# Lineare Algebra (Informatik)

Dr. Paula Reichert

22. Dezember 2023

## Inhaltsverzeichnis

# 1 Einführung

Die (lineare) Algebra beschäftigt sich mit dem Lösen von (linearen) Gleichungssystemen und den damit verbundenen Rechenoperationen.

Herkunft des Begriffs: Das Wort „Algebra" stammt aus dem Arabischen von *al-ǧabr*, zu deutsch in etwa „Ergänzen". Das Wort war ursprünglich Teil des Titels eines berühmten Mathematikbuchs von *al-Chwarizmi* (9. Jhd., Bagdad). Der vollständige Titel lautet:

*al-Kitab al-Muchtasar fi hisab al-dschabr wa-l-muqabala*
„Das kurz gefasste Buch über die Rechenverfahren durch Ergänzen und Ausgleichen"

## 1.1 Historische Einführung

Frühgeschichte in Babylon (Gebiet des heutigen Irak) und Ägypten:

Babylonier (ca. 2000 v. Chr.): Gleichungssysteme der folgenden Form werden in erster Näherung gelöst. Hier sind $a, b$ gegeben, $x, y$ gesucht:

$$\left.\begin{array}{c} x + y = a \\ x \cdot y = b \end{array}\right\} \Rightarrow x^2 - ax + b = 0$$

Dies entspricht dem Lösen quadratischen Gleichungen, aber: keine irrationalen Zahlen, keine negativen Zahlen, nur erste Näherung.

Ägypter (ca. 1700 v. Chr.): Lineare Gleichungen der Form

$$x + a \cdot x = b$$
$$x + c \cdot x + d \cdot x = e$$

mit $a, b, c, d, e$ gegeben, $x$ gesucht, werden mit geometrischen Methoden gelöst.

Höhepunkt der Entwicklung im antiken Griechenland:

Pythagoreer (ca. 600 v. Chr.): Entdeckung der Existenz inkommensurabler Strecken und damit irrationaler Zahlen durch *Hippasos von Metapont*.

Euklid (ca. 300 v. Chr.): Verfasst die 13 Bücher der „Elemente". Dies bleibt das Standardwerk der Geometrie und Arithmetik bis ins 19. Jhd (und das am zweitmeisten verbreitetste Buch nach der Bibel). Euklid wird u.a. der Beweis der Irrationalität von $\sqrt{2}$ zugeschrieben.

Diophantes (zwischen 100 v. Chr. und 300 n. Chr.): Verfasst die 13 Bücher der „Arithmetica", von denen nur noch sechs überliefert sind. Löst die Arithmetik und Algebra von der Geometrie ab.

Es folgt eine Weiterentwicklung der Mathematik im arabisch-persischen und indischen Raum mit Höhepunkt im 7. - 9. Jahrhundert.

Moderne Entwicklung:
Erst in der Renaissance im Italien des 16. Jahrhunderts kommt es wieder zu einer Weiterentwicklung der Mathematik (und aller anderer Wissenschaften) im europäischen Raum. Ab dem 18. Jhd. bestimmen bedeutende Mathematiker wie Euler, Lagrange, Gauss, Galois, Abel, Jordan, Hölder, Lie, Killing, Cartan, Grassmann, Hamilton, Clifford, Cauchy, Ricci, Levi-Civita die weitere Entwicklung.

Aussicht für diese Vorlesung (kurzes Intermezzo): Wir werden uns in diesem Semester mit linearen Gleichungssytemen der folgenden Art beschäftigen. Hier sind $a_{11}, ..., a_{mn}$ und $y_1, ..., y_m$ gegeben und $x_1, ..., x_n$ gesucht.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= y_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= y_m \end{aligned}$$

Dieses Gleichungssystem kann man auch schreiben als:

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Oder in Kurzschreibweise:

$$A \cdot \mathbf{x} = \mathbf{y}$$

Dabei ist $A = (a_{ij})_{\substack{i=1,...,m, \\ j=1,...,n}}$ eine $m \times n$ Matrix ($m$ Zeilen, $n$ Spalten), $\mathbf{x}$ ist ein Spaltenvektor mit $n$ Komponenten und $\mathbf{y}$ ein Spaltenvektor mit $m$ Komponenten.

## 1.2 Euklidische Mathematik

Im antiken Griechenland (vertreten insbesondere durch Euklid) basieren Zahlen und Rechenoperationen auf der geometrischen Anschauung. Mathematische Beweise werden über die Geometrie geführt. Wir betrachten einen Beweis der Irrationalität von $\sqrt{2}$.

**Definition 1.2.1** (Inkommensurabilität)**.** Zwei Strecken heißen *inkommensurabel*, wenn sie kein gemeinsames Maß besitzen. Sie heißen *kommensurabel*, wenn sie ein gemeinsames Maß besitzen.

Besitzen zwei Strecken ein gemeinsames Maß, so besitzen sie ein größtes gemeinsames Maß und dieses ist eindeutig (Beweis: Wechselwegnahme/euklidischer Algorithmus).

**Definition 1.2.2** (ggT)**.** Das größte gemeinsame Maß zweier natürlicher Zahlen ist der ggT (*größter gemeinsamer Teiler*).

Der *Euklidische Algorithmus* zur Konstruktion des größten gemeinsamen Maßes ist die sog. Wechselwegnahme. Betrachte dazu zwei Strecken X und Y.

Sei o. B. d. A. $Y < X$. Um das gemeinsame Maß (bzw. bei natürlichen Zahlen $X, Y$ den ggT) zu bestimmen, trage $Y$ so oft auf $X$ ab, bis ein Rest $r_1$ bleibt, $0 < r_1 < Y$. Wenn $r_1 = 0$, dann gilt: $Y = ggT(X, Y)$. Wenn $r_1 > 0$, trage $r_1$ auf $Y$ so oft ab, bis ein Rest $r_2$ bleibt, $0 < r_2 < r_1$. Wenn $r_2 = 0$, dann gilt: $r_1 = ggT(X, Y)$. Wenn $r_2 > 0$, trage $r_2$ auf $r_1$ so oft ab, bis ein Rest $r_3$ bleibt usw.
Damit ergibt sich das folgendes Gleichungssystem (mit $n_i \in \mathbb{N}\ \forall i = 1, ..., k$, setze $Y = r_0$):

$$\begin{aligned} X &= n_0 Y + r_1 \\ Y &= n_1 r_1 + r_2 \\ r_1 &= n_2 r_2 + r_3 \\ &\vdots \\ r_{k-1} &= n_k r_k + r_{k+1} \end{aligned}$$

Falls $r_{k+1} = 0 \Rightarrow r_k = ggT(X, Y)$.

Die Wechselwegnahme ist eine eindeutiges Charakterisierung des Verhältnisses $X/Y$ durch die natürlichen Zahlen $n_0, n_1, ..., n_k$:

$$\frac{X}{Y} = n_0 + \frac{r_1}{Y} = n_0 + \frac{1}{\frac{Y}{r_1}} = n_0 + \frac{1}{n_1 + \frac{r_2}{r_1}} = n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \frac{r_3}{r_2}}} = ... = n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{... + \frac{1}{n_k}}}}$$

Falls der Kettenbruch abbricht, sind $X$ und $Y$ kommensurabel. Falls nicht, sind $X$ und $Y$ inkommensurabel.

**Definition 1.2.3** (Irrationale Zahl)**.** Das Verhältnis zweier inkommensurabler Strecken ist eine *irrationale* Zahl, das Verhältnis zweier kommensurabler Strecken eine *rationale* Zahl.

**Beispiel 1.2.1.** Kettenbruchdarstellung von $\frac{5}{3}$ und $ggT(5, 3)$.

$$\begin{aligned} 5 &= 1 \cdot 3 + 2 \\ 3 &= 1 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0 \end{aligned}$$

$$\Rightarrow \frac{5}{3} = 1 + \frac{1}{1 + \frac{1}{2}}$$

und $ggT(5, 3) = 1$.

Wir betrachten nun den Beweis der Inkommensurabilität von Seite und Diagonale im Quadrat. Daraus folgt die Irrationalität von $\sqrt{2}$.

**Satz 1.2.1** (Irrationalität von $\sqrt{2}$)**.** *Seite und Diagonale eines Quadrats sind inkommensurabel.*

*Beweis.* Mittels Wechselwegnahme trage man die Seite $a$ des Quadrats an der Diagonalen $d$ ab. Dabei bleibt ein Rest $a_1$ (d.h. $d = a + a_1$). Dieser bildet nach Konstruktion die Seite eines kleineren Quadrats mit Diagonale $d_1$ (nach Konstruktion bildet $d_1$ einen Abschnitt der Seite $a$ mit $a = a_1 + d_1$). Man trage nun $a_1$ an $d_1$ ab. Dabei bleibt ein Rest $a_2$ (d.h. $a = 2a_1 + a_2$). Dieser bildet nach Konstruktion die Seite eines noch kleineren Quadrats mit Diagonale $d_2$ usw. Diese Konstruktion kann endlos fortgesetzt werden; sie bricht niemals ab. Zur geometrischen Konstruktion: siehe Tafel bzw. Übung.

Insbesondere gilt gemäß Konstruktion (mit $k \in \mathbb{N}_0$, $a = a_0$, $d = d_0$):

$$\begin{aligned} a_1 &= d_0 - a_0 \\ d_1 &= a_0 - a_1 \\ &\vdots \\ a_k &= d_{k-1} - a_{k-1} \\ d_k &= a_{k-1} - a_k \\ &\vdots \end{aligned}$$

Für die Wechselwegnahme gilt gemäß Konstruktion (mit $r_k = a_k$):

$$\begin{aligned} d_0 &= 1 \cdot a_0 + a_1 \\ a_0 &= 2 \cdot a_1 + a_2 \\ a_1 &= 2 \cdot a_2 + a_3 \\ &\vdots \\ a_{k-2} &= 2 \cdot a_{k-1} + a_k \\ &\vdots \end{aligned}$$

Daraus folgt:

$$\frac{d}{a} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \dots}}}$$

□

Um zu sehen, dass diese Kettenbruchdarstellung der Zahl $\sqrt{2}$ entspricht, brauchen wir den Satz von Pythagoras.

**Satz 1.2.2** (Satz des Pythagoras)**.** *Seien $a, b, c$ die Seiten eines rechtwinkligen Dreiecks. Dabei sei $c$ die Seite, die dem rechten Winkel gegenüberliegt. Dann gilt:*

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

*Beweis.* Z. B. von Euklid (Übung). □

Aus dem Satz von Pythagoras folgt, dass im gleichschenkligen rechtwinkligen Dreieck gilt: $a^2 + a^2 = d^2$, also $2a^2 = d^2$ bzw. $2 = \frac{d^2}{a^2}$. Demnach ist $\sqrt{2} \equiv \frac{d}{a}$. Wir haben also mittels Satz 1.2.1 und

Satz 1.2.2 bewiesen, dass $\sqrt{2}$ irrational ist und dass gilt:

$$\sqrt{2} \equiv 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2+\ldots}}}$$

Bricht man diesen Kettenbruch (die Iteration bzw. Rekursion) an einer bestimmten Stelle $k \in \mathbb{N}_0$ ab, indem man willkürlich $a_{k+1} = 0$ setzt, so erhält man eine Näherung von $\sqrt{2}$ (der $k$-ten Ordnung):

z.B.
$$\begin{aligned}
k = 0: \quad & \frac{d}{a} = 1 \\
k = 1: \quad & \frac{d}{a} = 1 + \frac{1}{2} = \frac{3}{2} = 1,5 \\
k = 2: \quad & \frac{d}{a} = 1 + \frac{1}{2 + \frac{1}{2}} = \frac{7}{5} = 1,4 \\
k = 3: \quad & \frac{d}{a} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}} = \frac{17}{12} = 1,41\bar{6}
\end{aligned}$$

# 2 Grundbegriffe

## 2.1 Logik

Die Aussagenlogik untersucht die Struktur von Sätzen (Aussagen) $\mathcal{A}, \mathcal{B}, ...$ hinsichtlich ihres Wahrheitswerts. Sätze (Aussagen) können entweder *wahr* $(w)$ oder *falsch* $(f)$ sein.

**Beispiel 2.1.1.** Wahrheitswert von Aussagen

$\mathcal{A}$ : 4 ist eine gerade Zahl $(w)$

$\mathcal{B}$ : 15 ist durch 7 teilbar $(f)$

$\mathcal{C}$ : Seite und Diagonale des Quadrats sind inkommensurabel $(w)$

Die Aussagenlogik arbeitet mit *Junktoren* (logischen Operatoren):

$\wedge$ : „und"

$\vee$ : „oder" (Achtung: *nicht* „entweder ... oder ..."!)

$\neg$ : „nicht"

$\Rightarrow$: „wenn ..., dann ..." bzw. „aus ... folgt ..."

$\Leftrightarrow$: „genau dann, wenn"

Mit Hilfe von Junktoren erhalten wir zusammengesetzte Aussagen, die wiederum auf ihren Wahrheitsgehalt untersucht werden können. Dies ist wichtig für die mathematische Beweisführung. Wir untersuchen den Wahrheitsgehalt mit Hilfe von *Wahrheitstafeln*.

**Beispiel 2.1.2.** Wahrheitstafel für $\mathcal{A} \wedge \mathcal{B}$ und $\mathcal{A} \vee \mathcal{B}$

| $\mathcal{A}$ | $\mathcal{B}$ | $\mathcal{A} \wedge \mathcal{B}$ | $\mathcal{A} \vee \mathcal{B}$ |
|---|---|---|---|
| $w$ | $w$ | $w$ | $w$ |
| $w$ | $f$ | $f$ | $w$ |
| $f$ | $w$ | $f$ | $w$ |
| $f$ | $f$ | $f$ | $f$ |

Für die Junktoren gelten bestimmte Rechenregeln:

**Satz 2.1.1** (Rechenregeln)**.** *Für $\wedge$ und $\vee$ gelten das Kommutativgesetz, das Assoziativgesetz und das Distributivgesetz,*

$$\begin{aligned} \mathcal{A} \wedge \mathcal{B} &\Leftrightarrow \mathcal{B} \wedge \mathcal{A} \\ (\mathcal{A} \wedge \mathcal{B}) \wedge \mathcal{C} &\Leftrightarrow \mathcal{A} \wedge (\mathcal{B} \wedge \mathcal{C}) \\ (\mathcal{A} \vee \mathcal{B}) \wedge \mathcal{C} &\Leftrightarrow (\mathcal{A} \wedge \mathcal{C}) \vee (\mathcal{B} \wedge \mathcal{C}) \end{aligned}$$

*und analog für $\vee$ (bzw. $\vee$ und $\wedge$ vertauscht im DG).*
*Weiter gelten die* De-Morgan-Regeln*:*

$$\begin{aligned} \neg(\mathcal{A} \wedge \mathcal{B}) &\Leftrightarrow \neg\mathcal{A} \vee \neg\mathcal{B} \\ \neg(\mathcal{A} \vee \mathcal{B}) &\Leftrightarrow \neg\mathcal{A} \wedge \neg\mathcal{B} \end{aligned}$$

*Zuletzt gilt für die doppelte Verneinung:*

$$\neg(\neg\mathcal{A}) \Leftrightarrow \mathcal{A}$$

*Beweis.* Mittels Wahrheitstafeln zeigt man jeweils die Äquivalenz der links und rechts von „⇔" stehenden Aussagen (Übung). □

**Definition 2.1.1** (Äquivalenz)**.** Zwei Aussagen $\mathcal{A}_1$ und $\mathcal{A}_2$ heißen (logisch) *äquivalent*, wenn sie dieselben Einträge in der Wahrheitstafel besitzen. Sind zwei Aussagen $\mathcal{A}_1$ und $\mathcal{A}_2$ äquivalent, dann ist die Aussage $\mathcal{A}_1 \Leftrightarrow \mathcal{A}_2$ eine Tautologie.

**Definition 2.1.2** (Tautologie)**.** Eine Aussage nennt man eine *Tautologie* (= allgemein gültige Aussage), wenn alle ihre Einträge in der Wahrheitstafel wahr ($w$) sind. Eine Aussage nennt man ein *Paradoxon*, wenn alle ihre Einträge in der Wahrheitstafel falsch ($f$) sind.

**Beispiel 2.1.3.** Tautologie $\neg\mathcal{A} \vee \mathcal{A}$ und Paradoxon $\neg\mathcal{A} \wedge \mathcal{A}$

| $\mathcal{A}$ | $\neg\mathcal{A}$ | $\neg\mathcal{A} \vee \mathcal{A}$ | $\neg\mathcal{A} \wedge \mathcal{A}$ |
|---|---|---|---|
| $w$ | $f$ | $w$ | $f$ |
| $f$ | $w$ | $w$ | $f$ |

**Satz 2.1.2.** $(\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A})$ *und* $\mathcal{A} \Leftrightarrow \mathcal{B}$ *sind äquivalent.*

*Beweis.* Wir beweisen dies mittels Wahrheitstafel:

| $\mathcal{A}$ | $\mathcal{B}$ | $\mathcal{A} \Rightarrow \mathcal{B}$ | $\mathcal{B} \Rightarrow \mathcal{A}$ | $(\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A})$ | $\mathcal{A} \Leftrightarrow \mathcal{B}$ | $((\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A})) \Leftrightarrow (\mathcal{A} \Leftrightarrow \mathcal{B})$ |
|---|---|---|---|---|---|---|
| $w$ | $w$ | $w$ | $w$ | $w$ | $w$ | $w$ |
| $w$ | $f$ | $f$ | $w$ | $f$ | $f$ | $w$ |
| $f$ | $w$ | $w$ | $f$ | $f$ | $f$ | $w$ |
| $f$ | $f$ | $w$ | $w$ | $w$ | $w$ | $w$ |

Zum Erstellen dieser Wahrheitstafel ist es wichtig zu erkennen, dass die Implikation $\mathcal{A} \Rightarrow \mathcal{B}$ nicht bedeutet, dass $\mathcal{B}$ wahr sein muss, sondern lediglich dass, *wenn* $\mathcal{A}$ wahr ist, *dann* $\mathcal{B}$ wahr ist. Darüber hinaus gilt die allgemeine Regel: aus Falschem kann alles folgen (*ex falso quodlibet*). □

**Bemerkung 2.1.1.** Die Äquivalenz von $(\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A})$ und $\mathcal{A} \Leftrightarrow \mathcal{B}$ ist wichtig, um Beweise zu führen. Wir beweisen in der Regel, dass $\mathcal{A} \Leftrightarrow \mathcal{B}$, indem wir $\mathcal{A} \Rightarrow \mathcal{B}$ und $\mathcal{B} \Rightarrow \mathcal{A}$ zeigen.

**Satz 2.1.3** (Kontrapositionsgesetz)**.** *Seien* $\mathcal{A}, \mathcal{B}$ *Aussagen. Dann ist allgemein gültig:*

$$(\mathcal{A} \Rightarrow \mathcal{B}) \Leftrightarrow (\neg\mathcal{B} \Rightarrow \neg\mathcal{A})$$

*Beweis.* Mittels Wahrheitstafel (Übung). □

**Beispiel 2.1.4.** Wir nehmen zur Veranschaulichung ein Alltagsbeispiel (auch wenn Alltagsbeispiele nie ganz präzise sind). Sei $\mathcal{A}$ die Aussage „der Junge trifft mit dem Ball die Scheibe" und $\mathcal{B}$ die Aussage „die Scheibe zerbricht". Dann gilt: $(\mathcal{A} \Rightarrow \mathcal{B})$. Dies impliziert aber auch $(\neg\mathcal{B} \Rightarrow \neg\mathcal{A})$, denn „wenn die Scheibe nicht gebrochen ist, folgt, dass der Junge mit dem Ball nicht die Scheibe getroffen hat" und andersherum.

**Bemerkung 2.1.2.** Statt eine Aussage *direkt* zu beweisen, können wir sie also auch durch *Kontraposition* beweisen.

**Bemerkung 2.1.3** (Sprechweise)**.** Statt $\mathcal{A} \Leftrightarrow \mathcal{B}$ sagen wir auch „$\mathcal{A}$ ist *hinreichend und notwendig* für $\mathcal{B}$" bzw. „$\mathcal{B}$ ist *hinreichend und notwendig* für $\mathcal{A}$". Statt $\mathcal{A} \Rightarrow \mathcal{B}$ sagen wir auch „$\mathcal{A}$ ist *hinreichend* für $\mathcal{B}$" oder äquivalent „$\mathcal{B}$ ist *notwendig* für $\mathcal{A}$" (siehe Kontraposition).

Achtung: Die umgekehrte Implikation $\neg A \Rightarrow \neg B$ ist nicht äquivalent zu der Aussage $\mathcal{A} \Rightarrow \mathcal{B}$!

**Satz 2.1.4.** *Es ist* nicht *allgemein gültig, dass*

$$(\mathcal{A} \Rightarrow \mathcal{B}) \Leftrightarrow (\neg\mathcal{A} \Rightarrow \neg\mathcal{B})$$

*Beweis.* Wir geben ein Gegenbeispiel an. Sei $\mathcal{A}$ die Aussage „eine ganze Zahl $k$ ist durch 9 teilbar" und $\mathcal{B}$ die Aussage „eine ganze Zahl $k$ ist durch 3 teilbar". Dann gilt: $\mathcal{A} \Rightarrow \mathcal{B}$ (die linke Seite von „$\Leftrightarrow$" ist also wahr). Es gilt aber nicht: $\neg\mathcal{A} \Rightarrow \neg\mathcal{B}$, denn wenn eine Zahl $k$ nicht durch 9 teilbar ist, folgt daraus nicht, dass sie nicht durch 3 teilbar ist, betrachte z.B. $k = 6$ (die rechte Seite von „$\Leftrightarrow$" ist also falsch und damit auch die Implikation $(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\neg\mathcal{A} \Rightarrow \neg\mathcal{B})$). □

**Bemerkung 2.1.4** (Beweisverfahren allgemein)**.** Die Aussage $\mathcal{A} \Rightarrow \mathcal{B}$ beweisen wir in der Regel durch einen Widerspruchsbeweis. Wir nehmen dazu an, dass $\mathcal{A}$ gilt, aber $\mathcal{B}$ nicht gilt, formal $\mathcal{A} \wedge \neg\mathcal{B}$, und zeigen dann mittels korrekter Schlüsse, dass dies zu einem Widerspruch führt.

**Beispiel 2.1.5.** Sei $\mathcal{A}$ wieder die Aussage „eine ganze Zahl $k$ ist durch 9 teilbar" und $\mathcal{B}$ die Aussage „eine ganze Zahl $k$ ist durch 3 teilbar". Wir nehmen nun an, dass es eine ganze Zahl $k$ gibt, die durch 9 teilbar ist, aber nicht durch 3, also $\mathcal{A} \wedge \neg\mathcal{B}$. Wenn aber $k$ durch 9 teilbar ist, dann ist $k = m \cdot 9$ mit $m \in \mathbb{Z}$. Dies gilt genau dann, wenn $k = (3 \cdot m) \cdot 3 = p \cdot 3$ mit $p \in \mathbb{Z}$. Somit ist $k$ durch 3 teilbar und dies ist im Widerspruch zur Annahme.

Weitere wichtige Schlussregeln für Beweise ergeben sich aus den folgenden Eigenschaften von „$\Rightarrow$" und „$\Leftrightarrow$" (auch *Transitivität* von „$\Rightarrow$" und „$\Leftrightarrow$" genannt).

**Satz 2.1.5.** *Seien $\mathcal{A}, \mathcal{B}, \mathcal{C}$ Aussagen. Dann ist allgemein gültig:*

$$\begin{aligned} (\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{C}) &\Rightarrow (\mathcal{A} \Rightarrow \mathcal{C}) \\ (\mathcal{A} \Leftrightarrow \mathcal{B}) \wedge (\mathcal{B} \Leftrightarrow \mathcal{C}) &\Rightarrow (\mathcal{A} \Leftrightarrow \mathcal{C}) \end{aligned}$$

*Beweis.* Mittels Wahrheitstafeln (Übung). □

**Bemerkung 2.1.5.** Dieser Satz kann (später) mittels vollständiger Induktion auf eine beliebige Anzahl an Aussagen $\mathcal{A}_1, ..., \mathcal{A}_n$, $n \in \mathbb{N}$, erweitert werden.

Zum Abschluss noch ein Exkurs von der Aussagen- zur Prädikatenlogik. In der Logik unterscheiden wir formal die Aussagenlogik von der Prädikatenlogik. Wechselt man zur Prädikatenlogik kommen zu den Junktoren noch (zwei) *Quantoren* dazu:

$\exists$ : „es existiert"

$\forall$ : „für alle"

Aussagen werden nun durch *Prädikate* $\mathcal{P}, \mathcal{Q}, ...$ ausgedrückt. Ist z.B. $\mathcal{P}$ das Prädikat „... ist grün" und $\mathcal{Q}$ das Prädikat „... ist stachlig", dann entspricht der Term $\exists x : \mathcal{P}(x) \wedge \mathcal{Q}(x)$ der Aussage „es existiert ein $x$, das grün und stachlig ist". Damit sind wir bereits sehr nah an der Mengenlehre.

Allgemein gültige Regeln zur Verneinung der Quantoren:

$$\neg(\forall x : \mathcal{P}(x)) \quad \Leftrightarrow \quad \exists x : \neg\mathcal{P}(x)$$

$$\neg(\exists x : \mathcal{P}(x)) \quad \Leftrightarrow \quad \forall x : \neg\mathcal{P}(x)$$

In der Mengenlehre werden wir in der Regel nicht mehr (wie hier) mit *universellen* Quantoren arbeiten, sondern mit *Quantoren über Mengen*. Diese sind wie folgt definiert. Sei $M$ eine Menge.

$$\forall x \in M : \mathcal{P}(x) \quad :\Leftrightarrow \quad \forall x : (x \in M \Rightarrow \mathcal{P}(x))$$

$$\exists x \in M : \mathcal{P}(x) \quad :\Leftrightarrow \quad \exists x : (x \in M \wedge \mathcal{P}(x))$$

**Bemerkung 2.1.6.** Wichtig: Die Reihenfolge von Quantoren darf nicht vertauscht werden!

**Beispiel 2.1.6** (Reihenfolge Quantoren)**.** Die Aussage $\forall k \in \mathbb{Z} : \exists l \in \mathbb{Z} : l \leq k$ ist wahr, denn für jede ganze Zahl $k \in \mathbb{Z}$ gibt es eine Zahl $l \in \mathbb{Z}$ mit $l \leq k$. Setze z.B. $l := k$ oder $l := k - 1$. Anders herum ist die Aussage $\exists l \in \mathbb{Z} : \forall k \in \mathbb{Z} : l \leq k$ falsch, denn es existiert keine Zahl $l \in \mathbb{Z}$, die kleiner oder gleich jeder ganzen Zahl $k \in \mathbb{Z}$ ist ($\mathbb{Z}$ hat kein kleinstes Element!).

## 2.2 Vollständige Induktion

Ein sehr nützliches Beweisverfahren ist die *vollständige Induktion*. Wir benutzen sie, um zu zeigen, dass eine Aussage $\mathcal{A}(n)$ für alle $n \in \mathbb{N}$ gilt (oder für alle $n \in \mathbb{N}_0$, je nach Aufgabe). Dabei ist $\mathcal{A}(n)$ in der Regel eine Gleichung, in der ein freier Parameter $n \in \mathbb{N}$ vorkommt. Dafür müssen wir zwei Dinge zeigen:

1) $\mathcal{A}(1)$ ist wahr (Induktionsanfang)

2) $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ ist wahr (Induktionsschritt)

Daraus folgt: $\mathcal{A}(n)$ ist wahr $\forall n \in \mathbb{N}$

Ausführlicher: Schritt 2) ist der Beweis dafür, dass, wenn $A(n)$ wahr ist (Induktionsannahme), $A(n+1)$ wahr ist (Induktionsbehauptung).

