

Lineare Algebra (Informatik)

Dr. Paula Reichert

3. Februar 2026

Inhaltsverzeichnis

1	Einführung	3
1.1	Historische Einführung	3
1.2	Euklidische Mathematik	4
2	Grundbegriffe	8
2.1	Logik	8
2.2	Vollständige Induktion	11
2.3	Mengenlehre	16
2.4	Relationen	22
2.5	Abbildungen (Funktionen)	25
2.6	Unendliche Mengen	29
3	Gruppen, Ringe, Körper	33
3.1	Gruppen	33
3.2	Ringe	36
3.3	Körper	38
4	Vektorräume und lineare Abbildungen	43
4.1	Vektorräume	43
4.2	Lineare Abbildungen	46
4.3	Untervektorräume	49
4.4	Matrizen	51
4.5	Erzeugendensystem, lineare Unabhängigkeit, Basis	61
4.6	Dimension, Koordinatendarstellung	67
5	$\mathbb{R}^n$ als euklidischer Vektorraum	73
5.1	Skalarprodukt, Längen, Winkel	73
5.2	Weihnachtsvorlesung: nicht-euklidische Geometrie	75
5.3	ONS, ONB, Orthogonalmatrizen	76

6	Lineare Gleichungssysteme	80
6.1	Elementare Zeilen-/Spaltenumformungen von Matrizen	80
6.2	Lösung von Gleichungssystemen der Form $Ax = y$ mit $x, y \in K^n$	83
6.3	Rang einer Matrix	88
6.4	Lösung von Gleichungssystemen der Form $Ax = y$ mit $x \in K^n, y \in K^m$	89
7	Determinanten	93
8	Eigenwerte und Eigenvektoren	98
9	Ausblick: Lineare Algebra 2	108

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-ğabr*, 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 (in etwa Gebiet des heutigen Irak) und Ägypten:

Babylonier (ca. 1800 v. Chr.): Gleichungssysteme der folgenden Form werden in erster Näherung gelöst. Hier sind a, b gegeben, x, y gesucht:

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

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

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

$$\begin{aligned} x + a \cdot x &= b \\ x + c \cdot x + d \cdot x &= e \end{aligned}$$

mit a, b, c, d, e gegeben und 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.

Es folgt eine Weiterentwicklung der Mathematik im arabisch-persischen Raum mit Höhepunkt im 9. Jahrhundert. Die Araber greifen dabei sowohl auf die indische als auch die griechische Mathematik zurück. Die Werke werden ins Arabische übersetzt, weiterentwickelt und gelangen durch die (Rück-)Eroberung Spaniens wieder zurück nach Europa.

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, Gauss, Laplace, Cantor etc. die weitere Entwicklung.

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

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

Dieses Gleichungssystem kann man auch schreiben als:

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

Oder in Kurzschreibweise:

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

Dabei ist $A = (a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}$ 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.

Bemerkung 1.2.1 (Größtes gemeinsames Maß). Besitzen zwei Strecken ein gemeinsames Maß, so besitzen sie ein *größtes gemeinsames Maß* (ggM), und dieses ist *eindeutig*. Beweis: Wechselwegnahme/euklidischer Algorithmus.

Definition 1.2.2 (ggT). Das größte gemeinsame Maß zweier natürlicher Zahlen $m, n \in \mathbb{N}$ ist der ggT (*größter gemeinsamer Teiler*).

Der *Euklidische Algorithmus* zur Konstruktion des größten gemeinsamen Maßes ist die sogenannte Wechselwegnahme.

Bemerkung 1.2.2 (Euklidischer Algorithmus). Man betrachte dazu zwei Strecken X und Y . Sei o.B.d.A. $Y \leq X$. Um das größte gemeinsame Maß (bzw. bei natürlichen Zahlen den ggT) zu bestimmen, trage man Y so oft auf X ab, bis ein Rest r_1 bleibt, $0 \leq r_1 < Y$. Wenn $r_1 = 0$, dann gilt: $Y = \text{ggM}(X, Y)$. Wenn $r_1 > 0$, trage man r_1 auf Y so oft ab, bis ein Rest r_2 bleibt, $0 \leq r_2 < r_1$. Wenn $r_2 = 0$, dann gilt: $r_1 = \text{ggM}(X, Y)$. Wenn $r_2 > 0$, trage man r_2 auf r_1 so oft ab, bis ein Rest r_3 bleibt usw.

Damit ergibt sich das folgendes Gleichungssystem (mit $n_i \in \mathbb{N} \forall i = 0, \dots, k; Y = r_0$):

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

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

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

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

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 (Kettenbruch). Kettenbruchdarstellung von $\frac{5}{3}$ und $\text{ggT}(5, 3)$.

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

und $\text{ggT}(5, 3) = 1$.

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

Satz 1.2.1 (Diagonale im Quadrat). *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. $d = a + a_1$). An dieser Stelle fällt man das Lot auf d . Die Strecke, die sich von dort bis zum Schnittpunkt mit der Seite a ergibt, hat die Länge a_1 (gleichschenkliges Dreieck!). Die Strecke bildet nach Konstruktion die Seite eines kleineren Quadrats mit Diagonale d_1 . Dabei bildet d_1 einen Abschnitt der Seite a mit $a = a_1 + d_1$ (man beachte das Drachenviereck, an dem man erkennt, dass $a - d_1 = a_1$!). Nun beginnt die ganze Konstruktion von vorne (nur diesmal am kleineren Quadrat). Man trage also 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 $k \in \mathbb{N}$, $a = a_0$, $d = d_0$):

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

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:

$$a^2 + a^2 = d^2 \quad \Leftrightarrow \quad 2a^2 = d^2 \quad \Leftrightarrow \quad 2 = \frac{d^2}{a^2} \quad \Leftrightarrow \quad \sqrt{2} \equiv \frac{d}{a}.$$

Aus dem Satz des Pythagoras erkennen wir also, dass die Zahl $\sqrt{2}$ genau dem Verhältnis aus Diagonale und Seite im Quadrat entspricht (wie Euklid ja feststellte, sind reelle Zahlen Verhältnisse von Strecken). Wichtig ist dabei: Egal von welcher Länge a das Quadrat ist, das Verhältnis $\frac{d}{a}$ ist immer $\sqrt{2}$.

Und wir wissen nun bereits aus Satz 1.2.1, was $\sqrt{2}$ für eine Zahl ist. Wir wissen, dass $\frac{d}{a} = \sqrt{2}$ irrational ist und dass gilt:

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

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):

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

2 Grundbegriffe

2.1 Logik

Die Aussagenlogik untersucht die Struktur von Sätzen (Aussagen) $\mathcal{A}, \mathcal{B}, \dots$ 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 $\mathcal{A} \vee \mathcal{B}$).

Seien $\mathcal{A}$ und $\mathcal{B}$ Aussagen.

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \wedge \mathcal{B}$	$\mathcal{A} \vee \mathcal{B}$
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,*

$$\mathcal{A} \wedge \mathcal{B} \Leftrightarrow \mathcal{B} \wedge \mathcal{A}$$

$$(\mathcal{A} \wedge \mathcal{B}) \wedge \mathcal{C} \Leftrightarrow \mathcal{A} \wedge (\mathcal{B} \wedge \mathcal{C})$$

$$(\mathcal{A} \vee \mathcal{B}) \wedge \mathcal{C} \Leftrightarrow (\mathcal{A} \wedge \mathcal{C}) \vee (\mathcal{B} \wedge \mathcal{C})$$

und analog für $\vee$ (bzw. $\vee$ und $\wedge$ vertauscht im DG).

Weiter gelten die De-Morgan-Regeln:

$$\neg(\mathcal{A} \wedge \mathcal{B}) \Leftrightarrow \neg\mathcal{A} \vee \neg\mathcal{B}$$

$$\neg(\mathcal{A} \vee \mathcal{B}) \Leftrightarrow \neg\mathcal{A} \wedge \neg\mathcal{B}$$

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 „ $\Leftrightarrow$ “ stehenden Aussagen (Übung). $\square$

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.

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

Bemerkung 2.1.1. Sind zwei Aussagen $\mathcal{A}_1$ und $\mathcal{A}_2$ äquivalent, dann ist die Aussage $\mathcal{A}_1 \Leftrightarrow \mathcal{A}_2$ eine Tautologie.

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

Sei $\mathcal{A}$ eine Aussage.

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

Satz 2.1.2. Seien $\mathcal{A}$ und $\mathcal{B}$ Aussagen. Die Aussagen $(\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A})$ 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*). $\square$

Bemerkung 2.1.2. 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 (Ball). Wir nehmen zur Veranschaulichung ein Alltagsbeispiel, auch wenn Alltagsbeispiele nie ganz präzise sind. Seien $\mathcal{A}, \mathcal{B}$ die folgenden Aussagen:

$\mathcal{A}$: „der Junge trifft mit dem Ball die Scheibe“

$\mathcal{B}$: „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 daraus, dass der Junge mit dem Ball nicht die Scheibe getroffen hat“.

Bemerkung 2.1.3. Statt eine Aussage *direkt* zu beweisen, können wir sie also auch durch *Kontraposition* ($\neg \mathcal{B} \Rightarrow \neg \mathcal{A}$) beweisen.

Bemerkung 2.1.4 (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: Aus $\mathcal{A} \Rightarrow \mathcal{B}$ folgt *nicht* die Aussage $\neg \mathcal{A} \Rightarrow \neg \mathcal{B}$, wie das folgende Beispiel zeigt.

Beispiel 2.1.5. Sei $k \in \mathbb{Z}$. Seien $\mathcal{A}$ und $\mathcal{B}$ die folgenden Aussagen:

$\mathcal{A}$: „ k ist durch 9 teilbar“

$\mathcal{B}$: „ k ist durch 3 teilbar“

Dann gilt: $\mathcal{A} \Rightarrow \mathcal{B}$. 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. $k = 6$.

In diesem Beispiel haben wir die Wahrheit der Aussage $\mathcal{A} \Rightarrow \mathcal{B}$ bereits angenommen. Was machen wir, wenn wir diese noch beweisen müssen?

Bemerkung 2.1.5 (Widerspruchsbeweis). Die Aussage $\mathcal{A} \Rightarrow \mathcal{B}$ beweisen wir in der Regel durch einen Widerspruch. Wir nehmen dazu an, dass $\mathcal{A}$ gilt, aber $\mathcal{B}$ *nicht*, formal $\mathcal{A} \wedge \neg \mathcal{B}$, und zeigen dann mittels korrekter Schlüsse, dass dies zu einem Widerspruch führt.

Beispiel 2.1.6 (Widerspruchsbeweis). 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“ (siehe oben).

Zu zeigen: $\mathcal{A} \Rightarrow \mathcal{B}$

Beweis. Wir nehmen 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 $k = (3 \cdot m) \cdot 3$. Daraus folgt, dass es ein $p \in \mathbb{Z}$ gibt (nämlich $p = 3 \cdot m$) mit $k = p \cdot 3$. Somit ist k durch 3 teilbar und dies ist im Widerspruch zur Annahme. Somit gilt: $\mathcal{A} \Rightarrow \mathcal{B}$.

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

Satz 2.1.4 (Transitivität von „ $\Rightarrow$ “ und „ $\Leftrightarrow$ “). Seien $\mathcal{A}, \mathcal{B}, \mathcal{C}$ Aussagen. Dann ist allgemein gültig:

$$(\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{C}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})$$

$$(\mathcal{A} \Leftrightarrow \mathcal{B}) \wedge (\mathcal{B} \Leftrightarrow \mathcal{C}) \Rightarrow (\mathcal{A} \Leftrightarrow \mathcal{C})$$

Beweis. Mittels Wahrheitstafeln (Übung). □

Bemerkung 2.1.6. Dieser Satz kann (später) mittels vollständiger Induktion auf eine beliebige Anzahl an Aussagen $\mathcal{A}_1, \dots, \mathcal{A}_n$, $n \in \mathbb{N}$, erweitert werden.

Bemerkung 2.1.7 (Prädikatenlogik). Zum Abschluss noch ein Exkurs zur Prädikatenlogik. In der Logik unterscheiden wir formal die Aussagenlogik von der Prädikatenlogik. Wechseln wir zur Prädikatenlogik kommen zusätzlich noch Variablen $x, y, \dots$ und (zwei) *Quantoren* dazu:

$\exists$: „es existiert“

$\forall$: „für alle“

Aussagen werden nun durch *Prädikate* $\mathcal{P}, \mathcal{Q}, \dots$ 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“.

In der Prädikatenlogik gibt es allgemein gültige Regeln zur Verneinung der Quantoren:

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

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

Wir sind nun bereits sehr nah an der Mengenlehre. 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.

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

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

Wichtig: Die Reihenfolge von Quantoren darf nicht vertauscht werden!

Beispiel 2.1.7 (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 $n \in \mathbb{N}$ (oder für alle $n \in \mathbb{N}_0$) gilt. Dabei ist $\mathcal{A}(n)$ in der Regel eine Gleichung, in der ein Parameter $n \in \mathbb{N}$ vorkommt.

Für die Induktion 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 $\mathcal{A}(n)$ wahr ist (Induktionsannahme), $\mathcal{A}(n+1)$ 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.

Es ist wichtig, bei 1) mit dem ersten nicht-trivialen Fall der Aussage zu beginnen. Dann können wir von $\mathcal{A}(1)$ ausgehend mit Hilfe von 2) iterativ die Wahrheit der nachfolgenden Aussagen feststellen:

$$\mathcal{A}(1) \text{ wahr} \Rightarrow \mathcal{A}(2) \text{ wahr} \Rightarrow \mathcal{A}(3) \text{ wahr} \Rightarrow \dots$$

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 vollständigen Induktion basiert auf der Tatsache, dass wir die natürlichen Zahlen anordnen können (*Wohlordnungsprinzip* von $\mathbb{N}$, siehe *axiomatische Mengenlehre*).

Bemerkung 2.2.3 (Minimum, Maximum). Insbesondere gibt es wegen des Wohlordnungsprinzips von $\mathbb{N}$ für jede nicht-leere endliche Teilmenge A von $\mathbb{N}$ ein kleinstes Element, d.h.

$$\exists k \in A, \text{ sodass } \forall n \in A : k \leq n,$$

und ein größtes Element, d.h.

$$\exists g \in A, \text{ sodass } \forall n \in A : g \geq n.$$

Wir nennen k das *Minimum* von A und g das *Maximum* von A .

Wir schreiben: $k =: \min A$ und $g =: \max A$.

Achtung: Nicht jede Teilmenge von $\mathbb{R}$ enthält ein kleinstes und größtes Element!

Bemerkung 2.2.4 (Offenes, abgeschlossenes Intervall). Seien $a, b \in \mathbb{R}$. Mit $[a, b] \subset \mathbb{R}$ bezeichnet man das *abgeschlossene* Intervall. Es enthält seine Randpunkte a, b . Mit $]a, b[\subset \mathbb{R}$ bezeichnet man das *offene* Intervall. Es enthält seine Randpunkte a, b nicht (siehe *Analysis*).

Bemerkung 2.2.5 (Supremum, Infimum). Man betrachte nun das offene Intervall $]0, 1[\subset \mathbb{R}$. Es gilt: $0 \notin]0, 1[$ und $1 \notin]0, 1[$. In diesem Fall gibt es kein Maximum bzw. Minimum, aber alle aufsteigenden bzw. absteigenden Folgen, die in $]0, 1[$ liegen, nähern sich 1 bzw. 0 an. Wir nennen 1 das *Supremum*, $\sup]0, 1[= 1$, und 0 das *Infimum*, $\inf]0, 1[= 0$. Falls es ein Maximum bzw. Minimum gibt, gilt: „ $\sup = \max$ “ bzw. „ $\inf = \min$ “ (siehe *Analysis*).

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:

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

und insbesondere

$$2) \quad \sum_{k=0}^{\infty} q^k = \frac{1}{1 - q}.$$

Bemerkung 2.2.6 (Summenzeichen). Zunächst ein Hinweis zur Schreibweise. Σ ist das sog. „Summenzeichen“. Wir summieren damit über die verschiedenen Werte von k wie im Folgenden angegeben:

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

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

$$\sum_{k=0}^n q^k = \sum_{k=0}^n \left(\frac{1}{2}\right)^k = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots + \left(\frac{1}{2}\right)^n$$

Beweis (von Satz 2.2.1). Wir beweisen die erste Aussage mittels vollständiger Induktion. Die zweite Aussage folgt daraus mit Hilfe der Regeln für die Grenzwertbildung.

Induktionsanfang ($n = 0$):

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

Die Aussage ist also für $n = 0$ wahr. Wir können uns interessehalber noch $n = 1$ anschauen:

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

Die Aussage ist also auch für $n = 1$ 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 ($n \Rightarrow n + 1$):

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

Hier haben wir im zweiten Schritt die Induktionsannahme verwendet. Wir haben also gezeigt: Aus $\mathcal{A}(n)$ wahr, folgt, $\mathcal{A}(n + 1)$ wahr.

Daraus folgt: Aussage 1) ist für alle $n \in \mathbb{N}_0$ wahr.

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

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

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

Beispiel 2.2.2 (Geometrische Reihe). Betrachte $q = \frac{1}{2}$. Dann gilt:

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

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

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

Beweis. Mittels vollständiger Induktion (Übung). $\square$

Bemerkung 2.2.7 (Historisch). Es heißt, Gauß (1777 – 1855) habe die Summenformel in der Grundschule entdeckt, nachdem er von seinem Lehrer aufgefordert wurde, die natürlichen Zahlen von 1 bis 100 zusammen zu zählen. Rechnung:

$$1 + 2 + \dots + 99 + 100 = (1 + 100) + (2 + 99) + \dots + (50 + 51) = 50 \cdot (101) = 5050.$$

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

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

Bemerkung 2.2.8. Man kennt den binomischen Lehrsatz aus der Wahrscheinlichkeitsrechnung. Dort ist $x = p$ die Wahrscheinlichkeit, dass ein gewisses Ereignis eintritt und $y = (1 - p)$ die Wahrscheinlichkeit, dass es nicht eintritt. Die Summe $x + y = p + 1 - p = 1$ ist die Gesamtwahrscheinlichkeit. Beispiel (n -facher Münzwurf): Hier ist $p = 1/2$ die Wahrscheinlichkeit für „Zahl“ und $1 - p = 1/2$ die Wahrscheinlichkeit für „Kopf“. Der Binomialkoeffizient $\binom{n}{k}$ gibt die Zahl der Möglichkeiten an, bei denen k aus n Münzen „Zahl“ zeigen. Σ summiert über alle Möglichkeiten, wie oft „Zahl“ unter n Würfeln geworfen werden kann (von $k = 0$ bis $k = n$).

Beweis (von Satz 2.2.3). 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 ($n \Rightarrow n + 1$):

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

Hier haben wir im zweiten Schritt die Induktionsannahme verwendet 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 $k \geq 1$ gilt:

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

Damit folgt:

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

Wir haben also gezeigt: Aus $\mathcal{A}(n)$ wahr, folgt, $\mathcal{A}(n + 1)$ wahr.

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

□

Bemerkung 2.2.9 (Vollständige Induktion). Statt des Induktionsschrittes $n \Rightarrow n+1$, 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 Induktionsschritt. In Worten: Man zeigt, dass die Aussage, wenn sie für $k < n$ gilt, auch für n gilt.

2.3 Mengenlehre

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:

$x \in M$: „ x ist ein Element von M “

$x \notin M$: „ x ist kein Element von M “

Bemerkung 2.3.3 (Russel'sche Antinomie). Man kann zunächst versuchen, die Mengendefinition wie folgt zu verstehen (*naive Mengenlehre*). Sei G die Menge aller Objekte (unserer Anschauung oder unseres Denkens). Sei $\mathcal{E}(x)$ eine Eigenschaft von Objekten x aus G . Dann bildet

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

wieder eine Menge, nämlich die Menge aller Objekte $x \in G$ mit der Eigenschaft $\mathcal{E}(x)$. 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 die sog. *Russel'sche Antinomie*, das ist

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

Formal:

$$R = \{x \mid x \notin x\}$$

Wäre $R \in R \Rightarrow R \notin R$, was ein Widerspruch ist. Andersherum, falls $R \notin R \Rightarrow R \in R$, 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: $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}$... (siehe *axiomatische Mengenlehre*).

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

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

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)
- $\mathbb{N} = \{1, 2, 3, \dots\}$: die Menge der natürlichen Zahlen
- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$: die Menge der ganzen Zahlen
- $\mathbb{Q} = \{\frac{p}{q} | p \in \mathbb{Z} \wedge q \in \mathbb{Z} \setminus \{0\}\}$: 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

Im Folgenden seien A, B jeweils Mengen.

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 ,

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

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:

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

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 Ω 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:

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

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:

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

Satz 2.3.2 (Rechenregeln für $\cap$ und $\cup$). 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 Rechenregeln für $\wedge$ und $\vee$ (Übung). $\square$

Bemerkung 2.3.6 (Notation). Wir können die Definitionen von $\cap$ und $\cup$ auf endlich und sogar abzählbar unendlich viele – mehr zu diesen Begriffen später – Mengen ausdehnen. Seien $A_1, A_2, \dots, A_n$ Mengen, $n \in \mathbb{N}$. Dann ist:

$$\begin{aligned} \bigcap_{i=1}^n A_i &= A_1 \cap A_2 \cap \dots \cap A_n \quad \text{und} \quad \bigcap_{n=1}^{\infty} A_n = \lim_{n \rightarrow \infty} \bigcap_{i=1}^n A_i \\ \bigcup_{i=1}^n A_i &= A_1 \cup A_2 \cup \dots \cup A_n \quad \text{und} \quad \bigcup_{n=1}^{\infty} A_n = \lim_{n \rightarrow \infty} \bigcup_{i=1}^n A_i \end{aligned}$$

Beispiel 2.3.6. Man betrachte reelle Intervalle $A_n, B_n \subseteq \mathbb{R}$ (auch Intervalle sind Mengen, nämlich Teilmengen von $\mathbb{R}$). Seien $a, b \in \mathbb{R}, a < b$, und $n \in \mathbb{N}$. Sei $A_n =]a - 1/n, b + 1/n[$. Dann ist

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

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

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

Definition 2.3.8 (Kartesisches Produkt). Als *kartesisches Produkt* nicht-leerer Mengen $A_1, \dots, A_n$ bezeichnen wir die Menge aller Folgen der Länge n , sog. *n-Tupel* $(a_1, \dots, a_n)$, mit $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$:

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

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

Bemerkung 2.3.7. 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 2.3.8 (Reelle Folgen). Man kann die Definition des kart. Produkts auf abzählbar unendlich viele Mengen $A_k, k \in \mathbb{N}$, ausdehnen. So ist z.B. $\mathbb{R}^{\mathbb{N}}$ die Menge aller reellen Folgen $(a_1, a_2, \dots)$ mit $a_k \in \mathbb{R}$ für alle $k \in \mathbb{N}$.

