung von A_1 nach A_2 gibt. Satz 10.33 sagt dann, dass für Basen B, C eines Vektorraums V stets $|B| = |C|$ gilt.

Beweisskizze für Satz 10.33. Wenn B oder C endlich ist, so liefert Satz 10.32 das Ergebnis. Wir nehmen nun an, dass B, C beide unendlich sind. Für jedes $b \in B$ gibt es eine endliche Teilmenge $C(b)$ von C, sodass $b \in L(C(b))$. Sei $C' := \bigcup_{b \in B} C(b)$. Dann gilt $L(C') \subseteq L(B) = V$. Wenn nun $C' \neq C$, so gibt es ein $c \in C$, das in der linearen Hülle von $C \setminus \{c\}$ liegt. Dann ist C linear abhängig, im Widerspruch zur Annahme, dass C eine Basis ist. Also gilt $C = \bigcup_{b \in B} C(b)$. Eine Vereinigung von $|B|$ endlichen Mengen hat, wenn B unendlich ist, auch höchstens Kardinalität $|B|$, und somit gilt $|C| \leq |B|$. Genauso begründet man $|B| \leq |C|$ und wegen Satz 8.9 gilt daher $|B| = |C|$. □

Definition 10.34. Sei V ein Vektorraum, und sei B eine Basis von V. Die *Dimension* von V ist die Anzahl der Elemente von B.

Wegen Satz 10.30 besitzt V eine Basis, und wegen Satz 10.33 sind alle Basen gleichmächtig. Daher ist die Dimension sinnvoll definiert.

Satz 10.35. *Sei T ein endlichdimensionaler Vektorraum, und sei S ein Unterraum von T. Dann gilt:*

(1) $\dim(S) \leq \dim(T)$.
(2) *Wenn* $\dim(S) = \dim(T)$, *so gilt* $S = T$.

Beweis. Sei $(t_1, \ldots, t_n)$ eine Basis von T, und sei $(s_1, \ldots, s_k)$ eine Basis von S.

(1) Wegen Satz 10.19 gilt dann $k \leq n$.

(2) Wir nehmen an, dass $k = n$. Wenn $L(s_1, \ldots, s_k) = T$, so gilt $S = T$. Wenn es ein $t \in T \setminus L(s_1, \ldots, s_k)$ gibt, so ist nach Satz 10.18 die Folge $(s_1, \ldots, s_k, t)$ linear unabhängig. Dann haben wir $n + 1$ linear unabhängige Vektoren in einem Vektorraum mit einer n-elementigen Basis gefunden. Das widerspricht Satz 10.19. □

Übungsaufgaben 10.36.

(1) Beweisen Sie:
Seien $k \in \mathbb{N}_0$, und sei M eine Teilmenge des Vektorraums V. Sei
$$B = (b_1, b_2, \ldots, b_k)$$
eine linear unabhängige Folge von Vektoren aus M. Wir nehmen an, dass B so ist, dass es kein $m \in M$ gibt, sodass $(b_1, b_2, \ldots, b_k, m)$ 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 $M = (m_1, m_2, \ldots, m_k)$ eine Folge von Vektoren aus dem Vektorraum V. Dann gibt es $r \in \{0, \ldots, k\}$ und $i_1, i_2, \ldots, i_r \in \{1, \ldots, k\}$, sodass folgendes gilt: $i_1 < i_2 < \cdots < i_r$, und $(m_{i_1}, m_{i_2}, \ldots, m_{i_r})$ ist eine Basis von $L(M)$.
Hinweis: Wählen Sie eine linear unabhängige Teilfolge maximaler Länge.

(3) Zeigen Sie:
Sei $k \in \mathbb{N}_0$, $n \in \mathbb{N}$, und sei $M = (m_1, m_2, \ldots, m_k)$ eine Folge von Vektoren aus dem Vektorraum V. Dann gilt $\dim(L(M)) \leq k$.

6. Matrizen in Treppenform

In diesem Abschnitt überlegen wir uns, wie wir eine Basis eines Unterraums U von K^n bestimmen können, der in der Form $U = L(b_1, \ldots, b_m)$ gegeben ist. Wir wollen also zum Beispiel eine Basis von

$$V = L(\begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -5 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 11 \\ 18 \end{pmatrix}, \begin{pmatrix} -3 \\ 10 \\ 12 \end{pmatrix})$$

bestimmen. Wir wollen also eine Basis des Zeilenraums der Matrix

$$A = \begin{pmatrix} -1 & 3 & 2 \\ 2 & -5 & 2 \\ -3 & 11 & 18 \\ -3 & 10 & 12 \end{pmatrix}$$

Einen Algorithmus liefert der folgende Satz:

Satz 10.37. *Seien $m, n \in \mathbb{N}$, und seien*

$$\begin{aligned} b_1 &= (a_{1,1}, \ldots, a_{1,n}), \\ b_2 &= (a_{2,1}, \ldots, a_{2,n}), \\ & \ldots, \\ b_m &= (a_{m,1}, \ldots, a_{m,n}) \end{aligned}$$

Vektoren im K^n. Wir nehmen an, dass $a_{1,1} \neq 0$. Sei $V := L(b_1, \ldots, b_n)$ und $V_1 := V \cap (\{0\} \times K^{n-1})$. Dann gilt:

(1) $L(b_2 - \frac{a_{2,1}}{a_{1,1}} b_1, b_3 - \frac{a_{3,1}}{a_{1,1}} b_1, \ldots, b_m - \frac{a_{m,1}}{a_{1,1}} b_1) = V_1$.
(2) *Wenn $(d_2, \ldots, d_k)$ eine Basis von V_1 ist, so ist $(b_1, d_2, \ldots, d_k)$ eine Basis von V.*

Beweis. (1) Für $i \in \{2, \ldots, m\}$ sei $c_i := b_i - \frac{a_{i,1}}{a_{1,1}} b_1$. Für den ersten Eintrag von c_i gilt $c_i[1] = b_i[1] - \frac{a_{i,1}}{a_{1,1}} b_1[1] = a_{i,1} - \frac{a_{i,1}}{a_{1,1}} a_{1,1} = 0$, also gilt $c_i \in V_1$. Somit gilt auch $L(c_2, \ldots, c_m) \subseteq V_1$.

Wir beweisen nun $V_1 \subseteq L(c_2, \ldots, c_m)$: Sei dazu $v \in V_1$. Da $v \in L(b_1, \ldots, b_n)$, gibt es $\alpha_1, \ldots, \alpha_n \in K$, sodass

$$v = \sum_{i=1}^{n} \alpha_i b_i.$$

Wir wissen, dass $v \in V_1$, also gilt $v[1] = 0$. Also gilt

$$0 = \sum_{i=1}^{n} \alpha_i b_i[1] = \sum_{i=1}^{n} \alpha_i a_{i,1}. \tag{10.3}$$

Es gilt daher

$$\begin{aligned} v &= \sum_{i=1}^{n} \alpha_i b_i \\ &= \sum_{i=1}^{n} \alpha_i (b_i - \frac{a_{i,1}}{a_{1,1}} b_1) + (\sum_{i=1}^{n} \alpha_i \frac{a_{i,1}}{a_{1,1}})\, b_1 \\ &= \sum_{i=1}^{n} \alpha_i c_i + \frac{1}{a_{1,1}} (\sum_{i=1}^{n} \alpha_i a_{i,1})\, b_1. \end{aligned}$$

Wegen (10.3) ist der zweite Summand gleich 0, also gilt

$$v = \sum_{i=1}^{n} \alpha_i c_i,$$

und somit $v \in L(c_1, \ldots, c_m)$.

(2) Wir zeigen als erstes, dass $L(b_1, d_2, \ldots, d_k) = V$. Da $b_1, d_2, \ldots, d_k$ alle in V liegen, gilt $\subseteq$. Sei nun $v = (v_1, \ldots, v_n) \in V$. Dann gilt $v - \frac{v_1}{a_{1,1}} b_1 \in V_1$, und daher gibt es $\alpha_2, \ldots, \alpha_k \in K$ mit

$$v - \frac{v_1}{a_{1,1}} b_1 = \sum_{i=2}^{m} \alpha_i d_i.$$

Somit gilt $v = \frac{v_1}{a_{1,1}} b_1 + \sum_{i=2}^{m} \alpha_i d_i$, also $v \in L(b_1, d_2, \ldots, d_k) = V$.

Wir zeigen nun, dass $(b_1, d_2, \ldots, d_k)$ linear unabhängig ist. Seien $(\lambda_1, \ldots, \lambda_k) \in K^k$ so, dass $\lambda_1 b_1 + \sum_{i=2}^{k} \lambda_i d_i = 0$. Für alle $i \in \{2, \ldots, k\}$ gilt $d_i \in V_1$, also $d_i[1] = 0$, und somit gilt $\lambda_1 b_1[1] = 0$, also $\lambda_1 a_{1,1} = 0$. Wegen $a_{1,1} \neq 0$ gilt daher $\lambda_1 = 0$. Somit gilt $\sum_{i=2}^{k} \lambda_i d_i = 0$ und wegen der linearen Unabhängigkeit von $(d_2, \ldots, d_k)$ auch $\lambda_2 = \cdots = \lambda_k = 0$. $\square$

Nach Satz 10.37 können wir uns eine Basis von $Z(A)$ für

$$A = \begin{pmatrix} -1 & 3 & 2 \\ 2 & -5 & 2 \\ -3 & 11 & 18 \\ -3 & 10 & 12 \end{pmatrix}$$

berechnen, indem wir zunächst eine Basis für den Zeilenraum von

$$A_1 = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 2 & 12 \\ 0 & 1 & 6 \end{pmatrix}$$

ausrechnen. Wir lassen die erste Nullspalte weg und berechnen eine Basis für den Zeilenraum von

$$A_2 = \begin{pmatrix} 1 & 6 \\ 2 & 12 \\ 1 & 6 \end{pmatrix}.$$

Nach Satz 10.37 können wir uns eine eine Basis des Zeilenraums von A_2 ausrechnen, indem wir zunächst eine Basis des Zeilenraums von

$$A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ausrechnen. Die leere Folge $\varnothing$ ist eine solche Basis, also ist nach Satz 10.37 $((1, 6))$ eine Basis von $Z(A_2)$, und somit $((0, 1, 6))$ eine Basis von $Z(A_1)$. Wegen Satz 10.37 ist $((-1, 3, 2), (0, 1, 6))$ eine Basis von $Z(A)$.

Diese Vorgangsweise kann man (iterativ anstatt rekursiv) so sehen: 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[19]:= A = {{-1,3,2},{2,-5,2},{-3,11,18},{-3,10,12}};

In[20]:= << RowRed12.m

In[21]:= RowEchelonForm [A]

(-1    3    2
  2   -5    2
 -3   11   18
```

```
 -3   10   12
)

(-1   3   2
  0   1   6
  0   2   12
  0   1   6
)

(-1   3   2
  0   1   6
  0   0   0
  0   0   0
)

Out[21]= {{{1,0,0,0},{2,1,0,0},{-7,-2,1,0},{-5,-1,0,1}},
   {{-1,3,2},{0,1,6},{0,0,0},{0,0,0}}}
```

Wir haben also eine Matrix in Treppenform erzeugt, deren Zeilenraum der gleiche wie der Zeilenraum der gegebenen Matrix ist. Somit ist $\left(\begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 6 \end{pmatrix}\right)$ eine Basis für $Z(A)$, also für V.

ALGORITHMUS 10.38 (Treppenform).
Eingabe: Eine $m \times n$-Matrix A.
Ausgabe: Eine $m \times n$-Matrix B, sodass B in Treppenform ist, und $Z(A) = Z(B)$.

$B \leftarrow A$
zeile $\leftarrow 1$
spalte $\leftarrow 1$
while *zeile* $\leq m$ **do**
    while *spalte* $\leq n$ und $B(i, \mathit{spalte}) = 0$ für alle i mit *zeile* $\leq i \leq m$ **do**
        spalte $\leftarrow$ *spalte* $+ 1$
    if *spalte* $\leq n$ **then**
        gewaehlteZeile $\leftarrow$ ein i, sodass *zeile* $\leq i \leq m$ und $B(i, \mathit{spalte}) \neq 0$
        if *gewaehlteZeile* $\neq$ *zeile* **then**
            Vertausche die *gewaehlteZeile*-te mit der *zeile*-ten Zeile von B
        $i \leftarrow$ *zeile* $+ 1$
        while $i \leq m$ **do**
            Addiere passendes Vielfaches der *zeile*-ten Zeile zur i-ten Zeile von B, sodass $B(i, \mathit{spalte}) = 0$
            $i \leftarrow i + 1$
    zeile $\leftarrow$ *zeile* $+ 1$
    spalte $\leftarrow$ *spalte* $+ 1$
return B

Dieser Algorithmus liefert also folgenden Satz.

SATZ 10.39. *Seien $m, n \in \mathbb{N}$, und sei A eine $m \times n$-Matrix. Dann gibt es eine Matrix B in Treppenform, sodass $Z(A) = Z(B)$.*

Die Zeilen einer Matrix in Treppenform, die nicht 0 sind, sind linear unabhängig.

Wir fassen zusammen:

Algorithmus 10.40 (Basis eines explizit gegebenen Unterraums).
Eingabe: Vektoren $v_1, \dots, v_m \in \mathbb{R}^n$.
Ausgabe: Eine Basis $b_1, \dots, b_k$ von $L(v_1, \dots, v_m)$.

Bilde die $m \times n$-Matrix V, in deren Zeilen die Vektoren $v_1, \dots, v_m$ stehen.
Berechne eine Matrix B in Treppenform, sodass $Z(V) = Z(B)$.
return $(b_1, \dots, b_k)$ als jene Zeilen von B, die nicht 0 sind.

Was wir noch nicht begründet haben ist, dass die Zeilen einer Matrix in Treppenform, 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. Allgemein gilt:

Satz 10.41. *Sei A eine $m \times n$-Matrix in Treppenform. Dann sind die Zeilen von A, die nicht 0 sind, linear unabhängige Vektoren im $\mathbb{R}^n$.*

Übungsaufgaben 10.42.

(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) $\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + 3\,y + z = 0\}$.
  (b) $L(\begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix})$.
  (c) $L(\binom{1}{2}, \binom{2}{4}, \binom{-2}{4})$.

Eine Treppenform mit zusätzlichen Eigenschaften ist die *Treppennormalform*. Sie löst 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 $V = L(v_1, \dots, v_m)$. Gefragt ist, ob $U = V$.

Definition 10.43 (Treppennormalform). Sei A eine $m \times n$-Matrix. A ist in *Treppennormalform*, 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 10.44 (Treppennormalform).
Eingabe: Eine $m \times n$-Matrix A.
Ausgabe: Eine $m \times n$-Matrix B, sodass B in Treppennormalform ist, und $Z(A) = Z(B)$.

$B \leftarrow A$
zeile $\leftarrow 1$
spalte $\leftarrow 1$
while *zeile* $\leq m$ **do**
  (* Die ersten *spalte* $- 1$ Spalten sind in Treppennormalform *)
  while *spalte* $\leq n$ und $B(i, \mathit{spalte}) = 0$ für alle i mit *zeile* $\leq i \leq m$ **do**
    spalte $\leftarrow$ *spalte* $+ 1$
  if *spalte* $\leq n$ **then**
    gewaehlteZeile $\leftarrow$ ein i, sodass *zeile* $\leq i \leq m$ und $B(i, \mathit{spalte}) \neq 0$
    if *gewaehlteZeile* $\neq$ *zeile* **then**
      Vertausche die *gewaehlteZeile*-te mit der *zeile*-ten Zeile von B
    Multipliziere die *zeile*-te Zeile von B mit $\frac{1}{B(\mathit{zeile},\mathit{spalte})}$
    $i \leftarrow 1$
    while $i \leq m$ **do**
      if $i \neq$ *zeile* **then**
        Addiere passendes Vielfaches der *zeile*-ten Zeile zur i-ten Zeile von B, sodass $B(i, \mathit{spalte}) = 0$
      $i \leftarrow i + 1$
  zeile $\leftarrow$ *zeile* $+ 1$
  spalte $\leftarrow$ *spalte* $+ 1$
return B

Wir geben ein Beispiel für das Berechnen einer Matrix in Treppennormalform, die den gleichen Zeilenraum wie die Ausgangsmatrix besitzt.

Beispiel 10.45.

```
In[24]:= << RowRed12.m

In[25]:= A = {{1,-5,8,2,-2},{1,-4,6,-2,0},{-1,0,2,2,0},{5,-8,6,0,-5}}
Out[25]= {{1,-5,8,2,-2},{1,-4,6,-2,0},{-1,0,2,2,0},{5,-8,6,0,-5}}

In[26]:= RowEchelonNormalForm [A]
 (1  -5   8   2   -2
  1  -4   6  -2    0
 -1   0   2   2    0
  5  -8   6   0   -5
)

(1   -5   8     2  -2
 0    1  -2    -4   2
 0   -5   10    4  -2
 0   17  -34  -10   5
)

(1   0  -2  -18   8
 0   1  -2   -4   2
 0   0   0  -16   8
```

```
 0    0    0    58 -29
 )

(1    0   -2   -18    8
 0    1   -2   -4     2
 0    0    0    1   -(1/2)
 0    0    0    58   -29
)

(1    0   -2    0   -1
 0    1   -2    0    0
 0    0    0    1   -(1/2)
 0    0    0    0    0
 )

Out[26]= {{{1/2,-(5/8),-(9/8),0},{0,-(1/4),-(1/4),0},
          {1/4,-(5/16),-(1/16),0},{-(5/2),9/8,29/8,1}},
         {{1,0,-2,0,-1},{0,1,-2,0,0},{0,0,0,1,-(1/2)},{0,0,0,0,0}}}
```

Der Algorithmus liefert den folgenden Satz.

Satz 10.46. *Seien $m, n \in \mathbb{N}$, und sei $A \in \mathbb{R}^{m\times n}$. Dann gibt es eine Matrix B in Treppennormalform, sodass $Z(A) = Z(B)$.*

Die Treppennormalform ist eindeutig:

Satz 10.47. *Seien $m, n \in \mathbb{N}$, und seien B, C zwei $m \times n$-Matrizen in Treppennormalform. Wir nehmen an, dass $Z(B) = Z(C)$. Dann gilt $B = C$.*

Beweisskizze. Induktion nach n. Wenn $n = 1$, so ist $Z(B) = K^1$ oder $Z(B) = \{(0)\}$. Im ersten Fall gilt $B = C = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, im zweiten Fall $B = C = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

Sei nun $n \geq 2$. Sei $\pi : K^n \to K^{n-1}$, $\pi((v_1, \ldots, v_n)) := (v_1, \ldots, v_{n-1})$, und sei B' die Matrix B ohne deren letzte Spalte; es gilt dann $B' \in K^{m\times(n-1)}$. Der Zeilenraum von B' ist $\pi(Z(B))$. Genauso bilden wir aus C die Matrix C'; ihr Zeilenraum ist $\pi(Z(C))$.

Aus der Definition der Treppennormalform sieht man, dass B' und C' wieder in Treppennormalform sind. Wegen $Z(B') = \pi(Z(B)) = \pi(Z(C)) = Z(C')$ und der Induktionsvoraussetzung gilt $B' = C'$.

Wir überprüfen nun, ob auch die letzte Spalte der Matrix gleich ist. Wir betrachten zuerst den Fall, dass $e_n = (0, \ldots, 0, 1) \in Z(B)$. Da B in Treppennormalform ist, muss dann $(0, \ldots, 0, 1)$ eine Zeile von B sein. Wenn B' genau s Nichtnullzeilen hat, ist also e_n die $s + 1$-te Zeile von B, und alle Einträge in der n-ten Spalte von B in den anderen Zeilen sind gleich 0. Wegen $e_n \in Z(C)$ gilt das gleiche für C, und somit ist auch die n-te Spalte beider Matrizen gleich e_{s+1}; also gilt $B = C$.

Wenn $(0, \ldots, 0, 1) \notin Z(B)$, so wählen wir $k \in \{1, \ldots, m\}$ und betrachten die k-te Zeile b_k von B und die k-te Zeile c_k von C. Da diese beiden Zeilen nach Induktionsvoraussetzung in den ersten $n - 1$ Einträgen übereinstimmen, gilt $b_k - c_k \in Z(B) \cap L(e_n)$. Da $L(e_n) \not\subseteq Z(B)$, gilt $Z(B) \cap L(e_n) = \{0\}$. Also gilt $b_k = c_k$. $\square$

Mathematica hat einen eigenen Befehl zur Berechnung der Treppennormalform, `RowReduce`.