**Bemerkung 2.2.1.** Falls $n \in \mathbb{N}_0$, dann beginnen wir beim Induktionsanfang mit $n = 0$, also zeigen in 1), dass $\mathcal{A}(0)$ wahr ist, bevor wir in 2) zeigen, dass $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ wahr ist.

Es ist wichtig, bei 1) mit dem ersten nicht-trivialen Fall der Aussage zu beginnen. Dann können wir von $\mathcal{A}(1)$ aus mit Hilfe von 2) iterativ die Wahrheit der nachfolgenden Aussagen feststellen: $\mathcal{A}(1)$ wahr $\Rightarrow$ $\mathcal{A}(2)$ wahr $\Rightarrow$ $\mathcal{A}(3)$ wahr $\Rightarrow$ ... So können wir letztlich darauf schließen, dass die Aussage $\mathcal{A}(n)$ für alle $n \in \mathbb{N}$ gilt.

**Bemerkung 2.2.2** (Wohlordnungsprinzip von $\mathbb{N}$)**.** Das Verfahren der vollständige Induktion basiert auf der Tatsache, dass wir die natürlichen Zahlen anordnen können (*Wohlordnungsprinzip* von $\mathbb{N}$).

**Bemerkung 2.2.3** (Nebenbemerkung zu Minimum, Maximum)**.** Insbesondere gibt es wegen des Wohlordnungsprinzips von $\mathbb{N}$ für jede endliche Teilmenge $A$ von $\mathbb{N}$ ein kleinstes Element ($\exists k \in A$, sodass $\forall n \in A : k \leq n$) und ein größtes Element ($\exists g \in A$, sodass $\forall n \in A : g \geq n$). Wir sagen, $k$ ist das *Minimum* und $g$ ist das *Maximum* von $A$. Wir schreiben: $k = \min A$ und $g = \max A$.

Wir betrachten im Folgenden ein paar Beispiele zur Induktion.

**Satz 2.2.1** (Geometrische Reihe)**.** *Sei* $|q| < 1$, $n \in \mathbb{N}_0$. *Dann gilt:*

$$\sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}$$

*und insbesondere*

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1 - q}.$$

*Beweis.* Wir beweisen die erste Aussage (= die erste Gleichung) mittels vollständiger Induktion. Die zweite Aussage (Gleichung) folgt daraus mit Hilfe der Regeln für die Grenzwertbildung.

Zunächst ein Hinweis zur Schreibweise. $\Sigma$ ist das „Summenzeichen". Wir summieren damit über die verschiedenen Werte von $k$ wie im Folgenden angegeben:

$$\sum_{k=0}^{n} q^k = q^0 + q^1 + q^2 + ... + q^{n-1} + q^n.$$

Induktionsanfang ($n = 0$):

$$\sum_{k=0}^{0} q^k = q^0 = 1 = \frac{1 - q}{1 - q} = \frac{1 - q^{0+1}}{1 - q}$$

ist wahr. Wir können uns interessehalber noch $n = 1$ anschauen:

$$\sum_{k=0}^{1} q^k = q^0 + q^1 = 1 + q = \frac{(1 + q)(1 - q)}{1 - q} = \frac{1 - q^2}{1 - q}$$

ist also auch wahr.

Induktionsannahme (IA): Wir nehmen an, die Aussage $\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}$ gilt für festes $n$.

Induktionsschritt ($n \Rightarrow n+1$):

$$\sum_{k=0}^{n+1} q^k = \sum_{k=0}^{n} q^k + q^{n+1} = \frac{1-q^{n+1}}{1-q} + q^{n+1} = \frac{1-q^{n+1}+q^{n+1}(1-q)}{1-q} = \frac{1-q^{n+2}}{1-q}$$

ist wahr. Hier haben wir im zweiten Schritt die Induktionsannahme verwendet.

Damit ist die Aussage für alle $n \in \mathbb{N}_0$ wahr.

Insbesondere gilt nun im Limes $n \to \infty$ :

$$\sum_{k=0}^{\infty} q^k = \lim_{n\to\infty} \sum_{k=0}^{n} q^k = \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \frac{1-\lim_{n\to\infty} q^{n+1}}{1-q} = \frac{1}{1-q}.$$

Hier folgt die vorletzte Gleichung aus den Rechenregeln für Limiten (ohne Beweis verwenden wir hier, dass wir den Limes in den Quotienten und die Differenz „reinziehen" können) und die letzte Gleichung daraus, dass $|q| < 1$, also $\lim_{n\to\infty} q^{n+1} = \lim_{n\to\infty} q^n = 0$. □

**Beispiel 2.2.1** (Geometrische Reihe)**.** Betrachte $q = \frac{1}{2}$. Dann gilt:

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + ... = \frac{1}{1-\frac{1}{2}} = 2.$$

**Satz 2.2.2** (Arithmetische Reihen)**.** *Für die Summe der ersten $n$ natürlichen Zahlen gilt die Gaußsche Summenformel:*

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

*Für die Summe der ersten $n$ ungeraden Zahlen gilt:*

$$\sum_{k=1}^{n} (2k-1) = n^2.$$

*Beweis.* Mittels vollständiger Induktion (Übung). □

**Bemerkung 2.2.4** (Historisch)**.** Es heißt, Gauß (1777 – 1855) habe die Summenformel in der Grundschule entdeckt, als er von seinem Lehrer aufgefordert wurde, die natürlichen Zahlen von 1 bis 100 zusammen zu zählen. Rechnung: $(1+99) + (2+98) + ... + (49+51) + 50 + 100 = 50 \cdot 100 + 50 = 5050$.

**Satz 2.2.3** (Binomischer Lehrsatz)**.** *Seien $x, y \in \mathbb{R}, n \in \mathbb{N}_0$. Es gilt der Binomische Lehrsatz:*

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

*Dabei ist*

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

*Beweis.* Wir führen die Induktion über $n \in \mathbb{N}_0$. Induktionsanfang ($n = 0$):

$$(x+y)^0 = 1 = \binom{0}{0} x^0 y^0 = \sum_{k=0}^{0} \binom{0}{k} x^{0-k} y^k$$

Der Fall $n = 0$ ist also wahr. Wir können uns interessehalber noch $n = 1$ anschauen:

$$(x+y)^1 = x + y = \binom{1}{0} x^1 y^0 + \binom{1}{1} x^0 y^1 = \sum_{k=0}^{1} \binom{1}{k} x^{1-k} y^k$$

Der Fall $n = 1$ ist also auch wahr.

Induktionsannahme (IA): Wir nehmen nun an, die Aussage $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$ gilt.

Induktionsschritt ($n \Rightarrow n+1$):

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)^n (x+y) = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k (x+y) \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k x + \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k y = \sum_{k=0}^{n} \binom{n}{k} x^{n+1-k} y^k + \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^{k+1} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{n+1-k} y^k + \sum_{l=1}^{n+1} \binom{n}{l-1} x^{n+1-l} y^l,
\end{aligned}$$

wobei wir im zweiten Schritt die Induktionsannahme verwendet haben und im letzten Schritt die Substitution $l := k+1$ (wobei wir nun, da wir richtig substituiert haben, das $l$ auch wieder $k$ nennen können). Nun nutzen wir, dass für jedes $k \geq 1$:

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

Damit folgt:

$$\begin{aligned}
(x+y)^{n+1} &= \sum_{k=0}^{n} \binom{n}{k} x^{n+1-k} y^k + \sum_{k=1}^{n+1} \binom{n}{k-1} x^{n+1-k} y^k \\
&= \binom{n}{0} x^{n+1} y^0 + \sum_{k=1}^{n} \left( \binom{n}{k} + \binom{n}{k-1} \right) x^{n+1-k} y^k + \binom{n}{n} x^0 y^{n+1} \\
&= \binom{n+1}{0} x^{n+1} y^0 + \sum_{k=1}^{n} \binom{n+1}{k} x^{n+1-k} y^k + \binom{n+1}{n+1} x^0 y^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^{n+1-k} y^k
\end{aligned}$$

Aus $\mathcal{A}(n)$ wahr folgt also $\mathcal{A}(n+1)$ wahr. Damit ist die Aussage für alle $n \in \mathbb{N}_0$ wahr.

□

**Bemerkung 2.2.5.** Man kennt den binomischen Lehrsatz aus der Wahrscheinlichkeitsrechnung. Dabei ist $x = p$ die Wahrscheinlichkeit, dass ein gewisses Ereignis eintritt, $y = (1 - p)$ die Wahrscheinlichkeit, dass es nicht eintritt, und $(x+y) = p+1-p = 1$ die Gesamtwahrscheinlichkeit.

**Bemerkung 2.2.6** (Vollständige Induktion)**.** Statt des Induktionsschrittes $n \to n+1$, also statt zu zeigen, dass für gegebenes $n$ die Implikation $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ wahr ist, kann man auch zeigen, dass die Implikation $\mathcal{A}(k) : k < n \Rightarrow \mathcal{A}(n)$ wahr ist. Dies ist auch ein gültiger Induktionschritt. In Worten: Man zeigt, dass die Aussage, wenn sie für $k < n$ gilt, auch für $n$ gilt.

## 2.3 Mengen

Man kann die Mengenlehre als den Rahmen oder die Sprache betrachten, in der sich die (moderne) Mathematik begründen und betreiben lässt. Die Mengenlehre wurde durch Cantor eingeführt. Er gab 1895 die folgende Definition einer Menge:

**Definition 2.3.1** (Menge)**.** Unter eine *Menge* verstehen wir jede Zusammenfassung $M$ von bestimmten wohlunterschiedenen Objekten $m$ unserer Anschauung oder unseres Denkens (welche die *Elemente* von $M$ genannt werden) zu einem Ganzen.

**Bemerkung 2.3.1** (Extensionalitätsprinzip)**.** Eine Menge ist durch ihre Elemente vollständig bestimmt.

**Bemerkung 2.3.2** (Schreibweise)**.** Wir schreiben:
$x \in M$ für „x ist ein Element von $M$"
$x \notin M$ für „x ist kein Element von $M$"

**Bemerkung 2.3.3** (Mengenparadoxon)**.** Man kann zunächst versuchen, die Mengendefinition wie folgt zu verstehen (= *naive Mengenlehre*). Sei G die Menge aller Objekte (unserer Anschauung oder unseres Denkens). Sei $E(x)$ eine Eigenschaft von Objekten $x$ aus $G$. Dann bildet

$$M = \{x \in G | E(x)\}$$

wieder eine Menge, nämlich die Menge aller Objekte $x \in G$ mit der Eigenschaft $E(x)$. Man beachte, dass auch $M \in G$, denn $G$ ist ja die Menge aller Objekte.

Eine solches Mengenverständnis führt zu Widersprüchen, wie Russel gezeigt hat. Er veröffentlichte 1903 ein Paradoxon, die sog. *Russel'sche Antinomie*, bekannt als

„die Menge aller Mengen, die sich selbst nicht enthalten"

Formal:

$$R = \{x | Mg(x) \wedge x \notin x\}$$

Wäre $R \in R \Rightarrow Mg(R)$ und $R \notin R$, was ein Widerspruch ist. Andersherum, falls $Mg(R)$ und $R \notin R \Rightarrow R \in R$, was auch ein Widerspruch ist.

Dieses Paradoxon zeigt, dass der naive Mengenbegriff nicht widerspruchsfrei ist. Problem ist, dass er von einer fertigen Gesamtheit von Mengen ausgeht. Tatsächlich brauchen wir einen solchen Mengenbegriff in der Mathematik nicht. Es reicht, Mengen aus Elementen zu bilden, die bereits *gegeben* sind. Dies kann man stufenweise tun, z.B. 1. Stufe: $\emptyset$, 2. Stufe: $\emptyset, \{\emptyset\}$, 3. Stufe: $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}$ ...

Beginnen wir mit *gegebenen* Objekten $x$, können wir eine Menge $M$ durchaus über ihre Eigenschaften $E(x)$ definieren und erhalten damit einen direkt Zusammenhang zu den Aussagen.

**Beispiel 2.3.1.** Sei $\mathcal{A}$ die Aussage „... ist eine Quadratzahl", dann ist die zugehörige Menge

$$A = \{x \in \mathbb{N} | \mathcal{A}(x)\} = \{x \in \mathbb{N} | x = k^2, k \in \mathbb{N}\} = \{1, 4, 9, 16, ...\}$$

**Bemerkung 2.3.4.** Bei Mengen kommt es weder auf die Reihenfolge der Elemente an noch darauf, ob Elemente mehrfach aufgezählt werden. So sind die folgenden Mengen ein- und dieselbe: $\{1, 2, 3\} = \{3, 1, 2\} = \{1, 2, 2, 3\}$. Mengen können auch selbst wieder Mengen enthalten, wie z.B. die Menge $M = \{1, \{1\}, \{2, 3\}\}$. Dabei sind 1 und $\{1\}$ verschieden.

**Beispiel 2.3.2.** Bekannte Mengen sind:

- $\emptyset$ oder $\{\}$: die leere Menge (sie beinhaltet kein Element)
- $\mathbb{N} = \{1, 2, 3, ...\}$: die Menge der natürlichen Zahlen
- $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$: die Menge der ganzen Zahlen
- $\mathbb{Q} = \left\{\frac{p}{q} | p \in \mathbb{Z} \wedge q \in \mathbb{Z} \setminus \{0\}\right\}$: die Menge der rationalen Zahlen
- $\mathbb{R}$ : die Menge der reellen Zahlen
- $\mathbb{C} = \{x + iy | x \in \mathbb{R} \wedge y \in \mathbb{R}\}$: die Menge der komplexen Zahlen

**Definition 2.3.2** (Gleichheit von Mengen)**.** Zwei Mengen $A$ und $B$ sind *gleich*, wenn sie dieselben Elemente enthalten:

$$A = B \quad \Leftrightarrow \quad \forall x : x \in A \Leftrightarrow x \in B$$

**Definition 2.3.3** (Teilmenge)**.** $A$ ist eine *Teilmenge* von $B$, wenn alle Elemente, die in $A$ enthalten sind, auch in $B$ enthalten sind:

$$A \subseteq B \quad \Leftrightarrow \quad \forall x : x \in A \Rightarrow x \in B$$

Wir nennen eine Menge $A \subset B$ eine *echte* Teilmenge von $B$, wenn $B$ Elemente enthält, die in $A$ nicht vorkommen ($A \neq B$). Wir schreiben dann $A \subset B$ oder explizit $A \subsetneq B$.

**Bemerkung 2.3.5.** Insbesondere gilt für jede Menge $B$: $\emptyset \subseteq B$, $B \subseteq B$.

**Satz 2.3.1.** *Für Mengen $A, B, C$ gilt:*

$$a) \quad A \subseteq B \wedge B \subseteq A \quad \Leftrightarrow \quad A = B$$

$$b) \quad A \subseteq B \wedge B \subseteq C \quad \Rightarrow \quad A \subseteq C$$

*Beweis.* Zu a):

$$\begin{aligned} A \subseteq B \wedge B \subseteq A &\Leftrightarrow (\forall x : x \in A \Rightarrow x \in B) \wedge (\forall x : x \in B \Rightarrow x \in A) \\ &\Leftrightarrow \forall x : (x \in A \Rightarrow x \in B) \wedge (x \in B \Rightarrow x \in A) \\ &\Leftrightarrow \forall x : (x \in A \Leftrightarrow x \in B) \\ &\Leftrightarrow A = B \end{aligned}$$

Zu b): Analog. □

**Definition 2.3.4** (Potenzmenge)**.** Als Potenzmenge $\mathcal{P}(A)$ bezeichnen wir die Menge aller Teilmengen von $A$,

$$\mathcal{P}(A) = \{M | M \subseteq A\}.$$

**Beispiel 2.3.3.** Sei $A = \{1, 2, 4\}$.

$$\mathcal{P}(A) = \mathcal{P}(\{1, 2, 4\}) = \{\emptyset, \{1\}, \{2\}, \{4\}, \{1, 2\}, \{1, 4\}, \{2, 4\}, \{1, 2, 4\}\}$$

**Definition 2.3.5** (Differenz)**.** Die *Differenz* $B \setminus A$ der Mengen $B$ und $A$ ist die Menge aller Elemente, die in $B$, aber nicht in $A$ liegen:

$$B \setminus A = \{x | x \in B \wedge x \notin A\} = \{x \in B | x \notin A\}.$$

Ist $A$ eine Teilmenge von $B$, dann nennen wir $B \setminus A$ das *Komplement* von $A$ in $B$. Ist $B$ eine nicht eigens zu spezifizierende Grundmenge (im Allgemeinen $\Omega$ genannt), dann schreiben wir für das Komplement $A^c$.

**Beispiel 2.3.4.** Seien $B = \{1, 3, 5\}$ und $A = \{1, 5\}$, dann ist $B \setminus A = \{3\}$.

**Definition 2.3.6** (Schnitt)**.** Der *Schnitt* der Mengen $A$ und $B$ ist die Menge aller Elemente, die in $A$ und $B$ liegen:

$$A \cap B = \{x | x \in A \wedge x \in B\} = \{x \in A | x \in B\}.$$

Ist $A \cap B = \emptyset$, so nennen wir $A$ und $B$ *disjunkt*.

**Beispiel 2.3.5.** Seien $A = \{1, 2\}, B = \{2, 7, 8\}$ und $C = \{7, 8, 9\}$.
Dann ist $A \cap B = \{2\}, B \cap C = \{7, 8\}$ und $A \cap C = \emptyset$, also $A$ und $C$ disjunkt.

**Definition 2.3.7** (Vereinigung)**.** Die *Vereinigung* der Mengen $A$ und $B$ ist die Menge aller Elemente, die in $A$ oder $B$ liegen:

$$A \cup B = \{x | x \in A \vee x \in B\}.$$

**Satz 2.3.2** (Rechenregeln). *Seien $A, B$ Mengen. Für $\cap$ und $\cup$ gelten Kommutativ-, Assoziativ- und Distributivgesetze sowie die De-Morgan-Regeln*

$$(A \cap B)^c = A^c \cup B^c \quad \text{und} \quad (A \cup B)^c = A^c \cap B^c.$$

*Beweis.* Dies folgt aus den entsprechenden Gesetzen für $\wedge$ und $\vee$ (Übung). □

**Bemerkung 2.3.6.** Wir können die letzten beiden Definitionen auf endliche und sogar abzählbar unendliche (mehr zu diesen Begriffen später) Schnitte und Vereinigungen ausdehnen. Wir schreiben dann, mit $n \in \mathbb{N}$:

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap ... \cap A_n \quad \text{und} \quad \bigcap_{n\in\mathbb{N}} A_n = \lim_{n\to\infty} \bigcap_{i=1}^{n} A_i$$

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup ... \cup A_n \quad \text{und} \bigcup_{n\in\mathbb{N}} A_n = \lim_{n\to\infty} \bigcup_{i=1}^{n} A_i$$

**Beispiel 2.3.6.** Betrachte reelle Intervalle $A_n \subseteq \mathbb{R}$ (auch dies sind Mengen, mehr dazu später). Seien $a, b \in \mathbb{R}, a < b$, und $n \in \mathbb{N}$. Sei $A_n = ]a - 1/n, b + 1/n[$. Dann ist

$$\bigcap_{n\in\mathbb{N}} A_n = \lim_{n\to\infty} \bigcap_{i=1}^{n} \left]a - \frac{1}{n}, b + \frac{1}{n}\right[ = [a, b].$$

Sei weiter $B_n = [a + 1/n, b - 1/n]$. Dann ist

$$\bigcup_{n\in\mathbb{N}} B_n = \lim_{n\to\infty} \bigcup_{i=1}^{n} \left[a + \frac{1}{n}, b - \frac{1}{n}\right] = ]a, b[.$$

**Bemerkung 2.3.7.** Hier bezeichnet $[a, b] \subset \mathbb{R}$ das *abgeschlossene* Intervall (das seine Randpunkte $a, b$ enthält) und $]a, b[ \subset \mathbb{R}$ das *offene* Intervall (das seine Randpunkte $a, b$ nicht enthält).

**Definition 2.3.8** (Kartesisches Produkt). Als *kartesisches Produkt* nicht-leerer Mengen $A_1, ..., A_n$ bezeichnen wir die Menge aller Folgen der Länge $n$, so genannte *n-Tupel* $(a_1, .., a_n)$, mit $a_1 \in A_1, ..., a_n \in A_n$:

$$A_1 \times A_2 \times ... \times A_n = \{(a_1, a_2, ..., a_n) | a_i \in A_i \ \forall i = 1, .., n\}.$$

Ist $A_i = A \ \forall i = 1, ..., n$ so schreiben wir $A^n := A \times A \times .... \times A$ ($n$-mal).

**Bemerkung 2.3.8.** Bei $n$-Tupeln kommt es im Unterschied zu Mengen auf die Reihenfolge der Einträge an. So ist

$$(1, 2, 3) \neq (2, 1, 3) \quad \text{aber} \quad \{1, 2, 3\} = \{2, 1, 3\}.$$

**Bemerkung 2.3.9** (Reelle Folgen). Auch die letzte Definition kann auf abzählbar unendlich viele Mengen $A_k, k \in \mathbb{N}$, ausgedehnt werden. So ist z.B. $\mathbb{R}^{\mathbb{N}}$ die Menge aller reellen Folgen $(a_1, a_2, ...)$ mit $a_k \in \mathbb{R}$ für alle $k \in \mathbb{N}$.

**Bemerkung 2.3.10** (Historisch)**.** Wir nennen das kartesische Produkt „kartesisch" in Erinnerung an Descartes, der erkannt hat, dass man die Punkte der euklidischen Ebene durch Zahlenpaare (2-Tupel) darstellen kann und die Ebene als kart. Produkt

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) | x, y \in \mathbb{R}\}.$$

Analog kann man die Punkte des 3-dim. euklidischen Raums als 3-Tupel darstellen und den 3-dim. Raum als kart. Produkt

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) | x, y, z \in \mathbb{R}\}.$$

**Bemerkung 2.3.11** (Paar)**.** Ein 2-Tupel $(\cdot, \cdot)$ nennt man auch ein *Paar*. Paare können auch direkt über (Mengen von) Mengen definiert werden. Seien $A, B$ Mengen. Sei $A \times B = \{(a, b) : a \in A, b \in B\}$. Dann definieren wir

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

Damit ist $(a, b) \subset \mathcal{P}(A \cup B)$ bzw. $(a, b) \in \mathcal{P}(\mathcal{P}(A \cup B))$. Explizit:

$$A \times B := \{x \in \mathcal{P}(\mathcal{P}(A \cup B)) | x = \{\{a\}, \{a, b\}\} \text{ mit } a \in A, b \in B\}.$$

Achtung: Es gibt keine analoge Definition für n-Tupel mit $n \geq 3$!

**Definition 2.3.9** (Endliche Menge)**.** Eine Menge $A$ heißt endlich, wenn sie nur endlich viele Elemente enthält. Wir bezeichnen die Anzahl der Elemente mit $|A|$.

**Beispiel 2.3.7.** Sei $A = \{1, 2, 4, 7, 9\}$. Dann ist $|A| = 5$.

**Bemerkung 2.3.12** (Mächtigkeit)**.** Später betrachten wir auch unendliche Mengen wie $\mathbb{N}$, $\mathbb{Q}$ und $\mathbb{R}$. Dann verallgemeinert sich der Ausdruck „Anzahl der Elemente" auf den Begriff der „Mächtigkeit". Auch für ihn schreiben wir $|A|$.

**Satz 2.3.3.** *Seien $A_1, ..., A_n$ endliche Mengen. Dann gilt:*

$$|A_1 \times A_2 ... \times A_n| = |A_1| \cdot |A_2| \cdot ... \cdot |A_n|.$$

*Beweis.* Wir suchen die Anzahl an $n$-Tupeln $(a_1, ..., a_n)$ mit $a_i \in A_i \ \forall i = 1, ..., n$. Dabei gibt es $|A_1|$ Möglichkeiten $a_1$ zu wählen. Für jede dieser $|A_1|$ Möglichkeiten gibt es $|A_2|$ Möglichkeiten $a_2$ zu wählen, für jede dieser $|A_1| \cdot |A_2|$ Möglichkeiten gibt es $|A_3|$ Möglichkeiten $a_3$ zu wählen usw. Damit erhält man genau die oben stehende Produktregel. □

**Satz 2.3.4.** *Seien $A, B$ und $C$ endliche Mengen. Dann gilt:*

$$a) \quad |A \cup B| = |A| + |B| - |A \cap B|.$$

*Insbesondere gilt damit, falls $A, B$ disjunkt:* $|A \cup B| = |A| + |B|$.

$$b) \quad |A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

*Seien $A_1, ..., A_n, n \in \mathbb{N}$ endliche Mengen. Dann gilt das sog.* Inklusions-Exklusions-Prinzip*:*

$$c) \quad \left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\substack{I: \\ \emptyset \neq I \subseteq \{1,...,n\}}} (-1)^{|I|-1} \left| \bigcap_{i \in I} A_i \right|.$$

*Beweis.* Zu a) Anschaulich: Wenn wir alle Elemente aus $A$ und aus $B$ zusammenzählen, zählen wir genau die Elemente doppelt, die in $A \cap B$ enthalten sind. Die müssen wir wieder abziehen, wenn wir die Anzahl an Elementen bestimmen wollen, die in $A \cup B$ enthalten sind. Formal: Für zwei endliche Mengen $A$ und $B$ gilt: $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$, wobei $A \setminus B, A \cap B$ und $B \setminus A$ disjunkte Mengen sind. Das heißt:

$$|A \cup B| = |A \setminus B| + |A \cap B| + |B \setminus A|$$

Weiter sind $A \setminus B$ und $A \cap B$ disjunkt und $(A \setminus B) \cup (A \cap B) = A$, d.h. $|A \setminus B| + |A \cap B| = |A|$ und analog für $B$, d.h. $|B| = |B \setminus A| + |A \cap B|$. Es folgt:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Falls $A, B$ disjunkt, dann ist $A \cap B = \emptyset$ und $|\emptyset| = 0$. Folglich ist dann: $|A \cup B| = |A| + |B|$.

Zu b): Die Aussage b) folgt direkt aus a) und den Rechengesetzen:

$$\begin{aligned} |A \cup B \cup C| &= |A \cup B| + |C| - |(A \cup B) \cap C| \\ &= |A| + |B| - |A \cap B| + |C| - |(A \cap C) \cup (B \cap C)| \\ &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|. \end{aligned}$$

Zu c): Wir beweisen nun mittels vollständiger Induktion, dass sich die Relation auf beliebige, ja sogar abzählbar unendliche Vereinigungen von Mengen ausgeweitet werden kann.

Wir führen die Induktion über $n \in \mathbb{N}$ und beginnen mit dem Induktionsanfang $n = 1$. Für $n = 1$ gilt: $|\bigcup_{i=1} A_i| = |A_1| = (-1)^{1-1} |\bigcap_{i=1} A_i|$. Das stimmt also.