Bemerkung 2.3.9 (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 2.3.10 (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 2.3.11 (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 $A_1, \dots, A_n$ endliche Mengen. Dann gilt:

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

Beweis. Wir suchen die Anzahl an n -Tupeln $(a_1, \dots, a_n)$ mit $a_i \in A_i \forall i = 1, \dots, n$. 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. $\square$

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

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

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

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

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

Beispiel 2.3.8 (Zu c)). Sei $n = 7$. Dann sind mögliche Indexmengen jeweils Teilmengen von $\{1, 2, 3, \dots, 6, 7\}$, z.B. $I_1 = \{2, 5\}$, $I_2 = \{1, 2, 4, 6, 7\}$ und $I_3 = \{3\}$.

In diesem Fall ist $|I_1| = 2$, $|I_2| = 5$ sowie $|I_3| = 1$. Weiter ist $\bigcap_{i \in I_1} A_i = A_2 \cap A_5$ und $\bigcap_{i \in I_2} A_i = A_1 \cap A_2 \cap A_4 \cap A_6 \cap A_7$ und $\bigcap_{i \in I_3} A_i = A_3$.

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: $|\bigcup_{i=1}^1 A_i| = |A_1| = (-1)^{1-1} |\bigcap_{i=1}^1 A_i|$. Das stimmt also.

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

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

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“. $\square$

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 $R_1 = \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}$, $R_2 = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x > y\}$ und $R_3 = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x = y\}$.

Eine wichtige Art von Relationen sind Äquivalenzrelationen.

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

- 1) $\forall x \in X : x \sim x$ (reflexiv)
- 2) $\forall x, y \in X : x \sim y \Leftrightarrow y \sim x$ (symmetrisch)
- 3) $\forall x, y, z \in X : x \sim y \wedge y \sim z \Rightarrow x \sim z$ (transitiv)

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:

- 1) $\forall x \in X : x = x$ (reflexiv)
- 2) $\forall x, y \in X : x = y \Leftrightarrow y = x$ (symmetrisch)
- 3) $\forall x, y, z \in X : x = y \wedge y = z \Rightarrow x = z$ (transitiv)

□

Wir betrachten ein weiteres wichtiges Beispiel.

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

$$x \sim y \Leftrightarrow m \mid (x - y)$$

ist eine Äquivalenzrelation. Hier steht $m \mid (x - y)$ für „ m teilt $(x - y)$ “.

Beweis. Wir prüfen die drei Bedingungen:

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

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

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

$$x \equiv y \pmod{m} :\Leftrightarrow m \mid (x - y)$$

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

Beispiel 2.4.2. Es ist $7 \equiv 2 \pmod{5}$ und $7 \equiv -2 \pmod{3}$.

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. zu repräsentieren).

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 $a \sim b \Leftrightarrow m | (a - b)$. In dem Fall 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

$$m\mathbb{Z} := \{k \cdot m | k \in \mathbb{Z}\} = \{\dots, -2m, -m, 0, m, 2m, \dots\}$$

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:

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

Man bezeichnet diese Klassen auch als „Restklassen“ (denn die Zahlen einer Klasse liefern bei Division durch m denselben Rest!) und schreibt für ihre Menge:

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

Hat man einmal festgelegt, dass man auf $\mathbb{Z}_m$ rechnet, kann man die eckigen Klammern auch wieder weglassen und schreibt dann wieder 1 statt $[1]$ oder $[1]_m$.

Bemerkung 2.4.2 (Quotientenraum). $\mathbb{Z}_m$ definiert einen sog. *Quotientenraum*,

$$\mathbb{Z}_m =: \mathbb{Z}/m\mathbb{Z}.$$

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 $a \not\sim b$).

Beweis. Es gibt nur zwei mögliche Fälle, entweder $a \sim b$ oder $a \not\sim b$.

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 $a \not\sim b$. 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. $\square$

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

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

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 $f : X \rightarrow Y$. Wir nennen X den *Definitionsbereich*, Y den *Wertebereich* und $R = G_f$ den *Graph* von f .

Definition 2.5.2 (Bild, Urbild). Sei $f : X \rightarrow Y$ 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 2.5.1. Insbesondere ist demnach $f(A) \subseteq Y$ und $f^{-1}(B) \subseteq X$.

Beispiel 2.5.1 (Bild, Urbild). Betrachte $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$. 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 2.5.3 (Verkettung). Seien $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ Abbildungen. Dann heißt die Abbildung $X \rightarrow Z, x \rightarrow g(f(x))$ *Verkettung* von g und f . Wir schreiben dafür $g \circ f$.

Satz 2.5.1 (Assoziativgesetz). Seien $f : W \rightarrow X, g : X \rightarrow Y, h : Y \rightarrow Z$ 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 2.5.2 (Verkettung). Sei $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2 + 2$ und $g : \mathbb{R} \rightarrow \mathbb{R}, g(x) = 3x - 5$. Dann ist:

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

Weiter gilt:

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

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 2.5.4 (Injektivität, Surjektivität, Bijektivität). Sei $f : X \rightarrow Y$ 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 2.5.2. Wir nennen eine bijektive Abbildung auch eine *1:1-Abbildung*.

Beispiel 2.5.3 (Injektivität, Surjektivität, Bijektivität). Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}_0^+, x \mapsto x^2$ ist nicht injektiv, denn $f(1) = f(-1)$ (genauer: für $x = 1, x' = -1$ gilt: $x \neq x'$, aber $f(x) = f(x')$, also f nicht injektiv), aber f 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 $y \in \mathbb{R}_0^+$).

Wir führen noch drei spezielle Funktionen ein.

Definition 2.5.5 (Identität, Einbettung, Einschränkung). Sei X eine Menge. Die Abbildung

$$\text{id}_X : X \rightarrow X, \quad x \rightarrow x$$

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

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

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

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

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

Bemerkung 2.5.3. Insbesondere ist demnach $i_A = \text{id}_X|_A$.

Beispiel 2.5.4 (Einschränkung). Man betrachte wieder die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}_0^+, x \mapsto x^2$. Diese Funktion ist nicht injektiv, aber die Einschränkung $f|_{\mathbb{R}_0^+} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+, x \mapsto x^2$ ist injektiv und surjektiv, also bijektiv.

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

- 1) $g \circ f$ injektiv $\Rightarrow f$ injektiv
- 2) $g \circ f$ surjektiv $\Rightarrow g$ surjektiv

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 $g : Y \rightarrow Z$. 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. $\square$

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

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

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

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

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

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

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

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 id_X bijektiv $\Rightarrow f^{-1} \circ f$ bijektiv $\Rightarrow f$ injektiv, f^{-1} surjektiv.

Weil 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 $f_1^{-1} : Y \rightarrow X$ und $f_2^{-1} : Y \rightarrow X$ 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 $f_1^{-1} = f_2^{-1}$. $\square$

Satz 2.5.4 (Eigenschaften von f^{-1}). *Seien $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ 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) „ $\Leftarrow$ “ wurde bereits gezeigt (siehe Satz 2.5.3). „ $\Rightarrow$ “: 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 $f^{-1} : Y \rightarrow X, f^{-1}(y) := x$. 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 $f : X \rightarrow Y$, 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:

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

Anders herum gilt:

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

$f^{-1} \circ g^{-1}$ 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). $\square$

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 $\phi : M \rightarrow N$ 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 (d.h. $|M| = n$) $\Leftrightarrow \exists \phi : \{1, \dots, n\} \mapsto M$ 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 $\mathbb{N}$ (d.h. $|M| = |\mathbb{N}|$)
- 5) M heißt *überabzählbar unendlich* $\Leftrightarrow M$ nicht endlich und M nicht abzählbar unendlich

Bemerkung 2.6.1 ($\mathbb{N}$ bzw. $\mathbb{N}_0$). Sei M gleichmächtig zu $\mathbb{N}$. Ist $|M| = |\mathbb{N}|$, so ist $|M| = |\mathbb{N}_0|$, denn weil $\phi : M \rightarrow \mathbb{N}$ bijektiv und $\phi' : \mathbb{N} \rightarrow \mathbb{N}_0, n \mapsto n - 1$ bijektiv, so ist auch $\phi' \circ \phi : M \rightarrow \mathbb{N}_0$ bijektiv. 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 $3\mathbb{N} = \{n \in \mathbb{N} | n = 3k, k \in \mathbb{N}\} = \{3, 6, 9, 12, \dots\}$. Dann gilt: $3\mathbb{N} \subset \mathbb{N}$ und $|3\mathbb{N}| = |\mathbb{N}|$ (denn es existiert $\phi : \mathbb{N} \rightarrow 3\mathbb{N}, n \rightarrow 3n$ 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 trivialerweise abzählbar (unendlich), denn $\phi = \text{id}_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{N}, n \rightarrow n$ ist bijektiv.
- $\mathbb{Z}$ ist abzählbar (unendlich), denn es existiert eine bijektive Abbildung

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

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:

$$\left(\begin{array}{ccccccc} 1 & \rightarrow & \frac{1}{2} & & \frac{1}{3} & \rightarrow & \frac{1}{4} & \dots \\ & \swarrow & & \nearrow & & \swarrow & & \\ \frac{2}{1} & & \frac{2}{2} & & \frac{2}{3} & & & \dots \\ \downarrow & \nearrow & & \swarrow & & & & \\ \frac{3}{1} & & \frac{3}{2} & & & \dots & \dots & \\ & \swarrow & & \swarrow & & & & \\ \frac{4}{1} & & & \dots & \dots & \dots & \dots & \\ \downarrow & & & & & & & \\ \vdots & & \vdots & & \vdots & & \vdots & \end{array} \right)$$

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	$+\frac{3}{2}$	-3	$+\frac{1}{3}$	$-\frac{1}{3}$	$+\frac{1}{4}$	$-\frac{1}{4}$	$+\frac{2}{3}$	$-\frac{2}{3}$	...

□

Die Menge der Primzahlen $\mathbb{P}$ ist eine weitere bekannte Menge, die abzählbar unendlich ist.

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, d.h. $|\mathbb{P}| = |\mathbb{N}|$.

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, siehe Übung). Daraus folgt: $|\mathbb{P}| = |\mathbb{N}|$. Insbesondere können wir die Primzahlen $\mathbb{P}$ als Teilmenge von $\mathbb{N}$ der Größe nach anordnen (siehe Bem. 2.2.2 zum Wohlordnungsprinzip von $\mathbb{N}$) mit $p_1 < p_2 < p_3 < \dots$. Damit können wir die Primzahlen „abzählen“, d.h. $\exists \phi$ bijektiv: $\mathbb{N} \rightarrow \mathbb{P}, i \rightarrow p_i$. $\square$

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

Satz 2.6.3. Die Menge der reellen Zahlen $\mathbb{R}$ ist überabzählbar unendlich, d.h. $|\mathbb{R}| > |\mathbb{N}|$.

Beweis. Cantors zweites Diagonalargument (Übung). $\square$

Bemerkung 2.6.3 ($\aleph$, Kontinuumshypothese). Cantor zeigt also: $|\mathbb{R}| > |\mathbb{N}|$.

Wir bezeichnen 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? Diese These nennt man die *Kontinuumshypothese*. Sie stammt von Cantor selbst. 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. Dies ist ein Beispiel für die Gültigkeit der Gödelschen Unvollständigkeitssätze.

Satz 2.6.4. 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 reeller Zahlen im Dezimalsystem zur Darstellung reeller Zahlen im Binärsystem. Jedes $r \in (0, 1)$ lässt sich im Binärsystem in der Form $0, d_1 d_2 d_3 \dots$ mit $d_k \in \{0, 1\}$ für alle $k \in \mathbb{N}$ schreiben. Explizit:

$$r = \sum_{k=1}^{\infty} d_k 2^{-k} = d_1 \cdot \frac{1}{2} + d_2 \cdot \frac{1}{4} + d_3 \cdot \frac{1}{8} + d_4 \cdot \frac{1}{16} + \dots$$

Eine im Binärsystem dargestellte Zahl ist z. B. $0,00101100001\dots$. 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 Folgenglieder hinzu). Es gibt also eine 1:1-Abbildung zwischen $\mathbb{R}$ und der Menge der 0-1-Folgen. $\square$

Bemerkung 2.6.4 (Exkurs zum Binärsystem). Betrachten wir zuerst $\mathbb{N}$. Heutzutage werden natürliche Zahlen $n \in \mathbb{N}$ im Dezimalsystem dargestellt, d.h. in der Form $c_m \dots c_3 c_2 c_1 c_0$ mit $c_i \in$

$\{0, \dots, 9\}$ 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 $m \in \mathbb{N}, c_i \in \{0, \dots, 9\}$.

Dieselben Zahlen können im Binärsystem dargestellt werden in der Form $d_l \dots d_3 d_2 d_1 d_0$ 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 für die Zahlen:

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

Im Unterschied zu natürlichen Zahlen gehen bei reellen Zahlen in der obigen Darstellung als Summe auch negativen Indizes ein, 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 $n \rightarrow \infty$.

3 Gruppen, Ringe, Körper

3.1 Gruppen

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

Bemerkung 3.1.1. Verknüpfungen sind etwa die Addition „+“ und Multiplikation „ $\cdot$ “ 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

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) := (a_1 + b_1, \dots, a_n + b_n)$$

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

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 $M = \{1, \dots, n\}$. Man nennt die Menge der bijektiven Abbildungen auf M , i.e. $S_n := \{\phi : M \rightarrow M \text{ bijektiv}\}$, zusammen mit der Funktionsverknüpfung $\circ$ die *Permutationsgruppe*. Die bijektiven Abbildungen $\phi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}, i \rightarrow \phi(i)$ 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 $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}, i \rightarrow \pi(i)$ schreiben wir:

$$\left\{ \begin{array}{ccccc} 1 & 2 & 3 & \dots & n \\ \pi(1) & \pi(2) & \pi(3) & \dots & \pi(n) \end{array} \right\}$$

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

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

$$\rho : \left\{ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{array} \right\}, \quad \sigma : \left\{ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \end{array} \right\}$$

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

$$\rho \circ \sigma : \left\{ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{array} \right\}, \quad \sigma \circ \rho : \left\{ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{array} \right\}$$

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): $\mathbb{Z}_m = \{[0]_m, [1]_m, \dots, [m-1]_m\}$, d.h. $\mathbb{Z}_m$ hat m Elemente. Dabei ist $[a]_m = a + m \cdot \mathbb{Z} = \{a + m \cdot k \mid k \in \mathbb{Z}\}$.

Zuerst zeigen wir, dass $[a]_m + [b]_m = [a + b]_m$ tatsächlich wieder auf $\mathbb{Z}_m$ abbildet, d. h. dass $+: \mathbb{Z}_m \times \mathbb{Z}_m \rightarrow \mathbb{Z}_m$. 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 $[a + b]_m = [a + km + b + lm]_m \Leftrightarrow m \mid (a + km + b + lm - (a + b)) \Leftrightarrow m \mid (km + lm)$ 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. Im folgenden Beispiel schreiben wir $\bar{a} := [a]_m$, wobei $a \in \{0, 1, \dots, m-1\}$ und $[a]_m \in \mathbb{Z}_m$.

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

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

Definition 3.1.4 (Homomorphismus, Isomorphismus, Automorphismus). Seien $(G, \circ)$ und $(H, \star)$ Gruppen. Eine Abbildung $\phi: G \rightarrow H$ 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 ϕ Homomorphismus und ϕ bijektiv ist. In dem Fall nennen wir $(G, \circ)$ und $(H, \star)$ isomorphe Gruppen.
- 3) Ist $\phi : G \rightarrow G$ ein Isomorphismus einer Gruppe G auf sich selbst, dann nennen wir ϕ einen *Automorphismus*.

Beispiel 3.1.4. $(\mathbb{R}, +)$ und $(\mathbb{R}^+, \cdot)$ sind isomorph, denn es existiert eine bijektive Abbildung $\phi : \mathbb{R} \rightarrow \mathbb{R}^+, a \rightarrow e^a$ 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 3.1.5. $(\mathbb{Z}_4, +)$ und $(\mathbb{Z}_5 \setminus \{\bar{0}\}, \cdot)$ sind isomorph, denn es existiert eine bijektive Abbildung

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

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.

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 $\text{id} : x \rightarrow x$. 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 $z = x + iy \rightarrow \bar{z} = x - iy$.

3.2 Ringe

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

- i) $+: R \times R \rightarrow R, (a, b) \rightarrow a + b$ („Addition“)
- ii) $\cdot: R \times R \rightarrow R, (a, b) \rightarrow a \cdot b$ („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. $\forall a, b, c \in R : (a + b) \cdot c = (a \cdot c) + (b \cdot c)$ und $c \cdot (a + b) = (c \cdot a) + (c \cdot b)$

$(R, +, \cdot)$ 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 wir 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. Beweis der geometrischen Reihe $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ für $|q| < 1$, $q \in \mathbb{R}$).

Bemerkung 3.3.2 (Gauloisfeld). 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.

Bemerkung 3.3.3 (Komplexe Zahlen). 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 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).

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

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

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

wird als *Körper der komplexen Zahlen* bezeichnet.

Bemerkung 3.3.4. $(\mathbb{C}, +, \cdot)$ erfüllt 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 und den oben definierten Vorschriften der Addition und Multiplikation können wir wie folgt rechnen:

$$i \cdot i = (0, 1) \cdot (0, 1) = (-1, 0),$$

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

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

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

Demnach ist $z \in \mathbb{C}$ genau dann, wenn $z := x + iy$ 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 $i \cdot i = -1$ und wenn wir dies berücksichtigen, können wir bei der Addition und Multiplikation jetzt wie folgt „rechnen“:

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2),$$

$$z_1 \cdot z_2 = (x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1 \cdot x_2 - y_1 \cdot y_2) + i(x_1 \cdot y_2 + y_1 \cdot x_2).$$

Definition 3.3.4 (Real- und Imaginärteil, Betrag, komplex konjugierte Zahl). Sei $z = x + iy \in \mathbb{C}$. Wir setzen:

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

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 ϕ :

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

Hier haben wir im letzten Schritt die Eulersche Formel

$$e^{i\phi} = \cos \phi + i \sin \phi$$

verwendet (Beweis über Reihenentwicklung von $\cos$, $\sin$ und $\exp$, siehe *Analysis*).

Wir können nun noch einige wichtige Gleichungen herleiten.

Satz 3.3.1. 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):

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{(x_1 + iy_1) + (x_2 + iy_2)} = \overline{(x_1 + x_2) + i(y_1 + y_2)} = (x_1 + x_2) - i(y_1 + y_2) \\ &= (x_1 - iy_1) + (x_2 - iy_2) = \bar{z}_1 + \bar{z}_2 \end{aligned}$$

Zu 2):

$$\begin{aligned}\overline{z_1 \cdot z_2} &= \overline{(x_1 + iy_1) \cdot (x_2 + iy_2)} = \overline{(x_1 \cdot x_2 - y_1 \cdot y_2) + i(x_1 \cdot y_2 + y_1 \cdot x_2)} \\ &= (x_1 \cdot x_2 - y_1 \cdot y_2) - i(x_1 \cdot y_2 + y_1 \cdot x_2) = x_1(x_2 - iy_2) - iy_1(x_2 - iy_2) \\ &= (x_1 - iy_1) \cdot (x_2 - iy_2) = \bar{z}_1 \cdot \bar{z}_2\end{aligned}$$

Zu 3):

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

Zu 4):

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

□

Wir kommen nochmal zurück zu Körpern im Allgemeinen.

Satz 3.3.2. *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 3.3.7 (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) $+: V \times V \rightarrow V, (v, w) \rightarrow v + w$ („Vektoraddition“)
- ii) $\cdot: K \times V \rightarrow V, (\lambda, v) \rightarrow \lambda \cdot v$ („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: $K = \mathbb{R}$ oder $K = \mathbb{C}$. (Wenn wir wissen, welcher Körper K gemeint ist, lassen wir den Bezug auf K manchmal auch ganz weg).

Bemerkung 4.1.2. Ohne die letzte Bedingung in der Vektorraumdefinition (Bedingung 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. Wir brauchen also auch die letzte Bedingung.

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

- i) $(v_1, \dots, v_n) + (w_1, \dots, w_n) := (v_1 + w_1, \dots, v_n + w_n)$
- ii) $\lambda \cdot (v_1, \dots, v_n) := (\lambda \cdot v_1, \dots, \lambda \cdot v_n)$

ein K -Vektorraum.

Beweis.

Zu 1): $(K^n, +)$ ist kommutative Gruppe, weil $(K, +)$ kommutative Gruppe. Neutrales Element ist der *Nullvektor* $\mathbf{0} := (0, \dots, 0)$. Inverses Element zu $(v_1, \dots, v_n) \in K^n$ ist: $(-v_1, \dots, -v_n)$. 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):

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

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. □

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

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 für komplexe Zahlen 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 euklid. Ebene $\mathbb{R}^2$ oder des euklid. Raums $\mathbb{R}^3$) kommt zu der Vektorraumstruktur, die wir 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.

Auch Mengen von Funktionen oder 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 $\text{Abb}(M, V) := \{f : M \rightarrow V\}$ von M nach V mit

- i) $f + g : M \rightarrow V, \quad (f + g)(x) := f(x) + g(x)$
- ii) $\lambda f : M \rightarrow V, \quad (\lambda f)(x) := \lambda f(x)$

ist ein K -Vektorraum.

Beweis.

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

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 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 $\text{Abb}(K^n, K^m) = \{f : K^n \rightarrow K^m\}$, ist ein K -Vektorraum.

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

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

- Die Menge aller *reellen Folgen* $\text{Abb}(\mathbb{N}, \mathbb{R}) := \{(a_n)_{n \in \mathbb{N}} = (a_1, a_2, \dots) \mid a_n \in \mathbb{R} \forall n \in \mathbb{N}\}$ 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 kleiner oder gleich n ,

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

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

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

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

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

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

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

$$\text{Hom}(V, W) := \{f : V \rightarrow W \mid f \text{ linear}\}.$$

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

Ist f zusätzlich bijektiv, dann heißt $f : V \rightarrow W$ *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 zusammenfassen als *eine* Bedingung, die besagt, dass f linear ist genau dann, wenn:

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

Beispiel 4.2.1.

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

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

- $h : \mathbb{R}^2 \rightarrow \mathbb{R}^3, (x, y) \mapsto (x + y, y, x)$ 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 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 : $\text{Hom}(V, K) = \mathcal{L}(V, K)$. Wie nennt man lineare Abbildungen in dem Fall?

Definition 4.2.2 (Linearform, Dualraum). Sei V ein K -Vektorraum und $f : V \rightarrow K$ 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. $\text{Hom}(V, W)$ ist mit der auf der Menge $\text{Abb}(V, W)$ definierten Addition und skalaren Multiplikation ebenfalls ein K -Vektorraum. Insbesondere sind also für $f, g \in \text{Hom}(V, W)$ und $\lambda \in K$ auch

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

lineare Abbildungen.

Bemerkung 4.2.4 (Linearität). Seien V und W K -Vektorräume. Sei $f : V \rightarrow W$ eine lineare Abbildung. Seien $v_1, \dots, v_n \in V$ und $\lambda_1, \dots, \lambda_n \in K$. 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 $f : V \rightarrow W$ und $g : U \rightarrow V$ lineare Abbildungen. Dann ist die Funktionsverkettung $f \circ g : U \rightarrow W$ auch eine lineare Abbildung.

Beweis. Übung. □

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

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

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

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

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

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:

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

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

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

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

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 $U = \{\lambda v \mid \lambda \in K\}$ 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.