```
In[27]:= A = {{1,-5,8,2,-2},{1,-4,6,-2,0},{-1,0,2,2,0},{5,-8,6,0,-5}}
Out[27]= {{1,-5,8,2,-2},{1,-4,6,-2,0},{-1,0,2,2,0},{5,-8,6,0,-5}}

In[28]:= RowReduce[A]
Out[28]= {{1,0,-2,0,-1},{0,1,-2,0,0},{0,0,0,1,-(1/2)},{0,0,0,0,0}}
```

7. Nullraum einer Matrix

Sei $A \in K^{m\times n}$, und seien $a_1, \ldots, a_m \in K^n$ die Zeilen von A. Für $u = (u_1, \ldots, u_n)$ und $v = (v_1, \ldots, v_n) \in K^n$ definieren wir

$$\langle u, v\rangle := \sum_{i=1}^{n} u_i v_i.$$

In $\mathbb{R}$ erhalten wir unser gewohntes Skalarprodukt. In anderen Körpern hat dieses Produkt die ungewöhnliche Eigenschaft, dass etwa in $\mathbb{Z}_2$ gilt $\langle \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\rangle = 0$, oder in $\mathbb{C}$ gilt, dass $\langle \left(\begin{smallmatrix}1\\i\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\i\end{smallmatrix}\right)\rangle = 0$. Man betrachtet daher in $\mathbb{C}$ oft auch andere Skalarprodukte

Definition 10.48. Sei M eine Teilmenge des K^n. Dann definieren wir

$$M^\wedge := \{v \in \mathbb{R}^n \mid \langle m, v\rangle = 0 \text{ für alle } m \in M\}.$$

Übungsaufgaben 10.49.

(1) Seien $M, U, V \subseteq K^n$. Zeigen Sie: $M^\wedge = L(M)^\wedge$, $U \subseteq V \Rightarrow V^\wedge \subseteq U^\wedge$, $L(M) \subseteq (M^\wedge)^\wedge$, $U \subseteq (U^\wedge)^\wedge$, $U^\wedge = ((U^\wedge)^\wedge)^\wedge$.

Als Folgerung aus $M^\wedge = L(M)^\wedge$ ergibt sich:

Satz 10.50. *Sei A eine $m \times n$-Matrix. Dann gilt $N(A) = (Z(A))^\wedge$.*

Somit gilt

Korollar 10.51. *Sei A eine $l \times n$, und sei B eine $m \times n$-Matrix. Wenn $Z(A) = Z(B)$, dann gilt auch $N(A) = N(B)$.*

Das liefert eine Vorgangsweise für die Berechnung des Nullraums von A: Man bestimmt eine Matrix B in Treppenform oder Treppennormalform und berechnet dann den Nullraum dieser Matrix. Wenn eine Matrix B in Treppennormalform ist, dann kann man ihren Nullraum $N(B)$ besonders schnell bestimmen.

Wir bestimmen so den Nullraum der folgenden Matrix $A \in \mathbb{R}^{4\times 5}$.

$$A = \begin{pmatrix} 1 & -9 & 2 & 7 & 2 \\ 2 & -6 & 0 & 2 & 0 \\ 3 & 3 & -4 & -9 & -4 \\ 1 & 3 & -5 & -11 & -5 \end{pmatrix}$$

Wir berechnen den Nullraum $N(A) = \{x \in \mathbb{R}^5 \mid A \cdot x = 0\}$ mit dem Gauß-Algorithmus:

```
In[36]:= << RowRed12.m

In[37]:= A = {{1,-9,2,7,2},{2,-6,0,2,0},{3,3,-4,-9,-4},{1,3,-5,-11,-5}}
Out[37]= {{1,-9,2,7,2},{2,-6,0,2,0},{3,3,-4,-9,-4},{1,3,-5,-11,-5}}
```

```
In[38]:= Gauss [A, {0,0,0,0}]

(x1   x2  x3  x4  x5
  1  -9   2   7   2    |    0
  2  -6   0   2   0    |    0
  3   3  -4  -9  -4    |    0
  1   3  -5  -11 -5    |    0

)
 We use equation (1) of the last system to eliminate x1
 (x2   x3   x4   x5    |
  12   -4  -12   -4    |    0
  30  -10  -30  -10    |    0
  12   -7  -18   -7    |    0

)
 We use equation (1) of the last system to eliminate x2
 (x3   x4   x5  |
   0   0   0    |    0
  -3  -6  -3    |    0

)
 We use equation (2) of the last system to eliminate x3
 (x4   x5    |
  0     0    |    0

)
 x4 does not appear in any equation.
 (x5    |
   0    |    0

)
 x5 does not appear in any equation.
 x5:=t1
 x4:=t2
 We use
 (x3   x4   x5    |
  -3  -6   -3     |    0

)
 to compute x3
 x3 = 0 + -t1 -2 t2
 We use
 (x2   x3   x4   x5    |
  12  -4  -12    -4    |    0

)
 to compute x2
 x2 = 0 + t2/3
 We use
 (x1   x2   x3   x4   x5    |
```

```
   1  -9    2    7    2      | 0

)
 to compute x1
 x1 = 0 + 0

Out[38]= {{0,0,0,0,0},{{0,0,-1,0,1},{0,1/3,-2,1,0}}}

In[39]:= NullSpace [A]
Out[39]= {{0,0,-1,0,1},{0,1,-6,3,0}}
```

Somit ist der Nullraum von A die lineare Hülle von $(0, 0, -1, 0, 1)$ und $(0, 1, -6, 3, 0)$.

Den Nullraum kann man auch mithilfe der Treppenform ausrechnen: Sei B die Matrix, die man durch Nebeneinanderschreiben der Matrix A^T mit der Einheitsmatrix erhält, also

$$B = (A^T|E_5) = \begin{pmatrix} 1 & 2 & 3 & 1 & 1 & 0 & 0 & 0 & 0 \\ -9 & -6 & 3 & 3 & 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & -4 & -5 & 0 & 0 & 1 & 0 & 0 \\ 7 & 2 & -9 & -11 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & -4 & -5 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Wir berechen eine Treppenform von B mit `RowReduce[B]` und erhalten die Matrix C mit

$$C = \begin{pmatrix} 1 & 0 & -2 & 0 & -\frac{5}{6} & 0 & 0 & \frac{5}{6} & -2 \\ 0 & 1 & \frac{5}{2} & 0 & \frac{13}{12} & 0 & 0 & -\frac{7}{12} & \frac{3}{2} \\ 0 & 0 & 0 & 1 & -\frac{1}{3} & 0 & 0 & \frac{1}{3} & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix}.$$

Dann findet man in den Zeilen, in denen die Einträge in den ersten vier Spalten 0 sind, die Matrix

$$D := \begin{pmatrix} 0 & 1 & 0 & 3 & -6 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix},$$

in deren Zeilen eine Basis für den Nullraum von A steht. Wir rechtfertigen jetzt diese Vorgangsweise.

Definition 10.52. Seien $m, n \in \mathbb{N}$ und sei $B = \{b_1, \ldots, b_m\}$ eine Basis des Unterraums U von K^n. Die Basis B ist in *Treppenform*, wenn für alle $i \in \{0, \ldots, n-1\}$ gilt, dass

$$L(B \cap (\{0\}^i \times K^{n-i})) = U \cap (\{0\}^i \times K^{n-i}).$$

Die Nichtnullzeilen einer Matrix in Treppenform bilden eine Basis in Treppenform für den Zeilenraum dieser Matrix.

Satz 10.53. *Seien $m, n \in K$, $A \in K^{m\times n}$, und sei B die $n \times (n + m)$-Matrix, die man durch Nebeneinanderschreiben der Matrix A^T mit der Einheitsmatrix erhält, also $B = (A^T|E_n)$. Dann gilt*

$$Z(B) \cap (\{0\}^m \times K^n) = \{0\}^m \times N(A).$$

Beweis. Sei $y \in Z(B) \cap (\{0\}^m \times K^n)$. Dann gibt es ein $x \in K^n$ (den wir als Spaltenvektor schreiben), sodass $y = x^T \cdot (A^T|E_n) = (x^T \cdot A^T|x^T)$. Da $y \in \{0\}^m \times K^n$, gilt $x^T \cdot A^T = 0$, also $A \cdot x = 0$ und somit $x \in N(A)$. Somit gilt $y = (x^T \cdot A^T|x^T) \in \{0\}^m \times N(A)$.

Sei nun $y \in \{0\}^m \times N(A)$ und sei $x \in N(A)$ so, dass $y = (0, \ldots, 0, x_1, \ldots, x_n)$. Dann gilt $x^T \cdot (A^T|E_n) = (x^T A^T|x^T) = (0|x^T) = y$. Somit gilt $y \in Z(B)$. $\square$

Um den Nullraum von A auszurechnen, können wir also $Z(B) \cap (\{0\}^m \times K^n)$ bestimmen. Eine Basis für $Z(B) \cap (\{0\}^m \times K^n)$ können wir aus einer Matrix B' in Treppenform mit $Z(B') = Z(B)$ bestimmen.

Mithilfe der Treppennormalform können wir auch die inverse Matrix berechnen:

Satz 10.54. *Sei $A \in K^{n\times n}$, sei B die $n \times (2n)$-Matrix $(A|E_n)$ und seien $X, Y \in K^{n\times n}$ so, dass $(X|Y)$ eine Matrix in Treppennormalform mit $Z(B) = Z((X|Y))$ ist. Dann gilt:*

(1) *Wenn $X = E_n$, so ist A invertierbar und es gilt $A^{-1} = Y$.*
(2) *Wenn $X \neq E_n$, so ist A nicht invertierbar.*

Beweis.

(1) Die Treppennormalform hat dann die Form $(E_n|Y)$. Da $Z((A|E_n)) = Z((E_n|Y))$ gibt es Matrizen $X_1, X_2 \in K^{n\times n}$ mit $X_1 \cdot (A|E_n) = (E_n|Y)$ und $X_2 \cdot (E_n|Y) = (A|E_n)$. Folglich gilt $X_1 A = E_n$, $X_1 = Y$, $X_2 = A$, $X_2 Y = E_n$ und somit $YA = E_n$ und $AY = E_n$, also ist A invertierbar mit inverser Matrix Y.
(2) Wenn A invertierbar ist, so gibt es eine Matrix $X \in K^{n\times n}$ mit $XA = E_n$. Sei $\pi : K^{2n} \to K^n$, $\pi((x|y)) = x$. Wegen $XA = E_n$ gilt $Z(E_n) \subseteq Z(A)$, also $K^n = Z(A)$. Somit gilt
$$\pi[Z((A|E_n))] = Z(A) = K^n.$$
Wir nehmen nun, dass für die Treppennormalform $(X|Y)$ von $(A|E_n)$ gilt, dass $X \neq E_n$. Dann enthält X eine Nullzeile. Seien $x_1, \ldots, x_{n-1}$ die ersten $n-1$ Zeilen von X. Es gilt dann $Z(X) \subseteq L(x_1, \ldots, x_{n-1})$ und außerdem $Z(X) = \pi[Z((X|Y))] = \pi[Z((A|E_n))] = Z(A) = K^n$. Somit gilt $K^n \subseteq L(x_1, \ldots, x_{n-1})$, und das widerspricht Satz 10.19. Also gilt $X = E_n$.

□

Dieser Beweis liefert auch eine Begründung dafür, dass eine $n \times n$ Matrix A, für die es ein X mit $XA = E_n$ gibt, invertierbar ist. In Satz 11.27 bekommen wir dieses Resultat dann, ohne über Treppenformen zu argumentieren zu müssen.

8. Rang einer Matrix

Definition 10.55. Sei A eine $m \times n$-Matrix. Der *Rang* von A, abgekürzt $\mathrm{rk}(A)$, ist die Dimension des Zeilenraums von A.

Man berechnet den Rang, indem man aus der Matrix A eine Matrix B in Treppenform 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 10.56 (Rang einer Matrix).
Eingabe: eine $m \times n$-Matrix A.
Ausgabe: der Rang von A.

Berechne mit Algorithmus 10.38 eine Matrix B in Treppenform, sodass $Z(B) = Z(A)$.
return Anzahl der Zeilen von B, die nicht 0 sind.

Der folgende Satz besagt, dass A und A^T stets den gleichen Rang haben.

Satz 10.57. *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 sei $(b_1, \ldots, b_r)$ eine Basis von $S(A)$. Sei

$$B = \begin{pmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_r \\ | & | & & | \end{pmatrix}$$

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 $b_1, \ldots, b_r$ ausdrücken lässt, gibt es eine Matrix $C \in K^{r \times n}$, sodass

$$A = B \cdot C. \tag{10.4}$$

Aus der Gleichung (10.4) 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 $\dim(Z(C)) \leq r$, und somit $\dim(Z(A)) \leq r$. Insgesamt gilt also für die Matrix A, dass $\dim(Z(A)) \leq \dim(S(A))$.

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 $\dim(Z(A^T)) \leq \dim(S(A^T))$, und somit $\dim(S(A)) \leq \dim(Z(A))$. ☐

Satz 10.58 (Rangsatz). *Seien $m, n \in \mathbb{N}$, $A \in K^{m \times n}$, und sei k der Rang von A. Dann hat $N(A)$ die Dimension $n - k$.*

Beweis.Wegen Satz 10.57 ist auch die Dimension des Spaltenraums von A gleich k, und $S(A)$ besitzt daher eine Basis $(b_1, \ldots, b_k)$. Sei $(c_1, \ldots, c_l)$ eine Basis für den Nullraum von A. Wir finden für jedes $i \in \{1, \ldots, k\}$ einen Vektor $y_i \in K^n$ mit $A \cdot y_i = b_i$. Wir behaupten nun, dass

$$D = (c_1, \ldots, c_l, y_1, \ldots, y_k)$$

eine Basis von K^n ist. Wir zeigen als erstes, dass D linear unabhängig ist. Seien dazu $\lambda_1, \ldots, \lambda_l, \mu_1, \ldots, \mu_k \in K$ so, dass

$$\lambda_1 c_1 + \cdots + \lambda_l c_l + \mu_1 y_1 + \cdots + \mu_k y_k = 0. \tag{10.5}$$

Durch Multiplikation von links mit A erhalten wir

$$\lambda_1(A \cdot c_1) + \cdots + \lambda_l(A \cdot c_l) + \mu_1(A \cdot y_1) + \cdots + \mu_k(A \cdot y_k) = 0,$$

und da $A \cdot c_i = 0$ und $A \cdot y_i = b_i$ auch

$$0 + \mu_1 b_1 + \cdots + \mu_k b_k = 0.$$

Wegen der linearen Unabhängigkeit von $(b_1, \ldots, b_k)$ gilt also $\mu_1 = \cdots = \mu_k = 0$. Also gilt wegen (10.5) auch

$$\lambda_1 c_1 + \cdots + \lambda_l c_l = 0,$$

und wegen der Unabhängigkeit von $(c_1, \ldots, c_l)$ auch $\lambda_1 = \cdots = \lambda_l = 0$. Somit ist D linear unabhängig.

Wir zeigen nun $K^n \subseteq L(D)$. Da $A \cdot x \in S(A)$, gibt es $\mu_1, \ldots, \mu_k \in K$ mit

$$A \cdot x = \sum_{i=1}^{k} \mu_i b_i.$$

Sei $z := \sum_{i=1}^{k} \mu_i y_i$. Dann gilt $A \cdot z = \sum_{i=1}^{k} \mu_i (A \cdot y_i) = \sum_{i=1}^{k} \mu_i b_i = A \cdot x$, also $A \cdot (x - z) = 0$. Folglich gilt $x - z \in N(A)$ und es gibt daher $\lambda_1, \ldots, \lambda_l \in K$ mit

$$x - z = \sum_{i=1}^{l} \lambda_i c_i.$$

Folglich gilt

$$x = \sum_{i=1}^{l} \lambda_i c_i + \sum_{i=1}^{k} \mu_i y_i,$$

und daher $x \in L(D)$.

Folglich ist D eine Basis von K^n. Da K^n die Basis $(e_1, \dots, e_n)$ hat, gilt $\dim(K^n) = n$ und daher $n = k + l$ und somit $l = n - k$. □

9. Die Lösungsmenge inhomogener linearer Gleichungssysteme

Definition 10.59. Sei $A \in K^{m \times n}$, und sei $b \in K^m$. Das Gleichungssystem $A \cdot x = b$ heißt *homogen*, wenn $b = 0$, und *inhomogen*, wenn $b \neq 0$.

Definition 10.60. Sei V ein K-Vektorraum, und sei $M \subseteq V$. Die Menge M ist eine *lineare Mannigfaltigkeit* in V (oder ein *affiner Unterraum* von V), wenn für alle $\lambda \in K$ und $a, b, c \in M$ gilt:

(1) $\lambda a + (1 - \lambda) b \in M$,
(2) $a - b + c \in M$.

Satz 10.61. *Sei V ein K-Vektorraum und M ein affiner Unterraum von V mit $M \neq \emptyset$. Dann gibt es einen Unterraum U von V und ein $t_0 \in M$, sodass $M = \{t_0 + u \mid u \in U\}$.*

Für $\{t_0 + u \mid u \in U\}$ schreiben wir auch $t_0 + U$.

Beweis. Sei t_0 ein Element von M. Wir definieren

$$U := \{m - t_0 \mid m \in M\}.$$

Wir zeigen nun, dass U ein Unterraum von V ist. Wegen $0 \in U$ gilt $U \neq 0$. Seien nun $u, v \in U$ und $\alpha, \beta \in K$. Dann gibt es $x, y \in V$ mit $u = x - t_0$ und $v = y - t_0$. Es gilt nun

$$(\alpha x + (1 - \alpha) t_0) - t_0 + (\beta y + (1 - \beta) t_0) \in M,$$

also

$$\alpha x + (1 - \alpha) t_0 - t_0 + \beta y + (1 - \beta) t_0 - t_0 \in U.$$

Durch Ausmultiplizieren erhalten wir

$$\alpha x + t_0 - \alpha t_0 - t_0 + \beta y + t_0 - \beta t_0 - t_0 = \alpha x - \alpha t_0 + \beta y - \beta t_0 = \alpha(x - t_0) + \beta(y - t_0) = \alpha u + \beta v,$$

und somit gilt $\alpha u + \beta v \in U$.

Nun gilt $M = t_0 + U$. □

Satz 10.62. *Seien $m, n \in \mathbb{N}$, sei $A \in K^{m \times n}$ und sei $b \in K^m$. Dann ist $L := \{x \in K^n \mid Ax = b\}$ eine lineare Mannigfaltigkeit. Wenn $L \neq \emptyset$ und $x_0 \in L$, so gilt $L = x_0 + N(A)$.*

Beweis. Wenn $A \cdot a = A \cdot b = A \cdot c = 0$, so gilt auch $A \cdot (\lambda a + (1 - \lambda) b) = 0$ und $A \cdot (a - b + c) = 0$, also ist L eine lineare Mannigfaltigkeit.

Sei nun $x_0 \in L$. Für $x \in L$ gilt $x = x_0 + (-x_0 + x)$ und wegen $A \cdot (-x_0 + x) = -A \cdot x_0 + A \cdot x = -b + b = 0$ gilt $-x_0 + x \in N(A)$ und somit $x \in x_0 + N(A)$. Für $y \in N(A)$ gilt $A \cdot (x_0 + y) = A \cdot x_0 + A \cdot y = b + 0 = b$ und somit $x_0 + y \in L$. □

Beispiel 10.63. 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[58]:= A =  {{3,1,2,-5},{0,1,-3,8}}
Out[58]= {{3,1,2,-5},{0,1,-3,8}}

In[59]:= b = {4,1}
Out[59]= {4,1}

In[60]:= NullSpace [A]
Out[60]= {{13,-24,0,3},{-5,9,3,0}}

In[61]:= LinearSolve [A,b]
Out[61]= {1,1,0,0}
```

Die Lösungsmenge ist damit gegeben durch

$$L = \{(1,1,0,0) + s \cdot (13,-24,0,3) + t \cdot (-5,9,3,0) \mid s,t \in \mathbb{R}\}.$$

Übungsaufgaben 10.64.

(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.

Es gibt eine Möglichkeit, Gleichungssysteme direkt durch Ausrechnen der Treppennormalform zu lösen. Dazu berechnet man für die Lösung von $A \cdot x = b$ mit $A \in \mathbb{R}^{m\times n}$ die Treppennormalform von

$$B = \left(\begin{array}{c|c} -b^T & \\ A^T & E_{n+1} \end{array} \right)$$

und kann dann aus der Treppennormalform von B direkt die Lösungsmenge von $Ax = 0$ ablesen: Für $A = \begin{pmatrix} 3 & 1 & 2 & -5 \\ 0 & 1 & -3 & 8 \end{pmatrix}$ und $b = \binom{4}{1}$ erhalten wir

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

Diese Matrix hat Treppennormalform

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 8 & 3 \\ 0 & 1 & 0 & 0 & 0 & 5 & 2 \\ 0 & 0 & 1 & 0 & 0 & 37 & 14 \\ 0 & 0 & 0 & 1 & 0 & -24 & -9 \\ 0 & 0 & 0 & 0 & 1 & -13 & -5 \end{pmatrix},$$

und wir lesen daraus, dass die Lösungsmenge

$$L = \{(0,0,37,14) + s(1,0,-24,-9) + t(0,1,-13,-5) \mid s,t \in \mathbb{R}\}$$

ist.