Wir betrachten nun den Induktionsschritt von $n$ auf $n + 1$ unter der Induktionsannahme, dass die Relation für $n$ stimmt.

$$
\begin{aligned}
\Big|\bigcup_{i\in\{1,\dots,n+1\}} A_i\Big| &= \Big|\bigcup_{i\in\{1,\dots,n\}} A_i\Big| + |A_{n+1}| - \Big|\Big(\bigcup_{i\in\{1,\dots,n\}} A_i\Big)\cap A_{n+1}\Big| \\
&= \Big|\bigcup_{i\in\{1,\dots,n\}} A_i\Big| + |A_{n+1}| - \Big|\bigcup_{i\in\{1,\dots,n\}} (A_i\cap A_{n+1})\Big| \\
&= \sum_{\substack{I:\\ \emptyset\neq I\subseteq\{1,\dots,n\}}} (-1)^{|I|-1}\Big|\bigcap_{i\in I} A_i\Big| + |A_{n+1}| - \sum_{\substack{J:\\ \emptyset\neq J\subseteq\{1,\dots,n\}}} (-1)^{|J|-1}\Big|\bigcap_{j\in J}(A_j\cap A_{n+1})\Big| \\
&= \sum_{\substack{\emptyset\neq I\subseteq\{1,\dots,n+1\}\\ n+1\notin I}} (-1)^{|I|-1}\Big|\bigcap_{i\in I} A_i\Big| + (-1)^{|\{n+1\}|-1}|\bigcap_{i=n+1} A_i| \\
&+ \sum_{\substack{J:\\ \emptyset\neq J\subseteq\{1,\dots,n\}}} (-1)^{|J\cup\{n+1\}|-1}\Big|\bigcap_{j\in(J\cup\{n+1\})} A_j\Big| \\
&= \sum_{\substack{I:\\ \emptyset\neq I\subseteq\{1,\dots,n+1\}}} (-1)^{|I|-1}\Big|\bigcap_{i\in I} A_i\Big|.
\end{aligned}
$$

Hier haben wir im dritten Schritt die Induktionsannahme gleich doppelt verwendet. Im letzten Schritt haben wir erkannt, dass der erste der drei verbleibenden Terme über alle Indexmengen $I$ summiert, die $n+1$ nicht enthalten, der letzte über alle, die $n+1$ und mindestens ein weiteres $i \in I$ enthalten, während der mittlere Term genau den noch fehlenden Term darstellt, der nur über die einelementige Indexmenge $I = \{n+1\}$ „summiert". □

## 2.4 Relationen

**Definition 2.4.1** (Relation)**.** Seien $X, Y$ Mengen, $R \subseteq X \times Y$. Das Tripel $\mathcal{R} := (X, Y, R)$ heißt (zweistellige) *Relation* zwischen (Elementen von) X und (Elementen von) Y. Formal:

$$x \sim y \quad :\Leftrightarrow \quad (x, y) \in R.$$

Statt $(X, Y, R)$ schreiben wir auch $(X, Y, \sim)$. Falls $Y = X$, spricht man von einer Relation auf $X$. In dem Fall schreibt man $(X, \sim)$ statt $(X, X, \sim)$.

**Beispiel 2.4.1.** $\mathcal{R}_1 = (\mathbb{R}, \leq)$, $\mathcal{R}_2 = (\mathbb{R}, >)$ und $\mathcal{R}_3 = (\mathbb{R}, =)$ sind Relationen auf $\mathbb{R}$. Die zugehörigen Mengen sind $R_1 = \{(x, y) \in \mathbb{R}^2 | x \leq y\}$, $R_2 = \{(x, y) \in \mathbb{R} \times \mathbb{R} | x > y\}$ und $R_3 = \{(x, y) \in \mathbb{R} \times \mathbb{R} | x = y\}$.

Eine wichtige Art von Relationen sind Äquivalenzrelationen.

**Definition 2.4.2** (Äquivalenzrelation)**.** Eine Relation $(X, \sim)$ heißt *Äquivalenzrelation*, wenn gilt:

$$\begin{aligned}
&1) \quad \forall x \in X: \quad x \sim x \quad \text{(reflexiv)} \\
&2) \quad \forall x, y \in X: \quad x \sim y \Leftrightarrow y \sim x \quad \text{(symmetrisch)} \\
&3) \quad \forall x, y, z \in X: \quad x \sim y \wedge y \sim z \Rightarrow x \sim z \quad \text{(transitiv)}
\end{aligned}$$

Die Äquivalenzrelation ist eine Abschwächung der Gleichheitsrelation. Insbesondere ist die Gleichheitsrelation selbst eine Äquivalenzrelation, wie der folgende Satz zeigt.

**Satz 2.4.1.** *$(X, =)$ ist eine Äquivalenzrelation.*

*Beweis.* Alle drei Bedingungen sind erfüllt:

$$\begin{aligned}
&1) \quad \forall x \in X: \quad x = x \quad \text{(reflexiv)} \\
&2) \quad \forall x, y \in X: \quad x = y \Leftrightarrow y = x \quad \text{(symmetrisch)} \\
&3) \quad \forall x, y, z \in X: \quad x = y \wedge y = z \Rightarrow x = z \quad \text{(transitiv)}
\end{aligned}$$

□

Weiteres wichtiges Beispiel:

**Satz 2.4.2.** *Seien $x, y \in \mathbb{Z}$ und $m \in \mathbb{N}$. Die Relation $(\mathbb{Z}, \sim)$ mit*

$$x \sim y \quad :\Leftrightarrow \quad m|(x-y)$$

*ist eine Äquivalenzrelation. Hier steht $m|(x-y)$ für „$m$ teilt $(x-y)$".*

*Beweis.* Wir prüfen die drei Bedingungen:

1) $\forall x \in \mathbb{Z}: m|(x-x) \Leftrightarrow m|0$ und $m|0$ ist *wahr* (reflexiv)
2) $\forall x, y \in \mathbb{Z}: m|(x-y) \Leftrightarrow m|(-1)(x-y) \Leftrightarrow m|(y-x)$ (symmetrisch)
3) $\forall x, y, z \in \mathbb{Z}: m|(x-y) \wedge m|(y-z) \Rightarrow m|((x-y)+(y-z)) \Leftrightarrow m|(x-z)$ (transitiv)

Begründung für „⇒" in 3): Weil $m|(x-y)$, $\exists k \in \mathbb{N}$ mit $(x-y) = k \cdot m$ und weil $m|(y-z)$, $\exists l \in \mathbb{N}$ mit $(y-z) = l \cdot m$. Es folgt: $(x-y)+(y-z) = k \cdot m + l \cdot m = (k+l) \cdot m$. Folglich $\exists p \in \mathbb{N}$, nämlich $p := k + l$, mit $(x-y)+(y-z) = p \cdot m$ und somit $m|((x-y)+(y-z))$. □

Die im letzten Satz gezeigte Relation ist bekannt als *Kongruenz modulo*.

**Definition 2.4.3** (Kongruenz modulo)**.** Seien $x, y \in \mathbb{Z}$ und $m \in \mathbb{N}$. Die Relation

$$x \equiv y \bmod m \quad :\Leftrightarrow \quad m|(x-y)$$

heißt *Kongruenz modulo*. Wir sagen „$x$ kongruent $y$ modulo $m$".

**Beispiel 2.4.2.** Es ist $7 \equiv 2 \bmod 5$ und $7 \equiv -2 \bmod 3$.

Der Begriff der Äquivalenzrelation bringt mit sich den Begriff der Äquivalenzklasse.

**Definition 2.4.4** (Äquivalenzklasse)**.** Sei $(X, \sim)$ eine Äquivalenzrelation und $a \in X$. Die Menge

$$[a] := \{x \in X | x \sim a\}$$

heißt Äquivalenzklasse von $a$.

**Bemerkung 2.4.1** (Repräsentant der Äquivalenzklasse)**.** Es reicht also, für jede Menge zueinander äquivalenter Elemente (i.e. für jede Äquivalenzklasse) einen einzigen Repräsentanten $a$ anzugeben. (Hätte man etwa die Äquivalenzklassen „Hund" und „Katze", so würde je ein einzelner Hund bzw. eine einzelne Katze reichen, um die entsprechende Klasse zu spezifizieren bzw. um sie zu repräsentieren).

**Satz 2.4.3.** *Sei $(X, \sim)$ eine Äquivalenzrelation und $a, b \in X$. Dann ist entweder $[a] = [b]$ (falls $a \sim b$) oder $[a] \cap [b] = \emptyset$ (falls $a \nsim b$).*

*Beweis.* Es gibt nur zwei mögliche Fälle: $a \sim b$ und $a \nsim b$.

1. Fall: Sei $a \sim b$. Für jedes $x \in [a]$ gilt $x \sim a$ und $a \sim b$, also $x \sim b$ wegen Transitivität. Damit ist $x \in [b]$, also $[a] \subseteq [b]$. Analog folgt, mit $a \sim b \Leftrightarrow b \sim a$ (Symmetrie), dass $[b] \subseteq [a]$ und damit $[a] = [b]$.

2. Fall: Sei $a \nsim b$. Angenommen $[a] \cap [b] \neq \emptyset$. Dann $\exists x \in X$ mit $x \in [a] \cap [b]$. Also $x \sim a$ und $x \sim b$. Wegen Symmetrie auch $a \sim x$. Aus $a \sim x$ und $x \sim b$ folgt mit Transitivität: $a \sim b$. Dies ist ein Widerspruch, folglich muss $[a] \cap [b] = \emptyset$ gelten. □

**Beispiel 2.4.3.** Betrachte $(X, =)$. Die Äquivalenzklasse von $a \in X$ ist gegeben durch $[a] = \{a\}$.

**Beispiel 2.4.4.** Seien $a, b \in \mathbb{Z}, m \in \mathbb{N}$. Betrachte $(\mathbb{Z}, \sim)$ mit $a \sim b :\Leftrightarrow m|(a-b)$. Dann ist die Äquivalenzklasse zu $a$ gegeben durch

$$\begin{aligned} [a] &= \{z \in \mathbb{Z} |\ m|(z-a)\} = \{z \in \mathbb{Z} | \exists k \in \mathbb{Z} : z - a = k \cdot m\} \\ &= \{z \in \mathbb{Z} | \exists k \in \mathbb{Z} : z = a + k \cdot m\} \ = \{a + k \cdot m | k \in \mathbb{Z}\} =: a + m\mathbb{Z} \end{aligned}$$

Hier ist $m\mathbb{Z} := \{..., -2m, -m, 0, m, 2m, ...\}$ die Menge aller ganzzahligen Vielfachen von $m$. Weil $\forall k \in \mathbb{Z} : [a] = [a + k \cdot m]$ , gibt es nur $m$ (paarweise) verschiedene Äquivalenzklassen:

$$[0], [1], [2], ..., [m-1].$$

Man bezeichnet diese Klassen auch als „Restklassen modulo $m$" und schreibt für ihre Menge:

$$\mathbb{Z}_m := \{[0]_m, [1]_m, [2]_m, ..., [m-1]_m\}.$$

**Bemerkung 2.4.2** (Quotientenraum)**.** $\mathbb{Z}_m$ definiert einen sog. *Quotientenraum*, $\mathbb{Z}_m =: \mathbb{Z}/m\mathbb{Z}$. Quotientenräume erhält man z.B. auch in der Physik, wenn man die Symmetrien eines Systems berücksichtigt (z.B. Translations- und Rotationssymmetrie, $\mathbb{R}^3 \times SO(3)$).

**Definition 2.4.5** (Umkehrrelation)**.** Sei $\mathcal{R} = (X, Y, R)$ eine Relation. Das Tripel $\mathcal{R}^{-1} := (Y, X, R^{-1})$ mit

$$R^{-1} = \{(y, x) \in Y \times X | (x, y) \in R\}$$

heißt *Umkehrrelation*.

**Bemerkung 2.4.3** (Existenz der Umkehrrelation)**.** Die Umkehrrelation ist selbst wieder eine Relation. Im Unterschied zur Umkehrfunktion existiert die Umkehrrelation immer ($R^{-1}$ beinhaltet ja einfach die Paare mit „vertauschten" Einträgen.)

**Beispiel 2.4.5.** Sei $X = \{1, 2, 3\}$ und $Y = \{1, 2, 3, 4, 5\}$ und $\mathcal{R} = (X, Y, R)$ mit

$$R = \{(1, 5), (2, 3), (2, 4), (2, 5)\}.$$

Dann ist die Umkehrrelation gegeben durch $\mathcal{R}^{-1} = (Y, X, R^{-1})$ mit

$$R^{-1} = \{(5, 1), (3, 2), (4, 2), (5, 2)\}.$$

Man kann sich bildlich z.B. vorstellen, $X$ sei eine Menge von Personen und $Y$ eine Menge von Eissorten und $(x, y) \in R$ besagt „x mag Eissorte y". Dann gibt es zwei Eissorten, die von niemanden gemocht werden (1 und 2) und eine Person (3), die kein Eis mag. Dazu gibt es natürlich eine Umkehrrelation $\mathcal{R}^{-1}$ und $(y, x) \in R^{-1}$ besagt, dass „y von x gemocht wird". (Man beachte: Im Unterschied zu Funktionen beschreiben $\mathcal{R}$ und $\mathcal{R}^{-1}$ keine eindeutigen Abbildungen, z.B. wird $2 \in X$ von $\mathcal{R}$ auf $3, 4, 5 \in Y$ abgebildet und $5 \in Y$ von $\mathcal{R}^{-1}$ auf $1, 2 \in X$.)

## 2.5 Abbildungen (Funktionen)

**Definition 2.5.1** (Abbildung/Funktion)**.** Wir nennen eine Relation $f = (X, Y, R)$ eine *Abbildung* oder *Funktion* genau dann, wenn gilt:

$$\forall x \in X \quad \exists! \quad y \in Y : (x, y) \in R$$

Hier bedeutet $\exists!$: „es existiert *genau ein*". Wir nennen $y$ den *Funktionswert* von $f$ and der Stelle $x$ und schreiben $y = f(x)$. Statt $f = (X, Y, R)$ schreiben wir $f : X \to Y$. Wir nennen $X$ den *Definitionsbereich*, $Y$ den *Wertebereich* und $R = G_f$ den *Graph* von $f$.

**Beispiel 2.5.1.**

- $f : [-3, 3] \to \mathbb{R}, f(x) = x^3$ ist eine Abbildung (Funktion).
- Sei $n \in \mathbb{N}$. Jedes $n$-Tupel $(x_1, ..., x_n) \in X^n$ kann als Abbildung $\{1, ..., n\} \to X, k \to x_k$ angesehen werden.
- Reelle Folgen $(a_n)_{n \in \mathbb{N}}$ sind Abbildungen $\mathbb{N} \to \mathbb{R}, n \to a_n$.
- Jede Familie $(x_i)_{i \in I}$ ist eine Abbildung $I \to X, i \to x_i$. Man nennt dabei $I$ die *Indexmenge*.

- Seien $n, m \in \mathbb{N}$. Alle reellen $m \times n$ Matrizen $A = (a_{ij})_{ij}$ mit $i = 1, ..., m, j = 1, ..., n$ können als Abbildungen $\{1, ..., m\} \times \{1, ..., n\} \to \mathbb{R}$, $(i, j) \to a_{ij}$ aufgefasst werden.

**Definition 2.5.2** (Identität, Einbettung, Einschränkung)**.** Sei $X$ eine Menge. Die Abbildung

$$\mathrm{id}_X : X \to X, \quad x \to x$$

heißt *Identität*. Sei $A \subseteq X$. Die Abbildung

$$i_A : \quad A \to X, \quad x \to x$$

heißt *Einbettung* (Inklusion) von $A$ in $X$. Sei $Y$ eine Menge, $f : X \to Y$ eine Funktion. Die Abbildung

$$f|_A : \quad A \to Y, \quad x \to f(x)$$

heißt *Einschränkung* (Restriktion) von $f$ auf $A$.

**Bemerkung 2.5.1.** Insbesondere ist demnach $i_A = \mathrm{id}_X|_A$.

**Definition 2.5.3** (Bild, Urbild)**.** Sei $f : X \to Y$ eine Abbildung, $A \subseteq X, B \subseteq Y$. Die Menge

$$f(A) := \{f(x) | x \in A\} = \{y \in Y | \exists x \in A : f(x) = y\}$$

heißt *Bild* von $A$ unter $f$. Die Menge

$$f^{-1}(B) := \{x \in X | f(x) \in B\} = \{x \in X | \exists y \in B : f(x) = y\}$$

heißt *Urbild* von $B$ unter $f$.

**Bemerkung 2.5.2.** Insbesondere ist demnach $f(A) \subseteq Y$ unf $f^{-1}(B) \subseteq X$.

**Beispiel 2.5.2.** Betrachte $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$. Seien $A = [-1, 1] \subseteq \mathbb{R}$ und $B = [1, 4] \subseteq \mathbb{R}$. Dann ist $f(A) = [0, 1]$ das Bild von $A$ und $f^{-1}(B) = [-2, -1] \cup [1, 2]$ das Urbild von $B$ unter $f$.

**Definition 2.5.4** (Verkettung)**.** Seien $f : X \to Y$ und $g : Y \to Z$ Abbildungen. Dann heißt die Abbildung $X \to Z, x \to g(f(x))$ *Verkettung* von $g$ und $f$. Wir schreiben dafür $g \circ f$.

**Satz 2.5.1** (Assoziativgesetz)**.** *Seien $f : W \to X, g : X \to Y, h : Y \to Z$ Funktionen. Dann gilt:*

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

*und dies ist eine Abbildung von $W$ nach $Z$.*

*Beweis.* Wiederholte Anwendung der Definition (Übung). □

**Beispiel 2.5.3.** Sei $f : \mathbb{R} \to \mathbb{R}, \ f(x) = x^2 + 2$ und $g : \mathbb{R} \to \mathbb{R}, \ g(x) = 3x - 5$. Dann ist:

$$g \circ f : \mathbb{R} \to \mathbb{R}, \quad (g \circ f)(x) = g(f(x)) = 3(x^2 + 2) - 5 = 3x^2 + 1$$

Weiter gilt:

$$f \circ g : \mathbb{R} \to \mathbb{R}, \quad (f \circ g)(x) = f(g(x)) = (3x-5)^2 + 2 = 9x^2 - 30x + 27$$

Achtung: Auch wenn $X = Y = Z$ ist, so ist i. A. $f \circ g \neq g \circ f$ (also $\circ$ i. A. nicht kommutativ)!

**Definition 2.5.5** (Injektivität, Surjektivität, Bijektivität)**.** Sei $f : X \to Y$ eine Abbildung. $f$ heißt

- *injektiv* $\Leftrightarrow \forall x, x' \in X : x \neq x' \Rightarrow f(x) \neq f(x')$

  (oder äquivalent als Kontraposition: $f$ injektiv $\Leftrightarrow f(x) = f(x') \Rightarrow x = x'$)
- *surjektiv* $\Leftrightarrow \forall y \in Y : \exists x \in X$ mit $y = f(x)$ (oder kurz: $f(X) = Y$)
- *bijektiv* $\Leftrightarrow f$ injektiv und surjektiv

**Bemerkung 2.5.3.** Anschaulicher nennen wir eine bijektive Abbildung auch eine 1:1-Abbildung.

**Beispiel 2.5.4.** Die Funktion $f : \mathbb{R} \to \mathbb{R}_0^+, x \mapsto x^2$ ist nicht injektiv, denn $f(1) = f(-1)$ (genauer: für $x = 1, x' = -1$ gilt: $x \neq x'$, aber $f(x) = f(x')$, also nicht injektiv), aber surjektiv, denn $\forall y \in \mathbb{R}_0^+ : \exists x \in \mathbb{R}$ mit $x^2 = y$, nämlich $x = \sqrt{y}$ (und die Wurzel ist wohldefiniert, weil $y \in \mathbb{R}_0^+$). Dagegen ist die Einschränkung $f|_{\mathbb{R}_0^+} : \mathbb{R}_0^+ \to \mathbb{R}_0^+, x \mapsto x^2$ injektiv und surjekiv, also bijektiv.

**Satz 2.5.2.** *Seien $f : X \to Y$ und $g : Y \to Z$ Abbildungen. Dann gilt:*

$$\begin{array}{lllll} 1) & g \circ f \text{ injektiv} & \Rightarrow & f \text{ injektiv} \\ 2) & g \circ f \text{ surjektiv} & \Rightarrow & g \text{ surjektiv} \end{array}$$

*Beweis.*

1) $g \circ f$ injektiv. Aus $f(x) = f(x')$ folgt $g(f(x)) = g(f(x'))$, weil $g$ eine Funktion. Daraus folgt, weil $g \circ f$ injektiv, dass $x = x'$. Wir haben also gezeigt: $f(x) = f(x') \Rightarrow x = x'$, also $f$ injektiv.

2) $g \circ f$ surjektiv. Es gilt $g(Y) \subseteq Z$, weil $g : Y \to Z$. Noch z. zg.: $Z \subseteq g(Y)$. Weil $g \circ f$ surjektiv, ist $(g \circ f)(X) = g(f(X)) = Z$. Aber $f(X) \subseteq Y$ und damit $g(f(X)) \subseteq g(Y)$, weil $g$ Funktion, also $Z \subseteq g(Y)$. Aus $g(Y) \subseteq Z$ und $Z \subseteq g(Y)$ folgt $Z = g(Y)$, also $g$ surjektiv. □

**Definition 2.5.6** (Umkehrfunktion)**.** Eine Abbildung $g : Y \to X$ heißt *Umkehrabbildung* oder *Umkehrfunktion* von $f : X \to Y$ genau dann, wenn gilt:

$$1) \quad g \circ f = \mathrm{id}_X, \quad \text{d.h.} \quad g(f(x)) = x \quad \forall x \in X$$

$$\text{und } 2) \quad f \circ g = \mathrm{id}_Y, \quad \text{d.h.} \quad f(g(y)) = y \quad \forall y \in Y$$

Wir schreiben dann: $g =: f^{-1}$.

**Beispiel 2.5.5.** Betrachte $f : \mathbb{R} \to \mathbb{R}, f(x) = 3x + 5$. Für die Umkehrfunktion $f^{-1}$ muss gelten $f \circ f^{-1} = \mathrm{id}_Y$, also:

$$f(f^{-1}(y)) = y \quad \Rightarrow 3f^{-1}(y) + 5 = y \quad \Rightarrow f^{-1}(y) = \frac{y-5}{3}$$

Wir überprüfen, ob $f^{-1}$ das notwendige Kriterium $f^{-1} \circ f = \mathrm{id}_X$ erfüllt:

$$f^{-1}(f(x)) = \frac{f(x)-5}{3} = \frac{3x+5-5}{3} = x, \quad \text{also } f^{-1} \circ f = \mathrm{id}_X$$

**Satz 2.5.3** (Umkehrfunktion)**.** *Falls die Umkehrfunktion $f^{-1}$ existiert, so ist sie eindeutig und bijektiv. Auch $f$ ist in dem Fall bijektiv.*

*Beweis.* Bijektivität von $f, f^{-1}$:
Weil $\mathrm{id}_X$ bijektiv $\Rightarrow f^{-1} \circ f$ bijektiv $\Rightarrow f$ injektiv, $f^{-1}$ surjektiv.
Weil $\mathrm{id}_Y$ bijektiv $\Rightarrow f \circ f^{-1}$ bijektiv $\Rightarrow f^{-1}$ injektiv, $f$ surjektiv.
$\Rightarrow f, f^{-1}$ bijektiv
Eindeutigkeit von $f^{-1}$:
Seien $f_1^{-1} : Y \to X$ und $f_2^{-1} : Y \to X$ zwei verschiedene Umkehrfunktionen zu $f$. Dann gilt für jedes $x \in X$:

$$f_1^{-1}(f(x)) = x \quad \wedge \quad f_2^{-1}(f(x)) = x$$

Weil $f$ surjektiv, d.h. $f(X) = Y$, existiert für jedes $y \in Y$ ein $x \in X$ mit $y = f(x)$. Also ist $f_1^{-1}(y) = f_2^{-1}(y)$ für jedes $y \in Y$ (wäre $f$ nicht surjektiv, könnte es $y \in Y$ geben, auf denen $f_1^{-1}$ und $f_2^{-1}$ nicht übereinstimmen). Folglich $f_1^{-1} = f_2^{-1}$. □

**Satz 2.5.4** (Eigenschaften von $f^{-1}$)**.** *Seien $f : X \to Y$ und $g : Y \to Z$ Funktionen. Es gilt:*

1) $f$ bijektiv $\Leftrightarrow f$ besitzt eine Umkehrfunktion
2) $f$ bijektiv $\Rightarrow f^{-1}$ bijektiv und $(f^{-1})^{-1} = f$
3) $f, g$ bijektiv $\Rightarrow g \circ f$ bijektiv und $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

*Beweis.*

1) „⇐" wurde bereits gezeigt (2.5.3). „⇒": Weil $f$ surjektiv ist, existiert $\forall y \in Y$ ein $x \in X$ mit $f(x) = y$, und weil $f$ injektiv ist, ist dieses $x$ sogar eindeutig (denn für $x, x' \in X$ mit $f(x) = y$ und $f(x') = y$ folgt, wegen $f(x) = f(x')$, dass $x = x'$), also:

$$\forall y \in Y \, \exists! x \in X : f(x) = y.$$

Aber dies definiert genau eine Funktion $f^{-1} : Y \to X, f^{-1}(y) := x$. Dies ist die Umkehrfunktion, denn sie erfüllt: $f(f^{-1}(y)) = y$ für alle $y \in Y$ und $f^{-1}(f(x)) = x$ für alle $x \in X$.

Anschaulich: Im Fall einer bijektiven Funktion $f : X \to Y$, dargestellt durch Pfeile von $X$ nach $Y$, endet bei jedem Element in $Y$ genau *ein* Pfeil. Dreht man diesen Pfeil um, so

endet bei jedem Element in $X$ genau *ein* Pfeil. Dies ist die (wieder bijektive) Umkehrabbildung (Umkehrfunktion).

2) Dass $f^{-1}$ existiert, folgt aus 1). Daraus folgt, dass $f^{-1}$ bijektiv ist (siehe Satz 2.5.3). Wir erhalten außerdem aus Satz 2.5.3, dass $f^{-1}$ eindeutig ist. Daraus folgt, dass $(f^{-1})^{-1} = f$ (denn $f$ ist ja nach Definition der Umkehrfunktion bereits Umkehrfunktion zu $f^{-1}$).

3) Aus der Definition der Umkehrfunktion und der Assoziativität der Verkettung (siehe Satz 2.5.1) folgt:

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \mathrm{id}_Y \circ f = f^{-1} \circ f = \mathrm{id}_X.$$

Anders herum gilt:

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathrm{id}_Y \circ g^{-1} = g \circ g^{-1} = \mathrm{id}_Z.$$

$f^{-1} \circ g^{-1}$ ist also Umkehrfunktion von $g \circ f$ und weil sie existiert, ist sie auch eindeutig und bijektiv und $g \circ f$ ist auch bijektiv (siehe Satz 2.5.3). □

## 2.6 Unendliche Mengen

**Definition 2.6.1** (Gleichmächtigkeit)**.** Eine Menge $M$ heißt *gleichmächtig* zu einer anderen Menge $N$ genau dann, wenn eine bijektive Abbildung $\phi : M \to N$ existiert. Wir schreiben dann:

$$|M| = |N|.$$

Wir können nun den Begriff der endlichen Menge präzisieren und den Begriff der unendlichen Menge einführen:

**Definition 2.6.2** (Endliche, abzählbar und überabzählbar unendliche Mengen)**.** Sei $M$ eine nicht-leere Menge und $n \in \mathbb{N}$. Dann gilt:

1) $|\emptyset| = 0$

2) $M$ hat $n$ Elemente, $|M| = n \Leftrightarrow \exists \phi : \{1, ..., n\} \to M$ bijektiv

3) $M$ heißt *endlich* $\Leftrightarrow \exists n \in \mathbb{N}$ mit $|M| = n$ (sonst heißt $M$ unendlich)

4) $M$ heißt *abzählbar unendlich* $\Leftrightarrow M$ gleichmächtig zu $\mathbb{N}$

5) $M$ heißt *überabzählbar unendlich*, falls $M$ nicht endlich und $M$ nicht abzählbar unendlich

**Bemerkung 2.6.1** ($\mathbb{N}$ oder $\mathbb{N}_0$)**.** Ist $M$ gleichmächtig zu $\mathbb{N}$, so ist $M$ auch gleichmächtig zu $\mathbb{N}_0$, denn wenn $\phi : M \to \mathbb{N}$ bijektiv und $\phi' : \mathbb{N} \to \mathbb{N}_0, n \to n-1$ bijektiv, dann $\phi' \circ \phi : M \to \mathbb{N}_0$ bijektiv (und umgekehrt sind $\mathbb{N}_0$ und $\mathbb{N}$ dann auch gleichmächtig zu $M$). In dem Fall: $|M| = |\mathbb{N}| = |\mathbb{N}_0|$.

**Bemerkung 2.6.2.** Für endliche Mengen gilt: Wenn $M \subseteq N$ und $|M| = |N|$, dann folgt $M = N$. Achtung: Dies gilt nicht für (abzählbar oder überabzählbar) unendliche Mengen!