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

Eine alternative Beschreibung dieses Untervektorraums ist:

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

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

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

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 $U_1 = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$ und $U_2 = \left\{ \mu \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid \mu \in \mathbb{R} \right\}$. 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 $U_1, U_2, \dots, U_n$ mit $n \in \mathbb{N}$ Untervektorräume von V . Dann ist die Summe*

$$U_1 + U_2 + \dots + U_n := \{u_1 + u_2 + \dots + u_n \mid u_1 \in U_1, u_2 \in U_2, \dots, u_n \in U_n\}$$

ein Untervektorraum von V .

Beweis. Analog zum Schnitt (siehe Satz 4.3.1). □

Bemerkung 4.3.2. Sind die U_i, U_j paarweise disjunkt bis auf den Nullvektor, d.h. ist $U_i \cap U_j = \{0\}$ für alle $i, j = 1, \dots, n$, dann nennt man $U_1 + \dots + U_n$ die *direkte* (innere) Summe.

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

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

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 $f : V \rightarrow W$ 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 $v_1, v_2 \in \text{Kern } f$. 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 $v_1 + v_2 \in \text{Kern } f$.

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

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

Angenommen $w_1, w_2 \in \text{Bild } f$. 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 $w_1 + w_2 \in \text{Bild } f$.

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

□

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

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

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 $i = 1, \dots, m$ und $j = 1, \dots, n$. Das rechteckige Schema mit m Zeilen und n Spalten

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

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

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

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

Definition 4.4.2 (Matrizenaddition, skalare Multiplikation). Seien $m, n \in \mathbb{N}$. Sei K ein Körper.

Seien $A = (a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}, B = (b_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}} \in K^{m \times n}$, $\lambda \in K$. Wir definieren:

- 1) $A + B := (a_{ij} + b_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}$ („Matrizenaddition“)
- 2) $\lambda A := (\lambda a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}$ („skalare Multiplikation“)

Satz 4.4.1. Die Menge $K^{m \times n}$ mit der angegebenen 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 $\{1, \dots, m\} \times \{1, \dots, n\} \rightarrow K$, $(i, j) \rightarrow a_{ij}$ aufgefasst werden kann. Wir können also $K^{m \times n}$ mit der Menge $\text{Abb}(\{1, \dots, m\} \times \{1, \dots, n\}, K)$ identifizieren. Für Letztere haben wir bereits bewiesen, dass sie einen K -Vektorraum bilden (siehe Satz 4.1.2). $\square$

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

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

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

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

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

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

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:

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

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

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

Bemerkung 4.4.2 (Dyadisches Produkt). 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 $A = (a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}$, mit Einträgen $a_{ij} = x_i \cdot y_j$ und schreibt dafür:

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

Bemerkung 4.4.3 (Lineares GLS). Für $A = (a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}} \in K^{m \times n}$ und $x \in K^{n \times 1}$ ergibt sich

schließlich:

$$A \cdot x = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots \\ a_{21}x_1 + a_{22}x_2 + \dots \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

Der hier entstandene Vektor $A \cdot x$ fasst die eine Seite eines linearen Gleichungssystems zusammen (mit A gegeben, x gesucht). Setzt man ihn nun gleich einem gegebenen Vektor $y \in K^{m \times 1}$, also $A \cdot x = y$, so hat man das vollständige lineare Gleichungssystem dastehen (mit n gesuchten Variablen $x_1, x_2, \dots, x_n$ und m Gleichungen).

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

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

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

- Wir definieren die *Nullmatrix* $0 \in K^{m \times n}$ als

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

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

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

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

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

Bemerkung 4.4.4. Es ist $e_j = (\delta_{ij})_{i=1, \dots, n}$ und $(E_n)_{ij} = \delta_{ij}$.

Bemerkung 4.4.5 (Notation). Statt $A = (a_{ij})_{i=1, \dots, m, j=1, \dots, n}$, schreiben wir kurz einfach nur $A = (a_{ij})$. Die Klammer zeigt dabei an, dass wir 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}).

Satz 4.4.2 (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): Diese Eigenschaft lässt sich am leichtesten mit zwei Sätzen zeigen, die wir in Kürze beweisen. Demnach gilt: A, B und C sind darstellende Matrizen der Abbildungen $f : K^n \rightarrow K^m$, $g : K^p \rightarrow K^n$ und $h : K^q \rightarrow K^p$ (siehe Satz 4.4.5). 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 4.4.6), dass $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

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

Zu 3):

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

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

Zu 4): Analog.

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

Bemerkung 4.4.6 (Ring, 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 nullteilerfrei, 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 4.4.7. Die Zeilen von A^T sind also die Spalten von A und umgekehrt.

Beispiel 4.4.2. 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 $x^T = (1 \ 2) \in \mathbb{R}^{1 \times 2}$.

Satz 4.4.3 (Rechenregeln für transponierte Matrizen). Seien $m, n \in \mathbb{N}$. Sei K ein Körper. Dann

gilt:

- 1) $A \in K^{m \times n} : (A^T)^T = A$
- 2) $A, B \in K^{m \times n} : (A + B)^T = A^T + B^T$
- 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 $A \cdot B = (\sum_{j=1}^n a_{ij} b_{jk}) \in K^{m \times p}$ und $(A \cdot B)^T \in K^{p \times m}$ mit

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

□

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 4.4.8. 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 4.4.3. 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

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

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

Satz 4.4.4 (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. □

Der folgende wichtige Satz beweist nun, dass jede lineare Abbildung durch eine Matrix dargestellt wird. Daraus folgt, dass jedes lineare Gleichungssystem analysiert werden kann, indem wir die zugehörige Matrix analysieren.

Satz 4.4.5 (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 $f : K^n \rightarrow K^m$ existiert genau eine Matrix $A \in K^{m \times n}$, so dass*

$$f(x) = Ax$$

Insbesondere gilt $a_{ij} = f_i(e_j)$ mit $i = 1, \dots, m$, $j = 1, \dots, n$ 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 $f : K^n \rightarrow K^m$, $f(x) = Ax$ 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

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

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 $Ae_j = A'e_j \forall j = 1, \dots, n \Rightarrow a_j = a'_j \forall j = 1, \dots, n \Rightarrow a_{ij} = a'_{ij} \forall i = 1, \dots, m, j = 1, \dots, n \Rightarrow A = A'$

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

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

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

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

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. $f_i : K^n \rightarrow K$ ist eine lineare Abbildung. Daraus folgt, dass auch $f : K^n \rightarrow K^m$ 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 4.4.4. Betrachte die lineare Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ y \\ x+z \end{pmatrix}$.

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 $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 3x - \frac{1}{2}z \\ 0 \end{pmatrix}$.

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 4.4.6 (Darstellende Matrix der Verkettung linearer Abbildungen). *Seien $n, m, p \in \mathbb{N}$. Sei K ein Körper. Seien $f : K^n \rightarrow K^m$ und $g : K^p \rightarrow K^n$ 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 $f \circ g : K^p \rightarrow K^m$.*

Beweis.

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

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

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 $i = 1, \dots, m$, $j = 1, \dots, p$. 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 4.4.5 (Additionstheoreme der Trigonometrie). Man betrachte die 2×2 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_ϕ beschreibt Drehungen des Vektors r in mathematisch positiver Richtung (d.h. entgegen dem Uhrzeigersinn) um den Winkel ϕ 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 an dieser Stelle R_ϕ die lineare Abbildung bezeichnet (wir hätten auch eine andere Bezeichnung wählen können, z.B. f_{R_ϕ} statt R_ϕ , um den Unterschied zur darstellenden Matrix deutlicher zu machen). D.h. dreht man nacheinander um die Winkel ϕ und ψ , so ist es dasselbe, wie wenn man einmal um den Winkel $\phi + \psi$ dreht (oder um $\psi + \phi$, das Ganze ist symmetrisch in ψ, ϕ). 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:

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

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

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 durch folgendes Beispiel motivieren.

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 $e_1, \dots, e_n \in \mathbb{R}^n$.

2) Betrachte $V = \mathbb{R}^3$ 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 $\mathbb{R}^3$ (denn jeder Vektor $v \in V$ lässt sich durch *mehrere* 4-Tupel $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ spezifizieren) und sie sind auch *nicht linear unabhängig* (denn z.B. $a_4 = a_1 + a_2 + a_3$), aber sie bilden ein *Erzeugendensystem*, d.h. jeder Vektor $v \in \mathbb{R}^3$ lässt sich in Bezug auf a_1, a_2, a_3 und a_4 durch *mindestens ein* 4-Tupel $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ spezifizieren.

Definition 4.5.1 (Lineare Hülle, Erzeugendensystem). Sei $m \in \mathbb{N}$. Sei V ein K -Vektorraum und $b_1, \dots, b_m \in V$. Man bezeichnet die Menge der *Linearkombinationen* von $b_1, \dots, b_m$,

$$\text{span}(b_1, \dots, b_m) = \left\{ \sum_{i=1}^m \lambda_i b_i \mid \lambda_i \in K \quad \forall i = 1, \dots, m \right\},$$

als *lineare Hülle* oder *Spann* von $b_1, \dots, b_m$. Weiter heißt $b_1, \dots, b_m$ *Erzeugendensystem* von V genau dann, wenn

$$\text{span}(b_1, \dots, b_m) = V.$$

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 $b_1, \dots, b_m$ Erzeugendensystem von V ist, für jedes $v \in V$ *mindestens eine* Linearkombination, also mindestens ein $\lambda = (\lambda_1, \dots, \lambda_m) \in K^m$, existiert mit $v = \sum_{i=1}^m \lambda_i b_i$.

Lemma 4.5.1 (Kleinsten Untervektorraum). Sei $m \in \mathbb{N}$. Sei V ein K -Vektorraum und seien $b_1, \dots, b_m \in V$. Dann ist $\text{span}(b_1, \dots, b_m)$ der kleinste Untervektorraum, der $b_1, \dots, b_m$ enthält, das heißt, für jeden Untervektorraum $U \subseteq V$ mit $b_1, \dots, b_m \in U$ gilt:

$$\text{span}(b_1, \dots, b_m) \subseteq U.$$

Beweis. Zu zeigen: 1) $\text{span}(b_1, \dots, b_m)$ ist Untervektorraum und 2) $\text{span}(b_1, \dots, b_m)$ ist *kleinster* Untervektorraum.

Zu 1): Dass $\text{span}(b_1, \dots, b_m)$ 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 $b_1, \dots, b_m \in U$. Weil U Untervektorraum ist, gilt $\forall \lambda_1, \dots, \lambda_m \in K : \lambda_1 b_1 + \dots + \lambda_m b_m \in U \Rightarrow \text{span}(b_1, \dots, b_m) \subseteq U$. $\square$

Definition 4.5.2 (Lineare Unabhängigkeit). Sei V ein K -Vektorraum, $m \in \mathbb{N}$ und $b_1, \dots, b_m \in V$. Die Vektoren $b_1, \dots, b_m$ heißen *linear unabhängig* genau dann, wenn $\forall \lambda_1, \dots, \lambda_m \in K$ gilt:

$$\lambda_1 b_1 + \dots + \lambda_m b_m = 0 \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_m = 0.$$

Die Vektoren $b_1, \dots, b_m$ heißen *linear abhängig* genau dann, wenn sie *nicht* linear unabhängig sind.

Sei weiter $v \in V$. Der Vektor v heißt *linear abhängig von* $b_1, \dots, b_m$ genau dann, wenn:

$$v \in \text{span}(b_1, \dots, b_m).$$

Lemma 4.5.2 (Eindeutigkeit der Darstellung). Sei $m \in \mathbb{N}$, V ein K -Vektorraum und $b_1, \dots, b_m \in V$. Seien $\lambda := (\lambda_1, \dots, \lambda_m) \in K^m$ und $\mu := (\mu_1, \dots, \mu_m) \in K^m$. Die Vektoren $b_1, \dots, b_m$ sind linear unabhängig genau dann, wenn

$$\forall \lambda, \mu \in K^m : \sum_{i=1}^m \lambda_i b_i = \sum_{i=1}^m \mu_i b_i \Rightarrow \lambda_i = \mu_i \quad \forall i = 1, \dots, m.$$

Beweis. „ $\Rightarrow$ “: Falls die Vektoren linear unabhängig sind, gilt:

$$\sum_{i=1}^m \lambda_i b_i = \sum_{i=1}^m \mu_i b_i \Leftrightarrow \sum_{i=1}^m (\lambda_i - \mu_i) b_i = 0 \Rightarrow \lambda_i - \mu_i = 0 \quad \forall i = 1, \dots, m \Leftrightarrow \lambda_i = \mu_i \quad \forall i = 1, \dots, m$$

„ $\Leftarrow$ “: Wir wenden die Bedingung an auf den Nullvektor $\mu = 0 \in K^m$ (d.h. $\mu_1 = \dots = \mu_m = 0$). Also steht da $\sum_{i=1}^m \lambda_i b_i = 0 \Rightarrow \lambda_1 = \dots = \lambda_m = 0$, aber das ist genau die Aussage, dass die $b_1, \dots, b_m$ linear unabhängig sind. $\square$

Bemerkung 4.5.2 (Lineare Unabhängigkeit). Es gibt also für jedes $v \in V$ *höchstens eine* Darstellung als Linearkombination linear unabhängiger Vektoren $b_1, \dots, b_m \in V$, also höchstens ein m -Tupel $\lambda = (\lambda_1, \dots, \lambda_m) \in K^m$ 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 $b_1, \dots, b_m \in V$. Die Vektoren $b_1, \dots, b_m$ heißen *Basis* von V genau dann, wenn $b_1, \dots, b_m$ linear unabhängig sind und ein Erzeugendensystem von V bilden.

Beispiel 4.5.2.

- Die kanonischen Einheitsvektoren $e_1, \dots, e_n$ bilden eine Basis des $\mathbb{R}^n$.
- Die *Monome* $p_k = x^k$ mit $k = 0, \dots, n$ bilden eine Basis des Vektorraums der reellen Polynome vom Grad kleiner oder gleich n , $P_n = \{ \sum_{k=0}^n a_k x^k \mid a_k \in \mathbb{R} \forall k = 0, \dots, n \}$.

Bemerkung 4.5.3. Für den Vektorraum $V = \{0\}$ wird die leere Menge $\emptyset$ als Basis festgelegt.

Bemerkung 4.5.4. Wenn $b_1, \dots, b_m$ eine Basis von V ist, dann existiert also für jedes $v \in V$ genau ein m -Tupel $\lambda = (\lambda_1, \dots, \lambda_m)$ mit $v = \sum_{i=1}^m \lambda_i b_i$. (Siehe die Bemerkungen zu „Lineare Unabhängigkeit“ und „Erzeugendensystem“, wonach in dem einen Fall „höchstens 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 $v_1, \dots, v_{m+1} \in V$. Dann gilt:

- 1) $v \in V$ linear abhängig $\Leftrightarrow v = 0$
- 2) für $m \geq 2$:
 $v_1, \dots, v_m$ linear abhängig $\Leftrightarrow \exists k \in \{1, \dots, m\} : v_k \in \text{span}(v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_m)$
- 3) für $m \geq 1$:
 v_{m+1} linear abhängig von $v_1, \dots, v_m \Leftrightarrow \text{span}(v_1, \dots, v_{m+1}) = \text{span}(v_1, \dots, v_m)$
- 4) $v_1, \dots, v_m$ linear unabhängig und $v_{m+1} \notin \text{span}(v_1, \dots, v_m) \Leftrightarrow v_1, \dots, v_{m+1}$ linear unabhängig

Beweis. Wir wissen: $v_1, \dots, v_m \in V$ sind linear abhängig $\Leftrightarrow \exists \lambda_1, \dots, \lambda_m \in K$ und $k \in \{1, \dots, m\} : \sum_{i=1}^m \lambda_i v_i = 0 \wedge \lambda_k \neq 0$.

Zu 1): „ $\Rightarrow$ “: $v \in V$ linear abhängig, d.h. $\exists \lambda \in K : \lambda v = 0 \wedge \lambda \neq 0 \Rightarrow v = 0$.

„ $\Leftarrow$ “: $v = 0$, d.h. $1 \cdot v = 0 \wedge 1 \neq 0 \Rightarrow v$ linear abhängig.

Zu 2): „ $\Rightarrow$ “: $v_1, \dots, v_m$ linear abhängig $\Rightarrow \exists \lambda_1, \dots, \lambda_m \in K$ und $k \in \{1, \dots, m\} : \sum_{i=1, i \neq k}^m \lambda_i v_i + \lambda_k v_k = 0$ mit $\lambda_k \neq 0 \Rightarrow v_k = -\frac{1}{\lambda_k} \sum_{i=1, i \neq k}^m \lambda_i v_i \Rightarrow v_k \in \text{span}(v_1, \dots, v_{k-1}, \hat{v}_k, v_{k+1}, \dots, v_m)$. Hier bedeutet „ $\hat{v}_k$ “, dass v_k nicht enthalten ist.

„ $\Leftarrow$ “: $v_k \in \text{span}(v_1, \dots, v_{k-1}, \hat{v}_k, v_{k+1}, \dots, v_m) \Rightarrow v_k = \sum_{i=1, i \neq k}^m \lambda_i v_i$. Wähle nun $\lambda_k = -1$, dann ist $\sum_{i=1}^m \lambda_i v_i = 0$ und damit $v_1, \dots, v_m$ linear abhängig.

Zu 3): „ $\Rightarrow$ “: Sei v_{m+1} linear abhängig von $v_1, \dots, v_m$. Zu zg.: $\text{span}(v_1, \dots, v_{m+1}) \subseteq \text{span}(v_1, \dots, v_m)$ und $\text{span}(v_1, \dots, v_{m+1}) \supseteq \text{span}(v_1, \dots, v_m)$. Dabei ist die zweite Inklusion klar und für die erste Inklusion gilt: $v_{m+1} = \sum_{i=1}^m \lambda_i v_i$, weil $v_{m+1} \in \text{span}(v_1, \dots, v_m)$, und damit gilt für beliebiges $v \in \text{span}(v_1, \dots, v_{m+1})$:

$$\begin{aligned} v &= \sum_{i=1}^m \mu_i v_i + \mu_{m+1} v_{m+1} = \sum_{i=1}^m \mu_i v_i + \mu_{m+1} \sum_{i=1}^m \lambda_i v_i \\ &= \sum_{i=1}^m (\mu_i + \mu_{m+1} \lambda_i) v_i = \sum_{i=1}^m \kappa_i v_i \end{aligned}$$

mit $\kappa_i = \mu_i + \mu_{m+1}\lambda_i \in K$ für alle $i = 1, \dots, m$, also $v \in \text{span}(v_1, \dots, v_m)$.

„ $\Leftarrow$ “: Offensichtlich ist $v_{m+1} \in \text{span}(v_1, \dots, v_{m+1})$ und $\text{span}(v_1, \dots, v_{m+1}) = \text{span}(v_1, \dots, v_m) \Rightarrow v_{m+1} \in \text{span}(v_1, \dots, v_m)$, also v_{m+1} linear abhängig von $v_1, \dots, v_m$.

Zu 4): „ $\Leftarrow$ “: Offensichtlich.

„ $\Rightarrow$ “: Sei $\sum_{i=1}^{m+1} \lambda_i v_i = 0$. Z. zg.: $\lambda_i = 0$ für alle $i = 1, \dots, m+1$.

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 $v_{m+1} \in \text{span}(v_1, \dots, v_m) \Rightarrow$ Widerspruch

2. Fall: $\lambda_{m+1} = 0 \Rightarrow \sum_{i=1}^m \lambda_i v_i = 0$. Weil $v_1, \dots, v_m$ linear unabhängig, folgt daraus: $\lambda_i = 0$ für alle $i = 1, \dots, m$. Somit $\lambda_i = 0$ für alle $i = 1, \dots, m+1$.

□

Satz 4.5.1 (Basisauswahlsatz). *Sei $V \neq \{0\}$ ein endlich erzeugter K -Vektorraum, $m \in \mathbb{N}$ und $b_1, \dots, b_m$ ein Erzeugendensystem von V . Dann existiert ein $n \in \mathbb{N}$ mit $n \leq m$, so dass (nach geeigneter Umnummerierung) $b_1, \dots, b_n$ 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 $b_1, \dots, b_m$ ein Erzeugendensystem von $V \neq \{0\}$.

Sei o.B.d.A. $b_1, \dots, b_m \neq 0$ (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 $b_1, \dots, b_m = 0$ sein, da sonst $V = \{0\}$.)

Nun machen wir eine Fallunterscheidung:

1. Fall: $b_1, \dots, b_m$ sind linear unabhängig. Dann sind $b_1, \dots, b_m$ bereits Basis von V

2. Fall: $b_1, \dots, b_m$ sind linear abhängig. Nach Annahme sind $b_1, \dots, b_m \neq 0$. 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 $\exists p \in \{1, \dots, m\} : b_p \in \text{span}(b_1, \dots, b_{p-1}, b_{p+1}, \dots, b_m)$ (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 $b_1, \dots, b_{m-1}$. Weil aber $\text{span}(b_1, \dots, b_{m-1}) = \text{span}(b_1, \dots, b_m)$ (siehe Aussage 3 von Lemma 4.5.3), ist auch $b_1, \dots, b_{m-1}$ Erzeugendensystem von V . Diese Schritte können wir so oft wiederholen, bis wir zu einem Erzeugendensystem $b_1, \dots, b_n$ 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 $b_1, \dots, b_n$ eine Basis von V . Seien $a_1, \dots, a_p \in V$ linear unabhängig. Dann gilt:*

$$p \leq n$$

Beweis. Wir zeigen die Kontraposition der Aussage mittels vollständiger Induktion. Statt der

Aussage $a_1, \dots, a_p \in V$ linear unabhängig $\Rightarrow p \leq n$ zeigen wir also

$$p > n \quad \Rightarrow \quad a_1, \dots, a_p \in V \text{ linear abhängig.}$$

Induktionsanfang ($n = 1$): b_1 ist eine Basis von V , insbesondere wird also V von b_1 erzeugt. Für Vektoren $a_1, \dots, a_p \in V$, $p > 1$, gilt dann: $a_i = \alpha_i b_1$ mit $\alpha_i \in K \forall i = 1, \dots, p$. Falls eines der a_i der Nullvektor ist, so sind die $a_1, \dots, a_p$ linear abhängig (siehe Def. der Unabhängigkeit; wegen $a_i = 0$ folgt aus $\lambda_1 a_1 + \dots + \lambda_p a_p = 0$, dass λ_i beliebig und *nicht* $\lambda_i = 0$). Ansonsten ist z.B. $\alpha_2 a_1 + (-\alpha_1) a_2 = \alpha_2 \alpha_1 b_1 + (-\alpha_1) \alpha_2 b_1 = 0$ eine nichttriviale Linearkombination zu 0 und damit $a_1, \dots, a_p$ 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 $b_1, \dots, b_{n+1}$ Basis von V . Insbesondere wird also V von $b_1, \dots, b_{n+1}$ erzeugt. In $V = \text{span}(b_1, \dots, b_{n+1})$ seien Vektoren $a_1, \dots, a_p$ mit $p > n + 1$ gegeben. Das heißt, es gibt Zahlen $\alpha_{jk} \in K$, so dass $\forall j = 1, \dots, p$:

$$a_j = \sum_{k=1}^{n+1} \alpha_{jk} b_k.$$

Seien nicht alle $\alpha_{jk} = 0$ (sonst wären alle $a_j = 0$ und damit alle $a_1, \dots, a_p$ linear abhängig). Sei etwa $\alpha_{11} \neq 0$. Dann liegen die $p - 1$ Vektoren

$$c_j := a_j - \frac{\alpha_{j1}}{\alpha_{11}} a_1 = \sum_{k=1}^{n+1} \frac{\alpha_{11} \alpha_{jk} - \alpha_{j1} \alpha_{1k}}{\alpha_{11}} b_k \quad (j = 2, \dots, p)$$

im $\text{span}(b_2, \dots, b_{n+1})$, weil der Koeffizient von b_1 gleich 0 ist. Wegen $p - 1 > n$ sind nach Induktionsannahme, die $p - 1$ Vektoren $c_2, \dots, c_p$ linear abhängig. Also gibt es Zahlen $\lambda_2, \dots, \lambda_p \in K$, 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 $a_1, \dots, a_p$ 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 $b_1, \dots, b_m$ und $b'_1, \dots, b'_n$ 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 $b_1, \dots, b_m$ als Basis und $b'_1, \dots, b'_n$ als linear unabhängige Vektoren in V , woraus folgt, dass $n \leq m$. Das andere Mal betrachten wir $b'_1, \dots, b'_n$ als Basis und $b_1, \dots, b_m$ als linear unabhängige Vektoren in V , woraus folgt, dass $m \leq n$. Es folgt: $n = m$. $\square$

4.6 Dimension, Koordinatendarstellung

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 $e_1, \dots, e_n$ eine Basis).