10. Koordinaten

Satz 10.65. *Sei $B = (b_i)_{i\in I}$ eine Basis des Vektorraums V und sei $v \in V$. Dann gibt es genau eine Familie $(\lambda_i)_{i\in I}$, sodass gilt:*

(1) $\{i \in I \mid \lambda_i \neq 0\}$ *ist endlich, und*
(2) $v = \sum\limits_{\substack{i\in I\\ \lambda_i\neq 0}} \lambda_i b_i.$

Beweis. Da $v \in L(B)$ existiert eine solche Familie $(\lambda_i)_{i\in I}$. Seien nun $(\lambda_i)_{i\in I}$ und $(\mu_i)_{i\in I}$ zwei Familien, die die beiden Eigenschaften erfüllen. Dann gilt für $(\nu_i)_{i\in I}$ mit $\nu_i = \lambda_i - \mu_i$ für $i \in I$, dass nur endlich viele ν_i ungleich 0 sind, und dass

$$\sum_{\substack{i\in I\\ \lambda_i\neq 0 \text{ oder } \mu_i\neq 0}} \nu_i b_i = \sum_{\substack{i\in I\\ \lambda_i\neq 0 \text{ oder } \mu_i\neq 0}} \lambda_i b_i - \sum_{\substack{i\in I\\ \lambda_i\neq 0 \text{ oder } \mu_i\neq 0}} \mu_i b_i = v - v = 0.$$

Da $(b_i)_{i\in I}$ linear unabhängig ist, gilt daher $\nu_i = 0$ für alle $i \in I$. Somit gilt $(\lambda_i)_{i\in I} = (\mu_i)_{i\in I}$. □

Wir nennen $(v)_B := (\lambda_i)_{i\in I}$ die *Koordinaten* von v bezüglich B. Es gilt also stets für alle $v \in V$ die Gleichheit

$$v = \sum_{\substack{i\in I\\ \lambda_i\neq 0}} (v)_B[i]\ b_i.$$

Wenn $I = \{1, \dots, m\}$, also $B = (b_1, \dots, b_m)$, so liegt $(\lambda_1, \dots, \lambda_m)$ in K^m, und wir nennen $(\lambda_1, \dots, \lambda_m)$ das *Koordinatentupel* von v bezüglich B.

Im Spezialfall, dass $B = (b_1, \dots, b_m)$ eine Basis des Unterraums T von K^n ist, kann man das Koordinatentupel mithilfe eines Gleichungssystems ausrechnen. Sei dazu $\overline{B}$ die $n \times m$-Matrix, in deren Spalten die Vektoren $(b_1, \dots, b_m)$ stehen. Sei $t \in T$. Dann erfüllt $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} := (t)_B$ das Gleichungssystem

$$\overline{B} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = t.$$

Das Koordinatentupel $(t)_B$ erfüllt also die Gleichung

$$\overline{B} \cdot (t)_B = t.$$

Beispiele 10.66.

(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 10.67.

(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 ε 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 ε 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 $\begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}$ die Koordinaten $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ hat.

KAPITEL 11

Lineare Abbildungen

1. Beispiele

Wir betrachten einige Funktionen, die von einem Vektorraum in einen Vektorraum gehen.

Beispiel 11.1. 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:

$$\begin{array}{rccl} s & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) & \longmapsto & \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right). \end{array}$$

Beispiel 11.2. Wir überlegen uns, wo der Punkt $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ nach einer Drehung um den Nullpunkt um 60° gegen den Uhrzeigersinn landet. Sei d diese Drehung. Dann lässt sich d so schreiben:

$$\begin{array}{rccl} d & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) & \longmapsto & \begin{pmatrix} \cos(\frac{\pi}{3}) & -\sin(\frac{\pi}{3}) \\ \sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3}) \end{pmatrix} \cdot \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right). \end{array}$$

Wenn man dreidimensionale Objekte auf einem Bildschirm zeichnen möchte, kann man dazu die folgende Abbildung verwenden:

$$\begin{array}{rccl} p & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & & \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right) & \longmapsto & \begin{pmatrix} 1 & 0 & \frac{1}{2\sqrt{2}} \\ 0 & 1 & \frac{1}{2\sqrt{2}} \end{pmatrix} \cdot \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right). \end{array}$$

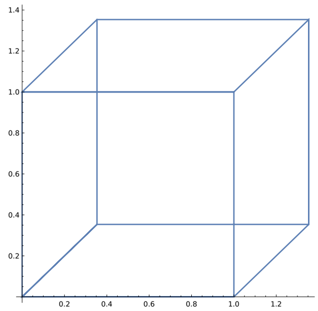
Die Punkte

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

des Würfels werden dadurch auf

$$(0,0),(1.000,0),(1.000,1.000),(0,1.000),$$
$$(0.3536,1.354),(1.354,1.354),(1.354,0.3536),(0.3536,0.3536)$$

abgebildet, und es ergibt sich folgende Zeichnung des Würfels (sie heißt *Kabinettprojektion*).



Solche Abbildungen, die man durch Matrizen beschreiben kann, werden der Inhalt dieses Kapitels sein.

2. Definition linearer Abbildungen

Definition 11.3. Seien U und V Vektorräume über denm Körper K. Eine Funktion $h : U \to V$ ist eine *lineare Abbildung*, wenn:

(1) für alle $u_1, u_2 \in U : h(u_1 + u_2) = h(u_1) + h(u_2)$, und
(2) für alle $u \in U$ und für alle $\lambda \in \mathbb{R} : h(\lambda\, u) = \lambda\, h(u)$.

Alle Abbildungen aus Abschnitt 1 sind lineare Abbildungen.

Beispiel 11.4. Sei $V = C[0,1]$, die Menge aller stetigen Funktion von $[0,1]$ nach $\mathbb{R}$. Die Abbildung

$$\begin{array}{rcrcl} I & : & C[0,1] & \longrightarrow & \mathbb{R} \\ & & f & \longmapsto & \int_0^1 f(x)\,dx \end{array}$$

ist linear.

Beispiel 11.5. Sei $V = \mathbb{R}^{\mathbb{N}}$ und sei

$$\begin{array}{rcrcl} E & : & \mathbb{R}^{\mathbb{N}} & \longrightarrow & \mathbb{R}^{\mathbb{N}} \\ & & (a_i)_{i\in\mathbb{N}} & \longmapsto & (a_{i+1})_{i\in\mathbb{N}} \end{array}.$$

Es gilt also für jede Folge a, dass $E(a)\,(i) := a(i+1)$. Die Fibonacci-Folge, die durch $f_1 = f_2 = 1$, $f_{i+2} = f_{i+1} + f_i$ für $i \in \mathbb{N}$ definiert ist, ist also eine Lösung der Gleichung $E(E(f)) - E(f) - E = 0$.

Beispiel 11.6. 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.

Übungsaufgaben 11.7.

(1) Eine lineare Abbildung 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

Aus Matrizen gewinnt man lineare Abbildungen:

Satz 11.8. *Sei K ein Körper, und sei $A \in K^{n\times m}$. Dann ist die Abbildung h, die durch*

$$\begin{array}{rcrcl} h & : & K^m & \longrightarrow & K^m \\ & & x & \longmapsto & A \cdot x \end{array}$$

definiert ist, eine lineare Abbildung vom K-Vektorraum K^m in den K-Vektorraum K^n.

In diesem Abschnitt werden wir zeigen, dass alle linearen Abbildungen von K^m nach K^n von dieser Form sind.

Satz 11.9. *Seien U, V Vektorräume über dem Körper K, und sei h eine bijektive lineare Abbildung von U nach V. Dann ist h^{-1} eine lineare Abbildung von V nach U.*

Beweis. Seien $x, y \in V$ und $u := h^{-1}(x)$, $v := h^{-1}(y)$. Dann gilt $h^{-1}(x+y) = h^{-1}(h(u) + h(v)) = h^{-1}(h(u+v)) = u + v = h^{-1}(x) + h^{-1}(y)$ und $h^{-1}(\alpha x) = h^{-1}(\alpha h(u)) = h^{-1}(h(\alpha u)) = \alpha u = \alpha h^{-1}(x)$. $\square$

Für eine Menge I und einen Körper K definieren wir

$$K^{(I)} := \{(k_i)_{i\in I} \in K^I \mid \{i \in I \mid k_i \neq 0\} \text{ ist endlich}\}.$$

Die Menge $K^{(I)}$ ist unter Addition und Vervielfachung abgeschlossen und daher ein Unterraum von K^I.

Satz 11.10. *Sei V ein K-Vektorraum mit Basis $(b_i)_{i\in I}$, und sei s definiert durch*

$$\begin{array}{rccl} s & : & K^{(I)} & \longrightarrow & V \\ & & (\lambda_i)_{i\in I} & \longmapsto & \sum_{i\in I, \lambda_i \neq 0} \lambda_i b_i. \end{array}$$

Dann ist s eine bijektive lineare Abbildung.

Beweis. Satz 10.65 liefert, dass λ surjektiv und injektiv ist. Für die Linearität beobachten wir

$$\begin{aligned} s((\lambda)_{i\in I} + (\mu)_{i\in I}) &= s((\lambda_i + \mu_i)_{i\in I}) \\ &= \sum_{\substack{i\in I \\ \lambda_i\neq 0 \vee \mu_i \neq 0}} (\lambda_i + \mu_i)\, b_i \\ &= \sum_{\substack{i\in I \\ \lambda_i\neq 0 \vee \mu_i \neq 0}} \lambda_i\, b_i + \sum_{\substack{i\in I \\ \lambda_i\neq 0 \vee \mu_i \neq 0}} \mu_i\, b_i \\ &= s((\lambda)_{i\in I}) + s((\mu)_{i\in I}) \end{aligned}$$

und

$$s(\alpha \cdot (\lambda)_{i\in I}) = s((\alpha\,\lambda)_{i\in I}) = \sum_{\substack{i\in I \\ \alpha\,\lambda_i\neq 0}} \alpha\lambda_i\, b_i = \alpha \cdot \sum_{\substack{i\in I \\ \alpha\,\lambda_i\neq 0}} \lambda_i\, b_i = \alpha \cdot s((\lambda)_{i\in I}).$$

$\square$

Korollar 11.11. *Sei V ein K-Vektorraum mit Basis $(b_i)_{i\in I}$, und sei $c : V \to K^{(I)}$ mit $c(v) := (v)_B$. Dann ist c eine bijektive lineare Abbildung.*

Beweis. Die Abbildung c ist die inverse der in Satz 11.10 definierten linearen Abbildung s und somit bijektiv und wegen Satz 11.9 ebenfalls linear. $\square$

Lineare Abbildungen zwischen endlichdimensionalen Vektorräumen kann man mit Matrizen darstellen. Sei $h : U \to V$ eine lineare Abbildung und sei $B = (b_1, \ldots, b_m)$ 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 fassen jetzt die Bilder der Basisvektoren in einer Matrix, der *Darstellungsmatrix* von h zusammen.

Definition 11.12. Seien U, V Vektorräume über dem Körper K, seien $m, n \in \mathbb{N}$, sei $B = (b_1, \dots, b_m)$ eine Basis von U, sei $C = (c_1, \dots, c_n)$ eine Basis von V, und sei h eine lineare Abbildung von U nach V. Die $n \times m$-Matrix $M_{B,C}(h)$ ist definiert durch

$$M_{B,C}(h)\,[i,j] := (h(b_j))_C\,[i].$$

für $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, m\}$. Diese Matrix ist die *Abbildungsmatrix* oder *Darstellungsmatrix* von h bezüglich der Basen B und C.

In der j-ten Spalte von $M_{B,C}(h)$ steht der Vektor $(h(b_j))_C$. Es gilt also

$$M_{B,C}(h) = \begin{pmatrix} | & | & & | \\ (h(b_1))_C & (h(b_2))_C & \cdots & (h(b_m))_C \\ | & | & & | \end{pmatrix}.$$

Mithilfe dieser Matrizen kann man das Bild einer linearen Abbildung ausrechnen.

Satz 11.13. *Seien U, V Vektorräume über dem Körper K, seien $m, n \in \mathbb{N}$, sei $B = (b_1, \dots, b_m)$ eine Basis von U und $C = (c_1, \dots, c_n)$ eine Basis von V, und Sei h eine lineare Abbildung von U nach V. Dann gilt*

$$(h(u))_C = M_{B,C}(h) \cdot (u)_B \text{ für alle } u \in U. \tag{11.1}$$

Beweis. Sei $u \in U$ mit $u = \sum_{j=1}^{m} \lambda_j\, b_j$, und sei $i \in \{1, \dots, n\}$. Dann gilt

$$\begin{aligned}
(M_{B,C}(h) \cdot (u)_B)\ [i] &= \sum_{j=1}^{m} M_{B,C}(h)[i,j] \cdot (u)_B[j] \\
&= \sum_{j=1}^{m} (h(b_j))_C[i] \cdot \lambda_j \\
&= \big(\sum_{j=1}^{m} \lambda_j \cdot (h(b_j))_C\big)\,[i] \\
&= (h(\sum_{j=1}^{m} \lambda_j b_j))_C\,[i] \\
&= (h(u))_C\,[i].
\end{aligned}$$

□

Beispiel 11.14. Sei $h : \mathbb{R}^2 \to \mathbb{R}^3$ die folgende lineare Abbildung.

$$\begin{array}{rccc} h & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & & \binom{x}{y} & \longmapsto & \begin{pmatrix} 3x-2y \\ 2x+y \\ -x+4y \end{pmatrix}. \end{array}$$

Sei B die Basis $(\binom{-1}{2}, \binom{4}{5})$ 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 $M_{B,C}(h)$.

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
$$\begin{pmatrix} -9 & 20 & 0 \\ -13 & 39 & 1 \\ -7 & 30 & 0 \end{pmatrix} \cdot x = \left(\begin{smallmatrix} -7 \\ 0 \\ 9 \end{smallmatrix}\right)$$
erhalten wir $(h(b_1))_C = \left(\begin{smallmatrix} 3 \\ 1 \\ 0 \end{smallmatrix}\right)$ und ebenso $(h(b_2))_C = \left(\begin{smallmatrix} 2 \\ 1 \\ 0 \end{smallmatrix}\right)$. Insgesamt erhalten wir
$$M_{B,C}(h) = \begin{pmatrix} 3 & 2 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Beispiel 11.15. Die lineare Abbildung $h : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch
$$h((x, y)) = 3\,x - 2\,y$$
für alle $x, y \in \mathbb{R}$, sei $B := ((1, 0), (0, 1))$ und $C = (1)$. In diesem Fall ist $(h(b_1))_C = (3)$ und $(h(b_2))_C = (-2)$. Somit gilt
$$M_{B,C}(h) = (3 \;\; -2).$$

Beispiel 11.16. Sei $U = \mathbb{R}^2, V = \mathbb{R}$ 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 $M_{B,C}(h) = (0 \;\; 6.5)$.

Die Abbildungsmatrix ist eindeutig bestimmt:

Satz 11.17. *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 $u \in U$:*
$$(h(u))_C = A \cdot (u)_B. \tag{11.2}$$
Dann gilt $A = M_{B,C}(h)$.

Beweis. Sei $i \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, n\}$. Wir setzen $u := b_j$ und erhalten
$$M_{B,C}(h)[i, j] = (h(b_j))_C\,[i] = (A \cdot (b_j)_B)\,[i] = (A \cdot e_j)\,[i] = A[i, j].$$
□

Übungsaufgaben 11.18.

(1) Eine lineare Abbildung $h : \mathbb{R}^2 \to \mathbb{R}^2$ sei gegeben durch $h(\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)) = \left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right)$ und $h(\left(\begin{smallmatrix} -2 \\ 1 \end{smallmatrix}\right)) = \left(\begin{smallmatrix} 1 \\ 7 \end{smallmatrix}\right)$. Bestimmen Sie die Abbildungsmatrix $M_{E,E}(h)$, wobei $E = (\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right))$.

(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 $M_{E_3,E_2}(\sigma)$ an.

4. Matrixmultiplikation und Komposition linearer Abbildungen

Die Matrizenmultiplikation haben wir wie in Definition 2.7 definiert, damit der folgende Satz gilt.

Satz 11.19. *Seien U, V, W endlichdimensionale 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*
$$M_{A,C}(g \circ f) = M_{B,C}(g) \cdot M_{A,B}(f). \tag{11.3}$$

Beweis. Wir zeigen, dass für alle $u \in U$ gilt:

$$(M_{B,C}(g) \cdot M_{A,B}(f)) \cdot (u)_A = ((g \circ f)\ (u))_C.$$

Dann folgt aus Satz 11.17 die Gleichung (11.3). Sei $u \in U$. Dann gilt $(M_{B,C}(g) \cdot M_{A,B}(f)) \cdot (u)_A = M_{B,C}(g) \cdot (M_{A,B}(f) \cdot (u)_A) = M_{B,C}(g) \cdot (f(u))_B = \big(g(f(u))\big)_C = \big(g \circ f\ (u)\big)_C$. □

Beispiel 11.20. Man spiegle den Punkt $(2,3)$ an der x-Achse, und drehe anschließend den gespiegelten Punkt um 90° gegen den Uhrzeigersinn um den Nullpunkt:

Lösung: Die Spiegelung σ an der x-Achse hat bzgl. der kanonischen Basis E des $\mathbb{R}^2$ die Darstellungsmatrix $M_{E,E}(\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Die Drehung δ um 90° hat die Darstellungsmatrix $M_{E,E}(\delta) = \begin{pmatrix} \cos(90°) & -\sin(90°) \\ \sin(90°) & \cos(90°) \end{pmatrix}$. Die Darstellungsmatrix der Spiegelung mit anschließender Drehung ist also

$$M_{E,E}(\delta \circ \sigma) = M_{E,E}(\delta) \cdot M_{E,E}(\sigma) = \begin{pmatrix} \cos(90°) & \sin(90°) \\ \sin(90°) & -\cos(90°) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Wir erhalten $\delta(\sigma((2,3))) = (3,2)$.

Definition 11.21. Seien U, V Vektorräume über dem Körper K, und sei $h : U \to V$ eine lineare Abbildung. Dann definieren wir den *Kern von h* durch

$$\ker(h) := \{u \in U \,|\, h(u) = 0\}.$$

Wir definieren das *Bild von h* oder *Image von h* durch

$$\operatorname{im}(h) := \{h(u) \,|\, u \in U\}.$$

Lemma 11.22. *Seien U, V Vektorräume über dem Körper K, und sei $h : U \to V$ eine lineare Abbildung Dann ist der Kern von h ein Unterraum von U, und das Image von h ein Unterraum von V.*

Satz 11.23. *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* $\operatorname{im}(h) = V$.

Beweis. (1): Wenn h injektiv ist und $y \in \ker(h)$, so gilt $h(y) = h(0) = 0$ und da h injektiv ist, auch $y = 0$.

Wenn $\ker(h) = \{0\}$ und $h(x) = h(y)$, so gilt $0 = h(x) - h(y) = h(x) + (-1)h(y) = h(x) + h(-y) = h(x - y)$, und somit $x - y \in \ker(h)$, also $x - y = 0$. □

Satz 11.24. *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) $\operatorname{im}(h_A) = S(A)$.
(3) *Der Rang von A ist die Dimension von* $\operatorname{im}(h_A)$.

Satz 11.25. *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.*

Beweis. Seien $a_1, \dots, a_n \in K^m$ die Spalten von A.

(2)⇒(3): Sei $\lambda \in K$ so, dass $\sum_{i=1}^n \lambda_i a_i = 0$. Dann gilt $A \cdot \lambda = 0$. Wegen (2) gilt daher $\lambda = 0$ und somit $a_1, \dots, a_n$ linear unabhängig.

(3)⇒(2): Sei $x \in N(A)$. Dann gilt $A \cdot x = 0$ und somit $\sum_{i=1}^n x_i a_i = 0$. Wegen (3) gilt daher $x = 0$. □

Satz 11.26. *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.*

Beweis. Seien $z_1, \dots, z_m$ die Zeilen von A.

(1)⇒(2) Wenn h_A surjektiv ist, so gilt $\mathrm{im}(h_A) = K^m$, also $m = \dim(S(A)) = \dim(Z(A)) = \mathrm{rk}(A)$.