**Beispiel 2.6.1.** Betrachte $\mathbb{N}$ und $3\mathbb{N} = \{n \in \mathbb{N} | n = 3k, k \in \mathbb{N}\} = \{3, 6, 9, 12, ..\}$. Dann gilt: $3\mathbb{N} \subset \mathbb{N}$ und $|3\mathbb{N}| = |\mathbb{N}|$ (denn es existiert $\phi : \mathbb{N} \to 3\mathbb{N}, n \to 3n$ bijektiv), aber $3\mathbb{N} \neq \mathbb{N}$.

**Satz 2.6.1.** $\mathbb{N}, \mathbb{Z}$ *und* $\mathbb{Q}$ *sind abzählbar unendlich.*

*Beweis.*

- $\mathbb{N}$ ist trivialer Weise abzählbar (unendlich), denn $\phi = \mathrm{id}_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}, n \to n$ ist bijektiv.
- $\mathbb{Z}$ ist abzählbar (unendlich), denn es existiert eine bijektive Abbildung

$$\phi' : \mathbb{N}_0 \to \mathbb{Z} : n \to \frac{1 - (-1)^n(2n+1)}{4}.$$

  Damit werden die ganzen Zahlen anschaulich wie folgt „abgezählt":

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | ... |
|---|---|---|---|---|---|---|---|
| 0 | +1 | –1 | +2 | –2 | +3 | –3 | ... |

- Dass $\mathbb{Q}$ abzählbar (unendlich) ist, folgt aus einer ähnlichen Überlegung wie im Fall von $\mathbb{Z}$ sowie einer besonderen „Abzählung". Diese Abzählung ist bekannt als Cantors erstes Diagonalargument. Sie geht wie folgt:

$$\begin{pmatrix}
1 & \to & \frac{1}{2} & & \frac{1}{3} & \to & \frac{1}{4} & \dots \\
 & \swarrow & & \nearrow & & \swarrow & & \\
\frac{2}{1} & & \frac{2}{2} & & \frac{2}{3} & & & \dots \\
\downarrow & \nearrow & & \swarrow & & & & \\
\frac{3}{1} & & \frac{3}{2} & & & \dots & & \dots \\
 & \swarrow & & & & & & \\
\frac{4}{1} & & & \dots & & \dots & & \dots \\
\downarrow & & & & & & & \\
\vdots & & \vdots & & \vdots & & \vdots & 
\end{pmatrix}$$

  Zur Abzählung aller rationalen Zahlen $\mathbb{Q}$ werden die nicht vollständig gekürzten Brüche übersprungen, die 0 wird hinzugefügt und nach jeder positiven rationalen Zahl wird die entsprechend negative rationale Zahl eingefügt. Damit ergibt sich:

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | +1 | –1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | +2 | –2 | + 3 | –3 | $+\frac{1}{3}$ | $-\frac{1}{3}$ | $+\frac{1}{4}$ | $-\frac{1}{4}$ | $+\frac{2}{3}$ | $-\frac{2}{3}$ | ... |

□

Es gibt noch eine weitere bekannte Menge, die abzählbar unendlich ist und das ist die Menge der Primzahlen $\mathbb{P}$.

**Definition 2.6.3** (Primzahl)**.** $p$ ist Primzahl $:\Leftrightarrow p \in \mathbb{N}, p \geq 2$ und $\forall a \in \mathbb{N} : a|p \Rightarrow a = 1 \vee a = p$.

**Satz 2.6.2.** *Die Menge der Primzahlen* $\mathbb{P}$ *ist abzählbar unendlich.*

*Beweis.* Einerseits gilt: $\mathbb{P} \subset \mathbb{N}$ (echte Teilmenge), d.h. $|\mathbb{P}| \leq |\mathbb{N}|$. Andererseits gilt: Es gibt unendlich viele Primzahlen (Beweis von Euklid, Übung). Daraus folgt: $|\mathbb{P}| = |\mathbb{N}|$. Insbesondere können wir die Primzahlen $\mathbb{P}$ als Teilmenge von $\mathbb{N}$ der Größe nach anordnen (siehe Bemerkung 2.2.2 zum Wohlordnungsprinzip von $\mathbb{N}$) mit $p_1 < p_2 < p_3 < ...$ . Damit können wir sie „abzählen", d.h. $\exists\ \phi$ bijektiv: $\mathbb{N} \to \mathbb{P}, i \to p_i$. □

**Bemerkung 2.6.3.** Der Beweis von Euklid, dass es unendlich viele Primzahlen gibt, stützt sich darauf, dass jede natürliche Zahl $n \in \mathbb{N}$, $n \geq 2$, als Produkt von Primzahlen geschrieben werden kann (dabei sprechen wir auch bei einem Faktor von einem Produkt). Daraus folgt auch der Fundamentalsatz der Arithmetik.

**Satz 2.6.3** (Fundamentalsatz der Arithmetik)**.** *Jede natürliche Zahl $n \in \mathbb{N}, n \geq 2$, ist eindeutig als Produkt endlich vieler Primzahlen darstellbar.*

*Beweis.* Existenz und Eindeutigkeit (Übung). □

Tatsächlich gibt es aber auch Mengen, die überabzählbar unendlich sind. Welche Mengen sind überabzählbar unendlich? Wie kann man das zeigen?

**Satz 2.6.4.** *Die Menge der reellen Zahlen $\mathbb{R}$ ist überabzählbar unendlich.*

*Beweis.* Cantors zweites Diagonalargument (Übung). □

**Bemerkung 2.6.4** ($\aleph$, Kontinuumshypothese)**.** Cantor zeigt also: $|\mathbb{R}| > |\mathbb{N}|$.

Wir bezeichen die Mächtigkeit von unendlichen Mengen mit dem hebräischen Buchstaben „Aleph". Für die kleinste solche Menge schreiben wir: $|\mathbb{N}| = \aleph_0$.

Wie groß ist $|\mathbb{R}|$? Ist $|\mathbb{R}| = \aleph_1$ (also die nächstgrößere Menge = *Kontinuumshypothese*)? Oder gibt es dazwischen noch etwas, also eine Menge $M$, deren Mächtigkeit zwischen der von $\mathbb{N}$ und $\mathbb{R}$ liegt: $|\mathbb{N}| < |M| < |\mathbb{R}|$?

Gödel beweist 1938: Die Kontinuumshypothese lässt sich auf Basis der axiomatischen Mengenlehre nicht widerlegen. Cohen zeigt 1960: Die Kontinuumshypothese lässt sich auf Basis der axiomatischen Mengenlehre nicht beweisen.

**Satz 2.6.5.** *Die Mächtigkeit von $\mathbb{R}$ ist gleich der Mächtigkeit der Menge der 0-1-Folgen. Insbesondere ist die Menge der 0-1-Folgen also auch überabzählbar unendlich.*

*Beweis.* Man zeigt Gleichmächtigkeit, indem man zeigt, dass es eine 1:1-Abbildung zwischen beiden Mengen gibt. Dazu wechselt man von der Darstellung reller Zahlen im Dezimalsystem zur Darstellung reeller Zahlen $r \in (0,1)$ im Binärsystem, also in der Form $0, d_1 d_2 d_3 ....$ mit $d_k \in \{0,1\}$ für alle $k \in \mathbb{N}$ und $r = \sum_{k=1}^{\infty} d_k 2^{-k} = d_1 \cdot \frac{1}{2} + d_2 \cdot \frac{1}{4} + d_3 \cdot \frac{1}{8} + d_4 \cdot \frac{1}{16} + ...$ In dieser Darstellung ist z. B. $0,00101100001....$ eine reelle Zahl $r \in (0,1)$. So kann man jede reelle Zahl $r \in (0,1)$ mit einer 0-1-Folge identifizieren (und dieses Argument erweitert sich auf ganz $\mathbb{R}$, denn für $r \notin (0,1)$ kommen zu der unendlichen 0-1-Folge nur endlich viele Folgeglieder hinzu). □

**Bemerkung 2.6.5** (Exkurs zum Binärsystem)**.** Betrachten wir zuerst $\mathbb{N}$. Normalerweise werden Zahlen $n \in \mathbb{N}$ im Dezimalsystem dargestellt, d.h. in der Form $c_m ... c_3 c_2 c_1 c_0$ mit $c_i \in \{0, ..., 9\}$ für alle $0 \leq i \leq m$ (hier ist $c_i$ die Ziffer der Zahl an der $i$-ten Stelle). Der Wert der Zahl ist in dem Fall $n = \sum_{i=0}^{m} c_i \cdot 10^i$ mit $m \in \mathbb{N}, c_i \in \{0, ..., 9\}$.

Dieselben Zahlen können im Binärsystem dargestellt werden in der Form $d_l ... d_3 d_2 d_1 d_0$ mit $d_k \in \{0, 1\}$ für alle $0 \leq k \leq l$. Der Wert der Zahl ist in dem Fall $n = \sum_{k=0}^{l} d_k \cdot 2^k$ mit $l \in \mathbb{N}, d_k \in \{0, 1\}$. Im Vergleich ergibt sich:

| Dezimalsystem | 0 | 1 | 2 | 3 | 4 | 5 | 6 | ... |
|---|---|---|---|---|---|---|---|---|
| Binärsystem | 0 | 1 | 10 | 11 | 100 | 101 | 110 | ... |

Reelle Zahlen haben in dieser Darstellung auch Summanden mit negativen Koeffizienten, wobei die Summe bei rationalen Zahlen nach endlich vielen (negativen) Gliedern abbricht oder sich wiederholt (z. B. ist $\frac{1}{5}$ im Binärsystem $0, \overline{0011}$). Für irrationale Zahlen bricht die Summe nie ab und wiederholt sich nie. D.h. allgemein für $r \in \mathbb{R}_0^+$: $r = \sum_{k=-n}^{m} d_k \cdot 2^k$ (mit $n \in \mathbb{N}$ oder $n \to \infty$).

# 3 Gruppen, Ringe, Körper

## 3.1 Gruppen

**Definition 3.1.1** (Verknüpfung)**.** Sei $M$ eine Menge. Eine Verknüpfung $\circ$ auf $M$ ist eine Abbildung $\circ : M \times M \to M, (a, b) \to a \circ b$.

**Bemerkung 3.1.1.** Verknüpfungen sind etwa die Addition „+" und Multiplikation „·" auf $\mathbb{N}$. Achtung: Die Subtraktion „–" und Division „:" sind *keine* Verknüpfungen auf $\mathbb{N}$!

**Definition 3.1.2** (Gruppe)**.** Eine Menge $G$ mit einer Verknüpfung $\circ$ heißt *Gruppe* $(G, \circ)$ genau dann, wenn gilt:

1) $\forall a, b, c \in G : (a \circ b) \circ c = a \circ (b \circ c)$ (Assoziativgesetz)

2) $\exists e \in G$, sodass $\forall a \in G : a \circ e = e \circ a = a$ (Existenz eines neutralen Elements)

3) $\forall a \in G \, \exists a^{-1} \in G$: $a \circ a^{-1} = a^{-1} \circ a = e$ (Existenz eines inversen Elements zu $a$)

$(G, \circ)$ heißt *abelsche* oder *kommutative* Gruppe, wenn zusätzlich gilt:

4) $\forall a, b \in G : a \circ b = b \circ a$ (Kommutativgesetz)

**Bemerkung 3.1.2.** Tatsächlich genügt in dieser Definition die Existenz eines *rechtsneutralen* und *rechtsinversen* Elements (oder analog eines linksneutralen und linksinversen Elements):

2') $\exists e \in G$, sodass $\forall a \in G : a \circ e = a$ (Existenz eines rechtsneutralen Elements)

3') $\forall a \in G \, \exists a^{-1} \in G : a \circ a^{-1} = e$ (Existenz eines rechtsinversen Elements)

**Beispiel 3.1.1.**

- $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +)$ sind abelsche Gruppen. Neutrales Element ist jeweils $e = 0$ und inverses Element zu $a$ ist $-a$. Dagegen ist $(\mathbb{N}, +)$ *keine* Gruppe ($\nexists$ inverses Element).
- $(\mathbb{Q} \setminus \{0\}, \cdot), (\mathbb{R} \setminus \{0\}, \cdot)$ sind abelsche Gruppen. Neutrales Element ist jeweils $e = 1$ und inverses Element zu $a$ ist $\frac{1}{a}$. Dagegen ist $(\mathbb{Z} \setminus \{0\}, \cdot)$ *keine* Gruppe ($\nexists$ inverses Element).
- $(\mathbb{Z}^n, +), (\mathbb{Q}^n, +), (\mathbb{R}^n, +)$ mit der komponentenweise definierten Addition

  $$(a_1, ..., a_n) + (b_1, ..., b_n) := (a_1 + b_1, ..., a_n + b_n)$$

  sind abelsche Gruppen. Neutrales Element ist $e = (0, ..., 0)$ und inverses Element zu $(a_1, ..., a_n)$ ist $(-a_1, ..., -a_n)$.

**Satz 3.1.1** (Eindeutigkeit $e$ und $a^{-1}$)**.** *Das neutrale Element $e$ einer Gruppe ist, sofern es existiert, eindeutig. Ebenso ist das inverse Element $a^{-1}$ zu $a$, sofern es existiert, eindeutig.*

*Beweis.* Angenommen $e$ und $e'$ sind neutrale Elemente. Dann gilt: $e \circ e' = e' \circ e = e'$ (weil $e$ neutrales Element) und $e' \circ e = e \circ e' = e$ (weil $e'$ neutrales Element). Folglich: $e = e'$.

Sei $a^{-1}$ inverses Element zu $a$. Angenommen $a'^{-1}$ ist weiteres inverses Element zu $a$. Dann gilt: $a \circ a'^{-1} = a'^{-1} \circ a = e$. Damit folgt (unter Verwendung der Gruppenaxiome):

$$a^{-1} = a^{-1} \circ e = a^{-1} \circ (a \circ a'^{-1}) = (a^{-1} \circ a) \circ a'^{-1} = e \circ a'^{-1} = a'^{-1}.$$

□

**Definition 3.1.3** (Permutationsgruppe)**.** Sei $M = \{1, ..., n\}$. Man nennt die Menge der bijektiven Abbildungen auf $M$, i.e. $S_n := \{\phi : M \to M \text{ bijektiv}\}$, zusammen mit der Funktionsverkettung $\circ$ die *Permutationsgruppe*.

Die bijektiven Abbildungen $\phi : \{1, ..., n\} \to \{1, ..., n\}, i \to \phi(i)$ nennt man *Permutationen*.

**Bemerkung 3.1.3.** Die Permutationsgruppe erfüllt die Gruppenaxiome (Übung).

**Satz 3.1.2.** *Für die Mächtigkeit von $S_n$ gilt:*

$$|S_n| = n!$$

*Beweis.* Es gibt insgesamt $n$ Möglichkeiten, welchen Wert $\phi(1)$ annehmen kann, mal $(n-1)$ Möglichkeiten, welchen Wert $\phi(2)$ annehmen kann, usw. □

**Bemerkung 3.1.4.** Für eine gegebene Permutation $\pi : \{1, ..., n\} \to \{1, ..., n\}, i \to \pi(i)$ schreiben wir:

$$\begin{Bmatrix} 1 & 2 & 3 & ... & n \\ \pi(1) & \pi(2) & \pi(3) & ... & \pi(n) \end{Bmatrix}$$

Tatsächlich reicht es zur Spezifikation einer Permutation auch aus, nur die untere Reihe dieser Klammer in Form eines $n$-Tupels anzugeben, also $(\pi(1), \pi(2), ..., \pi(n))$.

**Beispiel 3.1.2** (Nicht-kommutative Gruppe)**.** Seien $\rho : \{1, ..., 5\} \to \{1, ..., 5\}, i \to \rho(i)$ und $\sigma : \{1, ..., 5\} \to \{1, ..., 5\}, j \to \sigma(j)$ Permutationen mit

$$\rho : \begin{Bmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{Bmatrix}, \qquad \sigma : \begin{Bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \end{Bmatrix}$$

Dann sind $\rho \circ \sigma$ und $\sigma \circ \rho$ Permutationen mit

$$\rho \circ \sigma : \begin{Bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{Bmatrix}, \qquad \sigma \circ \rho : \begin{Bmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{Bmatrix}$$

Insbesondere ist also $\rho \circ \sigma \neq \sigma \circ \rho$, d.h. die Permutationsgruppe ist nicht kommutativ!

**Satz 3.1.3.**

1) *Sei $m \in \mathbb{N}$. $(\mathbb{Z}_m, +)$ mit der Verknüpfung $[a]_m + [b]_m = [a+b]_m$ ist eine kommutative Gruppe mit $m$ Elementen.*

2) *Sei $p$ Primzahl. $(\mathbb{Z}_p \setminus \{[0]_p\}, \cdot)$ mit der Verknüpfung $[a]_p \cdot [b]_p = [a \cdot b]_p$ ist eine kommutative Gruppe mit $p-1$ Elementen.*

*Beweis.*

Zu 1): $\mathbb{Z}_m = \{[0]_m, [1]_m, ..., [m-1]_m\}$, d.h. $\mathbb{Z}_m$ hat $m$ Elemente. Dabei ist $[a]_m = a + m \cdot \mathbb{Z} = \{a + m \cdot k | k \in \mathbb{Z}\}$.
Zuerst zeigen wir, dass $[a]_m + [b]_m = [a+b]_m$ tatsächlich wieder auf $\mathbb{Z}_m$ abbildet, d. h. dass $+ : \mathbb{Z}_m \times \mathbb{Z}_m \to \mathbb{Z}_m$. Dies sieht man wie folgt: $[a]_m = [a + km]_m \ \forall k \in \mathbb{Z}$ und $[b]_m = [b + lm]_m$ $\forall l \in \mathbb{Z}$. Um auf $\mathbb{Z}_m$ abzubilden, muss gelten, dass $[a+b]_m = [a + km + b + lm]_m \ \forall k, l \in \mathbb{Z}$. Dies ist tatsächlich der Fall, denn $[a+b]_m = [a+km+b+lm]_m \Leftrightarrow m \big| (a+km+b+lm-(a+b)) \Leftrightarrow m \big| (km+lm)$ und letztere ist eine wahre Aussage.

Nachweis der Gruppeneigenschaften:

Assoziativität:

$$\begin{aligned} ([a]_m + [b]_m) + [c]_m &= [a+b]_m + [c]_m = [(a+b)+c]_m = [a+(b+c)]_m \\ &= [a]_m + [b+c]_m = [a]_m + ([b]_m + [c]_m) \end{aligned}$$

Neutrales Element: $[0]_m$

$$[0]_m + [a]_m = [0+a]_m = [a]_m = [a+0]_m = [a]_m + [0]_m$$

Inverses Element: $[-a]_m$

$$[-a]_m + [a]_m = [-a+a]_m = [0]_m = [a+(-a)]_m = [a]_m + [-a]_m$$

Kommutativität:

$$[a]_m + [b]_m = [a+b]_m = [b+a]_m = [b]_m + [a]_m$$

Zu 2): Weitgehend analog. Man muss sich allerdings genau überlegen, wie das inverse Element aussieht (Übung).

□

**Bemerkung 3.1.5.** In den folgenden Beispielen schreiben wir $\bar{a} := [a]_m$, falls $[a]_m \in \mathbb{Z}_m$ und $a \in \{0, 1, ..., m-1\}$.

**Beispiel 3.1.3.** Verknüpfungstafel von $(\mathbb{Z}_4, +)$.

| $+$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ |
|---|---|---|---|---|
| $\bar{0}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ |
| $\bar{1}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{0}$ |
| $\bar{2}$ | $\bar{2}$ | $\bar{3}$ | $\bar{0}$ | $\bar{1}$ |
| $\bar{3}$ | $\bar{3}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ |

**Beispiel 3.1.4.** Verknüpfungstafel von $(\mathbb{Z}_5 \setminus \{\bar{0}\}, \cdot)$.

| $\cdot$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
|---|---|---|---|---|
| $\bar{1}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
| $\bar{2}$ | $\bar{2}$ | $\bar{4}$ | $\bar{1}$ | $\bar{3}$ |
| $\bar{3}$ | $\bar{3}$ | $\bar{1}$ | $\bar{4}$ | $\bar{2}$ |
| $\bar{4}$ | $\bar{4}$ | $\bar{3}$ | $\bar{2}$ | $\bar{1}$ |

**Definition 3.1.4** (Homomorphismus, Isomorphismus, Automorphismus)**.** Seien $(G, \circ)$ und $(H, \star)$ Gruppen. Eine Abbildung $\phi : G \to H$ heißt

1) *Homomorphismus* genau dann, wenn $\forall a, b \in G : \phi(a \circ b) = \phi(a) \star \phi(b)$,

2) *Isomorphismus* genau dann, wenn $\phi$ Homomorphismus und $\phi$ bijektiv ist. In dem Fall nennen wir $(G, \circ)$ und $(H, \star)$ isomorphe Gruppen.

3) Ist $\phi : G \to G$ ein Isomorphismus einer Gruppe $G$ auf sich selbst, dann nennen wir $\phi$ einen *Automorphismus*.

**Bemerkung 3.1.6** (Isomorphismus)**.** Man bezeichnet zwei isomorphe Strukturen $S$ und $S'$ (Gruppen, später auch Körper, Vektorräume etc.) als „strukturgleich" und schreibt: $S \cong S'$.

**Bemerkung 3.1.7** (Automorphismus)**.** Jede Struktur $S$ hat einen trivialen Automorphismus, die Identität $\mathrm{id} : x \to x$. Vorwegnahme: Tatsächlich gibt es auf den Körpern $\mathbb{Q}$ und $\mathbb{R}$ nur den trivialen Automorphismus, erst auf $\mathbb{C}$ gibt es einen nichttrivialen Automorphismus, nämlich die komplexe Konjugation $z = x + iy \to \bar{z} = x - iy$.

**Beispiel 3.1.5.** $(\mathbb{R}, +)$ und $(\mathbb{R}^+, \cdot)$ sind isomorph, denn es existiert eine bijektive Abbildung $\phi : \mathbb{R} \to \mathbb{R}^+, a \to e^a$ mit

$$\phi(a + b) = e^{a+b} = e^a \cdot e^b = \phi(a) \cdot \phi(b)$$

für alle $a, b \in \mathbb{R}$.

**Beispiel 3.1.6.** $(\mathbb{Z}_4, +)$ und $(\mathbb{Z}_5 \setminus \{\bar{0}\}, \cdot)$ sind isomorph, denn es existiert eine bijektive Abbildung

$$\phi : (\mathbb{Z}_4, +) \to (\mathbb{Z}_5 \setminus \{\bar{0}\}, \cdot), \quad [k]_4 \to [2^k]_5 \quad \text{mit} \quad k \in \{0, 1, 2, 3\},$$

welche die Gleichung $\phi([k]_4 + [l]_4) = \phi([k]_4) \cdot \phi([l]_4)$ für alle $k, l \in \{0, 1, 2, 3\}$ erfüllt.

## 3.2 Ringe

**Definition 3.2.1** (Ring)**.** Sei $R$ eine Menge mit zwei Verknüpfungen:

i) $+ : R \times R \to R, \quad (a, b) \to a + b$ („Addition")

ii) $\cdot : R \times R \to R, \quad (a, b) \to a \cdot b$ („Multiplikation")

$(R, +, \cdot)$ heißt *Ring* genau dann, wenn gilt:

1) $(R, +)$ ist eine kommutative Gruppe

2) Die Multiplikation $\cdot$ ist assoziativ (oder äquivalent: $(R, \cdot)$ ist „Halbgruppe"), d.h. es gilt $\forall a, b, c \in R : (a \cdot b) \cdot c = a \cdot (b \cdot c)$

3) Es gelten die Distributivgesetze, d.h.

$\forall a, b, c \in R : (a + b) \cdot c = (a \cdot c) + (b \cdot c) \quad \text{und} \quad c \cdot (a + b) = (c \cdot a) + (c \cdot b)$

$(R, +, \cdot)$ heißt *kommutativer Ring*, wenn zusätzlich gilt:

4) $\forall a, b \in R$: $a \cdot b = b \cdot a$

**Bemerkung 3.2.1** (Nullelement, inverses Element)**.** Das neutrale Element der Addition nennen wir *Nullelement* und schreiben 0. Für das inverse Element der Addition zu $a$ schreiben wir $-a$. Damit können wir die *Subtraktion* definieren als $a - b := a + (-b)$. (Achtung: Die Division kommt erst später auf Körpern!)

**Bemerkung 3.2.2.** Statt $(a \cdot c) + (b \cdot c)$ schreiben wir kurz $a \cdot c + b \cdot c$ oder $ac + bc$. Dabei bedeutet unsere Schreibweise: „Punkt vor Strich" und „*kein* Zeichen gleich Punkt".

**Definition 3.2.2** (Einselement)**.** Sei $(R, +, \cdot)$ ein Ring. Ein neutrales Element der Multiplikation $e \in R$ heißt *Einselement*. Statt $e$ schreiben wir in dem Fall 1 und dann gilt $\forall a \in R$:

$$1 \cdot a = a \cdot 1 = a$$

**Bemerkung 3.2.3** (Eindeutigkeit)**.** Falls ein Einselement existiert, so ist es eindeutig (Beweis dazu ist analog zum neutralen Element bei Gruppen, siehe Satz 3.1.1).

**Definition 3.2.3** (Nullteilerfrei)**.** Sei $(R, +, \cdot)$ ein Ring. Wir nennen $R$ *nullteilerfrei* genau dann, wenn $\forall a, b \in R$ gilt:

$$a \cdot b = 0 \Rightarrow a = 0 \vee b = 0$$

**Beispiel 3.2.1.**

- $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer, nullteilerfreier Ring mit Einselement.
- $(m\mathbb{Z}, +, \cdot)$ mit $m \geq 2, m \in \mathbb{N}$ ist ein kommutativer, nullteilerfreier Ring *ohne* Einselement.
- $(\mathbb{N}, +, \cdot)$ ist *kein* Ring (Erinnerung: $(\mathbb{N}, +)$ ist ja bereits keine Gruppe mangels inversen Elements!)
- $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind kommutative, nullteilerfreie Ringe mit Einselement und es sind sogar Körper (mehr dazu später)!
- Ein Beispiel für einen kommutativen Ring mit Einselement, der *nicht* nullteilerfrei ist, ist der Quotientenraum $(\mathbb{Z}_m, +, \cdot)$ mit $m \geq 2$ und $m$ keine Primzahl

**Satz 3.2.1.** *$(\mathbb{Z}_m, +, \cdot)$ mit $m \geq 2$ ist ein kommutativer Ring mit Einselement. Er ist nullteilerfrei genau dann, wenn $m$ eine Primzahl ist.*

*Beweis.* $(\mathbb{Z}_m, +)$ ist kommutative Gruppe (siehe Satz 3.1.3). Das Nullelement ist $[0]_m$.

$(\mathbb{Z}_m, \cdot)$ ist Halbgruppe, d.h. die Multiplikation ist assoziativ. Sie ist auch kommutativ und die Distributivgesetze gelten (siehe Satz 3.1.3 bzw. Übung). Das Einselement ist $[1]_m$.

1. Fall: Angenommen $m$ ist keine Primzahl. Dann $\exists a, b \in \mathbb{Z}, a, b \geq 2$ mit $m = a \cdot b$. Daraus folgt:

$$[a]_m \cdot [b]_m = [a \cdot b]_m = [m]_m = [0]_m.$$

Wäre nun $(\mathbb{Z}_m, +, \cdot)$ nullteilerfrei, dann müste aus dieser Bedingung folgen, dass $[a]_m = [0]_m$ oder $[b]_m = [0]_m$. Dies ist aber nicht der Fall. Tatsächlich gilt: $[a]_m \neq [0]_m$, denn $m$ teilt nicht $a$, und $[b]_m \neq [0]_m$, denn $m$ teilt nicht $b$. $(\mathbb{Z}_m, +, \cdot)$ ist also nicht nullteilerfrei.