Satz 4.6.1 (Basisergänzungssatz). Seien $n, p \in \mathbb{N}$. Sei V ein K -Vektorraum und $b_1, \dots, b_n$ eine Basis von V . Seien $a_1, \dots, a_p \in V$ linear unabhängig. Dann gibt es $a_{p+1}, \dots, a_n \in \{b_1, \dots, b_n\}$, so dass $a_1, \dots, a_n$ 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 $b_1, \dots, b_n$ (statt einer Basis $b_1, \dots, b_n$). Auch Erzeugendensysteme $b_1, \dots, b_n$ kann man also nutzen, um eine Menge linear unabhängiger Vektoren $a_1, \dots, a_p$ zu einer Basis zu ergänzen.

Beweis (von Satz 4.6.1). Wir können zwei Fälle unterscheiden.

1) $\forall i = 1, \dots, n: b_i \in \text{span}(a_1, \dots, a_p)$. Dann ist $\text{span}(b_1, \dots, b_n) \subseteq \text{span}(a_1, \dots, a_p)$ und damit $a_1, \dots, a_p$ Erzeugendensystem von V und $a_1, \dots, a_p$ linear unabhängig, also $a_1, \dots, a_p$ Basis. Weil jede Basis dieselbe Anzahl an Vektoren hat (Satz 4.5.2), gilt in diesem Fall: $p = n$.

2) $\exists i \in 1, \dots, n: b_i \notin \text{span}(a_1, \dots, a_p)$. Wir setzen $a_{p+1} = b_i$. Nach Aussage 4) von Lemma 4.5.3 sind $a_1, \dots, a_p, a_{p+1}$ linear unabhängig und insbesondere ist $b_i \in \text{span}(a_1, \dots, a_{p+1})$. Dieses Vorgehen können wir unter Hinzunahme geeigneter Basisvektoren wiederholen, bis wir eine Menge von Vektoren $a_{p+1}, \dots, a_{p+m}$ mit $m \in \mathbb{N} (m < n)$ gefunden haben, so dass $a_1, \dots, a_{p+m}$ linear unabhängig und $\forall i = 1, \dots, n: b_i \in \text{span}(a_1, \dots, a_{p+m})$. Dann ist $\text{span}(b_1, \dots, b_n) \subseteq \text{span}(a_1, \dots, a_{p+m})$, also $a_1, \dots, a_{p+m}$ Erzeugendensystem von V und $a_1, \dots, a_{p+m}$ linear unabhängig, also $a_1, \dots, a_{p+m}$ Basis. Insbesondere ist in diesem Fall $p + m = n$, also $p < n$. $\square$

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 $b_1, \dots, b_m \in V$. Dann gilt:

$$b_1, \dots, b_m \text{ Basis von } V \Rightarrow \exists f : K^m \rightarrow V \text{ linear und bijektiv, } f(e_i) = b_i \forall i = 1, \dots, m.$$

Beweis. Wir definieren $f : K^m \rightarrow V, f(x) := \sum_{i=1}^m x_i b_i$ 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 $f(x) = 0 \Leftrightarrow \sum_{i=1}^m x_i b_i = 0 \Rightarrow x_i = 0 \forall i = 1, \dots, m \Leftrightarrow x = 0$, wobei der vorletzte Schritt folgt, weil die b_i unabhängig sind. Somit ist $\text{Kern } f = \{0\}$ und damit f injektiv (siehe Übung). Zudem ist f surjektiv, weil $b_1, \dots, b_m$ ein Erzeugendensystem von V ist $\Rightarrow f$ ist bijektiv.

Zuletzt gilt nach Definition von f : $f(e_i) = b_i \forall i = 1, \dots, m$. □

Lemma 4.6.2 (Umkehrfunktion). *Seien V und W K -Vektorräume. Sei $f : V \rightarrow W$ linear und bijektiv. Dann ist $f^{-1} : W \rightarrow V$ 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 $f : V \rightarrow W$ gibt. Weil $\dim V = n$, existiert eine Basis $b_1, \dots, b_n$ von V (Satz 4.5.2). Damit existiert eine lineare und bijektive Funktion $f : K^n \rightarrow V$ (siehe Lemma 4.6.1) und somit eine lineare und bijektive Umkehrfunktion $f^{-1} : V \rightarrow K^n$ (siehe Lemma 4.6.2), d.h. jeder K -Vektorraum V mit $\dim V = n$ ist isomorph zu K^n . □

Wir haben ja bereits festgestellt, dass ein Vektorraum verschiedene Basen besitzen kann. Die Wahl der Basis verändert aber einen gegebenen Vektor $v \in V$ nicht, lediglich seine Darstellung. Für eine beliebige Basis $\mathcal{B}$ kann man nun, mithilfe des eben betrachteten Isomorphismus von V zu K^n , die Koordinatendarstellung des Vektors $v \in V$ bezüglich der Basis $\mathcal{B}$ angeben.

Definition 4.6.2 (Koordinatendarstellung). Sei V ein K -Vektorraum und $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis von V . Man nennt den Isomorphismus

$$\phi_{\mathcal{B}}^{-1} : V \rightarrow K^n, \quad \phi_{\mathcal{B}}^{-1}(v) = \phi_{\mathcal{B}}^{-1}\left(\sum_{j=1}^n v_j b_j\right) := \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

die *Koordinatendarstellung* von v bezüglich der Basis $\mathcal{B}$.

Bemerkung 4.6.4. Für die (wieder lineare, bijektive) Umkehrfunktion $\phi_{\mathcal{B}}$ zu $\phi_{\mathcal{B}}^{-1} : V \rightarrow K^n$ gilt:

$$\phi_{\mathcal{B}} : K^n \rightarrow V, \quad \phi_{\mathcal{B}}((v_1, \dots, v_n)) = \sum_{j=1}^n v_j b_j$$

und damit insbesondere: $\phi_{\mathcal{B}}(e_j) = b_j$.

Bemerkung 4.6.5. Wir unterscheiden an dieser Stelle in der Schreibweise nicht zwischen n -Tupeln und Spaltenvektoren.

Bemerkung 4.6.6. Bislang haben wir nicht zwischen verschiedenen Basen unterschieden. Erst mit unterschiedlichen Basen wird die Unterscheidung (und damit die kompliziertere Notation) nötig.

Wir können nun einen Basiswechsel beschreiben.

Satz 4.6.3 (Basiswechsel). Sei V ein K -Vektorraum und seien $\mathcal{A}$ und $\mathcal{B}$ Basen von V . Dann gibt es eine invertierbare Matrix $M_{\mathcal{B}}^{\mathcal{A}}(id_V)$, die den Basiswechsel $\mathcal{A} \rightarrow \mathcal{B}$ beschreibt, die also die Koordinatendarstellungen $\phi_{\mathcal{B}}^{-1}(v)$ und $\phi_{\mathcal{A}}^{-1}(v)$ eines Vektors $v \in V$ ineinander überführt:

$$\phi_{\mathcal{B}}^{-1}(v) = M_{\mathcal{B}}^{\mathcal{A}}(id_V) \cdot \phi_{\mathcal{A}}^{-1}(v) \quad \text{sowie} \quad \phi_{\mathcal{A}}^{-1}(v) = M_{\mathcal{A}}^{\mathcal{B}}(id_V) \cdot \phi_{\mathcal{B}}^{-1}(v)$$

Dabei gilt: $M_{\mathcal{A}}^{\mathcal{B}}(id_V) = \left(M_{\mathcal{B}}^{\mathcal{A}}(id_V)\right)^{-1}$.

Beweis.

Zur Existenz: Es gibt einen Isomorphismus zwischen K^n und V und eine Umkehrfunktion von V nach K^n , die selbst wieder ein Isomorphismus ist. Damit gibt es auch einen Isomorphismus (= Verkettung von Isomorphismen) zwischen den Koordinatendarstellungen von v bzgl. $\mathcal{A}$ und $\mathcal{B}$ mit eindeutiger darstellender Matrix $M_{\mathcal{B}}^{\mathcal{A}}(id_V)$.

Zur Eindeutigkeit von $M_{\mathcal{A}}^{\mathcal{B}}(id_V)$: Die darstellende Matrix der Umkehrabbildung ist wegen der Eindeutigkeit des Inversen eindeutig als die inverse Matrix bestimmt. $\square$

Beispiel 4.6.2 (Basiswechsel). Betrachte $V = \mathbb{R}^2$. Seien

$$\mathcal{A} = \left\{ a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, a_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \quad \text{und} \quad \mathcal{B} = \left\{ b_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, b_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

zwei Basen des $\mathbb{R}^2$. Betrachte den Vektor $v \in V$ gegeben durch $v = 3b_1 + 5b_2$.

Dann ist $v = 5a_1 - 2a_2$ und damit

$$\phi_{\mathcal{B}}^{-1}(v) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \quad \phi_{\mathcal{A}}^{-1}(v) = \begin{pmatrix} 5 \\ -2 \end{pmatrix}, \quad M_{\mathcal{B}}^{\mathcal{A}}(id_V) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad M_{\mathcal{A}}^{\mathcal{B}}(id_V) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

und wir rechnen nach:

$$M_{\mathcal{B}}^{\mathcal{A}}(id_V) \phi_{\mathcal{A}}^{-1}(v) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix},$$

$$M_{\mathcal{A}}^{\mathcal{B}}(id_V) \phi_{\mathcal{B}}^{-1}(v) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}.$$

Dabei ist die eine Matrix die Inverse der anderen, denn es gilt:

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

Insbesondere gilt:

$$id_V(a_1) = 1b_1 + 1b_2, \quad id_V(a_2) = 1b_1 + 0b_2,$$

$$id_V(b_1) = 0a_1 + 1a_2, \quad id_V(b_2) = 1a_1 - 1a_2.$$

Wir können also an den Koordinaten, die in diesen Gleichungen auftauchen, die Spaltenvektoren der Matrizen $M_{\mathcal{B}}^{\mathcal{A}}(id_V)$ und $M_{\mathcal{A}}^{\mathcal{B}}(id_V)$ ablesen!

Definition 4.6.3 (Koordinatendarstellung linearer Abbildung). Seien V, W K -Vektorräume, $f : V \rightarrow W$ eine lineare Abbildung (d.h. $f \in \text{Hom}(V, W)$), $\mathcal{A}$ eine Basis von V und $\mathcal{B}$ eine Basis von W . Sei $v \in V$. Wir nennen $M_{\mathcal{B}}^{\mathcal{A}}(f)$ die *Koordinatendarstellung* von f bezüglich der Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W und schreiben:

$$\phi_{\mathcal{B}}^{-1}(f(v)) = M_{\mathcal{B}}^{\mathcal{A}}(f) \phi_{\mathcal{A}}^{-1}(v)$$

Bemerkung 4.6.7. Statt $f(v) = Av$ schreiben wir nun $\phi_{\mathcal{B}}^{-1}(f(v)) = M_{\mathcal{B}}^{\mathcal{A}}(f) \phi_{\mathcal{A}}^{-1}(v)$. Dies ist zwar komplizierter, aber aussagekräftiger. Nur so können wir zwischen den Koordinatendarstellungen (= darstellenden Matrizen) einer linearen Abbildung bezüglich unterschiedlicher Basen unterscheiden.

Definition 4.6.4 (Ähnliche Matrizen). Zwei Matrizen $A, B \in K^{n \times n}$ heißen *ähnlich* genau dann, wenn eine invertierbare Matrix $S \in K^{n \times n}$ existiert mit:

$$A = S^{-1}BS$$

Bemerkung 4.6.8 (Reguläre Matrix). Man nennt eine invertierbare Matrix auch *regulär*.

Bemerkung 4.6.9 (Basiswechsel). Sei V ein K -Vektorraum. Wir können ähnliche Matrizen

A, B als Matrixdarstellungen derselben linearen Abbildung $f : V \rightarrow V$ bezüglich verschiedener Basen $\mathcal{A}, \mathcal{B}$ auffassen. Seien $A := M_{\mathcal{A}}^{\mathcal{A}}(f)$ und $B := M_{\mathcal{B}}^{\mathcal{B}}(f)$ die Matrixdarstellungen von f bezüglich der Basen $\mathcal{A}$ bzw. $\mathcal{B}$. Dann ist:

$$B = M_{\mathcal{B}}^{\mathcal{B}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(id_V \circ f \circ id_V) = M_{\mathcal{B}}^{\mathcal{A}}(id_V) M_{\mathcal{A}}^{\mathcal{A}}(f) M_{\mathcal{A}}^{\mathcal{B}}(id_V) =: S^{-1} A S.$$

Dabei haben wir benutzt, dass die darstellende Matrix einer Verkettung linearer Abbildungen gerade das Produkt der darstellenden Matrizen dieser Abbildungen ist. Wir werden das Schema $A = S^{-1} D S$ wiedersehen, wenn wir Matrizen diagonalisieren. Die Diagonalmatrix D ist in dem Fall die Darstellung der Abbildung bezüglich der Basis aus Eigenvektoren.

Wir zeigen noch einen wichtigen Satz zur Dimension von Kern und Bild einer linearen Abbildung $f : V \rightarrow W$. Dabei sind sowohl $\text{Kern} f$ als auch $\text{Bild} f$ Untervektorräume (von V bzw. W).

Lemma 4.6.3. *Sei V ein endlich erzeugter K -Vektorraum und $U \subseteq V$ ein Untervektorraum 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 Untervektorraum 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 $b_1, \dots, b_n$, d.h. es existiert ein $n \in \mathbb{N}$ mit $\dim V = n$. Für jede Menge linear unabhängiger Vektoren $u_1, \dots, u_p \in U \subseteq V$ gilt dann (Lemma 4.5.4): $p \leq n = \dim V$. Wegen $U \neq \{0\}$ existieren linear unabhängige $u_1, \dots, u_p$ mit $1 \leq p \leq n$ und insbesondere gibt es demnach eine maximale Anzahl $p_{\max}$ linear unabhängiger Vektoren: $u_1, \dots, u_{p_{\max}}$.

Wir behaupten nun: $\forall u \in U : u \in \text{span}(u_1, \dots, u_{p_{\max}})$. Anderenfalls gäbe es ein $u' \notin \text{span}(u_1, \dots, u_{p_{\max}})$ und damit $u_1, \dots, u_{p_{\max}}, u'$ linear unabhängig (siehe Lemma 4.5.3) im Widerspruch zur Annahme, dass $p_{\max}$ maximal ist. Folglich ist also $u_1, \dots, u_{p_{\max}}$ ein Erzeugendensystem von U (d.h. insbesondere: U ist endlich erzeugt) und $\dim U = p_{\max} \leq n = \dim V$. $\square$

Satz 4.6.4 (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. $\square$

Satz 4.6.5 (Dimensionsformel lineare Abbildungen). *Seien V, W K -Vektorräume, $f : V \rightarrow W$ 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 $b_1, \dots, b_m$ die Basis von $\text{Kern } f$ (der Fall $m = 0$ bezeichne die leere Basis). Eine Basis existiert, weil $\text{Kern } f$ Untervektorraum von V und damit selbst ein Vektorraum ist (siehe Satz 4.3.4). Weil $\text{Kern } f \subseteq V$: $\dim \text{Kern } f = m \leq n = \dim V$.

Weil sie die Basis von $\text{Kern } f$ bilden, sind die $b_1, \dots, b_m$ linear unabhängige Vektoren aus V und so können wir sie laut Basisergänzungssatz durch Hinzunahme von $n - m$ Basisvektoren $b_{m+1}, \dots, b_n$ zu einer Basis $b_1, \dots, b_n$ von V ergänzen. Bleibt zu zeigen, dass $f(b_{m+1}), \dots, f(b_n)$ eine Basis von $\text{Bild } f$ bilden (dann ist $n - m = \dim \text{Bild } f$ und damit: $\dim \text{Kern } f + \dim \text{Bild } f = m + (n - m) = n = \dim V$).

Wir müssen also zeigen: $f(b_{m+1}), \dots, f(b_n)$ Erzeugendensystem vom $\text{Bild } f$ und $f(b_{m+1}), \dots, f(b_n)$ linear unabhängig.

1) $f(b_{m+1}), \dots, f(b_n)$ Erzeugendensystem: Zunächst gilt: $f(b_1), \dots, f(b_n)$ Erzeugendensystem von $\text{Bild } f$ (denn für jedes $v \in V$: $\sum_{i=1}^n \lambda_i b_i = v$ und damit gilt für jedes $w \in \text{Bild } f$: $w = f(v) = f(\sum_i \lambda_i b_i) = \sum_i \lambda_i f(b_i)$, weil f linear). Nun ist aber $f(b_j) = 0 \forall j = 1, \dots, m$ (weil $b_1, \dots, b_m \in \text{Kern } f$) $\Rightarrow f(b_{m+1}), \dots, f(b_n)$ Erzeugendensystem von $\text{Bild } f$.

2) $f(b_{m+1}), \dots, f(b_n)$ linear unabhängig: Z. zg.: $\sum_{i=m+1}^n \lambda_i f(b_i) = 0 \Rightarrow \lambda_i = 0 \forall i = m+1, \dots, n$. Nun ist wegen Linearität von f :

$$\begin{aligned} \sum_{i=m+1}^n \lambda_i f(b_i) = 0 &\Leftrightarrow f\left(\sum_{i=m+1}^n \lambda_i b_i\right) = 0. \\ \Rightarrow \sum_{i=m+1}^n \lambda_i b_i &\in \text{Kern } f \Rightarrow \sum_{i=m+1}^n \lambda_i b_i = \sum_{i=1}^m \mu_i b_i \\ &\Leftrightarrow \sum_{i=m+1}^n \lambda_i b_i + \sum_{i=1}^m (-\mu_i) b_i = 0 \end{aligned}$$

Nun sind aber die b_i ($i = 1, \dots, n$) linear unabhängig (weil Basis von V) und damit folgt aus dieser Gleichung, dass $(-\mu_i) = 0 \forall i = 1, \dots, m \wedge \lambda_i = 0 \forall i = m+1, \dots, n$. Dies zeigt die Aussage. $\square$

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.

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 im $\mathbb{R}^n$). 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 (Skalarprodukt). *Die Abbildung*

$$\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, (x, y) \mapsto \langle x, y \rangle$$

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.

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:

$$0 \leq \langle x, x \rangle - \frac{\langle x, y \rangle^2}{\langle y, y \rangle} \Leftrightarrow \langle x, y \rangle^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle = |x|^2 \cdot |y|^2 \Leftrightarrow |\langle x, y \rangle| \leq |x| \cdot |y|.$$

Hier haben wir im letzten Schritt benutzt, dass $f(x) = x^2$ auf $\mathbb{R}_0^+$ monoton steigend ist (d.h. $x^2 \leq y^2 \Leftrightarrow x \leq y$). $\square$

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 $\phi = \frac{\pi}{2}$ ($n \geq 2$; 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. Um es aufzuschreiben, brauchen wir zunächst folgende Definition:

Definition 5.2.1 (Parallele Geraden). 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 fällen am Punkt P .

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, Lobatschewski-Geometrie, Minkowski-Raumzeit, Klein'sche Kreisscheibe, Poincaré-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. Geometrische Objekte (z.B. Geraden, Kreise) schauen anders aus, als wir es gewohnt sind.

Historische Entwicklung:

- Nikolai Lobatschewskij, János Bolyai (1826): hyperbolische Geometrie
- Carl Friedrich Gauß (~ 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

5.3 ONS, ONB, Orthogonalmatrizen

Das Skalarprodukt ermöglicht uns nun, eine Basis aus orthogonalen Vektoren zu konstruieren.

Definition 5.3.1 (ONS, ONB). Gegeben eine endliche Menge von Vektoren $\{b_1, \dots, b_m\}$ mit $b_i \in \mathbb{R}^n$ für alle $i = 1, \dots, m$ ($m \leq n$; $m, n \in \mathbb{N}$). Ist

$$\langle b_i, b_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

so nennen wir $\{b_1, \dots, b_m\}$ ein *Orthonormalsystem* (kurz ONS). Ist $n = m$, so nennen wir $\{b_1, \dots, b_n\}$ eine *Orthonormalbasis* (kurz ONB) des $\mathbb{R}^n$.

Bemerkung 5.3.1. Man nennt $\{b_1, \dots, b_m\}$ ein Orthonormalsystem, weil die Vektoren $b_1, \dots, b_m$ auf 1 *normiert* (d.h. $\forall i = 1, \dots, m : \langle b_i, b_i \rangle = 1$) und *orthogonal* (d.h. $\forall i \neq j : \langle b_i, b_j \rangle = 0$) sind.

Bemerkung 5.3.2. $\{b_1, \dots, b_m\}$ ist eine Basis des $\mathbb{R}^m$, weil die $b_1, \dots, b_m$ linear unabhängig sind. Hier sieht man die Unabhängigkeit der $b_1, \dots, b_m$ wie folgt: Sei $\sum_{i=1}^m \lambda_i b_i = 0$. Daraus folgt, dass alle $\lambda_i = 0$ sind, denn

$$0 = \langle b_j, \sum_{i=1}^m \lambda_i b_i \rangle = \sum_{i=1}^m \lambda_i \langle b_j, b_i \rangle = \sum_{i=1}^m \lambda_i \delta_{ij} = \lambda_j$$

für alle $j = 1, \dots, m$.

Beispiel 5.3.1 (ONS, ONB). Die Menge der Einheitsvektoren $\{e_1, \dots, e_m\}$ ist ein Orthonormalsystem (ONS) und eine Orthonormalbasis (ONB) des $\mathbb{R}^m$.