(2)⇒(1) Wenn $\mathrm{rk}(A) = m$, so gilt $\dim(\mathrm{im}(h_A)) = \dim(S(A)) = \dim(Z(A)) = m$, und somit gilt wegen Satz 10.35, dass $\mathrm{im}(h_A) = K^m$.

(2)⇒(3) Wenn die Zeilen von A linear abhängig sind, so gibt es wegen Satz 10.18 ein $j \in \{1, \dots, m\}$ mit $z_j \in L(z_1, \dots, z_{j-1})$, und somit gilt $L(z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_m) = L(z_1, \dots, z_m)$. Sei $(b_1, \dots, b_l)$ eine Basis von $Z(A)$. Dann ist $(b_1, \dots, b_l)$ eine linear unabhängige Folge in $L(z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_m)$ und somit gilt wegen Satz 10.19, dass $l \leq m - 1$, im Widerspruch zu $\mathrm{rk}(A) = m$.

(3)⇒(2) Wenn $(z_1, \dots, z_m)$ linear unabhängig ist, so ist $(z_!, \dots, z_m)$ eine Basis von $Z(A)$. Somit gilt $\dim(Z(A)) = m$, also $\mathrm{rk}(A) = m$. □

Satz 11.27. *Sei K ein Körper, und sei $A \in K^{n\times n}$, und sei $h_A : K^n \to K^n$, $h_A(x) := A \cdot x$. Dann sind äquivalent:*

(1) *Es gibt eine Matrix B in $K^{n\times n}$ mit $A \cdot B = E_n$.*
(2) *Es gibt eine Matrix C in $K^{n\times n}$ mit $C \cdot A = E_n$.*
(3) *Die Matrix A ist invertierbar.*
(4) *h_A ist injektiv.*
(5) *h_A ist surjektiv.*

Beweis. (4)⇒(5): Da h_A injektiv ist, gilt $\dim(N(A)) = 0$ und folglich wegen Satz 10.58 $\dim(S(A)) = n$. Aus Satz 10.35 ergibt sich daher, dass $S(A) = K^n$. Wegen Satz 11.24 gilt daher $\mathrm{im}(h_A) = K^n$, also ist h_A surjektiv.

(5)⇒(4) Da h_A surjektiv ist, gilt $\dim(S(A)) = n$, daher wegen des Rangsatzes $\dim(N(A)) = 0$, also $\ker(h_A) = \{0\}$. Somit ist h_A injektiv.

(4)⇒(3) Sei h_A injektiv. Wir haben bereits gezeigt, dass h_A dann surjektiv ist, und folglich ist h_A bijektiv. Sei E die kanonische Basis von K^n und E_n die $n \times n$-Einheitsmatrix. Wegen Satz 11.9 ist die inverse Abbildung h_A^{-1} ebenfalls linear, und wegen Satz 11.19 gilt

$$M_{E,E}(h_A^{-1}) \cdot M_{E,E}(h_A) = M_{E,E}(\mathrm{id}) \text{ und } M_{E,E}(h_A) \cdot M_{E,E}(h_A^{-1}) = M_{E,E}(\mathrm{id}).$$

Sei $X := M_{E,E}(h_A^{-1})$. Wegen Satz 11.17 gilt $M_{E,E}(h_A) = A$ und $M_{E,E}(\mathrm{id}) = E_n$, also gilt

$$X \cdot A = E_n \text{ und } A \cdot X = E_n.$$

Folglich ist A invertierbar.

(3)⇒(1) und (3)⇒(2): offensichtlich.

(1)⇒(5): Wegen $A \cdot B = E$ gilt $h_A \circ h_B = \mathrm{id}$. Da id surjektiv ist, ist daher wegen Satz 7.25(1) h_A surjektiv.

(2)⇒(4) Wegen $C \cdot A = E$ gilt $h_C \circ h_A = \mathrm{id}$. Da id injektiv ist, ist daher wegen Satz 7.25(2) h_A injektiv. □

5. Lineare Abbildungen als Homomorphismen

Definition 11.28. 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.

Lemma 11.29. *Seien U, V Vektorräume über dem Körper K, und sei f ein Homomorphismus von U nach V, und sei $(u_i)_{i\in I}$ eine Familie aus U. Dann gilt:*

(1) $f[L((u_i)_{i\in I})] = L((f(u_i))_{i\in I})$.
(2) *Wenn $(u_i)_{i\in I}$ linear abhängig ist, so ist auch $(f(u_i))_{i\in I}$ linear abhängig.*
(3) *Wenn f injektiv und $(u_i)_{i\in I}$ linear unabhängig ist, so ist auch $(f(u_i))_{i\in I}$ linear unabhängig.*

Beweis. (1) Sei $x \in f[L((u_i)_{i\in I})]$. Dann gibt es $y \in L((u_i)_{i\in I})$ mit $x = f(y)$. Da $y \in L((u_i)_{i\in I})$, gibt es $(\alpha_i)_{i\in I} \in K^{(I)}$ mit $y = \sum_{i\in I, \alpha_i \neq 0} \alpha_i u_i$ und somit $x = f(y) = f(\sum_{i\in I} \alpha_i u_i) = \sum_{i\in I} \alpha_i f(u_i) \in L((f(u_i))_{i\in I})$.

Wenn $x \in L((f(u_i))_{i\in I})$, so gibt es $(\beta_i)_{i\in I} \in K^{(I)}$ mit $x = \sum_{i\in I} \beta_i f(u_i)$, also gilt $x = f(\sum_{i\in I} \beta_i u_i) \in f[L((u_i)_{i\in I})]$.

(2) Sei $(\lambda_i)_{i\in I} \in K^{(I)} \setminus \{0\}$ so, dass $\sum_{i\in I} \lambda_i u_i = 0$. Dann gilt auch $0 = f(0) = f(\sum_{i\in I} \lambda_i u_i) = 0 = \sum_{i\in I} \lambda_i f(u_i) = 0$, und somit ist $(f(u_i))_{i\in I}$ linear abhängig.

(3) Wir nehmen an, dass $(f(u_i))_{i\in I}$ linear abhängig ist. Dann gibt es $(\lambda_i)_{i\in I} \in K^{(I)} \setminus \{0\}$ mit $\sum_{i\in I} \lambda_i f(u_i) = 0$. Dann gilt $0 = \sum_{i\in I} \lambda_i f(u_i) = f(\sum_{i\in I} \lambda_i u_i) = 0$. Da f injektiv ist, gilt auch $\sum_{i\in I} \lambda_i u_i = 0$. Das belegt, dass $(u_i)_{i\in I}$ linear abhängig ist. □

Satz 11.30. *Seien K ein Körper und U, V K-Vektorräume. Wir nehmen an, dass U eine endliche Basis $B = (b_1, \ldots, b_m)$ besitzt, und dass es eine bijektive lineare Abbildung f von U nach V gibt. Dann hat auch V eine endliche Basis $C = (c_1, \ldots, c_n)$, es gilt $m = n$, und $M_{C,B}(f^{-1}) = M_{B,C}(f)^{-1}$*

Beweis. Für $i \in \{1, \ldots, m\}$ definieren wir $c_i := f(b_i)$. Wegen Lemma 11.29(1) gilt

$$L(c_1, \ldots, c_m) = f[L(b_1, \ldots, b_m)] = f[U] = V,$$

und wegen Lemma 11.29(3) ist $C := (c_1, \ldots, c_m)$ linear unabhängig. Somit ist C eine Basis von V.

Es gilt nun $M_{C,B}(f^{-1}) \cdot M_{B,C}(f) = M_{B,B}(\mathrm{id}_U) = E_m$ und $M_{B,C}(f) \cdot M_{C,B}(f^{-1}) = M_{C,C}(\mathrm{id}_U) = E_m$. Somit ist $M_{B,C}(f)$ invertierbar, und es gilt $(M_{B,C}(f))^{-1} = M_{C,B}(f^{-1})$. □

Wir bezeichen zwei Vektorräume U, V, zwischen denen es eine bijektive lineare Abbildung, also einen Isomorphismus, gibt, als *isomorph*. Aus Lemma 11.29 folgt, dass U und V dann Basen gleicher Mächtigkeit haben.

Satz 11.31 (Rangsatz für lineare Abbildungen). *Seien V, W Vektorräume, und sei $f : V \to W$. Seien $B = (b_i)_{i\in I}$ eine Basis von $\operatorname{im}(f)$ und $C = (c_j)_{j\in J}$ eine Basis von $\ker(f)$ mit $I \cap J = \varnothing$. Für $j \in J$ sei $y_j \in V$ so, dass $f(y_j) = b_j$. Dann ist $(b_i)_{i\in I} \cup (y_j)_{j\in J}$ eine Basis von V.*

Wenn V endlichdimensional ist, so gilt also $\dim(V) = \dim(\operatorname{im}(f)) + \dim(\ker(f))$.

Beweis. Wie in Satz 10.58 zeigt man, dass $(b_i)_{i\in I} \cup (y_j)_{j\in J}$ eine Basis von U ist.

Im endlichdimensionalen Fall gilt also $\dim(U) = \dim(\operatorname{im}(f)) + \dim(\ker(f))$. (Diese Gleichheit gilt auch im unendlichdimensionalen, aber dazu müsste man die *Addition von Kardinalzahlen* definieren.) □

Lemma 11.32. *Seien V, W Vektorräume über dem Körper K, sei h ein Epimorphismus von V nach W, und sei U ein Unterraum von W. Dann gilt $\dim(h^{-1}[U]) = \dim(U) + \dim(\ker(h))$.*

Beweis. Die Abbildung $\hat{h} : h^{-1}[U] \to W$ erfüllt $\ker(\hat{h}) = \ker(h)$ und $\operatorname{im}(\hat{h}) = U$. Also gilt wegen Satz 11.31, dass $\dim(h^{-1}(U)) = \dim(\ker(h)) + \dim(U)$. □

Wir können damit einige Ungleichungen für den Rang einer Matrix beweisen.

Satz 11.33. *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) $\operatorname{rk}(A) \le k$ *und* $\operatorname{rk}(A) \le l$.
(2) $\operatorname{rk}(A \cdot B) \le \min(\operatorname{rk}(A), \operatorname{rk}(B))$.
(3) *(Sylvesters Rangungleichung)* $\operatorname{rk}(A \cdot B) \ge \operatorname{rk}(A) + \operatorname{rk}(B) - l$.
(4) $\operatorname{rk}(A_1 + A_2) \le \operatorname{rk}(A_1) + \operatorname{rk}(A_2)$.
(5) $\operatorname{rk}(A_1 + A_2) \ge \operatorname{rk}(A_1) - \operatorname{rk}(A_2)$.

Beweis. (1) Da $Z(A) \subseteq K^l$, gilt wegen Satz 10.19, dass $\operatorname{rk}(A) = \dim(Z(A)) \le l$ und wegen $S(A) \subseteq K^k$ gilt $\operatorname{rk}(A) \le k$. (2) Es gilt $\operatorname{rk}(A \cdot B) = \dim(\operatorname{im}(h_{A\cdot B})) = \dim(\operatorname{im}(h_A \circ h_B))$. Nun gilt $\operatorname{im}(h_A \circ h_B) \subseteq \operatorname{im}(h_A)$, also $\dim(\operatorname{im}(h_A \circ h_B)) \le \operatorname{rk}(A)$.

Sei $C = (c_1, \dots, c_n)$ eine Basis von $\operatorname{im}(h_B)$. Dann gilt

$$\operatorname{im}(h_A \circ h_B)) = h_A[\operatorname{im}(h_B)] = h[L(c_1, \dots, c_n)] = L(h(c_1), \dots, h(c_n)).$$

Wegen Satz 10.19 gilt daher $\dim(\operatorname{im}(h_A \circ h_B)) \le n = \dim(\operatorname{im}(h_B))$, also $\operatorname{rk}(A \cdot B) \le \operatorname{rk}(B)$.

(3)

$$\begin{aligned}\dim(\operatorname{im}(h_A \circ h_B)) &= m - \dim(\ker(h_A \circ h_B)) \\ &= m - \dim(\{x \in K^m \mid h_B(x) \in \ker(h_A)\}) \\ &= m - \dim(\{x \in K^m \mid h_B(x) \in \ker(h_A) \cap S(B)\}) \\ &= m - \dim(h_B^{-1}[\ker(h_A) \cap S(B)]).\end{aligned}$$

Da h_B surjektiv von K^m auf $\operatorname{im}(h_B) = S(B)$ ist, gilt wegen Lemma 11.32, dass

$$\begin{aligned}\dim(h_B^{-1}[\ker(h_A) \cap S(B)]) &= \dim(\ker(h_A) \cap S(B)]) + \dim(\ker(h_B)) \\ &\le \dim(\ker(h_A)) + \dim(\ker(h_B)).\end{aligned}$$

Somit gilt

$$\begin{aligned} m - \dim(h_B^{-1}[\ker(h_A) \cap S(B)]) &\geq m - (\dim(\ker(h_A)) + \dim(\ker(h_B))) \\ &= m - \dim(\ker(h_A)) - \dim(\ker(h_B)) \\ &= m - (l - \dim(\operatorname{im}(h_A)) - (m - \dim(\operatorname{im}(h_B))) \\ &= \dim(\operatorname{im}(h_A)) + \dim(\operatorname{im}(h_B)) - l \\ &= \operatorname{rk}(B) + \operatorname{rk}(A) - l. \end{aligned}$$

(4) Sei C_1 eine Basis von $\operatorname{im}(h_{A_1})$ und C_2 eine Basis von $\operatorname{im}(h_{A_2})$. Dann gilt $\operatorname{im}(h_{A_1+A_2}) \subseteq L(C_1 \cup C_2)$, und somit $\dim(\operatorname{im}(h_{A_1+A_2})) \leq |C_1| + |C_2| = \operatorname{rk}(A_1) + \operatorname{rk}(A_2)$.

(5) $\operatorname{rk}(A) = \operatorname{rk}((A+B)+(-B)) \leq \operatorname{rk}(A+B)+\operatorname{rk}(-B) = \operatorname{rk}(A+B)+\operatorname{rk}(B)$, also $\operatorname{rk}(A+B) \geq \operatorname{rk}(A) - \operatorname{rk}(B)$. □ □

6. Basistransformationen

Wir lösen folgendes Problem: Gegeben sind zwei Basen B, C des gleichen Vektorraums V und die Koordinaten $(v)_B$ eines Vektors $v \in V$. Gesucht sind die Koordinaten von v bezüglich C, also $(v)_C$.

Beispiel 11.34. 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

$$x = B \cdot (x)_B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

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 $(x)_C = (-0.2, 0.6)$.

Wenn V endlichdimensional ist, können wir die Basistransformation auch durch eine Matrixmultiplikation realisieren.

Definition 11.35. Sei $B = (b_1, \ldots, b_n)$ und $C = (c_1, \ldots, c_n)$. Wir definieren die *Basistransformationsmatrix* $T_{B\to C}$ *von* B *nach* C durch

$$T_{B\to C} := M_{B,C}(\mathrm{id}_V).$$

Es gilt dann

$$T_{B\to C} \cdot (v)_B = M_{B,C}(\mathrm{id}_V) = (\mathrm{id}_V(v))_C = (v)_C$$

für alle $v \in V$, und wegen Satz 11.30 auch $T_{C\to B} = T_{B\to C}^{-1}$.

Wenn $V = K^n$ und $E = (e_1, \ldots, e_n)$ die kanonische Basis ist, so gilt für jedes $i \in \{1, \ldots, n\}$, dass $T_{B\to E} \cdot (b_j)_B = (b_j)_E$, also $T_{B\to E} \cdot e_j = b_j$. Somit ist die j-te Spalte von $T_{B\to E}$ gleich b_j. Dann gilt $T_{B\to C} = T_{E\to C} \cdot T_{B\to E} = T_{C\to E}^{-1} \cdot T_{B\to E}$. Wenn in den Spalten von $\overline{B}$ die Vektoren der Basis B und in den Spalten von $\overline{C}$ die Vektoren aus C stehen, gilt also $T_{B\to E} = \overline{B}$, $T_{E\to C} = \overline{C}^{-1}$ und $T_{B\to C} = \overline{C}^{-1}\overline{B}$.

Übungsaufgaben 11.36.

(1) (Koordinaten) Die Ebene ε 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*).

Wir wollen nun die Abbildungsmatrizen für bestimmte Drehungen und Spiegelungen bestimmen. Wir betrachten folgende Beispiele:

Beispiel 11.37. Sei e die Ebene $L(\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix})$, und sei σ 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 $M_{E,E}(\sigma)$ dieser Spiegelung bezüglich der kanonischen Basis E des $\mathbb{R}^3$.

Beispiel 11.38. Sei g die Gerade mit der Gleichung $5\,x - 2\,y = 0$ im $\mathbb{R}^2$. Sei σ 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 $M_{E,E}(\sigma)$ dieser Spiegelung bezüglich der kanonischen Basis E des $\mathbb{R}^2$.

Beispiel 11.39. 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 $60°$ um die Gerade g landet. Gesucht ist $M_{E,E}(\delta)$. Wir müssen noch die Richtung der Drehung festlegen: Wenn wir vom Punkt $\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$ auf die Ebene $x + 2\,y - 2\,z = 0$ 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 $M_{B,B}(h)$ leicht zu bestimmen ist.
(2) Bestimme $M_{B,B}(h)$.
(3) Bestimme $M_{E,E}(h)$ aus $M_{B,B}(h)$.

Wir starten mit Beispiel 11.37. 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 $(\sigma(b_1))_B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $(\sigma(b_2))_B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ und $(\sigma(b_3))_B = (-b_3)_B = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$. Dann gilt

$$M_{B,B}(\sigma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Wir bestimmen nun eine Basis B mit den gewünschten Eigenschaften. Dazu wählen wir $b_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ und $b_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Wir bestimmen einen Vektor, der auf beide normal steht: Das Gleichungssystem $1x + 2y - 1z = 0$, $y - z = 0$ hat die Lösungsmenge $L((-1, 1, 1))$. Also wählen wir $b_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, und somit

$$B = (\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}).$$

Es gilt nun

$$M_{E,E}(\sigma) = T_{B\to E} \cdot M_{B,B}(\sigma) \cdot T_{E\to B}.$$

Sei $\overline{B}$ die Matrix, in deren Spalten die Vektoren aus B stehen. Dann gilt

$$M_{E,E}(\sigma) = \overline{B} \cdot M_{B,B}(\sigma) \cdot \overline{B}^{-1}.$$

```
In[82]:= B=Transpose[{{1,2,-1},{0,1,-1},{-1,1,1}}]
Out[82]= {{1,0,-1},{2,1,1},{-1,-1,1}}

In[83]:= MatrixForm [B]
Out[83]//MatrixForm=
(1  0  -1
 2  1   1
-1 -1   1

)
In[84]:= ShBB={{1,0,0},{0,1,0},{0,0,-1}}
Out[84]= {{1,0,0},{0,1,0},{0,0,-1}}

In[85]:= ETB=B
Out[85]= {{1,0,-1},{2,1,1},{-1,-1,1}}

In[86]:= BTE=Inverse[B]
Out[86]= {{2/3,1/3,1/3},{-1,0,-1},{-(1/3),1/3,1/3}}

In[87]:= ShEE=ETB.ShBB.BTE
Out[87]= {{1/3,2/3,2/3},{2/3,1/3,-(2/3)},{2/3,-(2/3),1/3}}

In[88]:= MatrixForm[ShEE]
Out[88]//MatrixForm=
(1/3      2/3     2/3
 2/3      1/3  -(2/3)
 2/3  -(2/3)      1/3)
```

Beispiel 11.38: Wir wählen eine Basis $B = (b_1, b_2)$ 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

$$M_{B,B}(\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

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

$$B = (\left(\begin{smallmatrix}2\\5\end{smallmatrix}\right), \left(\begin{smallmatrix}5\\-2\end{smallmatrix}\right)).$$

```
In[89]:= B=Transpose[{{2,5},{5,-2}}]
Out[89]= {{2,5},{5,-2}}

In[90]:= MatrixForm [B]
Out[90]//MatrixForm=
(2  5
```

```
5  -2)

In[91]:= ShBB={{1,0},{0,-1}}
Out[91]= {{1,0},{0,-1}}

In[92]:= ETB=B
Out[92]= {{2,5},{5,-2}}

In[94]:= BTE=Inverse[B]
Out[94]= {{2/29,5/29},{5/29,-(2/29)}}

In[95]:= ShEE=ETB.ShBB.BTE
Out[95]= {{-(21/29),20/29},{20/29,21/29}}

In[96]:= MatrixForm[ShEE]
Out[96]//MatrixForm=
(-(21/29)  20/29
   20/29   21/29)
```

Beispiel 11.39: Wir bestimmen zunächst B und $M_{B,B}(\delta)$. Dazu bestimmen wir eine Basis $B = (b_1, b_2, b_3)$ von $\mathbb{R}^3$, sodass wir die Bilder der Basisvektoren gut ausrechnen können. Da $\delta(\left(\begin{smallmatrix}1\\2\\-2\end{smallmatrix}\right)) = \left(\begin{smallmatrix}1\\2\\-2\end{smallmatrix}\right)$, wählen wir $b_3 = \left(\begin{smallmatrix}1\\2\\-2\end{smallmatrix}\right)$. Die anderen beiden Basisvektoren b_1 und b_2 wählen wir so, dass sie die Ebene durch 0 aufspannen, die auf b_3 normal ist. Die Bilder unter δ können wir dann leicht ausrechnen, wenn b_1, b_2 nicht nur in der Ebene $\{x \in \mathbb{R}^3 \mid \langle x, b_3\rangle = 0\}$ liegen, sondern zusätzlich noch $||b_1|| = ||b_2||$ erfüllen und aufeinander normal sind.