2. Fall: Angenommen $m$ ist eine Primzahl. Es gilt $[a]_m \cdot [b]_m = [0]_m$ genau dann, wenn $[a \cdot b]_m = [0]_m$ und dies gilt genau dann, wenn $m|(a \cdot b)$. Daraus folgt, weil $m$ Primzahl ist (siehe Übung): $m|a \vee m|b$. Dies ist genau dann der Fall, wenn $a$ oder $b$ (oder beide) Vielfache von $m$ sind, wenn also $[a]_m = [0]_m \vee [b]_m = [0]_m$. $(\mathbb{Z}_m, +, \cdot)$ ist also nullteilerfrei.

□

**Satz 3.2.2** (Rechenregeln). *Sei $(R, +, \cdot)$ ein Ring, $a, b, c \in R$, dann gelten folgende Rechenregeln:*

1) $0 \cdot a = a \cdot 0 = 0$
2) $-(a \cdot b) = (-a) \cdot b = a \cdot (-b)$
3) $(-a) \cdot (-b) = a \cdot b$
4) *Falls ein Einselement $1 \in R$ existiert:*

$$-a = (-1) \cdot a = 1 \cdot (-a)$$

5) *Falls $R$ nullteilerfrei ist:*

$$(c \neq 0 \wedge a \cdot c = b \cdot c) \Rightarrow a = b \quad \text{und} \quad (c \neq 0 \wedge c \cdot a = c \cdot b) \Rightarrow a = b$$

Aus Rechenregel 5) folgt insbesondere, dass man wie gewohnt beidseitig kürzen kann.

*Beweis.* Zu 1): Sei $a \in R$. Aus dem Distributivgesetz und der Tatsache, dass 0 das neutrale Element der Addition ist, folgt:

$$\begin{aligned} 0 \cdot a + 0 \cdot a &= (0 + 0) \cdot a = 0 \cdot a \\ 0 \cdot a + 0 &= 0 \cdot a \end{aligned}$$

$\Rightarrow 0 \cdot a = 0$ und analog $a \cdot 0 = 0$.

Zu 2): Seien $a, b \in R$. Aus dem Distributivgesetz und der Tatsache, dass $-a$ das inverse Element der Addition zu $a$ ist, folgt:

$$\begin{aligned} a \cdot b + (-a) \cdot b &= (a + (-a)) \cdot b = 0 \cdot b = 0 \\ a \cdot b + (-(a \cdot b)) &= 0 \end{aligned}$$

$\Rightarrow -(a \cdot b) = (-a) \cdot b$ und für die zweite Gleichung analog.

Zu 3) – 5): Übung.

□

## 3.3 Körper

**Definition 3.3.1** (Körper)**.** $(K, +, \cdot)$ heißt Körper genau dann, wenn gilt:

1) $(K, +, \cdot)$ ist ein Ring

2) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe

Für das neutrale Element der Multiplikation schreiben wie 1. Für das inverse Element der Multiplikation zu $a$ schreiben wir $\frac{1}{a}$ oder $a^{-1}$. Damit können wir die *Division* definieren als $\frac{a}{b} := a \cdot \frac{1}{b} = a \cdot b^{-1}$.

**Beispiel 3.3.1.**

- $(\mathbb{Q}, +, \cdot)$ ist ein Körper (der Körper der rationalen Zahlen)
- $(\mathbb{R}, +, \cdot)$ ist ein Körper (der Körper der reellen Zahlen)
- $(\mathbb{Z}, +, \cdot)$ ist *kein* Körper, denn $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe ($\nexists$ inverses Element)

**Bemerkung 3.3.1.** In Körpern gelten die uns gewohnten Rechenregeln der Addition, Subtraktion, Multiplikation und Division!

Insbesondere gelten demnach auf Körpern Sätze bzw. können Beweise geführt werden, wie wir sie etwa bereits im Kapitel zur vollständigen Induktion kennengelernt haben (z.B. geometrische Reihe $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ für $|q| < 1$, $q \in \mathbb{R}$).

**Satz 3.3.1.** *Jeder Körper ist ein kommutativer, nullteilerfreier Ring mit 1 als Einselement.*

*Beweis.* Wir zeigen nacheinander jede der genannten Eigenschaften:

- „kommutativer Ring": Weil $(K \setminus \{0\}, \cdot)$ eine kommutative Gruppe ist, gilt

  $$\forall a, b \in K \setminus \{0\} : a \cdot b = b \cdot a$$

  und weil $(K, +, \cdot)$ Ring, gilt $\forall a \in K : 0 \cdot a = a \cdot 0$ (siehe Aussage 1) von Satz 3.2.2)

  $\Rightarrow (K, +, \cdot)$ kommutativer Ring

- „1 ist Einselement": Weil $(K \setminus \{0\}, \cdot)$ eine kommutative Gruppe ist, gilt weiter

  $$\forall a \in K \setminus \{0\} : a \cdot 1 = 1 \cdot a = a$$

  und $0 \cdot 1 = 1 \cdot 0 = 0$ (siehe wieder Aussage 1) von Satz 3.2.2) $\Rightarrow (K, +, \cdot)$ ist Ring mit 1 als Einselement

- „nullteilerfrei": Seien $a, b \in K$. Sei $a \cdot b = 0$. Z. zg.: $a = 0 \vee b = 0$.

  Wir machen eine Fallunterscheidung. 1. Fall: $a = 0 \Rightarrow$ fertig

  2. Fall: $a \neq 0 \Rightarrow b = 1 \cdot b = (a^{-1} \cdot a) \cdot b = a^{-1} \cdot (a \cdot b) = 0$

  Hier folgt der letzte Schritt daraus, dass $a \cdot b = 0$.

□

**Bemerkung 3.3.2.** Jeder Körper hat mindestens 2 Elemente, das neutrale Element der Addition „0" und das neutrale Element der Multiplikation „1". Wir bezeichnen den Körper, der aus genau diesen zwei Elementen besteht, als Gauloisfeld $GF(2) = \{0, 1\}$.

Auch die komplexen Zahlen $\mathbb{C}$ bilden einen Körper, wenn man sie als $\mathbb{R}^2$ mit den entsprechenden Verknüpfungen der Addition und Multiplikation auffasst.

**Bemerkung 3.3.3** (Historisch)**.** Historisch sind die komplexen Zahlen entstanden als Antwort auf die Frage, wie man Gleichungen der Form $x^2 = -1$ lösen soll. Dieses Problem wurde bereits im 8. Jhd. von al-Chwarizmi in seinem Algebra-Buch benannt. Eingeführt wurden die komplexen Zahlen im 16. Jhd. von Cardano. Erstmals mit der imaginären Zahl $i$ gerechnet hat Euler (1748). Cauchy entwickelte die komplexe Analysis (= Integration und Differentiation im Komplexen), auch *Funktionentheorie* genannt (1814). Der Begriff „komplexe Zahl" stammt von Gauß (1831).

**Definition 3.3.2** (Körper der komplexen Zahlen)**.** Eine komplexe Zahl $z$ ist ein Paar $z = (x, y)$ reeller Zahlen $x, y \in \mathbb{R}$. Die Menge $\mathbb{C} = \{(x, y) | x, y \in \mathbb{R}\} = \mathbb{R}^2$ mit den Verknüpfungen

$$\begin{aligned} + &: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \quad (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2) \\ \cdot &: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \quad (x_1, y_1) \cdot (x_2, y_2) := (x_1 \cdot x_2 - y_1 \cdot y_2, x_1 \cdot y_2 + x_2 \cdot y_1) \end{aligned}$$

wird als Körper der komplexen Zahlen bezeichnet.

**Bemerkung 3.3.4.** Die komplexen Zahlen $\mathbb{C}$ erfüllen die Körperaxiome (Übung).

**Definition 3.3.3** (Imaginäre Einheit)**.** Wir definieren die *imaginäre Einheit* $i := (0, 1)$.

**Bemerkung 3.3.5.** Seien $x, y \in \mathbb{R}$. Mit Hilfe der imaginären Einheit können wir $(x, 0) \in \mathbb{R}^2$ mit $x \in \mathbb{R}$ identifizieren. Dazu bemerken wir, dass gilt:

$$\mathrm{i} \cdot \mathrm{i} = (0, 1) \cdot (0, 1) = (-1, 0)$$

und

$$(x, 0) + \mathrm{i}(y, 0) = (x, 0) + (0, 1) \cdot (y, 0) = (x, 0) + (0, y) = (x, y).$$

Damit kommen wir zu der üblichen Schreibweise für die komplexen Zahlen:

$$\mathbb{C} = \{x + \mathrm{i}y | x, y \in \mathbb{R}\}$$

Demnach ist $z \in \mathbb{C}$ genau dann, wenn $z := x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$. Dabei nennen wir $x$ den Real- und $y$ den Imaginärteil. Die reellen Zahlen sind demnach diejenigen komplexen Zahlen, deren Imaginärteil $y = 0$ ist.

In dieser Schreibweise ist $\mathrm{i}\cdot\mathrm{i} = -1$ und wenn wir dies berücksichtigen, können wir bei der Addition und Multiplikation jetzt wie gewohnt „rechnen":

$$z_1 + z_2 = (x_1 + \mathrm{i}y_1) + (x_2 + \mathrm{i}y_2) = (x_1 + x_2) + \mathrm{i}(y_1 + y_2)$$

$$z_1 \cdot z_2 = (x_1 + \mathrm{i}y_1) \cdot (x_2 + \mathrm{i}y_2) = (x_1 \cdot x_2 - y_1 \cdot y_2) + \mathrm{i}(x_1 \cdot y_2 + y_1 \cdot x_2)$$

Um mit komplexen Zahlen besser umgehen zu können, sind noch ein paar Definitionen nützlich.

**Definition 3.3.4** (Real- und Imaginärteil, Betrag, komplex konjugierte Zahl)**.** Sei $z = x+iy \in \mathbb{C}$. Dann definieren wir wie folgt:

1) $\mathrm{Re}z := x$ (Realteil von $z$)
2) $\mathrm{Im}z := y$ (Imaginärteil von $z$)
3) $|z| := \sqrt{x^2 + y^2}$ (Betrag von $z$)
4) $\bar{z} := x - \mathrm{i}y$ (komplex konjugierte Zahl)

Mit Hilfe dieser Definitionen kann man einige wichtige Gleichungen herleiten.

**Satz 3.3.2.** *Seien $z_1, z_2, z \in \mathbb{C}$. Dann gilt:*

1) $\overline{z_1 + z_2} = \bar{z_1} + \bar{z_2}$
2) $\overline{z_1 \cdot z_2} = \bar{z_1} \cdot \bar{z_2}$
3) $|z| = \sqrt{z \cdot \bar{z}}$
4) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$

*Beweis.*

Zu 1):

$$\begin{aligned}\overline{z_1 + z_2} &= \overline{(x_1 + iy_1) + (x_2 + iy_2)} = \overline{(x_1 + x_2) + i(y_1 + y_2)} = (x_1 + x_2) - i(y_1 + y_2)\\ &= (x_1 - iy_1) + (x_2 - iy_2) = \overline{z_1} + \overline{z_2}\end{aligned}$$

Zu 2):

$$\begin{aligned}\overline{z_1 \cdot z_2} &= \overline{(x_1 + iy_1) \cdot (x_2 + iy_2)} = \overline{(x_1 \cdot x_2 - y_1 \cdot y_2) + \mathrm{i}(x_1 \cdot y_2 + y_1 \cdot x_2)}\\ &= (x_1 \cdot x_2 - y_1 \cdot y_2) - \mathrm{i}(x_1 \cdot y_2 + y_1 \cdot x_2) = x_1(x_2 - iy_2) - iy_1(x_2 - iy_2)\\ &= (x_1 - iy_1) \cdot (x_2 - iy_2) = \bar{z_1} \cdot \bar{z_2}\end{aligned}$$

Zu 3):

$$\sqrt{z \cdot \bar{z}} = \sqrt{(x + iy) \cdot (x - iy)} = \sqrt{x^2 + y^2} = |z|$$

Zu 4):

$$\begin{aligned}|z_1 \cdot z_2| &= |(x_1 \cdot x_2 - y_1 \cdot y_2) + \mathrm{i}(x_1 \cdot y_2 + y_1 \cdot x_2)| = \sqrt{(x_1 \cdot x_2 - y_1 \cdot y_2)^2 + (x_1 \cdot y_2 + y_1 \cdot x_2)^2}\\ &= \sqrt{(x_1 \cdot x_2)^2 + (y_1 \cdot y_2)^2 + (x_1 \cdot y_2)^2 + (y_1 \cdot x_2)^2} = \sqrt{x_1^2 \cdot (x_2^2 + y_2^2) + y_1^2 \cdot (x_2^2 + y_2^2)}\\ &= \sqrt{(x_1^2 + y_1^2) \cdot (x_2^2 + y_2^2)} = \sqrt{x_1^2 + y_1^2} \cdot \sqrt{x_2^2 + y_2^2} = |z_1| \cdot |z_2|\end{aligned}$$

□

**Bemerkung 3.3.6** (Polardarstellung der komplexen Zahlen)**.** Geometrisch kann man die komplexen Zahlen als Vektoren in der Zahlenebene $\mathbb{R}^2$ auffassen, wobei auf der „$x$-Achse" der Realteil von $z$, auf der „$y$-Achse" der Imaginärteil von $z$ angetragen wird (die „y-Achse" trägt demnach die imaginäre Einheit $i$).

Die Länge des Vektors ist gegeben durch den Betrag $r := |z| = \sqrt{x^2 + y^2}$ und der Winkel, unter dem der Vektor angetragen wird, durch $\phi := \arctan y/x$. Daraus ergibt sich die sog. *Polardarstellung* der komplexen Zahlen mit Polarkoordinaten $r$ und $\phi$:

$$z = x + \mathrm{i}y = r\cos\phi + \mathrm{i}r\sin\phi = re^{\mathrm{i}\phi}.$$

Hier haben wir im letzten Schritt die Eulersche Formel $e^{\mathrm{i}\phi} = \cos\phi + \mathrm{i}\sin\phi$ verwendet (Beweis über Reihenentwicklung, siehe *Analysis*).

**Bemerkung 3.3.7** (Quaternionenschiefkörper)**.** Lässt man die Bedingung der Kommutativität der Multiplikation weg, erhält man statt eines Körpers einen sog. *Schiefkörper*. In einem Schiefkörper gibt es also Elemente $a, b$ mit $a \cdot b \neq b \cdot a$. Ein Beispiel dafür bildet der Schiefkörper der Quaternionen $\mathbb{H}$.

# 4 Vektorräume und lineare Abbildungen

## 4.1 Vektorräume

**Definition 4.1.1** (Vektorraum)**.** Sei $K$ ein Körper. Eine Menge $V$ mit zwei Verknüpfungen

i) $+ : V \times V \to V, \quad (v, w) \to v + w$ (Vektoraddition)

ii) $\cdot : K \times V \to V, \quad (\lambda, v) \to \lambda \cdot v$ (skalare Multiplikation)

heißt *Vektorraum* über $K$ oder $K$-Vektorraum genau dann, wenn gilt:

1) $(V, +)$ ist eine kommutative Gruppe

2) $\forall \lambda \in K$ und $v, w \in V$:

$$\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$$

3) $\forall \lambda, \mu \in K, v \in V$:

$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$$

4) $\forall \lambda, \mu \in K, v \in V$:

$$(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$$

5) $\forall v \in V$:

$$1 \cdot v = v$$

Die Elemente von $V$ heißen *Vektoren*, die Elemente von $K$ heißen *Skalare*.

**Bemerkung 4.1.1.** Statt vom $K$-Vektorraum $(V, +, \cdot)$ sprechen wir kurz vom $K$-Vektorraum $V$ und wissen, was gemeint ist. Meist gilt: $K = \mathbb{R}$ (in dem Fall lassen wir den Bezug auf $K$ manchmal auch ganz weg).

**Bemerkung 4.1.2.** Ohne die letzte Forderung in der Vektorraumdefinition (Forderung 5, wonach $1 \cdot v = v$) würde „$\lambda \cdot v = 0$ für alle $\lambda \in K$ und alle $v \in V$" alle anderen Forderungen erfüllen und das ist nicht, was wir wollen.

**Satz 4.1.1.** *Sei $K$ ein Körper, $n \in \mathbb{N}$. Seien $v := (v_1, ..., v_n), w := (w_1, ..., w_n) \in K^n, \lambda \in K$. Dann ist $K^n$ mit*

i) $(v_1, ..., v_n) + (w_1, ..., w_n) := (v_1 + w_1, ..., v_n + w_n)$

ii) $\lambda \cdot (v_1, ..., v_n) := (\lambda \cdot v_1, ..., \lambda \cdot v_n)$

*ein $K$-Vektorraum.*

**Bemerkung 4.1.3.** Insbesondere ist demnach jeder Körper $K$ selbst ein $K$-Vektorraum.

*Beweis.*

Zu 1): $(K^n, +)$ ist kommutative Gruppe, weil $(K, +)$ kommutative Gruppe. Neutrales Element ist der *Nullvektor* $\mathbf{0} := (0, ..., 0)$. Inverses Element zu $(v_1, ..., v_n) \in K^n$ ist: $(-v_1, ..., -v_n)$. Hier wie im Folgenden schreiben wir der Einfachheit halber 0 statt $\mathbf{0}$ (es sei denn, wir wollen explizit betonen, dass es sich um einen Vektor handelt).

Zu 2):

$$\begin{aligned}\lambda(v+w) &= \lambda((v_1,...,v_n)+(w_1,...,w_n)) = \lambda(v_1+w_1,...,v_n+w_n)\\ &= (\lambda(v_1+w_1),...,\lambda(v_n+w_n)) = (\lambda v_1+\lambda w_1,...,\lambda v_n+\lambda w_n)\\ &= (\lambda v_1,...,\lambda v_n)+(\lambda w_1,...,\lambda w_n) = \lambda(v_1,...,v_n)+\lambda(w_1,...,w_n)\end{aligned}$$

Hier haben wir im zweiten und vorletzten Schritt die Definition der Vektoraddition, im dritten und letzten Schritt die Definition der skalaren Multiplikation und im vierten Schritt das Distributivgesetz auf $K$ verwendet.

Zu 3) – 5): Analog. □

**Beispiel 4.1.1.**

- $\mathbb{R}$ ist $\mathbb{R}$-Vektorraum, $\mathbb{C}$ ist $\mathbb{C}$-Vektorraum
- $\mathbb{C} = \{(x,y)|x,y \in \mathbb{R}\} = \mathbb{R}^2$ mit dem darauf definierten „+" und „·" (siehe Def. 3.3.2) ist $\mathbb{R}$-Vektorraum
- $\mathbb{R}^n$ ist $\mathbb{R}$-Vektorraum, $\mathbb{C}^n$ ist $\mathbb{C}$-Vektorraum

**Bemerkung 4.1.4** (Euklidische Geometrie)**.** Im Fall eines *euklidischen Vektorraums* $\mathbb{R}^n$ (z.B. der euklidischen Ebene $\mathbb{R}^2$ oder des euklidischen Raums $\mathbb{R}^3$) kommt zu der Vektorraumstruktur, die wie bislang kennengelernt haben, noch ein Skalarprodukt (i.e. eine Vorschrift zur Multiplikation von Vektoren) hinzu. Erst dadurch können wir Winkel und Längen von Vektoren messen.

**Bemerkung 4.1.5.** Auch Mengen von Funktionen (Abbildungen) oder Mengen von Folgen können Vektorräume sein.

**Satz 4.1.2.** *Sei $M$ eine nicht-leere Menge und $V$ ein $K$-Vektorraum, $\lambda \in K$. Die Menge aller Abbildungen* $\mathrm{Abb}(M,V) := \{f : M \to V\}$ *von $M$ nach $V$ mit*

i) $f+g : M \to V, \quad (f+g)(x) := f(x)+g(x)$

ii) $\lambda f : M \to V, \quad (\lambda f)(x) := \lambda f(x)$

*ist ein $K$-Vektorraum.*

*Beweis.*

Zu 1): $(\mathrm{Abb}(M,V),+)$ ist Gruppe. Dabei ist die *Nullfunktion* $f_0 : M \to V, f_0(x) = 0$ das neutrale Element und das inverse Element zu $f$ ist $-f : M \to V, x \to -f(x)$.

Zu 2): Analog zu oben (siehe Satz 4.1.1).

Zu 3):

$$\begin{aligned}&((\lambda+\mu)f)(x) = (\lambda+\mu)f(x) = \lambda f(x)+\mu f(x) = (\lambda f)(x)+(\mu f)(x) = (\lambda f+\mu f)(x)\\ &\Rightarrow (\lambda+\mu)f = \lambda f+\mu f\end{aligned}$$

Hier haben wir im ersten und dritten Schritt die Definition der skalaren Multiplikation und im zweiten und vierten Schritt die Aussage 3) der Vektorraumdefinition von $V$ benutzt.

Zu 4) und 5): Analog. □

**Beispiel 4.1.2** (Vektorräume)**.**

- Die Menge aller *reellwertigen Funktionen* auf $\mathbb{R}$, i.e. die Menge $\text{Abb}(\mathbb{R}, \mathbb{R})$, ist ein $\mathbb{R}$-Vektorraum. Man nennt ihn auch den *reellen Funktionenraum*.

  (Analog ist der *komplexe Funktionenraum* $\text{Abb}(\mathbb{C}, \mathbb{C})$ ein $\mathbb{C}$-Vektorraum.)

- Sei $K$ ein Körper. Seien $n, m \in \mathbb{N}$. Die Menge aller Abbildungen von $K^n$ nach $K^m$, i.e. die Menge $\text{Abb}(K^n, K^m) = \{f : K^n \to K^m\}$, ist ein $K$-Vektorraum.

  Allgemein schreiben wir: $f : K^n \to K^m, (x_1, ..., x_n) \mapsto (f_1(x_1, ..., x_n), ..., f_m(x_1, ..., x_n))$.

  Sei z.B. $K = \mathbb{R}$. Für $n = 3, m = 2$ liegen z.B. die folgenden Funktionen in $\text{Abb}(\mathbb{R}^3, \mathbb{R}^2)$: $g : \mathbb{R}^3 \to \mathbb{R}^2, (x, y, z) \mapsto (x^2 + y^2, 2z - x)$ oder $h : \mathbb{R}^3 \to \mathbb{R}^2, (x, y, z) \mapsto (x, x + y)$.

- Die Menge aller *reellen Folgen* $\text{Abb}(\mathbb{N}, \mathbb{R}) := \{(a_n)_{n\in\mathbb{N}} = (a_1, a_2, ...) | a_n \in \mathbb{R}\, \forall n \in \mathbb{N}\}$ mit

  $+ : (a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} := (a_n + b_n)_{n\in\mathbb{N}}$ und

  $\cdot : \lambda \cdot (a_n)_{n\in\mathbb{N}} := (\lambda \cdot a_n)_{n\in\mathbb{N}}$

  ist ein $\mathbb{R}$-Vektorraum.

  (Analog bilden die *komplexen Folgen* $\text{Abb}(\mathbb{N}, \mathbb{C})$ einen $\mathbb{C}$-Vektorraum).

- Die Menge aller reellen Polynome vom Grad $n$,

$$P_n = \{P(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n | a_k \in \mathbb{R}\, \forall\, 0 \leq k \leq n\},$$

  bildet einen $\mathbb{R}$-Vektorraum.

**Satz 4.1.3** (Rechenregeln)**.** *Sei $V$ ein $K$-Vektorraum. Dann gilt:*

1) $\forall v \in V : \quad 0 \cdot v = 0$
2) $\forall \lambda \in K : \quad \lambda \cdot 0 = 0$
3) $\forall \lambda \in K, v \in V : \quad \lambda \cdot v = 0 \Rightarrow \lambda = 0 \vee v = 0$
4) $\forall v \in V : \quad (-1) \cdot v = -v$

*Beweis.*

Zu 1): Sei $v \in V$. Wegen Aussage 3) der Vektorraumdefinition und der Tatsache, dass 0 das neutrale Element der Vektorraumaddition ist, gilt:

$$\begin{aligned} 0 \cdot v + 0 \cdot v &= (0 + 0) \cdot v = 0 \cdot v \\ 0 \cdot v + 0 &= 0 \cdot v \end{aligned}$$

Folglich ist $0 \cdot v = 0$.

Zu 2): Sei $\lambda \in K$. Wegen Aussage 2) der Vektorraumdefinition und der Tatsache, dass 0 das neutrale Element der Vektorraumaddition ist, gilt:

$$\begin{aligned} \lambda \cdot 0 + \lambda \cdot 0 &= \lambda \cdot (0+0) = \lambda \cdot 0 \\ \lambda \cdot 0 + 0 &= \lambda \cdot 0 \end{aligned}$$

Folglich ist $\lambda \cdot 0 = 0$.

Zu 3): Sei $\lambda \cdot v = 0$. Wir machen eine Fallunterscheidung. Falls $\lambda = 0$, dann sind wir fertig. Falls $\lambda \neq 0$, dann existiert ein inverses Element $\frac{1}{\lambda} \in K$ (denn $K$ ist Körper) mit $\frac{1}{\lambda} \cdot \lambda = 1$ und damit:

$$v = 1 \cdot v = \left(\frac{1}{\lambda} \cdot \lambda\right) \cdot v = \frac{1}{\lambda} \cdot (\lambda \cdot v) = 0$$

Hier haben wir im ersten und dritten Schritt Aussage 5) und 4) der Vektorraumdefinition verwendet und im letzten Schritt, dass nach Annahme $\lambda \cdot v = 0$.

Zu 4): Analog. □

## 4.2 Lineare Abbildungen

**Definition 4.2.1** (Lineare Abbildung, $\mathrm{Hom}(V,W)$, Isomorphie)**.** Seien $V$ und $W$ $K$-Vektorräume. Eine Abbildung $f : V \to W$ heißt *lineare Abbildung* oder *Vektorraumhomomorphismus* genau dann, wenn gilt:

1) $\forall u, v \in V : \quad f(u+v) = f(u) + f(v)$
2) $\forall \lambda \in K : \quad f(\lambda v) = \lambda f(v)$.

Für die Menge der linearen Abbildungen $f$ von $V$ nach $W$ schreiben wir:

$$\mathrm{Hom}(V,W) := \{f : V \to W | f \text{ linear}\}.$$

Alternativ bezeichnet man die Menge $\mathrm{Hom}(V,W)$ auch als $\mathcal{L}(V,W)$.

Ist $f$ zusätzlich bijektiv, dann heißt $f : V \to W$ *Vektorraumisomorphismus*. In dem Fall sagen wir, $V$ und $W$ sind *isomorph*.

**Bemerkung 4.2.1.** Man kann die beiden Bedingungen für Linearität (1 und 2 in Def. 4.2.1) auch zusammen fassen als *eine* Bedingung, die besagt, $f$ ist linear genau dann, wenn:

$$\forall \lambda \in K, \forall u, v \in V : \quad f(\lambda u + v) = \lambda f(u) + f(v).$$

**Beispiel 4.2.1.**

- $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist *nicht* linear, denn für alle $x, x' \neq 0$ :

$$f(x+x') = (x+x')^2 = x^2 + 2xx' + x'^2 \neq x^2 + x'^2 = f(x) + f(x').$$

- $h : \mathbb{R}^2 \to \mathbb{R}^3, (x, y) \mapsto (x + y, y, x)$ ist linear, denn $\forall u := (x, y), v := (x', y') \in \mathbb{R}^2$:

$$\begin{aligned} h(\lambda u + v) &= h(\lambda(x, y) + (x', y')) = h((\lambda x, \lambda y) + (x', y')) \\ &= h(\lambda x + x', \lambda y + y') \\ &= (\lambda x + x' + \lambda y + y', \lambda y + y', \lambda x + x') \\ &= (\lambda x + \lambda y, \lambda y, \lambda x) + (x' + y', y', x') \\ &= \lambda(x + y, y, x) + (x' + y', y', x') = \lambda h(x, y) + h(x', y') = \lambda h(u) + h(v) \end{aligned}$$

**Bemerkung 4.2.2** (Operatoren)**.** Im Fall von unendlichdimensionalen Vektorräumen spricht man bei linearen Abbildungen auch oft von *Operatoren* (siehe z.B. Quantenmechanik).