Bemerkung 5.3.3. Ist $\{b_1, \dots, b_m\}$ eine ONB, so gilt für jedes $x \in \mathbb{R}^m$:

$$x = \sum_{i=1}^m \langle b_i, x \rangle \cdot b_i$$

Dabei projiziert $\langle b_i, x \rangle$ genau auf den Anteil von x in b_i -Richtung. Es findet also eine orthogonale Zerlegung (Dekomposition) des Vektors x in Bezug auf die Vektoren b_i statt. Dies sieht man wie folgt: Wenn $\{b_1, \dots, b_m\}$ Basis von $\mathbb{R}^m$, dann existiert für jedes $x \in \mathbb{R}^m$ genau ein $\lambda = (\lambda_1, \dots, \lambda_m) \in K^m$ mit

$$x = \sum_{i=1}^m \lambda_i b_i$$

$$\Rightarrow \langle b_j, x \rangle = \langle b_j, \sum_{i=1}^m \lambda_i b_i \rangle = \sum_{i=1}^m \lambda_i \langle b_j, b_i \rangle = \sum_{i=1}^m \lambda_i \delta_{ij} = \lambda_j.$$

Bemerkung 5.3.4 (Gram-Schmidt-Verfahren). Mit dem folgenden Algorithmus, dem sog. *Gram-Schmidt Orthogonalisierungsverfahren*, kann man aus n linear unabhängigen Vektoren $w_1, \dots, w_n$ ein System aus n paarweise orthogonalen Vektoren $v_1, \dots, v_n$ erzeugen.

Man beginnt wie folgt:

$$v_1 = w_1$$

$$v_2 = w_2 - \frac{\langle v_1, w_2 \rangle}{\langle v_1, v_1 \rangle} v_1$$

$$v_3 = w_3 - \frac{\langle v_1, w_3 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, w_3 \rangle}{\langle v_2, v_2 \rangle} v_2$$

Dies führt rekursiv zu der Vorschrift:

$$v_i = w_i - \sum_{j=1}^{i-1} \frac{\langle v_j, w_i \rangle}{\langle v_j, v_j \rangle} v_j$$

Mit einem erweiterten Algorithmus, dem sog. *Gram-Schmidt Orthonormalisierungsverfahren*, kann man aus n linear unabhängigen Vektoren $w_1, \dots, w_n$ n normierte, paarweise orthogonale Vektoren $b_1, \dots, b_n$ (also ein ONS) erhalten:

$$b'_1 = w_1$$

$$b_1 = \frac{b'_1}{|b'_1|}$$

$$b'_2 = w_2 - \langle b_1, w_2 \rangle \cdot b_1$$

$$b_2 = \frac{b'_2}{|b'_2|}$$

$$b'_3 = w_3 - \langle b_1, w_3 \rangle \cdot b_1 - \langle b_2, w_3 \rangle \cdot b_2$$

$\vdots$

Wir benötigen also jeweils einen Zwischenschritt, deshalb die gestrichene Variable. D.h. erst wird der erste Vektor orthogonalisiert, dann auf 1 normiert, dann der zweite Vektor orthogonalisiert, dann auf 1 normiert usw. Dies führt rekursiv zu der Vorschrift:

$$b'_i = w_i - \sum_{j=1}^{i-1} \langle b_j, w_i \rangle \cdot b_j, \quad b_i = \frac{b'_i}{|b'_i|}$$

Beispiel 5.3.2. Auf $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ seien zwei Vektoren w_1, w_2 gegeben:

$$w_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Mit dem Gram-Schmidt-Verfahren kann aus w_1, w_2 eine ONB des $\mathbb{R}^2$ konstruiert werden:

$$b'_1 = w_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad b_1 = \frac{b'_1}{|b'_1|} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$b'_2 = w_2 - \langle b_1, w_2 \rangle \cdot b_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \frac{1}{5} \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \frac{5}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$b_2 = \frac{b'_2}{|b'_2|} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

Wir kommen nochmal zurück zum Begriff der Norm. Es gibt neben der euklidischen Norm noch viele weitere Normen.

Bemerkung 5.3.5 (p -Normen). Sei $x \in \mathbb{R}^n$. Man nennt die euklidische Norm

$$|x| = \sqrt{\langle x, x \rangle} = \sqrt{|x_1|^2 + \dots + |x_n|^2} = \|x\|_2$$

auch 2-Norm. Sie wird vom Skalarprodukt (auch inneres Produkt genannt) induziert. Man unterscheidet sie von anderen Vektornormen, wie z.B. der Maximumsnorm $\|x\|_\infty = \max_{i=1, \dots, n} |x_i|$ oder der 3-Norm $\|x\|_3 = (|x_1|^3 + |x_2|^3 + \dots + |x_n|^3)^{1/3}$ oder der 1-Norm $\|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$ oder allgemein von p -Normen der Form

$$\|\cdot\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

Die euklidische Norm $|\cdot|$ hat folgende Eigenschaften, die sie überhaupt erst zur Norm machen. Mit anderen Worten, alle Normen müssen diese Eigenschaften erfüllen.

Lemma 5.3.1 (Eigenschaften der Norm). *Die Abbildung*

$$\mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto |x|$$

ist definit, absolut homogen und erfüllt die Dreiecksungleichung.

- 1) Für $x \in \mathbb{R}^n$: $|x| = 0 \Leftrightarrow x = 0$ (Definitheit)
- 2) Für $x \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$: $|\lambda x| = |\lambda| \cdot |x|$ (absolute Homogenität)
- 3) Für $x, y \in \mathbb{R}^n$: $|x + y| \leq |x| + |y|$ (Dreiecksungleichung)

Beweis.

Zu 1) – 2): Folgt direkt aus der Definition von $|\cdot|$.

Zu 3):

$$\begin{aligned} |x + y|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &\leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2 \Rightarrow |x + y| \leq |x| + |y| \end{aligned}$$

Hier haben wir die Bilinearität des Skalarprodukts und die Cauchy-Schwarz-Ungleichung benutzt sowie im letzten Schritt die Tatsache, dass $f(x) = x^2$ auf $\mathbb{R}_0^+$ monoton steigend ist. $\square$

Bemerkung 5.3.6. Aus den Eigenschaften der Norm 1) – 3) folgt insbesondere:

$$|x| \geq 0 \quad \forall x \in \mathbb{R}^n.$$

Dies gilt, weil $0 = |0| = \left| \frac{x}{2} + \left(-\frac{x}{2}\right) \right| \leq \left| \frac{x}{2} \right| + \left| \left(-\frac{x}{2}\right) \right| = \frac{1}{2}|x| + \frac{1}{2}|x| = |x|$. Insbesondere ist also die Norm eines Vektors immer positiv (wie zu erwarten, denn sie definiert ja die Länge eines Vektors bzw. den Abstand zwischen Punkten im $\mathbb{R}^n$).

Neben den ONS und ONB, die über das Kronecker-Delta-Symbol definiert sind, gibt es sog. Orthogonalmatrizen, die über die Einheitsmatrix definiert sind.

Definition 5.3.2 (Orthogonalmatrix). Sei $n \in \mathbb{N}$. Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt *Orthogonalmatrix* oder *orthogonale Matrix* genau dann, wenn:

$$A^T \cdot A = E_n.$$

Bemerkung 5.3.7. Wenn A orthogonal, dann ist auch A^T orthogonal, d.h. $A \cdot A^T = E_n$, und A invertierbar mit $A^{-1} = A^T$.

Satz 5.3.1. Sei $n \in \mathbb{N}$. Sei $A \in \mathbb{R}^{n \times n}$. Dann sind folgende Aussagen äquivalent:

- 1) A ist orthogonal
- 2) für $x \in K^n$: $|Ax| = |x|$ (d.h. A ist eine Isometrie, also längenerhaltend)
- 3) für $x, y \in K^n$: $\langle Ax, Ay \rangle = \langle x, y \rangle$

Beweis.

Zu 1) $\Rightarrow$ 2): A ist orthogonal, also $A^T \cdot A = E_n$. Daraus folgt:

$$|Ax| = \sqrt{\langle Ax, Ax \rangle} = \sqrt{(Ax)^T \cdot (Ax)} = \sqrt{x^T A^T A x} = \sqrt{x^T E_n x} = \sqrt{x^T x} = \sqrt{\langle x, x \rangle} = |x|$$

Der weitere Beweis ist Übung. □

Beispiel 5.3.3 (Orthogonalmatrizen). 2×2 -Orthogonalmatrizen sind z.B.

$$E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_\phi = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Wir können dies z.B. für die Drehmatrizen $R_\phi^T \cdot R_\phi$ nachrechnen:

$$\begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \cdot \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} = \begin{pmatrix} \cos^2 \phi + \sin^2 \phi & 0 \\ 0 & \sin^2 \phi + \cos^2 \phi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

6 Lineare Gleichungssysteme

6.1 Elementare Zeilen-/Spaltenumformungen von Matrizen

Definition 6.1.1 (EZU/ESU). Sei $\lambda \in K \setminus \{0\}$, $A \in K^{m \times n}$. Es gibt drei verschiedene *elementare Zeilen-/Spaltenumformungen* von Matrizen:

EZU (ESU) I: Die i -te Zeile (Spalte) wird mit λ multipliziert.

EZU (ESU) II: Zur i -ten Zeile (Spalte) wird das λ -fache der j -ten Zeile (Spalte) addiert.

EZU (ESU) III: Die i -te und j -te Zeile (Spalte) werden vertauscht.

Wie werden EZU/ESU vom Typ I–III umgesetzt? Dazu betrachte man die folgenden drei *Elementarmatrizen* (= quadratische Matrizen, die der Einheitsmatrix ähneln).

1) Die Matrix $P_i(\lambda)$, die sich nur an der i -ten Stelle von der Einheitsmatrix unterscheidet (dort steht λ statt 1):

$$P_i(\lambda) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & \lambda & \vdots \\ & & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

2) Die Matrix $Q_{ij}(\lambda)$, die sich nur in der i -ten Zeile, j -ten Spalte ($i \neq j$) von der Einheitsmatrix unterscheidet (dort steht λ statt 0):

$$Q_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \lambda & \vdots \\ & & & 1 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

3) Die Matrix S_{ij} , die sich nur in der i -ten und j -ten Zeile und Spalte ($i \neq j$) von der Einheitsmatrix unterscheidet (dort steht jeweils eine 1, $s_{ij} = s_{ji} = 1$, dafür steht eine 0 auf der Diagonalen, $s_{ii} = s_{jj} = 0$):

$$S_{ij} = \begin{pmatrix} 1 & & & \dots & 0 \\ & \ddots & & & \\ & & 0 & & 1 & \vdots \\ & & & \ddots & & \\ \vdots & & 1 & & 0 & \\ & & & & & \ddots \\ 0 & \dots & & & & 1 \end{pmatrix}$$

Wie wirken die Elementarmatrizen? Dazu betrachte man eine beliebige Matrix $A \in K^{m \times n}$,

$$A = (a_1, \dots, a_n) = \begin{pmatrix} a^1 \\ \vdots \\ a^m \end{pmatrix}.$$

Hier bezeichnet wieder a^i mit $i = 1, \dots, m$ die i -te Zeile und a_j mit $j = 1, \dots, n$ die j -te Spalte. Statt $(a_1 \dots a_n)$ schreiben wir der Deutlichkeit halber $(a_1, \dots, a_n)$.

Zeilenumformungen: Zeilenumformungen erreicht man durch Multiplikation *von links* mit einer der Matrizen $P_i(\lambda), Q_{ij}(\lambda), S_{ij} \in K^{m \times m}$. Dabei gilt:

Typ	Matrix	Inverse Matrix	Produkt
EZU I	$P_i(\lambda)$	$P_i(\frac{1}{\lambda})$	$P_i(\lambda) \cdot A = \begin{pmatrix} a^1 \\ \vdots \\ \lambda a^i \\ \vdots \\ a^m \end{pmatrix}$
EZU II	$Q_{ij}(\lambda)$	$Q_{ij}(-\lambda)$	$Q_{ij}(\lambda) \cdot A = \begin{pmatrix} a^1 \\ \vdots \\ a^i + \lambda a^j \\ \vdots \\ a^m \end{pmatrix}$
EZU III	S_{ij}	S_{ij}	$S_{ij} \cdot A = \begin{pmatrix} a^1 \\ \vdots \\ a^j \\ \vdots \\ a^i \\ \vdots \\ a^m \end{pmatrix}$

Wir können die jeweilige Zeilenumformung rückgängig machen, indem wir *von links mit der inversen Matrix* multiplizieren (z. B. $P_i^{-1}(\lambda) \cdot P_i(\lambda) \cdot A = P_i(\frac{1}{\lambda}) \cdot P_i(\lambda) \cdot A = E_m \cdot A = A$).

Spaltenumformungen: Spaltenumformungen erreicht man durch Multiplikation *von rechts* mit einer der Matrizen $P_i(\lambda), Q_{ij}(\lambda), S_{ij} \in K^{n \times n}$. Dabei gilt:

Typ	Matrix	Inverse Matrix	Produkt
ESU I	$P_i(\lambda)$	$P_i(\frac{1}{\lambda})$	$A \cdot P_i(\lambda) = (a_1, \dots, \lambda a_i, \dots, a_n)$
ESU II	$Q_{ij}(\lambda)$	$Q_{ij}(-\lambda)$	$A \cdot Q_{ij}(\lambda) = (a_1, \dots, \lambda a_i + a_j, \dots, a_n)$
ESU III	S_{ij}	S_{ij}	$A \cdot S_{ij} = (a_1, \dots, a_j, \dots, a_i, \dots, a_n)$

Achtung: Bei den Spalten wird mit $Q_{ij}(\lambda)$ das λ -fache der i -ten Spalte zur j -ten Spalte addiert!

Dies ist genau umgekehrt wie bei den Zeilen. Analog zu den Zeilenumformungen können wir die jeweilige Spaltenumformung rückgängig machen, indem wir *von rechts mit der inversen Matrix* multiplizieren (z. B. $A \cdot P_i(\lambda) \cdot P_i^{-1}(\lambda) = A \cdot E_n = A$).

Lemma 6.1.1. *Seien $m, n \in \mathbb{N}$. Sei K ein Körper, $A \in K^{m \times n}$ und $a_{11} \neq 0$. Dann gilt:*

1) *Es existieren $G \in K^{m \times m}$ invertierbar und $A' \in K^{(m-1) \times (n-1)}$, so dass*

$$A = G \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

2) *Es existieren $G \in K^{m \times m}$ invertierbar, $H \in K^{n \times n}$ invertierbar und $\tilde{A} \in K^{(m-1) \times (n-1)}$, so dass:*

$$A = G \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{pmatrix} \cdot H$$

Beweis.

Zu 1): Wir brauchen EZU vom Typ II für die Zeilen $2 \leq i \leq m$. Wir wählen $\lambda = -\frac{a_{i1}}{a_{11}}$ und $Q_{i1}(-\frac{a_{i1}}{a_{11}})$. Dann ergibt das Produkt $Q_{i1}(-\frac{a_{i1}}{a_{11}}) \cdot A$ eine Matrix, in der das erste Element der i -ten Zeile gleich 0 ist. Dies wiederholen wir für alle Zeilen, d.h. wir multiplizieren A insgesamt von links mit $Q_{m1}(-\frac{a_{m1}}{a_{11}}) \cdot \dots \cdot Q_{21}(-\frac{a_{21}}{a_{11}})$ und erhalten damit eine Matrix von der Form, wie wir sie wollen. Dabei ist das Produkt $Q_{m1}(-\frac{a_{m1}}{a_{11}}) \cdot \dots \cdot Q_{21}(-\frac{a_{21}}{a_{11}}) =: G^{-1}$ wohldefiniert und invertierbar, weil jedes der Q_{ij} invertierbar ist. Weil aber $G^{-1}A = B \Rightarrow GG^{-1}A = GB \Rightarrow A = GB$ und $(G^{-1})^{-1} = G$, setzen wir

$$G = \left[Q_{m1}\left(-\frac{a_{m1}}{a_{11}}\right) \cdot \dots \cdot Q_{21}\left(-\frac{a_{21}}{a_{11}}\right) \right]^{-1}$$

Zu 2): Wir brauchen zusätzlich ESU vom Typ II für die Spalten $2 \leq j \leq n$ und vom Typ I für die Spalte $j = 1$. Multipliziert man $G^{-1}A$ von rechts mit $Q_{12}(-\frac{a_{12}}{a_{11}})$ bis $Q_{1n}(-\frac{a_{1n}}{a_{11}})$, verschwinden alle außer dem ersten Matrixeintrag der ersten Zeile. Multipliziert man zum Schluss von rechts mit $P_1(\frac{1}{a_{11}})$, wird der erste Matrixeintrag (vormals a_{11}) auf 1 skaliert. Damit ist:

$$H = \left[Q_{12}\left(-\frac{a_{12}}{a_{11}}\right) \cdot \dots \cdot Q_{1n}\left(-\frac{a_{1n}}{a_{11}}\right) \cdot P_1\left(\frac{1}{a_{11}}\right) \right]^{-1}$$

□

Satz 6.1.1 (Äquivalenznormalform). *Seien $m, n \in \mathbb{N}$. Sei K ein Körper und $A \in K^{m \times n}$. Dann existieren $G \in K^{m \times m}$ invertierbar und $H \in K^{n \times n}$ invertierbar und $p \in \{0, \dots, \min(m, n)\}$, so*

dass

$$A = G \cdot \begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix} \cdot H$$

mit $\begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix} \in K^{m \times n}$ und Einheitsmatrix $E_p = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$. Diese Darstellung nennt man

die Äquivalenznormalform von A .

Beweis. Übung.

Beweisidee: Wiederholte Anwendung von Aussage 2) von Lemma 6.1.1. □

6.2 Lösung von Gleichungssystemen der Form $Ax = y$ mit $x, y \in K^n$

Wir können dieses Kapitel durch zwei Fragen motivieren:

1. Frage: Seien $A \in K^{m \times n}$, $y \in \mathbb{R}^m$ gegeben und $x \in \mathbb{R}^n$ gesucht. Wann besitzt das lineare Gleichungssystem $Ax = y$ für jedes $y \in \mathbb{R}^m$ genau eine Lösung $x \in \mathbb{R}^n$? Die Antwort lautet: Genau dann, wenn $m = n$ und $A \in K^{n \times n}$ invertierbar ist.

2. Frage: Lässt sich eine invertierbare Matrix $A \in K^{n \times n}$ allein durch elementare Zeilen- und Spaltenumformungen (EZU/ESU) auf die Einheitsmatrix E_n umformen? Die Antwort lautet: Ja. Dabei wird A zunächst auf Dreiecksform (den Algorithmus dazu nennt man das *Gaußsche Eliminationsverfahren*) und dann auf die Form der Einheitsmatrix gebracht. Wozu ist das nützlich? Die Transformation zur Dreiecksform nutzen wir, um das lineare Gleichungssystem $Ax = y$ zu lösen. Die Transformation zur Einheitsmatrix nutzen wir, um A^{-1} zu konstruieren.

Bevor wir uns der 1. Frage zuwenden, betrachten wir die 2. Frage.

Definition 6.2.1 (Dreiecksmatrix). Sei $n \in \mathbb{N}$. Sei K ein Körper.

$R = (r_{ij}) \in K^{n \times n}$ heißt *rechte* oder *obere Dreiecksmatrix*, wenn gilt:

$$r_{ij} = 0 \quad \forall i > j \quad (i, j = 1, \dots, n).$$

$L = (l_{ij}) \in K^{n \times n}$ heißt *linke* oder *untere Dreiecksmatrix*, wenn gilt:

$$l_{ij} = 0 \quad \forall j > i \quad (i, j = 1, \dots, n).$$

Satz 6.2.1 (Transformation zur Dreiecksmatrix). Sei $n \in \mathbb{N}$. Sei K ein Körper und $A \in K^{n \times n}$ invertierbar. Dann gilt: A kann durch elementare Zeilenumformungen (EZU) vom Typ II und III in die Form einer oberen Dreiecksmatrix R gebracht werden und dabei sind alle Diagonalelemente $r_{ii} \neq 0$.

Beweis. Algorithmus des Gaußschen Eliminationsverfahrens (kurz: Gaußsches Verfahren).

Dass man mit diesem Algorithmus für invertierbare Matrizen tatsächlich auf Dreiecksform kommt, in der alle Diagonalelemente $r_{ii} \neq 0$ sind, folgt daraus, dass es im Fall einer invertierbaren Matrix genau eine Lösung und damit keine Nullspalten/-zeilen gibt (Satz und Beweis folgt).

Der Algorithmus des Gaußschen Verfahrens ist wie folgt:

1) Man betrachte die erste Zeile von A . Ist $a_{11} = 0$, vertauscht man die erste Zeile mit einer anderen Zeile (EZU vom Typ III). So ist auf jeden Fall der erste linke obere Matrixeintrag $\neq 0$ (und es gibt auf jeden Fall eine Zeile a^i mit $a_{i1} \neq 0$, da sonst die ganze erste Spalte $= 0$ und damit A nicht invertierbar wäre).

2) Nun addiert man jeweils ein geeignetes Vielfaches λ der ersten Zeile zu jeder der anderen Zeilen a^j mit $j = 2, \dots, n$, so dass jeweils deren erster Eintrag verschwindet: $a_{j1} = 0$ für alle $j = 2, \dots, n$ (mehrfache EZU vom Typ II). Dadurch verschwinden in der ersten Spalte alle bis auf den obersten Matrixeintrag, $a_{11} \neq 0$.

3) Dieses Verfahren wiederholt man mit der zweiten Zeile bzw. Spalte und den „darunter liegenden“ Zeilen a^j mit $j = 3, \dots, n$ (d.h. mit der Untermatrix, die entsteht, wenn man die erste Zeile und Spalte streicht) usw. bis man bei der letzten Zeile bzw. Spalte angelangt ist und die Matrix vollständig in eine Dreiecksmatrix umgewandelt wurde. $\square$

Bemerkung 6.2.1 (Algorithmus zur Lösung von $Ax = y$). Betrachte $Ax = y$ mit $A \in K^{n \times n}$ invertierbar, $y \in R^n$ gegeben und $x \in R^n$ gesucht. Den Algorithmus zur Lösung dieses Gleichungssystems besteht aus zwei Schritten:

- 1) Konstruktion der Dreiecksform mittels Gaußschem Eliminationsverfahren.
- 2) Lösung des Gleichungssystems durch Rückwärtseinsetzen.

Beispiel 6.2.1 (Gaußsches Verfahren). Wir betrachten das Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 3 \\ x_1 + 2x_2 - x_3 &= 4 \end{aligned}$$

In Matrixnotation mit „erweiterter Matrix“:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 3 \\ 1 & 2 & -1 & 4 \end{array} \right)$$

Gaußsches Verfahren:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 3 \\ 1 & 2 & -1 & 4 \end{array} \right) \xrightarrow{EZU II} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & -2 & 3 \end{array} \right) \xrightarrow{EZU II} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & -3 & 4 \end{array} \right)$$

Daraus erhält man das Gleichungssystem,

$$-3x_3 = 4, \quad -x_2 - x_3 = 1, \quad x_1 + x_2 + x_3 = 1,$$

das gelöst wird durch $x_3 = -\frac{4}{3}, x_2 = \frac{1}{3}, x_1 = 2$.