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.“ Dann ist nämlich

$$\delta(b_1) = \cos(\pi/3) * b_1 + \sin(\pi/3) * b_2$$

und

$$\delta(b_2) = -\sin(\pi/3) * b_1 + \cos(\pi/3) * b_2.$$

Mit $c := \cos(60°)$ und $s := \sin(60°)$ lässt sich $M_{B,B}(\delta)$ also durch

$$M_{B,B}(\delta) = \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

angeben. Jetzt bestimmen eine Basis B des $\mathbb{R}^3$ mit den gewünschten Eigenschaften. Durch Lösen des Gleichungssystems $\langle b_3, \left(\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right)\rangle = 0$ finden wir zum Beispiel den Vektor $c_1 = \left(\begin{smallmatrix}2\\0\\1\end{smallmatrix}\right)$, der auf b_3 normal steht. Nun suchen wir einen Vektor c_2, der auf b_3 und c_1 normal steht. Einen solchen Vektor können wir mit $c_2 = b_3 \times c_1 = \left(\begin{smallmatrix}2\\-5\\-4\end{smallmatrix}\right)$ bestimmen. Es gilt $||c_1|| = \sqrt{5}$ und $||c_2|| = 3\sqrt{5}$. Also ist $B := (b_1, b_2, b_3)$ mit $b_1 = 3c_1$ und $b_2 = c_2$ eine Basis, in der jeder der drei Vektoren auf die beiden anderen normal steht, und in der $||b_1|| = ||b_2||$. Eine gesuchte Basis sind also die Spalten von

$$\overline{B} = \begin{pmatrix} 6 & 2 & 1 \\ 0 & -5 & 2 \\ 3 & -4 & -2 \end{pmatrix}.$$

Da $b_1 \times b_2 = \left(\begin{smallmatrix} 15 \\ 30 \\ -30 \end{smallmatrix}\right)$ und $b_3 = \left(\begin{smallmatrix} 1 \\ 2 \\ -2 \end{smallmatrix}\right)$ in die gleiche Richtung zeigen, ist (b_1, b_2, b_3) positiv orientiert. Somit haben wir auch für Beispiel 11.39 eine Basis B und die Matrix $M_{B,B}(\delta)$ gefunden.

```
In[100]:= B = Transpose [{{6,0,3},{2,-5,-4},{1,2,-2}}]
Out[100]= {{6,2,1},{0,-5,2},{3,-4,-2}}

In[101]:= MatrixForm [B]
Out[101]//MatrixForm=
(6   2  1
 0  -5  2
 3  -4 -2)

In[102]:= c = Cos [Pi/3]; s = Sin [Pi/3]; ShBB = {{c,-s,0}, {s,c,0},{0,0,1}}
Out[102]= {{1/2,-(Sqrt[3]/2),0},{Sqrt[3]/2,1/2,0},{0,0,1}}

In[103]:= ETB=B
Out[103]= {{6,2,1},{0,-5,2},{3,-4,-2}}

In[104]:= BTE=Inverse[B]
Out[104]= {{2/15,0,1/15},{2/45,-(1/9),-(4/45)},{1/9,2/9,-(2/9)}}

In[106]:= ShEE=FullSimplify [ETB.ShBB.BTE]
Out[106]= {{5/9,1/9+1/Sqrt[3],-(1/9)+1/Sqrt[3]},
           {1/9-1/Sqrt[3],13/18,1/18 (-4-3 Sqrt[3])},
           {-(1/9)-1/Sqrt[3],1/18 (-4+3 Sqrt[3]),13/18}}

In[108]:= MatrixForm[N [ShEE, 6]]
Out[108]//MatrixForm=
(0.555556  0.688461   0.466239
-0.466239  0.722222  -0.510897
-0.688461  0.0664529  0.722222)
```

Übungsaufgaben 11.40.

(1) Finden Sie eine Basis B, bezüglich der die Spiegelung δ an der Ebene $x + y + z = 0$ die Abbildungsmatrix

$$M_{B,B}(\delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

hat.

(2) Bestimmen Sie eine geeignete Basis B des $\mathbb{R}^2$, für die die Abbildungsmatrix der Spiegelung s an der Geraden $g : 7\,x - 4\,y$ folgende Form besitzt:

$$M_{B,B}(s) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right).$$

(3) (Spiegelung an einer Geraden in der Ebene) Wir bezeichnen mit σ jene Spiegelung, die jeden Punkt der Ebene $\mathbb{R}^2$ an der Geraden $-3\,x + 4\,y = 0$ spiegelt.

(a) Bestimmen Sie eine Basis B des $\mathbb{R}^2$, sodass für die Abbildungsmatrix $M_{B,B}(\sigma)$ der Spiegelung σ bezüglich der Basis B folgende Gleichung gilt.

$$M_{B,B}(\sigma) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right).$$

(b) Berechnen Sie die Abbildungsmatrix $M_{E,E}(\sigma)$ der Spiegelung σ bezüglich der kanonischen Basis E.

   (c) Wo landet der Punkt (a, b) nach dieser Spiegelung σ?

(4) Sei h die lineare Abbildung, die jeden Punkt an der Ebene $e : x + 2\,y + 2\,z = 0$ 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 $M_{B,B}(h)$ 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 $M_{E,E}(\sigma)$ 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 $2\,x - y + z = 0$ 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 σ jene Spiegelung, die jeden Punkt an der Geraden
$$x + y = 0$$
   spiegelt. Wo landet der Punkt (a, b) nach dieser Spiegelung σ? 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 $e : x + 2\,y + 2\,z = 0$ spiegelt. Wir suchen in diesem Beispiel $M_{E,E}(h)$, wobei E die kanonische Basis des $\mathbb{R}^3$ ist. Gehen Sie dazu so vor:
   (a) Bestimmen Sie $M_{E,B}(\text{id})$. Dabei ist id die identische Abbildung. $M_{E,B}(\text{id})$ ist also eine Basistransformationsmatrix, und erfüllt die Eigenschaft $(v)_B = M_{E,B}(\text{id}) \cdot (v)_E$ für alle $v \in \mathbb{R}^3$.)
   (b) Bestimmen Sie $M_{B,E}(\text{id})$.
   (c) Bauen Sie aus diesen beiden Matrizen und $M_{B,B}(h)$ die Matrix $M_{E,E}(h)$ 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 σ an der Ebene
$$-2\,x - y + z = 0\ ?$$
   Bestimmen Sie die Abbildungsmatrix dieser Spiegelung!

(10) (Spiegelung an einer Geraden in der Ebene) Wir bezeichnen mit σ jene Spiegelung, die jeden Punkt an der Geraden $12\,x + 5\,y = 0$ spiegelt. Wo landet der Punkt (a, b) nach dieser Spiegelung σ? Bestimmen Sie die Abbildungsmatrix dieser Spiegelung bezüglich der kanonischen Basis.

(11) Wir bezeichnen mit σ jene Spiegelung, die jeden Punkt an der Geraden $15\,x - 8\,y = 0$ spiegelt. Wo landet der Punkt (a, b) nach dieser Spiegelung σ? Bestimmen Sie die Abbildungsmatrix dieser Spiegelung bezüglich der kanonischen Basis.

(12) Geben Sie die Darstellungsmatrix für die Spiegelung an der Geraden $x - 2\,y = 0$ 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 δ um $90°$ 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 δ um $90°$ 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!

7. Existenz linearer Erweiterungen

In diesem Abschnitt werden wir Basen eines Vektorraums als Mengen ansehen. V ist stets ein Vektorraum über dem Körper K.

SATZ 11.41. *Sei M eine linear unabhängige Teilmenge von V. Dann gibt es eine Basis B von V mit $M \subseteq B$.*

BEWEIS. Wir zeigen mithilfe des Zornschen Lemmas wie in Satz 10.30, dass es eine maximale lineare unabhängige Teilmenge B von V gibt, die M als Teilmenge enthält. Diese Menge ist dann wegen Lemma 10.29 eine Basis von V. $\square$

LEMMA 11.42. *Seien V, W Vektorräume über dem Körper K, sei X eine Teilmenge von V, und sei $f : X \to V$ eine Funktion, die folgende Eigenschaft erfüllt:*

für alle $n \in \mathbb{N}$, für alle $x_1, \ldots, x_n \in X$ und

$$\text{für alle } \lambda_1, \ldots, \lambda_n \in K \text{ gilt } \sum_{i=1}^{n} \lambda_i x_i = 0 \Rightarrow \sum_{i=1}^{n} \lambda_i f(x_i) = 0. \quad (11.4)$$

Dann gibt es einen Homomorphismus $g : L(X) \to W$, sodass $g(x) = f(x)$ für alle $x \in X$.

BEWEIS. Wir definieren die Relation h durch

$$h := \Big\{\big(\sum_{b \in T} \lambda(b) * b, \sum_{b \in T} \lambda(b) * f(b)\big) \mid T \text{ ist eine endliche Teilmenge von } X, \lambda : T \to K\Big\}.$$

Dann ist h ein Homomorphismus von $L(X)$ nach W und $h|_X = f$. Wir zeigen als erstes, dass h eine Funktion ist. Seien dazu T_1, T_2 endliche Teilmengen von X, 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 μ 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 Eigenschaft (11.4) auch $\sum_{b \in T_1 \cup T_2} (\lambda_1(b) - \mu_1(b)) * f(b) = 0$. 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 λ_1, μ_1 wieder die Erweiterungen von λ und μ wie oben. Die Eigenschaft $h(\alpha * v) = \alpha * h(v)$ zeigt man genauso.

Sei nun $b \in X$. Dann gilt $h(b) = h(1 * b) = 1 * f(b) = f(b)$, also gilt $h|_X = f$. $\square$

SATZ 11.43. *Seien V, W Vektorräume über dem Körper K, sei U ein Unterraum von V und sei $g : U \to W$ ein Homomorphismus. Dann gibt es einen Homomorphismus $h : V \to W$ mit $h(u) = g(u)$ für alle $u \in U$.*

BEWEIS. Wegen Satz 10.30 gibt es eine Basis B von U. Wegen Satz 11.41 gibt es eine Basis C von V mit $B \subseteq C$. Wir definieren nun $f(c) = g(c)$ für $c \in B$ und $f(c) = 0$ für $c \in C \setminus B$. Da C linear unabhängig ist, erfüllen f und C die Bedingung (11.4). Folglich gibt es einen Homomorphismus $h : L(C) \to W$ mit $h|_C = g|_C$. Da g und $h|_U$ Homomorphismen von U nach W sind, die auf der Basis B übereinstimmen, gilt $g = h|_U$. $\square$

SATZ 11.44. *Seien V, W Vektorräume über dem Körper K, sei X eine linear unabhängige Teilmenge von V, und sei $f : X \to V$ eine Funktion. Dann gibt es einen Homomorphismus $h : V \to W$, sodass $h(x) = f(x)$ für alle $x \in X$.*

Wenn X eine Basis von V ist, so gibt es genau einen solchen Homomorphismus h.

BEWEIS. Wegen Lemma 11.42 gibt es einen Homomorphismus $g : L(X) \to W$ mit $g|_X = f$. Wegen Satz 11.43 gibt es einen Homomorphismus $h : V \to W$ mit $h|_{L(X)} = g$.

Nehmen wir nun an, dass X eine Basis von V ist und $h_1, h_2 \in \mathrm{Hom}_K(V, W)$ mit $h_1|_X = f = h_2|_X$. Wegen $(h_1 - h_2)|_X = 0$ folgt aus der Linearität, dass $(h_1 - h_2)_{L(X)} = 0$ und somit $h_1 = h_2$. $\square$

KOROLLAR 11.45. *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 10.30, 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. $\square$

ÜBUNGSAUFGABEN 11.46.

(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.

(3) Sei M eine Teilmenge von V mit $L(M) = V$. Zeigen Sie, dass es eine Teilmenge B von M gibt, die eine Basis von V ist.

KAPITEL 12

Determinanten

1. Volumen eines Parallelepipeds

Sei $A = \begin{pmatrix} \text{—}z_1\text{—} \\ \text{—}z_2\text{—} \\ \vdots \\ \text{—}z_m\text{—} \end{pmatrix}$ eine reelle $m \times m$-Matrix. Das *von den Zeilen von A aufgespannte Parallelepiped* ist die Teilmenge

$$\{\sum_{i=1}^{m} \lambda_i \cdot z_i \mid \lambda_1, \ldots, \lambda_m \in [0,1]\}$$

von $\mathbb{R}^m$. Wir versuchen jetzt, das Volumen dieses Parallelepipeds zu messen, und schreiben $D(z_1, z_2, \ldots, z_m)$ 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) *D ist multilinear*: Für alle $z_1, \ldots, z_m, y \in K^m$, $\alpha, \beta \in K$ und $i \in \{1, \ldots, m\}$ gilt

$$D(z_1, \ldots, z_{i-1}, \alpha * z_i + \beta * y, z_{i+1}, \ldots, z_m)$$
$$= \alpha D(z_1, \ldots, z_{i-1}, z_i, z_{i+1}, \ldots, z_m) + \beta D(z_1, \ldots, z_{i-1}, y, z_{i+1}, \ldots, z_m).$$

(D2) *D ist alternierend*: Für alle $z_1, \ldots, z_m$ und $i, j \in \{1, \ldots, m\}$ mit $i \neq j$ gilt: Wenn $z_i = z_j$, so gilt $D(z_1, \ldots, z_m) = 0$.

(D3) *D ist normiert*: Für Einheitsvektoren $e_1, \ldots, e_m)$ im K^m gilt $D(e_1, \ldots, e_m) = 1$.

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 $z_1, \ldots, z_m \in K^m$, $\alpha \in K$, und seien $i, j \in \{1, \ldots, m\}$ mit $i \neq j$. Dann gilt*

(1) $D(z_1, \ldots, z_{j-1},\ \alpha * z_i + z_j,\ z_{j+1}, \ldots, z_m) = D(z_1, \ldots, z_m)$.

(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

$$\begin{aligned}
0 &= D(z_1, \dots, z_{i-1},\ z_i,\ z_{i+1}, \dots, z_{j-1},\ z_i, z_{j+1}, \dots, z_m) \\
&\quad + D(z_1, \dots, z_{i-1},\ z_i,\ z_{i+1}, \dots, z_{j-1},\ z_j, z_{j+1}, \dots, z_m) \\
&\quad + D(z_1, \dots, z_{i-1},\ z_j,\ z_{i+1}, \dots, z_{j-1},\ z_i,\ z_{j+1}, \dots, z_m) \\
&\quad + D(z_1, \dots, z_{i-1},\ z_j,\ z_{i+1}, \dots, z_{j-1},\ z_j, z_{j+1}, \dots, z_m) \\
&= 0 + D(z_1, \dots, z_{i-1},\ z_i,\ z_{i+1}, \dots, z_{j-1},\ z_j, z_{j+1}, \dots, z_m) \\
&\quad + D(z_1, \dots, z_{i-1},\ z_j,\ z_{i+1}, \dots, z_{j-1},\ z_i,\ z_{j+1}, \dots, z_m) + 0.
\end{aligned}$$

□

Übungsaufgaben 12.2.

Sei D eine Funktion von $(\mathbb{R}^4)^4$ nach $\mathbb{R}$, die die Eigenschaften (D1), (D2) und (D3) erfüllt, also multlinear, alternierend und normiert ist. Seien $e_1 = (1, 0, 0, 0)$, …, $e_4 = (0, 0, 0, 1)$.

(1) Bestimmen Sie $D(e_4, e_3, e_2, e_1)$, indem Sie Satz 12.1 (2) verwenden, also durch Vertauschen der Argumente $D(e_1, e_2, e_3, e_4)$ erreichen.
(2) Bestimmen Sie
$$\begin{array}{ll} D(& (0,0,5,0), \\ & (0,1,0,0), \\ & (0,0,0,1), \\ & (-5,0,4,0) \quad), \end{array}$$
indem Sie verwenden, dass D multilinear, alternierend und normiert ist. Die Determinante welcher Matrix haben Sie dadurch berechnet?

Satz 12.3. *Sei K ein kommutativer Ring mit Eins, und sei $D : K^2 \times K^2 \to K$ so, dass D die Eigenschaften (D1), (D2), (D3) erfüllt. Dann gilt für alle $a, b, c, d \in K$: $D((a, b), (c, d) = ad - bc$.*

Beweis. Es gilt

$$\begin{aligned}
D((a, b), (c, d)) &= a\, D((1, 0), (c, d)) + b\, D((0, 1), (c, d)) \\
&= ac\, D((1, 0), (1, 0)) + ad\, D((1, 0), (0, 1)) \\
&\quad + bc\, D((0, 1), (1, 0)) + bd\, D((0, 1), (0, 1)) \\
&= 0\, ac + ad - bc + 0\, bd = ad - bc.
\end{aligned}$$

□

Man rechnet leicht nach, dass diese Funktion $D : K^2 \times K^2 \to K$ auch wirklich multilinear, alternierend und normiert ist.

Für $m > 2$ führt die Auswertung von $D(z_1, \dots, z_m)$ auf die Auswertung von $D(e_{i_1}, \dots, e_{i_m})$ mit $i_1, \dots, i_m \in \{1, \dots, m\}$. Wenn es k, l mit $k \neq l$ und $i_k = i_l$ gibt, so gilt $D(e_{i_1}, \dots, e_{i_m}) = 0$, weil D alterniernd ist. Für alle anderen Summanden müssen wir uns überlegen, ob und durch wieviele Vertauschungen man aus $D(e_{i_1}, \dots, e_{i_m})$ den Ausdruck $D(e_1, \dots, e_m)$ bilden kann. Es gilt zum Beispiel $D(e_2, e_3, e_1) = -D(e_1, e_3, e_2) = D(e_1, e_2, e_3) = 1$ und $D(e_3, e_2, e_1) = -D(e_1, e_2, e_3) = -1$.

2. Permutationen und Signatur

Definition 12.4. Sei X eine Menge, $m \in \mathbb{N}$. Eine *Permutation von X* ist eine bijektive Abbildung von X nach X. Die Menge aller Bijektionen von X kürzt man auch mit S_X, die Menge aller Bijektionen von $\{1, \dots, m\}$ 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 $f(1) = 3, \ldots, f(7) = 5$. Im allgemeinen beschreibt

$$\begin{pmatrix} 1 & 2 & \ldots & m \\ a_1 & a_2 & \ldots & a_m \end{pmatrix}$$

also die Funktion f mit $f(i) = a_i$ für $i \in \{1, \ldots, m\}$.