Es gibt einen Spezialfall linearer Abbildungen, nämlich den Fall $W = K$ (siehe Def. 4.2.1), also die Menge der linearen Abbildungen von $V$ nach $K$: $\mathrm{Hom}(V, K) = \mathcal{L}(V, K)$. Wie nennt man lineare Abbildungen in dem Fall?

**Definition 4.2.2** (Linearform, Dualraum)**.** Sei $V$ ein $K$-Vektorraum und $f : V \to K$ eine lineare Abbildung. In diesem Fall nennt man $f$ *Linearform*, 1-Form oder lineares Funktional auf $V$. Die Menge aller Linearformen auf $V$ bezeichnet man mit $V^*$. Man nennt $V^*$ den *Dualraum* von $V$.

**Bemerkung 4.2.3.** $\mathrm{Hom}(V, W)$ ist mit der auf der Menge $\mathrm{Abb}(V, W)$ definierten Addition und skalaren Multiplikation ebenfalls ein $K$-Vektorraum. Insbesondere sind also für $f, g \in \mathrm{Hom}(V, W)$ und $\lambda \in K$ auch

$$f + g : V \to W \quad \text{und} \quad \lambda f : V \to W$$

lineare Abbildungen.

**Bemerkung 4.2.4** (Linearität)**.** Seien $V$ und $W$ $K$-Vektorräume. Sei $f : V \to W$ eine lineare Abbildung. Seien $v_1, ..., v_n \in V$ und $\lambda_1, ..., \lambda_n \in K$. Dann gilt:

$$f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i f(v_i)$$

und

$$f(0) = 0.$$

Die erste Aussage folgt unmittelbar aus der Definition einer linearen Abbildung, die zweite Aussage daraus, dass nach Definition:

$$f(0) = f(0 \cdot 0) = 0 \cdot f(0) = 0.$$

**Satz 4.2.1.** *Seien $U, V$ und $W$ $K$-Vektorräume. Seien $f : V \to W$ und $g : U \to V$ lineare Abbildungen. Dann ist die Funktionsverkettung $f \circ g : U \to W$ auch eine lineare Abbildung.*

*Beweis.* Übung. □

**Bemerkung 4.2.5.** Wir werden bald beweisen, dass sich jede lineare Abbildung $f : K^n \to K^m$ durch eine $m \times n$ Matrix $A \in K^{m \times n}$ darstellen läßt, d.h. $A = \left(a_{ij}\right)_{\substack{i=1,...,m,\\ j=1,...,n}}$ mit $a_{ij} \in K$ bzw. ausgeschrieben

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

Die Matrix $A$ wirkt auf Vektoren $x = (x_1, ..., x_n) \in K^n$ und bildet sie ab auf Vektoren $y = (y_1, ..., y_m) \in K^m$ gemäß der Vorschrift $y = Ax$. Die Analyse der Matrizen werden wir brauchen, um lineare Gleichungssysteme zu lösen.

**Definition 4.2.3** (Lineares Gleichungssystem)**.** Sei $x = (x_1, ..., x_n) \in K^n, y = (y_1, ..., y_m) \in K^m$. Ein *lineares Gleichungssystem* ist ein Gleichungssystem der folgenden Form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= y_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= y_m \end{aligned}$$

Man nennt das Gleichungssystem *homogen*, falls $y = 0$, und *inhomogen*, falls $y \neq 0$.

**Bemerkung 4.2.6** ($n$-Tupel, Zeilen- und Spaltenvektoren)**.** Bislang haben wir Vektoren $v \in K^n$ nur als $n$-Tupel kennengelernt:

$$v = (v_1, v_2, ..., v_n) \in K^n.$$

Später werden wir Vektoren auch als spezielle Form von Matrizen auffassen und unterscheiden dann zwischen *Zeilenvektoren* (Linearformen) der Form

$$v = (v_1 \quad v_2 \quad ... \quad v_n) \in K^{1 \times n}$$

und *Spaltenvektoren* der Form

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in K^{n \times 1}.$$

Tatsächlich sind die Vektorräume $K^n, K^{1 \times n}$ und $K^{n \times 1}$ isomorph (denn offensichtlich gibt es eine bijektive, lineare Abbildung zwischen ihnen), so dass es oft reicht, eine einzige Schreibweise für Vektoren einzuführen (z.B. Spaltenvektoren). Das bedeutet nicht, dass Zeilen- und Spaltenvektoren dasselbe sind. Zeilenvektoren $v \in K^{1 \times n}$ sind, wie wir noch sehen werden, Linearformen, d.h. sie bilden Vektoren $v \in K^{n \times 1}$ auf Elemente von $K$ (Skalare bzw. Zahlen) ab.

## 4.3 Untervektorräume

**Definition 4.3.1** (Untervektorraum)**.** Sei $V$ ein $K$-Vektorraum. Eine nicht-leere Menge $U \subseteq V$ heißt *Untervektorraum* von $V$ genau dann, wenn gilt:

1) $\forall v, w \in U: \quad v + w \in U$
2) $\forall \lambda \in K, v \in U: \quad \lambda \cdot v \in U$

**Bemerkung 4.3.1.** Jeder Untervektorraum $U$ ist demnach selbst wieder ein Vektorraum. Insbesondere ist der Nullvektor in jedem Untervektorraum enthalten, d.h. $0 \in U$, und zu jedem Element $v \in U$ ist auch das inverse Element $-v \in U$.

**Beispiel 4.3.1** (Untervektorräume)**.**

- Sei $V$ ein $K$-Vektorraum. $\{0\}$ und $V$ sind Untervektorräume von $V$.
- Sei $V$ ein $K$-Vektorraum und $v \in V$. Dann ist $U = \{\lambda v |\ \lambda \in K\}$ ein Untervektorraum von $V$. Dies ist offensichtlich der Fall, denn für alle $u, w \in U$: $\exists \alpha, \beta \in K$ mit $u = \alpha v, w = \beta v$ und damit $u + w = \alpha v + \beta v = (\alpha + \beta) \cdot v \in U$, weil $\alpha + \beta \in K$, und $\forall \lambda \in K, v \in U : \lambda \cdot v \in U$.

  Wir können dieses Beispiel geometrisch deuten, indem wir z.B. $V = \mathbb{R}^2$ und

  $$v = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

  betrachten. Wir sehen, die Gerade, die durch den Ursprung geht und den Vektor $v$ enthält, ist der spezifizierte Untervektorraum, d.h.

  $$U = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} \middle| \lambda \in \mathbb{R} \right\}$$

  Eine alternative Beschreibung dieses Untervektorraums ist:

  $$U = \left\{ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2 \middle| \ -3v_1 + v_2 = 0 \right\}$$

- Vorigen Beispiel lässt sich verallgemeinern. Sei $V$ ein $K$-Vektorraum und $v_1, ..., v_m \in V$. Dann ist $U' = \{\lambda_1 v_1 + ... + \lambda_m v_m | \lambda_1, ..., \lambda_m \in K\}$ ein Untervektorraum von $V$.

  Wir können dieses Beispiel wieder geometrisch deuten, indem wir z.B. $V = \mathbb{R}^3$ und

  $$v_1 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

  betrachten. Wir sehen, die Ebene, die durch den Ursprung geht und von den Vektoren $v_1$

und $v_2$ aufgespannt wird, ist der spezifizierte Untervektorraum, d.h.

$$U' = \left\{ \lambda_1 \cdot \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \middle| \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

**Satz 4.3.1** (Schnitt ist Unterraum). *Sei $V$ ein $K$-Vektorraum und $(U_i)_{i \in I}$ eine Familie von Untervektorräumen von $V$. Dann ist der Schnitt $\bigcap_{i \in I} U_i$ ein Untervektorraum von $V$*

*Beweis.* Betrachte $\bigcap_{i \in I} U_i$. Seien $u_1, u_2 \in \bigcap_{i \in I} U_i$. Dann gilt $\forall i \in I$: $u_1, u_2 \in U_i$. Jedes $U_i$ ist aber Untervektorraum, folglich $u_1 + u_2 \in U_i$. Dies gilt $\forall i \in I$. Damit

$$u_1 + u_2 \in \bigcap_{i \in I} U_i.$$

Sei weiter $u \in \bigcap_{i \in I} U_i, \lambda \in K$. Dann gilt $\forall i \in I$: $u \in U_i$. Nun ist aber jedes $U_i$ Untervektorraum, folglich auch $\lambda u \in U_i$. Dies gilt $\forall i \in I$. Damit

$$\lambda u \in \bigcap_{i \in I} U_i.$$

$\Rightarrow \bigcap_{i \in I} U_i$ Untervektorraum. □

**Satz 4.3.2** (Vereinigung kein Unterraum). *Dagegen ist die Vereinigung $\bigcup_{i \in I} U_i$ im Allgemeinen* kein *Untervektorraum von $V$.*

*Beweis.* Wir geben ein Beispiel. Betrachte $U_1 = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \middle| \lambda \in \mathbb{R} \right\}$ und $U_2 = \left\{ \mu \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \middle| \mu \in \mathbb{R} \right\}$. Offensichtlich sind $U_1$ und $U_2$ Untervektorräume von $\mathbb{R}^2$.

Betrachte nun z. B. $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \in U_1 \cup U_2$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix} \in U_1 \cup U_2$. Obwohl beide Vektoren in der Vereinigung sind, gilt:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \notin U_1 \cup U_2.$$

$\Rightarrow U_1 \cup U_2$ ist kein Untervektorraum von $\mathbb{R}^2$. □

**Satz 4.3.3.** *Sei $V$ ein $K$-Vektorraum und seien $U_1, U_2, ..., U_n$ mit $n \in \mathbb{N}$ Untervektorräume von $V$. Dann ist die Summe*

$$U_1 + U_2 + ... + U_n := \{u_1 + u_2 + ... + u_n | u_1 \in U_1, u_2 \in U_2, ..., u_n \in U_n\}$$

*ein Untervektorraum von $V$.*

*Beweis.* Analog zu Beweis vom Schnitt von Untervektorräumen (siehe Satz 4.3.1). □

**Bemerkung 4.3.2.** Sind die $U_i, U_j$ paarweise disjunkt, d.h. ist $U_i \cap U_j = \emptyset$ für alle $i, j = 1, ..., n$, dann nennt man $U_1 + ... + U_n$ die *direkte* Summe.

**Definition 4.3.2** (Kern, Bild)**.** Seien $V$ und $W$ $K$-Vektorräume. Sei $f : V \to W$ eine lineare Abbildung. Dann bezeichnen wir als *Kern* von $f$ die Menge

$$\text{Kern} f := \{v \in V | f(v) = 0\}.$$

Als *Bild* von $f$ bezeichnen wir die Menge

$$\text{Bild} f := f(V).$$

**Satz 4.3.4** (Kern und Bild sind Untervektorräume)**.** *Seien $V$ und $W$ $K$-Vektorräume. Sei $f : V \to W$ eine lineare Abbildung. Dann gilt:*

1) *Kern $f$ ist Untervektorraum von $V$,*
2) *Bild $f$ ist Untervektorraum von $W$.*

*Beweis.*

Zu 1): Kern $f \neq \emptyset$, denn $f(0) = 0$ und $0 \in V$.

Angenommen $v_1, v_2 \in$ Kern $f$. Dann ist also $f(v_1) = 0$ und $f(v_2) = 0$. Wegen Linearität von $f$ gilt: $f(v_1 + v_2) = f(v_1) + f(v_2) = 0 + 0 = 0$, also auch $v_1 + v_2 \in$ Kern $f$.

Angenommen $v \in$ Kern $f$, $\lambda \in K$. Dann gilt, wegen Linearität von $f$ und weil $v \in$ Kern $f$: $f(\lambda v) = \lambda f(v) = 0$ und damit $\lambda v \in$ Kern $f$.

Zu 2): Bild $f \neq \emptyset$, denn $V \neq \emptyset$.

Angenommen $w_1, w_2 \in$ Bild $f$. Dann existieren also $v_1, v_2 \in V$ mit $w_1 = f(v_1)$ und $w_2 = f(v_2)$. Wegen Linearität von $f$ gilt: $w_1 + w_2 = f(v_1) + f(v_2) = f(v_1 + v_2)$, also auch $w_1 + w_2 \in$ Bild $f$.

Angenommen $w \in$ Bild $f$, $\lambda \in K$. Dann gilt, wegen Linearität von $f$ und weil $w \in$ Kern $f$ (d.h. $\exists v \in V$: $w = f(v)$): $\lambda w = \lambda f(v) = f(\lambda v)$ und damit $\lambda w \in$ Bild $f$.

□

**Lemma 4.3.1.** *Seien $V, W$ $K$-Vektorräume und sei $f : V \to W$ eine lineare Abbildung. Dann gilt:*

$$f \text{ injektiv} \quad \Leftrightarrow \quad \text{Kern} f = \{0\}$$

*Beweis.* Übung. □

## 4.4 Matrizen

**Definition 4.4.1** ($m \times n$ Matrix über $K$, $K^{m \times n}$)**.** Seien $m, n \in \mathbb{N}$ und $a_{ij} \in K$ für $i = 1, ..., m$ und $j = 1, .., n$. Das rechteckige Schema mit $m$ Zeilen und $n$ Spalten

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

heißt $m \times n$ *Matrix* (über $K$). Wir schreiben kurz: $A = \big(a_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}}$.
Die Menge aller $m \times n$ Matrizen über $K$ wird mit $K^{m\times n}$ bezeichnet.

$$a_j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = (a_{ij})_{i=1,\dots,m} \in K^{m\times 1} \text{ heißt } j\text{-te Spalte von } A$$

$$a^i := \begin{pmatrix} a_{i1} & \dots & a_{in} \end{pmatrix} = (a_{ij})_{j=1,\dots,n} \in K^{1\times n} \text{ heißt } i\text{-te Zeile von } A$$

**Definition 4.4.2** (Matrizenaddition, skalare Multiplikation)**.** Seien $m, n \in \mathbb{N}$. Sei $K$ ein Körper. Seien $A = \big(a_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}}, B = \big(b_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}} \in K^{m\times n}$, $\lambda \in K$. Wir definieren:

1) $A + B := \big(a_{ij} + b_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}}$ (Matrizenaddition)

2) $\lambda A := \big(\lambda a_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}}$ (skalare Multiplikation)

**Satz 4.4.1.** *Die Menge $K^{m\times n}$ mit der angegeben Matrizenaddition und skalaren Multiplikation (siehe Def. 4.4.2) ist ein $K$-Vektorraum.*

*Beweis.* Jede Matrix $A \in K^{m\times n}$ kann als eine Funktion $\{1, ..., m\} \times \{1, .., n\} \to K$, $(i, j) \to a_{ij}$ aufgefasst werden kann. Wir können also $K^{m\times n}$ mit der Menge $\mathrm{Abb}(\{1, ..., m\} \times \{1, .., n\}, K)$ identifizieren. Für Letztere haben wir bereits bewiesen, dass sie einen $K$-Vektorraum bilden (siehe Satz 4.1.2). □

**Definition 4.4.3** (Matrizenmultiplikation)**.** Seien $m, n, p \in \mathbb{N}$. Sei $K$ ein Körper. Seien $A = \big(a_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}} \in K^{m\times n}, B = \big(b_{jk}\big)_{\substack{j=1,\dots,n,\\ k=1,\dots,p}} \in K^{n\times p}$. Wir definieren:

$$A \cdot B := \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right)_{\substack{i=1,\dots,m,\\ k=1,\dots,p}} \in K^{m\times p} \quad \text{(Matrizenmultiplikation)}$$

**Bemerkung 4.4.1.** Die Matrizenmultiplikation ist somit nur definiert, wenn die Spaltenzahl von $A$ gleich der Zeilenzahl von $B$ ist.
Ausgeschrieben erhalten wir für $A = \big(a_{ij}\big)_{\substack{i=1,\dots,m,\\ j=1,\dots,n}} \in K^{m\times n}, B = \big(b_{jk}\big)_{\substack{j=1,\dots,n,\\ k=1,\dots,p}} \in K^{n\times p}$:

$$A \cdot B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ & & & \\ \vdots & & \ddots & \vdots \\ & & & \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{pmatrix}$$

mit $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$, also z. B. $c_{23} = \sum_{j=1}^{n} a_{2j} b_{j3} = a_{21} b_{13} + a_{22} b_{23} + ... + a_{2n} b_{n3}$.

Für den Spezialfall $a = \begin{pmatrix} a_1 & ... & a_n \end{pmatrix} \in K^{1\times n}, b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in K^{n\times 1}$ ergibt sich demnach:

$$\begin{pmatrix} a_1 & ... & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + ... + a_n b_n \in K^{1\times 1}$$

Man nennt dies auch das *Skalarprodukt* der Vektoren $a$ und $b$ und schreibt kurz $< a, b >$ (diese Schreibweise impliziert, dass $a$ ein Zeilen- und $b$ ein Spaltenvektor ist).

Steht umgekehrt links ein Spaltenvektor $x$ und rechts ein Zeilenvektor $y$, so nennt man dies das *dyadische Produkt* der Vektoren $x$ und $y$. Man erhält daraus eine $m \times n$ Matrix $A = \left(a_{ij}\right)_{\substack{i=1,...,m,\\ j=1,...,n}}$ mit Einträgen $a_{ij} = x_i \cdot y_j$ und schreibt dafür:

$$x \otimes y := \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \cdot \begin{pmatrix} y_1 & ... & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \dots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_m y_1 & \dots & x_m y_n \end{pmatrix} \in K^{m\times n}$$

Für $A = \left(a_{ij}\right)_{\substack{i=1,...,m,\\ j=1,...,n}} \in K^{m\times n}$ und $x \in K^{n\times 1}$ ergibt sich schließlich:

$$A \cdot x = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{n} a_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{pmatrix}$$

Der hier entstandene Vektor fasst die eine Seite eines linearen Gleichungssystems zusammen (mit $A$ gegeben, $x$ gesucht). Setzt man ihn nun gleich einem gegebenen Vektor $y \in K^{n\times 1}$, also $A \cdot x = y$, so hat man das vollständige lineare Gleichungssystem dastehen.

**Beispiel 4.4.1** (Matrizenmultiplikation).

- Gegeben $A = \begin{pmatrix} 2 & 1 & 2 & 4 \\ 1 & 3 & 1 & 2 \\ -2 & 1 & 4 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 4}$ und $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{4\times 2}$, dann ist:

$$A \cdot B = \begin{pmatrix} 2 & 1 & 2 & 4 \\ 1 & 3 & 1 & 2 \\ -2 & 1 & 4 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 13 & 7 \\ 9 & 11 \\ 4 & 14 \end{pmatrix} \quad \in \mathbb{R}^{3\times 2}$$

- Gegeben $(1 \quad 2 \quad -1) \in \mathbb{R}^{1\times n}$, $\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^{n\times 1}$:

$$(1 \quad 2 \quad -1) \cdot \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = 4 \quad \in \mathbb{R}^{1\times 1}$$

**Definition 4.4.4** (Kronecker-Delta-Symbol). Das *Kronecker-Delta-Symbol* ist

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

**Definition 4.4.5** (Nullmatrix, Einheitsmatrix, Einheitsvektor). Sei $m, n \in \mathbb{N}$, $K$ ein Körper.

- Wir definieren die *Nullmatrix* $0 \in K^{m\times n}$ (oder explizit $0_{mn} \in K^{m\times n}$) als

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

- Wir definieren die *Einheitsmatrixatrix* $E_n \in K^{n\times n}$ als die quadratische Matrix, die auf der Diagonalen nur 1er und sonst nur 0er hat,

$$E_n := \begin{pmatrix} 1 & 0 & \dots & & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & \vdots \\ & & & 1 & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix} \quad \in K^{n\times n}$$

- Wir definieren den $j$-ten *Einheitsvektor* $e_j$ von $K^n$ als $j$-te Spalte von $E_n$:

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \in K^{n\times 1}$$

**Bemerkung 4.4.2.** Es ist $e_j = (\delta_{ij})_{i=1,\dots,n}$.

**Satz 4.4.2** (Darstellende Matrix einer linearen Abbildung)**.** *Seien $m, n \in \mathbb{N}$. Sei $K$ ein Körper und $x \in K^n$. Dann gilt:*

1) *Zu jeder linearen Abbildung $f : K^n \to K^m$ existiert genau eine Matrix $A \in K^{m\times n}$, so dass*

$$f(x) = Ax$$

*Insbesondere gilt $a_{ij} = f_i(e_j)$ mit $i = 1, ..., m$, $j = 1, ..., n$ und $a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = f(e_j)$.*

*$A$ heißt* darstellende *Matrix von $f$.*

2) *Für jedes $A \in K^{m\times n}$ ist die Abbildung $f : K^n \to K^m$, $f(x) = Ax$ linear.*

*Beweis.*

Zu 1): Wir müssen *Existenz* und *Eindeutigkeit* der darstellenden Matrix zeigen.
Zunächst zur Existenz: Wir definieren $a_{ij} := f_i(e_j)$ und damit

$$a_j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} f_1(e_j) \\ \vdots \\ f_m(e_j) \end{pmatrix} = f(e_j).$$

Nun betrachte $x \in K^n$. Es ist

$$x = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + ... + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + ... + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{j=1}^{n} x_j e_j$$

Wir setzen dieses Ergebnis in $f(\cdot)$ ein und benutzen die Linearität von $f$. Dann ist:

$$f(x) = f\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j f(e_j) = \sum_{j=1}^{n} x_j a_j$$

Nun ersetzen wir $f(x)$ auf der linken Seite durch $f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$. Folglich gilt:

$$\begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \sum_{j=1}^{n} x_j a_j = \sum_{j=1}^{n} x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \sum_{j=1}^{n} \begin{pmatrix} a_{1j}x_j \\ \vdots \\ a_{mj}x_j \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix}$$

Also ist $f_i(x) = \sum_{j=1}^{n} a_{ij}x_j = (Ax)_i$ und damit $f(x) = Ax$.

Nun zur Eindeutigkeit: Angenommen es gibt $A, A' \in K^{m\times n}$ mit $f(x) = Ax$ und $f(x) = A'x$ ($x \in K^n$). Dann ist $Ax = A'x$. Insbesondere ist also $Ae_j = A'e_j \, \forall j = 1, ..., n \Rightarrow a_j = a'_j \, \forall j = 1, ..., n \Rightarrow a_{ij} = a'_{ij} \, \forall i = 1, ..., m, j = 1, ..., n \Rightarrow A = A'$

Zu 2): Betrachte $A \in K^{m\times n}$. Dann ist

$$Ax = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix}$$

Nun definieren wir $f_i : K^n \to K$, $f_i(x) := \sum_{j=1}^{n} a_{ij}x_j$ und damit

$$f : K^n \to K^m, \, f(x) := \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

Wir prüfen, ob diese Abbildung linear ist. Betrachte dazu $x, y \in K^n, \lambda \in K$. Dann ist:

$$f_i(x+y) = \sum_{j=1}^{n} a_{ij}(x_j + y_j) = \sum_{j=1}^{n} (a_{ij}x_j + a_{ij}y_j) = \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ij}y_j = f_i(x) + f_i(y)$$

$$\lambda f_i(x) = \lambda \sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} \lambda a_{ij}x_j = \sum_{j=1}^{n} a_{ij}\lambda x_j = f_i(\lambda x)$$

D.h. $f_i : K^n \to K$ ist eine lineare Abbildung. Daraus folgt, dass auch $f : K^n \to K^m$ eine lineare Abbildung ist, denn:

$$f(x+y) = \begin{pmatrix} f_1(x+y) \\ \vdots \\ f_m(x+y) \end{pmatrix} = \begin{pmatrix} f_1(x) + f_1(y) \\ \vdots \\ f_m(x) + f_m(y) \end{pmatrix} = f(x) + f(y)$$

$$\lambda f(x) = \lambda \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} f_1(\lambda x) \\ \vdots \\ f_m(\lambda x) \end{pmatrix} = f(\lambda x)$$

□

**Beispiel 4.4.2.** Betrachte die lineare Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^3$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ y \\ x+z \end{pmatrix}$.

Die darstellende Matrix $A \in \mathbb{R}^{3\times 3}$ der Abbildung $f$ ist:

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

Betrachte als nächstes die lineare Abbildung $g : \mathbb{R}^3 \to \mathbb{R}^2$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 3x - \frac{1}{2}z \\ 0 \end{pmatrix}$.

Die darstellende Matrix $B \in \mathbb{R}^{2\times 3}$ der Abbildung $g$ ist:

$$B = \begin{pmatrix} 3 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

**Satz 4.4.3** (Darstellende Matrix der Verkettung linearer Abbildungen)**.** *Seien $n, m, p \in \mathbb{N}$. Sei $K$ ein Körper. Seien $f : K^n \to K^m$ und $g : K^p \to K^n$ lineare Abbildungen mit darstellenden Matrizen $A \in K^{m\times n}$ und $B \in K^{n\times p}$. Dann ist $C := A \cdot B \in K^{m\times p}$ die darstellende Matrix der linearen Abbildung $f \circ g : K^p \to K^m$.*

*Beweis.*
Erinnerung: Wenn $f, g$ linear ist, dann ist auch die Verkettung $f \circ g$ linear (siehe Satz 4.2.1).

Nun ist $\forall j = 1, ...p : (f \circ g)(e_j) = f(g(e_j)) = f(B \cdot e_j) = A \cdot (B \cdot e_j)$.

Sei nun $C$ die darstellende Matrix von $f \circ g$. Dann ist

$$c_{ij} = (f \circ g)_i(e_j) = (A \cdot (B \cdot e_j))_i = \sum_{k=1}^{n} a_{ik}(Be_j)_k = \sum_{k=1}^{n} a_{ik}b_{kj} = (A \cdot B)_{ij}$$

für alle $i = 1, .., m$, $j = 1, ..., p$. Hier haben wir im dritten Schritt benutzt, dass $(Ax)_i = \sum_{k=1}^{n} a_{ik}x_k$. Somit ist $C = A \cdot B$ die darstellende Matrix. □

**Beispiel 4.4.3** (Additionstheoreme der Trigonometrie)**.** Man betrachte die $2 \times 2$ *Drehmatrix*

$$R_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

mit $\phi \in \mathbb{R}$ und einen Vektor $r = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

$R_\phi$ beschreibt Drehungen des Vektors $r$ in mathematisch positiver Richtung (d.h. entgegen dem Uhrzeigersinn) um den Winkel $\phi$ um den Ursprung.