Bemerkung 6.2.2 (Anwendbarkeit des Algorithmus). Man kann das Gaußsche Verfahren auch auf beliebige Gleichungssysteme der Form $Ax = y$ mit $A \in K^{m \times n}, x \in \mathbb{R}^n, y \in \mathbb{R}^m$ anwenden (wobei i. A. die Matrix A nicht invertierbar ist oder $m \neq n$), muss dann aber manchmal Nullspalten überspringen und es kann sein, dass am Ende Nullzeilen stehen bleiben.

Statt Dreiecksform erhält man dann i. A. die sog. Zeilenstufenform (mehr dazu später). Entsprechend erhält man i. A. nicht *genau eine* Lösung x für jedes y , sondern entweder *mehrere* Lösungen, d.h. einen sog. *Lösungsraum* (dies erkennt man daran, dass am Ende des Algorithmus Zeilen der Form „ $0 = 0$ “ bleiben), oder einen Widerspruch. Das GLS ist dann entweder *universell lösbar* oder *nicht lösbar*.

Bemerkung 6.2.3 (Invertierbarkeit von A). Wir benutzen das Gaußsche Verfahren nicht nur, um Gleichungssysteme zu lösen, sondern auch, um festzustellen, ob eine Matrix $A \in K^{n \times n}$ invertierbar ist. Lässt sich eine Matrix A *nicht* in die Form einer Dreiecksmatrix R bringen, in der alle Diagonalelemente $r_{ii} \neq 0$ sind, so ist A nicht invertierbar.

Satz 6.2.2 (Transformation zur Einheitsmatrix). Sei $n \in \mathbb{N}$. Sei K ein Körper und $R \in K^{n \times n}$ eine obere Dreiecksmatrix mit allen Diagonalelementen $r_{ii} \neq 0$. Dann gilt: R kann durch elementare Zeilenumformungen (EZU) vom Typ I und II in die Form einer Einheitsmatrix E_n gebracht werden.

Beweis. Man wendet auf R mehrere EZU vom Typ II und eine EZU vom Typ I an, um zunächst die letzte Spalte wie folgt zu transformieren:

$$\begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & r_{(n-1)n} \\ 0 & \dots & 0 & r_{nn} \end{pmatrix} \xrightarrow{\text{mehrere EZU II}} \begin{pmatrix} r_{11} & r_{12} & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & r_{nn} \end{pmatrix} \xrightarrow{\text{EZU I}} \begin{pmatrix} r_{11} & r_{12} & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Dies wiederholt man für alle anderen Spalten r_j mit $j = n - 1$ bis $j = 1$. □

Satz 6.2.3 (Bestimmung von A^{-1}). Sei $n \in \mathbb{N}$. Sei K ein Körper und $A \in K^{n \times n}$. Dann gilt: Lässt sich A durch elementare Zeilenumformungen (EZU) in eine Einheitsmatrix E_n verwandeln, dann ist A invertierbar und A^{-1} entsteht aus der Einheitsmatrix E_n durch dieselben Zeilenumformungen (EZU). Kurz:

$$A \xrightarrow{\text{bestimmte EZU}} E_n \Rightarrow E_n \xrightarrow{\text{bestimmte EZU}} A^{-1}$$

Beweis. Kann man A durch bestimmte EZU in E_n umformen, so gibt es eine bestimmte invertierbare Matrix G (ein Produkt aus invertierbaren Elementarmatrizen) mit $G \cdot A = E_n$. Wegen der Eindeutigkeit des inversen Elements und weil bei der Matrixmultiplikation das linksinverse Element gleich das inverse Element ist, gilt: $G = A^{-1}$. Daraus folgt:

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

□

Beispiel 6.2.2 (Berechnung von A^{-1}). Zur leichteren Notation verwenden wir „Doppelmatrizen“:

$$\begin{aligned}
 \left(A \mid E_3 \right) &= \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{\text{bestimmte EZU}} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array} \right) \\
 &\xrightarrow{\text{bestimmte EZU}} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -3 & -3 & 1 & 1 \end{array} \right) \xrightarrow{\text{bestimmte EZU}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -3 & -3 & 1 & 1 \end{array} \right) \\
 &\xrightarrow{\text{bestimmte EZU}} \dots \xrightarrow{\text{bestimmte EZU}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \end{array} \mid A^{-1} \right) = \left(E_3 \mid A^{-1} \right)
 \end{aligned}$$

Wir formen solange um, bis links die Einheitsmatrix steht. Sobald dies der Fall ist, steht rechts die inverse Matrix zu unserer Ausgangsmatrix.

Bemerkung 6.2.4 (2×2 -Matrix). Im Fall einer 2×2 -Matrix M können wir uns dieses Verfahren sparen, denn wir wissen bereits aus der Übung, dass gilt:

$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist invertierbar genau dann, wenn $ad - bc \neq 0$, und die inverse Matrix ist

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

Wir kommen nun zur Antwort auf die eingangs gestellte 1. Frage nach der Existenz einer eindeutigen Lösung. Dabei gilt: $\exists$ genau eine Lösung x zu jedem $y \Leftrightarrow A$ invertierbar $\Leftrightarrow f$ bijektiv.

Satz 6.2.4. Seien $m, n \in \mathbb{N}$. Sei K ein Körper, $f : K^n \rightarrow K^m$ eine lineare Abbildung und $A \in K^{m \times n}$ ihre darstellende Matrix. Dann gilt:

$$f \text{ bijektiv} \Leftrightarrow A \text{ invertierbar} \wedge m = n$$

Beweis.

„ $\Leftarrow$ “: A kann nur invertierbar sein, wenn $m = n$. Wenn $A \in K^{n \times n}$ invertierbar, dann hat die lineare Abbildung $f(x) = Ax$ die Umkehrfunktion $f^{-1}(x) = A^{-1}x$ (denn $x = A^{-1}Ax = f^{-1}(f(x))$ und umgekehrt). Diese ist, weil sie existiert, eindeutig und bijektiv, ebenso wie f dann bijektiv ist.

„ $\Rightarrow$ “: Wir nutzen die Äquivalenznormalform von $A \in K^{m \times n}$ (siehe Satz 6.1.1). Wir wissen

also, es gibt $G \in K^{m \times m}$ und $H \in K^{n \times n}$ invertierbar mit

$$A = G \cdot \begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix} \cdot H.$$

Da G und H invertierbar, sind die zugehörigen linearen Abbildungen $g : K^m \rightarrow K^m, g(x) = Gx$ und $h : K^n \rightarrow K^n, h(x) = Hx$ bijektiv und somit die lineare Abbildung $f' : K^n \rightarrow K^m, f' = g^{-1} \circ f \circ h^{-1}$ ebenfalls bijektiv mit darstellender Matrix $\begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix}$. Weil f' bijektiv ist, kann diese Matrix keine Nullzeile und keine Nullspalte haben. Dies sieht man wie folgt:

Ist die i -te Zeile eine Nullzeile, gilt: $f'_i(x) = 0$ für $x \in K^n$ beliebig. Somit gilt für $e_i \in K^n$, dass $e_i \notin f'(K^n)$ und damit $f'(K^n) \neq K^m$, also f' nicht surjektiv.

Ist die j -te Spalte eine Nullspalte, gilt: $f'(e_j) = 0 = f'(0)$, aber $e_j \neq 0$. Somit ist f' nicht injektiv.

Damit folgt: $p = m$ und $p = n$. Somit ist $m = n$ und $\begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix} = E_n$. Weil G, H invertierbar, ist demnach auch $A = G \cdot E_n \cdot H = G \cdot H$ invertierbar. $\square$

Lemma 6.2.1. Sei $n \in \mathbb{N}$. Sei K ein Körper und $f : K^n \rightarrow K^n$ linear. Dann sind äquivalent:

- 1) f ist bijektiv
- 2) f ist injektiv
- 3) f ist surjektiv

Beweis. Offensichtlich gilt: 1) $\Rightarrow$ 2) und 1) $\Rightarrow$ 3). Bleibt zu zeigen, dass für $f : K^n \rightarrow K^n$ linear gilt: 2) $\Rightarrow$ 1) und 3) $\Rightarrow$ 1) (dann auch 2) $\Rightarrow$ 3) und umgekehrt).

Wir betrachten wieder die lineare Abbildung $f' = g^{-1} \circ f \circ h^{-1}$ aus dem vorhergehenden Beweis. Weil g und h bijektiv sind, gilt: f' injektiv $\Leftrightarrow f$ injektiv und f' surjektiv $\Leftrightarrow f$ surjektiv. Es genügt also 2) $\Rightarrow$ 1) und 3) $\Rightarrow$ 1) für f' zu zeigen (dann folgt dasselbe für f). Die darstellende Matrix von f' ist $\begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix}$.

Wir zeigen zuerst: f' injektiv $\Rightarrow f'$ bijektiv.

Wenn f' injektiv, dann hat die Matrix $\begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix}$ keine Nullspalte (siehe vorhergehenden Beweis). Weil aber die Matrix quadratisch ist, kann sie somit auch keine Nullzeile haben $\Rightarrow E_n$ ist darstellende Matrix von f' und damit f' bijektiv.

Analog zeigt man: f' surjektiv $\Rightarrow f'$ bijektiv. $\square$

Satz 6.2.5. Sei $n \in \mathbb{N}$. Sei K ein Körper und $A \in K^{n \times n}$. Dann sind äquivalent:

- 1) A ist invertierbar
- 2) Sei $x \in K^n$. Dann gilt: $Ax = 0 \Rightarrow x = 0$. D. h. das homogene lineare Gleichungssystem $Ax = 0$ hat nur die triviale Lösung $x = 0$.
- 3) $\forall y \in K^n$ existiert genau ein $x \in K^n : Ax = y$. D. h. das inhomogene lineare Gleichungssystem $Ax = y$ hat für jedes $y \in K^n$ genau eine Lösung $x \in K^n$.

Bemerkung 6.2.5 (Fredholmsche Alternative). Das impliziert insbesondere, dass entweder das inhomogene Gleichungssystem $Ax = y$ für jedes y eindeutig lösbar ist oder das homogene Gleichungssystem $Ax = 0$ auch eine nichttriviale Lösung $x \neq 0$ besitzt (*Fredholmsche Alternative*).

Beweis (von Satz 6.2.5). Wir betrachten $f : K^n \rightarrow K^n$ mit $f(x) = Ax$.

Zu 3) $\Leftrightarrow$ 1): Es gibt für jedes $y \in K^n$ genau ein $x \in K^n$ mit $f(x) = y \Leftrightarrow f$ bijektiv $\Leftrightarrow A$ invertierbar (siehe Satz 6.2.4).

Zu 3) $\Leftrightarrow$ 2): Es gibt für jedes $y \in K^n$ genau ein $x \in K^n$ mit $f(x) = y \Leftrightarrow f$ bijektiv $\Leftrightarrow f$ injektiv (siehe Lemma 6.2.1). Nun wissen wir bereits aus der Übung: f injektiv $\Leftrightarrow \text{Kern } f = \{0\}$, d.h. nur die 0 bildet auf die 0 ab, somit hat das homogene GLS $f(x) = Ax = 0$ nur die triviale Lösung $x = 0$. $\square$

6.3 Rang einer Matrix

Definition 6.3.1 (Zeilenrang, Spaltenrang). Sei K ein Körper und $m, n \in \mathbb{N}$. Sei $A \in K^{m \times n}$ mit Zeilen $a^1, \dots, a^m$ und Spalten $a_1, \dots, a_n$, d.h.

$$A = (a_1, \dots, a_n) = \begin{pmatrix} a^1 \\ \vdots \\ a^m \end{pmatrix}.$$

1) Wir bezeichnen die lineare Hülle $\text{span}(a_1, \dots, a_n)$ als *Spaltenraum* von A und die lineare Hülle $\text{span}(a^1, \dots, a^m)$ als *Zeilenraum* von A .

2) Der *Spaltenrang* von A ist die Dimension vom Spaltenraum von A und der *Zeilenrang* von A die Dimension vom Zeilenraum von A .

Bemerkung 6.3.1. Der Spaltenrang bestimmt also die maximale Anzahl linear unabhängiger Spalten von A , der Zeilenrang die maximale Anzahl linear unabhängiger Zeilen von A .

Satz 6.3.1 (Rang). Seien $m, n \in \mathbb{N}$, K ein Körper und $A \in K^{m \times n}$. Dann gilt: Der Zeilenrang von A ist gleich dem Spaltenrang von A . Man nennt diese Zahl den *Rang* von A , kurz: $\text{Rang } A$.

Beweis. Evtl. Übung.

Beweisidee: Man benutze die Äquivalenznormalform $A = G \cdot \begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix} \cdot H$ und zeigt, dass durch „Dranmultiplizieren“ von invertierbaren Matrizen G und H der Zeilenrang und Spaltenrang der ursprünglichen Matrix nicht verändert wird. Zuletzt sieht man, dass der Zeilenrang und Spaltenrang von $\begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix}$ jeweils p , also gleich ist. Der Rang von A ist also p . $\square$

Satz 6.3.2 (Rang = Dimension vom Bild). Seien $m, n \in \mathbb{N}$, K ein Körper und $A \in K^{m \times n}$. Der Rang von $A \in K^{m \times n}$ ist die Dimension vom Bild der zugehörigen linearen Abbildung $f(x) = Ax$, $f : K^n \rightarrow K^m$, d.h.:

$$\text{Rang } A = \dim \text{Bild } f.$$

Beweis. Dies sieht man wie folgt. Der Rang von A ist nach Definition die Dimension des Spaltenraums von A , wobei der Spaltenraum wiederum definiert ist als $\text{span}(a_1, \dots, a_n)$. Nun ist aber

$$\begin{aligned}\text{span}(a_1, \dots, a_n) &= \left\{ \sum_{i=1}^n \lambda_i \cdot a_i \mid \lambda_i \in K \forall i = 1, \dots, n \right\} = \left\{ \sum_{i=1}^n \lambda_i \cdot (A \cdot e_i) \mid \lambda_i \in K \forall i = 1, \dots, n \right\} \\ &= \left\{ \sum_{i=1}^n A \cdot (\lambda_i \cdot e_i) \mid \lambda_i \in K \forall i = 1, \dots, n \right\} = \left\{ A \cdot \lambda \mid \lambda \in K^n \right\}\end{aligned}$$

Dies ist aber gerade das Bild von f , denn $\text{Bild } f = \{f(x) = Ax \in K^m \mid x \in K^n\}$. $\square$

Beispiel 6.3.1 (Zeilenrang gleich Spaltenrang).

$$M = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 4 & 12 \\ 0 & 2 & 6 \end{pmatrix}, \quad \text{Zeilenrang } M = 2 = \text{Spaltenrang } M \quad (\text{also } \text{Rang } M = 2)$$

Wir sehen, es gibt jeweils zwei linear unabhängige Zeilen und Spalten.

6.4 Lösung von Gleichungssystemen der Form $Ax = y$ mit $x \in K^n$, $y \in K^m$

Definition 6.4.1 (Zeilenstufenform). Seien K ein Körper, $m, n \in \mathbb{N}$ und $A \in K^{m \times n}$, $A \neq 0$. Die Matrix A hat *Zeilenstufenform*, wenn es eine Anzahl von Zeilen $r \in \{1, \dots, \min(m, n)\}$ und Spalten $j_1, \dots, j_r \in \mathbb{N}$ gibt mit $1 \leq j_1 < j_2 < \dots < j_r \leq n$, so dass

$$\begin{aligned}a_{ij} &\neq 0 & i = k, j = j_k & \quad (k = 1, \dots, r) \\ a_{ij} &= 0 & i \geq k, j_{k-1} \leq j < j_k & \quad (k = 1, \dots, r+1)\end{aligned}$$

mit $j_0 := 1$ und $j_{r+1} := n+1$.

Satz 6.4.1. Sei K ein Körper und $m, n \in \mathbb{N}$. Sei r wie in Def. 6.4.1. Jede Matrix $A \in K^{m \times n}$, $A \neq 0$, kann durch elementare Zeilenumformungen (EZU) vom Typ II und III in Zeilenstufenform gebracht werden. Dabei gilt:

$$\text{Rang } A = r.$$

Beweis.

Zur Existenz der Zeilenstufenform: Gaußsches Verfahren. Der Unterschied zu invertierbaren Matrizen (und der dort erhaltenen Dreiecksform mit Diagonalelementen $r_{ii} \neq 0$) besteht lediglich darin, dass in der jeweils zu betrachtenden Untermatrix „führende“ (d.h. links stehende) Nullspalten auftreten können. Diese werden im Algorithmus einfach übersprungen. Zum Schluss können außerdem Nullzeilen stehen bleiben.

Zum Rang: Da der Zeilenrang von A gleich dem Spaltenrang von A (und damit gleich dem Rang von A) ist, brauchen wir in der Zeilenstufenform nur die Zeilen zählen, die keine Nullzeilen sind. Diese „beginnen“ jeweils mit einem Eintrag $a_{ij} \neq 0$ in unterschiedlichen Spalten $j_1 < j_2 < \dots < j_r$, sind also linear unabhängig. Die Anzahl solcher Zeilen ist nach Definition gerade r . $\square$

Beispiel 6.4.1 (Zeilenstufenform, Rang). Wir betrachten die Matrix M von zuvor (Bsp. 6.4.1):

$$\begin{pmatrix} 2 & 1 & 3 \\ 0 & 4 & 12 \\ 0 & 2 & 6 \end{pmatrix} \xrightarrow{EZU} \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 6 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \text{Rang } M = 2$$

Ist die Matrix einmal in Zeilenstufenform, können wir den Rang also direkt an der Anzahl der Zeilen, die keine Nullzeilen sind, ablesen.

Lemma 6.4.1. Seien $m, n \in \mathbb{N}$, K ein Körper, $A \in K^{m \times n}$ und $y \in K^m$. Es bezeichne $\text{Lös}(A, y) := \{x \in K^n | Ax = y\}$ den Lösungsraum des Gleichungssystems $Ax = y$. Dann gilt:

- 1) $\dim \text{Lös}(A, 0) = n - \text{Rang } A$.
- 2) Falls $x' \in K^n$ existiert mit $Ax' = y$, dann gilt:
 $\text{Lös}(A, y) = \{x \in K^n | Ax = y\} = \{x' + x_0 | x_0 \in K^n, Ax_0 = 0\}$.

Bemerkung 6.4.1. Die Lösung des inhomogenen GLS setzt sich also aus der Lösung des homogenen GLS plus einer (beliebigen) speziellen Lösung x' des inhomogenen GLS zusammen.

Beweis.

Zu 1): Dies folgt direkt aus der Dimensionsformel für Kern und Bild. Sei $f(x) = Ax$, dann:

$$n = \dim K^n = \dim \text{Kern } f + \dim \text{Bild } f = \dim \{x \in K^n | Ax = 0\} + \text{Rang } A.$$

Zu 2): Sei $Ax' = y$. Dann gilt: $Ax = y \Leftrightarrow A(x' + (x - x')) = y \Leftrightarrow A(x - x') = 0$ (denn $Ax' = y \Leftrightarrow x_0 := x - x'$ und $Ax_0 = 0 \Leftrightarrow x = x' + x_0$ und $Ax_0 = 0$). $\square$

Satz 6.4.2 (Allgemeine Lösung von GLS). Seien $m, n \in \mathbb{N}$, K ein Körper, $A \in K^{m \times n}$ und $y \in K^m$. Dann gilt:

- 1) $\text{Rang}(A|y) = \text{Rang } A \Leftrightarrow Ax = y$ besitzt mindestens eine Lösung $x \in K^n$
- 2) $\text{Rang } A = n \Rightarrow Ax = y$ besitzt höchstens eine Lösung $x \in K^n$
- 3) $\text{Rang}(A|y) = \text{Rang } A = n \Leftrightarrow Ax = y$ besitzt genau eine Lösung $x \in K^n$

Bemerkung 6.4.2. Dieser Satz impliziert, dass das GLS genau dann *keine* Lösung hat, wenn $\text{Rang}(A|y) \neq \text{Rang } A$ (dies folgt durch Kontraposition direkt aus Aussage 1).

Beweis.

Zu 1): $\text{Rang}(A|y) = \text{Rang } A \Leftrightarrow \dim \text{span}(a_1, \dots, a_n, y) = \dim \text{span}(a_1, \dots, a_n)$
 $\Leftrightarrow \text{span}(a_1, \dots, a_n, y) = \text{span}(a_1, \dots, a_n) \Leftrightarrow y \in \text{span}(a_1, \dots, a_n) \Leftrightarrow \exists \lambda \in K^n : y = \sum_{i=1}^n \lambda_i a_i$
 $\Leftrightarrow y = Ax = \sum_{i=1}^n a_i x_i$ besitzt mindestens eine Lösung, nämlich $x = \lambda$.

Zu 2): $\text{Rang } A = n \Leftrightarrow a_1, \dots, a_n$ linear unabhängig $\Rightarrow$ es existiert höchstens eine Linearkombination mit $\sum_{i=1}^n \lambda_i a_i = y$ und damit höchstens eine Lösung $x \in K^n$ (hier $x = \lambda$).

Zu 3): „ $\Rightarrow$ “: Folgt direkt aus 1) und 2).

„ $\Leftarrow$ “: $Ax = y$ besitzt genau eine Lösung x zu jedem $y \Leftrightarrow Ax = 0$ besitzt nur die Lösung $x = 0 \Leftrightarrow a_1, \dots, a_n$ linear unabhängig $\Leftrightarrow \text{Rang } A = n$. Aus 1) folgt außerdem (weil mind. eine Lösung existiert): $\text{Rang}(A|y) = \text{Rang } A$. $\square$

Beispiel 6.4.2 (Allgemeine Lösung von GLS).

$$x_1 + 2x_2 + 3x_3 = 1$$

$$4x_1 + 5x_2 + 6x_3 = 3$$

$$7x_1 + 8x_2 + 9x_3 = 5$$

Die zugehörige erweiterte Matrix ist:

$$(A|y) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 5 \end{array} \right)$$

Aus elementaren Zeilenumformungen bekommt man:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 5 \end{array} \right) \xrightarrow{EZU} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -1 \\ 0 & -6 & -12 & -2 \end{array} \right) \xrightarrow{EZU} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -1 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Wäre in der letzten Zeile nicht „0=0“ gestanden, sondern z.B. „0=1“, hätten wir einen Widerspruch und das GLS wäre nicht lösbar (in dem Fall wäre $\text{Rang } A = 2 \neq 3 = \text{Rang } (A|y)$). So haben wir aber eine Lösung.

Rückübersetzen in Gleichungen: Die dritte Zeile ergibt $0 \cdot x_3 = 0$, d.h. wir können x_3 frei wählen. Wir setzen $x_3 = \mu$.