Definition 12.5 (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 12.6. *Sei $m \in \mathbb{N}$, und seien $f, g \in S_m$, und sei*

$$F(f) := \{(i,j) \in \{1, 2, \ldots, m\}^2 \mid i < j \text{ und } f(i) > f(j)\}$$

die Menge der Fehlstellen *von f. Dann gilt:*

(1) $\operatorname{sgn}(f) = (-1)^{|F(f)|}$.
(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) $\operatorname{sgn}(f) = \dfrac{\displaystyle\prod_{\substack{(i,j) \in \{1,\ldots,m\}^2 \\ i<j}} f(i) - f(j)}{\displaystyle\prod_{\substack{(i,j) \in \{1,\ldots,m\}^2 \\ i<j}} i - j}$. Für den Zähler dieses Produkts gilt:

$$\prod_{\substack{(i,j) \in \{1,\ldots,m\}^2 \\ i<j}} f(i) - f(j)$$

$$= \Big(\prod_{\substack{(i,j) \in \{1,\ldots,m\}^2 \\ i<j \text{ und } f(i)<f(j)}} f(i) - f(j) \Big) \cdot \Big(\prod_{\substack{(i,j) \in \{1,\ldots,m\}^2 \\ i<j \text{ und } f(i)>f(j)}} f(j) - f(i) \Big) \cdot (-1)^{F(f)}.$$

Wir behaupten nun, dass dieses Produkt gleich $(-1)^{F(f)} \displaystyle\prod_{\substack{(s,t) \in \{1,\ldots,m\}^2 \\ s<t}} (s - t)$ 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

$$\operatorname{sgn}(f\circ g) = \frac{\displaystyle\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} f(g(i))-f(g(j))}{\displaystyle\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} i-j}$$

$$= \frac{\displaystyle\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} f(g(i))-f(g(j))}{\displaystyle\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} g(i)-g(j)} \cdot \frac{\displaystyle\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} g(i)-g(j)}{\displaystyle\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} i-j}. \tag{12.1}$$

Für den ersten Bruch in der rechten Seite von (12.1) gilt

$$\prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)}$$

$$= \prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j \text{ und } g(i)<g(j)}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)} \cdot \prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j \text{ und } g(i)>g(j)}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)}$$

$$= \prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j \text{ und } g(i)<g(j)}} \frac{f(g(i))-f(g(j))}{g(i)-g(j)} \cdot \prod_{\substack{(i,j)\in\{1,\ldots,m\}^2\\ i<j \text{ und } g(i)>g(j)}} \frac{f(g(j))-f(g(i))}{g(j)-g(i)}$$

$$= \prod_{\substack{(s,t)\in\{1,\ldots,m\}^2\\ s<t}} \frac{f(s)-f(t)}{s-t} = \operatorname{sgn}(f).$$

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 12.7 (Zyklen). Seien $i_1, i_2, \ldots, i_k$ paarweise verschiedene Zahlen in $\{1,\ldots,m\}$. Dann bezeichnen wir mit $(i_1\ i_2\ i_3\ \ldots\ i_k)$ die Permutation f mit $f(i_1) = i_2, \ldots, f(i_{k-1}) = i_k$, $f(i_k) = i_1$, $f(j) = j$ für alle $j \in \{1,\ldots,m\}\setminus\{i_1,\ldots,i_k\}$. Diese Abbildung ist ein *Zyklus der Länge k* oder *k-Zyklus*. Eine *Transposition* ist ein Zyklus der Länge 2.

Satz 12.8. *Sei $m\in\mathbb{N}$, $a,b\in\mathbb{N}_0$, $i,j\in\{1,\ldots,m\}$, und sei $\sigma\in S_m$.*

(1) *σ 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 $\rho_1,\ldots,\rho_a$ und $\tau_1,\ldots,\tau_b$ mit $\sigma = \rho_1\circ\cdots\circ\rho_a = \tau_1\circ\cdots\circ\tau_b$ teilt 2 die Differenz $a-b$.*
(4) *Für $n\in\mathbb{N}$ und paarweise verschiedene $i_1,\ldots,i_n\in\{1,\ldots,m\}$ gilt*
$$\operatorname{sgn}((i_1\ i_2\ \ldots\ i_n)) = (-1)^{n+1}.$$

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 $\sigma = \mathrm{id}$; σ ist dann also das Produkt von 0 Transpositionen. Sei nun $n \geq 1$, und sei σ 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 σ ebenfalls ein Produkt von Transpositionen.

(2) Für den Fall, dass $i = 1$ und $j = 2$ bestimmen wir
$$\mathrm{sgn}((1\,2)) = (-1)^{|F((1\,2))|} = (-1)^1 = -1.$$
Seien nun $i, j \in \underline{n}$ mit $i \neq j$, sei $\tau := (i\,j)$, und sei f eine Permutation mit $f(1) = i$ und $f(2) = j$. Dann gilt
$$(i\,j) = f \circ (1\,2) \circ f^{-1}.$$
Sei dazu $x \in \underline{n}$. Wenn $x \notin \{i, j\}$, so gilt $f^{-1}(x) \notin \{1, 2\}$, und somit $\tau(f^{-1}(x)) = f^{-1}(x)$, also $f(\tau(f^{-1}(x))) = x$. Ausserdem gilt $f \circ (1\,2) \circ f^{-1}\,(i) = j$ und $f \circ (1\,2) \circ f^{-1}\,(j) = i$. Also gilt wegen Satz 12.6
$$\mathrm{sgn}((i\,j)) = \mathrm{sgn}(f)^2 \cdot \mathrm{sgn}((1\,2)) = -1.$$

(3) Die Signatur von σ ist $(-1)^a = (-1)^b$.

(4) Es gilt $(i_1\ i_2\ \dots\ i_n) = (i_1\ i_n) \circ (i_1\ i_2\ \dots,\ i_{n-1})$, somit folgt die behauptete Gleichheit aus (2) durch Induktion nach n. □

Satz 12.9. *Sei $m \geq 2$, und seien $i, j \in \{1, \dots, m\}$ 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 $A_m \cap ((i\ j) \circ A_m) = \varnothing$ 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 φ 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 φ bijektiv. □

Übungsaufgaben 12.10.

(1) Der Beweis von Satz 12.8 (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 12.11. *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 $z_1, \dots, z_m \in K^m$, und sei $f \in S_m$. Dann gilt*
$$D(z_{f(1)}, \dots, z_{f(m)}) = \mathrm{sgn}(f)\, D(z_1, \dots, z_m). \tag{12.2}$$

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 τ_i Transpositionen sind, und sei $g := \tau_n \circ \cdots \circ \tau_2$. Dann gilt wegen Satz 12.1
$$D(z_{g(\tau_1(1))}, \dots, z_{g(\tau_1(m))}) = -D(z_{g(1)}, \dots, z_{g(m)}).$$

Nach Induktionsvoraussetzung ist dieser Ausdruck gleich $-\operatorname{sgn}(g)\, D(z_1, \ldots, z_m)$. Da $f = g \circ \tau_1$, gilt $\operatorname{sgn}(f) = -\operatorname{sgn}(g)$. Also gilt $D(z_{f(1)}, \ldots, z_{f(m)}) = \operatorname{sgn}(f)\, D(z_1, \ldots, z_m)$. Das beschließt den Induktionsbeweis.

Da sich jede Permutation als Produkt von endlich vielen Transpositionen schreiben lässt, gilt (12.2) also für alle $f \in S_m$. $\square$

3. Determinante einer quadratischen Matrix

Eine Funktion D mit den in Sektion 1 geforderten Eigenschaften erhält man mithilfe der *Determinante.*

Definition 12.12. 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

$$\det(A) := \sum_{f \in S_m} \operatorname{sgn}(f) \prod_{i=1}^{m} A(i, f(i)).$$

Satz 12.13. *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*

$$\begin{array}{cccc} \mathbf{D} & : & (K^m)^m & \longrightarrow & K \\ & & (z_1, z_2, \ldots, z_m) & \longmapsto & \det\left(\begin{pmatrix} -z_1- \\ -z_2- \\ \vdots \\ -z_m- \end{pmatrix}\right). \end{array}$$

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 12.14. *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

$$\begin{aligned}
\det(A \cdot B) = \mathbf{D}\big(\quad & A(1,1)b_1 + A(1,2)b_2 + \cdots + A(1,m)b_m, \\
& A(2,1)b_1 + A(2,2)b_2 + \cdots + A(2,m)b_m, \\
& \ldots \\
& A(m,1)b_1 + A(m,2)b_2 + \cdots + A(m,m)b_m \big).
\end{aligned}$$

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 12.11 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

$$\det(A \cdot B) = \sum_{f \in S_m} \Big(\big(\prod_{i=1}^{m} A(i, f(i))\big) \operatorname{sgn}(f) \, \det(B) \Big) = \det(A) \cdot \det(B).$$

□

SATZ 12.15. *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 Treppenform zu berechnen:

SATZ 12.16. *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 Treppennormalform 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[1]:= << RowRed12.m
In[2]:= DeterminantenDemo [{{1,2,-3},{2,29,-16}, {-3,-16,22}}]
Det[(1    2     -3
     2    29    -16
    -3   -16     22)]

 Wir addieren das -2 fache der 1. Zeile zum 1 fachen der 2.Zeile.
  = 1 * 1 * det[(1    2    -3
                 0    25  -10
                -3   -16   22)]

 Wir addieren das 3 fache der 1. Zeile zum 1 fachen der 3.Zeile.
  = 1 * 1 * det[(1    2    -3
                 0    25  -10
                 0   -10   13)]

 Wir addieren das 2 fache der 2. Zeile zum 5 fachen der 3.Zeile.
  = 1 * 1/5 * det[(1    2    -3
                   0   25    -10
                   0    0     45)]
  = 225
```

```
Out[2]= 225


In[4]:= DeterminantenDemo [{{1,2,-3,4},{0,-5,12,-7},{4,3,0,9},{8,0,0,1}}]
Det[(1    2    -3   4
     0   -5    12  -7
     4    3    0    9
     8    0    0    1)]

 Wir addieren das -4 fache der 1. Zeile zum 1 fachen der 3.Zeile.
  = 1 * 1 * det[(1    2    -3    4
                 0    -5    12  -7
                 0    -5    12  -7
                 8     0    0    1)]

 Wir addieren das -8 fache der 1. Zeile zum 1 fachen der 4.Zeile.
  = 1 * 1 * det[(1     2    -3     4
                 0    -5    12    -7
                 0    -5    12    -7
                 0    -16   24   -31)]

 Wir addieren das -1 fache der 2. Zeile zum 1 fachen der 3.Zeile.
  = 1 * 1 * det[(1    2    -3    4
                0    -5    12   -7
                 0    0    0     0
                 0  -16   24   -31)]

 Wir addieren das -16 fache der 2. Zeile zum 5 fachen der 4.Zeile.
  = 1 * 1/5 * det[(1    2    -3     4
                   0    -5    12    -7
                   0     0     0     0
                   0     0   -72   -43)]

 = -(1/5) det[(1    2     -3     4
               0   -5     12    -7
               0    0    -72   -43
               0    0      0     0)]
 = 0