Betrachte z.B. $r = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\phi = \pi/2$. Dann ist

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad R_\phi r = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Für Drehungen um den Ursprung gilt $\forall \phi, \psi \in \mathbb{R}$, dass $R_\phi \circ R_\psi = R_{\phi+\psi}$, wobei hier $R_\phi$ die lineare Abbildung bezeichnet (wir hätten auch eine andere Bezeichnung wählen können, z.B. $f_{R_\phi}$ statt $R_\phi$, um den Unterschied zur darstellenden Matrix deutlicher zu machen). D.h. dreht man nacheinander um die Winkel $\phi$ und $\psi$, so ist es dasselbe, wie wenn man einmal um den Winkel $\phi + \psi$ dreht (oder um $\psi + \phi$, das Ganze ist symmetrisch in $\psi, \phi$). Benutzen wir den vorhergehenden Satz, so finden wir, dass für die Drehmatrizen gilt:

$$R_{\phi+\psi} = R_\phi \cdot R_\psi,$$

also

$$\begin{aligned} \begin{pmatrix} \cos(\phi+\psi) & -\sin(\phi+\psi) \\ \sin(\phi+\psi) & \cos(\phi+\psi) \end{pmatrix} &= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \cdot \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \\ &= \begin{pmatrix} \cos\phi\cos\psi - \sin\phi\sin\psi & -\cos\phi\sin\psi - \sin\phi\cos\psi \\ \sin\phi\cos\psi + \cos\phi\sin\psi & -\sin\phi\sin\psi + \cos\phi\cos\psi \end{pmatrix} \end{aligned}$$

Daraus ergeben sich durch Vergleich der Matrixeinträge die bekannten Additionstheoreme. D.h. $\forall \phi, \psi \in \mathbb{R}$ gilt:

$$\cos(\phi+\psi) = \cos\phi\cos\psi - \sin\phi\sin\psi$$

$$\sin(\phi+\psi) = \sin\phi\cos\psi + \cos\phi\sin\psi$$

**Satz 4.4.4** (Rechenregeln für Matrizen)**.** *Seien $m, n, p, q \in \mathbb{N}$. Sei $K$ ein Körper. Dann gilt:*

1) *falls $A \in K^{m\times n}, B \in K^{n\times p}, C \in K^{p\times q}$:*

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

2) *falls $A \in K^{m\times n}$:*

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

3) *falls $A \in K^{m\times n}$, $B, C \in K^{n\times p}$:*

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

4) *falls* $A, B \in K^{m\times n}$, $C \in K^{n\times p}$*:*

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

5) *falls* $\lambda \in K, A \in K^{m\times n}, B \in K^{n\times p}$*:*

$$\lambda \cdot (A \cdot B) = (\lambda \cdot A) \cdot B = A \cdot (\lambda \cdot B)$$

*Beweis.*

Zu 1): $A$, $B$ und $C$ sind darstellende Matrizen der Abbildungen $f : K^n \to K^m$, $g : K^p \to K^n$ und $h : K^q \to K^p$. Nun folgt aus $(f \circ g) \circ h = f \circ (g \circ h)$ (Assoziativität der Verkettung, siehe Satz 2.5.1) und der Tatsache, dass die Verkettung von Funktionen $f, g$ der Multiplikation ihrer darstellenden Matrizen $A, B$ entspricht (siehe Satz 4.4.3), dass $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Zu 2): $(A \cdot E_n)_{ik} = \sum_{j=1}^{n} a_{ij}\ \delta_{jk} = a_{ik} = (A)_{ik}\ \forall i = 1, ..., m,\ k = 1, .., n \Rightarrow (A \cdot E_n) = A$, und analog für $E_m \cdot A$.

Zu 3):

$$\begin{aligned}(A \cdot (B + C))_{ik} &= \sum_{j=1}^{n} (A)_{ij} \cdot (B + C)_{jk} = \sum_{j=1}^{n} a_{ij} \cdot (b_{jk} + c_{jk}) = \sum_{j=1}^{n} (a_{ij} \cdot b_{jk} + a_{ij} \cdot c_{jk}) \\ &= \sum_{j=1}^{n} a_{ij} \cdot b_{jk} + \sum_{j=1}^{n} a_{ij} \cdot c_{jk} = (A \cdot B)_{ik} + (A \cdot C)_{ik}\end{aligned}$$

$\forall i = 1, ..., m,\ k = 1, .., p \Rightarrow A \cdot (B + C) = A \cdot B + A \cdot C$.

Zu 4): Analog.

Zu 5): Es ist $(\lambda \cdot (A \cdot B))_{ik} = \lambda \cdot (A \cdot B)_{ik} = \lambda \cdot \sum_{j=1}^{n} a_{ij} b_{jk} = \sum_{j=1}^{n} \lambda a_{ij} b_{jk} = ((\lambda \cdot A) \cdot B)_{ik}$
$\forall i = 1, ..., m,\ k = 1, .., p \Rightarrow \lambda \cdot (A \cdot B) = (\lambda \cdot A) \cdot B$ und analog für $\lambda \cdot (A \cdot B) = A \cdot (\lambda \cdot B)$. □

**Bemerkung 4.4.3** (Nicht kommutativ, nicht nullteilerfrei)**.** Die quadratischen Matrizen $K^{n\times n}$ bilden mit der Matrizenaddition und -multiplikation einen Ring mit $E_n$ als Einselement.
Dieser Ring ist jedoch für $n \geq 2$ nicht kommutativ, denn es gibt $A, B \in K^{n\times n}$, für die gilt: $A \cdot B \neq B \cdot A$. Betrachte z.B.:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Der Ring ist für $n \geq 2$ auch nicht nullteierfrei, denn es gibt $A, B \in K^{n\times n}$, für die aus $A \cdot B = 0$ *nicht* folgt, dass $A = 0$ oder $B = 0$. Betrachte z.B.:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

**Definition 4.4.6** (Transponierte Matrix)**.** Seien $m, n \in \mathbb{N}$. Sei $K$ ein Körper und die Matrix $A = (a_{ij})_{\substack{i=1,\dots,m,\\ j=1,\dots,n}} \in K^{m\times n}$. Wir nennen

$$A^T := (a_{ji})_{\substack{j=1,\dots,n,\\ i=1,\dots,m}} \in K^{n\times m}$$

die zu $A$ *transponierte* Matrix.

**Bemerkung 4.4.4.** Die Zeilen von $A^T$ sind also die Spalten von $A$ und umgekehrt.

**Bemerkung 4.4.5.** Statt $A = (a_{ij})_{\substack{i=1,\dots,m,\\ j=1,\dots,n}}$ und $A^T = (a_{ji})_{\substack{j=1,\dots,n,\\ i=1,\dots,m}}$ schreiben wir oft kurz einfach nur: $A = (a_{ij})$, und $A^T = (a_{ji})$. Die Klammer zeigt dabei an, dass wir jeweils die ganze Matrix meinen. Achtung: Dies ist nicht zu verwechseln mit der Schreibweise für *einzelne Matrixeinträge*. So bezeichnet z.B. $(A)_{ik} = a_{ik}$ den Eintrag der Matrix $A$ in der $i$-ten Zeile und $k$-ten Spalte (dies ist einfach eine Zahl, nämlich die Zahl $a_{ik}$).

**Beispiel 4.4.4.** Betrachte z.B. $A = \begin{pmatrix} 1 & 5 & 1 \\ 2 & 3 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 3}$. Dann ist

$$A^T = \begin{pmatrix} 1 & 2 \\ 5 & 3 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

Für $x = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbb{R}^{2\times 1}$ ist $x^T = (1 \quad 2) \in \mathbb{R}^{1\times 2}$.

**Satz 4.4.5** (Rechenregeln für transponierte Matrizen)**.** *Seien $m, n \in \mathbb{N}$. Sei $K$ ein Körper. Dann gilt:*

1) $A \in K^{m\times n}: \quad (A^T)^T = A$
2) $A, B \in K^{m\times n}: \quad (A+B)^T = A^T + B^T$
3) $\lambda \in K, A \in K^{m\times n}: (\lambda A)^T = \lambda A^T$
4) $A \in K^{m\times n}, B \in K^{n\times p}: (A \cdot B)^T = B^T \cdot A^T$

*Beweis.* Der Beweis von 1) – 3) ist trivial (man überlege sich, was passiert, wenn Zeilen und Spalten vertauschen).

Zu 4: Seien $A = (a_{ij}) \in K^{m\times n}, B = (b_{jk}) \in K^{n\times p}$. Dann ist $A \cdot B = \left(\sum_{j=1}^n a_{ij}b_{jk}\right) \in K^{m\times p}$ und $(A \cdot B)^T \in K^{p\times m}$ mit

$$(A \cdot B)^T_{ki} = (A \cdot B)_{ik} = \sum_{j=1}^n (A)_{ij}(B)_{jk} = \sum_{j=1}^n (A^T)_{ji}(B^T)_{kj} = \sum_{j=1}^n (B^T)_{kj}(A^T)_{ji} = \left(B^T \cdot A^T\right)_{ki}$$

□

**Definition 4.4.7** (Inverse Matrix)**.** Sei $n \in \mathbb{N}$. Sei $K$ ein Körper und $A \in K^{n\times n}$. $A$ heißt invertierbar, wenn $A^{-1} \in K^{n\times n}$ existiert mit

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

$A^{-1}$ ist eindeutig definiert und heißt inverse Matrix zu $A$.

**Bemerkung 4.4.6.** Die Menge der invertierbaren Matrizen $A \in K^{n\times n}$ bildet mit der Matrixmultiplikation eine Gruppe. Wie bezeichnen sie als $GL(n, K)$ („general linear group").

**Beispiel 4.4.5.** Betrachte $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Die inverse Matrix $A^{-1} \in \mathbb{R}^{2\times 2}$ muss die Gleichung $A \cdot A^{-1} = A^{-1} \cdot A = E_2$ erfüllen. Nun ist

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

Man kann nachrechnen, dass damit auch die Gleichung $A^{-1} \cdot A = E_2$ erfüllt ist.

**Satz 4.4.6** (Rechenregeln für inverse Matrizen)**.** *Sei $n \in \mathbb{N}$. Sei $K$ ein Körper. Dann gilt:*

1) $A \in K^{n\times n}$ invertierbar $\Rightarrow$ $A^{-1}$ invertierbar und $(A^{-1})^{-1} = A$
2) $A, B \in K^{n\times n}$ invertierbar $\Rightarrow$ $A \cdot B$ invertierbar und $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$
3) $A \in K^{n\times n}$ invertierbar $\Rightarrow$ $A^T$ invertierbar und $(A^{-1})^T = (A^T)^{-1}$

*Beweis.*

Zu 1) $A$ ist invertierbar, d.h. es existiert $A^{-1}$ mit $A \cdot A^{-1} = A^{-1} \cdot A = E_n$. Aus dieser Gleichung folgt: $A^{-1}$ ist invertierbar und die inverse Matrix $(A^{-1})^{-1}$ ist gerade $A$: $(A^{-1})^{-1} = A$.

Zu 2) und 3): Übung. □

## 4.5 Erzeugendensystem, lineare Unabhängigkeit, Basis

Im Folgenden werden wir die Begriffe der „Basis", der „Linearen Unabhängigkeit" und des „Erzeugendensystems" kennenlernen. Diese drei Begriffe kann man wie folgt motivieren.

Sei $V$ ein (endlich erzeugter) $K$-Vektorraum und seien (beliebige, aber fixe) Vektoren $b_1, ..., b_m \in V$ gegeben. Dann kann man folgende Fragen stellen:

- Gibt es für jedes $v \in V$ *genau ein* $m$-Tupel $(\lambda_1, ..., \lambda_m) \in K^m$, so dass $v = \lambda_1 b_1 + \lambda_2 b_2 + ... + \lambda_m b_m$? In dem Fall bilden die Vektoren $b_1, ..., b_m$ eine *Basis* von $V$.
- Gibt es für jedes $v \in V$ *höchstens ein* $m$-Tupel $(\lambda_1, ..., \lambda_m) \in K^m$, so dass $v = \lambda_1 b_1 + \lambda_2 b_2 + ... + \lambda_m b_m$? In dem Fall sind die Vektoren $b_1, ..., b_m$ *linear unabhängig*.
- Gibt es für jedes $v \in V$ *mindestens ein* $m$-Tupel $(\lambda_1, ..., \lambda_m) \in K^m$, so dass $v = \lambda_1 b_1 + \lambda_2 b_2 + ... + \lambda_m b_m$? In dem Fall bilden die $b_1, ..., b_m$ ein *Erzeugendensystem* von $V$.

**Beispiel 4.5.1.**

1) Betrachte z. B. $V = \mathbb{R}^3$ und die *Einheitsvektoren*

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

$e_1, e_2$ und $e_3$ bilden eine *Basis* des $\mathbb{R}^3$. Das heißt, jeder Vektor $v \in \mathbb{R}^3$ lässt sich in Bezug auf diese Basis durch *genau ein* 3-Tupel $(\lambda_1, \lambda_2, \lambda_3)$ spezifizieren, denn

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \stackrel{!}{=} \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$\Leftrightarrow v_1 = \lambda_1, v_2 = \lambda_2, v_3 = \lambda_3$.

Bemerkung: Dies lässt sich verallgemeinern auf $n$ Einheitsvektoren $e_1, ..., e_n \in \mathbb{R}^n$.

2) Betrachte wieder $V = \mathbb{R}^3$ und die Einheitsvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

In diesem Fall ist

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \stackrel{!}{=} \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

genau dann, wenn $v_1 = \lambda_1$, $v_2 = \lambda_2$ und $v_3 = 0$.

In diesem Fall lassen sich nicht alle $v \in V$ durch $e_1$ und $e_2$ darstellen ($e_1$ und $e_2$ formen also *keine* Basis). Für jedes $v \in V$ gibt es *höchstens ein* (explizit: ein oder kein) darstellendes 2-Tupel $(\lambda_1, \lambda_2)$, die Vektoren $e_1$ und $e_2$ nennt man linear unabhängig.

3) Betrachte wieder $V = \mathbb{R}^3$ und die Vektoren

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad a_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die Vektoren $a_1, a_2, a_3$ und $a_4$ bilden *keine* Basis des $\mathbb{R}^3$ und sind auch *nicht* linear unabhängig, aber sie bilden ein Erzeugendensystem, denn jeder Vektor $v \in \mathbb{R}^3$ lässt sich in Bezug auf $a_1, a_2, a_3$ und $a_4$ durch *mindestens ein* (hier explizit: durch mehrere) 4-Tupel $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ spezifizieren, wie man sich leicht überlegt.

**Definition 4.5.1** (Lineare Hülle, Erzeugendensystem)**.** Sei $m \in \mathbb{N}$. Sei $V$ ein $K$-Vektorraum und $b_1, ..., b_m \in V$. Man bezeichnet die Menge der *Linearkombinationen* von $b_1, ..., b_m$,

$$\text{span}(b_1, ..., b_m) = \left\{ \sum_{i=1}^{m} \lambda_i b_i \middle| \lambda_i \in K \quad \forall i = 1, ..., m \right\},$$

als *lineare Hülle* oder *Spann* von $b_1, ..., b_m$. Weiter heißt $b_1, ..., b_m$ *Erzeugendensystem* von $V$ genau dann, wenn

$$\text{span}(b_1, ..., b_m) = V.$$

Existiert ein Erzeugendensystem von $V$, so sagen wir, $V$ ist *endlich erzeugt*.

**Bemerkung 4.5.1** (Erzeugendensystem)**.** Es folgt direkt aus dieser Definition, dass, wenn $b_1, ..., b_m$ Erzeugendensystem von $V$ ist, für jedes $v \in V$ *mindestens eine* Linearkombination, also mindestens ein $\lambda = (\lambda_1, ..., \lambda_m) \in K^m$, existiert mit $v = \sum_{i=1}^{m} \lambda_i b_i$.

**Lemma 4.5.1** (Kleinster Untervektorraum)**.** *Sei $m \in \mathbb{N}$. Sei $V$ ein $K$-Vektorraum und seien $b_1, ..., b_m \in V$. Dann ist* $\text{span}(b_1, ..., b_m)$ *der kleinste Untervektorraum, der $b_1, .., b_m$ enthält, das heißt, für jeden Untervektorraum $U \subseteq V$ mit $b_1, ..., b_m \in U$ gilt:*

$$\text{span}(b_1, ..., b_m) \subseteq U.$$

*Beweis.* Zu zeigen: 1) $\text{span}(b_1, ..., b_m)$ ist Untervektorraum und 2) $\text{span}(b_1, ..., b_m)$ ist *kleinster* Untervektorraum.

Zu 1): Dass $\text{span}(b_1, ..., b_m)$ Untervektorraum von $V$ ist, folgt direkt aus der Definition (siehe Definitionen 4.5.1 und 4.3.1).

Zu 2): Sei nun $U$ ein beliebiger Untervektorraum von $V$. Seien $b_1, ..., b_m \in U$. Weil $U$ Untervektorraum ist, gilt $\forall \lambda_1, ..., \lambda_m \in K : \lambda_1 b_1 + ... + \lambda_m b_m \in U \Rightarrow \text{span}(b_1, ..., b_m) \subseteq U$. □

**Definition 4.5.2** (Lineare Unabhängigkeit)**.** Sei $V$ ein $K$-Vektorraum, $m \in \mathbb{N}$ und $b_1, ..., b_m \in V$. Die Vektoren $b_1, ..., b_m$ heißen *linear unabhängig* genau dann, wenn $\forall \lambda_1, ..., \lambda_m \in K$ gilt:

$$\lambda_1 b_1 + \ldots + \lambda_m b_m = 0 \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_m = 0.$$

Die Vektoren $b_1, ..., b_m$ heißen *linear abhängig* genau dann, wenn sie *nicht* linear unabhängig sind.

Sei weiter $v \in V$. Der Vektor $v$ heißt *linear abhängig von* $b_1, ..., b_m$ genau dann, wenn:

$$v \in \text{span}(b_1, ..., b_m).$$

**Lemma 4.5.2** (Eindeutigkeit der Darstellung)**.** *Sei $m \in \mathbb{N}$, $V$ ein $K$-Vektorraum und $b_1, ..., b_m \in V$. Seien $\lambda := (\lambda_1, ..., \lambda_m) \in K^m$ und $\mu := (\mu_1, ..., \mu_m) \in K^m$. Die Vektoren $b_1, .., b_m$ sind linear unabhängig genau dann, wenn*

$$\forall \lambda,\ \mu \in K^m : \quad \sum_{i=1}^{m} \lambda_i b_i = \sum_{i=1}^{m} \mu_i b_i \Rightarrow \lambda_i = \mu_i \quad \forall i = 1, ..., m.$$

*Beweis.* „⇒": Falls die Vektoren linear unabhängig sind, gilt:

$$\sum_{i=1}^{m} \lambda_i b_i = \sum_{i=1}^{m} \mu_i b_i \Leftrightarrow \sum_{i=1}^{m} (\lambda_i - \mu_i) b_i = 0 \Rightarrow \lambda_i - \mu_i = 0 \, \forall i = 1, ..., m \Leftrightarrow \lambda_i = \mu_i \, \forall i = 1, ..., m$$

„⇐": Wir wenden die Bedingung an auf den Nullvektor $\mu = 0 \in K^m$ (d.h. $\mu_1 = ... = \mu_m = 0$). Also steht da $\sum_{i=1}^{m} \lambda_i b_i = 0 \Rightarrow \lambda_1 = ... = \lambda_m = 0$, aber das ist genau die Aussage, dass die $b_1, ..., b_m$ linear unabhängig sind. □

**Bemerkung 4.5.2** (Lineare Unabhängigkeit)**.** Es gibt also für jedes $v \in V$ *höchstens eine* Darstellung als Linearkombination linear unabhängiger Vektoren $b_1, ..., b_m \in V$, also höchstens ein $m$-Tupel $\lambda = (\lambda_1, ..., \lambda_m) \in K^m$ mit $v = \sum_{i=1}^{m} \lambda_i b_i$.

**Definition 4.5.3** (Basis)**.** Sei $V$ ein $K$-Vektorraum, $m \in \mathbb{N}$ und $b_1, ..., b_m \in V$. Die Vektoren $b_1, ..., b_m$ heißen *Basis* von $V$ genau dann, wenn $b_1, ..., b_m$ linear unabhängig sind und ein Erzeugendensystem von $V$ bilden.

**Beispiel 4.5.2.**

- Die kanonischen Einheitsvektoren $e_1, ..., e_n$ bilden eine Basis des $\mathbb{R}^n$.
- Die *Monome* $p_k = x^k$ mit $k = 0, ..., n$ bilden eine Basis des Vektorraums der reellen Polynome vom Grad $n$, $P_n = \left\{ \sum_{k=0}^{n} a_k x^k | a_k \in \mathbb{R} \, \forall k = 0, ..., n \right\}$.

**Bemerkung 4.5.3.** Für den Vektorraum $V = \{0\}$ wird die leere Menge $\emptyset$ als Basis festgelegt.

**Bemerkung 4.5.4.** Wenn $b_1, ..., b_m$ eine Basis von $V$ ist, dann existiert also für jedes $v \in V$ *genau ein* $m$-Tupel $\lambda = (\lambda_1, ..., \lambda_m)$ mit $v = \sum_{i=1}^{m} \lambda_i b_i$. (Siehe die Bemerkungen zu „Lineare Unabhängigkeit" und „Erzeugendensystem", wonach in dem einen Fall „höchsten eine", im anderen Fall „mindestens eine" Linearkombination existiert – sollte beides erfüllt sein, also „genau eine".)

**Lemma 4.5.3.** *Sei $V$ ein $K$-Vektorraum. Sei $m \in \mathbb{N}$ und $v_1, ..., v_{m+1} \in V$. Dann gilt:*

1) *$v \in V$ linear abhängig* $\Leftrightarrow v = 0$

2) *für $m \geq 2$:*
*$v_1, ..., v_m$ linear abhängig* $\Leftrightarrow \exists k \in \{1, ..., m\} : v_k \in \text{span}(v_1, ..., v_{k-1}, v_{k+1}, ..., v_m)$

3) *für $m \geq 1$:*
*$v_{m+1}$ linear abhängig von $v_1, ..., v_m$* $\Leftrightarrow \text{span}(v_1, ..., v_{m+1}) = \text{span}(v_1, ..., v_m)$

4) *$v_1, ..., v_m$ linear unabhängig und $v_{m+1} \notin \text{span}(v_1, ..., v_m)$* $\Leftrightarrow$ *$v_1, ..., v_{m+1}$ linear unabhängig*

*Beweis.* Wir wissen: $v_1, ..., v_m \in V$ sind linear abhängig $\Leftrightarrow \exists \lambda_1, ..., \lambda_m \in K$ und $k \in \{1, ..., m\}$ : $\sum_{i=1}^{m} \lambda_i v_i = 0 \wedge \lambda_k \neq 0$.

Zu 1): „⇒": $v \in V$ linear abhängig, d.h. $\exists \lambda \in K : \lambda v = 0 \wedge \lambda \neq 0 \Rightarrow v = 0$.

„⇐": $v = 0$, d.h. $1 \cdot v = 0 \wedge 1 \neq 0 \Rightarrow v$ linear abhängig.

Zu 2): „⇒": $v_1, ..., v_m$ linear abhängig $\Rightarrow \exists \lambda_1, ..., \lambda_m \in K$ und $k \in \{1, ..., m\} : \sum_{i=1, i\neq k}^{m} \lambda_i v_i + \lambda_k v_k = 0$ mit $\lambda_k \neq 0 \Rightarrow v_k = -\frac{1}{\lambda_k} \sum_{i=1, i\neq k}^{m} \lambda_i v_i \Rightarrow v_k \in \text{span}(v_1, ..., v_{k-1}, \hat{v}_k, v_{k+1}, ...., v_m)$. Hier bedeutet „$\hat{v}_k$", dass $v_k$ nicht enthalten ist.

„⇐": $v_k \in \text{span}(v_1, ..., v_{k-1}, \hat{v}_k, v_{k+1}, ...., v_m) \Rightarrow v_k = \sum_{i=1, i\neq k}^{m} \lambda_i v_i$. Wähle nun $\lambda_k = -1$, dann ist $\sum_{i=1}^{m} \lambda_i v_i = 0$ und damit $v_1, ..., v_m$ linear abhängig.

Zu 3): „⇒": Sei $v_{m+1}$ linear abhängig von $v_1, ..., v_m$. Zu zg.: $\text{span}(v_1, ..., v_{m+1}) \subseteq \text{span}(v_1, ..., v_m)$ und $\text{span}(v_1, ..., v_{m+1}) \supseteq \text{span}(v_1, ..., v_m)$. Dabei ist die zweite Inklusion klar und für die erste Inklusion gilt: $v_{m+1} = \sum_{i=1}^{m} \lambda_i v_i$, weil $v_{m+1} \in \text{span}(v_1, ..., v_m)$, und damit gilt für beliebiges $v \in \text{span}(v_1, ..., v_{m+1})$ :

$$\begin{aligned} v &= \sum_{i=1}^{m} \mu_i v_i + \mu_{m+1} v_{m+1} = \sum_{i=1}^{m} \mu_i v_i + \mu_{m+1} \sum_{i=1}^{m} \lambda_i v_i \\ &= \sum_{i=1}^{m} (\mu_i + \mu_{m+1}\lambda_i) v_i = \sum_{i=1}^{m} \kappa_i v_i \end{aligned}$$

mit $\kappa_i = \mu_i + \mu_{m+1}\lambda_i \in K$ für alle $i = 1, ..., m$, also $v \in \text{span}(v_1, ..., v_m)$.

„⇐": Offensichtlich ist $v_{m+1} \in \text{span}(v_1, ..., v_{m+1})$ und $\text{span}(v_1, ..., v_{m+1}) = \text{span}(v_1, ..., v_m) \Rightarrow v_{m+1} \in \text{span}(v_1, ..., v_m)$, also $v_{m+1}$ linear abhängig von $v_1, ..., v_m$.

Zu 4): „⇐": Offensichtlich.

„⇒": Sei $\sum_{i=1}^{m+1} \lambda_i v_i = 0$. Z. zg.: $\lambda_i = 0$ für alle $i = 1, ..., m+1$.

1. Fall: $\lambda_{m+1} \neq 0 \Rightarrow v_{m+1} = \sum_{i=1}^{m} \left( -\frac{\lambda_i}{\lambda_{m+1}} \right) v_i$, also $v_{m+1} \in \text{span}(v_1, ..., v_m) \Rightarrow$ Widerspruch

2. Fall: $\lambda_{m+1} = 0 \Rightarrow \sum_{i=1}^{m} \lambda_i v_i = 0$. Weil $v_1, ..., v_m$ linear unabhängig, folgt daraus: $\lambda_i = 0$ für alle $i = 1, ..., m$. Somit $\lambda_i = 0$ für alle $i = 1, ..., m+1$.

□

**Satz 4.5.1** (Basisauswahlsatz). *Sei $V \neq \{0\}$ ein endlich erzeugter $K$-Vektorraum, $m \in \mathbb{N}$ und $b_1, ..., b_m$ ein Erzeugendensystem von $V$. Dann existiert ein $n \in \mathbb{N}$ mit $n \leq m$, so dass (nach geeigneter Umnummerierung) $b_1, ..., b_n$ eine Basis von $V$ bilden.*

**Bemerkung 4.5.5.** Das heißt, aus jedem Erzeugendensystem kann man durch geschicktes Weglassen von Vektoren eine Basis konstruieren.

*Beweis.* Sei $b_1, ..., b_m$ ein Erzeugendensystem von $V \neq \{0\}$.

Sei o.B.d.A. $b_1, ..., b_m \neq 0$ (Wir können im Erzeugendensystem alle Nullvektoren weglassen, da diese keinen Beitrag zur Linearkombination eines beliebigen Vektors $v \neq 0, v \in V$ leisten; gleichzeitig können nicht alle $b_1, ..., b_m = 0$ sein, da sonst $V = \{0\}$.)