Es verbleiben die Gleichungen aus der ersten und zweiten Zeile (mit $x_3 = \mu$):

$$x_1 + 2x_2 + 3\mu = 1$$

$$3x_2 + 6\mu = 1$$

Aus $3x_2 + 6\mu = 1 \Rightarrow x_2 = 1/3 - 2\mu$ und aus $x_1 + 2x_2 + 3\mu = 1 \Rightarrow x_1 = 1 - 3\mu - 2x_2 = 1 - 3\mu - 2/3 + 4\mu = 1/3 + \mu$. Damit ist die allgemeine Lösung:

$$\text{Lös}(A, y) = \left\{ x = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \mid \mu \in \mathbb{R} \right\}$$

Hier ist $x' = (1/3, 1/3, 0)$ eine spezielle Lösung und für die Lösung des homogenen GLS gilt:

$$\text{Lös}(A, 0) = \left\{ x = \mu \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \mid \mu \in \mathbb{R} \right\}.$$

Es ist also $\dim \text{Lös}(A, 0) = 1$ und $\text{Rang } A = 2$ und damit $\dim \text{Lös}(A, 0) + \text{Rang } A = 1 + 2 = 3$, wie erwartet.

Bemerkung 6.4.3 (Affiner Unterraum). Geometrisch ist der Lösungsraum $\text{Lös}(A, y)$ ein *affiner Unterraum* von K^n , d.h. ein um einen Vektor x' verschobener Unterraum.

7 Determinanten

Beispiel 7.0.1. Sei $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine 2×2 -Matrix. Dann gilt:

$$\det M := ad - bc = \begin{pmatrix} a \\ c \end{pmatrix} \times \begin{pmatrix} b \\ d \end{pmatrix}.$$

Hier bezeichnet „ $\times$ “ das aus der Schule bekannte Kreuzprodukt.

Geometrische Anschauung: Die Determinante misst also die (orientierte) *Fläche* des durch die Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ aufgespannten Parallelogramms.

Beispiel 7.0.2. Sei $M' = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ eine 3×3 -Matrix. Dann gilt:

$$\det M' := a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - c_1b_2) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \left[\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \times \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right] =: a \cdot (b \times c).$$

Geometrische Anschauung: Die Determinante misst somit das (orientierte) *Volumen* des durch die Vektoren $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ und $c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ aufgespannten Körpers.

Bemerkung 7.0.1 (Determinante als Volumenform). Man sieht bereits im Fall der 2×2 - und 3×3 -Matrizen (ohne dass wir dies an dieser Stelle beweisen): die Determinante misst das n -dimensionale (orientierte) *Volumen* des durch die n Spaltenvektoren einer $n \times n$ -Matrix aufgespannten geometrischen Körpers.

Bemerkung 7.0.2 (Invertierbarkeit, $\det M \neq 0$). Man sieht bereits im Fall der 2×2 - und 3×3 -Matrizen (ohne dass wir dies an dieser Stelle beweisen): Die Spaltenvektoren der Matrix $M \in K^{n \times n}$ sind genau dann linear unabhängig, wenn das von ihnen aufgespannte Volumen von 0 verschieden, die Determinante also ungleich 0 ist ($\det M \neq 0$). Dabei gilt: Die Spaltenvektoren der Matrix M sind linear unabhängig genau dann, wenn der Rang der Matrix n ist genau dann, wenn die Abbildung $f(x) = Mx$ bijektiv und damit die Matrix invertierbar ist. D.h.:

$$\det M \neq 0 \quad \Leftrightarrow \quad M \text{ invertierbar}$$

Wir betrachten im Folgenden wieder eine allgemeine $n \times n$ -Matrix A mit Spaltenvektoren a_1 bis a_n , d.h. $A = (a_1, \dots, a_n)$.

Definition 7.0.1 (Determinante). Sei $n \in \mathbb{N}$ und K ein Körper. Wir nennen eine Abbildung $\det : (K^n)^n \rightarrow K$ eine *Determinante*, wenn gilt:

1) $\det$ ist *Multilinearform* auf K^n , d.h. $\det$ ist linear in jedem seiner n Argumente, d.h. für alle $a_i, y \in K^n, \lambda \in K, i = 1, \dots, n$ gilt:

$$\det(a_1, \dots, a_i + \lambda y, \dots, a_n) = \det(a_1, \dots, a_i, \dots, a_n) + \lambda \det(a_1, \dots, y, \dots, a_n),$$

2) $\det$ ist *alternierend*, d.h. ihr Vorzeichen ändert sich unter Vertauschung zweier Argumente, d.h. für alle $i \neq j$ mit $i, j \in \{1, \dots, n\}$ gilt:

$$\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -\det(a_1, \dots, a_j, \dots, a_i, \dots, a_n),$$

3) $\det$ ist *normiert*, d.h. falls wir Einheitsvektoren $e_1, \dots, e_n$ einsetzen, gilt:

$$\det(e_1, \dots, e_n) = 1.$$

Für $A \in K^{n \times n}$ setzen wir: $\det A = \det(a_1, \dots, a_n)$.

Bemerkung 7.0.3. Statt „ $\det A$ “ schreibt man manchmal auch: $|A|$. (Dies soll darauf hinweisen, dass die Determinante das Volumen misst).

Bemerkung 7.0.4 (Existenz und Eindeutigkeit der Determinante). Man kann zeigen, dass für jede Matrix $A \in K^{n \times n}$ eine Determinante existiert und diese eindeutig ist (wir sparen uns hier den Beweis).

Bemerkung 7.0.5. Aus Aussage 2) folgt, dass $\det(a_1, \dots, a_n) = 0$, falls zwei der n Vektoren a_1 bis a_n gleich sind. Begründung: Sei o.B.d.A. $a_j = a_i$. Dann folgt aus 2): $\det(a_1, \dots, a_i, \dots, a_i, \dots, a_n) = -\det(a_1, \dots, a_i, \dots, a_i, \dots, a_n)$, was nur erfüllt sein kann (die Determinante ist ein Skalar!), wenn $\det(a_1, \dots, a_n) = 0$.

Andersherum folgt aus der (Multi-)Linearität sowie der Aussage, dass $\det(a_1, \dots, a_n) = 0$, falls zwei der n Vektoren a_1 bis a_n gleich sind, dass die Determinante alternierend ist. Begründung:

$$\begin{aligned} 0 &= \det(a_1, \dots, a_i + a_j, \dots, a_i + a_j, \dots, a_n) = \det(a_1, \dots, a_i, \dots, a_i, \dots, a_n) \\ &\quad + \det(a_1, \dots, a_j, \dots, a_i, \dots, a_n) + \det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) + \det(a_1, \dots, a_j, \dots, a_j, \dots, a_n) \\ &= 0 + \det(a_1, \dots, a_j, \dots, a_i, \dots, a_n) + \det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) + 0 \\ &\Rightarrow \det(a_1, \dots, a_j, \dots, a_i, \dots, a_n) = -\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n). \end{aligned}$$

Satz 7.0.1 (Laplacescher Entwicklungssatz). Sei $n \geq 2 \in \mathbb{N}$, K ein Körper und $A \in K^{n \times n}$. Dann gilt:

$$1) \det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A^{(ij)} \quad (i = 1, \dots, n) \quad \text{„Entwicklung nach der } i\text{-ten Zeile“}$$

$$2) \det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A^{(ij)} \quad (j = 1, \dots, n) \quad \text{„Entwicklung nach der } j\text{-ten Spalte“}$$

Dabei bezeichnet $A^{(ij)} \in K^{(n-1) \times (n-1)}$ die Untermatrix, die entsteht, wenn man die i -te Zeile und j -te Spalte streicht.

Beweis. Evtl. Übung. □

Beispiel 7.0.3. Sei $M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Wir entwickeln z.B. nach der dritten Spalte:

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= +a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} - a_{23} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix} + a_{33} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &= a_{13}(a_{21}a_{32} - a_{31}a_{22}) - a_{23}(a_{11}a_{32} - a_{31}a_{12}) + a_{33}(a_{11}a_{22} - a_{21}a_{12}) \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) + a_{21}(a_{32}a_{13} - a_{33}a_{12}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \\ &= a_1 \cdot (a_2 \times a_3) \end{aligned}$$

Bemerkung 7.0.6 (Geometrische Anschauung). Es ist $|a_2 \times a_3| = |a_2||a_3|\sin\phi$ (dabei ist ϕ der Winkel zwischen a_2 und a_3). $|a_2 \times a_3|$ ist also maximal, wenn $\phi = \pi/2$, d.h. wenn a_2 und a_3 senkrecht aufeinander stehen (und $|a_2 \times a_3| = 0$, falls $\phi = 0$ oder $\phi = \pi$). Weiter gilt: $a_2 \times a_3 = |a_2||a_3|\sin\phi \cdot \hat{n}$, wobei $\hat{n}$ der Einheitsvektor ist, der senkrecht auf die Fläche steht, die von a_2 und a_3 aufgespannt wird. Zuletzt gilt: $a_1 \cdot (a_2 \times a_3) = |a_1||a_2 \times a_3|\cos\psi$ (dabei ist ψ der Winkel zwischen a_1 und $a_2 \times a_3$, also der Winkel zwischen a_1 und $\hat{n}$), d.h. $a_1 \cdot (a_2 \times a_3)$ ist maximal, falls $\psi = 0$ oder $\psi = \pi$, d.h. falls a_1 und $\hat{n}$ parallel oder antiparallel sind (und $a_1 \cdot (a_2 \times a_3) = 0$, falls $\psi = \pi/2$).

Bemerkung 7.0.7. Für eine Dreiecksmatrix R mit Diagonalelementen $r_{ii} \neq 0$ gilt also:

$$\det R = r_{11} \cdot r_{22} \cdot \dots \cdot r_{nn}.$$

Dies sieht man sofort, wenn man R (und die Untermatrizen) nach der 1. Spalte entwickelt.

Satz 7.0.2 (Rechenregeln für Determinanten). Seien $n \in \mathbb{N}$, K ein Körper, $A, B \in K^{n \times n}$ und $\lambda \in K$. Dann gilt:

- 1) $\det(\lambda A) = \lambda^n \det A$,
- 2) $\det(A \cdot B) = \det A \cdot \det B$,
- 3) $\det A^T = \det A$,
- 4) A invertierbar $\Leftrightarrow \det A \neq 0$,
- 5) A invertierbar $\Rightarrow \det A^{-1} = \frac{1}{\det A}$.

Beweis. Evtl. Übung. □

Bemerkung 7.0.8. Wichtig: Es gilt nicht, dass $\det(A + B) = \det A + \det B$!

Beispiel 7.0.4 (Cramersche Regel für $n = 2$). Es sei ein inhomogenes lineares Gleichungssystem aus zwei Gleichungen gegeben ($a_{11}, a_{12}, a_{21}, a_{22}, y_1, y_2 \in \mathbb{R}$ bekannt, x_1, x_2 gesucht):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= y_1 \\ a_{21}x_1 + a_{22}x_2 &= y_2 \end{aligned}$$

In Matrixnotation:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Wir wissen bereits: Dieses Gleichungssystem ist nur eindeutig lösbar, falls $\det A := a_{11}a_{22} - a_{12}a_{21} \neq 0$. Wir zeigen im Folgenden, dass dann für die Lösung gilt:

$$x_1 = \frac{y_1 a_{22} - y_2 a_{12}}{a_{11}a_{22} - a_{12}a_{21}} = \frac{\det \begin{pmatrix} y_1 & a_{12} \\ y_2 & a_{22} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

und

$$x_2 = \frac{a_{11}y_2 - a_{21}y_1}{a_{11}a_{22} - a_{12}a_{21}} = \frac{\det \begin{pmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

Satz 7.0.3 (Cramersche Regel). Seien $n \in \mathbb{N}, K$ ein Körper, $A = (a_1, \dots, a_n) \in K^{n \times n}$ invertierbar, $y \in K^n$ und $x \in K^n$ die eindeutige Lösung des Gleichungssystems $Ax = y$. Dann gilt $\forall i = 1, \dots, n$:

$$x_i = \frac{\det(a_1, \dots, a_{i-1}, y, a_{i+1}, \dots, a_n)}{\det A}$$

Beweis. Aus $Ax = y$ folgt $\sum_{k=1}^n a_k x_k = y$ und damit:

$$\begin{aligned} \det(a_1, \dots, a_{i-1}, y, a_{i+1}, \dots, a_n) &= \det(a_1, \dots, a_{i-1}, \sum_{k=1}^n a_k x_k, a_{i+1}, \dots, a_n) \\ &= \sum_{k=1}^n x_k \det(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n) = x_i \det A. \end{aligned}$$

Hier haben wir im letzten Schritt benutzt, dass $\det(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n) = 0$ für alle $k \neq i$ (siehe Bemerkung 7.0.5). $\square$

Satz 7.0.4 (Leibnizformel). Seien $n \in \mathbb{N}, K$ ein Körper, $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$. Dann gilt:

$$\det A = \sum_{\pi \in S_n} \sigma(\pi) \cdot a_{\pi(1)1} \cdot \dots \cdot a_{\pi(n)n}$$

Dabei bezeichnet $\sigma(\pi) := \det(e_{\pi(1)}, \dots, e_{\pi(n)}) \in \{-1, 1\}$ das Signum der Permutation $\pi \in S_n$.

Beweis. Evtl. Übung. □

Beispiel 7.0.5. Sei $n = 2$. Wir betrachten also S_2 , d.h. die Gruppe der Permutationen von $\{1, 2\}$. Diese Gruppe hat $|S_2| = 2! = 2$ Elemente. Explizit, $S_2 = \{id, \rho\}$ mit

$$id : \begin{Bmatrix} 1 & 2 \\ 1 & 2 \end{Bmatrix} \quad \text{und} \quad \rho : \begin{Bmatrix} 1 & 2 \\ 2 & 1 \end{Bmatrix}.$$

Damit ist:

$$\det A = \sum_{\pi \in S_2} \sigma(\pi) \cdot a_{\pi(1)1} \cdot a_{\pi(2)2} = \sigma(id) \cdot a_{11} \cdot a_{22} + \sigma(\rho) \cdot a_{21} \cdot a_{12} = +a_{11}a_{22} - a_{21}a_{12}.$$

Hier haben wir benutzt, dass

$$\sigma(id) = \det(e_1, e_2) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$$

und

$$\sigma(\rho) = \det(e_2, e_1) = \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1.$$

8 Eigenwerte und Eigenvektoren

Definition 8.0.1 (Eigenwert, Eigenvektor). Seien $n \in \mathbb{N}$, K ein Körper und $A \in K^{n \times n}$. Dann gilt: $\lambda \in K$ heißt *Eigenwert* von A , wenn es einen Vektor $v \in K^n \setminus \{0\}$ gibt, so dass

$$Av = \lambda v.$$

In diesem Fall heißt v *Eigenvektor* von A zum Eigenwert λ .

Beispiel 8.0.1. Sei $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. Dann gilt (ohne dass wir dies an dieser Stelle herleiten):

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow 1 \text{ ist Eigenwert und } \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ ist Eigenvektor zu } 1$$

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow 2 \text{ ist Eigenwert und } \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ ist Eigenvektor zu } 2$$

Satz 8.0.1 (Bestimmung der Eigenwerte). Seien $n \in \mathbb{N}$, K ein Körper, $A \in K^{n \times n}$ und $\lambda \in K$. Dann gilt:

$$\lambda \text{ Eigenwert von } A \Leftrightarrow \det(A - \lambda E_n) = 0.$$

Beweis. λ ist Eigenwert von $A \Leftrightarrow \exists v \in K^n \setminus \{0\}$, so dass $Av = \lambda v \Leftrightarrow \exists v \in K^n \setminus \{0\}$, so dass $(A - \lambda E_n)v = 0 \Leftrightarrow A - \lambda E_n$ nicht invertierbar (siehe Satz 6.2.5, wonach eine Matrix M invertierbar ist genau dann, wenn gilt: $Mv = 0 \Rightarrow v = 0$, hier aber: $Mv = 0 \wedge v \neq 0$, also M nicht invertierbar) $\Leftrightarrow \det(A - \lambda E_n) = 0$ (siehe Satz 7.0.2). $\square$

Beispiel 8.0.2. Wir betrachten das vorige Beispiel, also $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. Dann gilt:

$$\begin{aligned} \det(A - \lambda E_n) = 0 &\Leftrightarrow \det \left[\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \det \begin{pmatrix} 1-\lambda & 2 \\ 0 & 2-\lambda \end{pmatrix} \\ &= (1-\lambda)(2-\lambda) = \lambda^2 - 3\lambda + 2 = 0. \end{aligned}$$

Diese Gleichung besitzt zwei Lösungen: $\lambda_1 = 1$ und $\lambda_2 = 2$. Dies können wir direkt an der Form $(1-\lambda)(2-\lambda) = 0$ ablesen.

Sei nun $v = (v_1, v_2)$ der Eigenvektor zu $\lambda_1 = 1$ und $u = (u_1, u_2)$ der Eigenvektor zu $\lambda_2 = 2$. Wir bestimmen v durch Auflösen des Gleichungssystems $Av = \lambda_1 v$ (und analog für u), d.h.

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 1 \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \Leftrightarrow v_1 + 2v_2 = v_1 \text{ und } 2v_2 = v_2 \Rightarrow v = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 2 \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \Leftrightarrow u_1 + 2u_2 = 2u_1 \text{ und } 2u_2 = 2u_2 \Rightarrow u = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Bemerkung 8.0.1. Wir sehen an diesem Beispiel: auch Vielfache von Eigenvektoren sind selbst wieder Eigenvektoren, d.h.: v Eigenvektor $\Rightarrow \alpha v$ Eigenvektor, $\alpha \in K$.

Bemerkung 8.0.2. Eigenwerte können auch mehrfach vorkommen und 0 ist ein möglicher Eigenwert (allerdings ist 0 kein möglicher Eigenvektor!).

Definition 8.0.2 (Charakteristisches Polynom). Seien $n \in \mathbb{N}$, K ein Körper, $A \in K^{n \times n}$ und $\lambda \in K$. Die Funktion

$$\chi_A : K \rightarrow K, \quad \chi_A(\lambda) := \det(A - \lambda E_n)$$

heißt *charakteristisches Polynom* von A .

Beispiel 8.0.3. Wir betrachten wieder das vorige Beispiel, also $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. In dem Fall ist

$$\chi_A(\lambda) = \det(A - \lambda E_2) = \lambda^2 - 3\lambda + 2$$

das *charakteristische Polynom* von A .

Bemerkung 8.0.3. Das charakteristische Polynom einer $n \times n$ -Matrix ist allgemein von der Form $c_n \lambda^n + c_{n-1} \lambda^{n-1} + c_1 \lambda + c_0$ mit Koeffizienten $c_i \in K$ für $i = 0, \dots, n-1$ und $c_n = (-1)^n$. Es handelt sich also um ein Polynom vom Grad n . Dieses hat n (ggf. komplexe) Nullstellen. (Mehrfache Nullstellen werden hier mehrfach gezählt.) Um die Art und Vielfachheit der Nullstellen zu bestimmen, müssen wir nochmal ausführlicher über Polynome sprechen. Zunächst betrachten wir aber die Diagonalisierbarkeit von A .

Definition 8.0.3 (Diagonalmatrix). Wir bezeichnen eine Matrix $D = (d_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$ als *Diagonalmatrix*, wenn ihre Matriceinträge außerhalb der Diagonalen alle gleich 0 sind, d.h. wenn $d_{ij} = 0$ für $i \neq j$. Wir schreiben auch:

$$D = \text{diag}(d_{11}, \dots, d_{nn}).$$

Definition 8.0.4 (Diagonalisierbarkeit). Sei $n \in \mathbb{N}$ und K ein Körper. Eine Matrix $A \in K^{n \times n}$ heißt *diagonalisierbar*, wenn sie ähnlich zu einer Diagonalmatrix D ist, d.h. wenn es eine invertierbare Matrix $S \in K^{n \times n}$ gibt, so dass:

$$D = S^{-1}AS.$$

Bemerkung 8.0.4. Ist A diagonalisierbar, so stehen in der Diagonalmatrix D auf der Diagonalen die Eigenwerte von A , d.h.

$$D = S^{-1}A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Satz 8.0.2 (Diagonalisierbarkeit). Sei $n \in \mathbb{N}$, K ein Körper und $A \in K^{n \times n}$. A ist diagonalisierbar genau dann, wenn K^n eine Basis aus Eigenvektoren von A besitzt.

Beweis (von Satz 8.0.2).

„ $\Leftarrow$ “: Sei $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis von K^n mit $b_1, \dots, b_n$ Eigenvektoren von A , d.h. für alle $j = 1, \dots, n$: $A b_j = \lambda_j b_j$ (λ_j Eigenwerte).

$\mathcal{B} = \{b_1, \dots, b_n\}$ Basis von $K^n \Rightarrow \exists S : K^n \rightarrow K^n, S := (b_1, \dots, b_n)$ invertierbar mit $S e_j = b_j$ für alle $j = 1, \dots, n$.

Erinnerung: S ist die Matrix, die den Basiswechsel von $\mathcal{B}$ zur Einheitsbasis $\mathcal{A} := \{e_1, \dots, e_n\}$ beschreibt, d.h. explizit: $S := M_{\mathcal{A}}^{\mathcal{B}}(id_{K^n})$. Entsprechend ist: $S^{-1} = M_{\mathcal{B}}^{\mathcal{A}}(id_{K^n})$.

$b_1, \dots, b_n$ Eigenvektoren von $A \Rightarrow A S e_j = \lambda_j b_j = \lambda_j S e_j$ für alle $j = 1, \dots, n \Rightarrow S^{-1} A S e_j = S^{-1} \lambda_j S e_j = \lambda_j e_j$ für alle $j = 1, \dots, n$

$$\Rightarrow S^{-1} A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

„ $\Rightarrow$ “: A diagonalisierbar $\Leftrightarrow \exists S \in K^{n \times n}$ und $\exists \lambda_j, j = 1, \dots, n$, so dass:

$$S^{-1} A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

$$\Leftrightarrow S^{-1} A S e_j = \lambda_j e_j \text{ für alle } j = 1, \dots, n$$

$$\Leftrightarrow A S e_j = S \lambda_j e_j \text{ für alle } j = 1, \dots, n \text{ (wir haben beidseitig von links mit } S \text{ multipliziert)}$$

$$\Leftrightarrow A (S e_j) = \lambda_j (S e_j) \text{ für alle } j = 1, \dots, n$$

Weil $S \in K^{n \times n}$ invertierbar ist und $\{e_1, \dots, e_n\}$ eine Basis von K^n , folgt, dass $\{b_1, \dots, b_n\}$ mit $b_j := S e_j$ eine Basis von K^n ist.

Begründung: Dies gilt, weil die zu S gehörige lineare Abbildung $f : K^n \rightarrow K^n$ bijektiv ist. Explizit: Weil f surjektiv ist, bildet f auf ganz K^n ab, d.h. für das Erzeugendensystem $e_1, \dots, e_n$ von K^n gilt: $f(e_1), \dots, f(e_n)$ bilden ein Erzeugendensystem von K^n . Weil f injektiv ist, gilt für linear unabhängige Vektoren $e_1, \dots, e_n$: $f(e_1), \dots, f(e_n)$ sind linear unabhängig. Folglich ist $f(e_1) =: b_1, \dots, f(e_n) =: b_n$ eine Basis von K^n .

Insbesondere ist also $b_j = f(e_j) = S e_j \neq 0$ und damit $b_j = S e_j$ Eigenvektor von A (denn wir hatten bereits gezeigt, dass $A b_j = \lambda_j b_j$). $\square$

Beispiel 8.0.4. Wir betrachten wieder das vorige Beispiel, also $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. In dem Fall ist

$$S = \begin{pmatrix} v_1 & u_1 \\ v_2 & u_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

und

$$S^{-1}AS = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Wir erhalten also wie erwartet eine Diagonalmatrix mit den Eigenwerten auf der Diagonalen.