Out[4]= 0
```

---

5. Die adjungierte Matrix

Definition 12.17. 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 12.18. *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 12.19. *Sei $n \geq 2$, sei A eine $n \times n$-Matrix, und seien $i, j \in \{1, \ldots, n\}$. 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

$$\det(A^{[n,n]}) = \sum_{f \in S_n} \operatorname{sgn}(f) \prod_{k=1}^{n} A^{[n,n]}(k, f(k))$$
$$= \sum_{\substack{f \in S_n \\ f(n)=n}} \operatorname{sgn}(f) \prod_{k=1}^{n-1} A(k, f(k)).$$

Sei f eine Permutation von $\{1, \ldots, n\}$ mit $f(n) = n$. Dann ist $g := f|_{\{1,\ldots,n-1\}}$ eine Permutation von $\{1, \ldots, n-1\}$. 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

$$\sum_{\substack{f \in S_n \\ f(n)=n}} \operatorname{sgn}(f) \prod_{k=1}^{n-1} A(k, f(k))$$
$$= \sum_{g \in S_{n-1}} \operatorname{sgn}(g) \prod_{k=1}^{n-1} A(k, f(k))$$
$$= \det(A^{(i,j)}).$$

Wir beweisen nun den Fall $(i, j) \neq (n, n)$. Sei σ der Zyklus $(i \;\; i+1 \;\; \ldots \;\; n)$, und $\tau := (j \;\; j+1 \;\; \ldots \;\; n)$

$$B(k, l) := A(\sigma(k), \tau(l).)$$

Dann gilt $B^{(n,n)} = A^{(i,j)}$. Außderdem gilt für alle $k, l \in \{1, \ldots, n\}$ auch

$$B^{[n,n]}(k, l) = A^{[i,j]}(\sigma(k), \tau(l)).$$

Wir berechnen nun $\det(A^{[i,j]})$. Nach Lemma 12.18 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 12.8 (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 12.20. Sei $n \in \mathbb{N}$, und sei A eine $n \times n$-Matrix. Die zu A adjungierte (oder adjunkte[1]) Matrix A^{ad} ist wie folgt definiert: wenn $n \geq 2$, so gilt

$$A^{\mathrm{ad}}(i, j) := (-1)^{i+j} \det(A^{(j,i)}) \text{ für } i, j \in \{1, \ldots, n\}.$$

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 12.21. *Sei A eine $n \times n$-Matrix. Dann gilt $(A^{\mathrm{ad}})^T = (A^T)^{\mathrm{ad}}$.*

Beweis: Für $i, j \in \{1, \ldots, n\}$ gilt $(A^{\mathrm{ad}})^T(i, j) = A^{\mathrm{ad}}(j, i) = \det(A^{[i,j]}) = \det((A^{[i,j]})^T) = \det((A^T)^{[j,i]}) = (A^T)^{\mathrm{ad}}(i, j)$. □

[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^{ad} zu tun hat, sondern mit A^T. Der Begriff der zu A *adjunkten Matrix* vermeidet diese Inkongruenz.

Satz 12.22. *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.

Wir nehmen an, dass $n \geq 2$. Für $j \in \{1, \dots, n\}$ sei

$$e_j := (0, 0, \dots, 0, 1, 0, \dots, 0)$$

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 $i \in \{1, \dots, n\}$. Wir bilden nun Matrizen $A_1, \dots, A_n$ 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

$$A_j = \begin{pmatrix} A(1,1) & \dots & A(1,j-1) & A(1,j) & A(1,j+1) & \dots & A(1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(i-1,1) & \dots & A(i-1,j-1) & A(i-1,j) & A(i-1,j+1) & \dots & A(i-1,n) \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ A(i+1,1) & \dots & A(i+1,j-1) & A(i+1,j) & A(i+1,j+1) & \dots & A(i+1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(n,1) & \dots & A(n,j-1) & A(n,j) & A(n,j+1) & \dots & A(n,n) \end{pmatrix}.$$

Aus Satz 12.1 (1) und Satz 12.13 erhalten wir, dass für alle $j \in \{1, \dots, n\}$ gilt:

$$\det(A_j) = \det(A^{[i,j]}).$$

Sei nun $k \in \{1, \dots, n\}$ 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

$$(A(k,1), A(k,2), \dots, A(k,n)),$$

also durch die k-te Zeile von A ersetzen. Es gilt also

$$B = \begin{pmatrix} A(1,1) & \dots & A(1,j-1) & A(1,j) & A(1,j+1) & \dots & A(1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(i-1,1) & \dots & A(i-1,j-1) & A(i-1,j) & A(i-1,j+1) & \dots & A(i-1,n) \\ A(k,1) & \dots & A(k,j-1) & A(k,j) & A(k,j+1) & \dots & A(k,n) \\ A(i+1,1) & \dots & A(i+1,j-1) & A(i+1,j) & A(i+1,j+1) & \dots & A(i+1,n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A(n,1) & \dots & A(n,j-1) & A(n,j) & A(n,j+1) & \dots & A(n,n) \end{pmatrix}.$$

Die Determinante dieser Matrix B ist wegen Satz 12.13 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 $A \cdot A^{\text{ad}} = \det(A) * E_n$ gilt. Dann gilt auch $A^{\text{ad}} \cdot A = (A^T \cdot (A^{\text{ad}})^T)^T = (A^T \cdot (A^T)^{\text{ad}})^T = (\det(A^T) * E_n)^T = \det(A) * E_n$. ☐

Korollar 12.23 (Entwicklungssatz von Laplace). *Sei $n \geq 2$, sei A eine $n \times n$-Matrix über einem kommutativen Ring mit Eins K, und seien $i, j \in \{1, \dots, n\}$. 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 12.22 gilt

$$\begin{aligned} \det(A) &= (A \cdot A^{\text{ad}})\ (i,i) \\ &= \sum_{k=1}^{n} A(i,k) A^{\text{ad}}(k,i) \\ &= \sum_{k=1}^{n} A(i,k)(-1)^{i+k} \det(A^{(i,k)}). \end{aligned}$$

(2) Nach Satz 12.22 gilt

$$\begin{aligned} \det(A) &= (A^{\text{ad}} \cdot A)\ (j,j) \\ &= \sum_{k=1}^{n} A^{\text{ad}}(j,k) A(k,j) \\ &= \sum_{k=1}^{n} (-1)^{j+k} \det(A^{(k,j)}) A(k,j). \end{aligned}$$

6. Determinanten und Rang

Satz 12.24. *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 12.14 gilt $1 = \det(E_n) = \det(A)\det(B)$, also leistet $y := \det(B)$ das Gewünschte. (3)⇒(1): Sei $B := y * A^{\text{ad}}$. Dann gilt $A \cdot B = y * (A \cdot A^{\text{ad}})$, und das ist wegen

Satz 12.22 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, dass eine ganzzahlige Matrix genau dann eine ganzzahlige inverse Matrix hat, wenn ihre Determinante $+1$ oder -1 ist.

SATZ 12.25. *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): $B := \frac{1}{\det(A)} A^{\mathrm{ad}}$ ist nach Satz 12.22 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 $\det(A) \neq 0$. □

DEFINITION 12.26. Seien K ein Körper, $m, n, k \in \mathbb{N}$, und sei $A \in K^{m \times n}$. Seien $i_1 < \cdots < i_k$ in $\{1, \dots, m\}$ und $j_1 < \cdots < j_k$ in $\{1, \dots, n\}$. Dann ist $B \in K^{k \times k}$ mit $B[r, s] := A[i_r, j_s]$ die zu $(i_1, \cdots, i_k)$ und $(j_1, \dots, j_k)$ gehörende *Untermatrix* von A, und ihre Determinante $\det(B)$ der zu $(i_1, \dots, i_k)$ und $(j_1, \dots, j_k)$ gehörende *Minor vom Grad* k von A.

SATZ 12.27. *Sei A eine $m \times n$-Matrix, und sei $r \in \mathbb{N}$. Dann sind äquivalent:*

(1) $r \leq \mathrm{rk}(A)$.
(2) *A besitzt eine reguläre Untermatrix vom Format $r \times r$.*

Beweis: Seien die $i_1, \dots, i_r$-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 $j_1, \dots, j_r$. (2)⇒(1): Wenn die zu $(i_1, \dots, i_k)$ und $(j_1, \dots, j_k)$ gehörende Untermatrix B von A regulär ist, so sind die $i_1, \dots, i_r$-te Zeile von A linear unabhängig, also gilt $\mathrm{rk}(A) \geq r$.

KOROLLAR 12.28. *Sei K ein Körper und $A \in R^{m \times n}$. Dann ist der Rang der Matrix das größte $k \in \mathbb{N}$, für das A einen Minor vom Grad k ungleich 0 besitzt.*

7. Gleichungssysteme

Satz 12.29. *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 $\det(A) \cdot \det(X) = \det(A_i^b)$. Durch Entwicklung nach der i-ten Zeile erhält man $\det(X) = y_i$. $\square$

Satz 12.30. *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 12.25 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. $\square$

Satz 12.31 (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 12.30 hat $A \cdot x = b$ genau eine Lösung, und wegen Satz 12.29 berechnet sich diese Lösung durch $y_i := \frac{\det(A_i^b)}{\det(A)}$. $\square$

KAPITEL 13

Konstruktionen von Vektorräumen

1. Faktorräume

DEFINITION 13.1. 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 13.2. 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 $\boldsymbol{+}$ und $\boldsymbol{*}$ 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 13.3. Sei V ein Vektorraum über dem Körper K, und sei U ein Unterraum von V. Dann ist $\langle V/U, \boldsymbol{+}, \boldsymbol{-}, 0 + U, \boldsymbol{*}\rangle$ ein Vektorraum über K.

ÜBUNGSAUFGABEN 13.4. 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 13.5. *Sei V ein Vektorraum über dem Körper K, sei $(v_i)_{i\in I}$ eine Familie aus V und sei $(u_j)_{j\in J}$ eine Familie aus U. Wir nehmen an, dass $I \cap J = \varnothing$, und dass $(v_i + U)_{i\in I}$ eine Basis von V/U und $(u_j)_{j\in J}$ eine Basis von U ist. Dann ist $B := (v_i)_{i\in I} \cup (u_j)_{j\in J}$ eine Basis von V.*

BEWEIS. Sei $f : V \to V/U$, $f(v) = v + U$ für $v \in V$. Dann gilt $\operatorname{im}(f) = V/U$ und $\ker(f) = U$. Wegen des Rangsatzes für lineare Abbildungen (Satz 11.31) ist $(v_i)_{i\in I} \cup (y_j)_{j\in J}$ eine Basis von V. □

KOROLLAR 13.6. *Sei V ein Vektorraum über K und U ein Unterraum von V. Dann gilt $\dim(V) = \dim(V/U) + \dim U$.*

Für unendlichdimensionale Vektorräume müssen wir dazu erklären, wie Gleichungen der Form $\dim(U) + \dim(V) = \dim(W)$ zu lesen sind. Man kann $\dim(U)$ als jene Kardinalzahl definieren, die die Mächtigkeit einer Basis von U beschreibt und dann die Addition von Kardinalzahlen verwenden. Eine weniger aufwändige Möglichkeit, 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 13.7 (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 $\operatorname{im}(h)$.

Anstelle "Dann ist die Abbildung [...] wohldefiniert" kann man sagen: "Dann ist die Relation $H = \{(v + U, h(v)) \,|\, v \in V\}$ eine Funktion von V/U nach W."

Aus Satz 11.31 kann man folgern, dass

$$\dim(V) = \dim(\operatorname{im}(h)) + \dim(\ker(h)),$$

und aus Korollar 13.6, dass

$$\dim(V) = \dim(V/\ker(h)) + \dim(\ker(h)).$$

Wenn V endlichdimensional ist, folgt daraus $\dim(V/\ker(h)) = \dim(\operatorname{im}(h))$. Aus Satz 13.7 ergibt sich, dass diese Gleichung auch für unendlichdimensionale Vektorräume gilt.

2. Summe von Vektorräumen

Definition 13.8. Seien U, V Vektorräume über dem Körper K. Dann definieren wir auf $U \times V$ die Operationen $\boldsymbol{+}$ durch $(u_1, v_1) \boldsymbol{+} (u_2, v_2) := (u_1 +_U u_2, v_1 +_V v_2)$ und $\boldsymbol{*}$ durch $\alpha \boldsymbol{*} (u, v) := (\alpha *_U u, \alpha *_V v)$ für alle $u, u_1, u_2 \in U$, $v, v_1, v_2 \in V$ und $\alpha \in K$ und bezeichnen $U \times V$ mit diesen Operationen als das *Produkt* der Vektorräume U und V.

Satz 13.9. *Seien U, V Vektorräume über dem Körper K. Dann gilt* $\dim(U \times V) = \dim(U) + \dim(V)$.

Beweis. Wenn $(b_i)_{i\in I}$ eine Basis von U und $(c_j)_{j\in J}$ eine Basis von V ist, so ist $((b_i, 0))_{i\in I} \cup ((0, c_j))_{j\in J}$ eine Basis von $U \times V$. □

Definition 13.10. Sei K ein Körper. Wir definieren die *Summe* der Unterräume $U_1, \ldots, U_n$ des K-Vektorraums V durch

$$U_1 + U_2 + \cdots + U_n := \Big\{\sum_{i=1}^{n} u_i \,|\, u_i \in U_i \text{ für alle } i \in \{1, \ldots, n\}\Big\}.$$

Satz 13.11. *Seien U, V Unterräume des K-Vektorraums W. Dann gilt*

(1) $\dim(U) + \dim(V) = \dim(U + V) + \dim(U \cap V)$.
(2) *Die Vektorräume $(U + V)/V$ und $U/(U \cap V)$ sind isomorph.*

Beweis. (1) Sei $f : U \times V \to U + V$ definiert durch $f((u, v)) := u + v$. Dann ist f eine lineare Abbildung, und folglich gilt $\dim(U + V) = \dim(\operatorname{im}(f)) + \dim(\ker(f))$. Es gilt $\operatorname{im}(f) = U + V$ und $\ker(f) = \{(x, -x) \mid x \in U \cap V\}$. Daher ist $\ker(f)$ isomorph zu $U \cap V$, und es gilt $\dim(\ker(f)) = \dim(U \cap V)$. Aus Satz 11.31 folgt nun $\dim(U \times V) = \dim(\operatorname{im}(f)) + \dim(\ker(f))$.

(2) Sei $f : U \to (U + V)/V$ definiert durch $f(u) := u + V$. Dann gilt $\operatorname{im}(f) = (U + V)/V$ und $\ker(f) = U \cap V$, und somit wegen des Homomorphiesatzes (Satz 13.7), dass $U/(U \cap V)$ und $(U + V)/V$ isomorph sind. □

Übungsaufgaben 13.12.

(1) Zeigen Sie, dass für Unterräume U, V, W eines Vektorraums X die Gleichheit $(U+V)\cap W = U + (V \cap W)$ im allgemeinen nicht gilt, aber richtig ist, wenn $U \subseteq W$.

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 13.13. Seien $U_1, \ldots, U_n$ Unterräume des Vektorraums V. Der Unterraum U ist genau dann die *direkte Summe* von $U_1, \ldots, U_n$, wenn

(1) $U = U_1 + \cdots + U_n$,
(2) für jedes $u \in U$ gibt es genau ein $(u_1, \ldots, u_n) \in U_1 \times \cdots \times U_n$, sodass $u = u_1 + \cdots + u_n$.

Wenn U die direkte Summe von $U_1, \ldots, U_n$ ist, so schreiben wir $U = U_1 \dotplus \cdots \dotplus U_n$.

Satz 13.14. *Seien $U_1, \ldots, U_n$ 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 $(u_1, \ldots, u_n) \in U_1 \times \cdots \times U_n$, sodass $u = u_1+\cdots+u_n$.*
(2) *Für jedes $i \in \{1, \ldots, n\}$ gilt*

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_n) = \{0\}.$$

Beweis. (1)⇒(2): Wenn $0 \neq u_i = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_n$, so besitzt u_i zwei Darstellungen, im Widerspruch zu (1).

(2)⇒(1): Sei $u_1 + \cdots + u_n = v_1 + \cdots + v_n$ mit $u_i \neq v_i$. Dann gilt $u_i - v_i = \sum_{\substack{j=1 \\ j\neq i}} (v_j - u_j)$ und infolgedessen liegt $u_i - v_i \in U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_n)$. Also gilt wegen (2), dass $u_i - v_i = 0$ im Widerspruch zu $u_i \neq v_i$. □

Satz 13.15. *Seien $n \in \mathbb{N}$ und $U, U_1, \ldots, U_n$ Unterräume des Vektorraums V. Wir nehmen an, dass $U = U_1 \dotplus \cdots \dotplus U_n$, und dass für jedes $i \in \{1, \ldots, n\}$ 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.

3. Vektorräume von Homomorphismen

Definition 13.16. Seien V, W Vektorräume über dem Körper K. Mit $\operatorname{Hom}_K(V, W)$ bezeichnen wir die Menge aller Homomorphismen von V nach W.

Für $g, h \in \operatorname{Hom}_K(V, W)$ 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 $\operatorname{Hom}_K(V, W)$ ein Vektorraum über K.

Satz 13.17. *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 Φ ist definiert durch*

$$\begin{array}{rcrcl} \Phi & : & \operatorname{Hom}_K(V, W) & \longrightarrow & K^{n\times m} \\ & & g & \longmapsto & M_{B,C}(g) \end{array}$$

Dann ist Φ ein K-Vektorraum-Isomorphismus.

KOROLLAR 13.18. *Seien V, W endlichdimensionale Vektorräume über K. Dann gilt*

$$\dim(\mathrm{Hom}_K(V, W)) = \dim(V) \cdot \dim(W).$$

SATZ 13.19. *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.*

BEWEIS. (1) Sei $\sum_{b\in B} \sum_{c\in C} \lambda_{b,c}\varphi[b, c] = 0$. Sei $b_0 \in B$. Dann gilt

$$0 = \sum_{b\in B} \sum_{c\in C} \lambda_{b,c}\varphi[b, c]\,(b_0) = \sum_{c\in C} \lambda_{b_0,c}\, c.$$

Da C linear unabängig ist, gilt daher für alle $c \in C$, dass $\lambda_{b_0,c} = 0$.

(2) Wenn $W = \{0\}$, so gilt $D = \varnothing$ und $\mathrm{Hom}_K(V, W) = \{0\}$, also ist D eine Basis für $\mathrm{Hom}_K(V, W)$. Wir nehmen nun an, dass V endlichdimensional ist und $h \in \mathrm{Hom}(V, W)$. Für jedes $b \in B$ wählen wir eine endliche Teilmenge $C(b)$ von C und $\alpha_{b,c} \in K$ so dass $h(b) = \sum_{c\in C(b)} \alpha_{b,c}c$. Sei $g := \sum_{b\in B} \sum_{c\in C(b)} \alpha_{b,c}\varphi[b, c]$. Durch Auswerten an $b_0 \in B$ erkennt man, dass g und h auf B übereinstimmen, und somit gleich sind.

Wir nehmen nun an, dass V unendlichdimensional und $W \neq \{0\}$ ist, und zeigen, dass die lineare Hülle von D nicht $\mathrm{Hom}_K(V, W)$ ist. Sei $w \in W \setminus \{0\}$. Wir definieren eine Funktion $f : B \to W$ durch $f(b) = w$ für alle $b \in B$. Wegen Satz 11.44 gibt es ein $h \in \mathrm{Hom}(V, W)$ mit $h(b) = w$ für alle $b \in B$. Für alle $g \in L(D)$ ist die Menge $\{b \in B \mid g(b) \neq 0\}$ endlich, also gilt $h \notin L(D)$. ☐

DEFINITION 13.20. 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 13.21. Sei V ein endlichdimensionaler Vektorraum mit Basis B (als Menge), und sei $b \in B$. Das zu b und B gehörende Element im Dualraum $\delta(b, B)$ sei jener Homomorphismus von V nach K, der

$$\delta(b, B)(b) = 1,\ \ \delta(b, B)(c) = 0 \text{ für } c \in B \setminus \{b\}$$

erfüllt. Wir bezeichnen diesen Homomorphismus auch mit $b^\star$.

Bei der Schreibweise $b^\star$ für $\delta(b, B)$ ist Vorsicht geboten, weil $b^\star$ auch von B abhängt. Es muss also bei der Verwendung von $b^\star$ klar sein, mit welcher Basis B wir $b^\star$ definieren. Wenn die Basis B als Familie $(b_i)_{i\in I}$ gegeben ist, dann gilt $b_i^\star(b_i) = 1$ und $b_i^\star(b_j) = 0$ für $i \neq j$.

LEMMA 13.22. *Sei V ein endlichdimensionaler Vektorraum mit Basis $B = (b_1, \dots, b_n)$, sei $B^\star = (b_1^\star, \dots, b_n^\star)$ wie in Definition 13.21, und sei $x \in V$. Dann gilt:*

(1) $b_i^\star(x) = (x)_B[i]$.
(2) $\varphi(x) = \displaystyle\sum_{i=1}^{n} \varphi(b_i) * b_i^\star(x)$.

Beweis: (1) Sei $(y_1, \ldots, y_n) := (x)_B$. Dann gilt

$$\begin{aligned}\sum_{i=1}^{n} b_i^\star(x) * b_i &= \sum_{i=1}^{n} b_i^\star\big(\sum_{j=1}^{n} y_j * b_j\big) * b_i \\ &= \sum_{i=1}^{n}\big(\sum_{j=1}^{n} y_j * b_i^\star(b_j)\big) * b_i \\ &= \sum_{i=1}^{n} y_i * b_i \\ &= x.\end{aligned}$$

Also gilt $(b_1^\star(x), \ldots, b_n^\star(x)) = (x)_B$.

(2): Es gilt

$$\begin{aligned}\sum_{i=1}^{n} \varphi(b_i) \cdot b_i^\star(x) &= \sum_{i=1}^{n} \varphi(b_i) \cdot (x)_B[i] \\ &= \varphi\big(\sum_{i=1}^{n} (x)_B[i] * b_i\big) \\ &= \varphi(x).\end{aligned}$$

□

SATZ 13.23. *Sei V ein endlichdimensionaler Vektorraum, B eine Basis von V, und $B^\star$ wie in Definition 13.21. Dann ist $B^\star$ eine Basis von $V^\star$.*

Beweis: Wir zeigen als erstes, dass $B^\star$ linear unabhängig ist. Seien $\lambda_1, \ldots, \lambda_n \in K$ so, dass $\sum_{i=1}^{n} \lambda_i b_i^\star = 0$. Sei $j \in \{1, \ldots, n\}$. Es gilt $0 = \sum_{i=1}^{n} \lambda_i b_i^\star(b_j) = \lambda_j$. Also ist $B^\star$ linear unabhängig. Wegen Lemma 13.22 (2) gilt für $\varphi \in V^\star$, dass $\varphi = \sum_{i=1}^{n} \varphi(b_i) * b_i^\star$, 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 13.22 berechnen: Für jedes $\varphi \in V^\star$ gilt $(\varphi)_{B^\star} = (\varphi(b_1), \ldots, \varphi(b_n))$, oder, anders geschrieben,

$$(\varphi)_{B^\star}[i] = \varphi(b_i). \tag{13.1}$$

Wir halten noch eine Folgerung von Satz 13.23 fest:

KOROLLAR 13.24. *Sei V ein endlichdimensionaler Vektorraum. Dann gilt* $\dim(V) = \dim(V^\star)$.

DEFINITION 13.25. 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

$$\begin{array}{rcrcl} f^\star & : & W^\star & \longrightarrow & V^\star \\ & & \varphi & \longmapsto & f^\star(\varphi), \\ f^\star(\varphi) & : & V & \longrightarrow & K \\ & & v & \longmapsto & \varphi(f(v)). \end{array}$$

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 13.26. *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 $f^\star(\alpha * \varphi)(v) = (\alpha * f^\star(\varphi))\,(v)$. $\square$

Satz 13.27. *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*

$$M_{C^\star,B^\star}(f^\star) = M_{B,C}(f)^T.$$