Nun machen wir eine Fallunterscheidung:

1. Fall: $b_1, ..., b_m$ sind linear unabhängig. Dann sind $b_1, ..., b_m$ bereits Basis von $V$

2. Fall: $b_1, ..., b_m$ sind linear abhängig. Nach Annahme sind $b_1, ..., b_m \neq 0$. Dann ist $m \geq 2$, denn anderenfalls wäre $b_1$ linear abhängig und damit $b_1 = 0$ (siehe Aussage 1 von Lemma 4.5.3),

was ein Widerspruch zur Annahme ist. Nun $\exists p \in \{1, ..., m\} : b_p \in \text{span}(b_1, ..., b_{p-1}, b_{p+1}, ..., b_m)$ (siehe Aussage 2 von Lemma 4.5.3). Wir können nun umbenennen. Wir nennen $b_p$ nun $b_m$ und $b_m$ nennen wir $b_p$. Damit ist das neue $b_m$ linear abhängig von $b_1, ..., b_{m-1}$. Weil aber $\text{span}(b_1, ..., b_{m-1}) = \text{span}(b_1, ..., b_m)$ (siehe Aussage 3 von Lemma 4.5.3), ist auch $b_1, ..., b_{m-1}$ Erzeugendensystem von $V$. Diese Schritte können wir so oft wiederholen, bis wir zu einem Erzeugendensystem $b_1, ..., b_n$ aus linear unabhängigen Vektoren kommen. Dies ist spätestens bei $n = 1$ der Fall, denn $b_1 \neq 0$ (und damit $b_1$ linear unabhängig, siehe wieder Aussage 1 von Lemma 4.5.3). □

**Lemma 4.5.4.** *Seien $n, p \in \mathbb{N}$. Sei $V$ ein $K$-Vektorraum und $b_1, ..., b_n$ eine Basis von $V$. Seien $a_1, ..., a_p \in V$ linear unabhängig. Dann gilt:*

$$p \leq n$$

*Beweis.* Wir zeigen die Kontraposition der Aussage mittels vollständiger Induktion. Statt der Aussage $a_1, ..., a_p \in V$ linear unabhängig $\Rightarrow p \leq n$ zeigen wir also

$$p > n \quad \Rightarrow \quad a_1, ..., a_p \in V \text{ linear abhängig.}$$

Induktionsanfang ($n = 1$): $b_1$ ist eine Basis von $V$, insbesondere wird also $V$ von $b_1$ erzeugt. Für Vektoren $a_1, .., a_p \in V$, $p > 1$, gilt dann: $a_i = \alpha_i b_1$ mit $\alpha_i \in K$ $\forall i = 1, .., p$. Falls eines der $a_i$ der Nullvektor ist, so sind die $a_1, ..., a_p$ linear abhängig (siehe Def. der Unabhängigkeit; wegen $a_i = 0$ folgt aus $\lambda_1 a_1 + ... + \lambda_p a_p = 0$, dass $\lambda_i$ beliebig und *nicht* $\lambda_i = 0$). Ansonsten ist z.B. $\alpha_2 a_1 + (-\alpha_1) a_2 = \alpha_2 \alpha_1 b_1 + (-\alpha_1) \alpha_2 b_2 = 0$ eine nichttriviale Linearkombination zu 0 und damit $a_1, ..., a_p$ linear abhängig.

Induktionsschritt ($n \Rightarrow n + 1$): Wir nehmen nun an, die Aussage gilt für festes $n$. Wir betrachten den Fall $n+1$. Sei also $b_1, ..., b_{n+1}$ Basis von $V$. Insbesondere wird also $V$ von $b_1, ..., b_{n+1}$ erzeugt. In $V = \text{span}(b_1, ..., b_{n+1})$ seien Vektoren $a_1, .., a_p$ mit $p > n + 1$ gegeben. Das heißt, es gibt Zahlen $\alpha_{jk} \in K$, so dass $\forall j = 1, ..., p$:

$$a_j = \sum_{k=1}^{n+1} \alpha_{jk} b_k.$$

Seien nicht alle $\alpha_{jk} = 0$ (sonst wären alle $a_j = 0$ und damit alle $a_1, ..., a_p$ linear abhängig). Sei etwa $\alpha_{11} \neq 0$. Dann liegen die $p - 1$ Vektoren

$$c_j := a_j - \frac{\alpha_{j1}}{\alpha_{11}} a_1 = \sum_{k=1}^{n+1} \frac{\alpha_{11}\alpha_{jk} - \alpha_{j1}\alpha_{1k}}{\alpha_{11}} b_k \quad (j = 2, ..., p)$$

im $\text{span}(b_2, .., b_{n+1})$, weil der Koeffizient von $b_1$ gleich 0 ist. Wegen $p - 1 > n$ sind nach Induktionsannahme, die $p - 1$ Vektoren $c_2, ..., c_p$ linear abhängig. Also gibt es Zahlen $\lambda_2, ...\lambda_p \in K$, die

nicht alle 0 sind und für die gilt:

$$0 = \sum_{j=2}^{p} \lambda_j c_j = \sum_{j=2}^{p} \lambda_j \left( a_j - \frac{\alpha_{j1}}{\alpha_{11}} a_1 \right) = \sum_{j=1}^{p} \lambda_j a_j \quad \text{mit } \lambda_1 := -\frac{1}{\alpha_{11}} \sum_{j=2}^{p} \lambda_j \alpha_{j1}.$$

Also sind auch die Vektoren $a_1, ..., a_p$ linear abhängig. □

**Satz 4.5.2** (Existenz einer Basis)**.** *Jeder endlich erzeugte $K$-Vektorraum $V$ besitzt eine Basis. Alle Basen von $V$ besitzen dieselbe Anzahl an Vektoren.*

*Beweis.* Wir machen eine Fallunterscheidung.

1. Fall: $V = \{0\}$. Basis von $V$ ist die leere Menge $\emptyset$. Da $0 \in V$ linear abhängig ist, kann es keine Basis geben, die irgendeinen Vektor enthält. Die Anzahl an Vektoren ist also gleich 0.

2. Fall: $V \neq \{0\}$. Da $V$ endlich erzeugt ist, existiert gemäß Basisauswahlsatz eine Basis von $V$ (und dies ist nicht die leere Menge).

Seien nun $b_1, ..., b_m$ und $b'_1, ..., b'_n$ zwei Basen von $V$ mit $n, m \in \mathbb{N}$. Wir können nun Lemma 4.5.4 benutzen, um zu zeigen, dass beide Basen dieselbe Anzahl an Elementen haben, d.h. $m = n$. Dafür betrachten wir einmal $b_1, ..., b_m$ als Basis und $b'_1, ..., b'_n$ als linear unabhängige Vektoren in $V$, woraus folgt, dass $n \leq m$. Das andere Mal betrachten wir $b'_1, ..., b'_n$ als Basis und $b_1, ..., b_m$ als linear unabhängige Vektoren in $V$, woraus folgt, dass $m \leq n$. Es folgt: $n = m$. □

## 4.6 Dimension

**Definition 4.6.1** (Dimension)**.** Sei $V$ ein endlich erzeugter $K$-Vektorraum.

1) Die Anzahl $k \in \mathbb{N}_0$ der Vektoren einer Basis von $V$ nennt man die *Dimension* von $V$. Wir schreiben: $\dim V = k$.

2) Endlich erzeugte Vektorräume nennt man *endlichdimensional*. Wir schreiben: $\dim V < \infty$.

3) Ist ein Vektorraum $V$ nicht endlich erzeugt, nennt man ihn *unendlichdimensional*. Wir schreiben: $\dim V = \infty$.

**Beispiel 4.6.1.** Sei $V = \{0\}$. Es ist $\dim \{0\} = 0$ (es gibt keinen linear unabhängigen Basisvektor). Sei $V = K^n$. Es ist $\dim K^n = n$ (z. B. sind die Einheitsvektoren $e_1, ..., e_n$ eine Basis).

**Satz 4.6.1** (Basisergänzungssatz)**.** *Seien $n, p \in \mathbb{N}$. Sei $V$ ein $K$-Vektorraum und $b_1, ..., b_n$ eine Basis von $V$. Seien $a_1, ..., a_p \in V$ linear unabhängig. Dann gibt es $a_{p+1}, ..., a_n \in \{b_1, ..., b_n\}$, so dass $a_1, ..., a_n$ eine Basis von $V$ bildet.*

**Bemerkung 4.6.1.** Das heißt, durch Hinzunahme geeigneter Basisvektoren kann jede Menge von linear unabhängigen Vektoren zu einer Basis ergänzt werden.

**Bemerkung 4.6.2** (Verallgemeinerung)**.** Der Basisergänzungssatz gilt sogar noch allgemeiner und zwar für jedes Erzeugendensystem $b_1, ..., b_n$ (statt einer Basis $b_1, ..., b_n$). Auch Erzeugendensysteme $b_1, ..., b_n$ kann man also nutzen, um eine Menge linear unabhängiger Vektoren $a_1, ..., a_p$ zu einer Basis zu ergänzen.

*Beweis (von Satz 4.6.1).* Wir können zwei Fälle unterscheiden.

1) $\forall i = 1, ..., n$: $b_i \in \text{span}(a_1, ..., a_p)$. Dann ist $\text{span}(b_1, ..., b_n) \subseteq \text{span}(a_1, ..., a_p)$ und damit $a_1, ..., a_p$ Erzeugendensystem von $V$ und $a_1, ..., a_p$ linear unabhängig, also $a_1, ..., a_p$ Basis. Weil jede Basis dieselbe Anzahl an Vektoren hat (Satz 4.5.2), gilt in diesem Fall: $p = n$.

2) $\exists i \in 1, ..., n$: $b_i \notin \text{span}(a_1, ..., a_p)$. Wir setzen $a_{p+1} = b_i$. Nach Aussage 4) von Lemma 4.5.3 sind $a_1, ..., a_p, a_{p+1}$ linear unabhängig und insbesondere ist $b_i \in \text{span}(a_1, ..., a_{p+1})$. Dieses Vorgehen können wir unter Hinzunahme geeigneter Basisvektoren wiederholen, bis wir eine Menge von Vektoren $a_{p+1}, ..., a_{p+m}$ mit $m \in N (m < n)$ gefunden haben, so dass $a_1, ..., a_{p+m}$ linear unabhängig und $\forall i = 1, ..., n : b_i \in \text{span}\ (a_1, ..., a_{p+m})$. Dann ist $\text{span}(b_1, ..., b_n) \subseteq \text{span}(a_1, ..., a_{p+m})$, also $a_1, ..., a_{p+m}$ Erzeugendensystem von $V$ und $a_1, ..., a_{p+m}$ linear unabhängig, also $a_1, ..., a_{p+m}$ Basis. Insbesondere ist in diesem Fall $p + m = n$, also $p < n$. □

**Bemerkung 4.6.3.** Aus dem Basisergänzungssatz folgt unmittelbar: In einem $n$-dimensionalen Vektorraum bilden je $n$ linear unabhängige Vektoren eine Basis.

**Lemma 4.6.1.** *Sei $m \in \mathbb{N}$. Sei $V$ ein $K$-Vektorraum und $b_1, ..., b_m \in V$. Dann gilt:*

$$b_1, ..., b_m \text{ Basis von } V \Rightarrow \exists f : K^m \to V \text{ linear und bijektiv, } f(e_i) = b_i \, \forall i = 1, ..., m.$$

*Beweis.* Wir definieren $f : K^m \to V, f(x) := \sum_{i=1}^m x_i b_i$ und zeigen, dass dieses $f$ die gesuchten Eigenschaften hat.

Linearität: $f$ ist linear, denn

$$\forall x, y \in K^m : f(x + y) = \sum_{i=1}^m (x_i + y_i) b_i = \sum_{i=1}^m x_i b_i + \sum_{i=1}^m y_i b_i = f(x) + f(y)$$

und

$$\forall x \in K^m, \lambda \in K : f(\lambda x) = \sum_{i=1}^m \lambda x_i b_i = \lambda \sum_{i=1}^m x_i b_i = \lambda f(x)$$

Injektivität: $f$ ist injektiv, denn $f(x) = 0 \Leftrightarrow \sum_{i=1}^m x_i b_i = 0 \Rightarrow x_i = 0 \, \forall i = 1, ..., m \Leftrightarrow x = 0$, wobei der vorletzte Schritt folgt, weil die $b_i$ unabhängig sind. Somit ist Kern $f = \{0\}$ und damit $f$ injektiv (siehe Übung). Zudem ist $f$ surjektiv, weil $b_1, ..., b_m$ ein Erzeugendensystem von $V$ ist $\Rightarrow f$ ist bijektiv.

Zuletzt gilt nach Definition von $f$: $f(e_i) = b_i \ \forall i = 1, ..., m$. □

**Lemma 4.6.2** (Umkehrfunktion)**.** *Seien $V$ und $W$ $K$-Vektorräume. Sei $f : V \to W$ linear und bijektiv. Dann ist $f^{-1} : W \to V$ linear und bijektiv.*

*Beweis.* Die Bijektivität von $f^{-1}$ haben wir bereits gezeigt. Zu zeigen ist die Linearität:

Seien $w_1, w_2 \in W$. Dann ist

$$f^{-1}(w_1 + w_2) = f^{-1}(f(f^{-1}(w_1)) + f(f^{-1}(w_2))) = f^{-1}(f(f^{-1}(w_1) + f^{-1}(w_2))) = f^{-1}(w_1) + f^{-1}(w_2)$$

Dabei haben wir im ersten und letzten Schritt die Definition der Umkehrfunktion und im zweiten Schritt die Linearität von $f$ benutzt haben.

Sei $w \in W, \lambda \in K$. Dann ist

$$f^{-1}(\lambda \cdot w) = f^{-1}(\lambda \cdot f(f^{-1}(w)) = f^{-1}(f\lambda \cdot (f^{-1}(w)) = \lambda f^{-1}(w)$$

$\Rightarrow f^{-1}$ ist linear □

**Satz 4.6.2.** *Jeder $K$-Vektorraum $V$ mit $n = \dim V$ ist isomorph zu $K^n$.*

*Beweis.* Wir erinnern uns, dass ein Vektorraum $V$ isomorph zu einem Vektorraum $W$ ist, wenn es eine lineare, bijektive Abbildung $f : V \to W$ gibt. Weil $\dim V = n$, existiert eine Basis $b_1, ..., b_n$ von $V$ (Satz 4.5.2). Damit existiert eine lineare und bijektive Funktion $f : K^n \to V$ (siehe Lemma 4.6.1) und somit eine lineare und bijektive Umkehrfunktion $f^{-1} : V \to K^n$ (siehe Lemma 4.6.2), d.h. jeder $K$-Vektorraum $V$ mit $\dim V = n$ ist isomorph zu $K^n$. □

**Lemma 4.6.3.** *Sei $V$ ein endlich erzeugter $K$-Vektorraum und $U \subseteq V$ ein Unterraum von $V$. Dann ist $U$ endlich erzeugt und es gilt:*

$$\dim U \leq \dim V$$

*Beweis.* Die Fälle $V = \{0\}, U = \{0\}$ und $V \neq \{0\}, U = \{0\}$ sind klar (denn $\{0\}$ ist Unterraum von jedem Vektorraum und $\dim\{0\} = 0$).

Wir betrachten im Folgenden den allgemeinen Fall $V \neq \{0\}, U \neq \{0\}$. Weil $V$ endlich erzeugt ist, existiert eine Basis $b_1, ..., b_n$, d.h. es existiert ein $n \in \mathbb{N}$ mit $\dim V = n$. Für jede Menge linear unabhängiger Vektoren $u_1, ..., u_p \in U \subseteq V$ gilt dann (Lemma 4.5.4): $p \leq n = \dim V$. Wegen $U \neq \{0\}$ existieren linear unabhängige $u_1, ..., u_p$ mit $1 \leq p \leq n$ und insbesondere gibt es demnach eine maximale Anzahl $p_{max}$ linear unabhängiger Vektoren: $u_1, ..., u_{p_{max}}$.

Wir behaupten nun: $\forall u \in U : u \in \text{span}(u_1, ..., u_{p_{max}})$. Anderenfalls gäbe es ein $u' \notin \text{span}(u_1, ..., u_{p_{max}})$ und damit $u_1, ..., u_{p_{max}}, u'$ linear unabhängig (siehe Lemma 4.5.3) im Widerspruch zur Annahme, dass $p_{max}$ maximal ist. Folglich ist also $u_1, ..., u_{p_{max}}$ ein Erzeugendensystem von $U$ (d.h. insbesondere: $U$ ist endlich erzeugt) und $\dim U = p_{max} \leq n = \dim V$. □

**Satz 4.6.3** (Dimensionsformel Untervektorräume)**.** *Sei $V$ ein $K$-Vektorraum mit* $\dim V < \infty$. *Seien $U_1, U_2$ Untervektorräume von $V$. Dann gilt:*

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

*Beweis.* Übung. □

**Satz 4.6.4** (Dimensionsformel lineare Abbildungen)**.** *Seien $V, W$ $K$-Vektorräume, $f : V \to W$ eine lineare Abbildung und* $\dim V < \infty$. *Dann gilt:*

$$\dim\ \text{Kern} f + \dim \text{Bild} f = \dim V$$

*Beweis.* Weil dim $V < \infty$, gibt es ein $n \in \mathbb{N}$ mit $n := \dim V$.

Sei nun $b_1, ..., b_m$ die Basis von Kern$f$ (der Fall $m = 0$ bezeichne die leere Basis). Eine Basis existiert, weil Kern$f$ Untervektorraum von $V$ und damit selbst ein Vektorraum ist (siehe Satz 4.3.4). Weil Kern$f \subseteq V$: dim Kern$f = m \leq n = \dim V$.

Weil sie die Basis von Kern$f$ bilden, sind die $b_1, ..., b_m$ linear unabhängige Vektoren aus $V$ und so können wir sie laut Basisergänzungssatz durch Hinzunahme von $n - m$ Basisvektoren $b_{m+1}, ..., b_n$ zu einer Basis $b_1, ..., b_n$ von $V$ ergänzen. Bleibt zu zeigen, dass $f(b_{m+1}), ..., f(b_n)$ eine Basis von Bild$f$ bilden (dann ist $n - m = \dim$ Bild$f$ und damit: dim Kern$f$ + dim Bild$f$ = $m + (n - m) = n = \dim V$).

Wir müssen also zeigen: $f(b_{m+1}), ..., f(b_n)$ Erzeugendensystem vom Bild$f$ und $f(b_{m+1}), ..., f(b_n)$ linear unabhängig

1) $f(b_{m+1}), ..., f(b_n)$ Erzeugendensystem: Zunächst gilt: $f(b_1), ..., f(b_n)$ Erzeugendensystem von Bild$f$ (denn für jedes $v \in V : \sum_{i=1}^{n} \lambda_i b_i = v$ und damit gilt für jedes $w \in$ Bild$f : w = f(v) = f(\sum_i \lambda_i b_i) = \sum_i \lambda_i f(b_i)$, weil $f$ linear). Nun ist aber $f(b_j) = 0 \, \forall j = 1, ..., m$ (weil $b_1, ..., b_m \in$ Kern$f) \Rightarrow f(b_{m+1}), ..., f(b_n)$ Erzeugendensystem von Bild$f$.

2) $f(b_{m+1}), ..., f(b_n)$ linear unabhängig: Z. zg.:$\sum_{i=m+1}^{n} \lambda_i f(b_i) = 0 \Rightarrow \lambda_i = 0 \, \forall i = m+1, ..., n$. Nun ist wegen Linearität von $f$:

$$\sum_{i=m+1}^{n} \lambda_i f(b_i) = 0 \Leftrightarrow f\left( \sum_{i=m+1}^{n} \lambda_i b_i \right) = 0.$$

$$\Rightarrow \sum_{i=m+1}^{n} \lambda_i b_i \in \text{Kern} f \;\Rightarrow \sum_{i=m+1}^{n} \lambda_i b_i = \sum_{i=1}^{m} \mu_i b_i$$

$$\Leftrightarrow \sum_{i=m+1}^{n} \lambda_i b_i + \sum_{i=1}^{m} (-\mu_i) b_i = 0$$

Nun sind aber die $b_i$ $(i = 1, ..., n)$ linear unabhängig (weil Basis von $V$) und damit folgt aus dieser Gleichung, dass $(-\mu_i) = 0 \, \forall i = 1, ..., m \wedge \lambda_i = 0 \, \forall i = m + 1, ..., n$. Dies zeigt die Aussage. □

# 5 $\mathbb{R}^n$ als euklidischer Vektorraum

## 5.1 Skalarprodukt, Längen, Winkel

Der Vektorraum $\mathbb{R}^n$ mit darauf definiertem Skalarprodukt ist ein euklidischer Vektorraum.

**Bemerkung 5.1.1.** Im Unterschied zu nicht-euklidischen Räumen gilt in euklidischen Räumen das Parallelenpostulat von Euklid (siehe den Exkurs zu nicht-euklidischer Geometrie).

Mit dem Skalarprodukt können wir Längen, Abstände und Winkel messen. Mit anderen Worten, erst mit dem Skalarprodukt kommen wir zur Geometrie.

**Definition 5.1.1** (Skalarprodukt)**.** Seien $x, y \in \mathbb{R}^n$. Wir nennen

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i$$

das *Skalarprodukt* von $x$ und $y$. Wir schreiben auch: $\langle x, y\rangle = x^T y$.

**Lemma 5.1.1.** *Die Abbildung*

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto \langle x, y\rangle$$

*ist bilinear, symmetrisch und positiv definit.*

1) *Für* $x, x', y, y' \in \mathbb{R}^n$ *und* $\lambda \in \mathbb{R}$*:* $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$ *und* $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ *sowie* $\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle$ *und* $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$ (bilinear)

2) *Für* $x, y \in \mathbb{R}^n$*:* $\langle x, y\rangle = \langle y, x\rangle$ (symmetrisch)

3) *Für* $x \in \mathbb{R}^n$*:* $\langle x, x\rangle \geq 0$ *und* $\langle x, x\rangle = 0 \Leftrightarrow x = 0$ (positiv definit)

**Bemerkung 5.1.2.** Oft wird das Skalarprodukt direkt über diese Eigenschaften (und nicht wie oben für den speziellen Fall $\mathbb{R}^n$) definiert.

**Definition 5.1.2** (Orthogonalität)**.** Seien $x, y \in \mathbb{R}^n$. Falls

$$\langle x, y\rangle = 0,$$

so nennen wir $x$ und $y$ *orthogonal*. Wir schreiben: $x \perp y$.

**Definition 5.1.3** (Länge, Abstand)**.** Sei $x \in \mathbb{R}^n$. Wir nennen

$$|x| := \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

die *euklidische Norm* oder *Länge* von $x$. Alternativ schreiben wir: $|x| = ||x||_2$. Mit der euklidischen Norm können wir auch den *Abstand* $d(X, Y)$ zwischen zwei Punkten $X, Y \in \mathbb{R}^n$ definieren:

$$d(X, Y) := |x - y| = \sqrt{\langle x - y, x - y\rangle} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

**Bemerkung 5.1.3** (Notation)**.** In der Regel unterscheiden wir nicht zwischen Punkten $X \in \mathbb{R}^n$ und Vektoren $x \in \mathbb{R}^n$. Tun wir dies doch, bezeichnen wir Punkte mit Großbuchstaben, Vektoren mit Kleinbuchstaben.

Das Skalarprodukt und die euklidische Norm erfüllen die Cauchy-Schwarz-Ungleichung.

**Satz 5.1.1** (Cauchy-Schwarz-Ungleichung)**.** *Seien $x, y \in \mathbb{R}^n$. Dann gilt:*

$$|\langle x, y\rangle| \leq |x| \cdot |y|.$$

*Hier gilt „=" genau dann, wenn $x$ und $y$ linear abhängig sind.*

*Beweis.* Im Spezialfall $x = 0$ oder $y = 0$ sind $x$ und $y$ linear abhängig und es gilt:

$$|\langle x, y\rangle| = 0 = |x| \cdot |y|.$$

Im allgemeinen Fall $x \neq 0, y \neq 0$ gilt für alle $\lambda \in \mathbb{R}$:

$$0 \leq \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - 2\lambda\langle x, y\rangle + \lambda^2\langle y, y\rangle.$$

Wähle nun $\lambda := \frac{\langle x,y\rangle}{\langle y,y\rangle}$, dann ist:

$$0 \leq \langle x, x\rangle - \frac{\langle x, y\rangle^2}{\langle y, y\rangle} \quad \Leftrightarrow \quad \langle x, y\rangle^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle = |x|^2 \cdot |y|^2 \quad \Leftrightarrow \quad |\langle x, y\rangle| \leq |x| \cdot |y|.$$

Hier haben wir im letzten Schritt benutzt, dass $f(x) = x^2$ auf $\mathbb{R}_0^+$ monoton steigend ist (d.h. $x^2 \leq y^2 \Leftrightarrow x \leq y$). □

**Definition 5.1.4** (Winkel)**.** Seien $x, y \in \mathbb{R}^n \setminus \{0\}$. Wir nennen $\phi \in [0, \pi]$ mit

$$\cos\phi = \frac{\langle x, y\rangle}{|x| \cdot |y|} \quad \left(\text{bzw.} \quad \phi = \arccos\frac{\langle x, y\rangle}{|x| \cdot |y|}\right)$$

den Winkel zwischen $x$ und $y$. Wir schreiben: $\phi = \angle(x, y)$.

**Bemerkung 5.1.4** (Eindeutigkeit des Winkels)**.** Für je zwei Vektoren $x, y$ gibt es genau ein $\phi \in [0, \pi]$, das die Gleichung $\cos\phi = \frac{\langle x,y\rangle}{|x|\cdot|y|}$ erfüllt. Dies folgt aus der Cauchy-Schwarz-Ungleichung, wonach $|\langle x, y\rangle| \leq |x| \cdot |y|$, also $-1 \leq \frac{\langle x,y\rangle}{|x|\cdot|y|} \leq 1$, sowie daraus, dass der Kosinus $f(\phi) = \cos(\phi)$ das Intervall $[0, \pi]$ bijektiv auf $[-1, 1]$ abbildet.

**Bemerkung 5.1.5.** Aus der Definition des Winkels folgt unmittelbar, dass

$$|x \pm y|^2 = \langle x \pm y, x \pm y\rangle = \langle x, x\rangle \pm 2\langle x, y\rangle + \langle y, y\rangle = |x|^2 + |y|^2 \pm 2\,|x||y|\cos\phi,$$

wobei dies, für $\phi = \frac{\pi}{2}$ (dies ist der Fall $x \perp y$), den Satz des Pythagoras ergibt:

$$|x \pm y|^2 = |x|^2 + |y|^2.$$

## 5.2 Weihnachtsvorlesung: nicht-euklidische Geometrie

Wir machen nun einen Exkurs in die nicht-euklidische Geometrie. Wie bereits erwähnt gilt in euklidischen Vektorräumen (wie dem $\mathbb{R}^n$) das Parallelenpostulat. Dazu brauchen wir folgende Definition:

**Definition 5.2.1.** Zwei Geraden $g, h$ in einer Ebene heißen parallel, wenn sie keinen Schnittpunkt haben.

Das Parallelenpostulat von Euklid besagt nun (in moderner Form):

> „Zu jeder Geraden $g$ und jedem Punkt $P$ außerhalb von $g$ existiert *genau eine* zu $g$ parallele Gerade $h$, die durch $P$ verläuft."

Das Parallelenpostulat gilt im euklidischen Raum $\mathbb{R}^n$. Explizit können wir dort die Parallele $h$ durch $P$ konstruieren, indem wir das Lot auf $g$ durch $P$ fällen und darauf wiederum das Lot nehmen, das durch $P$ geht.

Im Unterschied dazu gilt das Parallelenpostulat in nicht-euklidischen Räumen *nicht*. Dabei gibt es zwei Möglichkeiten, wie es verletzt sein kann:

1) Es gibt *keine* parallele Gerade zu $g$ durch $P \Rightarrow$ *Elliptische* oder *sphärische Geometrie* (z.B. Kugeloberfläche, Oberfläche unserer Erde, $n$-dimensionale Sphäre $S_n$). Hier ist die Innenwinkelsumme im Dreieck größer als 180 Grad. Anordnungsaxiome gelten nicht.

2) Es gibt *mindestens zwei* parallele Geraden zu $g$ durch $P \Rightarrow$ *Hyperbolische Geometrie* (z.B. Sattelfläche, Minkowski-Raumzeit, Klein'sche Kreisscheibe, Poincaré'sche Scheibe, Poincaré'sche Halbebene). Hier ist die Innenwinkelsumme im Dreieck kleiner als 180 Grad. Anordnungsaxiome gelten.

Da das Parallelenpostulat in nicht-euklidischen Räumen nicht gilt, gelten auch viele der geometrischen Sätze, die wir aus dem euklidischen $\mathbb{R}^n$ kennen, nicht, und geometrische Objekte schauen anders aus, als wir es gewohnt sind.

Historische Entwicklung:

- Nikolai Lobatschewskij, János Bolyai (1826): hyperbolische Geometrie
- Carl Friedrich Gauß ($\sim$ 1826): sphärische Geometrie
- Bernhard Riemann (1854): Differentialgeometrie gekrümmter Räume $\Rightarrow$ Grundlage zur Beschreibung der gekrümmten Raumzeit in der Allg. Relativitätstheorie
- Eugene Beltrami, Felix Klein (1868, 1871): Beltrami-Klein-Modell (Klein'sche Kreisscheibe)
- Henri Poincaré (1882): Poincaré'sches Scheibenmodell, Poincaré'sche Halbebene