Definition 8.0.5 (Eigenraum, geometrische Vielfachheit). Sei $n \in \mathbb{N}$, K ein Körper, $A \in K^{n \times n}$ und λ ein Eigenwert von A .

1) Wir nennen

$$\text{Eig}_A(\lambda) := \{x \in K^n \mid Ax = \lambda x\}$$

den *Eigenraum* von A zum Eigenwert λ .

2) Wir nennen seine Dimension, $\dim \text{Eig}_A(\lambda)$, die *geometrische Vielfachheit* von λ .

Bemerkung 8.0.5 (Eigenraum). Es gilt:

$$\text{Eig}_A(\lambda) = \{x \in K^n \mid Ax = \lambda x\} = \text{Kern}(A - \lambda E_n).$$

Hier steht $A - \lambda E_n$ für die lineare Abbildung $f(x) = (A - \lambda E_n)x$. Eigenräume sind also in der Tat Untervektorräume und besitzen damit eine Basis und Dimension. Nach Definition sind sie genau die lineare Hülle der zu einem bestimmten Eigenwert gehörenden Eigenvektoren.

Beispiel 8.0.5. Wir betrachten wieder das vorige Beispiel, also $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. In dem Fall sind die Eigenwerte $\lambda_1 = 1$ mit Eigenvektor $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\lambda_2 = 2$ mit Eigenvektor $u = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Für die zugehörigen Eigenräume gilt:

$$\text{Eig}_A(1) = \left\{ \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}, \quad \dim \text{Eig}_A(1) = 1$$

$$\text{Eig}_A(2) = \left\{ \beta \begin{pmatrix} 2 \\ 1 \end{pmatrix} \mid \beta \in \mathbb{R} \right\}, \quad \dim \text{Eig}_A(2) = 1.$$

Bemerkung 8.0.6. Nur wenn ein Eigenwert mehrfach vorkommt, kann der Eigenraum mehr als eine Dimension haben! Dabei kommt ein Eigenwert mehrfach vor, wenn er mehrfache Nullstelle des charakteristischen Polynoms ist (mehr dazu gleich).

Bevor wir die Nullstellen eines Polynoms genauer analysieren, betrachten wir zuerst ein Beispiel.

Beispiel 8.0.6 (Polynomdivision). Wir betrachten ein Polynom P vom Grad 3 mit

$$P : \mathbb{R} \rightarrow \mathbb{R}, P(x) = x^3 - 2x^2 - 11x + 12.$$

Wir suchen die Nullstellen des Polynoms, d.h. diejenigen $\lambda_i \in \mathbb{R}$, für die $P(\lambda_i) = 0$.

Erster Schritt: Wir müssen durch Betrachten des Polynoms eine Nullstelle „erraten“ oder besser gesagt finden. Wir sehen: $\lambda_1 = 1$ ist eine Nullstelle. Diese können wir faktorisieren. Dabei gilt:

$$P(x) = (x - 1) \cdot Q(x)$$

mit Q einem (eindeutig festgelegten) Polynom vom Grad 2.

Zweiter Schritt: Wir berechnen Q wie gewohnt durch Polynomdivision:

$$\begin{array}{r} Q(x) = P(x) : (x - 1) = (x^3 - 2x^2 - 11x + 12) : (x - 1) = x^2 - x - 12 \\ \quad - \underline{(x^3 - x^2)} \\ \quad \quad - x^2 - 11x + 12 \\ \quad \quad - \underline{(-x^2 + x)} \\ \quad \quad \quad - 12x + 12 \end{array}$$

Letzter Schritt: Zuletzt können wir die verbleibenden Nullstellen $\lambda_{2/3}$ über die Formel für quadratische Gleichungen der Form $ax^2 + bx + c$ berechnen:

$$\lambda_{2/3} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-(-1) \pm \sqrt{(-1)^2 - 4 \cdot (-12)}}{2} = \frac{1 \pm \sqrt{49}}{2} = \frac{1 \pm 7}{2}$$

$\Rightarrow \lambda_2 = 4$ und $\lambda_3 = -3$.

Bemerkung 8.0.7. Wir verstehen an diesem Beispiel. Ein Polynom mit reellen Koeffizienten muss keine reellen Nullstellen haben (es hat maximal n reelle Nullstellen!), z.B. hat das Polynom $x^2 + 1$ nur zwei komplexe Nullstellen, nämlich $\lambda_{1/2} = \pm i$. Außerdem kann eine Nullstelle mehrfach vorkommen, z.B. hat das Polynom $x^2 - 4x + 4 = (x - 2)(x - 2)$ die doppelte Nullstelle $\lambda = 2$.

Allgemein gilt für die Nullstellen von Polynomen vom Grad n Folgendes.

Definition 8.0.6 (Polynom, Grad, Höchstkoeffizient, Nullstelle). Sei K ein unendlicher Körper (z.B. $K = \mathbb{R}$ oder $K = \mathbb{C}$) und $n \in \mathbb{N}_0$. Seien $a_0, a_1, \dots, a_n \in K$ und $a_n \neq 0$.

1) Wir nennen

$$P : K \rightarrow K, P(x) = \sum_{k=0}^n a_k x^k$$

ein *Polynom* vom Grad n (kurz: $n = \text{grad } P$). Dabei nennen wir a_n den *Höchstkoeffizienten*.

2) Wir nennen λ eine *Nullstelle* des Polynoms, falls $P(\lambda) = 0$.

Lemma 8.0.1 (Faktorisieren einer Nullstelle). Sei K ein unendlicher Körper, $n \in \mathbb{N}$, $P : K \rightarrow K$ ein Polynom vom Grad n und λ eine Nullstelle, d.h. $P(\lambda) = 0$. Dann gibt es ein eindeutig bestimmtes Polynom $Q : K \rightarrow K$ vom Grad $n - 1$ (mit eindeutigen Koeffizienten $a_0, \dots, a_{n-1}$), so dass

$$P(x) = (x - \lambda) \cdot Q(x).$$

Beweis. Allgemein ist P von der Form $P(x) = \sum_{k=0}^n a_k x^k$ mit $a_n \neq 0$. Wir machen eine Fallunterscheidung:

1. Fall ($\lambda = 0$): Aus $\lambda = 0$ Nullstelle, also $P(0) = a_n \cdot 0^n + \dots a_1 \cdot 0 + a_0 = 0$, folgt $a_0 = 0$.
Damit gilt:

$$P(x) = \sum_{k=1}^n a_k x^k = x \sum_{k=1}^n a_k x^{k-1} \Rightarrow Q(x) = \sum_{k=1}^n a_k x^{k-1} = \sum_{l=0}^{n-1} a_{l+1} x^l$$

ist eindeutig bestimmtes Polynom vom Grad $n-1$ (denn der Koeffizient vor x^{n-1} ist $a_n \neq 0$).

2. Fall ($\lambda \neq 0$): Wir betrachten die Funktion $\tilde{P}(x) := P(x + \lambda)$. Auch $\tilde{P}$ ist ein Polynom vom Grad n (dies erhält man z.B. aus dem binomischen Lehrsatz, wonach $(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$). Nun ist $\tilde{P}(0) = P(\lambda) = 0$ (denn λ ist Nullstelle von P). Nun sind wir wieder beim 1. Fall und wir können folgern, dass ein eindeutig bestimmtes Polynom $\tilde{Q} : K \rightarrow K$ vom Grad $n-1$ existiert mit

$$\tilde{P}(x) = x \cdot \tilde{Q}(x).$$

Statt x setzen wir in diese Gleichung $x - \lambda$ ein und bekommen:

$$\tilde{P}(x - \lambda) = (x - \lambda) \cdot \tilde{Q}(x - \lambda)$$

mit $\tilde{P}(x - \lambda) = P(x)$. Somit $P(x) = (x - \lambda) \cdot \tilde{Q}(x - \lambda) =: (x - \lambda) \cdot Q(x)$ mit eindeutig bestimmtem Polynom Q vom Grad $n-1$. $\square$

Lemma 8.0.2 (Vielfachheit einer Nullstelle). *Sei K ein unendlicher Körper, $n \in \mathbb{N}$, $P : K \rightarrow K$ ein Polynom vom Grad $n \geq 1$ und λ eine Nullstelle, d.h. $P(\lambda) = 0$. Dann gibt es ein eindeutig bestimmtes $m \in \mathbb{N}$ und ein eindeutig bestimmtes Polynom $Q : K \rightarrow K$, $Q(\lambda) \neq 0$, so dass*

$$P(x) = (x - \lambda)^m \cdot Q(x).$$

Wir nennen m die Vielfachheit der Nullstelle λ .

Beweis. Wiederholte Anwendung von Lemma 8.0.1, bis wir bei einem Polynom Q ankommen, für das gilt: $Q(\lambda) \neq 0$. $\square$

Lemma 8.0.3. *Sei K ein unendlicher Körper, $n \in \mathbb{N}$ und $P : K \rightarrow K$ ein Polynom vom Grad n . Seien $\lambda_1, \dots, \lambda_l$ mit $l \in \mathbb{N}_0$ paarweise verschiedene Nullstellen. Dann gibt es eindeutige $m_1, \dots, m_l \in \mathbb{N}$ mit $m_1 + \dots + m_l \leq n$ und ein eindeutig bestimmtes Polynom $Q : K \rightarrow K$ vom Grad $n - (m_1 + \dots + m_l)$, $Q(x) \neq 0$, so dass*

$$P(x) = (x - \lambda_1)^{m_1} \cdot \dots \cdot (x - \lambda_l)^{m_l} \cdot Q(x).$$

Beweis. Folgt direkt als Verallgemeinerung von Lemma 8.0.2. $\square$

Damit kommen wir nun zu folgendem wichtigen Satz (der von Gauß bewiesen wurde).

Satz 8.0.3 (Fundamentalsatz der Algebra). *Jedes Polynom $P : \mathbb{C} \rightarrow \mathbb{C}$ vom Grad $n \in \mathbb{N}$ besitzt genau n Nullstellen in $\mathbb{C}$ (entsprechend ihrer Vielfachheit gezählt). Insbesondere gibt es ein $l \in \mathbb{N}$*

und $m_1, \dots, m_l \in \mathbb{N}$ mit $m_1 + \dots + m_l = n$ sowie paarweise verschiedene $\lambda_1, \dots, \lambda_l \in \mathbb{C}$ und ein $\alpha \in \mathbb{C} \setminus \{0\}$, so dass

$$P(x) = \alpha \cdot (x - \lambda_1)^{m_1} \cdot \dots \cdot (x - \lambda_l)^{m_l}.$$

Beweis. Folgt aus Lemma 8.0.2 zusammen mit dem Beweis, dass mindestens eine Nullstelle existiert (Gauß). $\square$

Wir kommen nun wieder zurück zum Eigenwertproblem.

Definition 8.0.7 (Algebraische Vielfachheit). Sei K ein unendlicher Körper, $n \in \mathbb{N}$, $A \in K^{n \times n}$. Sei $\lambda \in K$ ein Eigenwert von A . Wir nennen $m \in \mathbb{N}$ die *algebraische Vielfachheit* des Eigenwerts λ , wenn λ eine m -fache Nullstelle des charakteristischen Polynoms $\chi_A = \det(A - \lambda E_n)$ ist.

Bemerkung 8.0.8. Aus dem, was wir über Polynome gelernt haben, folgt unmittelbar:

- 1) Eine Matrix $A \in \mathbb{R}^{n \times n}$ besitzt höchstens n reelle Eigenwerte.
- 2) Eine Matrix $B \in \mathbb{C}^{n \times n}$ besitzt genau n komplexe Eigenwerte.

Dabei zählen wir die Eigenwerte entsprechend ihrer algebraischen Vielfachheit.

Wir haben zuvor bereits über die geometrische Vielfachheit von Eigenwerten gesprochen. Die algebraische und geometrische Vielfachheit hängen wie folgt zusammen.

Satz 8.0.4 (Algebraische und geometrische Vielfachheit). Sei K ein unendlicher Körper, $n \in \mathbb{N}$, $A \in K^{n \times n}$. Sei $\lambda \in K$ ein Eigenwert von A . Dann gilt: Die algebraische Vielfachheit von λ ist größer oder gleich der geometrischen Vielfachheit von λ .

Beweis. Müssen wir aus Zeitgründen weglassen. $\square$

Bemerkung 8.0.9. Wenn also ein Eigenwert λ eine m -fache Nullstelle des charakter. Polynoms ist, kann der Eigenraum maximal die Dimension m haben, d.h. es kann maximal m linear unabhängige Eigenvektoren zum Eigenwert λ geben. Dabei gilt: Die Eigenvektoren zu *unterschiedlichen* Eigenwerten sind ohnehin linear unabhängig.

Das heißt, die Eigenvektoren von A bilden genau dann eine Basis von K^n – und damit ist A genau dann diagonalisierbar –, wenn sich die geometrischen Vielfachheiten aller Eigenwerte zu n addieren (wenn es also n linear unabhängige Eigenvektoren gibt).

Damit ergibt sich der folgende Satz.

Satz 8.0.5. Sei K ein unendlicher Körper, $n \in \mathbb{N}$, $A \in K^{n \times n}$. Seien $l \in \mathbb{N}$ und $\lambda_1, \dots, \lambda_l$ paarweise verschiedene Eigenwerte von A mit geometrischen Vielfachheiten $k_1, \dots, k_l$. Dann gilt: A ist diagonalisierbar genau dann, wenn $k_1 + \dots + k_l = n$.

Beweis. Beweisidee: k_i beschreibt die Anzahl linear unabhängiger Eigenvektoren zum Eigenwert λ_i . Dies gilt $\forall i = 1, \dots, l$ und Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig, woraus folgt: Es gibt insgesamt $k_1 + \dots + k_l = n$ linear unabhängige Eigenvektoren, diese bilden Basis von K_n . $\square$

Wir betrachten zum Schluss noch das Beispiel eines Eigenwertproblems bei einer 3×3 -Matrix.

Beispiel 8.0.7. Sei $A \in \mathbb{R}^{3 \times 3}$ mit

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

Wir berechnen zunächst die Eigenwerte. Das charakteristische Polynom ist:

$$\begin{aligned} \chi_A(\lambda) &= \det(A - \lambda E_n) = \det \begin{pmatrix} 1 - \lambda & 0 & 3 \\ 0 & 2 - \lambda & 0 \\ 1 & 0 & -1 - \lambda \end{pmatrix} \\ &= (1 - \lambda)(2 - \lambda)(-1 - \lambda) + 1(-1)(2 - \lambda)3 \\ &= (2 - \lambda)[(1 - \lambda)(-1 - \lambda) - 3] = (2 - \lambda)[\lambda^2 - 4] \\ &= (2 - \lambda)(\lambda - 2)(\lambda + 2) = -(2 - \lambda)^2(\lambda + 2) \end{aligned}$$

Nun setzen wir $\chi_A(\lambda) = 0$ und erhalten damit die Eigenwerte.

$\Rightarrow \lambda_1 = 2$ ist doppelte Nullstelle (d.h. von algebraischer Vielfachheit 2),

$\lambda_2 = -2$ ist einfache Nullstelle (d.h. von algebraischer Vielfachheit 1).

Nun bestimmen wir die Eigenräume:

Zu $\lambda_2 = -2$:

$$\text{Eig}_A(-2) = \{u \in \mathbb{R}^3 \mid Au = -2u\} = \text{Kern}(A + 2E_3) = \text{Kern} \begin{pmatrix} 3 & 0 & 3 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \text{Kern} \begin{pmatrix} 3 & 0 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Damit verbleiben die folgenden zwei Gleichungen:

$$3u_1 + 3u_3 = 0, \quad 4u_2 = 0.$$

Die Lösung dieses Gleichungssystems gibt uns den Eigenraum zu $\lambda_2 = -2$ (dessen Dimension 1 ist, d.h. die geometrische Vielfachheit vom Eigenwert $\lambda_2 = -2$ ist 1) :

$$\Rightarrow \text{Eig}_A(-2) = \left\{ \alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

Zu $\lambda_2 = 2$:

$$\begin{aligned}\text{Eig}_A(2) &= \{v \in \mathbb{R}^3 \mid Av = 2v\} = \text{Kern}(A - 2E_3) = \text{Kern} \begin{pmatrix} -1 & 0 & 3 \\ 0 & 0 & 0 \\ 1 & 0 & -3 \end{pmatrix} \\ &= \text{Kern} \begin{pmatrix} -1 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: \text{Kern } B\end{aligned}$$

Wir sehen, dass $\text{Rang } B = 1$. Damit gilt für die Dimension des Kerns: $\dim \text{Kern } B = 2$. Der Eigenwert $\lambda_2 = 2$ hat also geometrische Vielfachheit 2.

Alternativ überlegen wir uns, dass es nur eine Gleichung und damit zwei freie Parameter $\alpha, \beta \in \mathbb{R}$ (also zwei linear unabhängige Vektoren) in der Lösung von $Bv = 0$ gibt:

$$-1v_1 + 3v_3 = 0 \quad \Rightarrow \quad v_1 = 3v_3.$$

Wähle z.B. $v_2 = \alpha$ und $v_3 = \beta$. Dann erhalten wir als Lösung den folgenden Eigenraum:

$$\Rightarrow \text{Eig}_A(2) = \left\{ \alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \mid \alpha, \beta \in \mathbb{R} \right\}.$$

Wir haben also drei linear unabhängige Eigenvektoren:

$$u = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Die drei Eigenvektoren bilden zusammen eine Basis des $\mathbb{R}^3$. Die Matrix $A \in \mathbb{R}^{3 \times 3}$ ist also diagonalisierbar und es gilt:

$$D = S^{-1}AS \quad \text{mit} \quad S = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Dabei sind die Eigenvektoren genau die Spalten der Transformationsmatrix S . Wir rechnen nach, ob das stimmt und tatsächlich eine Diagonalmatrix D mit den Eigenwerten auf der Diagonalen rauskommt. Da A nicht symmetrisch und damit S nicht orthogonal ist, müssen wir S^{-1} explizit berechnen (sonst könnten wir benutzen, dass für orthogonale Matrizen gilt: $S^{-1} = S^T$). Wir

berechnen also S^{-1} :

$$\begin{aligned}
 \left(S \mid E_3 \right) &= \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 1 & 0 & 1 \end{array} \right) \\
 &\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{4} & 0 & -\frac{3}{4} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 1 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{4} & 0 & -\frac{3}{4} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{4} & 0 & \frac{1}{4} \end{array} \right) = \left(E_3 \mid S^{-1} \right) \\
 &\Rightarrow S^{-1} = \frac{1}{4} \cdot \begin{pmatrix} 1 & 0 & -3 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

Damit ist:

$$\begin{aligned}
 S^{-1}AS &= \frac{1}{4} \cdot \begin{pmatrix} 1 & 0 & -3 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\
 &= \frac{1}{4} \cdot \begin{pmatrix} 1 & 0 & -3 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 6 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.
 \end{aligned}$$

Wir erhalten also tatsächlich eine Diagonalmatrix D mit den Eigenwerten (entsprechend ihrer algebraischen Vielfachheit) auf der Diagonalen.

9 Ausblick: Lineare Algebra 2

U. A. die folgenden Themen sind Teil einer Vorlesung „Lineare Algebra 2“.

- Jordannormalform einer Matrix

Kann man eine Matrix $A \in K^{n \times n}$ *nicht* diagonalisieren, so ist die „zweitbeste Option“ die sog. *Jordannormalform* J . Dies ist eine Blockdiagonalform mit Eigenwerten auf der Diagonalen und 1-en auf der oberen Nebendiagonalen von der Art

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 & \dots & 0 \\ 0 & \lambda_1 & 1 & & \\ \vdots & 0 & \lambda_1 & 0 & \vdots \\ 0 & & & \ddots & 0 \\ 0 & \dots & & & \lambda_j \end{pmatrix}$$

Das Prozedere, sie zu erhalten, ist ähnlich wie bei der Diagonalisierung. Ist die geometrische Vielfachheit eines Eigenwerts kleiner als seine algebraische (und die Matrix somit nicht diagonalisierbar), d.h. gibt es nicht genug Eigenvektoren, um eine Basis des K^n zu bilden, so füllt man mit sog. *Hauptvektoren* v_p auf. Diese erhält man wie folgt:

$$(A - \lambda E_n)^p \cdot v_p = 0, \quad p = 1, 2, 3, \dots$$

Erinnerung: Für $p = 1$ entspricht dies der Gleichung für Eigenvektoren (diese sind also ein Spezialfall der Hauptvektoren). Im Anschluss macht man alles wie immer, nur dass in den Spalten von S nun die Eigen- und Hauptvektoren stehen. Dann gilt:

$$S^{-1}AS = J.$$

- Matrizen im Komplexen

Wir haben bereits bestimmte spezielle Matrizen im Reellen kennengelernt. Analog gibt es bestimmte spezielle Matrizen im Komplexen.

– **Im Reellen:** Sei $n \in \mathbb{N}$. Wir nennen eine Matrix $A \in \mathbb{R}^{n \times n}$ *symmetrisch*, falls

$$A^T = A.$$

Dabei ist A^T die zu A *transponierte* Matrix. Wir nennen eine Matrix $A \in \mathbb{R}^{n \times n}$ *orthogonal*, falls

$$A^{-1} = A^T.$$

In dem Fall gilt insbesondere, dass $E_n = A^{-1}A = A^T A$ und $|\det A| = 1$.

- **Im Komplexen:** Sei $n \in \mathbb{N}$. Wir nennen eine Matrix $B \in \mathbb{C}^{n \times n}$ *selbstadjungiert*, falls

$$\overline{B^T} = B.$$

Dabei ist $\overline{B^T}$ die zu B *adjungierte* (= transponierte und komplex konjugierte) Matrix. Wir nennen eine Matrix $B \in \mathbb{C}^{n \times n}$ *unitär*, falls

$$B^{-1} = \overline{B^T}.$$

In dem Fall gilt, dass $E_n = B^{-1}B = \overline{B^T}B$ und $|\det B| = 1$.

- Skalarprodukt im Komplexen

Wir haben bereits das Skalarprodukt im Reellen kennengelernt. Analog gibt es das Skalarprodukt im Komplexen.

- **Im Reellen:** Seien $x \in \mathbb{R}^k, y \in \mathbb{R}^k$. Es ist

$$\langle x, y \rangle = \sum_{k=1}^n x_k y_k$$

das Standardskalarprodukt in $\mathbb{R}$. Es ist *bilinear*, *symmetrisch* und *positiv definit*.

- **Im Komplexen:** Seien $x \in \mathbb{C}^k, y \in \mathbb{C}^k$. Es ist

$$\langle x, y \rangle = \sum_{k=1}^n \bar{x}_k y_k$$

das Skalarprodukt in $\mathbb{C}$. Es ist *sesquilinear*, *hermitesch* und *positiv definit*. Dabei bedeutet sesquilinear, dass zwar $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$, aber $\langle \lambda x, y \rangle = \bar{\lambda} \langle x, y \rangle$, und hermitesch, dass $\langle x, y \rangle = \overline{\langle y, x \rangle}$.