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 $M_{C^\star,B^\star}(f^\star)\,[i,j] = (f^\star(c_j^\star))_{B^\star}[i]$. Wegen Gleichung (13.1) ist das gleich $f^\star(c_j^\star)\ (b_i)$. Aus der Definition der dualen Abbildung $f^\star$ erhalten wir $f^\star(c_j^\star)\ (b_i) = (c_j^\star \circ f)\ (b_i) = c_j^\star(f(b_i))$. Wegen Lemma 13.22(1) ist das gleich $(f(b_i))_C[j] = S_f(B,C)[j,i]$. $\square$

Definition 13.28. 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 13.29. *Sei V ein Vektorraum über K, und sei $\Phi : V \to V^{\star\star}$ gegeben durch*

$$\begin{array}{rccc} \Phi & : & V & \longrightarrow & V^{\star\star} \\ & & v & \longmapsto & \Phi(v), \end{array}$$

wobei

$$\begin{array}{rccc} \Phi(v) & : & V^\star & \longrightarrow & K \\ & & \varphi & \longmapsto & \varphi(v). \end{array}$$

Dann ist Φ ein Monomorphismus von V nach $V^{\star\star}$. Wenn V endlichdimensional ist, ist Φ 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 Φ linear ist, zeigen wir, dass für alle $v_1, v_2 \in V$ und $\varphi \in V^\star$ gilt, dass $\Phi(v_1 + v_2)\,(\varphi) = \Phi(v_1)\,(\varphi) + \Phi(v_2)\,(\varphi)$. $\square$

Für $v \in V$ und $\varphi \in V^\star$ gilt also $\Phi(v)\,(\varphi) = \varphi(v)$. Vektorräume, für die Φ ein Isomorphismus von V nach $V^{\star\star}$ ist, heißen *reflexiv*. Jeder endlichdimensionale Vektorraum ist also reflexiv.

4. Tensorprodukt

Definition 13.30. Seien $n \in \mathbb{N}$ und $U_1, \ldots, U_n, W$ Vektorräume über dem Körper K. Eine Abbildung

$$\varphi : U_1 \times \cdots \times U_n \to W$$

ist *multilinear* $:\Longleftrightarrow$ für alle $j \in \{1, \ldots, n\}$ und alle $u_1 \in U_1, \ldots, u_{j-1} \in U_{j-1}, u_{j+1} \in U_{j+1}, \ldots, u_n \in U_n$ ist die Abbildung

$$\varphi' : U_j \to W,\ \varphi'(x) = \varphi(u_1, \ldots, u_j, x, u_{j+1}, \ldots, u_n)$$

eine lineare Abbildung. $\mathrm{Mult}(U_1, \ldots, U_n; W)$ bezeichnet die Menge aller multlinearen Abbildungen von $\bigtimes_{i=1}^n U_i$ nach W.

$\mathrm{Mult}(U_1, \ldots, U_n; W)$ ist unter $+$ und Vervielfachung abgeschlossen und bildet daher einen Vektorraum. Für $(f_1, \ldots, f_n) \in \bigtimes_{i=1}^n U_i^\star$ ist

$$\pi(f_1, \ldots, f_n) : U_1 \times \cdots \times U_n \to K,\ \pi(f_1, \ldots, f_n)(u_1, \ldots, u_n) := \prod_{i=1}^n f_i(u_i)$$

ein Element von $\mathrm{Mult}(U_1, \ldots, U_n)$.

Lemma 13.31. *Seien $n \in \mathbb{N}$ und $U_1, \dots, U_n$ endlichdimensionale K-Vektorräume. Für alle $i \in \{1, \dots, n\}$ sei B_i eine Basis von U_i, und $B_i^\star$ die dazu duale Basis von $U_i^\star$. Dann ist die Familie*

$$C = (\pi(b_1^\star, \dots, b_n^\star))_{(b_1,\dots,b_n)\in B_1\times\cdots\times B_n}$$

eine Basis von $\mathrm{Mult}(U_1, \dots, U_n; K)$.

Beweis. Sei $\varphi \in \mathrm{Mult}(U_1, \dots, U_n; K)$. Wir zeigen nun:

$$\varphi = \sum_{(b_1,\dots,b_n)\in \bigtimes_{i=1}^n B_i} \varphi(b_1, \dots, b_n)\pi(b_1^\star, \dots, b_n^\star). \tag{13.2}$$

Wegen der Multilinearität beider Seiten reicht es, zu zeigen, dass die linke und die rechte Seite auf allen Punkten $(c_1, \dots, c_n) \in \bigtimes_{i=1}^n B_i$ den gleichen Wert annehmen. Tatsächlich gilt

$$\sum_{(b_1,\dots,b_n)\in \bigtimes_{i=1}^n B_i} \varphi(b_1, \dots, b_n)\pi(b_1^\star, \dots, b_n^\star)\,(c_1, \dots, c_n)$$
$$= \sum_{(b_1,\dots,b_n)\in \bigtimes_{i=1}^n B_i} \varphi(b_1, \dots, b_n)b_1^\star(c_1)\cdots b_n^\star(c_n) = \varphi(c_1, \dots, c_n), \tag{13.3}$$

und somit (13.2). Wegen (13.2) gilt $L(C) = \mathrm{Mult}(U_1, \dots, U_n; K)$. Für die lineare Unabhängigkeit wählen wir $\alpha : \bigtimes_{i=1}^n B_i \to K$ so, dass

$$\sum_{(b_1,\dots,b_n)\in \bigtimes_{i=1}^n B_i} \alpha(b_1, \dots, b_n)\pi(b_1^\star, \dots, b_n^\star) = 0.$$

Durch Auswerten an der Stelle $(c_1, \dots, c_n) \in \bigtimes_{i=1}^n B_i$ erhalten wir $\alpha(c_1, \dots, c_n) = 0$. $\square$

Korollar 13.32. *Seien $n \in \mathbb{N}$ und $U_1, \dots, U_n$ endlichdimensionale K-Vektorräume. Dann gilt* $\dim(\mathrm{Mult}(U_1, \dots, U_n; K)) = \prod_{i=1}^n \dim(U_i)$.

Definition 13.33. Seien $U_1, \dots, U_n$ K-Vektorräume, und sei $(u_1, \dots, u_n) \in U_1 \times \cdots \times U_n$. Dann definieren wir

$$u_1 \otimes \cdots \otimes u_n \in \mathrm{Mult}(U_1, \dots, U_n; K)^\star$$

durch

$$u_1 \otimes \cdots \otimes u_n\,(\varphi) := \varphi(u_1, \dots, u_n).$$

Das *Tensorprodukt* $U_1 \otimes \cdots \otimes U_n$ ist die lineare Hülle von

$$\{u_1 \otimes \cdots \otimes u_n \mid u_1 \in U_1, \dots, u_n \in U_n\}$$

in $\mathrm{Mult}(U_1, \dots, U_n; K)^\star$.

Satz 13.34. *Seien $U_1, \dots, U_n, V$ K-Vektorräume. Dann gilt:*

(1) *Die Abbildung $\bigotimes : \bigtimes_{i=1}^n U_i \to U_1 \otimes \cdots \otimes U_n$ mit*
$$\bigotimes(u_1, \dots, u_n) := u_1 \otimes \cdots \otimes u_n$$
ist eine multlineare Abbildung von $\bigtimes_{i=1}^n U_i$ nach $U_1 \otimes \cdots \otimes U_n$.

(2) *Für alle $(u_1, \dots, u_n), (v_1, \dots, v_n) \in \bigtimes_{i=1}^n U_i$ sind äquivalent:*
   (a) $u_1 \otimes \cdots \otimes u_n = v_1 \otimes \cdots \otimes v_n$.
   (b) *Für alle Vektorräume W und für alle $h \in \mathrm{Mult}(U_1, \dots, U_n, W)$ gilt $h(u_1, \dots, u_n) = h(v_1, \dots, v_n)$.*

(3) *Für jedes $h \in \mathrm{Mult}(U_1, \dots, U_n; V)$ gibt es ein $\hat{h} \in \mathrm{Hom}_K(U_1 \otimes \cdots \otimes U_n, V)$, sodass für alle $(u_1, \dots, u_n) \in \bigtimes_{i=1}^n U_i$ gilt:*
$$\hat{h}(u_1 \otimes \cdots \otimes u_n) = h(u_1, \dots, u_n).$$

Beweis. (1) Wir fixieren $(u_2, \dots, u_n) \in \bigtimes_{i=2}^n U_i$ und zeigen, dass $f(x) := x \otimes u_2 \otimes \cdots \otimes u_n$ linear ist. Seien dazu $x, y \in U_1$ und $\alpha \in K$. Für alle $\varphi \in \mathrm{Mult}(U_1, \dots, U_n; K)$ gilt

$$\begin{aligned}(x+y) \otimes u_2 \otimes \cdots \otimes u_n\,(\varphi) &= \varphi(x+y, u_2, \dots, u_n) = \varphi(x, u_2, \dots, u_n) + \varphi(y, u_2, \dots, u_n) \\ &= x \otimes u_2 \otimes \cdots \otimes u_n\,(\varphi) + y \otimes u_2 \otimes \cdots \otimes u_n\,(\varphi)\end{aligned}$$

und

$$\begin{aligned}(\alpha x) \otimes u_2 \otimes \cdots \otimes u_n\,(\varphi) &= \varphi(\alpha x, u_2, \dots, u_n) \\ &= \alpha\varphi(x, u_2, \dots, u_n) = (\alpha(x \otimes u_2, \dots, u_n))\,(\varphi).\end{aligned}$$

Die Linearität in den anderen Komponenten zeigt man genauso.

(2): (2a)⇒(2b) Um $h(u_1, \dots, u_n) = h(v_1, \dots, v_n)$ zu beweisen, zeigen wir, dass für alle $\gamma \in W^\star$ gilt, dass $\gamma(h(u_1, \dots, u_n)) = \gamma(h(v_1, \dots, v_n))$. Sei dazu $\gamma \in W^\star$. Dann gilt $\gamma \circ h \in \mathrm{Mult}(U_1, \dots, U_n; K)$. Nun gilt

$$(\gamma \circ h)\,(u_1, \dots, u_n) = (u_1 \otimes \cdots \otimes u_n)\,(\gamma \circ h) = (v_1 \otimes \cdots \otimes v_n)\,(\gamma \circ h) = (\gamma \circ h)\,(v_1, \dots, v_n).$$

Also gilt $h(u_1, \dots, u_n) = h(v_1, \dots, v_n)$.

(2b)⇒(2a) Wegen (1) gilt

$$\bigotimes \in \mathrm{Mult}(U_1, \dots, U_n; U_1 \otimes \cdots \otimes U_n)$$

und somit wegen (2b), dass

$$u_1 \otimes \cdots \otimes u_n = v_1 \otimes \cdots \otimes v_n.$$

(3) Sei

$$\hat{h} = \{\big((u_1 \otimes \cdots \otimes u_n), h(u_1, \dots, u_n)\big) \mid (u_1, \dots, u_n) \in \bigtimes_{i=1}^n U_i\}.$$

Wegen (2) ist $\hat{h}$ funktional. Die Menge $\hat{h}$ ist als Bild des Homomorphismus $H : \bigotimes_{i=1}^n U_i \to (\bigotimes_{i=1}^n U_i) \times V$, $H(\boldsymbol{u}) := (\boldsymbol{u}, v)$ ein Unterraum von $(\bigotimes_{i=1}^n U_i) \times V$. Unterräume, die gleichzeitg funktionale Relationen sind, sind Homomorphismen. Folglich ist $\hat{h}$ ein Homomorphismus. $\square$

Tensorprodukte können helfen, Algorithmen zum Multiplizieren von Matrizen zu verbessern. Wir fassen hier nur die Idee kurz zusammen und verweisen auf die detailliertere Darstellung in [**Kau23**].

Wenn wir zwei $n \times n$-Matrizen gemäß unserer Definition der Matrixmultiplikation multiplizieren, so führen wir n^3 Multiplikationen und $n^2(n-1)$ Additionen aus. Ziel ist, die Anzahl zu reduzieren. Dazu stellen wir die Matrixmultiplikation durch einen Tensor dar. Sei $V := K^{n \times n}$. Wir definieren auf V ein Skalarprodukt durch $\langle A, B\rangle := \sum_{i=1}^n \sum_{j=1}^n A_{i,j} B_{i,j}$. Mit $E^{(i,j)}$ bezeichen wir die Matrix, deren (i,j)-ter Eintrag 1 und deren andere Einträge alle 0 sind. Wir definieren nun eine multilineare Abbildung

$$\mu \in \mathrm{Mult}(V, V, V; \mathrm{Mult}(V, V; V))$$

durch

$$\mu(A, B, C)\,(X, Y) = (\langle A, X\rangle\langle B, Y\rangle) \cdot C.$$

Da μ multilinear ist, gibt es wegen Satz 13.34(3) eine lineare Abbildung

$$\nu : V \otimes V \otimes V \to \mathrm{Mult}(V, V; V),$$

sodass

$$\nu(A \otimes B \otimes C)(X, Y) = \mu(A, B, C)(X, Y)$$

für alle $A, B, C, X, Y \in V$. Sei nun $T = \sum_{k=1}^{l} A^{(k)} \otimes B^{(k)} \otimes C^{(k)} \in V \otimes V \otimes V$. Die zu T assoziierte Multiplikation von $n \times n$-Matrizen $\star_T$ ist definiert durch

$$X \star_T Y := \nu(T)\,(X, Y).$$

Es gilt dann

$$X \star_T Y = \sum_{k=1}^{l} (\langle A^{(k)}, X\rangle \langle B^{(k)}, Y\rangle) \cdot C^{(k)}$$

für alle $X, Y \in K^{n\times n}$. Die Multiplikation $\star_T$ hängt nicht von der Darstellung

$$(A^{(k)}, B^{(k)}, C^{(k)})_{k\in\{1,\ldots,l\}}$$

des Tensors ab; die die konkrete Darstellung liefert aber einen Algorithmus zur Berechnung von $\star_T$, der, falls alle $A^{(k)}, B^{(k)}, C^{(k)} \in \mathbb{Z}^{n\times n}$ sind, l Multiplikationen braucht.

Lemma 13.35. *Die mit dem Tensor $M_n := \sum_{a=1}^{n} \sum_{b=1}^{n} \sum_{c=1}^{n} E^{(a,b)} \otimes E^{(b,c)} \otimes E^{(a,c)}$ assoziierte Multiplikation ist die Matrixmultiplikation.*

Beweis. Es gilt

$$X \star_{M_n} Y = \sum_{a=1}^{n} \sum_{b=1}^{n} \sum_{c=1}^{n} (\langle E^{(a,b)}, X\rangle \langle E^{(b,c)}, Y\rangle) E^{(a,c)}.$$

$\square$

Wir betrachten nun den Fall $n = 2$. Sei der Tensor $S \in V \otimes V \otimes V$ gegeben durch

$$\begin{array}{rcccccc} S & = & (E^{(1,2)} - E^{(2,2)}) & \otimes & (E^{(2,1)} + E^{(2,2)}) & \otimes & E^{(1,1)} \\ & + & (E^{(1,1)} + E^{(1,2)}) & \otimes & E^{(2,2)} & \otimes & (-E^{(1,1)} + E^{(1,2)}) \\ & + & E^{(2,2)} & \otimes & (-E^{(1,1)} + E^{(2,1)}) & \otimes & (E^{(1,1)} + E^{(2,1)}) \\ & + & (E^{(2,1)} + E^{(2,2)}) & \otimes & E^{(1,1)} & \otimes & (E^{(2,1)} - E^{(2,2)}) \\ & + & (-E^{(1,1)} + E^{(2,1)}) & \otimes & (E^{(1,1)} + E^{(1,2)}) & \otimes & E^{(2,2)} \\ & + & (E^{(1,1)} + E^{(2,2)}) & \otimes & (E^{(1,1)} + E^{(2,2)}) & \otimes & (E^{(1,1)} + E^{(2,2)}) \\ & + & E^{(1,1)} & \otimes & (E^{(1,2)} - E^{(2,2)}) & \otimes & (E^{(1,2)} + E^{(2,2)}). \end{array} \tag{13.4}$$

Wenn man diesen Tensor mithilfe von Satz 13.34(1) ausmultipliziert, so erhält man

$$\begin{array}{rcccccc} S & = & E^{(1,1)} & \otimes & E^{(1,1)} & \otimes & E^{(1,1)} \\ & + & E^{(1,2)} & \otimes & E^{(2,1)} & \otimes & E^{(1,1)} \\ & + & E^{(1,1)} & \otimes & E^{(1,2)} & \otimes & E^{(1,2)} \\ & + & E^{(1,2)} & \otimes & E^{(2,2)} & \otimes & E^{(1,2)} \\ & + & E^{(2,1)} & \otimes & E^{(1,1)} & \otimes & E^{(2,1)} \\ & + & E^{(2,2)} & \otimes & E^{(2,1)} & \otimes & E^{(2,1)} \\ & + & E^{(2,1)} & \otimes & E^{(1,2)} & \otimes & E^{(2,2)} \\ & + & E^{(2,2)} & \otimes & E^{(2,2)} & \otimes & E^{(2,2)}. \end{array}$$

Also gilt $S = M_2$. Wie man auf diese, Strassens[1], Darstellung von M_2 in (13.4) kommt, ist keineswegs offensichtlich.

[1]Volker Strassen, *1936.

Somit ist die zu S gehörende bilineare Abbildung $\nu(S)$ die Matrixmultiplikation. Sie lässt sich nun mit

$$\begin{array}{rcl} m_1 & = & (a_{1,2} - a_{2,2}) \cdot (b_{2,1} + b_{2,2}) \\ m_2 & = & (a_{1,1} + a_{1,2}) \cdot b_{2,2} \\ m_3 & = & a_{2,2} \cdot (-b_{1,1} + b_{2,1}) \\ m_4 & = & (a_{2,1} + a_{2,2}) \cdot b_{1,1} \\ m_5 & = & (-a_{1,1} + a_{2,1}) \cdot (b_{1,1} + b_{1,2}) \\ m_6 & = & (a_{1,1} + a_{2,2}) \cdot (b_{1,1} + b_{2,2}) \\ m_7 & = & a_{1,1} \cdot (b_{1,2} - b_{2,2}) \end{array}$$

und

$$\begin{array}{rcl} c_{1,1} & = & m_1 - m_2 + m_3 + m_6 \\ c_{1,2} & = & m_2 + m_7 \\ c_{2,1} & = & m_3 + m_4 \\ c_{2,2} & = & -m_4 + m_5 + m_6 + m_7 \end{array}$$

ausrechnen. Somit brauchen wir nicht mehr 8, sondern 7 Multiplikationen und 8 Additionen/Subtraktionen. Wenn man nun zwei $2^k \times 2^k$-Matrizen miteinander multipliziert, so kann man diese Matrizen als 2×2-Matrizen über $K^{2^{k-1} \times 2^{k-1}}$ sehen und das oben angeführte Schema verwenden. Wenn $f(k)$ die Anzahl der Operationen dafür ist, so gilt

$$f(k) = 7f(k-1) + 18 \cdot 2^{2k-2} = 7f(k-1) + \frac{9}{2} \cdot 4^k.$$

Wir zeigen mit Induktion, dass dann $f(k) \le 7 \cdot 7^k - 6 \cdot 4^k$. Wegen $f(0) = 1$ gilt die Aussage für $k = 0$, und für $k \ge 1$ gilt $f(k) = 7f(k-1) + \frac{9}{2} \cdot 4^k \le 7 \cdot (7 \cdot 7^{k-1} - 6 \cdot 4^{k-1}) + \frac{9}{2} \cdot 4^k = 7 \cdot 7^k - \frac{42}{4} \cdot 4^k + \frac{9}{2} \cdot 4^k = 7 \cdot 7^k - 6 \cdot 4^k$. Also braucht man für die Multiplikation von zwei $n \times n$-Matrizen mit $n = 2^k$ höchstens $7\, n^{\log_2(7)} \le 7\, n^{2.808}$ Operationen.

Definition 13.36. Sei $T \in U_1 \otimes \cdots \otimes U_n$ ein Tensor. Der *Tensorrang* von T, $\operatorname{trk}(T)$, ist das minimale $k \in \mathbb{N}_0$, für das eine Familie $(u_1^{(j)}, \ldots, u_n^{(j)})_{j \in \{1,\ldots,k\}} \in (\bigtimes_{i=1}^n U_i)^k$ mit

$$T = \sum_{j=1}^{k} u_1^{(j)} \otimes \cdots \otimes u_n^{(j)}$$

existiert.

Der Tensorrang des Matrixmultiplikationstensors M_2 ist also wegen der Darstellung in (13.4) höchstens 7.

Satz 13.37. *Sei $n \in \mathbb{N}$, und seien $U_1, \ldots, U_n$ endlichdimensional Vektorräume, und sei $T \in U_1 \otimes \cdots \otimes U_n$. Dann gilt:*

(1) *Wenn $n = 1$, so gilt* $\operatorname{trk}(T) \le 1$.
(2) *Wenn $n = 2$, so sei $B = (b_1, \ldots, b_k)$ eine Basis von U_1, $C := (c_2, \ldots, c_m)$ eine Basis von U_2, und sei $M(T) \in K^{k \times m}$ definiert durch $M(T)[i,j] := T(\pi(b_i^\star, c_j^\star))$. Dann gilt* $\operatorname{trk}(T) = \operatorname{rk}(M(T))$.
(3) *Für alle $j \in \{1, \ldots, n\}$ gilt* $\operatorname{trk}(T) \le \prod_{\substack{i=1 \\ i \ne j}}^{n} \dim(U_i)$.

Beweis. (1) Es gilt $\operatorname{Hom}(U_1, K)^\star = U_1^{\star\star}$. Die Abbildung $h : U_1 \to U_1^{\star\star}$ mit $h(u)(\varphi) := \varphi(u)$ für $u \in U_1$, $\varphi \in U_1^\star$ ist injektiv und daher wegen $\dim(U_1) = \dim(U_1^{\star\star})$ surjektiv. Somit gibt es ein $u_1 \in U_1$ mit $T = h(u_1)$. Dann gilt $T = \bigotimes_{i=1}^{1} u_i$. Also gilt $\operatorname{rk}(f) \le 1$.

(2) Sei $\operatorname{trk}(T) = s$. Dann gibt es $u_1, \ldots, u_s \in U_1$ und $v_1, \ldots, v_s \in U_2$ mit $T = \sum_{i=1}^{s} u_i \otimes v_i$. Nun gilt für $u \in U$ und $v \in V$, dass $M(u \otimes v)[i,j] = u \otimes v\,(\pi(b_i^\star, c_j^\star)) = \pi(b_i^\star, c_j^\star)\,(u \otimes v) = b_i^\star(u) \cdot c_j^\star(v) = (u)_B[i] \cdot (v)_C[j]$. Folglich gilt

$$M(u \otimes v) = \begin{pmatrix} (u)_B[1] \\ \vdots \\ (u)_B[k] \end{pmatrix} \cdot ((u)_C[1] \cdots (u)_C[m])$$

und daher $\operatorname{rk}(M(u \otimes v)) \leq 1$. Somit gilt $M(T) = \sum_{i=1}^{s} M(u_i \otimes v_i)$ und folglich $\operatorname{rk}(M(T)) \leq s$. Das beweist $\operatorname{rk}(M(T)) \leq \operatorname{trk}(T)$.

Für die andere Ungleichung nehmen wir an, dass $\operatorname{rk}(M(T)) = r$. Dann gibt es $X \in K^{k \times r}$ und $Y \in K^{r \times m}$ mit $X \cdot Y = M(T)$. Sei $Y_i \in K^{r \times m}$ die Matrix, deren i-te Zeile die i-te Zeile von Y ist, und deren andere Zeilen 0 sind. Dann gilt $X \cdot Y = \sum_{i=1}^{r} X \cdot Y_i$. Für $i \in \{1, \ldots, r\}$ lässt sich die Matrix $X \cdot Y_i$ als $x_i \cdot y_i^T$ schreiben, wobei x_i die i-te Spalte von X und y_i^T die i-te Zeile von Y ist. Somit gilt $X \cdot Y = \sum_{i=1}^{r} x_i \cdot y_i^T$. Wenn wir nun $u^{(i)}$ und $v^{(i)}$ so wählen, dass $(u^{(i)})_B = x_i$ und $(v^{(i)})_C = y_i$, dann gilt $M(u^{(i)} \otimes v^{(i)}) = x_i \cdot y_i^T$ und somit $M(T) = M(\sum_{i=1}^{r} u^{(i)} \otimes v^{(i)})$, also $T = \sum_{i=1}^{r} u^{(i)} \otimes v^{(i)}$ und folglich $\operatorname{trk}(T) \leq r$. Das beweist $\operatorname{trk}(T) \leq \operatorname{rk}(M(T))$.

(3) Sei B_i eine Basis von U_i. T lässt sich als Linearkombination der $\dim(U_1) \cdots \dim(U_n)$ Tensoren $b_1 \otimes \cdots \otimes b_n$ schreiben. Dadurch erhalten wir $\dim(U_1) \cdots \dim(U_n)$ Summanden. Durch Zusammenfassen der Tensoren mit gleichem $b_1, \ldots, b_{j-1}, b_{j+1}, \ldots, b_n$ und Herausheben erhält man $\dim(U_1) \cdots \dim(U_{j-1}) \cdot \dim(U_{j+1}) \cdots \dim(U_n)$ Summanden. $\square$

Literaturverzeichnis

[BT09] M. Bramanti and G. Travaglini. *Matematica. Questione di metodo*. Zanichelli, Bologna, 2009.

[FS25] G. Fischer and B. Springborn. *Lineare Algebra. Eine Einführung für Studienanfänger*. Grundkurs Math. Berlin: Springer Spektrum, 20th edition, 2025. ISBN 978-3-662-71260-3; 978-3-662-71261-0. ISSN 2626-613X.

[Hal76] P. R. Halmos. *Naive Mengenlehre*. Vandenhoeck & Ruprecht, Göttingen, 1976. 132 pp. Vierte Auflage, Aus dem Englischen übersetzt von Manfred Armbrust und Fritz Ostermann, Moderne Mathematik in elementarer Darstellung, No. 6.

[Jän08] K. Jänich. *Lineare Algebra*. Springer-Lehrbuch. Berlin: Springer, 11th edition, 2008. ISBN 978-3-540-75501-2. ISSN 0937-7433.

[Kau23] M. Kauers. Lineare Algebra und Analytische Geometrie. Vorlesungsskriptum, `http://www.algebra.uni-linz.ac.at/people/mkauers/linalg/script.pdf`, September 2023.

[KS99] Karl-Heinz Kiyek and Friedrich Schwarz. *Lineare Algebra*. Teubner Studienbücher Mathematik. [Teubner Mathematical Textbooks]. B. G. Teubner, Stuttgart, 1999. ISBN 3-519-02390-3. 320 pp.

[Smi20] P. Smith. *An introduction to formal logic*. Camb. Introd. Philos. Cambridge: Cambridge University Press, 2nd edition edition, 2020. ISBN 978-1-108-41139-4; 978-1-108-42006-8; 978-1-108-32899-9.

[Wei80] P. Weiß. *Lineare Algebra und analytische Geometrie*. Rudolf Trauner, Linz, 1980. ISBN 3-85320-221-7. iv+386 pp. Eine anwendungsbezogene Einführung. [An application oriented introduction].

ANHANG A

Programme, die vorrechnen

Die Mathematica-Files `GaussDemo7.m` und `RowRed12.m` enthalten Mathematica-Funktionen, die folgende Probleme mit Zwischenschritten vorrechnen:

- ▷ Lösen eines linearen Gleichungssystems (`Gauss[A,b]`).
- ▷ Bestimmen einer Matrix in Treppenform, die den gleichen Zeilenraum wie die eingegebene Matrix hat (`RowEchelonForm[A]`).
- ▷ Bestimmen einer Matrix in Treppennormalform, 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 `<< GaussDemo7.m` und `<< RowRed12.m` geladen werden. Sie sind auf `http://www.algebra.uni-linz.ac.at/Students/MathInf/vlws05/MathematicaProgramme/` erhältlich.